

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 640 100**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**C12N 9/10** (2006.01)

**C11C 3/00** (2006.01)

**C10L 1/02** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **28.06.2011 PCT/AU2011/000794**

87 Fecha y número de publicación internacional: **05.01.2012 WO12000026**

96 Fecha de presentación y número de la solicitud europea: **28.06.2011 E 11799957 (3)**

97 Fecha y número de publicación de la concesión europea: **21.06.2017 EP 2585589**

54 Título: **Métodos para producir lípidos**

30 Prioridad:

**12.05.2011 US 201161485349 P**

**09.07.2010 US 399286 P**

**28.06.2010 AU 2010902841**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**31.10.2017**

73 Titular/es:

**COMMONWEALTH SCIENTIFIC AND INDUSTRIAL  
RESEARCH ORGANISATION (100.0%)**

**Black Mountain Science and Innovation Park,  
Clunies Ross Street  
Acton ACT 2601, AU**

72 Inventor/es:

**PETRIE, JAMES;  
VANHERCKE, THOMAS;  
SHRESTHA, PUSHKAR;  
LIU, QING;  
SINGH, SURINDER, PAL y  
ZHOU, XUE-RONG**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 640 100 T3

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).

## DESCRIPCIÓN

Métodos para producir lípidos

5 **Campo de la invención**

La presente invención se refiere a métodos para producir lípidos. En particular, la presente invención se relaciona con métodos para aumentar el nivel de uno o más lípidos no polares y/o el contenido total de lípidos no polares, incluyendo triacilglicerol (TAG), en un organismo transgénico o parte del mismo. En una realización particular, la presente invención se refiere al uso de una monoacilglicerol aciltransferasa (MGAT) para aumentar el nivel de uno o más lípidos no polares y/o el contenido total de lípidos no polares, incluyendo triacilglicerol (TAG), en plantas, semillas y/u hojas, algas, levaduras y hongos.

15 **Antecedentes de la invención**

Los lípidos vegetales, tales como los triacilgliceroles (TAG) del aceite de las semillas, tienen muchos usos, por ejemplo, usos culinarios (materia grasa, textura, sabor), usos industriales (en jabones, velas, perfumes, cosméticos, adecuados como agentes de secado, aislantes, lubricantes) y proporcionan valor nutricional. También existe un interés creciente en el uso de lípidos vegetales para la producción de biocombustible.

20 Biocombustible

La creciente demanda de fuentes de energía alternativas se puede satisfacer, al menos en parte, con un suministro renovable de biocombustibles derivados de plantas. Para que constituya una alternativa viable a los combustibles fósiles, el biocombustible deberá proporcionar una ganancia neta de energía en la producción, ofrecer beneficios para el medio ambiente, ser económicamente competitivo y producible en grandes cantidades sin reducir los suministros de alimento, una consecuencia no intencionada actual de la producción existente de biocombustible.

Las plantas representan una fuente de lípidos significativa porque muchas especies acumulan lípidos como principales componentes de almacenamiento en las semillas. La forma más importante de lípidos de almacenamiento vegetativo en las semillas, que representa, según la especie, un 15-50 % del peso de la semilla, es el triacilglicerol (TAG). Sin embargo, el sustrato primario para la síntesis de lípidos son los carbohidratos generados en los tejidos fotosintéticos verdes (hojas y tallos) que luego son metabolizados en los cloroplastos para producir unidades de ácidos grasos libres y acetil-coenzima A (acetil-CoA), los bloques de construcción básicos para el TAG. Por lo tanto, las hojas de las plantas son el principal lugar de síntesis del bloque de construcción de TAG. La cantidad de TAG acumulada en las semillas oleaginosas se puede determinar, en parte, por la cantidad de ácidos grasos producida en los plástidos (Bao y Ohlrogge, 1999). El almacenamiento final de TAG tiene lugar en las semillas, en pequeños orgánulos esféricos denominados cuerpos grasos. Tan solo un 0,2-0,3 % aproximadamente de la biomasa foliar está representada por TAG.

Las plantas de biomasa elevada, en particular plantas de hojas anchas de biomasa elevada, ofrecen un gran potencial de biocombustible. Las plantas con un rendimiento de entre 100-400 toneladas/acre de materiales de bajo costo, de gran valor de biomasa, son particularmente útiles, en especial cuando no existe ninguna de las restricciones de costos elevados, requerimientos de trabajo, agregado de compuestos químicos o geográficas asociadas con la producción de plantas de biomasa baja.

45 Monoacilglicerol aciltransferasas

La enzima monoacilglicerol aciltransferasa (MGAT) está asociada con mamíferos, principalmente con el intestino en mamíferos donde cataliza la síntesis de diacilglicerol (DAG) directamente a partir de monoacilglicerol (MAG) y acil-CoA grasa. Por el contrario, la vía primaria de síntesis de TAG presente en plantas es la vía de Kennedy o glicerol fosfato (Figura 1) que no incluye una etapa de MGAT. En la vía de Kennedy, se forma DAG a partir de una cadena principal de glicerol acilado en una reacción en dos etapas que consiste en una acilación inicial por la aciltransferasa del ácido lisofosfatídico (LPAAT) que agrega un acil-CoA grasa a un sustrato de ácido lisofosfatídico (LisoPA; LPA) y la posterior eliminación de un grupo fosfato del producto, ácido fosfatídico (PA), para obtener fosfato inorgánico (Pi) y DAG. Por el contrario, MGAT cataliza la formación de DAG directamente por acilación de un MAG con un grupo acilo proveniente de la acil-CoA grasa. Después de la síntesis de DAG, otra enzima, la diacilglicerol aciltransferasa (DGAT), acila DAG para formar TAG.

60 El primer gen de MGAT que se aisló era de ratón (MGAT1) y este gen codifica una enzima no soluble unida a membrana (Yen et al., 2002). Se han caracterizado otros genes MGAT similares en animales, incluyendo un segundo gen de MGAT de ratón (MGAT2) y tres genes humanos, pero no se ha confirmado la clonación de genes que codifican MGAT de plantas (Cao et al., 2003; Cheng et al., 2003).

65

Diacilglicerol aciltransferasas

La DGAT es una proteína integral de membrana que cataliza la etapa final enzimática en la producción de TAG en plantas, hongos y mamíferos. Esta enzima es responsable de la transferencia de un grupo acilo de la acil-coenzima A (acil-CoA) a DAG para formar TAG. La DGAT está asociada con fracciones de membrana y de cuerpos lipídicos en plantas y hongos, en particular, en las semillas oleaginosas donde contribuye al almacenamiento de carbono usado como reserva de energía. Se sabe que la DGAT regula la estructura de TAG y dirige la síntesis de TAG. Además, se sabe que la reacción de DGAT es específica de la síntesis de lípidos. La sobreexpresión de la DGAT dependiente de acil-CoA de manera específica en las semillas en plantas de tipo silvestre da como resultado un aumento de la dposición de aceite de semilla y del peso medio de las semillas (Jako et al., 2001).

Para maximizar los rendimientos para la producción comercial de lípidos, existe la necesidad de medios adicionales para aumentar los niveles de lípidos, en particular lípidos no polares tales como DAG y TAG, en organismos transgénicos o partes de los mismos, tales como plantas, semillas, hojas, algas y hongos.

**Sumario de la invención**

Los presentes inventores han demostrado sorprendentemente que la expresión transgénica de un gen MGAT o de un gen relacionado da como resultado aumentos significativos en el rendimiento de lípidos en células tales como células vegetales. Los presentes inventores también han identificado una nueva vía para la síntesis de DAG y TAG en organismos transgénicos, tales como plantas, que es diferente de la bien conocida vía de Kennedy.

Por consiguiente, la presente invención proporciona un método para producir lípidos extraídos, comprendiendo el método las etapas de:

- i) obtener un organismo transgénico no humano o una parte del mismo, que comprende uno o más polinucleótidos exógenos que codifican una monoacilglicerol aciltransferasa (MGAT), en donde el organismo transgénico no humano o una parte del mismo tiene un nivel aumentado de uno o más lípidos no polares en comparación con un organismo correspondiente o una parte del mismo que no contiene dichos uno o más polinucleótidos exógenos, en donde el organismo transgénico no humano es una planta, alga, levadura u hongo y los uno o más lípidos no polares incluyen triacilglicerol (TAG) y
- ii) extraer el lípido del organismo transgénico no humano o una parte del mismo, produciendo de esa manera los lípidos extraídos.

En una realización, el contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo, está aumentado en comparación con el organismo correspondiente o una parte del mismo.

El organismo transgénico no humano o una parte del mismo, puede definirse además por las características (i), (ii), (iii), ya sea solas o combinadas: la característica (i) cuantifica el grado del nivel aumentado de dichos uno o más lípidos no polares o del contenido total de lípidos no polares, que se puede expresar como el grado de aumento sobre una base en peso o como el aumento relativo comparado con el nivel en el organismo correspondiente no humano o una parte del mismo y/o la característica (ii) especifica el género o la especie de planta o la especie de hongo o alga u otro tipo celular y la característica (iii) especifica que dichos uno o más lípidos específicos que están aumentados.

En el caso de la característica (i), en una realización, el grado de aumento de dichos uno o más lípidos no polares es de al menos un 0,5 %, al menos un 1 %, al menos un 2 %, al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 6 %, al menos un 7 %, al menos un 8 %, al menos un 9 %, al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 % o al menos un 24 % (p/p) mayor sobre una base en peso que en el organismo correspondiente no humano o una parte del mismo, preferentemente hasta un aumento máximo de aproximadamente un 25 % (p/p) sobre una base en peso.

Además, en el caso de la característica (i), en una realización preferida, el contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo, está aumentado en comparación con el organismo correspondiente o una parte del mismo. En una realización, el contenido total de lípidos está aumentado en al menos un 0,5 %, al menos un 1 %, al menos un 2 %, al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 6 %, al menos un 7 %, al menos un 8 %, al menos un 9 %, al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 % o al menos un 24 % (p/p) mayor sobre una base en peso que en el organismo correspondiente no humano o una parte del mismo, preferentemente hasta un aumento máximo de aproximadamente un 25 % (p/p) sobre una base en peso.

Además, para la característica (i), en una realización, el nivel de dichos uno o más lípidos no polares y/o el contenido total de lípidos no polares es de al menos un 1 %, al menos un 2 %, al menos un 3 %, al menos un 4 %, al menos un 5 %, al menos un 6 %, al menos un 7 %, al menos un 8 %, al menos un 9 %, al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 % o al menos un 24 % (p/p) mayor sobre una base en peso que en el organismo correspondiente no humano o una parte del mismo, preferentemente hasta un aumento máximo de aproximadamente un 25 % (p/p) sobre una base en peso.

al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 %, al menos un 24 %, al menos un 25 %, al menos un 30 %, al menos un 35 %, al menos un 40 %, al menos un 45 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 % o al menos un 90 % mayor sobre una base relativa que el organismo correspondiente no humano o una parte del mismo.

Además, para la característica (i), el grado de aumento en el nivel de dichos uno o más lípidos no polares y/o el contenido total de lípidos no polares puede ser al menos 2 veces, al menos 3 veces, al menos 4 veces, al menos 5 veces, al menos 6 veces, al menos 7 veces, al menos 8 veces, al menos 9 veces o al menos 10 veces, preferentemente hasta un máximo de aproximadamente 12 veces mayor sobre una base relativa que en el organismo correspondiente no humano o una parte del mismo.

Para la característica (ii), el organismo transgénico no humano es una planta, un alga o un organismo adecuado para una fermentación, tal como una levadura u otro hongo, preferentemente una levadura oleaginosa u otro hongo. La planta puede ser, por ejemplo, *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Trifolium sp.*, *Elaeisis guineensis*, *Nicotiana benthamiana*, *Hordeum vulgare*, *Lupinus angustifolius*, *Oryza sativa*, *Oryza glaberrima*, *Camelina sativa*, *Miscanthus x giganteus* o *Miscanthus sinensis*.

En el caso de la característica (iii), TAG está aumentado. También pueden estar aumentados DAG, TAG y DAG, MAG, ácidos grasos poliinsaturados (PUFA) o un PUFA específico (tal como ácido eicosadienoico (EDA), ácido araquidónico (ARA), ácido alfa linolénico (ALA), ácido estearidónico (SDA), ácido eicosatrienoico (ETE), ácido eicosatetraenoico (ETA), ácido eicosapentaenoico (EPA), ácido docosapentaenoico (DPA), ácido docosahexaenoico (DHA)) o una combinación de dos de más de los mismos. El grado de aumento de TAG, DAG, TAG y DAG, MAG, PUFA o de un PUFA específico, puede ser como se definió previamente en la característica (i). Preferentemente, está aumentado el total de DAG y TAG.

Tal como se divulga en el presente documento, los uno o más lípidos no polares y/o el contenido total de lípidos no polares se define por la combinación de las características (i), (ii) y (iii) o las características (i) y (ii) o las características (i) y (iii).

En una realización, la parte es una semilla, un fruto, un tubérculo, una raíz o una parte vegetativa de una planta. La parte vegetativa de la planta puede ser una parte de planta aérea o una parte verde, tal como una hoja o un tallo. En otra realización, la parte es una célula de un organismo multicelular. El grado de aumento del nivel de dichos uno o más lípidos no polares y/o el contenido total de lípidos no polares de la parte de planta específica en esta realización puede ser como se definió previamente en la característica (i). La parte de planta de *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Oryza sativa*, *Oryza glaberrima*, *Camelina sativa*, *Glycine max* o *Zea mays* preferentemente es la semilla, en tanto la parte preferida de *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Trifolium sp.*, *Elaeisis guineensis*, *Nicotiana benthamiana*, *Hordeum vulgare*, *Lupinus angustifolius*, *Miscanthus x giganteus* o *Miscanthus sinensis* es una parte vegetativa, en particular hojas y tallos.

En una realización, la parte es una semilla y los lípidos extraídos son de aceites de las semillas. El método de la invención puede comprender además cosechar la semilla de la planta transgénica, pensar el aceite de la semilla y/o purificar el aceite de las semillas en una o más etapas. La semilla puede ser, por ejemplo, de una planta de colza, una planta de maíz, una planta de soja, una planta de lupina, una planta de cacahuete, una planta de girasol, una planta de algodón, una planta de cártamo o una planta de lino.

En una realización, el contenido total de aceite o el contenido total de ácidos grasos de la semilla es al menos de un 0,5 % (p/p) a un 25 % (p/p) mayor sobre una base en peso que en una semilla correspondiente que no contiene dichos uno o más polinucleótidos exógenos.

En una realización, el contenido relativo de DAG del aceite de las semillas es al menos un 10 %, al menos un 10,5 %, al menos un 11 %, al menos un 11,5 %, al menos un 12 %, al menos un 12,5 %, al menos un 13 %, al menos un 13,5 %, al menos un 14 %, al menos un 14,5 %, al menos un 15 %, al menos un 15,5 %, al menos un 16 %, al menos un 16,5 %, al menos un 17 %, al menos un 17,5 %, al menos un 18 %, al menos un 18,5 %, al menos un 19 %, al menos un 19,5 %, al menos un 20 % (p/p) mayor que en una semilla correspondiente. En una realización, el contenido de DAG de la semilla aumenta en una cantidad definida en la característica (i) y la semilla es de un género y/o de una especie definida en la característica (ii).

En una realización, el contenido relativo de TAG del aceite de las semillas es al menos un 5 %, al menos un 5,5 %, al menos un 6 %, al menos un 6,5 %, al menos un 7 %, al menos un 7,5 %, al menos un 8 %, al menos un 8,5 %, al menos un 9 %, al menos un 9,5 % o al menos un 10 % (p/p) mayor que en una semilla correspondiente. En una realización, el contenido de TAG de la semilla aumenta en una cantidad definida en la característica (i) y la semilla es de un género y/o de una especie definida en la característica (ii). En una realización, la semilla es una semilla de colza que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 45 %, al menos un 46 %, al menos un 47 %, al menos un 48 %, al menos un 49 %, al menos un 50 %, al menos un 51 %, al menos un 52 %, al

## ES 2 640 100 T3

menos un 53 %, al menos un 54 % o al menos un 55 %.

En una realización, la semilla es una semilla de maíz que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 5 %, al menos un 6 %, al menos un 7 %, al menos un 8 %, al menos un 9 % o al menos un 10 %.

5 En una realización, la semilla es una semilla de soja que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 %, al menos un 24 %, al menos un 25 %, al menos un 26 %, al menos un 27 %, al menos un 28 %, al menos un 29 % o al menos un 30 %.

10 En una realización, la semilla es una semilla de lupina que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 % o al menos un 16 %.

15 En una realización, la semilla es una semilla de cacahuete que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 50 %, al menos un 51 %, al menos un 52 %, al menos un 53 %, al menos un 54 % o al menos un 55 %.

20 En una realización, la semilla es una semilla de girasol que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 50 %, al menos un 51 %, al menos un 52 %, al menos un 53 %, al menos un 54 % o al menos un 55 %.

25 En una realización, la semilla es una semilla de algodón que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 41 %, al menos un 42 %, al menos un 43 %, al menos un 44 %, al menos un 45 %, al menos un 46 %, al menos un 47 %, al menos un 48 %, al menos un 49 %, al menos un 50 % o al menos un 30 %.

En una realización, la semilla es una semilla de cártamo que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 35 %, al menos un 36 %, al menos un 37 %, al menos un 38 %, al menos un 39 %, al menos un 40 %, al menos un 41 %, al menos un 42 %, al menos un 43 %, al menos un 44 % o al menos un 45 %.

30 En una realización, la semilla es una semilla de lino que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 36 %, al menos un 37 %, al menos un 38 %, al menos un 39 % o al menos un 40 %.

35 En una realización, la semilla es una semilla de *Camelina sativa* que tiene un contenido de aceite sobre una base en peso (p/p) de al menos un 36 %, al menos un 37 %, al menos un 38 %, al menos un 39 %, al menos un 40 %, al menos un 41 %, al menos un 42 %, al menos un 43 %, al menos un 44 % o al menos un 45 %.

40 En otra realización, la parte de organismo es una parte de planta vegetativa y el contenido de TAG, DAG, TAG y DAG o MAG de la parte de planta vegetativa es al menos un 10 %, al menos un 11 %, al menos un 12 %, al menos un 13 %, al menos un 14 %, al menos un 15 %, al menos un 16 %, al menos un 17 %, al menos un 18 %, al menos un 19 %, al menos un 20 %, al menos un 21 %, al menos un 22 %, al menos un 23 %, al menos un 24 %, al menos un 25 %, al menos un 30 % al menos un 35 %, al menos un 40 %, al menos un 45 %, al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 80 % o al menos un 90 % (p/p) mayor sobre una base relativa que el contenido de TAG, DAG, TAG y DAG o MAG de una parte de planta vegetativa correspondiente que no contiene dichos uno o más polinucleótidos exógenos. En una realización preferida, el MAG es 2-MAG. En una realización, el contenido de TAG, DAG, TAG y DAG o MAG de la parte de planta vegetativa se determina a partir de la cantidad de estos componentes de lípidos en los lípidos extraíbles de la parte de planta vegetativa. En una realización adicional, el contenido de TAG, DAG, TAG y DAG o MAG de la parte de planta vegetativa transgénica está aumentado en una cantidad según se define en la característica (i).

50 En una realización, al menos un 60 % (% en moles) del contenido de ácidos grasos del contenido total de lípidos no polares del organismo o una parte del mismo o de los lípidos extraídos del mismo, es ácido oleico.

55 En otra realización, el contenido de PUFA del organismo o una parte del mismo, está aumentado en comparación con el organismo correspondiente o una parte del mismo. En este contexto, el contenido de PUFA incluye PUFA esterificados (incluyendo TAG, DAG, etc.) y PUFA no esterificados. En una realización, el contenido de PUFA del organismo o una parte del mismo, preferentemente se determina a partir de la cantidad de PUFA en el lípido extraíble del organismo o una parte del mismo. El grado de aumento en el contenido de PUFA puede ser como se define en la característica (i). El contenido de PUFA puede comprender EDA, ARA, ALA, SDA, ETE, ETA, EPA, DPA, DHA o una combinación de dos de más de los mismos.

60 En otra realización, el nivel de un PUFA en el organismo o una parte del mismo o el lípido extraído del mismo está aumentado en comparación con el organismo correspondiente o una parte del mismo o el lípido extraído del mismo. El PUFA puede ser EDA, ARA, ALA, SDA, ETE, ETA, EPA, DPA, DHA o una combinación de dos o más de los mismos. El grado de aumento del PUFA puede ser como se define en la característica (i).

65 En una realización, el nivel de dichos uno o más lípidos no polares (tales como TAG, DAG, TAG y DAG, MAG, PUFA

o un PUFA específico) y/o el contenido total de lípidos no polares se puede determinar por análisis mediante cromatografía de gases de los metil ésteres de ácidos grasos obtenidos de los lípidos extraídos. En la técnica se conocen métodos alternativos para determinar cualquiera de estos contenidos e incluyen métodos que no requieren de extracción del lípido del organismo o una parte del mismo, por ejemplo, análisis por infrarrojo cercano (NIR) o resonancia magnética nuclear (RMN).

En una realización, dichos uno o más polinucleótidos exógenos codifican:

- i) una monoacilglicerol aciltransferasa (MGAT),
- ii) una MGAT y una glicerol-3-fosfato aciltransferasa (GPAT), o
- iii) una MGAT y una DGAT, o
- iv) una MGAT, una GPAT y una DGAT.

En una realización, el polinucleótido exógeno codifica una MGAT que cataliza la acilación de *sn*-1 MAG o *sn*-2 MAG para formar *sn*-1,3 DAG o *sn*-1,2/2,3-DAG, respectivamente. En una realización preferida, la MGAT cataliza la acilación de *sn*-2 MAG para formar *sn*-1,2/2,3-DAG. El polinucleótido exógeno que codifica la MGAT puede comprender una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 1 a 44,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 45 a 82 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas.

En una realización, el polinucleótido exógeno codifica una MGAT1, que comprende una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 1, 3 a 5 o 7 a 23,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 45 a 61 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas.

En otra realización, el polinucleótido exógeno codifica una MGAT2, que comprende una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 2, 6 o 24 a 37,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 62 a 75 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas.

En otra realización, el polinucleótido exógeno codifica una MGAT3, que comprende una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 38 a 44,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 76 a 82 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas.

En una realización adicional, el polinucleótido exógeno codifica una DGAT que cataliza la acilación de *sn*-1,3 DAG o de *sn*-1,2/2,3-DAG, preferentemente *sn*-1,2/2,3-DAG, para formar TAG. En una realización, el polinucleótido exógeno codifica una DGAT2 que comprende una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 204 a 211,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 212 a 219 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas. En una realización preferida, la DGAT2 comprende una secuencia de nucleótidos de la SEQ ID NO: 204 y/o una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en la SEQ ID NO: 212.

En una realización adicional, el polinucleótido exógeno codifica una glicerol-3-fosfato aciltransferasa (GPAT). En una realización preferida, la GPAT también tiene actividad fosfatasa y produce MAG (es decir, una GPAT que acila G-3-P para formar *sn*-1 LPA o bien, *sn*-2 LPA y elimina un grupo fosfato de LPA para formar MAG). En una realización adicional preferida, la GPAT es una *sn*-2 GPAT (es decir, con preferencia para producir *sn*-2 LPA a partir de G-3-P) y tiene actividad fosfatasa para producir 2-MAG, por ejemplo, GPAT4 o GPAT6 de *Arabidopsis*. El polinucleótido

exógeno que codifica la GPAT puede comprender una o más de:

- i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO: 84 a 141,
- ii) una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO: 144 a 201 o un fragmento biológicamente activo de la misma,
- iii) una secuencia de nucleótidos que es al menos un 50 % idéntica a i) o ii), o
- iv) una secuencia de nucleótidos que se hibrida con cualquiera de i) a iii) en condiciones rigurosas.

En una realización adicional, el polinucleótido exógeno codifica una GPAT que tiene actividad fosfatasa que comprende una o más secuencias de aminoácidos conservadas provistas en las SEQ ID NO: 225, 226 y 227 o una secuencia de aminoácidos que es al menos un 50 %, preferentemente al menos un 60 %, más preferentemente al menos un 65 % idénticas a las mismas.

En una realización, dichos uno o más polinucleótidos exógenos codifican una MGAT y posiblemente también DGAT y/o GPAT mutantes. Por ejemplo, dichos uno o más polinucleótidos exógenos pueden codificar una MGAT y posiblemente también DGAT y/o GPAT que tiene una sustitución de aminoácidos conservadora como se ejemplifica en la Tabla 1 con relación a una MGAT y/o DGAT y/o GPAT de tipo silvestre.

En una realización, el organismo transgénico no humano o una parte del mismo, comprende un primer polinucleótido exógeno que codifica una MGAT y un segundo polinucleótido exógeno que codifica una GPAT. El primer y segundo polinucleótidos se pueden proporcionar como moléculas separadas o se pueden proporcionar como una sola molécula contigua. En una realización preferida, la GPAT es una GPAT que tiene actividad fosfatasa tal como una GPAT4 o GPAT6 de *Arabidopsis*. La GPAT que tiene actividad fosfatasa actúa catalizando la formación de MAG a partir de G-3-P (es decir, acila G-3-P para formar LPA y posteriormente elimina un grupo fosfato para formar MAG) en el organismo transgénico no humano o una parte del mismo. La MGAT actúa luego catalizando la formación de DAG en el organismo transgénico no humano o una parte del mismo, por acilación de MAG con un grupo acilo derivado de acil-CoA grasa. La MGAT, tal como MGAT1 de *A. thaliana*, también puede actuar catalizando la formación de TAG en el organismo transgénico no humano o una parte del mismo, si también tiene actividad DGAT.

El organismo transgénico no humano o una parte del mismo, puede comprender un tercer polinucleótido exógeno que codifica, por ejemplo, una DGAT. El primer, el segundo y el tercer polinucleótidos se pueden proporcionar como moléculas separadas o se pueden proporcionar como a sola molécula contigua. La DGAT actúa catalizando la formación de TAG en el organismo transgénico no humano o una parte del mismo, por acilación de DAG (producido preferentemente por la vía MGAT) con un grupo acilo derivado de acil-CoA grasa.

En otra realización, el organismo transgénico no humano o una parte del mismo, comprende un primer polinucleótido exógeno que codifica una MGAT y un segundo polinucleótido exógeno que codifica una DGAT. El primer y segundo polinucleótidos se pueden proporcionar como moléculas separadas o se pueden proporcionar como una sola molécula contigua. El transgénico no humano puede comprender un tercer polinucleótido exógeno que codifica, por ejemplo, una GPAT, preferentemente una GPAT que tiene actividad fosfatasa tal como una GPAT4 o GPAT6 de *Arabidopsis*. El primer, el segundo y el tercer polinucleótidos se pueden proporcionar como moléculas separadas o se pueden proporcionar como a sola molécula contigua.

En una realización adicional, el nivel de dichos uno o más lípidos no polares y/o el contenido total de lípidos no polares del organismo transgénico o una parte del mismo, es al menos 0,5 % (p/p) mayor sobre una base en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que un organismo correspondiente o una parte del mismo, que no contiene dichos uno o más polinucleótidos exógenos pero que comprende un polinucleótido exógeno que codifica una DGAT1 de *Arabidopsis thaliana* (SEQ ID NO: 83).

En otra realización adicional más, el organismo transgénico no humano o una parte del mismo, comprende además una o más mutaciones introducidas y/o un polinucleótido exógeno que regula negativamente la producción y/o la actividad de una enzima endógena del organismo transgénico no humano o una parte del mismo, seleccionada entre DGAT, *sn-1* GPAT, 1-acil-glicerol-3-fosfato aciltransferasa (LPAAT), acil-CoA:lisofosfatidilcolina aciltransferasa (LPCAT), ácido fosfatídico fosfatasa (PAP) o una combinación de dos o más de las mismas. La *sn-1* GPAT puede ser una GPAT que en su estado de tipo silvestre no tiene una actividad fosfatasa detectable, por ejemplo, una GPAT1 o GPAT3. La GPAT1 puede tener una secuencia de aminoácidos provista en la SEQ ID NO: 202 o un homólogo de la misma. La GPAT3 puede tener una secuencia de aminoácidos que se muestra en la SEQ ID NO: 203 o un homólogo de la misma.

El polinucleótido exógeno se puede seleccionar, por ejemplo, entre: un polinucleótido antisentido, un polinucleótido con sentido, un polinucleótido catalítico, un microARN, un polinucleótido que codifica un polipéptido que se une a la enzima endógena y un ARN bicatenario.

En otro aspecto, se divulga en el presente documento un método para producir lípidos extraídos, comprendiendo el método las etapas de:

i) obtener un organismo fototrófico transgénico o una parte del mismo, que comprende una monoacilglicerol aciltransferasa (MGAT) exógena, en donde dicho organismo fototrófico transgénico o una parte del mismo, tiene un nivel aumentado de un lípido no polar en comparación con dicho organismo o una parte del mismo, que no contiene una MGAT exógena, y

5 ii) extraer el lípido del organismo fototrófico transgénico o una parte del mismo, produciendo de esa manera los lípidos extraídos.

En otro aspecto, la presente invención proporciona un organismo transgénico no humano o una parte del mismo, que comprende uno o más polinucleótidos exógenos que codifican una monoacilglicerol aciltransferasa (MGAT), en donde el organismo transgénico no humano o una parte del mismo, tiene un nivel aumentado de uno o más lípidos no polares en comparación con un correspondiente organismo o una parte del mismo, que no contiene dichos uno o más polinucleótidos exógenos, en donde el organismo transgénico no humano es una palnta, alga, levadura u hongo y los uno o más lípidos no polares incluyen triacilglicerol (TAG).

15 En una realización preferida, el contenido total de lípidos del organismo transgénico no humano o una parte del mismo, está aumentado en comparación con el organismo correspondiente o una parte del mismo. El grado de aumento de dichos uno o más lípidos y/o el contenido total de lípidos no polares pueden ser como se define en la característica (i). El organismo transgénico no humano o una parte del mismo, es como se define en la característica (ii). El lípido no polar se define según la característica (iii). El organismo transgénico no humano o una parte del mismo, puede definirse por la combinación de las características (i), (ii) y (iii) o de las características (i) y (ii) o de las características (i) y (iii).

25 En una realización, el nivel de dichos uno o más lípidos no polares del organismo transgénico no humano o una parte del mismo, es al menos un 0,5 % (p/p) mayor sobre una base en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en el organismo correspondiente o una parte del mismo o como también se define en la característica (i).

30 En una realización adicional, el contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo, es al menos un 0,5 % (p/p) mayor sobre una base en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa, que en el organismo correspondiente no humano o una parte del mismo o también como se define en la característica (i).

35 En una realización, el contenido de PUFA del organismo transgénico no humano o una parte del mismo o el lípido extraído del mismo, está aumentado en comparación con el organismo correspondiente o una parte del mismo o el lípido extraído del mismo. El grado de aumento en el contenido de PUFA puede ser como se define en la característica (i).

40 En otra realización, el nivel de un PUFA en el organismo transgénico no humano o una parte del mismo o el lípido extraído del mismo está aumentado en comparación con el correspondiente organismo o una parte del mismo o el lípido extraído del mismo. El grado de aumento del PUFA puede ser como se define en la característica (i).

45 En una realización adicional, el nivel de dichos uno o más lípidos no polares y/o el contenido total de lípidos del organismo transgénico o una parte del mismo, es al menos 0,5 % (p/p) mayor sobre una base en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en un organismo correspondiente o una parte del mismo, que no contiene dichos uno o más polinucleótidos exógenos pero que comprende un polinucleótido exógeno que codifica una DGAT1 de *Arabidopsis thaliana*.

50 De acuerdo con la invención, el organismo transgénico no humano es una planta, un alga o un organismo adecuado para la fermentación tal como una levadura o un hongo. La planta puede ser como se define en la característica (ii).

La presente invención también proporciona un organismo transgénico no humano que comprende uno o más polinucleótidos exógenos que codifican:

- 55 i) una MGAT,  
ii) una MGAT y una GPAT  
iii) una MGAT y una DGAT, o  
iv) una MGAT, una GPAT y una DGAT.

60 El organismo transgénico no humano se puede caracterizar además por una o más características definidas en el presente documento. Dichos uno o más polinucleótidos exógenos pueden comprender una secuencia según se definió anteriormente.

65 En otro aspecto, la presente invención proporciona un método para obtener una célula con mayor capacidad para producir uno o más lípidos no polares, donde dicho método comprende

- i) introducir en una célula uno o más polinucleótidos exógenos que codifican



- a) una MGAT,  
 b) una MGAT y una GPAT  
 c) una MGAT y una DGAT o  
 d) una MGAT, una GPAT y una DGAT,
- 5 en donde dichos uno o más polinucleótidos exógenos están ligados operativamente a uno o más promotores que tienen la capacidad de dirigir la expresión de dichos uno o más polinucleótidos exógenos en la célula, en donde la célula es una célula vegetal, de alga, de levadura o fúngica y en donde los uno o más lípidos no polares incluyen triacilglicerol,
- 10 ii) expresar dichos uno o más polinucleótidos exógenos en la célula,  
 iii) analizar el contenido de lípidos de la célula y  
 iv) seleccionar una célula que tenga un nivel aumentado de uno o más lípidos no polares en comparación con una célula correspondiente que no contiene a los polinucleótidos exógenos.
- 15 En una realización preferida, el contenido total de lípidos de la célula seleccionada está aumentado en comparación con la célula correspondiente.
- En una realización, dichos uno o más lípidos no polares y/o el contenido total de lípidos no polares de la célula seleccionada es al menos un 0,5 % (p/p) mayor sobre una base en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en la célula correspondiente o como se define adicionalmente en la característica (i).
- 20 En otra realización, el contenido de PUFA de la célula seleccionada está aumentado en comparación con la célula correspondiente. El grado de aumento en el contenido de PUFA puede ser como se define en la característica (i).
- 25 En otra realización, el nivel de un PUFA en la célula seleccionada está aumentado en comparación con la célula correspondiente. El grado de aumento del PUFA puede ser como se define en la característica (i).
- Dichos uno o más polinucleótidos exógenos pueden comprender en este aspecto una secuencia según se definió anteriormente. Además, puede ser desconocido antes de aplicar el método que dichos uno o más polinucleótidos exógenos codifiquen una MGAT, una DGAT2, una MGAT y una GPAT, una MGAT y una DGAT o una MGAT, una GPAT y una DGAT, pero pueden ser candidatos para ello. El método se puede usar entonces como un ensayo para identificar o seleccionar polinucleótidos que codifican una MGAT, una DGAT2, una MGAT y una GPAT (preferentemente una GPAT que también tiene actividad fosfatasa), una MGAT y una DGAT o una MGAT, una GPAT (preferentemente una GPAT que también tiene actividad fosfatasa) y una DGAT.
- 30  
 35 La expresión de GPAT puede dar como resultado un aumento en los niveles de MAG en la célula si la GPAT también tiene actividad fosfatasa, la expresión de una MGAT puede dar como resultado un aumento en los niveles de DAG en la célula, en tanto la expresión de una DGAT puede dar como resultado un aumento en los niveles de TAG en la célula. La expresión de una MGAT también puede dar como resultado un aumento en los niveles de TAG en la célula, si la MGAT también tiene actividad DGAT. La expresión de una MGAT y una DGAT puede dar como resultado un aumento en los niveles de DAG y/o TAG en la célula. La expresión de una MGAT y una GPAT puede dar como resultado un aumento en los niveles de MAG (si la GPAT también tiene actividad fosfatasa) y/o DAG en la célula. La expresión de una MGAT y una GPAT también puede dar como resultado un aumento en los niveles de TAG en la célula, si la MGAT también tiene actividad DGAT. La expresión de una MGAT, una GPAT y una DGAT puede dar como resultado un aumento de los niveles de MAG (si la GPAT también tiene actividad fosfatasa) y/o DAG y/o TAG en la célula.
- 40  
 45 En una realización, la célula seleccionada es una célula de acuerdo con la invención.
- 50 En una realización, los polinucleótidos exógenos están integrados de manera estable en el genoma de la célula.
- El método puede comprender además las etapas de regenerar un organismo transgénico a partir de la célula y/u obtener progenie de la célula, por ejemplo obtener semillas, donde dichas etapas pueden tener lugar en cualquier punto después de la etapa i).
- 55 En una realización alternativa, los polinucleótidos exógenos se expresan de manera transitoria en la célula.
- En otro aspecto, la presente invención proporciona una célula o planta transgénica obtenida usando el método de la invención o la progenie de la misma.
- 60 En otro aspecto, la presente invención proporciona el uso de uno o más polinucleótidos exógenos que codifican
- i) una MGAT,  
 ii) una MGAT y una GPAT  
 iii) una MGAT y una DGAT, o  
 iv) una MGAT, una GPAT y una DGAT,
- 65

para producir un organismo transgénico no humano o una parte del mismo, con mayor capacidad para producir uno o más lípidos no polares en comparación con un organismo correspondiente o una parte del mismo, que no contenga dichos uno o más polinucleótidos exógenos, en donde el organismo transgénico no humano es una planta, un alga, una levadura o un hongo y los uno o más lípidos no polares incluyen triacilglicerol (TAG). El grado de dicha mayor capacidad puede ser como se define en la característica (i). El organismo transgénico no humano o una parte del mismo, se define según la característica (ii) o se puede definir por la combinación de las características (i), (ii) y (iii) o de las características (i) y (ii) o de las características (i) y (iii). Dichos uno o más polinucleótidos exógenos pueden comprender una secuencia según se definió anteriormente.

En otro aspecto, la presente invención proporciona una semilla transgénica de una planta, donde la semilla comprende uno o más polinucleótidos exógenos que codifican una monoacilglicerol transferasa (MGAT) y tiene un nivel aumentado de uno o más lípidos no polares y/o un contenido total de lípidos no polares aumentado en comparación con una semilla correspondiente que no contiene a los polinucleótidos exógenos, en donde los uno o más lípidos no polares incluyen triacilglicerol (TAG). El grado de aumento puede ser como se define en la característica (i). La semilla transgénica puede ser de un género y/o de una especie según se define en la característica (ii). La semilla transgénica se puede definir por la combinación de las características (i), (ii) y (iii) o de las características (i) y (ii) o de las características (i) y (iii). Dichos uno o más polinucleótidos exógenos pueden comprender una secuencia según se definió anteriormente.

En otro aspecto, la presente invención proporciona una planta transgénica que produce una semilla de la invención. La planta transgénica puede ser, por ejemplo, *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Trifolium sp.*, *Elaeis guineensis*, *Nicotiana benthamiana*, *Hordeum vulgare*, *Lupinus angustifolius*, *Oryza sativa*, *Oryza glaberrima*, *Camelina sativa*, *Miscanthus x giganteus* o *Miscanthus sinensis*.

También se divulga en el presente documento un método para producir semillas, donde dicho método comprende:

- i) cultivar una planta transgénica de la invención, y
- ii) cosechar las semillas.

En otro aspecto, la presente invención proporciona un proceso de fermentación que comprende las etapas de:

- i) proporcionar un recipiente que contiene una composición líquida que comprende un organismo transgénico no humano de la invención que sea adecuado para la fermentación y los constituyentes necesarios para la fermentación y biosíntesis de ácidos grasos, y
- ii) proporcionar condiciones que conducen a la fermentación de la composición líquida contenida en dicho recipiente.

En otro aspecto, la presente invención proporciona lípidos extraídos que se pueden obtener con un método de la invención, un organismo transgénico no humano de la invención, una célula transgénica de la invención, una planta transgénica de la invención o una semilla transgénica de la invención. El lípido extraíble tiene un mayor contenido de TAG, y puede tener además un contenido de DAG, contenido de TAG y DAG, contenido de MAG, contenido de PUFA o contenido de un PUFA específico y/o contenido total de lípidos no polares aumentado. El grado de mayor contenido de TAG, contenido de DAG, contenido de TAG y DAG, contenido de MAG, contenido de PUFA, contenido de un PUFA específico y/o contenido total de lípidos no polares puede ser como se define en la característica (i).

También se divulga en el presente documento un lípido extraído que comprende un contenido de DAG que es de al menos un 1 % (p/p) o al menos un 2 % (p/p), sobre una base en peso de los lípidos totales extraídos. Preferentemente, los lípidos extraídos también comprenden MAG a niveles detectables. En una realización, el lípido extraído es aceite de colza, maíz, soja, lupina, cacahuete, girasol, algodón, cártamo o lino. El lípido puede comprender ácido erúico, ácidos grasos ciclopropanoides (CPFA) y/o glucosinolatos a niveles detectables.

La presente invención también proporciona el uso de un organismo transgénico no humano o una parte del mismo, de la invención, una célula transgénica de la invención, una planta transgénica de la invención, una semilla transgénica de la invención o un lípido extraído de la invención para la elaboración de un producto industrial. El producto industrial puede ser, por ejemplo, combustible. El combustible puede comprender un nivel mínimo de los lípidos extraídos de acuerdo con la invención tal como al menos un 10 %, al menos un 20 % o al menos un 30 % (p/p).

En otro aspecto, la presente invención proporciona un método para producir combustible, donde dicho método comprende:

- i) hacer reaccionar un lípido de la invención con un alcohol, opcionalmente, en la presencia de un catalizador, para producir alquil ésteres, y
- ii) opcionalmente, mezclar los alquil ésteres con combustible basado en petróleo.

En una realización, los alquil ésteres son metil ésteres. El combustible producido mediante el método puede comprender un nivel mínimo del lípido de la invención, tal como al menos un 10 %, al menos un 20 % o al menos un 30 % (p/p).

5 En otro aspecto, la presente invención proporciona un método para producir un alimento, donde dicho método comprende mezclar un organismo transgénico no humano o una parte del mismo, de la invención, una célula transgénica de la invención, una planta transgénica de la invención, una semilla transgénica de la invención o un lípido extraído de la invención o un extracto o una porción del mismo, con al menos un ingrediente alimentario adicional.

10 En otro aspecto, la presente invención proporciona alimentos, cosméticos o sustancias químicas que comprenden un organismo transgénico no humano o una parte del mismo, de la invención, una célula transgénica de la invención, una planta transgénica de la invención, una semilla transgénica de la invención o un lípido extraído de la invención o un extracto o una porción del mismo.

15 También se divulga en el presente documento un método para identificar una molécula de ácido nucleico que codifica una aciltransferasa que tiene mayor capacidad para producir MAG, DAG y/o TAG en una célula, donde dicho método comprende:

- 20 i) obtener una célula que comprende una molécula de ácido nucleico ligada operativamente a un promotor que es activo en la célula, en donde la molécula de ácido nucleico comprende una secuencia de nucleótidos definida en el presente documento y/o una secuencia de nucleótidos que codifica un polipéptido que tiene una o más de las secuencias de aminoácidos provistas en las SEQ ID NO: 220, 221, 222, 223, 224, 225, 226 y 227 o una secuencia de aminoácidos que es al menos un 50 %, preferentemente al menos un 60 %, más preferentemente
- 25 al menos un 65 %, idéntica a las mismas,
- ii) determinar si el nivel de MAG, DAG y/o TAG producido en la célula está aumentado en comparación con una célula correspondiente que no contiene al ácido nucleico, y
- iii) identificar una molécula de ácido nucleico que codifica una aciltransferasa que tiene mayor capacidad para producir MAG, DAG y/o TAG en la célula.

30 La molécula de ácido nucleico puede comprender una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos conservadas de DGAT2 y/o MGAT1/2 provistas en las SEQ ID NO: 220, 221, 222, 223 y 224. La molécula de ácido nucleico puede comprender una secuencia de nucleótidos que codifica las secuencias de aminoácidos conservadas provistas en la SEQ ID NO: 220 y/o en la SEQ ID NO: 224. La molécula de ácido nucleico puede comprender una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos conservadas de GPAT provistas en las SEQ ID NO: 225, 226 y 227 o una secuencia de aminoácidos que es al menos un 50 %, preferentemente al menos un 60 %, más preferentemente al menos un 65 %, idéntica a las mismas.

40 El método puede ser un método que sirve para identificar una MGAT que tiene mayor capacidad para producir DAG y/o TAG.

También se divulga en el presente documento un método que sirve para identificar una GPAT que tiene actividad fosfatasa y una mayor capacidad para producir MAG. El método puede servir para identificar una *sn-2* GPAT que tiene actividad fosfatasa y una mayor capacidad para producir 2-MAG.

45 El método puede ser un método que sirve para identificar una MGAT y una GPAT que juntas tienen una mayor capacidad para producir MAG y/o DAG y/o TAG, preferentemente TAG. Por ejemplo, el método sirve para identificar una *sn-2* GPAT que tiene actividad fosfatasa para producir 2-MAG y junto con la MGAT, una mayor capacidad para producir *sn-1,2/2,3* DAG y/o TAG.

50 El método puede ser un método que sirve para identificar una DGAT que tiene una mayor capacidad para producir TAG.

55 El método puede ser un método que sirve para identificar una MGAT y una DGAT que juntas tienen una mayor capacidad para producir DAG y/o TAG, preferentemente TAG.

60 El método puede ser un método que sirve para identificar una MGAT, una GPAT y una DGAT que juntas tienen una mayor capacidad para producir MAG y/o DAG y/o TAG, preferentemente TAG. La GPAT puede tener actividad fosfatasa para producir MAG. La GPAT puede ser una *sn-2* GPAT que tiene actividad fosfatasa para producir 2-MAG.

El método puede comprender además introducir la molécula de ácido nucleico en una célula antes de la etapa i).

65 El método puede comprender además la etapa de regenerar una planta transgénica a partir de la célula de la etapa i) y, opcionalmente, obtener semillas de la planta transgénica. La célula puede ser una célula vegetal. El grado de mayor capacidad para producir MAG y/o DAG y/o TAG puede ser como se define en la característica (i). La célula

puede ser de un género y/o de una especie según se define en la característica (ii). La célula se puede definir por la combinación de las características (i), (ii) y (iii) o de las características (i) y (ii) o de las características (i) y (iii).

5 Tal como se divulga en el presente documento, la MGAT y/o la GPAT y/o la DGAT puede aumentar la producción de TAG en la célula en una mayor cantidad que la DGAT1 de *Arabidopsis thaliana* (SEQ ID NO: 83).

También se divulga en el presente documento un polinucleótido aislado y/o recombinante que comprende:

- 10 i) una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO:1 a 6, o  
ii) una secuencia de nucleótidos que es al menos un 80 % idéntica a i).

También se divulga en el presente documento un vector que comprende un polinucleótido de la invención.

15 También se divulga en el presente documento una célula transgénica que comprende un polinucleótido de la invención o un vector de la invención.

20 En otro aspecto, la invención proporciona un organismo transgénico no humano o una parte del mismo, que comprende un polinucleótido de la invención, un vector de la invención o una célula transgénica de la invención. El organismo transgénico es una planta, un alga o un organismo adecuado para la fermentación, tal como una levadura o un hongo. En otra realización, la parte es una semilla, un fruto, un tubérculo, una raíz o una parte vegetativa de una planta.

25 Los expertos en la materia comprenderán que se pueden realizar numerosas variaciones y/o modificaciones a la invención, tal como se ilustra en las realizaciones específicas, sin apartarse del espíritu o alcance de la invención, descrita ampliamente. Las presentes realizaciones deben, por lo tanto, considerarse en todo sentido a modo ilustrativo y no restrictivo. Cada una de las realizaciones descritas en el presente documento se aplicará *mutatis mutandis* a cada una y a todas las demás realizaciones a menos que específicamente se defina de otra manera.

30 En toda la memoria descriptiva, la expresión "que comprende" o variaciones tales como "comprende" o "comprenden", incluye un elemento, un número o una etapa mencionada o un grupo de elementos, números o etapas, pero no excluyen cualquier otro elemento, número o etapa o grupo de elementos, números o etapas.

35 El término "y/o", por ejemplo, "X y/o Y" significarán ya sea "X e Y" o "X o Y" y se considerarán como un soporte explícito de ambos significados o de cualquiera de dichos significados.

#### Breve descripción de las figuras adjuntas

40 **Figura 1.** Una representación de diversas vías de síntesis de lípidos, la mayoría de las cuales convergen en DAG, una molécula central en la síntesis de lípidos. Este modelo incluye una ruta posible para la formación de *sn*-2 MAG que se podría usar con una MGAT/DGAT bifuncional para la formación de DAG a partir de glicerol-3-fosfato (G-3-P). Las abreviaturas empleadas son:

- 45 G-3-P; glicerol-3-fosfato  
LisoPA; ácido lisofosfatídico  
PA; ácido fosfatídico  
MAG; monoacilglicerol  
DAG; diacilglicerol  
TAG; triacilglicerol  
50 Acil-CoA y FA-CoA; acil-coenzima A y acilo graso-coenzima A  
PC; fosfatidilcolina  
GPAT; glicerol-3-fosfato aciltransferasa; glicerol-3-fosfato O-aciltransferasa; acil-CoA:*sn*-glicerol-3-fosfato 1-O-aciltransferasa; EC 2,3.1.15  
GPAT4; glicerol-3-fosfato aciltransferasa 4  
GPAT6; glicerol-3-fosfato aciltransferasa 6  
55 LPAAT; 1-acil-glicerol-3-fosfato aciltransferasa; 1-acilglicerol-3-fosfato O-aciltransferasa; acil-CoA:1-acil-*sn*-glicerol-3-fosfato 2-O-aciltransferasa; EC 2,3.1.51  
PAP; ácido fosfatídico fosfatasa; fosfatidato fosfatasa; ácido fosfático fosfohidrolasa; ácido fosfatídico fosfatasa; EC 3.1.3.4  
MGAT; una aciltransferasa que tiene actividad monoacilglicerol aciltransferasa (MGAT; 2-acilglicerol O-aciltransferasa acil-CoA:2-acilglicerol O-aciltransferasa; EC 2,3.1.22)  
60 M/DGAT; una aciltransferasa que tiene actividad monoacilglicerol aciltransferasa (MGAT; 2-acilglicerol O-aciltransferasa; acil-CoA:2-acilglicerol O-aciltransferasa; EC 2,3.1.22) y/o diacilglicerol aciltransferasa (DGAT; diacilglicerol O-aciltransferasa; acil-CoA:1,2-diacil-*sn*-glicerol O-aciltransferasa; EC 2,3.1.20)  
LPCAT; acil-CoA:lisofosfatidilcolina aciltransferasa; 1-acilglicerofosfocolina O-aciltransferasa; acil-CoA:1-acil-*sn*-glicerol-3-fosfocolina O-aciltransferasa; EC 2,3.1.23  
65 PLD-Z; Fosfolipasa D zeta; colina fosfatasa; lecitinasa D; lipofosfodiesterasa II; EC 3.1.4.4

CPT; CDP-colina:diacilglicerol colina fosfotransferasa; 1-alkil-2-acetilglicerol colina fosfotransferasa; alkilacilglicerol colina fosfotransferasa; colina fosfotransferasa; fosforilcolina-glicérido transferasa; EC 2.7.8.2  
 PDCT; fosfatidilcolina: diacilglicerol colina fosfotransferasa  
 PLC; fosfolipasa C; EC 3.1.4.3  
 5 PDAT; fosfolípido:diacilglicerol aciltransferasa; fosfolípido:1,2-diacil-*sn*-glicerol O-aciltransferasa; EC 2,3.1.158  
 Pi; fosfato inorgánico

**Figura 2.** Incrementos relativos de DAG y TAG en tejido foliar de *Nicotiana benthamiana* transformado con construcciones que codifican p19 (control negativo), DGAT1 de *Arabidopsis thaliana*, MGAT1 de *Mus musculus* y una combinación de DGAT1 y MGAT1, cada una expresada a partir del promotor 35S. La enzima MGAT1 era mucho más activa que la enzima DGAT1 en la promoción de la acumulación de ambos DAG y TAG en el tejido foliar. La expresión del gen de MGAT1 dio como resultado el doble de acumulación de DAG y TAG en el tejido foliar en comparación con la expresión de DGAT1 solamente.

**Figura 3.** Incrementos relativos de TAG en hojas de *N. benthamiana* transformadas con construcciones que codifican p19 (control negativo), DGAT1 de *A. thaliana*, MGAT2 de *M. musculus* y una combinación de MGAT2 y DGAT1. Las barras de error indican el error estándar de muestras triplicadas.

**Figura 4.** Radioactividad (DPM) en fracciones de MAG, DAG y TAG aisladas de lisados de hojas de *N. benthamiana* transformadas transitoriamente alimentadas con *sn*-2-MAG [<sup>14</sup>C] y ácido oleico no marcado en un transcurso de tiempo. Las construcciones usadas eran las mismas que las usadas para la Figura 3.

**Figura 5.** Igual que la Figura 4 pero alimentadas con [<sup>14</sup>C]G-3-P y ácido oleico no marcado.

**Figura 6.** Autorradiograma de una placa de TLC que muestra la formación de TAG por DGAT1 de *A. thaliana* y MGAT1 de *M. musculus* pero no MGAT2 de *M. musculus* en ensayos de levadura. El ensayo se describe en el Ejemplo 5.

### CLAVE PARA EL LISTADO DE SECUENCIAS

SEQ ID NO: 1 MGAT1 optimizada por codones de *Mus musculus*  
 SEQ ID NO: 2 MGAT2 optimizada por codones de *Mus musculus*  
 SEQ ID NO: 3 MGAT1 predicha optimizada por codones de *Ciona intestinalis*  
 35 SEQ ID NO: 4 MGAT1 predicha optimizada por codones de *Tribolium castaneum*  
 SEQ ID NO: 5 MGAT1 optimizada por codones de *Danio rerio*  
 SEQ ID NO: 6 MGAT2 optimizada por codones de *Danio rerio*  
 SEQ ID NO: 7 polinucleótido de MGAT1 de *Mus musculus* (AF384162)  
 SEQ ID NO: 8 polinucleótido de MGAT1 de *Homo sapiens* (AF384163)  
 40 SEQ ID NO: 9 variante de transcrito del polinucleótido predicho de MGAT1 de *Pan troglodytes* (XM\_001166055)  
 SEQ ID NO: 10 variante de transcrito del polinucleótido predicho de MGAT1 de *Pan troglodytes* (XM\_0526044.2)  
 SEQ ID NO: 11 polinucleótido predicho de MGAT1 de *Canis familiaris* (XM\_545667.2)  
 SEQ ID NO: 12 polinucleótido de MGAT1 de *Bos taurus* (NM\_001001153.2)  
 SEQ ID NO: 13 polinucleótido de MGAT1 de *Rattus norvegicus* (NM\_001108803.1)  
 45 SEQ ID NO: 14 polinucleótido de MGAT1 de *Danio rerio* (NM\_001122623.1)  
 SEQ ID NO: 15 polinucleótido predicho de MGAT1 de *Caenorhabditis elegans* (NM\_073012.4)  
 SEQ ID NO: 16 polinucleótido predicho de MGAT1 de *Caenorhabditis elegans* (NM\_182380.5)  
 SEQ ID NO: 17 polinucleótido predicho de MGAT1 de *Caenorhabditis elegans* (NM\_065258.3)  
 SEQ ID NO: 18 polinucleótido predicho de MGAT1 de *Caenorhabditis elegans* (NM\_075068.3)  
 50 SEQ ID NO: 19 polinucleótido predicho de MGAT1 de *Caenorhabditis elegans* (NM\_072248.3)  
 SEQ ID NO: 20 polinucleótido predicho de MGAT1 de *Kluyveromyces lactis* (XM\_455588.1)  
 SEQ ID NO: 21 polinucleótido predicho de MGAT1 de *Ashbya gossypii* (NM\_208895.1)  
 SEQ ID NO: 22 polinucleótido predicho de MGAT1 de *Magnaporthe oryzae* (XM\_368741.1)  
 SEQ ID NO: 23 polinucleótido predicho de MGAT1 de *Ciona intestinalis* (XM\_002120843.1)  
 55 SEQ ID NO: 24 polinucleótido de MGAT2 de *Mus musculus* (AY157609)  
 SEQ ID NO: 25 polinucleótido de MGAT2 de *Homo sapiens* (AY157608)  
 SEQ ID NO: 26 polinucleótido predicho de MGAT2 de *Pan troglodytes* (XM\_522112.2)  
 SEQ ID NO: 27 polinucleótido predicho de MGAT2 de *Canis familiaris* (XM\_542304.1)  
 SEQ ID NO: 28 polinucleótido de MGAT2 de *Bos taurus* (NM\_001099136.1)  
 60 SEQ ID NO: 29 polinucleótido de MGAT2 de *Rattus norvegicus* (NM\_001109436.2)  
 SEQ ID NO: 30 polinucleótido predicho de MGAT2 de *Gallus gallus* (XM\_424082.2)  
 SEQ ID NO: 31 polinucleótido de MGAT2 de *Danio rerio* (NM\_001006083.1)  
 SEQ ID NO: 32 polinucleótido de MGAT2 de *Drosophila melanogaster* (NM\_136474.2)  
 SEQ ID NO: 33 polinucleótido de MGAT2 de *Drosophila melanogaster* (NM\_136473.2)  
 65 SEQ ID NO: 34 polinucleótido de MGAT2 de *Drosophila melanogaster* (NM\_136475.2)  
 SEQ ID NO: 35 polinucleótido de MGAT2 de *Anopheles gambiae* (XM\_001688709.1)

## ES 2 640 100 T3

SEQ ID NO: 36 polinucleótido de MGAT2 de *Anopheles gambiae* (XM\_315985)  
 SEQ ID NO: 37 polinucleótido predicho de MGAT2 de *Tribolium castaneum* (XM\_970053.1)  
 SEQ ID NO: 38 polinucleótido de MGAT3 de *Homo sapiens* (AY229854)  
 SEQ ID NO: 39 variante de transcrito 1 del polinucleótido predicho de MGAT3 de *Pan troglodytes*  
 5 (XM\_001154107.1)  
 SEQ ID NO: 40 variante de transcrito 2 del polinucleótido predicho de MGAT3 de *Pan troglodytes*  
 (XM\_001154171.1)  
 SEQ ID NO: 41 variante de transcrito 3 del polinucleótido predicho de MGAT3 de *Pan troglodytes* (XM\_527842,2)  
 SEQ ID NO: 42 polinucleótido predicho de MGAT3 de *Canis familiaris* (XM\_845212.1)  
 10 SEQ ID NO: 43 polinucleótido predicho de MGAT3 de *Bos taurus* (XM\_870406.4)  
 SEQ ID NO: 44 polinucleótido predicho de MGAT3 de *Danio rerio* (XM\_688413.4)  
 SEQ ID NO: 45 polipéptido de MGAT1 de *Mus musculus* (AAK84177.1)  
 SEQ ID NO: 46 polipéptido de MGAT1 de *Homo sapiens* (AAK84178.1)  
 SEQ ID NO: 47 isoforma 1 del polipéptido predicho de MGAT1 de *Pan troglodytes* (XP\_001166055.1)  
 15 SEQ ID NO: 48 isoforma 2 del polipéptido predicho de MGAT1 de *Pan troglodytes* (XP\_526044,2)  
 SEQ ID NO: 49 polipéptido predicho de MGAT1 de *Canis familiaris* (XP\_545667,2)  
 SEQ ID NO: 50 polipéptido de MGAT1 de *Bos taurus* (NP\_001001153.1)  
 SEQ ID NO: 51 polipéptido de MGAT1 de *Rattus norvegicus* (NP\_001102273.1)  
 SEQ ID NO: 52 polipéptido de MGAT1 de *Danio rerio* (NP\_001116095.1)  
 20 SEQ ID NO: 53 polipéptido predicho de MGAT1 de *Caenorhabditis elegans* (NP\_505413.1)  
 SEQ ID NO: 54 polipéptido predicho de MGAT1 de *Caenorhabditis elegans* (NP\_872180.1)  
 SEQ ID NO: 55 polipéptido predicho de MGAT1 de *Caenorhabditis elegans* (NP\_497659.1)  
 SEQ ID NO: 56 polipéptido predicho de MGAT1 de *Caenorhabditis elegans* (NP\_507469.1)  
 SEQ ID NO: 57 polipéptido predicho de MGAT1 de *Caenorhabditis elegans* (NP\_504649.1)  
 25 SEQ ID NO: 58 polipéptido predicho de MGAT1 de *Kluyveromyces lactis* (XP\_455588.1)  
 SEQ ID NO: 59 polipéptido predicho de MGAT1 de *Ashbya gossypii* (NP\_983542.1)  
 SEQ ID NO: 60 polipéptido predicho de MGAT1 de *Magnaporthe oryzae* (XP\_368741.1)  
 SEQ ID NO: 61 polipéptido predicho de MGAT1 de *Ciona intestinalis* (XP\_002120879)  
 SEQ ID NO: 62 polipéptido de MGAT2 de *Mus musculus* (AA023673.1)  
 30 SEQ ID NO: 63 polipéptido de MGAT2 de *Homo sapiens* (AA023672.1)  
 SEQ ID NO: 64 polipéptido predicho de MGAT2 de *Pan troglodytes* (XP\_522112,2)  
 SEQ ID NO: 65 polipéptido predicho de MGAT2 de *Canis familiaris* (XP\_542304.1)  
 SEQ ID NO: 66 polipéptido de MGAT2 de *Bos taurus* (NP\_001092606.1)  
 SEQ ID NO: 67 polipéptido de MGAT2 de *Rattus norvegicus* (NP\_001102906,2)  
 35 SEQ ID NO: 68 polipéptido predicho de MGAT2 de *Gallus gallus* (XP\_424082,2)  
 SEQ ID NO: 69 polipéptido de MGAT2 de *Danio rerio* (NP\_001006083.1)  
 SEQ ID NO: 70 polipéptido de MGAT2 de *Drosophila melanogaster* (NP\_610318.1)  
 SEQ ID NO: 71 polipéptido de MGAT2 de *Drosophila melanogaster* (NP\_610317.1)  
 SEQ ID NO: 72 polipéptido de MGAT2 de *Drosophila melanogaster* (NP\_610319,2)  
 40 SEQ ID NO: 73 polipéptido de MGAT2 de *Anopheles gambiae* (XP\_001688761)  
 SEQ ID NO: 74 polipéptido de MGAT2 de *Anopheles gambiae* (XP\_315985,3)  
 SEQ ID NO: 75 polipéptido predicho de MGAT2 de *Tribolium castaneum* (XP\_975146)  
 SEQ ID NO: 76 polipéptido de MGAT3 de *Homo sapiens* (AA063579.1)  
 SEQ ID NO: 77 isoforma 1 del polipéptido predicho de MGAT3 de *Pan troglodytes* (XP\_001154107.1)  
 45 SEQ ID NO: 78 isoforma 2 del polipéptido predicho de MGAT3 de *Pan troglodytes* (XP\_001154171.1)  
 SEQ ID NO: 79 isoforma 3 de MGAT3 predicha de *Pan troglodytes* (XP\_527842,2)  
 SEQ ID NO: 80 polipéptido predicho de MGAT3 de *Canis familiaris* (XP\_850305.1)  
 SEQ ID NO: 81 polipéptido predicho de MGAT3 de *Bos taurus* (XP\_875499,3)  
 SEQ ID NO: 82 polipéptido predicho de MGAT3 de *Danio rerio* (XP\_693505.1)  
 50 SEQ ID NO: 83 polipéptido de DGAT1 de *Arabidopsis thaliana* (CAB44774.1)  
 SEQ ID NO: 84 polinucleótido de GPAT4 de *Arabidopsis thaliana* (NM\_100043.4)  
 SEQ ID NO: 85 polinucleótido de GPAT6 de *Arabidopsis thaliana* (NM\_129367,3)  
 SEQ ID NO: 86 polinucleótido de BAC F510 de *Arabidopsis thaliana* (AF195115.1)  
 SEQ ID NO: 87 polinucleótido de proteína desconocida de *Arabidopsis thaliana* (AY062466.1)  
 55 SEQ ID NO: 88 polinucleótido del cromosoma 3 de *Oryza sativa* (AC118133.4)  
 SEQ ID NO: 89 polinucleótido del clon WS0276\_F13 de *Picea sitchensis* (EF086095.1)  
 SEQ ID NO: 90 polinucleótido del clon ZM\_BFc0110A1 de *Zea mays* (BT067649.1)  
 SEQ ID NO: 91 polinucleótido del clon RAFL16-19-H05 de *Arabidopsis thaliana* (AK228870.1)  
 SEQ ID NO: 92 clon J065058J01 polinucleótido de *Oryza sativa* (AK241033.1)  
 60 SEQ ID NO: 93 polinucleótido del cromosoma 2 de *Oryza sativa* (CM000127.1)  
 SEQ ID NO: 94 cromosoma 5 polinucleótido de *Oryza sativa* (CM000130.1)  
 SEQ ID NO: 95 polinucleótido del cromosoma 2 de *Oryza sativa* (CM000139.1)  
 SEQ ID NO: 96 polinucleótido del cromosoma 1 de *Oryza sativa* (CM000126.1)  
 SEQ ID NO: 97 polinucleótido del cromosoma 3 de *Oryza sativa* (CM000128.1)  
 65 SEQ ID NO: 98 polinucleótido del cromosoma 3 de *Oryza sativa* (CM000140.1)  
 SEQ ID NO: 99 polinucleótido de SELMOscaffold\_102 de *Selaginella moellendorffii* (GL377667.1)

## ES 2 640 100 T3

SEQ ID NO: 100 polinucleótido de SELMOscaffold\_102 de *Selaginella moellendorffii* (GL377667.1)  
 SEQ ID NO: 101 polinucleótido de SELMOscaffold\_83 de *Selaginella moellendorffii* (GL377648.1)  
 SEQ ID NO: 102 polinucleótido de SELMOscaffold\_57 de *Selaginella moellendorffii* (GL377622.1)  
 SEQ ID NO: 103 polinucleótido de SELMOscaffold\_25 de *Selaginella moellendorffii* (GL377590.1)  
 5 SEQ ID NO: 104 polinucleótido de SELMOscaffold\_11 de *Selaginella moellendorffii* (GL377576.1)  
 SEQ ID NO: 105 polinucleótido de SELMOscaffold\_11 de *Selaginella moellendorffii* (GL377576.1)  
 SEQ ID NO: 106 polinucleótido de Os01g0855000 de *Oryza sativa* (NM\_001051374.2)  
 SEQ ID NO: 107 polinucleótido de Os02g0114400 de *Oryza sativa* (NM\_001052203.1)  
 SEQ ID NO: 108: polinucleótido de GPAT8 de *Zea mays* (NM\_001153970.1)  
 10 SEQ ID NO: 109: polinucleótido LOC100282930 de *Zea mays* (NM\_001155835.1)  
 SEQ ID NO: 110: polinucleótido LOC100382119 de *Zea mays* (NM\_001174880.1)  
 SEQ ID NO: 111 polinucleótido de GPAT6 de *Arabidopsis thaliana* (NM\_129367.3)  
 SEQ ID NO: 112 polinucleótido de GPAT8 de *Arabidopsis thaliana* (NM\_116264.5)  
 SEQ ID NO: 113 polinucleótido (PHYPADRAFT\_128739) de una proteína predicha de *Physcomitrella patens*  
 15 (XM\_001764949.1)  
 SEQ ID NO: 114 polinucleótido (PHYPADRAFT\_83824) de una proteína predicha de *Physcomitrella patens*  
 (XM\_001769619.1)  
 SEQ ID NO: 115 polinucleótido (PHYPADRAFT\_188308) de una proteína predicha de *Physcomitrella patens*  
 (XM\_001769672.1)  
 20 SEQ ID NO: 116 polinucleótido (PHYPADRAFT\_189499) de una proteína predicha de *Physcomitrella patens*  
 (XM\_001771134.1)  
 SEQ ID NO: 117 polinucleótido (PHYPADRAFT\_95487) de una proteína predicha de *Physcomitrella patens*  
 (XM\_001780481.1)  
 SEQ ID NO: 118 polinucleótido LOC100243321 de una proteína predicha de *Vitis Vinifera* (XM\_002268477.1)  
 25 SEQ ID NO: 119 polinucleótido LOC100243093 de una proteína predicha de *Vitis Vinifera* (XM\_002275312.1)  
 SEQ ID NO: 120 polinucleótido LOC100259433 de una proteína predicha de *Vitis Vinifera* (XM\_002275996.1)  
 SEQ ID NO: 121 polinucleótido LOC100264832 de una proteína predicha de *Vitis Vinifera* (XM\_002279055.1)  
 SEQ ID NO: 122 polinucleótido de una proteína predicha de *Populus trichocarpa* (XM\_002309088.1)  
 SEQ ID NO: 123 polinucleótido de una proteína predicha de *Populus trichocarpa* (XM\_002309240.1)  
 30 SEQ ID NO: 124 polinucleótido de una proteína predicha de *Populus trichocarpa* (XM\_002322716.1)  
 SEQ ID NO: 125 polinucleótido de una proteína predicha de *Populus trichocarpa* (XM\_002323527.1)  
 SEQ ID NO: 126 polinucleótido de una proteína de *Sorghum bicolor* (XM\_002439842.1)  
 SEQ ID NO: 127 polinucleótido de una proteína de *Sorghum bicolor* (XM\_002458741.1)  
 SEQ ID NO: 128 polinucleótido de una proteína de *Sorghum bicolor* (XM\_002463871.1)  
 35 SEQ ID NO: 129 polinucleótido de una proteína de *Sorghum bicolor* (XM\_002464585.1)  
 SEQ ID NO: 130 polinucleótido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis*  
 (XM\_002511827.1)  
 SEQ ID NO: 131 polinucleótido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis*  
 (XM\_002517392.1)  
 40 SEQ ID NO: 132 polinucleótido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis*  
 (XM\_002520125.1)  
 SEQ ID NO: 133 polinucleótido de una proteína de la familia de fosfolípido/glicerol aciltransferasa de *Arabidopsis*  
*lyrata* (XM\_002872909.1)  
 SEQ ID NO: 134 polinucleótido de GPAT6 de *Arabidopsis lyrata* (XM\_002881518.1)  
 45 SEQ ID NO: 135 polinucleótido de GPAT8 de *Vernicia fordii putative* (FJ479753.1)  
 SEQ ID NO: 136 polinucleótido Os03g0735900 de *Oryza sativa* (NM\_001057724.1)  
 SEQ ID NO: 137 polinucleótido de GPAT4 de *Arabidopsis thaliana* (NM\_100043.4)  
 SEQ ID NO: 138 polinucleótido de una proteína predicha de *Populus trichocarpa* (XM\_002320102.1)  
 SEQ ID NO: 139 polinucleótido de una proteína de *Sorghum bicolor* (XM\_002451332.1)  
 50 SEQ ID NO: 140 polinucleótido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis*  
 (XM\_002531304.1)  
 SEQ ID NO: 141 polinucleótido de GPAT4 de *Arabidopsis lyrata* (XM\_002889315.1)  
 SEQ ID NO: 142 polinucleótido de GPAT1 de *Arabidopsis thaliana* (NM\_100531.2)  
 SEQ ID NO: 143 polinucleótido de GPAT3 de *Arabidopsis thaliana* (NM\_116426.2)  
 55 SEQ ID NO: 144 polipéptido de GPAT4 de *Arabidopsis thaliana* (NP\_171667.1)  
 SEQ ID NO: 145 polipéptido de GPAT6 de *Arabidopsis thaliana* (NP\_181346.1)  
 SEQ ID NO: 146 polipéptido del producto genético F5I10.4 de *Arabidopsis thaliana* (AAF02784.1)  
 SEQ ID NO: 147 polipéptido de una proteína desconocida de *Arabidopsis thaliana* (AAL32544.1)  
 SEQ ID NO: 148 polipéptido de una proteína de *Oryza sativa* (AAP03413.1)  
 60 SEQ ID NO: 149 polipéptido desconocido de *Picea sitchensis* (ABK25381.1)  
 SEQ ID NO: 150 polipéptido desconocido de *Zea mays* (ACN34546.1)  
 SEQ ID NO: 151 polipéptido de proteína predicha de *Arabidopsis thaliana* (BAF00762.1)  
 SEQ ID NO: 152 polipéptido de producto proteico sin nombre de *Oryza sativa* (BAH00933.1)  
 SEQ ID NO: 153 polipéptido Osl\_05566 de una proteína predicha de *Oryza sativa* (EAY84189.1)  
 65 SEQ ID NO: 154 polipéptido Osl\_20155 de una proteína predicha de *Oryza sativa* (EAY98245.1)  
 SEQ ID NO: 155 polipéptido OsJ\_05094 de una proteína predicha de *Oryza sativa* (EAZ21484.1)

## ES 2 640 100 T3

SEQ ID NO: 156 polipéptido Osl\_04478 de una proteína predicha de *Oryza sativa* (EEC71826.1)  
 SEQ ID NO: 157 polipéptido Osl\_13423 de una proteína predicha de *Oryza sativa* (EEC76137.1)  
 SEQ ID NO: 158 polipéptido OsJ\_12482 de una proteína predicha de *Oryza sativa* (EEE59882.1)  
 5 SEQ ID NO: 159 polipéptido SELMODRAFT\_269600 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ08963.1)  
 SEQ ID NO: 160 polipéptido SELMODRAFT\_184962 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ08964.1)  
 SEQ ID NO: 161 polipéptido SELMODRAFT\_235331 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ11200.1)  
 10 SEQ ID NO: 162 polipéptido SELMODRAFT\_118155 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ15664.1)  
 SEQ ID NO: 163 polipéptido SELMODRAFT\_102257 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ24086.1)  
 SEQ ID NO: 164 polipéptido SELMODRAFT\_170164 de una proteína predicha de *Selaginella moellendorffii*  
 15 (EFJ29816.1)  
 SEQ ID NO: 165 polipéptido SELMODRAFT\_170165 de una proteína predicha de *Selaginella moellendorffii*  
 (EFJ29817.1)  
 SEQ ID NO: 166 polipéptido Os01g0855000 de *Oryza sativa* (NP\_001044839.1)  
 SEQ ID NO: 167 polipéptido Os02g0114400 de *Oryza sativa* (NP\_001045668.1)  
 20 SEQ ID NO: 168 polipéptido de GPAT 8 de *Zea mays* (NP\_001147442.1)  
 SEQ ID NO: 169 polipéptido LOC100282930 de *Zea mays* (NP\_001149307.1)  
 SEQ ID NO: 170 polipéptido de una proteína LOC100382119 de *Zea mays* (NP\_001168351.1)  
 SEQ ID NO: 171 polipéptido de GPAT6 de *Arabidopsis thaliana* (NP\_181346.1)  
 SEQ ID NO: 172 polipéptido de GPAT8 de *Arabidopsis thaliana* (NP\_191950,2)  
 25 SEQ ID NO: 173 polipéptido de una proteína de *Physcomitrella patens* (XP\_001765001.1)  
 SEQ ID NO: 174 polipéptido de una proteína de *Physcomitrella patens* (XP\_001769671.1)  
 SEQ ID NO: 175 polipéptido de una proteína de *Physcomitrella patens* (XP\_001769724.1)  
 SEQ ID NO: 176 polipéptido de una proteína de *Physcomitrella patens* (XP\_001771186.1)  
 SEQ ID NO: 177 polipéptido de una proteína de *Physcomitrella patens* (XP\_001780533.1)  
 30 SEQ ID NO: 178 polipéptido de una proteína de *Vitis vinifera* (XP\_002268513.1)  
 SEQ ID NO: 179 polipéptido de una proteína de *Vitis vinifera* (XP\_002275348.1)  
 SEQ ID NO: 180 polipéptido de una proteína de *Vitis vinifera* (XP\_002276032.1)  
 SEQ ID NO: 181 polipéptido de una proteína de *Vitis vinifera* (XP\_002279091.1)  
 SEQ ID NO: 182 polipéptido de una proteína de *Populus trichocarpa* (XP\_002309124.1)  
 35 SEQ ID NO: 183 polipéptido de una proteína de *Populus trichocarpa* (XP\_002309276.1)  
 SEQ ID NO: 184 polipéptido de una proteína de *Populus trichocarpa* (XP\_002322752.1)  
 SEQ ID NO: 185 polipéptido de una proteína de *Populus trichocarpa* (XP\_002323563.1)  
 SEQ ID NO: 186 polipéptido de una proteína SORBIDRAFT\_09g022020 de *Sorghum bicolor* (XP\_002439887.1)  
 SEQ ID NO: 187 polipéptido de una proteína SORBIDRAFT\_03g040260 de *Sorghum bicolor* (XP\_002458786.1)  
 40 SEQ ID NO: 188 polipéptido de una proteína SORBIDRAFT\_01g008880 de *Sorghum bicolor* (XP\_002463916.1)  
 SEQ ID NO: 189 polipéptido de una proteína SORBIDRAFT\_01g022140 de *Sorghum bicolor* (XP\_002464630.1)  
 SEQ ID NO: 190 polipéptido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis* (XP\_002511873.1)  
 SEQ ID NO: 191 polipéptido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis* (XP\_002517438.1)  
 SEQ ID NO: 192 polipéptido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis* (XP\_002520171.1)  
 45 SEQ ID NO: 193 polipéptido de una proteína de la familia de fosfolípido/glicerol aciltransferasa de *Arabidopsis*  
*lyrata* (XP\_002872955.1)  
 SEQ ID NO: 194 polipéptido de GPAT6 de *Arabidopsis lyrata* (XP\_002881564.1)  
 SEQ ID NO: 195 polipéptido de GPAT de *Vernicia fordii putative* (ACT32032.1)  
 SEQ ID NO: 196 polipéptido Os03g0735900 de *Oryza sativa* (NP\_001051189.1)  
 50 SEQ ID NO: 197 polipéptido de GPAT4 de *Arabidopsis thaliana* (NP\_171667.1)  
 SEQ ID NO: 198 polipéptido de una proteína de *Populus trichocarpa* (XP\_002320138.1)  
 SEQ ID NO: 199 polipéptido de una proteína SORBIDRAFT\_04g001060 de *Sorghum bicolor* (XP\_002451377.1)  
 SEQ ID NO: 200 polipéptido de glicerol-fosfato aciltransferasa de RE de *Ricinus communis* (XP\_002531350.1)  
 SEQ ID NO: 201 polipéptido de GPAT4 de *Arabidopsis lyrata* (XP\_002889361.1)  
 55 SEQ ID NO: 202 polipéptido de GPAT1 de *Arabidopsis thaliana* (NP\_563768.1)  
 SEQ ID NO: 203 polipéptido de GPAT3 de *Arabidopsis thaliana* (NP\_192104.1)  
 SEQ ID NO: 204 polinucleótido de DGAT2 de *Arabidopsis thaliana* (NM\_115011,3)  
 SEQ ID NO: 205 polinucleótido de DGAT2 de *Ricinus communis* (AY916129.1)  
 SEQ ID NO: 206 polinucleótido de DGAT2 de *Vernicia fordii* (DQ356682.1)  
 60 SEQ ID NO: 207 polipéptido de DGAT2 de *Mortierella ramanniana* (AF391089.1)  
 SEQ ID NO: 208 polipéptido de DGAT2 de *Homo sapiens* (NM\_032564.1)  
 SEQ ID NO: 209 polipéptido de DGAT2 de *Homo sapiens* (NM\_001013579,2)  
 SEQ ID NO: 210 polipéptido de DGAT2 de *Bos taurus* (NM\_205793,2)  
 SEQ ID NO: 211 polipéptido de DGAT2 de *Mus musculus* (AF384160.1)  
 65 SEQ ID NO: 212 polipéptido de DGAT2 *Arabidopsis thaliana* (NP\_566952.1)  
 SEQ ID NO: 213 polipéptido de DGAT2 *Ricinus communis* (AAY16324.1)



- SEQ ID NO: 214 polipéptido de DGAT2 de *Vernicia fordii* (ABC94474.1)  
 SEQ ID NO: 215 polipéptido de DGAT2 de *Mortierella ramanniana* (AAK84179.1)  
 SEQ ID NO: 216 polipéptido de DGAT2 de *Homo sapiens* (Q96PD7.2)  
 SEQ ID NO: 217 polipéptido de DGAT2 de *Homo sapiens* (Q58HT5.1)  
 5 SEQ ID NO: 218 polipéptido de DGAT2 de *Bos taurus* (Q70VZ8.1)  
 SEQ ID NO: 219 polipéptido de DGAT2 de *Mus musculus* (AAK84175.1)  
 SEQ ID NO: 220 tripéptido YFP –motivo de secuencia conservado de DGAT2 y/o MGAT1/2  
 SEQ ID NO: 221 tetrapéptido HPHG –motivo de secuencia conservado de DGAT2 y/o MGAT1/2  
 SEQ ID NO: 222 tetrapéptido EPHS –motivo de secuencia conservado de DGAT2 vegetal  
 10 SEQ ID NO: 223 RXGFX(K/R)XAXXXGXXX(L/V)VPXXXFG(E/Q) –motivo de secuencia conservado largo de DGAT2 que forma parte del dominio putativo de fosfolípido glicerol  
 SEQ ID NO: 224 FLXLXXXN – motivo de secuencia conservado de DGAT2 y MGAT1/2 de ratón que es un dominio de unión de lípidos neutro putativo  
 SEQ ID NO: 225 dominio plsC aciltransferasa (PF01553) de GPAT  
 15 SEQ ID NO: 226 dominio de la superfamilia de hidrolasa tipo HAD (PF12710) de GPAT  
 SEQ ID NO: 227 dominio de fosfoserina fosfatasa (PF00702). GPAT4-8 contiene una región N-terminal homóloga de este dominio  
 SEQ ID NO: 228 secuencia de aminoácidos GDLVICPEGTTTCREP conservada de GPAT  
 SEQ ID NO: 229 secuencia de aminoácidos conservada de GPAT/fosfatasa (Motivo I)  
 20 SEQ ID NO: 230 secuencia de aminoácidos conservada de GPAT/fosfatasa (Motivo III)

## DESCRIPCIÓN DETALLADA DE LA INVENCION

### Técnicas generales y definiciones

25 A menos que específicamente se defina de otra manera, se considera que todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el conocido comúnmente por el experto en la materia (por ejemplo, en cultivo celular, genética molecular, inmunología, inmunohistoquímica, química de proteínas, química de lípidos y ácidos grasos, producción de biocombustible y bioquímica).

30 A menos que se indique de otra manera, las técnicas de proteínas recombinantes, de cultivo celular e inmunológicas utilizadas en el presente documento invención son procedimientos estándar, bien conocidos por los expertos en la materia. Dichas técnicas se describen y explican en la bibliografía en fuentes tales como, J. Perbal, *A Practical Guide to Molecular Cloning*, John Wiley y Sons (1984), J. Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory Press (1989), T.A. Brown (editor), *Essential Molecular Biology: A Practical Approach*, Volúmenes 1 y 2, IRL Press (1991), D.M. Glover y B.D. Hames (editores), *DNA Cloning: A Practical Approach*, Volúmenes 1-4, IRL Press (1995 y 1996), F.M. Ausubel et al., (editores), *Current Protocols in Molecular Biology*, Greene Pub. Associates y Wiley-Interscience (1988, incluyendo todas las actualizaciones hasta el presente), Ed Harlow y David Lane (editores), *Antibodies: A Laboratory Manual*, Cold Spring Harbour Laboratory, 35 (1988) y J.E. Coligan et al., (editores), *Current Protocols in Immunology*, John Wiley & Sons (incluyendo todas las actualizaciones hasta el presente).

### Definiciones seleccionadas

45 El término "organismo transgénico no humano" se refiere, por ejemplo, a una planta completa, un alga, un animal no humano o un organismo adecuado para la fermentación tal como una levadura o un hongo, que comprende un polinucleótido exógeno (transgen) o un polipéptido exógeno. En una realización, el organismo transgénico no humano no es un animal o una parte del mismo. En una realización, el organismo transgénico no humano es un organismo fototrófico (por ejemplo, una planta o un alga) capaz de obtener energía de la luz solar para sintetizar 50 compuestos orgánicos para su nutrición. En otra realización, el organismo transgénico no humano es una bacteria fotosintética. El término "exógeno", en el contexto de un polinucleótido o un polipéptido, se refiere al polinucleótido o polipéptido cuando está presente en una célula en una cantidad alterada en comparación con su estado nativo. En una realización, la célula es una célula que no comprende naturalmente al polinucleótido o polipéptido. En otra realización, el polinucleótido exógeno o polipéptido es de un género diferente. En otra realización, el polinucleótido exógeno o polipéptido es de una especie diferente. En una realización el polinucleótido o polipéptido exógeno se expresa en una planta o célula vegetal hospedadora y el polinucleótido o polipéptido exógeno es de una especie o género diferente. En una realización, el polipéptido exógeno es una MGAT exógena. Tal como se usa en el presente documento, el término "lípidos extraído" se refiere a una composición extraída de un organismo transgénico o una 60 parte del mismo, que comprende al menos un 60 % (p/p) del lípido.

65 Tal como se usa en el presente documento, el término "lípidos no polar" se refiere a ácidos grasos y derivados de los mismos, que son solubles en solventes orgánicos pero insolubles en agua. Los ácidos grasos pueden ser ácidos grasos libres y/o se pueden encontrar en una forma esterificada. Los ejemplos de formas esterificadas incluyen, pero en un sentido no limitativo, triacilglicerol (TAG), diacilglicerol (DAG), monoacilglicerol (MAG). Los lípidos no polares también incluyen esteroides, ésteres de esteroles y ésteres de ceras. Los lípidos no polares también se conocen como "lípidos neutros". El lípido no polar normalmente es líquido a temperatura ambiente. Preferentemente, el lípido no

polar comprende predominantemente (>50 %) ácidos grasos que son de al menos 16 carbonos de longitud. Más preferentemente, al menos un 50 % de los ácidos grasos totales en el lípido no polar son ácidos grasos de C18 como por ejemplo ácido oleico. En una realización, al menos un 50 %, más preferentemente al menos un 70 %, más preferentemente al menos un 80 %, más preferentemente al menos un 90 %, más preferentemente al menos un 91 %  
 5 %, más preferentemente al menos un 92 %, más preferentemente al menos un 93 %, más preferentemente al menos un 94 %, más preferentemente al menos un 95 %, más preferentemente al menos un 96 %, más preferentemente al menos un 97 %, más preferentemente al menos un 98 %, más preferentemente al menos un 99 % de los ácidos grasos en el lípido no polar de la invención se pueden encontrar como TAG. El lípido no polar se puede purificar o tratar además, por ejemplo mediante hidrólisis con una base fuerte para liberar el ácido graso libre  
 10 o mediante fraccionamiento, destilación o semejantes. El lípido no polar de la invención puede formar parte del "aceite de las semillas" si se obtiene de semillas. El lípido no polar puede estar presente o se puede obtener de otras partes de plantas, incluyendo hojas o frutos, de células u organismos no humanos recombinantes, en cuyo caso el lípido no es un aceite de semillas según se define en el presente documento.

15 Los concentraciones de esteroles libres y esterificados (por ejemplo, sitoesterol, campesterol, estigmaesterol, brassicaesterol, D5-avenaesterol, sitostanol, campestanol y colesterol) en los lípidos extraídos pueden ser como se describe en Phillips et al., 2002.

Tal como se usa en el presente documento, el término "aceite de semillas" se refiere a una composición obtenida a partir de semillas/granos de una planta que comprende al menos un 60 % (p/p) de lípidos o que se puede obtener a partir de semillas/granos si el aceite de semillas aún está presente en dichas semillas/granos. Es decir, el aceite de semillas de la invención incluye un aceite de semillas que está presente en dichas semillas/granos o en una porción de los mismos, así como un aceite de semillas que ha sido extraído de las semillas/granos. El aceite de semillas preferentemente es un aceite de semillas extraído. **normalmente** el aceite de semillas es un líquido a temperatura ambiente. Preferentemente, el contenido de ácidos grasos totales (TFA) en el aceite de semillas comprende predominantemente (>50 %) ácidos grasos que son de al menos 16 carbonos de longitud. Más preferentemente, al menos un 50 % de los ácidos grasos totales en el aceite de semillas son ácidos grasos de C18 como por ejemplo ácido oleico. Los ácidos grasos normalmente se pueden encontrar en una forma esterificada tal como por ejemplo, TAG, DAG, acil-CoA o fosfolípido. Los ácidos grasos pueden ser ácidos grasos libres y/o se pueden encontrar en una forma esterificada. En una realización, al menos un 50 %, más preferentemente al menos un 70 %, más preferentemente al menos un 80 %, más preferentemente al menos un 90 %, más preferentemente al menos un 91 %  
 25 %, más preferentemente al menos un 92 %, más preferentemente al menos un 93 %, más preferentemente al menos un 94 %, más preferentemente al menos un 95 %, más preferentemente al menos un 96 %, más preferentemente al menos un 97 %, más preferentemente al menos un 98 %, más preferentemente al menos un 99 % de los ácidos grasos en el aceite de semillas se pueden encontrar como TAG. En una realización, el aceite de semillas de la invención es un aceite "sustancialmente purificado" o "purificado" que fue separado de uno o más lípidos adicionales, ácidos nucleicos, polipéptidos u otras moléculas contaminantes con los cuales está asociado en la semilla o en un extracto crudo. Se prefiere que el aceite de semillas sustancialmente purificado está al menos un 60 % libre, más preferentemente al menos un 75 % libre y, más preferentemente, al menos un 90 %  
 30 libre de otros componentes con los cuales está asociado en la semilla o el extracto. El aceite de semillas de la invención puede comprender además moléculas de ácidos no grasos tales como, en un sentido no limitativo, esteroides. En una realización, el aceite de semillas es aceite de colza (*Brassica napus*, *Brassica rapa* ssp.), aceite de mostaza (*Brassica juncea*), otro aceite de Brassica (por ejemplo, *Brassica napobrassica*, *Brassica camelina*), aceite de girasol (*Helianthus annuus*), aceite de lino (*Linum usitatissimum*), aceite de soja (*Glycine max*), aceite de cártamo (*Carthamus tinctorius*), aceite de maíz (*Zea mays*), aceite de tabaco (*Nicotiana tabacum*), aceite de cacahuete (*Arachis hypogaea*), aceite de palma (*Elaeis guineensis*), aceite de semillas de algodón (*Gossypium hirsutum*), aceite de coco (*Cocos nucifera*), aceite de palta (*Persea americana*), aceite de oliva (*Olea europaea*), aceite de cajú (*Anacardium occidentale*), aceite de macadamia (*Macadamia intergrifolia*), aceite de almendras (*Prunus amygdalus*),  
 45 aceite de avena (*Avena sativa*), aceite de arroz (*Oryza sativa* u *Oryza glaberrima*) o aceite de semillas de *Arabidopsis* (*Arabidopsis thaliana*). El aceite de semillas se puede extraer de las semillas/granos mediante cualquier método conocido en la técnica. Esto comprende normalmente extracción con solventes no polares tales como éter dietílico, éter de petróleo, mezclas de cloroformo/metanol o butanol, asociados generalmente con la primera molienda de las semillas. Los lípidos asociados con el almidón en el grano se puede extraer con butanol saturado con agua. El aceite de semillas se puede "desgomar" mediante métodos conocidos en la técnica para eliminar polisacáridos o se puede tratar de otras maneras para eliminar los contaminantes o mejorar la pureza, la estabilidad o el color. Los TAG y otros ésteres en el aceite de semillas se puede hidrolizar para liberar los ácidos grasos libres o el aceite de semillas se puede hidrogenar, tratar químicamente o enzimáticamente como es sabido en la técnica.  
 50

Tal como se usa en el presente documento, el término "ácido graso" se refiere a un ácido carboxílico con una cola alifática larga de al menos 8 átomos de carbono de longitud, ya sea saturado o insaturado. normalmente, los ácidos grasos tienen una cadena de carbono-carbono de al menos 12 carbonos de longitud. La mayoría de los ácidos grasos naturales tienen un número par de átomos de carbono porque su biosíntesis comprende acetato que tiene dos átomos de carbono. Los ácidos grasos se pueden encontrar en un estado libre (no esterificado) o en una forma esterificada tal como una parte de un triglicérido (TAG), diacilglicérido (DAG), monoacilglicérido (MAG), unido a acil-CoA (tio-éster) u otra forma unida. Cuando está unido de manera covalente en una forma esterificada, en el presente documento el ácido graso se conoce como un grupo "acilo". El ácido graso puede estar esterificado como un  
 60  
 65

fosfolípido tal como una fosfatidilcolina (PC), fosfatidiletanolamina, fosfatidilserina, fosfatidilglicerol, fosfatidilinositol o difosfatidilglicerol. Los "ácidos grasos saturados" no contienen ningún enlace doble o ningún grupo funcional a lo largo de la cadena. El término "saturado" se refiere a hidrógeno, en que todos los carbonos (además del grupo de ácido carboxílico [-COOH]) contiene tantos hidrógenos como sea posible. En otras palabras, el extremo omega ( $\omega$ )  
 5 contiene 3 hidrógenos (CH<sub>3</sub>-) y cada carbono dentro de la cadena contiene 2 hidrógenos (-CH<sub>2</sub>-). Los "ácidos grasos insaturados" son de una forma similar a los ácidos grasos saturados, excepto que existe uno o más grupos funcionales alqueno a lo largo de la cadena, donde cada alqueno sustituye una parte "-CH<sub>2</sub>-CH<sub>2</sub>-" de enlace simple de la cadena con una porción "-CH=CH-" unida mediante un enlace doble (es decir, un carbono unido mediante un enlace doble a otro carbono). Los dos átomos de carbono siguientes en la cadena que están unidos por cualquiera  
 10 de los lados del enlace doble pueden presentar una configuración *cis* o *trans*.

Tal como se usan en el presente documento, los términos "ácidos grasos poliinsaturados" o "PUFA" se refieren a un ácido graso que comprende al menos 12 átomos de carbono en su cadena de carbonos y al menos dos grupos alqueno (dobles enlaces carbono-carbono).

15 Un "monoacilglicérido" o "MAG" es un glicérido en el cual el glicerol está esterificado con un ácido graso. Tal como se usa en el presente documento, un MAG comprende un grupo hidroxilo en la posición *sn*-1/3 (en el presente documento también denominado *sn*-1 MAG o 1-MAG o 1/3-MAG) o *sn*-2 (en el presente documento también denominado 2-MAG) y por ello el MAG no incluye moléculas fosforiladas tal como PA o PC. Por lo tanto, MAG es un componente de lípidos neutros en una célula.

20 Un "diacilglicérido" o "DAG" es un glicérido en el cual el glicerol está esterificado con dos ácidos grasos. Tal como se usa en el presente documento, un DAG comprende un grupo hidroxilo en la posición *sn*-1,3 o *sn*-1,2/2,3 y por ello DAG no incluye moléculas fosforiladas tal como PA o PC. Por lo tanto, DAG es un componente de lípidos neutros en una célula. En la vía de Kennedy de la síntesis de DAG (Figura 1), el precursor *sn*-glicerol-3-fosfato (G-3-P) se esterifica a dos grupos acilo, cada uno proveniente de un éster de coenzima A de ácidos grasos, en una primera reacción catalizada por una glicerol-3-fosfato aciltransferasa (GPAT) en la posición *sn*-1 para formar LisoPA, seguido por una segunda acilación en la posición *sn*-2 catalizada por un ácido lisofosfatídico aciltransferasa (LPAAT) para formar ácido fosfatídico (PA). Este intermediario se desfosforila luego para formar DAG. En una vía anabólica alternativa (Figura 1), el DAG se puede formar por la acilación de ya sea *sn*-1 MAG o preferentemente *sn*-2 MAG, catalizada por MGAT. El DAG también se puede formar a partir de TAG por eliminación de un grupo acilo con una lipasa o a partir de PC esencialmente por eliminación de un grupo cabeza de colina con cualquiera de las enzimas CPT, PDCT o PLC (Figura 1).

30 Un "triacilglicérido" o "TAG" es un glicérido en el cual el glicerol está esterificado con tres ácidos grasos. En la vía de Kennedy de la síntesis de TAG, el DAG se forma como se describió previamente y luego se esterifica un tercer grupo acilo en el esqueleto de glicerol por la actividad de DGAT. Las vías alternativas para la formación de TAG incluyen una catalizada por la enzima PDAT y la vía MGAT que se describe en el presente documento.

40 Tal como se usa en el presente documento, el término "aciltransferasa" se refiere a una proteína que tiene la capacidad de transferir un grupo acilo de la acil-CoA sobre un sustrato e incluye MGAT, GPAT y DGAT.

45 Tal como se usa en el presente documento, el término "monoacilglicerol aciltransferasa" o "MGAT" se refiere a una proteína que transfiere un grupo de acilo graso de la acil-CoA a un sustrato MAG para producir DAG. Por lo tanto, el término "actividad monoacilglicerol aciltransferasa" se refiere al menos a la transferencia de un grupo acilo de la acil-CoA al MAG para producir DAG. La MGAT se conoce más por su rol en la absorción de grasa en el intestino de mamíferos, donde los ácidos grasos y *sn*-2 MAG generados por la digestión de la grasa de la dieta son resintetizados en TAG en los enterocitos para la síntesis y secreción de quilomicrones. La MGAT cataliza la primera etapa de este proceso, en donde se une covalentemente el grupo acilo de la acil-CoA graso, formada a partir de ácidos grasos y CoA, con *sn*-2 MAG. El término "MGAT" Tal como se usa en el presente documento incluye enzimas que actúan sobre sustratos *sn*-1/3 MAG y/o *sn*-2 MAG para formar *sn*-1,3 DAG y/o *sn*-1,2/2,3-DAG, respectivamente. En una realización preferida, la MGAT tiene preferencia por el sustrato *sn*-2 MAG con relación a *sn*-1 MAG o sustancialmente emplea solamente *sn*-2 MAG como sustrato (los ejemplos incluyen las MGAT que se describen en Cao et al., 2003 (especificidad de la MGAT1 de ratón por *sn*2-18:1-MAG > *sn*1/3-18:1-MAG (Figura 5));  
 50 Yen y Farese, 2003 (las actividades generales de MGAT1 de ratón y de MGAT2 humana sobre 2-MAG son mayores que sobre sustratos aceptores de acilo 1-MAG (Figura 5); y Cheng et al., 2003 (la actividad de la MGAT3 humana sobre 2-MAG es mucho mayor que sobre los sustratos 1/3-MAG (Figura 2D)).

60 Tal como se usa en el presente documento, la MGAT no incluye enzimas que transfieren preferencialmente un grupo acilo a LisoPA con relación a MAG, donde dichas enzimas se conocen como LPAAT. Es decir, una MGAT emplea preferencialmente sustratos de monoacilo no fosforilados, aún cuando pueden mostrar una actividad catalítica baja sobre LisoPA. Una MGAT preferida no tiene una actividad detectable en la acilación de LisoPA. Según se muestra en el presente documento, una MGAT (es decir, MGAT2 de *M. musculus*) también puede tener una función DGAT pero funciona predominantemente como una MGAT, es decir, tiene mayor actividad catalítica como una MGAT que como una DGAT cuando la actividad de la enzima se expresa en unidades de nmoles de producto/min/mg de proteína (véase también Yen et al., 2002).

Hay tres clases conocidas de MGAT, denominadas, MGAT1, MGAT2 y MGAT3, respectivamente. Los homólogos del gen MGAT1 humano (AF384163) están presentes (es decir, se conocen las secuencias) al menos en chimpancé, perro, vaca, ratón, rata, pez zebra, *Caenorhabditis elegans*, *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Eremothecium gossypii*, *Magnaporthe grisea* y *Neurospora crassa*. Los homólogos del gen MGAT2 humano (AY157608) están presentes al menos en chimpancé, perro, vaca, ratón, rata, pollo, pez zebra, mosca de la fruta y mosquito. Los homólogos del gen MGAT3 humano (AY229854) están presentes al menos en chimpancé, perro, vaca y pez zebra. Sin embargo, los homólogos de otros organismos se pueden identificar fácilmente mediante métodos conocidos en la técnica de identificación de secuencias homólogas.

Los ejemplos de polipéptidos de MGAT incluyen proteínas codificadas por genes de MGAT1 de *Homo sapiens* (AF384163), *Mus musculus* (AF384162), *Pan troglodytes* (XM\_001166055, XM\_0526044,2), *Canis familiaris* (XM\_545667,2), *Bos taurus* (NM\_001001153,2), *Rattus norvegicus* (NM\_001108803.1), *Danio rerio* MGAT1 (NM\_001122623.1), *Caenorhabditis elegans* (NM\_073012.4, NM\_182380.5, NM\_065258,3, NM\_075068,3, NM\_072248,3), *Kluyveromyces lactis* (XM\_455588.1), *Ashbya gossypii* (NM\_208895.1), *Magnaporthe oryzae* (XM\_368741.1), *Ciona intestinalis* predicho (XM\_002120843.1). Los ejemplos de polipéptidos de MGAT2 incluyen proteínas codificadas por genes de MGAT2 de *Homo sapiens* (AY157608), *Mus musculus* (AY157609), *Pan troglodytes* (XM\_522112,2), *Canis familiaris* (XM\_542304.1), *Bos taurus* (NM\_001099136.1), *Rattus norvegicus*, *Gallus gallus* (XM\_424082,2), *Danio rerio* (NM\_001006083.1), *Drosophila melanogaster* (NM\_136474,2, NM\_136473,2, NM\_136475,2), *Anopheles gambiae* (XM\_001688709.1, XM\_315985), *Tribolium castaneum* (XM\_970053.1). Los ejemplos de polipéptidos de MGAT3 incluyen proteínas codificadas por genes de MGAT3 de *Homo sapiens* (AY229854), *Pan troglodytes* (XM\_001154107.1, XM\_001154171.1, XM\_527842,2), *Canis familiaris* (XM\_845212.1), *Bos taurus* (XM\_870406.4), *Danio rerio* (XM\_688413.4).

Tal como se usa en el presente documento, la "vía de MGAT" se refiere a una vía anabólica, diferente de la vía de Kennedy, para la formación de TAG, en la cual se forma DAG por la acilación de *sn*-1 MAG o, preferentemente, *sn*-2 MAG, catalizada por MGAT. El DAG se puede usar luego para formar TAG u otros lípidos. La vía MGAT se muestra como ejemplo en la Figura 1. Tal como se usa en el presente documento, el término "diacilglicerol aciltransferasa" (DGAT) se refiere a una proteína que transfiere un grupo de acilo graso de la acil-CoA a un sustrato DAG para producir TAG. Por lo tanto, el término "actividad diacilglicerol aciltransferasa" se refiere a la transferencia de un grupo acilo de la acil-CoA a DAG para producir TAG. Una DGAT también puede tener una función MGAT pero funciona predominantemente como una DGAT, es decir, tiene mayor actividad catalítica como una DGAT que como una MGAT cuando la actividad de la enzima se expresa en unidades de nmoles de producto/min/mg de proteína (véase, por ejemplo, Yen et al., 2005).

Hay tres tipos conocidos de DGAT denominados DGAT1, DGAT2 y DGAT3 respectivamente. Los polipéptidos de DGAT1 normalmente tienen 10 dominios transmembrana, los polipéptidos de DGAT2 normalmente tienen 2 dominios transmembrana, en tanto los polipéptidos de DGAT3 normalmente no tienen ninguno y se cree que son solubles en el citoplasma, no estando integrados en las membranas. Los ejemplos de polipéptidos de DGAT1 incluyen proteínas codificadas por los genes DGAT1 de *Aspergillus fumigatus* (N° Acceso XP\_755172), de *Arabidopsis thaliana* (CAB44774), de *Ricinus communis* (AAR11479), de *Vernicia fordii* (ABC94472), de *Vernonia galamensis* (ABV21945, ABV21946), de *Euonymus alatus* (AAV31083), de *Caenorhabditis elegans* (AAF82410), de *Rattus norvegicus* (NP\_445889), de *Homo sapiens* (NP\_036211), así como variantes y/o mutantes de los mismos. Los ejemplos de polipéptidos de DGAT2 incluyen proteínas codificadas por genes de DGAT2 de *Arabidopsis thaliana* (NP\_566952.1; SEQ ID NO: 212), *Ricinus communis* (AAY16324.1; SEQ ID NO: 213), *Vernicia fordii* (ABC94474.1; SEQ ID NO: 214), *Mortierella ramanniana* (AAK84179.1; SEQ ID NO: 215), *Homo sapiens* (Q96PD7,2; SEQ ID NO: 216) (Q58HT5.1; SEQ ID NO: 217), *Bos taurus* (Q70VZ8.1; SEQ ID NO: 218), *Mus musculus* (AAK84175.1; SEQ ID NO: 219), así como variantes y/o mutantes de los mismos.

Los ejemplos de polipéptidos de DGAT3 incluyen proteínas codificadas por los genes DGAT3 de cacahuete (*Arachis hypogaea*, Saha, et al., 2006), así como variantes y/o mutantes de los mismos. Una DGAT tiene poca o ninguna actividad MGAT detectable, por ejemplo, menos que 300 pmol/min/mg de proteína, preferentemente menos que 200 pmol/min/mg de proteína, más preferentemente 100 pmol/min/mg de proteína.

La DGAT2, pero no DGAT1, comparte una gran homología de secuencia con las enzimas MGAT, lo cual sugiere que los genes de DGAT2 y MGAT probablemente comparten un origen genético común. Aunque se conocen múltiples isoformas relacionadas con la catálisis de la misma etapa de síntesis de TAG, pueden cumplir distintos roles funcionales, como lo sugiere la distribución tisular diferencial y la localización subcelular de la familia de enzimas DGAT/MGAT. En mamíferos, la MGAT1 se expresa principalmente en estómago, riñón, tejido adiposo, en tanto MGAT2 y MGAT3 muestran la mayor expresión en el intestino delgado. En mamíferos, DGAT1 se expresa de manera ubicua en muchos tejidos, pero con mayor expresión en el intestino delgado, en tanto DGAT2 es más abundante en el hígado. MGAT3 solamente existe en mamíferos superiores y seres humanos, pero no en roedores según un análisis bioinformático. MGAT3 comparte una mayor homología de secuencia con DGAT2 que MGAT1 y MGAT3. MGAT3 presenta una actividad DGAT significativamente mayor que las enzimas MGAT1 y MGAT2 (MGAT3 > MGAT1 > MGAT2) cuando se usó cualquier MAG o DAG como sustrato, lo cual sugiere que MGAT3 funciona como una TAG sintasa putativa.

Ambas MGAT1 y MGAT2 pertenecen a la misma clase de aciltransferasas que la DGAT2. Algunos de los motivos que han demostrado ser importantes para la actividad catalítica de DGAT2 en algunas DGAT2, también están conservados en las aciltransferasas MGAT. De particular interés es un dominio de unión a lípidos neutros putativo con la secuencia consenso FLXLXXXN (SEQ ID NO: 224), donde cada X es de manera independiente cualquier aminoácido distinto de prolina y N es cualquier aminoácido no polar, ubicado dentro de la región transmembrana N-terminal seguido por un dominio glicerol/fosfolípido aciltransferasa putativo. El motivo FLXLXXXN está presente en la DGAT2 (aminoácidos 81-88) y en la MGAT1/2 de ratón pero no en las DGAT2 de levaduras o plantas. Es importante para la actividad de la DGAT2 de ratón. Otros motivos de secuencias de DGAT2 y/o MGAT1/2 incluyen:

1. Un tripéptido YFP altamente conservado (SEQ ID NO: 220) en la mayoría de los polipéptidos de DGAT2 y también en MGAT1 y MGAT2, por ejemplo, presente como los aminoácidos 139-141 en DGAT2 de ratón. Una mutación de este motivo dentro de la DGAT2 de levadura con sustituciones no conservadoras volvió a la enzima no funcional.
2. Un tetrapéptido HPHG (SEQ ID NO: 221), altamente conservado en las MGAT así como en las secuencias de DGAT2 de animales y hongos, por ejemplo, presente como los aminoácidos 161-164 en DGAT2 de ratón, e importante para la actividad catalítica al menos en DGAT2 de levadura y ratón. Las aciltransferasas DGAT2 de plantas contienen en su lugar una secuencia conservada EPHS (SEQ ID NO: 222), de modo que se pueden tolerar cambios conservadores en el primer y cuarto aminoácidos.
3. Un motivo conservado más largo que forma parte del dominio putativo de fosfolípido glicerol. Un ejemplo de este motivo es RXGFX(K/R)AXXXGXXX(L/V)VPXXFG(E/Q) (SEQ ID NO: 223), que está presente como los aminoácidos 304-327 en la DGAT2 de ratón. Este motivo está menos conservado en una secuencia de aminoácidos que los demás, como sería de esperar por su longitud, pero se pueden reconocer homólogos mediante búsqueda de motivos. El espaciado puede variar entre los aminoácidos más conservados, es decir, puede existir X aminoácidos adicionales o X aminoácidos menos dentro del motivo en comparación con la secuencia anterior.

Tal como se usa en el presente documento, el término "glicerol-3-fosfato aciltransferasa" o "GPAT" se refiere a una proteína que acila glicerol-3-fosfato (G-3-P) para formar LisoPA y/o MAG, donde éste último producto se forma si la GPAT también tiene actividad fosfatasa sobre LisoPA. El grupo acilo transferido normalmente es de la acil-CoA. Por lo tanto, el término "actividad glicerol-3-fosfato aciltransferasa" se refiere a la acilación de G-3-P para formar LisoPA y/o MAG. El término "GPAT" abarca enzimas que acilan G-3-P para formar *sn*-1 LPA y/o *sn*-2 LPA, preferentemente *sn*-2 LPA. En una realización preferida, la GPAT tiene actividad fosfatasa. En una realización más preferida aún, la GPAT es una *sn*-2 GPAT que tiene una actividad fosfatasa que produce *sn*-2 MAG.

Tal como se usa en el presente documento, el término "*sn*-1 glicerol-3-fosfato aciltransferasa" (*sn*-1 GPAT) se refiere a una proteína que acila *sn*-glicerol-3-fosfato (G-3-P) para formar preferencialmente 1-acil-*sn*-glicerol-3-fosfato (*sn*-1 LPA). Por lo tanto, el término "actividad *sn*-1 glicerol-3-fosfato aciltransferasa" se refiere a la acilación de *sn*-glicerol-3-fosfato para formar 1-acil-*sn*-glicerol-3-fosfato (*sn*-1 LPA).

Tal como se usa en el presente documento, el término "*sn*-2 glicerol-3-fosfato aciltransferasa" (*sn*-2 GPAT) se refiere a una proteína que acila *sn*-glicerol-3-fosfato (G-3-P) para formar preferencialmente 2-acil-*sn*-glicerol-3-fosfato (*sn*-2 LPA). Por lo tanto, el término "actividad *sn*-2 glicerol-3-fosfato aciltransferasa" se refiere a la acilación de *sn*-glicerol-3-fosfato para formar 2-acil-*sn*-glicerol-3-fosfato (*sn*-2 LPA).

La familia GPAT es grande y todos los miembros conocidos contienen dos dominios conservados, un dominio plsC aciltransferasa (PF01553; SEQ ID NO: 225) y un dominio de la superfamilia de hidrolasas tipo HAD (PF12710; SEQ ID NO: 226). Además de esto, en *Arabidopsis thaliana*, la GPAT4-8 contiene una región N-terminal homóloga de un dominio fosfoserina fosfatasa (PF00702; SEQ ID NO: 227). Ambas GPAT4 y GPAT6 contienen restos conservados conocidos por ser críticos para la actividad fosfatasa, específicamente aminoácidos conservados (indicados en negritas) en el Motivo I (**DXDX**[T/V][L/V]; SEQ ID NO:229) y en el Motivo III (**K**-[G/S][D/S]XXX[D/N]; SEQ ID NO:330) ubicados por el extremo N-terminal (Yang et al., 2010). Preferentemente, la GPAT tiene preferencia por *sn*-2 y actividad fosfatasa para producir *sn*-2 MAG (también denominado "2-MAG" en el presente documento) a partir de glicerol-3-fosfato (G-3-P) (Figura 1), por ejemplo, (GPAT4 (NP\_171667.1) y GPAT6 (NP\_181346.1)) de *Arabidopsis*. Más preferentemente, la GPAT emplea acil-CoA como sustrato de ácidos grasos.

Los homólogos de GPAT4 (NP\_171667) y GPAT6 (NP\_181346) incluyen AAF02784, AAL32544, AAP03413, ABK25381, ACN34546, BAF00762, BAH00933, EAY84189, EAY98245, EAZ21484, EEC71826, EEC76137, EEE59882, EFJ08963, EFJ08964, EFJ11200, EFJ15664, EFJ24086, EFJ29816, EFJ29817, NP\_001044839, NP\_001045668, NP\_001147442, NP\_001149307, NP\_001168351, NP\_181346, NP\_191950, XP\_001765001, XP\_001769671, XP\_001769724, XP\_001771186, XP\_001780533, XP\_002268513, XP\_002275348, XP\_002276032, XP\_002279091, XP\_002309124, XP\_002309276, XP\_002322752, XP\_002323563, XP\_002439887, XP\_002458786, XP\_002463916, XP\_002464630, XP\_002511873, XP\_002517438, XP\_002520171, XP\_002872955, XP\_002881564, ACT32032, ACT32032, NP\_001051189, NP\_171667, XP\_002320138, XP\_002451377, XP\_002451377, XP\_002531350, XP\_002872955 y XP\_002889361.

Los motivos y/o restos conservados se pueden usar como elemento diagnóstico basado en secuencias para la

identificación de enzimas GPAT/fosfatasa bifuncionales. Como alternativa, se podría usar un ensayo basado en la función más severo. Dicho ensayo comprende, por ejemplo, alimentar glicerol-3-fosfato marcado a células o microsomas y cuantificar los niveles de productos marcados por cromatografía en capa delgada o una técnica similar. La actividad GPAT da como resultado la producción de LPA marcado en tanto la actividad GPAT/fosfatasa da como resultado la producción de MAG marcado.

5 Tal como se usa en el presente documento, el término "tipo silvestre" o variaciones del mismo se refiere a una célula o a un organismo no humano o a una parte del mismo que no ha sido modificado genéticamente.

10 El término "correspondiente" se refiere a una célula o a un organismo no humano o a una parte del mismo que tiene el mismo antecedente genético o uno similar, que una célula o un organismo no humano o una parte del mismo de la invención pero que ha sido modificado según se describe en el presente documento (por ejemplo, la célula o el organismo no humano o una parte del mismo no contiene un polinucleótido exógeno que codifica una MGAT o una MGAT exógena). Una célula o, un organismo no humano o una parte del mismo, correspondiente se puede usar como control para comparar los niveles de ácidos nucleicos o la expresión de proteínas o el grado y la naturaleza de la modificación de caracteres, por ejemplo la producción y/o el contenido de lípidos no polares, con una célula o un organismo no humano o una parte del mismo, modificado según se describe en el presente documento.

20 Tal como se usa en el presente documento "comparado con" se refiere a la comparación de niveles de un lípido no polar o del contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo, que expresa dichos uno o más polinucleótidos exógenos o polipéptidos exógenos con un organismo transgénico no humano o una parte del mismo, que no contiene a dichos uno o más polinucleótidos o polipéptidos exógenos.

25 Tal como se usa en el presente documento, "mayor capacidad para producir lípidos no polares" es un término relativo que se refiere a la cantidad total de lípidos no polares producida por una célula o por un organismo no humano o una parte del mismo, de la invención que es incrementada con relación a una célula o un organismo no humano o una parte del mismo, correspondiente. En una realización, aumenta el contenido de TAG y/o ácidos grasos poliinsaturados del lípido no polar.

30 Tal como se usa en el presente documento, el término "un polinucleótido aislado o recombinante que regula negativamente de la producción y/o actividad de una enzima endógena" o variaciones del mismo, se refiere a un polinucleótido que codifica una molécula de ARN que regula negativamente de la producción y/o actividad (por ejemplo, que codifica un ARNsi) o que por sí mismo regula negativamente de la producción y/o actividad (por ejemplo, es un ARNsi que se puede administrar directamente, por ejemplo, a una célula) de una enzima endógena por ejemplo, DGAT, *sn*-1 glicerol-3-fosfato aciltransferasa (GPAT), 1-acil-glicerol-3-fosfato aciltransferasa (LPAAT), acil-CoA:lisofosfatidilcolina aciltransferasa (LPCAT), ácido fosfatídico fosfatasa (PAP) o una combinación de dos o más de los mismos.

40 Tal como se usa en el presente documento, el término "sobre una base en peso" se refiere al peso de una sustancia (por ejemplo, TAG, DAG, ácidos grasos) como un porcentaje del peso de la composición que comprende la sustancia (por ejemplo, semillas, hojas). Por ejemplo, si una semilla transgénica contiene 25 µg de ácidos grasos totales por cada 120 µg de peso de la semilla; el porcentaje de ácidos grasos totales sobre una base en peso es del 20,8 %.

45 Tal como se usa en el presente documento, el término "sobre una base relativa" se refiere a la cantidad de una sustancia en una composición que comprende la sustancia en comparación con una composición correspondiente, como un porcentaje.

50 Tal como se usa en el presente documento, el término "el contenido no lipídico relativo" se refiere a la expresión del contenido de lípidos no polares de una célula, organismo o una parte del mismo o a los lípidos extraídos de los mismos, en comparación con una célula, organismo o una parte del mismo, correspondiente o al lípido extraído de la célula, organismo o una parte del mismo, correspondiente como un porcentaje. Por ejemplo, si una semilla transgénica contiene 25 µg de ácidos grasos totales, en tanto la semilla correspondiente tenía 20 µg de ácidos grasos totales; el incremento en el contenido de lípido no polar sobre una base relativa es igual al 25 %.

55 Producción de diacilgliceroles y triacilgliceroles

60 En una realización, el organismo transgénico no humano o una parte del mismo, de la invención produce niveles más altos de lípidos no polares, incluyendo TAG, preferentemente tanto TAG como DAG, que un organismo no humano correspondiente o una parte del mismo. En un ejemplo, las plantas transgénicas de la invención producen semillas y/u hojas que tienen un contenido aumentado de lípidos no polares de TAG, preferentemente tanto TAG como DAG, en comparación con las semillas y/u hojas correspondientes. El contenido de lípidos no polares del organismo no humano o una parte del mismo es un 0,5 % mayor sobre una base en peso en comparación con un organismo no humano correspondiente o una parte del mismo.

65 En otra realización, el organismo transgénico no humano o una parte del mismo, preferentemente una planta o

semilla, produce TAG y posiblemente también DAG que están enriquecidos en uno o más ácidos grasos particulares. Se puede incorporar un amplio espectro de ácidos grasos en los DAGs y/o TAGs, incluyendo ácidos grasos saturados e insaturados y ácidos grasos de cadena corta y de cadena larga. Algunos ejemplos no limitativos de ácidos grasos que se pueden incorporar en los DAGs y/o TAGs incluyen: ácidos grasos cáprico (10:0), láurico (12:0), mirístico (14:0), palmítico (16:0), palmitoleico (16:1), esteárico (18:0), oleico (18:1), vaccénico (18:1), linoleico (18:2), eleosteárico (18:3),  $\gamma$ -linolénico (18:3),  $\alpha$ -linolénico (18:3 $\omega$ 3), estearidónico (18:4 $\omega$ 3), araquídico (20:0), eicosadienoico (20:2), dihomo- $\gamma$ -linoleico (20:3), eicosatrienoico (20:3), araquidónico (20:4), eicosatetraenoico (20:4), eicosapentaenoico (20:5 $\omega$ 3), behénico (22:0), docosapentaenoico (22:5 $\omega$ ), docosahexaenoico (22:6 $\omega$ 3), lignocérico (24:0), nervónico (24:1), cerótico (26:0) y montánico (28:0). En una realización de la presente invención, el organismo transgénico o partes del mismo está enriquecido en DAGs y/o TAGs que comprenden ácidos grasos poliinsaturados.

En una realización adicional de la invención, el organismo transgénico no humano o una parte del mismo, preferentemente una planta o semilla, es transformado con un ADN quimérico que codifica una MGAT según se define en el presente documento que puede tener o no, actividad DGAT. La expresión de la MGAT preferentemente da como resultado niveles más altos de lípidos no polares tales como DAG o TAG y/o un mayor rendimiento de lípidos no polares en dicho organismo transgénico no humano o una parte del mismo. En una realización preferida, el organismo transgénico no humano es una planta.

En una realización adicional, el organismo transgénico no humano o una parte del mismo, es transformado con un ADN quimérico que codifica una GPAT o una DGAT. Preferentemente, el organismo es transformado con ambos ADN quiméricos, que preferentemente están ligados de manera covalente en una molécula de ADN tal como, por ejemplo, una sola molécula de ADN-T.

Yang et al., (2010) describen dos glicerol-3-fosfato aciltransferasas (GPAT4 y GPAT6) de *Arabidopsis* con preferencia por *sn*-2 y actividad fosfatasa que tienen capacidad para producir *sn*-2 MAG a partir de glicerol-3-fosfato (G-3-P) (Figura 1). Se propone que estas enzimas forman parte de la vía de síntesis de cutina. Se ha demostrado que las GPAT4 y GPAT6 de *Arabidopsis* emplean acil-CoA como sustrato de ácidos grasos (Zheng et al., 2003).

La combinación de una GPAT/fosfatasa bifuncional con una MGAT permite obtener una nueva vía de síntesis de DAG usando G-3-P como un sustrato y dos grupos acilo derivados de acil-CoA como los otros sustratos. De manera similar, la combinación de dicha GPAT/fosfatasa bifuncional con una MGAT que tiene actividad DGAT permite obtener una nueva vía de síntesis de TAG usando glicerol-3-fosfato como un sustrato y tres grupos acilo derivados de acil-CoA como los otros sustratos.

Por consiguiente, en una realización de la invención, el organismo transgénico no humano o una parte del mismo, es co-transformado con una GPAT/fosfatasa bifuncional y con una MGAT que no tiene actividad DGAT. Esto daría como resultado la producción de MAG por la GPAT/fosfatasa bifuncional que luego se podría convertir en DAG por la MGAT y luego en TAG por una DGAT nativa u otra actividad. La nueva producción de DAG se podría confirmar y seleccionar, por ejemplo, realizando dicha co-transformación en una cepa de levadura que contiene noqueos *SLC1* + *SLC4* letales tal como el que describen Benghezal et al., (2007; Figura 2). La Figura 2 de Benghezal et al., (2007) muestra que el noqueo de las dos LPATS de levadura (*SLC1* y *SLC4*) es letal. El doble mutante *SLC1* + *SLC4* de levadura solamente se puede mantener debido a un plásmido complementario que proporciona uno de los genes *sic* (*SLC1* en ese caso) en *trans*. La selección negativa por adición de FOA al medio da como resultado la pérdida de este plásmido complementario (contraselección del marcador de selección Ura) y vuelve no viables a las células.

En otra realización de la invención, el organismo transgénico no humano o una parte del mismo, preferentemente una planta o semilla, es co-transformado con ADN quiméricos que codifican una GPAT/fosfatasa bifuncional y una MGAT que tiene actividad DGAT. Esto daría como resultado la producción de MAG por la GPAT/fosfatasa bifuncional que luego se podría convertir en DAG y luego en TAG por la MGAT.

En una realización adicional más, se silencia una o más GPAT endógenas sin actividad fosfatasa detectable, por ejemplo se silencia uno o más genes que codifican GPAT que acilan glicerol-3-fosfato para formar LPA en la vía de Kennedy (por ejemplo, *Arabidopsis* GPAT1).

Las preferencias de sustrato se podrían introducir en las nuevas vías de síntesis DAG y TAG, por ejemplo, suministrando cepas H1246 de levadura transgénicas que expresan variantes MGAT con una concentración de un ácido graso libre particular (por ejemplo, DHA) que previene la complementación por el gen MGAT de tipo silvestre. Solamente crecerían las variantes capaces de usar los ácidos grasos libres suministrados. Varios ciclos de introducción de MGAT daría como resultado la producción de una MGAT con mayor preferencia por ácidos grasos particulares.

Los diversos complementos y suplementos de la vía de Kennedy descritos precedentemente se podría efectuar en cualquier tipo celular debido a la naturaleza ubicua del sustrato inicial glicerol-3-fosfato. En una realización, el uso de transgenes da como resultado mayores rendimientos de aceite.

Polinucleótidos

Los términos "polinucleótido" y "ácido nucleico" se usan indistintamente. Se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sea desoxirribonucleótidos o ribonucleótidos o análogos de los mismos. Un polinucleótido de la invención puede ser de origen genómico, ADNc, semisintético o sintético, de cadena doble o de cadena simple y en virtud de su origen o manipulación: (1) no está asociado con todo o una porción de un polinucleótido con el cual está asociado en la naturaleza, (2) está ligado a un polinucleótido distinto del polinucleótido al cual está ligado en la naturaleza o (3) no existe en la naturaleza. Los siguientes son ejemplos no limitativos de polinucleótidos: regiones de codificación o no codificantes de un gen o fragmento de gen, loci (locus) definidos por análisis de ligamiento, exones, intrones, ARN mensajero (ARNm), ARN de transferencia (ARNt), ARN ribosómico (ARNr), ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, ADN quimérico de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y análogos de nucleótidos. Si están presentes, las modificaciones de la estructura de nucleótidos pueden ser impartidas antes o después del ensamblaje del polímero. La secuencia de nucleótidos puede estar interrumpida por componentes no nucleotídicos. Un polinucleótido se puede modificar además después de la polimerización, tal como por conjugación, con un componente de marcación.

Un "polinucleótido aislado" significa un polinucleótido que generalmente fue separado de las secuencias de polinucleótidos con las cuales está asociado o ligado en su estado nativo. Preferentemente, el polinucleótido aislado está al menos un 60 % libre, más preferentemente al menos un 75 % libre y más preferentemente al menos un 90 % libre de las secuencias de polinucleótidos con las cuales está naturalmente asociado o ligado.

Tal como se usa en el presente documento, el término "gen" se interpreta en su sentido más amplio e incluye las secuencias de desoxirribonucleótidos que comprenden a la región transcrita y, si se traducen, la región codificante de la proteína, de un gen estructural e inclusive las secuencias ubicadas adyacentes a la región codificante por ambos extremos 5' y 3' para una distancia de al menos 2 kb aproximadamente por cualquier extremo y que están relacionadas con la expresión del gen. En este sentido, el gen incluye las señales de control, tales como promotores, potenciadores, señales de terminación y/o de poliadenilación, que están asociadas naturalmente con un gen dado o señales de control heterólogas en cuyo caso el gen se conoce como un "gen quimérico". Las secuencias ubicadas 5' con respecto a la región codificante de la proteína y que están presentes sobre el ARNm se denominan secuencias no traducidas 5'. Las secuencias ubicadas 3' con respecto a la región codificante de la proteína y que están presentes sobre el ARNm se denominan secuencias no traducidas 3'. El término "gen" abarca tanto al ADNc como las formas genómicas de un gen. Una forma genómica o clon de un gen contiene la región codificante que puede interrumpirse con secuencias no codificantes denominadas "intrones" o "regiones intervinientes" o "secuencias intervinientes". Los intrones son segmentos de un gen que se transcriben en un ARN nuclear (ARNn). Los intrones pueden contener elementos reguladores tales como potenciadores. Los intrones son removidos o "procesados fuera" del transcrito nuclear o primario; por ello no hay intrones en el transcrito de ARNm. El ARNm funciona durante la traducción especificando la secuencia o el orden de los aminoácidos en un polipéptido naciente. El término "gen" incluye una molécula sintética o de fusión que codifica todo o parte de las proteínas de la invención que se describen en el presente documento y una secuencia de nucleótidos complementaria de cualquiera de las mencionadas anteriormente.

Tal como se usa en el presente documento, "ADN quimérico" se refiere a cualquier molécula de ADN que no existe naturalmente en la naturaleza; también denominada "construcción de ADN" en el presente documento. Normalmente, un ADN quimérico comprende secuencias reguladoras y transcritas o codificantes de proteínas que no existen juntas en la naturaleza. Por consiguiente, un ADN quimérico puede comprender secuencias reguladoras y secuencias codificantes derivadas de distintas fuentes o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero en una disposición diferente de la hallada en la naturaleza. El marco de lectura abierto puede estar ligado o no, a sus elementos reguladores 5' y 3' naturales. El marco de lectura abierto se puede incorporar, por ejemplo, en el genoma de la planta, en una ubicación no natural o en un replicón o vector donde no existe naturalmente tal como en un plásmido bacteriano o un vector viral. El término "ADN quimérico" no se limita a moléculas de ADN que se pueden replicar en un hospedador, sino que incluye ADN capaz de ligarse en un replicón, por ejemplo, mediante secuencias adaptadoras específicas.

Un "transgen" es un gen que fue introducido en el genoma mediante un procedimiento de transformación. Los términos "modificado genéticamente", "transgénico" y variaciones de los mismos incluyen introducir un gen en una célula por transformación o transducción, mutar un gen en una célula y alterar o modular genéticamente la regulación de un gen en una célula o la progenie de cualquier célula modificada como se describió previamente.

Una "región genómica", Tal como se usa en el presente documento, se refiere a una posición dentro del genoma en donde se ha insertado un transgen o un grupo de transgenes (también denominado agrupación en el presente documento), en una célula o en un predecesor de la misma. Dichas regiones comprenden solamente nucleótidos que fueron incorporados por la intervención del hombre, tal como mediante los métodos que se describen en el presente documento.



Un "polinucleótido recombinante" de la invención se refiere a una molécula de ácido nucleico que se construyó o modificó mediante métodos recombinantes artificiales. El polinucleótido recombinante puede estar presente en una célula en una cantidad alterada o se puede expresar a un índice alterado (por ejemplo, en el caso del ARNm) en comparación con su estado nativo. En una realización, el polinucleótido se introduce en una célula que no comprende naturalmente al polinucleótido. normalmente se usa un ADN exógeno como templado para la transcripción del ARNm que luego es traducido en una secuencia de restos de aminoácidos continua que codifican un polipéptido de la invención dentro de la célula transformada. En otra realización, el polinucleótido es endógeno para la célula y su expresión es alterada por medios recombinantes, por ejemplo, se introduce una secuencia de control exógena 5' con respecto a un gen endógeno de interés para permitir la expresión del polipéptido codificado por el gen por parte de la célula transformada.

Un polinucleótido recombinante de la invención incluye polinucleótidos que no han sido separados de los demás componentes del sistema de expresión basado en células o sin células, en el cual está presente y los polinucleótidos producidos en tales sistemas celulares o no celulares que luego se separan por purificación de por lo menos algunos de los demás componentes. El polinucleótido puede ser una extensión contigua de nucleótidos existente en la naturaleza o comprende dos o más extensiones contiguas de nucleótidos de fuentes diferentes (naturales y/o sintéticas) unidas para formar un solo polinucleótido. normalmente dichos polinucleótidos quiméricos comprenden al menos un marco de lectura abierto que codifica un polipéptido de la invención ligado operativamente a un promotor adecuado para dirigir la transcripción del marco de lectura abierto en una célula de interés.

Con respecto a los polipéptidos definidos, se podrá apreciar que las cifras de % de identidad mayores que las provistas antes abarcarán formas de realización preferidas. Por consiguiente, cuando fuera aplicable, a la luz de las cifras mínimas de % de identidad, se prefiere que el polinucleótido comprenda una secuencia de polinucleótidos que es al menos un 60 %, más preferentemente al menos un 65 %, más preferentemente al menos un 70 %, más preferentemente al menos un 75 %, más preferentemente al menos un 80 %, más preferentemente al menos un 85 %, más preferentemente al menos un 90 %, más preferentemente al menos un 91 %, más preferentemente al menos un 92 %, más preferentemente al menos un 93 %, más preferentemente al menos un 94 %, más preferentemente al menos un 95 %, más preferentemente al menos un 96 %, más preferentemente al menos un 97 %, más preferentemente al menos un 98 %, más preferentemente al menos un 99 %, más preferentemente al menos un 99,1 %, más preferentemente al menos un 99,2 %, más preferentemente al menos un 99,3 %, más preferentemente al menos un 99,4 %, más preferentemente al menos un 99,5 %, más preferentemente al menos un 99,6 %, más preferentemente al menos un 99,7 %, más preferentemente al menos un 99,8 % y aún más preferentemente al menos un 99,9 % idéntica al SEQ ID N° relevante nominado.

Un polinucleótido de o de utilidad para, la presente invención se puede hibridar selectivamente, en condiciones rigurosas, con un polinucleótido definido en el presente documento. Tal como se usa en el presente documento, las condiciones rigurosas son aquellas que (1) durante la hibridación emplean un agente desnaturante tal como formamida, por ejemplo, formamida 50 % (v/v) con albúmina de suero bovino 0,1 % (p/v), Ficoll 0,1 %, polivinilpirrolidona 0,1 %, solución amortiguadora de fosfato de sodio 50 mM a pH 6,5 con NaCl 750 mM, citrato de sodio 75 mM a 42 °C; o (2) emplean formamida 50 %, SSC 5 x (NaCl 0,75 M, citrato de sodio 0,075 M), fosfato de sodio 50 mM (pH 6,8), pirofosfato de sodio 0,1 %, solución de Denhardt 5 x, ADN de esperma de salmón sonicado (50 g/ml), SDS 0,1 % y sulfato de dextrano 10 % a 42 °C en SSC 0,2 x y SDS 0,1 % y/o (3) emplean una fuerza iónica baja y temperatura alta para el lavado, por ejemplo, NaCl 0,015 M/citrato de sodio 0,0015 M/SDS 0,1 % a 50 °C.

Los polinucleótidos de la invención pueden contener, en comparación con las moléculas naturales, una o más mutaciones que son supresiones, inserciones o sustituciones de restos de nucleótidos. Los polinucleótidos que contienen mutaciones con relación a una secuencia de referencia pueden ser naturales (es decir, aislados de una fuente natural) o sintéticas (por ejemplo, obtenidos mediante mutagénesis dirigida al sitio o entremezclado de ADN del ácido nucleico descrito precedentemente).

#### Vectores recombinantes

En el presente documento se divulga un vector recombinante, que comprende al menos un polinucleótido definido en el presente documento, capaz de suministrar el polinucleótido en una célula hospedadora. Los vectores recombinantes incluyen vectores de expresión. Los vectores recombinantes contienen secuencias de polinucleótidos heterólogos, es decir, secuencias de polinucleótidos que no existen naturalmente adyacentes a un polinucleótido definido en el presente documento, que, preferentemente, deriva de una especie diferente. El vector puede ser de ARN o ADN, ya sea procariota o eucariota y normalmente es un vector viral, derivado de un virus o un plásmido. Los vectores de plásmidos incluyen normalmente secuencias de ácidos nucleicos adicionales que proporcionan una manera fácil de selección, amplificación y transformación del casete de expresión en células procariotas, por ejemplo, vectores derivados de pUC, vectores derivados de pSK, vectores derivados de pGEM, vectores derivados de pSP, vectores derivados de pBS o vectores binarios que contienen una o más regiones de ADN-T. Las secuencias de ácidos nucleicos adicionales incluyen orígenes de replicación para proporcionar una replicación autónoma del vector, genes marcadores de selección, que preferentemente codifican resistencia a antibióticos o herbicidas, múltiple sitios de clonación únicos que proporcionan múltiples sitios para insertar secuencias de ácidos

nucleicos o genes codificados en la construcción de ácido nucleico y secuencias que mejoran la transformación de células procariotas y eucariotas (en especial plantas).

El término "ligado operativamente", tal como se usa en el presente documento, se refiere a una relación funcional entre dos o más segmentos de ácidos nucleicos (por ejemplo, ADN). normalmente, se refiere a la relación funcional de un elemento regulador de la transcripción (promotor) con una secuencia transcrita. Por ejemplo, un promotor está ligado operativamente a una secuencia codificante de un polinucleótido definido en el presente documento, si estimula o modula la transcripción de la secuencia codificante en una célula apropiada. En general, los elementos promotores reguladores de la transcripción que están ligados operativamente a una secuencia transcrita se encuentran físicamente contiguos de la secuencia transcrita, es decir, actúan en *cis*. Sin embargo, algunos de los elementos reguladores de la transcripción, tales como potenciadores, no necesitan ubicarse físicamente contiguos o muy próximas de las secuencias codificantes cuya transcripción mejoran.

Cuando se encuentran presentes múltiples promotores, cada promotor puede ser igual o diferente.

Los vectores recombinantes también pueden contener (a) una o más señales secretoras que codifican secuencias de péptidos señal, para permitir la secreción de un polipéptido expresado definido en el presente documento de la célula que produce dicho polipéptido o que proporcione la localización del polipéptido expresado, por ejemplo para retener el polipéptido en el retículo endoplasmático (ER) de la célula o transferirlo a un plástido y/o (b) secuencias de fusión que conducen a la expresión de moléculas de ácidos nucleicos como proteínas de fusión. Los ejemplos de segmentos de señal adecuados incluyen cualquier segmento de señal capaz de dirigir la secreción o la localización de un polipéptido definido en el presente documento. Los segmentos señal preferidos incluyen, pero en un sentido no limitativo, el péptido señal de *Nicotiana glauca* (US 5.939.288), la señal de la extensina de tabaco o la señal de la proteína de unión a cuerpos oleosos de la oleosina de soja. Los vectores recombinantes también pueden incluir secuencias interviniendo y/o no traducidas que rodean y/o que se encuentran dentro de la secuencia de ácidos nucleicos de un polinucleótido definido en el presente documento.

Para facilitar la identificación de los transformantes, el vector recombinante comprende preferentemente un gen de un marcador de selección o rastreable como secuencia de ácidos nucleicos de un polinucleótido definido en el presente documento o además de la misma. Un "gen marcador" se refiere a un gen que imparte un fenotipo distinto a las células que expresan al gen marcador y por consiguiente permite diferenciar dicha célula transformada de las células que no contienen al marcador. Un gen marcador de selección confiere una característica que se puede "seleccionar" basado en la resistencia a un agente selectivo (por ejemplo, un herbicida, un antibiótico, radiación, calor u otro tratamiento capaz de dañar células no transformadas). Un gen marcador rastreable (o gen informante) confiere un rasgo que se pueda identificar por observación o evaluación, es decir, mediante un "examen" (por ejemplo,  $\beta$ -glucuronidasa, luciferasa, GFP u otra actividad enzimática que no está presente en células no transformadas). El gen marcador y la secuencia de nucleótidos interés no tienen que estar unidos, dado que la co-transformación de genes no ligados como se describe, por ejemplo, en el documento US 4.399.216 también constituye un proceso eficiente, por ejemplo, en la transformación de plantas. La elección real de un marcador no es crucial en tanto sea funcional (es decir, selectivo) en combinación con las células de elección, tal como una célula vegetal.

Los ejemplos de marcadores de selección bacterianos son marcadores que confieren resistencia a antibióticos tales como resistencia a ampicilina, eritromicina, cloramfenicol o tetraciclina, preferentemente resistencia a kanamicina. Los ejemplos de marcadores de selección para la selección de transformantes de plantas incluyen, pero en un sentido no limitativo, un gen *hyg* que codifica resistencia a higromicina B; un gen de neomicina fosfotransferasa (*nptII*) que confiere resistencia a kanamicina, paromomicina, G418; un gen de glutatión-S-transferasa de hígado de rata que confiere resistencia a herbicidas derivados de glutatión como se describe, por ejemplo, en el documento EP 256223; un gen de la glutamina sintetasa que confiere, con la sobreexpresión, resistencia a los inhibidores de la glutamina sintetasa tal como fosfinotricina como se describe, por ejemplo, en el documento WO 87/05327, un gen de la acetiltransferasa de *Streptomyces viridochromogenes* que confiere resistencia al agente selectivo de fosfinotricina como se describe, por ejemplo, en el documento EP 275957, un gen que codifica una 5-enolshiquimato-3-fosfato sintasa (EPSPS) que confiere tolerancia a N-fosfonometilglicina como describen, por ejemplo, Hinchee et al., (1988), un gen *bar* que confiere resistencia contra bialafos como se describe, por ejemplo, en el documento WO91/02071; un gen de nitrilasa tal como *bxn* de *Klebsiella ozaenae* que confiere resistencia a bromoxinilo (Stalker et al., 1988); un gen de la dihidrofolato reductasa (DHFR) que confiere resistencia a metotrexato (Thillet et al., 1988); un gen mutante de acetolactato sintasa (ALS), que confiere resistencia a imidazolinona, sulfonilurea u otras sustancias químicas inhibidoras de ALS (EP 154,204); un gen mutado de la antranilato sintasa que confiere resistencia a 5-metilriptofano; o un gen de dalapón deshalogenasa que confiere resistencia al herbicida.

Los marcadores rastreables preferidos incluyen, pero en un sentido no limitativo, un gen *uidA* que codifica una enzima  $\beta$ -glucuronidasa (GUS) para la cual se conocen diversos sustratos cromogénicos; un gen de  $\beta$ -galactosidasa que codifica una enzima para la cual se conocen sustratos cromogénicos; un gen de aecuatorina (Prasher et al., 1985) que se puede emplear en la detección de bioluminiscencia sensible a calcio; un gen de la proteína fluorescente verde (Niedz et al., 1995) o derivados de los mismos; o un gen de luciferasa (*luc*) (Ow et al., 1986) que permite la detección de bioluminiscencia. Una "molécula informante" significa una molécula que, por su naturaleza química,

proporciona una señal analíticamente identificable que facilita la determinación de la actividad del promotor por referencia al producto proteico.

- 5 Preferentemente, el vector recombinante se incorpora de manera estable en el genoma de la célula, tal como la célula vegetal. Por consiguiente, el vector recombinante puede comprender elementos apropiados que permiten incorporar al vector en el genoma o en un cromosoma de la célula.

#### Vector de expresión

- 10 Tal como se usa en el presente documento, un vector de expresión es un vector de ADN o ARN que tiene la capacidad de transformar una célula hospedadora y de lograr la expresión de uno o más polinucleótidos específicos. Preferentemente, el vector de expresión también permite la replicación dentro de una célula hospedadora. Los vectores de expresión pueden ser procariotas o eucariotas y normalmente son virus o plásmidos. Los vectores de expresión de la presente invención incluyen cualquier vector que funciona (es decir, expresión directa de genes) en células hospedadoras de la presente invención, incluyendo en células bacterianas, fúngicas, de endoparásitos, de artrópodos, de animales, algas y vegetales. Los vectores de expresión particularmente preferidos de la presente invención pueden dirigir la expresión de genes en células de levadura, algas y/o vegetales.

- 20 Los vectores de expresión de la presente invención contienen secuencias reguladoras tales como secuencias de control de transcripción, secuencias de control de la traducción, orígenes de replicación y otras secuencias reguladoras que son compatibles con la célula hospedadora y que controlan la expresión de los polinucleótidos de la presente invención. En particular, los vectores de expresión de la presente invención incluyen secuencias de control de la transcripción. Las secuencias de control de la transcripción son secuencias que controlan el comienzo, la elongación y terminación de la transcripción. Las secuencias de control de la transcripción particularmente importantes son aquellas que controlan el comienzo de la transcripción, tales como secuencias promotoras, potenciadoras, operadoras y represoras. Las secuencias de control de la transcripción adecuadas incluyen cualquier secuencia de control de la transcripción que pueda funcionar en al menos una de las células recombinantes de la presente invención. La elección de las secuencias reguladoras usadas depende del organismo blanco, tal como una planta y/u órgano blanco o tejido de interés. Dichas secuencias reguladoras se pueden obtener de cualquier organismo eucariota tal como plantas o virus de plantas o se pueden sintetizar químicamente. Los expertos en la materia conocen una gran variedad de tales secuencias de control de la transcripción. Las secuencias de control de la transcripción particularmente preferidas son promotores activos en la dirección de la transcripción en plantas, ya sea constitutivamente o específicos de etapa y/o tejidos, dependiendo del uso de la planta o de partes de la misma.

- 35 Se han descrito numerosos vectores adecuados para una transfección estable de células vegetales o para establecer plantas transgénicas, por ejemplo, en Pouwels *et al.*, *Cloning Vectors: A Laboratory Manual*, 1985, suplemento 1987; Weissbach y Weissbach, *Methods for Plant Molecular Biology*, Academic Press, 1989; y Gelvin *et al.*, *Plant Molecular Biology Manual*, Kluwer Academic Publishers, 1990. normalmente, los vectores de expresión en plantas incluyen, por ejemplo, uno o más genes vegetales clonados bajo el control de transcripción de secuencias reguladoras 5' y 3' y un marcador de selección dominante. Dichos vectores de expresión en plantas también pueden contener una región reguladora promotora (por ejemplo, una región reguladora que controla una expresión inducible o constitutiva, regulada por el medio ambiente o por el desarrollo o específica de células o tejidos), un sitio de inicio de la transcripción, un sitio de unión a ribosoma, una señal de procesamiento de ARN, un sitio de terminación de la transcripción y/o una señal de poliadenilación.

- 45 Se han descrito numerosos promotores constitutivos que son activos en células vegetales. Los promotores adecuados para una expresión constitutiva en plantas incluyen, pero en un sentido no limitativo, el promotor 35S del virus en mosaico de la coliflor (CaMV), el promotor del 35S virus en mosaico de la escrofularia (FMV), el promotor del virus baciliforme de la caña de azúcar, el promotor del virus moteado amarillo de la comelina, el promotor inducible por luz de la subunidad pequeña de la ribulosa-1,5-bisfosfatocarboxilasa, el promotor de la triosafosfato isomerasa citosólica de arroz, el promotor de la adenina fosforibosiltransferasa de *Arabidopsis*, el promotor del gen de la actina 1 de arroz, los promotores de la manopina sintasa y octopina sintasa, el promotor Adh, el promotor de la sacarosa sintasa, el promotor del complejo del gen R y el promotor del gen de la proteína de unión a clorofila  $\alpha/\beta$ . Estos promotores se han usado para crear vectores de ADN que se han expresado en plantas; véase, por ejemplo, WO 84/02913. Todos estos promotores se han usado para crear diversos tipos de vectores de ADN recombinantes de expresión en plantas.

- 60 Con el fin de obtener expresión en los tejidos fuente de la planta, tal como hojas, semillas, raíz o tallo, se prefiere que los promotores utilizados en el presente documento invención sean de una expresión relativamente alta en estos tejidos específicos. Para este fin, se puede elegir entre una cantidad de promotores para genes con expresión específica o mejorada para tejidos o células. Los ejemplos de tales promotores informados en la bibliografía incluyen el promotor GS2 de la glutamina sintetasa de cloroplasto de guisante, el promotor de la fructosa-1,6-bisfosfatasa de cloroplastos de trigo, el promotor ST-LS1 fotosintética nuclear de patata, el promotor de serina/treonina cinasa y el promotor de glucoamilasa (CHS) de *Arabidopsis thaliana*. También se informaron como activos en los tejidos fotosintéticamente activos al promotor de la ribulosa-1,5-bisfosfato carboxilasa del alerce del este (*Larix laricina*), el promotor del gen Cab, Cab6, de pino, el promotor del gen Cab-1 de trigo, el promotor del gen Cab-1 de espinaca, el

promotor del gen Cab 1R de arroz, el promotor del gen de piruvato, el promotor de ortofosfato dicinasa (PPDK) de *Zea mays*, el promotor del gen Lhcb1\*2 de tabaco, el promotor Suc2 sacarosa-H<sup>30</sup> symporter de *Arabidopsis thaliana* y el promotor de los genes de proteínas de la membrana tilacoide de espinaca (PsaD, PsaF, PsaE, PC, FNR, AtpC, AtpD, Cab, RbcS). También se pueden utilizar otros promotores de las proteínas de unión a clorofila  $\alpha/\beta$  en el presente documento invención, tales como los promotores del gen Lhcb y del gen PsbP de mostaza blanca (*Sinapis alba*).

Hay varios promotores de genes vegetales que son regulados como respuesta a señales del medio ambiente, hormonales, sustancias químicas y/o del desarrollo que también se pueden usar para la expresión de genes de proteínas de unión a ARN en células vegetales, incluyendo promotores regulados por (1) calor, (2) luz (por ejemplo, el promotor RbcS-3A de guisante, el promotor RbcS de maíz); (3) hormonas, tales como ácido abscísico, (4) lesiones (por ejemplo, Wun1); o (5) sustancias químicas, tal como jasmonato de metilo, ácido salicílico, hormonas esteroides, alcohol, protectores (WO 97/06269) o también puede ser ventajoso emplear (6) promotores específicos de órganos.

Tal como se usa en el presente documento, el término "promotor específico del órgano de almacenamiento de una planta" se refiere a un promotor que preferencialmente, en comparación con otros tejidos vegetales, dirige la transcripción de genes en un órgano de almacenamiento de una planta. Preferentemente, el promotor solamente dirige la expresión de un gen de interés en el órgano de almacenamiento y/o la expresión del gen de interés en otras partes de la planta, tal como las hojas, no es detectable mediante un análisis de transferencia Northern y/o RT-PCR. Normalmente, el promotor dirige la expresión de genes durante el crecimiento y desarrollo del órgano de almacenamiento, en particular durante la fase de síntesis y acumulación de compuestos de almacenamiento en el órgano de almacenamiento. Tales promotores pueden dirigir la expresión de genes en todo el órgano de almacenamiento de una planta o solamente en una parte del mismo, tal como el recubrimiento, el embrión o los cotiledones en semillas de plantas dicotiledóneas o el endosperma o la capa de aleurona de semillas de plantas monocotiledóneas.

Con el fin de obtener expresión en los tejidos de almacenamiento o reserva de la planta, tales como el tubérculo de la planta de patata, el fruto de tomate o las semillas de soja, colza, algodón, *Zea mays*, trigo, arroz y cebada, se prefiere que los promotores utilizados en la invención tengan una expresión relativamente alta en estos tejidos específicos. Se conocen numerosos promotores de genes con expresión específica o mejorada en tubérculos, incluyendo el promotor de patatina de clase I, el promotor para los genes ADPGPP de tubérculos de patata, ambas subunidades grande y pequeña, el promotor de sacarosa sintasa, el promotor de las proteínas mayores de tubérculos incluyendo los complejos proteicos de 22 kD e inhibidores de proteinasas, el promotor para el gen de la almidón sintasa unida a gránulos (GBSS) y otros promotores de patatinas de clase I y II. También se pueden emplear otros promotores para expresar una proteína en tejidos específicos, tales como semillas o frutas. Se puede usar el promotor de la  $\beta$ -conglucina u otros promotores específicos de semillas tales como los promotores de napina, zeína, linina y faseolina. También se pueden usar promotores específicos de raíces. Un ejemplo de un promotor tal es el promotor del gen de la ácido quitinasa. La expresión en los tejidos de la raíz también se podría lograr utilizando los subdominios específicos de la raíz del promotor CaMV 35S que han sido identificados.

En una realización particularmente preferida, el promotor dirige la expresión en tejidos y órganos en los cuales tiene lugar la biosíntesis de lípidos. Dichos promotores actúan en el desarrollo de las semillas en el momento adecuado para modificar la composición del lípido en las semillas.

En otra realización preferida particular, el promotor es promotor específico de los órganos de almacenamiento de una planta. En una realización, el promotor específico del órgano de almacenamiento vegetal es un promotor específico de semillas. En una realización más preferida, el promotor dirige preferencialmente la expresión en los cotiledones de una planta dicotiledónea o en el endosperma de una planta monocotiledónea, con relación a la expresión en el embrión de la semilla o con relación a otros órganos en la planta, tal como las hojas. Los promotores preferidos para una expresión específica de semillas incluyen: 1) promotores de genes que codifican enzimas relacionadas con la biosíntesis y la acumulación en semillas de lípidos tales como desaturasas y elongasas, 2) promotores de genes que codifican proteínas de almacenamiento en semillas y 3) promotores de genes que codifican enzimas relacionadas con la biosíntesis y acumulación en semillas de carbohidratos. Los promotores específicos de semillas que son adecuados comprenden el promotor del gen de napina de colza oleaginosa (US 5.608.152), el promotor de USP de *Vicia faba* (Baumlein et al., 1991), el promotor de oleosina de *Arabidopsis* (WO 98/45461), el promotor de faseolina de *Phaseolus vulgaris* (US 5.504.200), el promotor Bce4 de *Brassica* (WO 91/13980) o el promotor de legumina B4 (Baumlein et al., 1992) y los promotores que conducen a la expresión específica de semillas en monocotiledóneas tales como maíz, cebada, trigo, centeno, arroz y semejantes. Los promotores adecuados más importantes son el promotor del gen lpt2 o lpt1 de cebada (WO 95/15389 y WO 95/23230) o los promotores que se describen en el documento WO 99/16890 (promotores del gen de la hordeína de cebada, el gen de glutelina de arroz, el gen de orizina de arroz, el gen de prolamina de arroz, el gen de gliadina de trigo, el gen de glutelina de trigo, el gen de zeína de maíz, el gen de glutelina de avena, el gen de kasirina de sorgo, el gen de secalina de centeno). Otros promotores incluyen los que se describen en Broun et al., (1998), Potenza et al., (2004), US 20070192902 y US 20030159173. En una realización, el promotor específico de semillas se expresa preferencialmente en partes definidas de las semillas tales como en los cotiledones o en el endosperma. Los

ejemplos de promotores específicos de cotiledones incluyen, pero en un sentido no limitativo, el promotor de FP1 (Ellerstrom et al., 1996), el promotor de legumina de guisante (Perrin et al., 2000) y el promotor de fitohemaglutina de haba (Perrin et al., 2000). Los ejemplos de promotores específicos del endosperma incluyen, pero en un sentido no limitativo, el promotor de zeína-1 de maíz (Chikwamba et al., 2003), el promotor de glutelina-1 de arroz (Yang et al., 2003), el promotor de D-hordeína de cebada (Horvath et al., 2000) y los promotores de HMW glutenina de trigo (Alvarez et al., 2000). En una realización adicional, el promotor específico de semillas no se expresa o solamente se expresa a un nivel bajo, en el embrión y/o después que germinó la semilla.

En otra realización, el promotor específico del órgano de almacenamiento vegetal es un promotor específico de tubérculos. Los ejemplos incluyen, pero en un sentido no limitativo, los promotores B33, PAT21 y GBSS de patatina de patata, así como el promotor de esporamina de batata (por una revisión véase Potenza et al., 2004). En una realización preferida, el promotor dirige la expresión preferencialmente en la médula del tubérculo, con relación a las capas externas (piel, corteza) o el embrión del tubérculo.

En otra realización, el promotor específico del órgano de almacenamiento vegetal es un promotor específico de frutos. Los ejemplos incluyen, pero en un sentido no limitativo, los promotores E8 y Pds de la poligalacturonasa de tomate, así como el promotor ACC oxidasa de manzana (por una revisión véase Potenza et al., 2004). En una realización preferida, el promotor dirige preferencialmente la expresión en las partes comestibles de la fruta, por ejemplo la médula de la fruta, con relación a la piel de la fruta o las semillas dentro del fruto.

La secuencia líder no traducida 5' puede derivar del promotor seleccionado para expresar la secuencia del gen heterólogo del polinucleótido de la presente invención o puede ser heteróloga con respecto a la región codificante de la enzima a producir y se puede modificar específicamente, si se desea, para incrementar la traducción del ARNm. Por una revisión sobre la optimización de la expresión de transgenes, véase Koziel et al., (1996). Las regiones no traducidas 5' también se pueden obtener de los ARN virales de plantas (virus en mosaico del tabaco, virus del grabado de la hoja de tabaco, virus en mosaico enano de maíz, virus en mosaico enano de alfalfa, entre otros), de genes eucariotas adecuados, de genes vegetales (directriz del gen de la proteína de unión a clorofila a/b de trigo y maíz) o de una secuencia genética sintética. La presente invención no se limita a las construcciones en donde la región no traducida deriva de la secuencia no traducida 5' que acompaña a la secuencia del promotor. La secuencia líder también podría derivar de una secuencia codificante o de un promotor no relacionado. Las secuencias líder de utilidad en el contexto de la presente invención comprenden la directriz Hsp70 de maíz (US 5.362.865 y US 5.859.347) y el elemento TMV omega.

La terminación de transcripción se efectúa con una secuencia de ADN no traducida 3' ligada operativamente en el vector de expresión al polinucleótido de interés. La región no traducida 3' de una molécula de ADN recombinante contiene una señal de poliadenilación cuya función en plantas causa la adición de nucleótidos adenilato al extremo 3' del ARN. La región no traducida 3' se puede obtener de diversos genes que se expresan en células vegetales. La región no traducida 3' de la nopalina sintasa, la región no traducida 3' del gen de la subunidad pequeña Rubisco de guisante, la región no traducida 3' de la proteína de almacenamiento 7S de soja en genes de semillas son de uso común en este sentido. También son adecuadas las regiones no traducidas transcritas 3', que contienen la señal de poliadenilato de los genes del plásmido inductor de tumores (Ti) de *Agrobacterium*.

Se pueden usar tecnologías de ADN recombinante para mejorar la expresión de un polinucleótido transformado mediante manipulación, por ejemplo, la cantidad de copias del polinucleótido dentro de una célula hospedadora, la eficacia con la cual se transcriben dichos polinucleótidos, la eficacia con que se traducen los transcritos resultantes y la eficacia de las modificaciones de post-traducción. Las técnicas recombinantes de utilidad para incrementar la expresión de los polinucleótidos definidos en el presente documento incluyen, pero en un sentido no limitativo, unión operativa del polinucleótido a plásmidos con un gran número de copias, integración del polinucleótido en uno o más cromosomas de las células hospedadoras, adición de secuencias estabilizadoras del vector al plásmido, sustituciones o modificaciones de las señales de control de la transcripción (por ejemplo, promotores, operadores, potenciadores), sustituciones o modificaciones de las señales de control de la traducción (por ejemplo, sitios de unión a ribosomas, secuencias Shine-Dalgarno), modificación del polinucleótido para que se correspondan con la utilización de codones de la célula hospedadora y supresión de las secuencias que desestabilizan a los transcritos.

### Ácidos nucleicos de transferencia

Se pueden usar ácidos nucleicos de transferencia para administrar un polinucleótido exógeno a una célula y comprende una, preferentemente dos, secuencias de bordes y un polinucleótido de interés. El ácido nucleico de transferencia puede codificar o no, un marcador de selección. Preferentemente, el ácido nucleico de transferencia forma parte de un vector binario en la bacteria, donde dicho vector binario comprende además elementos que permiten la replicación del vector en la bacteria o permite la selección o conservación de células bacterianas que contienen al vector binario. Con la transferencia a una célula eucariota, el componente de ácido nucleico de transferencia del vector binario tiene la capacidad de integrarse en el genoma de la célula eucariota.

Tal como se usa en el presente documento, el término "ácido nucleico de transferencia extracromosómico" se refiere a una molécula de ácido nucleico que tiene la capacidad de poder ser transferida de una bacteria, tal como

*Agrobacterium sp.*, a una célula eucariota, tal como una célula foliar vegetal. Un ácido nucleico de transferencia extracromosómico es un elemento genético reconocido como un elemento que puede ser transferido con la subsiguiente integración de la secuencia de nucleótidos contenida dentro de sus bordes en el genoma de la célula receptora. En este sentido, un ácido nucleico de transferencia normalmente está flanqueado por dos secuencias de "bordes", aunque en algunos casos se puede usar un solo borde por uno de los extremos y el segundo extremo del ácido nucleico transferido es generado aleatoriamente en el proceso de transferencia. Un polinucleótido de interés se ubica normalmente entre la secuencia tipo borde izquierdo y la secuencia tipo borde derecho de un ácido nucleico de transferencia. El polinucleótido contenido dentro del ácido nucleico de transferencia se puede ligar operativamente a varios diferentes elementos reguladores promotores y terminadores que facilitan su expresión, es decir, la transcripción y/o traducción del polinucleótido. Los ADN-T de *Agrobacterium sp.* tal como de *Agrobacterium tumefaciens* o de *Agrobacterium rhizogenes* y variantes/mutantes de los mismos elaboradas por el hombre son probablemente los ejemplos mejor caracterizados de ácidos nucleicos de transferencia. Otro ejemplo es el ADN-P ("ADN de plantas") que comprende secuencias tipo borde de ADN-T de plantas.

Tal como se usa en el presente documento, "ADN-T" se refiere, por ejemplo, al ADN-T de un plásmido Ti de *Agrobacterium tumefaciens* o de un plásmido Ri de *Agrobacterium rhizogenes* o variantes de los mismos elaboradas por el hombre que funcionan como ADN-T. El ADN-T puede comprender un ADN-T completo incluyendo ambas secuencias de bordes derecho e izquierdo, pero solamente es necesario que comprenda las secuencias mínimas requeridas en *cis* para la transferencia, es decir, la secuencia del borde derecho y de ADN-T. En los ADN-T de la invención se han insertado, en cualquier parte entre las secuencias de bordes derecho e izquierdo (si está presente), el polinucleótido de interés flanqueado por sitios blanco para una recombinasa específica del sitio. Las secuencias que codifican los factores requeridos en *trans* para la transferencia del ADN-T en una célula vegetal, tal como los genes *vir*, se pueden insertar en el ADN-T o pueden estar presentes sobre el mismo replicón que el ADN-T o preferentemente se encuentran en *trans* sobre un replicón compatible en el hospedador de *Agrobacterium*. Dichos "sistemas de vectores binarios" son bien conocidos en la técnica.

Tal como se usa en el presente documento, el "ADN-P" se refiere a un ácido nucleico de transferencia aislado de un genoma vegetal o variantes/mutantes del mismo elaboradas por el hombre y comprende por cada extremo o por un solo extremo, una secuencia tipo borde de ADN-T. La secuencia tipo borde comparte preferentemente al menos un 50 %, al menos un 60 %, al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 90 % o al menos un 95 %, pero menos que un 100 % de identidad de secuencia, con una secuencia de borde de ADN-T de *Agrobacterium sp.*, tal como de *Agrobacterium tumefaciens* o de *Agrobacterium rhizogenes*. Por consiguiente, se pueden usar ADN-P en lugar de ADN-T para transferir una secuencia de nucleótidos contenida en el ADN-P, por ejemplo de *Agrobacterium* a otra célula. El ADN-P, antes de la inserción del polinucleótido exógeno que será transferido, puede modificarse para facilitar la clonación y preferentemente no debería codificar ninguna proteína. El ADN-P se caracteriza por contener al menos una secuencia de borde derecho y preferentemente también una secuencia de borde izquierdo.

Tal como se usa en el presente documento, las secuencias de "bordes" de un ácido nucleico de transferencia se pueden aislar de un organismo seleccionado, tal como una planta o bacteria o puede ser una variante/mutante de la misma elaborada por el hombre. La secuencia del borde promueve y facilita la transferencia del polinucleótido con el cual está ligado y puede facilitar su integración en el genoma de la célula receptora. En una realización, una secuencia de borde es de entre 5-100 pares de bases (pb) de longitud, 10-80 pb de longitud, 15-75 pb de longitud, 15-60 pb de longitud, 15-50 pb de longitud, 15-40 pb de longitud, 15-30 pb de longitud, 16-30 pb de longitud, 20-30 pb de longitud, 21-30 pb de longitud, 22-30 pb de longitud, 23-30 pb de longitud, 24-30 pb de longitud, 25-30 pb de longitud o 26-30 pb de longitud. Las secuencias de bordes del ADN-T de *Agrobacterium sp.* son bien conocidas en la técnica e incluyen los que se describen en Lacroix et al. (2008), Tzfira y Citovsky (2006) y Glevin (2003).

Si bien tradicionalmente sólo se ha usado *Agrobacterium sp.* para transferir genes a células vegetales, existe ahora una gran cantidad de sistemas que fueron identificados/desarrollados y que actúan de una manera similar a *Agrobacterium sp.* Recientemente se han modificado genéticamente varias especies distintas de *Agrobacterium* para que sean competentes para la transferencia de genes (Chung et al., 2006; Broothaerts et al., 2005). Estas especies incluyen *Rhizobium sp.* NGR234, *Sinorhizobium meliloti* y *Mezorhizobium loti*. Las bacterias se hacen competentes para la transferencia de genes proporcionándoles a las bacterias la maquinaria necesaria para el proceso de transformación: es decir un conjunto de genes de virulencia codificado por un plásmido Ti de *Agrobacterium* y el segmento de ADN-T que reside sobre un pequeño plásmido binario separado. Las bacterias manipuladas de esta manera tienen la capacidad de transformar diferentes tejidos vegetales (tejido de discos foliares, callos y ovario), de monocotiledóneas o dicotiledóneas y de diversas especies vegetales diferentes (por ejemplo, tabaco, arroz).

La transferencia directa de plásmidos de expresión eucariotas de bacterias a hospedadores eucariotas se logró por primera vez hace varias décadas por fusión de células de mamífero y protoplastos de *Escherichia coli* portadores de plásmidos (Schaffner, 1980). Desde entonces, ha ido aumentando de manera constante la cantidad de bacterias capaces de distribuir genes en células de mamífero (Weiss, 2003), habiendo sido descubiertos por cuatro grupos de forma independiente (Sizemore et al., 1995; Courvalin et al., 1995; Powell et al., 1996; Darji et al., 1997).

Se ha demostrado que *Shigella flexneri*, *Salmonella typhimurium* o *E. coli* atenuadas que se volvieron invasivas por

el plásmido de virulencia (pWR100) de *S. flexneri* pueden transferir plásmidos de expresión después de la invasión en células hospedadoras y muerte intracelular debido a una atenuación metabólica. La aplicación en mucosas, ya sea por vía nasal u oral, de tales organismos recombinantes de *Shigella* o *Salmonella* indujo respuestas inmunes contra el antígeno codificado por los plásmidos de expresión. Mientras tanto, se había más que duplicado la lista de bacterias capaces de transferir plásmidos de expresión una células hospedadoras de mamífero *in vitro* e *in vivo* y fue documentado para *S. typhi*, *S. choleraesuis*, *Listeria monocytogenes*, *Yersinia pseudotuberculosis* e *Y. enterocolitica* (Fennelly et al., 1999; Shiau et al., 2001; Dietrich et al., 1998; Hense et al., 2001; Al-Mariri et al., 2002).

En general, se podría asumir que todas las bacterias capaces de ingresar en el citosol de la célula hospedadora (como *S. flexneri* o *L. monocytogenes*) y lisarse dentro de este compartimento celular, deberían poder transferir ADN. Esto se conoce como una invasión 'abortiva' o 'suicida' ya que las bacterias deben lisarse para que tenga lugar la transferencia de ADN (Grillot-Courvalin et al., 1999). Además, aún muchas de las bacterias que permanecen en la vacuola fagocítica (como *S. typhimurium*) también podrían hacerlo. Por consiguiente, las cepas de laboratorio recombinantes de *E. coli* que han sido manipuladas para ser invasivas pero incapaces de un escape fagosómico, podrían no obstante distribuir su carga de plásmidos en el núcleo de la célula de mamífero infectada (Grillot-Courvalin et al., 1998). Además, recientemente también se ha demostrado que *Agrobacterium tumefaciens* puede introducir transgenes en células de mamífero (Kunik et al., 2001).

Tal como se usa en el presente documento, los términos "transfección", "transformación" y variaciones de los mismos en general se usan indistintamente. Las células "transfectadas" o "transformadas" pueden haber sido manipuladas para introducir el o los polinucleótidos de interés o pueden ser células de la progenie derivadas de ellas.

#### Células recombinantes

La invención también proporciona una célula recombinante, por ejemplo, una célula vegetal recombinante, que es una célula hospedadora transformada con uno o más polinucleótidos o vectores definidos en el presente documento o una combinación de los mismos. El término "célula recombinante" se usa indistintamente con el término "célula transgénica" en el presente documento. Las células adecuadas de la invención incluyen cualquier célula que se pueda transformar con un polinucleótido o un vector recombinante de la invención, que codifica, por ejemplo, un polipéptido o una enzima descrita en el presente documento. La célula es preferentemente una célula que de esa manera se pueda usar para producir lípidos. La célula recombinante puede ser una célula en cultivo, una célula *in vitro* o presente en un organismo, tal como por ejemplo una planta o en un órgano, tal como por ejemplo una semilla o una hoja. Preferentemente, la célula se encuentra en una planta, más preferentemente en la semilla de una planta.

Las células hospedadoras en las cuales se introducen los polinucleótido pueden ser células no transformadas o células que ya fueron transformadas con al menos un ácido nucleico. Dichos ácidos nucleicos pueden estar relacionados o no, con la síntesis de lípidos. Las células hospedadoras de la presente invención pueden tener capacidad endógena (es decir, natural) para producir los polipéptidos definidos en el presente documento, en cuyo caso la célula recombinante derivada de los mismos tienen una mayor capacidad para producir los polipéptidos o pueden tener la capacidad de producir tales polipéptidos solamente después de su transformación con al menos un polinucleótido de la invención. En una realización, la célula recombinante de la invención tiene una capacidad mejorada para producir lípidos no polares.

Las células hospedadoras de la presente invención pueden ser cualquier célula capaz de producir al menos una de las proteínas descritas en el presente documento, seleccionadas entre células fúngicas (incluyendo de levadura), de algas y vegetales. Las células hospedadoras preferidas son células de levadura, algas y vegetales. Los ejemplos de células de algas de utilidad como células hospedadoras de la presente invención incluyen, por ejemplo, *Chlamydomonas sp* (por ejemplo, *Chlamydomonas reinhardtii*), *Dunaliella sp.*, *Haematococcus sp.*, *Chlorella sp.*, *Thraustochytrium sp.*, *Schizochytrium sp.* y *Volvox sp.*

Las células hospedadoras para la expresión de los ácidos nucleicos de la presente pueden incluir hospedadores microbianos que crecen sobre varias alimentaciones, incluyendo carbohidratos simples o complejos, ácidos orgánicos y alcoholes y/o hidrocarburos, son un amplio rango de valores de temperatura y pH. Los hospedadores microbianos preferidos son organismos oleaginosos que tienen una capacidad natural de sintetizar lípidos no polares.

Las células hospedadoras pueden ser de un organismo adecuado para un proceso de fermentación, tal como, por ejemplo, *Yarrowia lipolytica* u otras levaduras.

#### Plantas transgénicas

La invención también proporciona una planta que comprende un polinucleótido o un polipéptido exógeno de la invención, una célula de la invención, un vector de la invención o una combinación de los mismos. El término "planta" se refiere a plantas completas, en tanto el término "parte de la misma" se refiere a órganos vegetales (por ejemplo, hojas, tallos, raíces, flores, frutos), células individuales (por ejemplo, polen), semillas, partes de semillas

tales como un embrión, endosperma, escutelo o recubrimiento de semillas, tejido vegetal tal como tejido vascular, células vegetales y la progenie de las mismas. Tal como se usa en el presente documento, las partes de plantas comprenden células vegetales.

- 5 Tal como se usa en el presente documento, el término "planta" se usa en su sentido más amplio. Incluye, pero en un sentido no limitativo, cualquier especie de pasto, planta ornamental o decorativa, cultivo o cereal (por ejemplo, semilla oleaginosa, maíz, soja), pienso o forraje, frutos o verduras, plantas herbáceas, plantas leñosas, plantas con flores o árboles. No pretende limitar una planta a ninguna estructura particular. También se refiere a una planta unicelular (por ejemplo, un microalga). El término "parte de la misma" con referencia a una planta se refiere a una
- 10 célula vegetal y a la progenie de la misma, a una pluralidad de células vegetales que se diferencian mayormente en una colonia (por ejemplo, volvox), una estructura que está presente en cualquier etapa del una desarrollo de una planta o un tejido vegetal. Dichas estructuras incluyen, pero en un sentido no limitativo, hojas, tallos, flores, frutos, nueces, raíces, semillas, recubrimiento de semillas, embriones. El término "tejido vegetal" incluye tejidos vegetales diferenciados y no diferenciados incluyendo los que están presentes en hojas, tallos, flores, frutos, nueces, raíces,
- 15 semillas, por ejemplo, tejido embrionario, endosperma, tejido dérmico (por ejemplo, epidermis, periderma), tejido vascular (por ejemplo, xilema, floema) o tejido básico (que comprende células de parénquima, colénquima y/o esclerénquima), así como células en cultivo (por ejemplo, células individuales, protoplastos, callos, embriones, etc.). El tejido vegetal puede estar *in planta*, en cultivo de órganos, en cultivo tisular o cultivo celular.
- 20 Una "planta transgénica", una "planta modificada genéticamente" o variaciones de las mismas, se refieren a una planta que contiene un transgen que no se encuentra en una planta de tipo silvestre de la misma especie, variedad o cultivar. Las plantas transgénicas definidas en el contexto de la presente invención incluyen plantas y su progenie que han sido modificadas genéticamente usando técnicas recombinantes para producir al menos un polipéptido definido en el presente documento en la planta o parte de la misma. Las partes de plantas transgénicas tienen un
- 25 significado correspondiente.

Los términos "semilla" y "grano" se usan indistintamente en el presente documento. Un "grano" se refiere a un grano maduro, tal como un grano cosechado o un grano que aún se encuentra sobre la planta pero que está listo para ser cosechado y también se refiere a un grano después de embeberlo o hacerlo germinar, según el contexto. Los granos

30 maduros comúnmente tienen un contenido de humedad menor que un 18-20 % aproximadamente. Una "semilla en desarrollo", tal como se usa en el presente documento, se refiere a una semilla antes de la madurez, normalmente presente en las estructuras reproductoras de la planta después de la fertilización o antesis, pero también se refieren a las semillas antes de la madurez que se pueden aislar de la planta.

- 35 Tal como se usa en el presente documento, el término "órgano de almacenamiento de una planta" se refiere a una parte de una planta especializada en guardar energía en la forma de, por ejemplo, proteínas, carbohidratos, lípidos. Los ejemplos de órganos de almacenamiento de una planta son semillas, frutos, raíces tuberosas y tubérculos. Un órgano de almacenamiento preferido de una planta de la invención es la semilla.
- 40 Tal como se usa en el presente documento, el término "fenotípicamente normal" se refiere a una planta o parte de la misma, en particular un órgano de almacenamiento tal como una semilla, tubérculo o fruto de la invención modificada genéticamente que no tiene una capacidad significativamente reducida para crecer y reproducirse en comparación con una planta o parte de la misma no modificada. En una realización, la planta o parte de la misma modificada genéticamente que es fenotípicamente normal comprende un polinucleótido recombinante que codifica
- 45 un supresor por silenciamiento ligado operativamente a un promotor específico del órgano de almacenamiento de una planta y que tiene la capacidad de crecer o reproducirse, esencialmente igual que una planta o parte de la misma correspondiente que no que comprende a dicho polinucleótido. Preferentemente, la biomasa, el índice de crecimiento, índice de germinación, el tamaño de los órganos de almacenamiento, el tamaño de las semillas y/o la cantidad de semillas viables producidas no son menores que un 90 % de los valores de una planta que no contiene
- 50 a dicho polinucleótido recombinante cuando se cultiva bajo condiciones idénticas. Este término no abarca las características de la planta que pueden ser diferentes de la planta de tipo silvestre pero que no afectan la utilidad de la planta para fines comerciales tal como, por ejemplo, un fenotipo *ballerina* de las hojas de las plántulas.

- Las plantas suministradas o contempladas para su uso en la práctica de la presente invención incluyen tanto
- 55 monocotiledóneas como dicotiledóneas. En formas de realización preferidas, las plantas de la presente invención son plantas de cultivo (por ejemplo, cereales y legumbres, maíz, trigo, patata, tapioca, arroz, sorgo, mijo, mandioca, cebada o guisante) u otras legumbres. Las plantas se pueden cultivar para la producción de raíces, tubérculos, hojas, tallos, flores o frutos comestibles. Las plantas pueden ser verduras o plantas ornamentales. Las plantas de la invención pueden ser: maíz (*Zea mays*), colza (*Brassica napus*, *Brassica rapa* ssp.), otras *Brassicaceae* tales como, por ejemplo, rutabaga (*Brassica napobrassica*) o *Brassica camelina*, remolacha azucarera (*Beta vulgaris*), trébol (*Trifolium* sp.), lino (*Linum usitatissimum*), alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*),
- 60 sorgo (*Sorghum bicolor*, *Sorghum vulgare*), girasol (*Helianthus annuus*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), cacahuete (*Arachis hypogaea*), algodón (*Gossypium hirsutum*), batata (*Lopmoea batatus*), mandioca (*Manihot esculenta*), café (*Coffea* spp.), coco (*Cocos nucifera*), piña (*Ananas comosus*), árboles cítricos (*Citrus* spp.), cacao (*Theobroma cacao*), té (*Camellia sinensis*), banana (*Musa* spp.), palta (*Persea americana*), higo (*Ficus casica*), guayava (*Psidium guajava*), mango (*Mangifer indica*), olivo
- 65



(*Olea europaea*), papaya (*Carica papaya*), cajú (*Anacardium occidentale*), macadamia (*Macadamia intergrifolia*), almendra (*Prunus amygdalus*), remolacha (*Beta vulgaris*), avena o cebada.

5 Otras plantas preferidas incluyen pastos C4 tales como *Andropogon gerardi*, *Bouteloua curtipendula*, *B. gracilis*, *Buchloe dactiloides*, *Panicum virgatum*, *Schizachyrium scoparium*, especies de *Miscanthus* por ejemplo, *Miscanthus x giganteus* y *Miscanthus sinensis*, *Sorghastrum nutans*, *Sporobolus cryptandrus*, pasto varilla (*Panicum virgatum*); pastos C3 tal como *Elymus canadensis*, las legumbres *Lespedeza capitata* y *Petalostemum villosum*, la margarita azul *Aster azureus*; y plantas leñosas tales como *Quercus ellipsoidalis* y *Q. macrocarpa*.

10 En una realización preferida, la planta es una angiosperma.

En una realización, la planta es una planta oleaginosa, preferentemente una planta de cultivo de semillas oleaginosas. Tal como se usa en el presente documento, una "planta oleaginosa" es una especie vegetal usada en la producción comercial de lípidos a partir de las semillas de la planta. La planta oleaginosa puede ser colza oleaginosa (tal como colza), maíz, girasol, cártamo, soja, sorgo, lino (semillas de lino) o remolacha. Además, la planta oleaginosa puede ser otras *Brassicas*, algodón, cacahuete, amapola, rutabaga, mostaza, semillas de ricino, sésamo, cártamo o plantas productoras de nueces. La planta puede producir niveles altos de lípidos en sus frutos, tal como olivo, palma oleaginosa o coco. Las plantas de horticultura en las cuales se puede aplicar la presente invención son lechuga, endivia o verduras de *brassicas* incluyendo coles, brócoli o coliflor. La presente invención se puede aplicar en tabaco, cucurbitáceas, zanahorias, fresas, tomate o pimienta.

En una realización preferida, la planta transgénica es homocigota para cada uno y todos los genes que fueron introducidos (transgen) de modo que su progenie no segregará el fenotipo deseado. La planta transgénica también puede ser heterocigota para los transgenes introducidos, preferentemente uniformemente heterocigotas para el transgen, tal como por ejemplo en la progenie F1 que fue cultivada a partir de semillas híbridas. Dichas plantas pueden ofrecer ventajas tales como el vigor híbrido, bien conocido en la técnica.

30 Cuando fuera relevante, las plantas transgénicas también pueden comprender transgenes adicionales que codifican enzimas relacionadas con la producción de lípidos no polares tales como, en un sentido no limitativo, LPAAT, LPCAT, PAP o una fosfolípido:diacilglicerol aciltransferasa (PDAT1, PDAT2 o PDAT3; véase por ejemplo, Ghosal et al., 2007) o una combinación de dos o más de los mismos. Las plantas transgénicas de la invención también pueden expresar oleosina a partir de un polinucleótido exógeno.

#### 35 *Transformación de plantas*

Las plantas transgénicas se pueden producir usando técnicas conocidas en la técnica, tales como las que se describen en general en A. Slater et al., *Plant Biotechnology - The Genetic Manipulation of Plants*, Oxford University Press (2003) y Christou y Klee, *Handbook of Plant Biotechnology*, John Wiley y Sons (2004).

40 Tal como se usa en el presente documento, los términos "transformación de manera estable", "transformado de manera estable" y variaciones de los mismos se refieren a la integración del polinucleótido en el genoma de la célula de modo tal que son transferidas a las células de la progenie durante la división celular sin necesidad de seleccionar positivamente su presencia. Los transformantes estables o la progenie de los mismos, se pueden seleccionar con cualquier medio conocido en la técnica tales como transferencias Southern con ADN cromosómico o hibridación *in situ* de ADN genómico.

La transferencia mediada por *Agrobacterium* es un sistema de amplia aplicación para introducir genes en células vegetales porque el ADN se puede introducir en las células en tejidos vegetales u órganos vegetales completos o en explantes en cultivo tisular, ya sea para una expresión transitoria o para una integración estable del ADN en el genoma de la célula vegetal. El uso de vectores de integración vegetales mediados por *Agrobacterium* para introducir ADN en células vegetales es bien conocido en la técnica (véase, por ejemplo, US 5177010, US 5104310, US 5004863 o US 5159135). La región de ADN a transferir está definida por las secuencias de los bordes y el ADN que interviene (ADN-T) comúnmente se inserta en el genoma vegetal. Además, la integración del ADN-T es un proceso relativamente preciso que da como resultado unos pocos reordenamientos. En aquellas variedades vegetales donde la transformación mediada por *Agrobacterium* es eficiente, es el método de elección debido a la naturaleza fácil y definida de la transferencia de genes. Los vectores preferidos para la transformación con *Agrobacterium* tienen la capacidad de replicarse en *E. coli*, así como en *Agrobacterium*, permitiendo manipulaciones convenientes como se describe en la bibliografía (Klee et al., En: *Plant DNA Infectious Agents*, Hohn y Schell (editores), Springer-Verlag, Nueva York, páginas 179-203 (1985)).

60 Los métodos de aceleración que se pueden usar incluyen, por ejemplo, bombardeo con microproyectiles y semejantes. Un ejemplo de un método para distribuir moléculas de ácidos nucleicos transformantes en células vegetales es el bombardeo con microproyectiles. Este método ha sido revisado por Yang et al., *Particle Bombardment Technology for Gene Transfer*, Oxford Press, Oxford, Inglaterra (1994). Se pueden recubrir partículas no biológicas (microproyectiles) con ácidos nucleicos y distribuir las en las células mediante una fuerza propulsora. Los ejemplos de partículas incluyen las que comprenden tungsteno, oro, platino y semejantes. Una ventaja particular

del bombardeo de microproyectiles, además de ser un medio eficaz para transformar monocotiledóneas de manera reproducible, es que no es necesario aislar los protoplastos ni susceptibilidad de una infección por *Agrobacterium*. Una realización ilustrativa de un método para distribuir ADN en células de *Zea mays* por aceleración es un sistema de distribución de partículas  $\alpha$  biolístico, que se puede usar para impulsar partículas recubiertas con ADN a través de una pantalla, tal como una pantalla de acero inoxidable o Nytex, sobre una superficie de filtro cubierta con células de maíz cultivadas en suspensión. Un sistema de distribución de partículas adecuado para su uso con la presente invención es la pistola de aceleración por helio PDS-1000/He disponible de Bio-Rad Laboratories.

Para el bombardeo, se pueden concentrar células en suspensión sobre filtros. Los filtros que contienen a las células que serán bombardeadas se ubican a una distancia apropiada debajo de la placa de detención de microproyectiles. Si se deseara, también se pueden colocar una o varias pantallas entre la pistola y las células a bombardear.

Como alternativa, se pueden disponer embriones inmaduros u otras células blanco sobre un medio de cultivo sólido. Las células que serán bombardeadas se ubican a una distancia apropiada debajo de la placa de detención de microproyectiles. Si se deseara, también se coloca una o varias pantallas entre el dispositivo de aceleración y las células que serán bombardeadas. El uso de las técnicas que se describen en este documento permitirá obtener hasta 1000 o más focos de células que expresan de forma transitoria un gen de un marcador rastreable o seleccionable. La cantidad de células de un foco que expresa el producto genético 48 horas después del bombardeo a menudo varía en un rango de uno a diez, con un promedio de uno a tres.

En la transformación por bombardeo, se pueden optimizar las condiciones de cultivo de pre-bombardeo y los parámetros de bombardeo con el fin de obtener la cantidad máxima de transformantes estables. Tanto los parámetros físicos como biológicos para el bombardeo son importantes en esta tecnología. Los factores físicos son aquellos que comprenden la manipulación del ADN/precipitado de microproyectiles o aquellos que afectan el vuelo y la velocidad de los macro- o microproyectiles. Los factores biológicos incluyen todas las etapas involucradas en la manipulación de las células antes e inmediatamente después del bombardeo, el ajuste osmótico de las células blanco para ayudar a aliviar el trauma asociado con el bombardeo y también la naturaleza del ADN transformante, tal como ADN alineado o plásmidos superenrollados intactos. Se considera que las manipulaciones de pre-bombardeo son especialmente importantes para una transformación exitosa de embriones inmaduros.

En otra realización alternativa, se pueden transformar plástidos de forma estable. Los métodos divulgados para la transformación de plástidos en plantas superiores incluyen distribución con pistola de partículas de ADN que contienen un marcador de selección y direccionamiento del ADN al genoma del plástido por medio de recombinación homóloga (US 451.513, US 5.545.818, US 5.877.402, US 5.932.479 y WO 99/05265).

Por consiguiente, se considera que puede ser necesario ajustar diversos aspectos de los parámetros de bombardeo en estudios a pequeña escala para optimizar completamente las condiciones. Puede ser particularmente deseable ajustar los parámetros físicos, tal como la distancia a la hendidura, la distancia de vuelo, la distancia al tejido y la presión de helio. También puede ser conveniente minimizar los factores de reducción de trauma modificando las condiciones que afectan el estado fisiológico de las células receptoras y que por ello pueden afectar la transformación y las eficacias de integración. Por ejemplo, se puede ajustar el estado osmótico, la hidratación del tejido y la etapa de subcultivo o el ciclo celular de las células receptoras para lograr una transformación óptima. La ejecución de otros ajustes de rutina constituye un aspecto conocido para los expertos en la materia a la luz de la presente descripción.

La transformación de protoplastos vegetales se puede efectuar usando métodos basados en precipitación por fosfato de calcio, tratamiento con polietilenglicol, electroporación y combinaciones de estos tratamientos. La aplicación de estos sistemas a diferentes variedades de plantas depende de la capacidad para regenerar dicha cepa vegetal particular a partir de protoplastos. Se han descrito métodos ilustrativos para la regeneración de cereales a partir de protoplastos (Fujimura et al., 1985; Toriyama et al., 1986; Abdullah et al., 1986).

También se pueden usar otros métodos de transformación de células, los cuales incluyen, por ejemplo, la introducción de ADN en plantas mediante la transferencia directa de ADN en el polen, la inyección directa de ADN en los órganos reproductores de una planta o la inyección directa de ADN en las células de embriones inmaduros, seguida por la rehidratación de embriones desecados.

La regeneración, el desarrollo y el cultivo de plantas a partir de protoplastos transformantes de plantas individuales o a partir de varios explantes transformados, son bien conocidos en la técnica (Weissbach et al., en: *Methods for Plant Molecular Biology*, Academic Press, San Diego, CA (1988)). Este proceso de regeneración y desarrollo normalmente incluye las etapas de selección de las células transformadas, cultivo de aquellas células individualizadas a través de las etapas usuales del desarrollo embrionario, hasta la etapa de plántula con raíz. Los embriones y las semillas transgénicas se regeneran de forma similar. Luego se siembran los brotes con raíces resultantes en un medio apropiado para el cultivo de plantas, tal como tierra.

El desarrollo o la regeneración de plantas que contienen al gen exógeno extraño son bien conocidos en la técnica. Preferentemente, las plantas regeneradas son autopolinizadas para proporcionar plantas transgénicas homocigotas.

Por otro lado, puede cruzarse el polen obtenido de las plantas regeneradas con plantas cultivadas a partir de semillas de líneas de importancia agrícola. Por otro lado, se usa el polen de plantas de estas líneas importantes para polinizar plantas regeneradas. Una planta transgénica de la presente invención que contiene un polinucleótido deseado se cultiva usando métodos bien conocidos por aquellos versados en la técnica.

5 Los métodos para transformar dicotiledóneas, primariamente mediante el uso de *Agrobacterium tumefaciens* y obtener plantas transgénicas fueron publicados para algodón (US 5.004.863, US 5.159.135, US 5.518.908); soja (US 5.569.834, US 5.416.011); *Brassica* (US 5.463.174); cacahuete (Cheng et al., 1996); y guisante (Grant et al., 1995).

10 Los métodos para transformar de plantas de cereales tales como trigo y cebada para introducir una variación genética en la planta por introducción de un ácido nucleico exógeno y para regenerar plantas a partir de protoplastos o de embriones inmaduros de las plantas son bien conocidos en la técnica, véase por ejemplo, CA 2.092.588, AU 61781/94, AU 667939, US 6.100.447, PCT/US97/10621, US 5.589.617, US 6.541.257 y otros métodos se indican en el documento WO99/14314. Preferentemente, las plantas transgénicas de trigo o cebada se producen mediante  
15 procedimientos de transformación mediados por *Agrobacterium tumefaciens*. Los vectores que contienen el polinucleótido deseado se pueden introducir en células de trigo regenerables de plantas de cultivo tisular o de explantes o de sistemas vegetales adecuados tales como protoplastos.

20 Las células de trigo regenerables son preferentemente del escutelo de embriones inmaduros, embriones maduros, callos derivados de los mismos o del tejido meristemático.

Para confirmar la presencia de los transgenes en células y plantas transgénicas, se puede conducir una  
25 amplificación con la reacción en cadena de la polimerasa (PCR) o un análisis de transferencia Southern usando métodos conocidos por los expertos en la materia. Los productos de expresión de los transgenes se pueden detectar en cualquiera de varias maneras, dependiendo de la naturaleza del producto e incluyen transferencia Western y ensayos enzimáticos. Una manera de particular utilidad para cuantificar la expresión de proteínas y para detectar la replicación en diferentes tejidos vegetales comprende usar un gen informante, tal como GUS. Una vez que se obtuvieron las plantas transgénicas, se pueden cultivar para producir tejidos vegetales o partes de plantas con el fenotipo deseado. Se puede cosechar el tejido vegetal o las partes de plantas y/o se pueden recolectar las semillas.  
30 La semilla puede servir como una fuente para cultivar plantas adicionales con tejidos o partes que tienen las características deseadas.

Una planta transgénica formada usando métodos de transformación con *Agrobacterium* u otros métodos contiene normalmente un solo locus genético sobre un cromosoma. Dichas plantas transgénicas pueden considerarse como  
35 hemicigotas para el o los genes agregados. Más preferida es una planta transgénica que es homocigota para el o los genes agregados; es decir, una planta transgénica que contiene dos genes agregados, un gen en el mismo locus de cada cromosoma de un par de cromosomas. Se puede obtener una planta transgénica homocigota por autofertilización de una planta transgénica hemicigota, germinación de algunas de las semillas producidas y análisis de las plantas resultantes para el gen de interés.

40 Se comprenderá asimismo que también se pueden cruzar (aparear) dos diferentes plantas transgénicas que contienen dos genes o loci exógenos que segregan de forma independiente para producir una descendencia que contiene a ambos juegos de genes o loci. La autocruza de la progenie F1 apropiada puede producir plantas que son homocigotas para ambos loci o genes exógenos agregados. También se contempla la retrocruza con una planta progenitora y la exocruza con una planta no transgénica, al igual que la propagación vegetativa. Se pueden consultar descripciones de otros métodos de cruzamiento usados comúnmente para distintos caracteres y cultivos en Fehr, en *Breeding Methods for Cultivar Development*, Wilcos J. ed., *American Society of Agronomy*, Madison, Wis.

## 50 TILLING

En una realización, se puede usar TILLING (lesiones locales inducidas dirigidas en genomas) para producir plantas en las cuales se han noqueado genes endógenos, por ejemplo genes que codifican una DGAT, *sn-1* glicerol-3-fosfato aciltransferasa (GPAT), 1-acil-glicerol-3-fosfato aciltransferasa (LPAAT), acil-CoA:lisofofatidilcolina aciltransferasa (LPCAT), ácido fosfatídico fosfatasa (PAP) o una combinación de dos o más de los mismos.  
55

En una primera etapa, se inducen las mutaciones introducidas, tales como cambios de pares de bases individuales novedosos, en una población de plantas mediante tratamiento de semillas (o polen) con un mutágeno químico y luego avanzar las plantas hasta una generación donde las mutaciones serán heredadas de manera estable. Se extrae el ADN y las semillas se guardan separadas de todos los miembros de la población para crear una fuente al cual se puede acceder repetidamente en el tiempo.  
60

Para un ensayo de TILLING, se diseñan cebadores para PCR para amplificar específicamente un solo gen blanco de interés. La especificidad es especialmente importante si el blanco es un miembro de una familia de genes o forma parte de un genoma poliploide. A continuación, los cebadores marcados con colorante se pueden usar para amplificar los productos de la PCR del ADN agrupado de múltiples individuos. Estos productos de la PCR se  
65

desnaturalizan y alinean nuevamente para permitir la formación de pares de bases no coincidentes. Las faltas de coincidencia o heteroduplas, representan tanto polimorfismos de un solo nucleótido (SNP) naturales (es decir, es probable que varias plantas de la población sean portadores del mismo polimorfismo) y SNP inducidos (es decir, solamente unas pocas plantas individuales tienen probabilidad de expresar la mutación). Una vez formada la heterodupla, el uso de una endonucleasa, tal como *CelI*, que reconoce y corta el ADN no coincidente, es la clave del descubrimiento de nuevos SNP dentro de una población TILLING.

Cuando se usa este enfoque, se pueden examinar miles de plantas para identificar cualquiera que tuviera un único cambio de base, así como inserciones o supresiones pequeñas (1-30 pb) en cualquier gen o región específica del genoma. Los fragmentos genómicos evaluados pueden variar en tamaño en un rango entre 0,3 y 1,6 kb. Con 8 agrupaciones, fragmentos de 1,4 kb (descontando los extremos de los fragmentos donde la detección de SNP es problemática debido al ruido) y 96 calles por ensayo, esta combinación permitió examinar hasta un millón de pares de bases de ADN genómico por cada ensayo, lo que vuelve al TILLING una técnica de gran rendimiento.

El TILLING se describe además en Slade y Knauf (2005) y en Henikoff et al., (2004).

Además de permitir una detección eficiente de las mutaciones, la tecnología TILLING de gran rendimiento es ideal para la detección de polimorfismos naturales. Por ello, la búsqueda de un ADN homólogo desconocidos por formación de heteroduplas con una secuencia conocida revela la cantidad y posición de sitios polimórficos. Se identifican tanto cambios de nucleótidos como inserciones y supresiones pequeñas, incluyendo al menos algunos polimorfismos de número repetido. Esto se ha denominado Ecotilling (Comai et al., 2004).

Cada SNP se registra según su posición aproximada dentro de unos pocos nucleótidos. Por lo tanto, cada haplotipo se puede obtener basándose en su movilidad. Los datos de secuencia se pueden obtener con un esfuerzo de incremento relativamente pequeño usando alícuotas del mismo ADN amplificado que se usa para el ensayo de falta de coincidencia-clivaje. El cebador de secuenciación izquierdo o derecho para una sola reacción se elige por su proximidad con el polimorfismo. El software Sequencher efectúa una alineación múltiple y descubre el cambio de bases, que en cada caso confirmó la banda de gel.

El ecotilling es más económico que la secuenciación completa, el método utilizado actualmente en el descubrimiento de la mayoría de los SNP. Se pueden examinar placas que contienen al ADN ecotípico agrupado en lugar de agrupaciones de ADN de las plantas mutadas. Dado que la detección se efectúa sobre geles con una resolución cercana al par de bases y que los patrones basales son uniformes en todas las calles, se pueden hacer coincidir las bandas que son de tamaño idéntico, por consiguiente el descubrimiento y la determinación del genotipo de los SNP se efectúa en una sola etapa. De esta manera, la secuenciación final del SNP es simple y eficiente, aún más por el hecho que las alícuotas de los mismos productos de PCR usados para la selección pueden someterse a secuenciación de ADN.

#### Aumento de los niveles de ARN exógeno y expresión estabilizada

El silenciamiento de genes de post-transcripción (PTGS) es un mecanismo de defensa específico de la secuencia de nucleótidos que puede ser dirigido a ARNm celulares y virales para la degradación. El PTGS tiene lugar en plantas u hongos transformados de manera estable o transitoria con uno o más polinucleótidos recombinantes y da como resultado la acumulación reducida de moléculas de ARN con similitud de secuencia con el polinucleótido introducido.

Los niveles de moléculas de ARN se pueden incrementar y/o los niveles de moléculas de ARN se pueden estabilizar sobre numerosas generaciones, limitando la expresión de un supresor por silenciamiento en un órgano de almacenamiento de una planta o una parte del mismo. Tal como se usa en el presente documento, un "supresor por silenciamiento" es cualquier polinucleótido o polipéptido que se puede expresar en una célula vegetal que aumente el nivel de los productos de expresión de un transgen diferente en la célula vegetal, en particular sobre repetidas generaciones a partir de la planta inicialmente transformada. En una realización, el supresor por silenciamiento es un supresor por silenciamiento viral o un mutante del mismo. Se conocen numerosos supresores por silenciamiento virales en la técnica que incluyen, pero en un sentido no limitativo, P19, V2, P38, Pe-Po y RPV-P0. Los ejemplos de supresores de silenciamiento virales adecuados incluyen los que se describen en el documento WO 2010/057246. Se puede expresar un supresor de silenciamiento de manera estable en una planta o en una parte del mismo, de la presente invención.

Tal como se usa en el presente documento, el término "expresado de forma estable" o variaciones del mismo se refiere al nivel de moléculas de ARN siendo esencialmente igual o mayor en las plantas de progenie sobre generaciones repetidas, por ejemplo, al menos tres, al menos cinco o al menos diez generaciones, en comparación con las plantas correspondiente que no contienen al polinucleótido exógeno que codifica al supresor de silenciamiento. Sin embargo, estos términos no excluyen la posibilidad que sobre repetidas generaciones haya una cierta pérdida de los niveles de la molécula de ARN en comparación con una generación anterior, por ejemplo no menos que un 10 % de pérdida por generación.

El supresor se puede seleccionar de cualquier fuente como por ejemplo vegetal, viral, de mamífero, etc. El supresor

puede ser, por ejemplo:

- virus B2 de los gallineros;
- virus P14 latente de *Pothos*;
- virus AC2 latente de *Pothos*;
- 5 virus en mosaico AC4 de mandioca africana;
- enfermedad por el virus en mosaico de amarillamiento de venas C2 de Bhendi;
- enfermedad por el virus en mosaico de amarillamiento de venas C4 de Bhendi;
- enfermedad por el virus en mosaico de amarillamiento de venas  $\beta$ C1 de Bhendi;
- virus de clorosis p22 de tomate;
- 10 virus de clorosis CP de tomate;
- virus de clorosis CPm de tomate;
- virus en mosaico dorado AL2 de tomate;
- virus de hoja enrollada Java  $\beta$ C1 de tomate;
- virus de amarillamiento de hojas enrolladas V2 de tomate;
- 15 virus de amarillamiento de hojas enrolladas China-C2 de tomate;
- forma aislada  $\beta$ C1 del virus Y10-China de amarillamiento de hojas enrolladas de tomate;
- forma aislada V2 israelí de amarillamiento de hojas enrolladas de tomate;
- virus en mosaico amarillo Vigna AC2 de habichuela;
- virus clorótico de manchas en anillo CP de hibisco;
- 20 virus rugoso P38 de nabo;
- virus rugoso CP de nabo;
- virus en mosaico de la coliflor P6;
- virus de amarillamiento p21 de remolacha;
- virus tristeza p20 de cítricos;
- 25 virus tristeza p23 de cítricos;
- virus tristeza CP de cítricos;
- virus en mosaico SCP de caupí;
- virus enano clorótico p22 de batata;
- virus en mosaico 2b de pepino;
- 30 virus HC-Pro de aspermia de tomate;
- virus L2 de punta enrollada de remolacha;
- virus en mosaico 19K del suelo de trigo;
- virus en mosaico estriado Gammab de cebada;
- virus Gammab semilatente de *Poa*;
- 35 pecluvirus P15 de aglomeraciones de maní;
- virus Pns10 enano de arroz;
- virus P0 de amarillamiento transmitido por áfidos de curcubitáceas;
- virus P0 de amarillamiento del oeste de la remolacha;
- virus X P25 de papa;
- 40 virus P1b del amarillamiento de venas de pepino;
- virus HC-Pro pox de ciruela;
- virus en mosaico HC-Pro de caña de azúcar
- virus Y cepa HC-Pro de papa;
- virus del grabado de la hoja P1/HC-Pro de tabaco;
- 45 virus en mosaico P1/HC-Pro de nabo;
- virus P1 moteado de pasto ovillo;
- forma aislada del virus noruego P1 moteado de pasto ovillo;
- virus P1 moteado amarillo de arroz;
- forma aislada del virus nigeriano P1 moteado amarillo de arroz;
- 50 virus NS3 de hoja blanca de arroz;
- virus NS3 estriado de arroz;
- virus en mosaico del tabaco 126K infectante de crucíferas;
- virus en mosaico del tabaco p122 infectante de crucíferas;
- virus en mosaico del tabaco p122;
- 55 virus en mosaico del tabaco 126
- virus en mosaico del tabaco 130K;
- virus del cascabel 16K de tabaco;
- virus enano arbustivo P19 de tomate;
- virus bandeado reticular NSs de tomate;
- 60 virus de manchas cloróticas de hojas P50 de manzana;
- virus A p10 de uva;
- virus 2 homólogo de BYV p21 asociado a enrollamiento de hojas de uva,

así como variantes/mutantes de los mismos. El listado anterior proporciona el virus del cual se puede obtener el supresor y la denominación de la proteína (por ejemplo, B2, P14 etc.) o región codificante del supresor de cada virus particular.

Se pueden usar múltiples copias de un supresor. Se pueden usar diferentes supresores juntos (por ejemplo, en tándem).

5 Se puede co-expresar esencialmente cualquier molécula de ARN que se deseara expresar en un órgano de almacenamiento de una planta con el supresor por silenciamiento. La molécula de ARN puede afectar una característica agronómica, resistencia a insectos, resistencia a enfermedades, resistencia a herbicidas, esterilidad, características de los granos y semejantes. Los polipéptidos codificados pueden estar relacionados con el metabolismo de lípidos, almidón, carbohidratos, nutrientes, etc. o pueden ser responsables de la síntesis de  
10 proteínas, péptidos, lípidos, ceras, almidones, azúcares, carbohidratos, sabores, olores, toxinas, carotenoides, hormonas, polímeros, flavonoides, proteínas de almacenamiento, ácidos fenólicos, alcaloides, ligninas, taninos, celulosas, glicoproteínas, glucolípidos, etc.

15 En un ejemplo particular, las plantas produjeron niveles aumentados de enzimas para la producción de lípidos en plantas tales como *Brassicas*, por ejemplo colza oleaginosa o girasol, cártamo, lino, algodón, soja o maíz.

#### Producción de lípidos no polares

20 Se pueden emplear las técnicas que son de práctica rutinaria en la técnica para extraer, procesar, purificar y analizar los lípidos no polares producidos por células, organismos o partes de los mismos, de la presente invención. Dichas técnicas se describen y explican en la bibliografía, en fuentes tales como Fereidoon Shahidi, *Current Protocols in Food Analytical Chemistry*, John Wiley & Sons, Inc. (2001) D1.1.1-D1.1.11 y Perez-Vich et al. (1998).

#### *Producción de aceite de semillas*

25 Normalmente, las semillas de las plantas se cocinan, se prensan y/o se extraen para producir un aceite de semillas crudo, que luego es desgomado, refinado, blanqueado y desodorizado. En general, las técnicas para triturar semillas son conocidas en la técnica. Por ejemplo, semillas oleaginosas se pueden temperar rociándolas con agua para elevar el contenido de humedad, por ejemplo, a un 8,5 % y cortar en copos usando un rodillo compactador liso con  
30 un ajuste de hueco definido en 0,23 a 0,27 mm. Según el tipo de semilla, se puede agregar agua o no, antes de la trituración. La aplicación de calor desactiva las enzimas, facilita la ruptura celular posterior, coalesce las gotas de lípidos y aglomera las partículas de proteínas, todo lo cual facilita el proceso de extracción.

35 La mayoría de los aceites de semillas se libera por medio de pasaje a través de una prensa de tornillo. Las tortas expulsadas de la prensa de tornillo son sometidas a extracción con solvente, por ejemplo, con hexano, usando una columna de temperatura controlada. Como alternativa, el aceite de semillas crudo producido por la operación de prensado se puede hacer pasar a través de un tanque de asentamiento con una vía de drenaje acanalada superior para remover los sólidos que se generan con el aceite de semillas durante la operación de prensado. El aceite de semillas clarificado se puede hacer pasar a través de un filtro tipo placa y marco para remover cualquier partícula  
40 sólida fina. Si se deseara, el aceite de semillas recuperado del proceso de extracción se puede combinar con el aceite de semillas clarificado para producir un aceite de semillas crudo mezclado.

Una vez que se deprivó el solvente del aceite de semillas crudo, se combinan las porciones prensadas y extraídas y luego son sometidas a los procedimientos normales de procesamiento de lípidos (es decir, desgomado, refinación cáustica, blanqueado y desodorización). El desgomado se puede realizar por adición de ácido fosfórico concentrado  
45 al aceite de semillas crudo para convertir los fosfátidos no hidratables en una forma hidratable y para quelar los metales menores que están presentes en el aceite. La goma se separa del aceite de semillas por centrifugación. El aceite de semillas se puede refinar por adición de una cantidad suficiente de una solución de hidróxido de sodio para titular todos los ácidos grasos y remover los jabones así formados.

50 La desodorización se puede realizar calentando el aceite de semillas a 260 °C bajo vacío e introduciendo lentamente vapor en el aceite de semillas a una velocidad de 0,1 ml/minuto/100 ml de aceite de semillas aproximadamente. Después de aproximadamente 30 minutos de burbujeo, se deja que el aceite de semillas se enfríe bajo vacío. El aceite de semillas es transferido normalmente a un recipiente de vidrio y se lava con argón antes de almacenarlo  
55 bajo refrigeración. Si la cantidad de aceite de semillas es limitada, el aceite de semillas se puede colocar bajo vacío, por ejemplo, en un reactor Parr y calentar a 260 °C por el mismo tiempo que hubiera sido desodorizado. Este tratamiento mejora el color del aceite de semillas y remueve la mayoría de las sustancias volátiles.

#### *Biomasa vegetal para la producción de lípidos*

60 Las partes de plantas relacionadas con la fotosíntesis (por ejemplo, tallos y hojas de plantas superiores y plantas acuáticas tales como algas) también se pueden usar para producir lípidos. Independientemente del tipo de planta, existen diversos métodos para extraer lípidos de la biomasa verde. Otra manera es la extracción física, que a menudo no emplea extracción con solventes. Es una manera "tradicional" que emplea diversos tipos diferentes de  
65 extracción mecánica. La extracción por prensa *Expeller* es un tipo común, como lo es la prensa de tornillo y los métodos de extracción con una prensa *Ram*. La cantidad de lípido extraído usando estos métodos varía

ampliamente, dependiendo del material vegetal y del proceso mecánico empleado. normalmente la extracción mecánica es menos eficiente que la extracción con solvente que se describe a continuación.

5 En la extracción con solvente, se mezcla un solvente orgánico (por ejemplo, hexano) con al menos la biomasa vegetal verde modificada genéticamente, preferentemente después de secar y moler dicha biomasa verde. Por supuesto, también se pueden moler y mezclar otras partes de la planta además de la biomasa verde (por ejemplo, semillas que contienen lípidos). El solvente disuelve el lípido en la biomasa y semejantes, luego se separa dicha solución de la biomasa por acción mecánica (por ejemplo, con los procesos de prensado anteriores). Esta etapa de separación también se puede efectuar por filtración (por ejemplo, con un filtro prensa o un dispositivo similar) o por  
10 centrifugación etc. El solvente orgánico se puede separar luego del lípido no polar (por ejemplo, por destilación). Esta segunda etapa de separación permite obtener lípidos no polares de la planta y puede dar como resultado un solvente reusable si se emplea una recuperación por vapor convencional.

#### 15 *Producción de combustible de algas*

El cultivo de algas es una forma de acuicultura que comprende la cría de especies de algas (incluyendo microalgas, también denominado fitopláncton, microfitas o algas planctónicas y macroalgas, comúnmente conocidas como sargazo). Los ejemplos de especies de algas de utilidad en el presente documento invención incluyen, por ejemplo,  
20 *Chlamydomonas sp* (por ejemplo, *Chlamydomonas reinhardtii*), *Dunaliella sp.*, *Haematococcus sp.*, *Chlorella sp.*, *Thraustochytrium sp.*, *Schizochytrium sp.* y *Volvox sp.*

El cultivo comercial e industrial de algas tiene numerosos usos, incluyendo producción de ingredientes alimenticios, alimento y combustible de algas.

25 Los cultivos de un alga o de algas mixtos se pueden cultivar en estanques abiertos (tales como estanques tipo *raceway* y lagos) o fotobiorreactores.

30 Las algas se pueden cosechar usando *Microscreens*, por centrifugación, por floculación (usando por ejemplo, quitosán, alúmina y cloruro férrico) y por flotación con espuma. La interrupción del suministro de dióxido de carbono puede causar la floculación de las algas por sí solos, que se denomina "autofloculación". En la flotación con espuma, el cultivador airea el agua formando una espuma y luego se limpian las algas desde la superficie. El ultrasonido y otros métodos de cosecha se encuentran actualmente bajo desarrollo.

35 Los lípidos se pueden separar de las algas por trituración mecánica. Cuando las algas se secan retienen su contenido de lípidos, que luego se puede extraer por "prensado" con una prensa para aceite. Dado que las diferentes cepas de algas varían ampliamente en cuanto a sus atributos físicos, hay distintas configuraciones de prensas (tornillo, *Expeller*, pistón, etc.) que funcionan mejor para tipos de algas específicos.

40 A menudo se emplea shock osmótico para liberar componentes celulares tales como lípidos de algas. El shock osmótico es una reducción repentina de la presión osmótica y puede causar la ruptura de las células en solución.

45 La extracción ultrasónica puede acelerar los procesos de extracción, en particular los procesos de extracción enzimática empleados para extraer lípidos de algas. Se usan ondas ultrasónicas para crear burbujas de cavitación en un material solvente. Cuando estas burbujas colapsan cerca de las paredes celulares, las ondas de shock resultantes y los chorros líquidos causan la ruptura de las paredes celulares y la liberación de su contenido en un solvente.

50 A menudo se usan solventes químicos (por ejemplo, hexano, benceno, éter de petróleo) en la extracción de los lípidos de las algas. Se puede usar extracción Soxhlet para extraer lípidos de algas por lavado repetido o percolación, con un solvente orgánico bajo reflujo en un material de vidrio especial.

55 La extracción enzimática se puede usar para extraer lípidos de algas. La extracción enzimática emplea enzimas para degradar las paredes celulares con agua que actúan como solvente. La extracción enzimática puede ser sustentada por ultrasonificación.

También se puede usar CO<sub>2</sub> supercrítico como solvente. En este método, se licúa CO<sub>2</sub> bajo presión y se calienta hasta el punto en que se vuelve supercrítico (con propiedades tanto de líquido como de gas), permitiendo que actúe como solvente.

#### 60 *Procesos de fermentación para la producción de lípidos*

65 Tal como se usa en el presente documento, el término "proceso de fermentación" se refiere a cualquier proceso de fermentación o a cualquier proceso que comprende una etapa de fermentación. Un proceso de fermentación incluye, en un sentido no limitativo, los procesos de fermentación usados para producir alcoholes (por ejemplo, etanol, metanol, butanol); ácidos orgánico (por ejemplo, ácido cítrico, ácido acético, ácido itacónico, ácido láctico, ácido

glucónico); cetonas (por ejemplo, acetona); aminoácidos (por ejemplo, ácido glutámico); gases (por ejemplo, H<sub>2</sub> y CO<sub>2</sub>); antibióticos (por ejemplo, penicilina y tetraciclina); enzimas; vitaminas (por ejemplo, riboflavina, beta-caroteno); y hormonas. Los procesos de fermentación también incluyen los procesos de fermentación usados en la industria de alcohol de consumo (por ejemplo, cerveza y vino), la industria láctea (por ejemplo, productos lácteos fermentados), la industria del cuero y la industria del tabaco. Los procesos de fermentación preferidos incluyen los procesos de fermentación de alcoholes, bien conocidos en la técnica. Los procesos de fermentación preferidos son procesos de fermentación anaeróbicos, bien conocidos en la técnica. Las células fermentadoras adecuadas, normalmente microorganismos, tienen la capacidad de fermentar, es decir, convertir, azúcares, tales como glucosa o maltosa, directa o indirectamente en el producto de fermentación deseado. Los ejemplos de microorganismos fermentadores incluyen organismos fúngicos, tal como levadura, preferentemente un organismo oleaginoso. Tal como se usa en el presente documento, un "organismo oleaginoso" es aquel que acumula al menos un 25 % de su peso seco como triglicéridos. Tal como se usa en el presente documento, una "levadura" incluye *Saccharomyces* spp., *Saccharomyces cerevisiae*, *Saccharomyces carlbergensis*, *Candida* spp., *Kluveromyces* spp., *Pichia* spp., *Hansenula* spp., *Trichoderma* spp., *Lipomyces starkey* y *Yarrowia lipolitica*. Las levaduras preferidas incluyen *Yarrowia lipolitica* u otras levaduras oleaginosas y cepas de *Saccharomyces* spp. Y, en particular, *Saccharomyces cerevisiae*.

El microorganismo transgénico preferentemente se cultiva bajo condiciones que optimizan la actividad de los genes de la biosíntesis de ácidos grasos y los genes de aciltransferasas. Esto conduce a la producción del mayor rendimiento y más económico de lípidos. En general, las condiciones de los medios que se pueden optimizar incluyen el tipo y la cantidad de fuentes de carbono, el tipo y la cantidad de fuentes de nitrógeno, la relación de carbono a nitrógeno, el nivel de oxígeno, la temperatura de crecimiento, pH, extensión de la fase de producción de biomasa, extensión de la fase de acumulación de lípidos y el momento de cosecha de las células.

El medio de fermentación debe contener una fuente de carbono adecuada. Las fuentes de carbono adecuadas pueden incluir, pero en un sentido no limitativo: monosacáridos (por ejemplo, glucosa, fructosa), disacáridos (por ejemplo, lactosa, sacarosa), oligosacáridos, polisacáridos (por ejemplo, almidón, celulosa o mezclas de los mismos), alcoholes de azúcares (por ejemplo, glicerol) o mezclas de alimentaciones renovables (por ejemplo, permeato de suero de queso, licor de maíz macerado, molasas de remolacha, malta de cebada). Además, las fuentes de carbono pueden incluir alcanos, ácidos grasos, ésteres de ácidos grasos, monoglicéridos, diglicéridos, triglicéridos, fosfolípidos y diversas fuentes de ácidos grasos comerciales incluyendo aceites vegetales (por ejemplo, aceite de soja) y grasas animales. Además, el sustrato de carbono puede incluir sustratos de un carbono (por ejemplo, dióxido de carbono, metanol, formaldehído, formiato, aminas que contienen carbono) para los cuales se haya demostrado conversión metabólica en los intermediarios bioquímicos clave. Por ende, se contempla que la fuente de carbono utilizada en el presente documento invención puede abarcar una amplia variedad de sustratos que contienen carbono y solamente queda limitada por la elección del microorganismo hospedador. Aunque se espera que todos los sustratos de carbono mencionados previamente y las mezclas de los mismos sean adecuados en el presente documento invención, los sustratos de carbono preferidas son azúcares y/o ácidos grasos. Se prefiere especialmente la glucosa y/o los ácidos grasos que contienen entre 10-22 carbonos.

El nitrógeno se puede suministrar de una fuente inorgánica (por ejemplo, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>) u orgánico (por ejemplo, urea, glutamato). Además de fuentes de carbono y nitrógeno apropiados, el medio de fermentación también puede contener minerales, sales, cofactores, soluciones amortiguadoras, vitaminas y otros componentes adecuados conocidos por los expertos en la materia que son adecuados para el crecimiento del microorganismo y la promoción de las vías enzimáticas necesarias para la producción de lípidos.

Un rango de pH adecuado para la fermentación comprende normalmente entre aproximadamente pH 4,0 y pH 8,0, en donde se prefiere entre pH 5,5 y pH 7,0 como rango para las condiciones de crecimiento iniciales. La fermentación se puede conducir bajo condiciones aeróbicas o anaeróbicas, siendo preferidas las condiciones microaeróbicas.

Normalmente, la acumulación de niveles elevados de lípidos en las células de microorganismos oleaginosos requiere de un proceso de dos etapas, dado que el estado metabólico debe "equilibrarse" entre el crecimiento y la síntesis/almacenamiento de grasas. Por lo tanto, con mayor preferencia, es necesario un proceso de fermentación de dos etapas para la producción de lípidos en los microorganismos. En este enfoque, la primera etapa de la fermentación está dedicada a la generación y acumulación de masa celular y se caracteriza por un crecimiento celular y división celular rápidos. En la segunda etapa de la fermentación, se prefiere establecer condiciones de privación de nitrógeno en el cultivo para promover niveles elevados de acumulación de lípidos. El efecto de esta privación de nitrógeno comprende reducir la concentración efectiva de AMP en las células, reduciendo así la actividad de la isocitrato deshidrogenasa dependiente de NAD de las mitocondrias. Cuando esto sucede, se acumulará ácido cítrico y por lo tanto se forman agrupaciones abundantes de acetil-CoA en el citoplasma y se ceba la síntesis de ácidos grasos. Por lo tanto, esta fase se caracteriza por el cese de la división celular seguido por la síntesis de ácidos grasos y acumulación de TAGs.

Aunque las células normalmente se cultivan a 30 °C aproximadamente, algunos estudios han mostrado una mayor síntesis de ácidos grasos insaturados a temperaturas inferiores. Sobre la base de la economía del proceso, este



cambio de temperatura probablemente debería tener lugar después de la primera fase de la fermentación de dos etapas, cuando ya se ha producido el crecimiento masivo del microorganismo.

Se contempla que se puede aplicar varios diseños de procesos de fermentación, cuando se desea una producción comercial de lípidos usando los ácidos nucleicos de la presente. Por ejemplo, la producción comercial de lípidos a partir de un hospedador recombinante microbiano se puede efectuar mediante un proceso de fermentación por lotes, por lote alimentado o continuo.

Un proceso de fermentación por lotes es un sistema cerrado en donde la composición del medio se define al comienzo del proceso y no se efectúan otras adiciones más allá de los requeridos para mantener el pH y el nivel de oxígeno durante el proceso. Por lo tanto, al comienzo del proceso de cultivo el medio es inoculado con el organismo deseado y se permite que tenga lugar el crecimiento o la actividad metabólica sin adición de otros sustratos (es decir, fuentes de carbono y nitrógeno) al medio. En los procesos por lotes, las composiciones de metabolitos y biomasa del sistema cambia constantemente hasta el momento de terminar el cultivo. En un proceso por lotes típico, las células pasan por una fase de retardo estático a una fase logarítmica de gran crecimiento y finalmente a una fase estacionaria, en donde el índice de crecimiento está disminuido o detenido. Si se dejan sin tratar, las células en la fase estacionaria finalmente morirán. Una variación del proceso por lotes estándar es el proceso por lote alimentado, en donde el sustrato se agrega de manera continua al fermentador durante el transcurso del proceso de fermentación. Un proceso por lote alimentado también es adecuado en el presente documento invención. Los procesos por lotes alimentados son de utilidad cuando la represión de catabolitos es adecuada para inhibir el metabolismo de las células o cuando fuera deseable contar con cantidades limitadas de sustrato en el medio en cualquier momento. La medición de la concentración de sustrato en los sistemas por lotes alimentados es difícil y por ello se puede estimar basándose en los cambios de factores mensurables tales como pH, oxígeno disuelto y la presión parcial de los desechos gaseosos (por ejemplo, CO<sub>2</sub>). Los métodos de cultivo por lotes y por lotes alimentados son comunes y bien conocidos en la técnica y se pueden consultar ejemplos de ellos en Brock, en *Biotechnology: A Textbook of Industrial Microbiology*, 2a ed., Sinauer Associates, Sunderland, Mass, (1989); o Deshpande y Mukund (1992).

La producción comercial de lípidos con las células de la presente también se puede efectuar mediante un proceso de fermentación continuo, en donde se agrega un medio definido de manera continua a un biorreactor al mismo tiempo que se retira una cantidad igual de volumen de cultivo para la recuperación de producto. Los cultivos continuos generalmente conservan las células en la fase de crecimiento logarítmica a una densidad celular constante. Los métodos de cultivo continuos o semi-continuos permiten modular un factor o cualquier número de factores que afectan el crecimiento celular o la concentración del producto final. Por ejemplo, un enfoque puede limitar la fuente de carbono y permitir los demás parámetros para moderar el metabolismo. En otros sistemas, se pueden alterar numerosos factores que afectan el crecimiento de manera continua en tanto la concentración celular, medida por la turbidez del medio, se mantiene constante. Los sistemas continuos buscan mantener un crecimiento de estado estable y por lo tanto se debe equilibrar el índice de crecimiento celular contra la pérdida de células debido al medio retirado del cultivo. Los métodos para modular los nutrientes y factores de crecimiento para los procesos de cultivo continuos, así como técnicas para maximizar el índice de formación de producto, son bien conocidos en la técnica de la microbiología industrial y varios métodos se describen con detalle en Brock, *supra*.

Los ácidos grasos, incluyendo los PUFA, presentes en el microorganismo hospedador se pueden encontrar como ácidos grasos libres o en formas esterificadas, tales como acilgliceroles, fosfolípidos, sulfolípidos o glucolípidos y se pueden extraer de la célula hospedadora mediante varios medios bien conocidos en la técnica.

En general, la purificación de ácidos grasos, incluyendo PUFA, puede incluir extracción con solventes orgánicos, sonicación, extracción por fluidos supercríticos (por ejemplo, usando dióxido de carbono), saponificación y medios físicos tales como prensas o combinaciones de los mismos. De particular interés es la extracción con metanol y cloroformo en la presencia de agua (Bligh y Dyer, 1959). Cuando fuera deseable, se puede acidificar la capa acuosa para protonar los grupos de carga negativa y de esa manera incrementar el particionamiento de los productos deseados en la capa orgánica. Después de la extracción, los solventes orgánicos se pueden eliminar por evaporación bajo flujo de nitrógeno. Cuando se aíslan en formas conjugadas, los productos se pueden cortar enzimáticamente o químicamente para liberar los ácidos grasos libres o un conjugado menos complejo de interés y luego pueden ser sometidos a manipulaciones adicionales para producir un producto final deseado. Preferentemente, las formas conjugadas de los ácidos grasos se cortan con hidróxido de potasio.

Si fuera necesaria una purificación adicional, se pueden emplear métodos estándar. Dichos métodos pueden incluir extracción, tratamiento con urea, cristalización fraccionada, HPLC, destilación fraccionada, cromatografía sobre gel de sílice, centrifugación o destilación a gran velocidad o combinaciones de estas técnicas. La protección de grupos reactivos, tales como grupos ácidos o alqueno, se puede efectuar en cualquier etapa mediante técnicas conocidas (por ejemplo, alquilación, iodación). Los métodos usados incluyen metilación de los ácidos grasos para producir metil ésteres. De manera similar, los grupos protectores se pueden eliminar en cualquier etapa. Preferentemente, la purificación de las fracciones que contienen GLA, STA, ARA, DHA y EPA se puede efectuar mediante tratamiento con urea y/o destilación fraccionada.

Usos de los lípidos

Los lípidos producidos mediante los métodos descritos tienen diversos usos. En algunas formas de realización, los lípidos se usan como aceites comestibles. En otras formas de realización, los lípidos se refinan y usan como lubricantes o para otros usos industriales tal como en la síntesis de plásticos. En algunas formas de realización preferidas, los lípidos se refinan para producir biodiesel.

*Biodiesel*

La producción de biodiesel o alquil ésteres, es bien conocida. Hay tres rutas básicas para producir ésteres a partir de lípidos: 1) Transesterificación catalizada por bases del lípido con alcohol; 2) Esterificación directa catalizada por ácidos del lípido con metanol; y 3) Conversión del lípido en ácidos grasos y luego en alquil ésteres con catalizadores ácidos.

En algunas formas de realización preferidas, los lípidos se transesterifican para producir metil ésteres y glicerol. En algunas formas de realización preferidas, los lípidos se hacen reaccionar con un alcohol (tal como metanol o etanol) en la presencia de un catalizador (hidróxido de potasio o sodio) para producir alquil ésteres. Los alquil ésteres se pueden para biodiesel o se pueden mezclar con combustibles basados en petróleo.

20 *Alimentos*

La presente invención incluye composiciones que se pueden usar como alimentos. A los efectos de la presente invención, el término "alimentos" incluye cualquier alimento o preparación destinada al consumo de humanos o animales (incluyendo un consumo por vía enteral y/o parenteral) que cuando es ingerido por el cuerpo (1) sirve para nutrir o aumentar los tejidos o suministrar energía; y/o (2) mantiene, restablece o sustenta un estado nutricional o función metabólica adecuados. Los alimentos de la invención incluyen composiciones nutricionales para bebés y/o niños.

Los alimentos de la invención comprenden, por ejemplo, una célula de la invención, una planta de la invención, la parte de planta de la invención, la semilla de la invención, un extracto de la invención, el producto de un método de la invención, el producto de un proceso de fermentación de la invención o una composición junto con uno o más vehículos adecuados. El término "vehículo" se usa en su sentido más amplio para abarcar cualquier componente que puede tener o no, valor nutricional. Como podrá apreciar el especialista, el vehículo debe ser adecuado para su uso (o usado a una concentración suficientemente baja) en un alimento de modo tal que no tenga efectos nocivos sobre un organismo que consume el alimento.

El alimento de la presente invención comprende un lípido producido directa o indirectamente mediante el uso de los métodos, células u organismos divulgados en el presente documento. La composición se puede encontrar en la forma de un gel o de un líquido. Además, la composición puede incluir macronutrientes comestibles, vitaminas y/o minerales en las cantidades deseadas para un uso particular. Las cantidades de estos ingredientes varían dependiendo de si la composición se usará con individuos normales o si se usará con individuos que tienen necesidades especiales, tales como los individuos que sufren de trastornos metabólicos y semejantes.

Los ejemplos de vehículos adecuados con valor nutricional incluyen, pero en un sentido no limitativo, macronutrientes tales como grasas comestibles, carbohidratos y proteínas. Los ejemplos de tales grasas comestibles incluyen, pero en un sentido no limitativo, aceite de coco, aceite de borraja, aceite de hongos, aceite de grosellas, aceite de soja y mono y diglicéridos. Los ejemplos de dichos hidratos de carbono incluyen, a modo de ejemplo ilustrativo, glucosa, lactosa comestible y almidón hidrolizado. Además, los ejemplos de proteínas que se pueden utilizar en la composición nutricional de la invención incluyen (pero en un sentido no limitativo) proteínas de soja, suero electrodiálizado, leche descremada electrodiálizada, suero de leche o los hidrolizados de estas proteínas.

Con respecto a las vitaminas y los minerales, se puede agregar lo siguiente a las composiciones de alimento de la presente invención: calcio, fósforo, potasio, sodio, cloruro, magnesio, manganeso, hierro, cobre, zinc, selenio, yodo y las vitaminas A, E, D, C y el complejo B. También se pueden agregar otras vitaminas y minerales.

Los componentes utilizados en las composiciones alimenticias de la presente invención pueden ser de origen semipurificado o purificado. Los términos semipurificado o purificado se refieren a un material que fue preparado por purificación de un material natural o mediante síntesis *de novo*.

La composición alimenticia de la presente invención también se puede agregar a un alimento aún cuando la dieta no requiera de suplementos. Por ejemplo, la composición se puede agregar a un alimento de cualquier tipo, incluyendo (pero en un sentido no limitativo): margarina, manteca modificada, quesos, leche, yogur, chocolate, caramelos, refrigerios, aceites para ensalada, aceites para cocinar, grasas para cocinar, carnes, pescado y bebidas.

El género *Saccharomyces* spp se usa tanto en la elaboración de cerveza como en la elaboración de vino y también

como un agente de horneado, en particular de pan. La levadura es uno de los constituyentes principales de los extractos vegetales. La levadura también se usa como un aditivo en los piensos. Resultará evidente que se pueden suministrar cepas de levadura modificadas que están adaptadas para sintetizar los lípidos que se describen en el presente documento. Estas cepas de levadura se pueden usar luego en alimentos y en la elaboración de vino y cerveza para proporcionar productos con un mayor contenido de lípidos.

Además, los lípidos producidos de acuerdo con la presente invención o células hospedadoras transformadas para contener y expresar los genes objeto de la presente también se pueden usar como suplementos de piensos para alterar el tejido de un animal o la composición de ácidos grasos de la leche para uno que fuera más deseable para consumo humano o animal. Los ejemplos de dichos animales incluyen ovejas, ganado, caballos y semejantes.

Además, los alimentos de la invención se pueden usar en acuicultura para aumentar los niveles de ácidos grasos en los peces para consumo humano o de animales.

Los alimentos preferidos de la invención son las plantas, semillas y otras partes de plantas tales como hojas, frutas y tallos que se pueden usar directamente como alimento o pienso para humanos u otros animales. Por ejemplo, los animales pueden pastar directamente sobre tales plantas cultivadas a campo o pueden ser alimentados con cantidades más medidas en una alimentación controlada. La invención incluye el uso de dichas plantas y partes de plantas como un alimento para aumentar los niveles de ácidos grasos poliinsaturados en humanos y otros animales.

### Composiciones

En el presente documento se divulgan composiciones, en particular composiciones farmacéuticas, que comprenden uno o más lípidos producidos usando los métodos de la invención.

Una composición farmacéutica puede comprender uno o más de los lípidos, en combinación con un transportador, adyuvante o vehículo estándar, bien conocido, no tóxico, farmacéuticamente aceptable, tal como solución amortiguadora fosfato salina, agua, etanol, polioles, aceites vegetales, un agente humectante o una emulsión tal como una emulsión de agua/aceite. La composición está tanto en forma líquida como sólida. Por ejemplo, la composición puede estar en la forma de una tableta, cápsula, líquido o polvo ingerible o ungüento tópico o crema. La fluidez adecuada se puede mantener, por ejemplo, por mantenimiento del tamaño de partícula requerido en el caso de dispersiones y por el uso de tensioactivos. También puede ser deseable incluir agentes isotónicos, por ejemplo, azúcares, cloruro de sodio y similares. Además de dichos diluyentes inertes, la composición también puede incluir adyuvantes, tales como agentes humectantes, agentes emulsificantes y de suspensión, agentes edulcorantes, agentes saborizantes y agentes aromatizantes.

Las suspensiones pueden comprender, además del ingrediente activo, agentes de suspensión, tales como alcoholes isoestearílicos etoxilados, polioxietilensorbitol y ésteres de sorbitán, celulosa microcristalina, metahidróxido de aluminio, bentonita, agar-agar y tragacanto y mezclas de estas sustancias.

Las formas de dosificación sólidas tales como tabletas y cápsulas se pueden preparar usando técnicas bien conocidas en la técnica. Por ejemplo, los lípidos producidos de acuerdo a la presente invención se pueden armar en tabletas con bases de tabletas convencionales tales como lactosa, sacarosa y almidón de maíz en combinación con aglutinantes tales como acacia, almidón de maíz o gelatina, agentes desintegrantes tales como almidón de patata o ácido algínico y un lubricante tal como ácido esteárico o estearato de magnesio. Las cápsulas se pueden preparar por incorporación de estos excipientes a una cápsula de gelatina junto con antioxidantes y el o los lípidos de interés.

Para una administración intravenosa, los lípidos producidos de acuerdo con la presente invención o derivados de los mismos, se pueden incorporar en formulaciones comerciales.

Una dosificación típica de un ácido graso particular comprende entre 0,1 mg y 20 g, ingeridas una a cinco veces por día (hasta 100 g por día) y preferentemente se encuentra en el rango de entre aproximadamente 10 mg y aproximadamente 1, 2, 5 o 10 g diarios (tomado en una o múltiples dosis). Según es de conocimiento en la técnica, resulta deseable un mínimo de aproximadamente 300 mg/día de ácidos grasos, en especial de ácidos grasos poliinsaturados. Sin embargo, se podrá apreciar que cualquier cantidad de ácidos grasos será beneficiosa para el sujeto.

Las posibles rutas de administración de las composiciones farmacéuticas de la presente invención incluyen, por ejemplo, la vía enteral y parenteral. Por ejemplo, una preparación líquida se puede administrar por vía oral. Además, se puede dispersar por completo una mezcla homogénea en agua, mezclar bajo condiciones estériles con diluyentes, conservantes, soluciones amortiguadoras o propelentes fisiológicamente aceptables para formar un aerosol o inhalante.

La dosificación de la composición que será administrada al sujeto puede ser determinada por el experto en la materia y depende de diversos factores tales como peso, edad del paciente, estado de salud general, antecedentes, estado inmune, etc del sujeto.

Además, las composiciones de la presente invención se pueden utilizar con fines de cosméticos. Las composiciones se pueden agregar a composiciones cosméticas pre-existentes de modo tal que se forma una mezcla o se puede usar un ácido graso producido de acuerdo con la presente invención como único ingrediente "activo" en una composición cosmética.

5

### Polipéptidos

Los términos "polipéptido" y "proteína" generalmente se usan indistintamente.

10 Un polipéptido o clase de polipéptidos se puede definir por el grado de identidad (% de identidad) de su secuencia de aminoácidos con una secuencia de aminoácidos de referencia o por tener un mayor % de identidad con una secuencia de aminoácidos de referencia que con otra. El % de identidad de un polipéptido con una secuencia de aminoácidos de referencia se determina normalmente mediante el análisis GAP (Needleman y Wunsch, 1970; programa GCG) con los parámetros de una penalidad por creación de hueco = 5 y una penalidad por extensión de hueco = 0,3. La secuencia de consulta es de al menos 100 aminoácidos de longitud y el análisis GAP alinea las dos secuencias sobre una región de al menos 100 aminoácidos. Aún más preferentemente, la secuencia de consulta es de al menos 250 aminoácidos de longitud y el análisis GAP alinea las dos secuencias sobre una región de al menos 250 aminoácidos. Más preferentemente aún, el análisis GAP alinea dos secuencias sobre su longitud completa. El polipéptido o la clase de polipéptidos pueden tener la misma actividad enzimática o una actividad diferente o una falta de actividad, con respecto al polipéptido de referencia. Preferentemente, el polipéptido tiene una actividad enzimática de al menos 10 % de la actividad del polipéptido de referencia.

25 Tal como se usa en el presente documento, un fragmento "fragmento biológicamente activo" es una porción de un polipéptido de la invención que mantiene una actividad definida de un polipéptido de referencia de longitud completa, por ejemplo, actividad MGAT. Los fragmentos biológicamente activos, Tal como se usan en el presente documento, excluyen al polipéptido de longitud completa. Los fragmentos biológicamente activos pueden ser una porción de cualquier tamaño, siempre que conserven la actividad definida. Preferentemente, el fragmento biológicamente activo conserva al menos un 10 % de la actividad del polipéptido de longitud completa.

30 Con respecto a un polipéptido o a una enzima definida, se podrá apreciar que las cifras de % de identidad mayores que las provistas en el presente documento abarcarán formas de realización preferidas. Por consiguiente, cuando fuera aplicable, a la luz de las cifras mínimas de % de identidad, se prefiere que el polipéptido/enzima comprenda un aminoácido que es al menos un 60 %, más preferentemente al menos un 65 %, más preferentemente al menos un 70 %, más preferentemente al menos un 75 %, más preferentemente al menos un 80 %, más preferentemente al menos un 85 %, más preferentemente al menos un 90 %, más preferentemente al menos un 91 %, más preferentemente al menos un 92 %, más preferentemente al menos un 93 %, más preferentemente al menos un 94 %, más preferentemente al menos un 95 %, más preferentemente al menos un 96 %, más preferentemente al menos un 97 %, más preferentemente al menos un 98 %, más preferentemente al menos un 99 %, más preferentemente al menos un 99,1 %, más preferentemente al menos un 99,2 %, más preferentemente al menos un 99,3 %, más preferentemente al menos un 99,4 %, más preferentemente al menos un 99,5 %, más preferentemente al menos un 99,6 %, más preferentemente al menos un 99,7 %, más preferentemente al menos un 99,8 % y aún más preferentemente al menos un 99,9 % idéntica al SEQ ID N° relevante nominado.

45 Los mutantes de las secuencias de aminoácidos de los polipéptidos definidos en el presente documento se pueden preparar introduciendo cambios de nucleótidos apropiados en un ácido nucleico definido en el presente documento o mediante síntesis *in vitro* del polipéptido deseado. Tales mutantes incluyen, por ejemplo, supresiones, inserciones o sustituciones de restos dentro de la secuencia de aminoácidos. Se puede efectuar una combinación de supresiones, inserciones y sustituciones para obtener la construcción final, siempre que el producto polipeptídico final posea las características deseadas.

50 Los polipéptidos mutantes (alterados) se pueden preparar usando cualquier técnica conocida en la técnica, por ejemplo, usando estrategias de evolución dirigida o de diseño racional (véase más adelante). Los productos derivados de ADN mutado/alterado se pueden examinar fácilmente usando las técnicas que se describen en el presente documento para determinar si poseen actividad aciltransferasa, por ejemplo, actividad MGAT, DGAT o GPAT/fosfatasa.

60 Cuando se diseñan secuencias de aminoácidos mutantes, la ubicación del sitio de mutación y la naturaleza de la mutación dependerá de la o las características a modificar. Los sitios para la mutación se pueden modificar individualmente o en serie, por ejemplo, por (1) sustitución primero con las elecciones de los aminoácidos conservadores y luego con selecciones más radicales dependiendo de los resultados obtenidos, (2) suprimir el resto de interés o (3) insertar otros restos adyacentes al sitio ubicado.

65 Las supresiones de secuencia de aminoácidos varían en general en un rango entre 1 y 15 restos aproximadamente, más preferentemente entre 1 y 10 restos aproximadamente y normalmente entre 1 y 5 restos contiguos aproximadamente.

En estos mutantes de sustitución se ha eliminado al menos un resto de aminoácido en el polipéptido y se ha insertado un resto diferente en su lugar. Los sitios de mayor interés para la mutagénesis de sustitución incluyen los sitios identificados como sitio(s) activo(s). Otros sitios de interés son aquellos en los cuales los restos particulares obtenidos de diversas cepas o especies son idénticos. Estas posiciones pueden ser importantes para la actividad biológica. Estos sitios, en especial los que se encuentran dentro de una secuencia de al menos tres sitios idénticamente conservados adicionales, se sustituyen preferentemente de una manera relativamente conservadora. Tales sustituciones conservadoras se muestran en la Tabla 1 en la columna con el encabezado de "Ejemplos de sustituciones".

5  
10 En una realización preferida, un polipéptido mutante/variante tiene solamente o no más de, uno o dos o tres o cuatro cambios de aminoácidos conservadores en comparación con un polipéptido natural. Los detalles de los cambios de aminoácidos conservadores se indican en la Tabla 1. Como comprenderá el especialista, tales cambios menores pueden ser razonablemente predichos para no alterar la actividad del polipéptido cuando se expresa en una célula recombinante.

15

**Tabla 1.** Ejemplos de sustituciones.

Resto original	Ejemplos de sustituciones
Ala (A)	val; leu; ile; gly
Arg (R)	lys
Asn (N)	gln; his
Asp (D)	glu
Cys (C)	ser
Gln (Q)	asn; his
Glu (E)	asp
Gly (G)	pro, ala
His (H)	asn; gln
Ile (I)	leu; val; ala
Leu (L)	ile; val; met; ala; phe
Lys (K)	arg
Met (M)	leu; phe
Phe (F)	leu; val; ala
Pro (P)	gly
Ser (S)	thr
Thr (T)	ser
Trp (W)	tyr
Tyr (Y)	trp; phe
Val (V)	ile; leu; met; phe, ala

#### *Evolución dirigida*

20 En la evolución dirigida, se aplica mutagénesis aleatoria a una proteína y se usa un régimen de selección para escoger variantes con las cualidades deseadas, por ejemplo, una mayor actividad aciltransferasa. Además, luego se aplican rondas de mutación y selección. Una típica estrategia de evolución dirigida comprende tres etapas:

25 1) *Diversificación*: El gen que codifica la proteína de interés es mutado y/o recombinado aleatoriamente para crear una gran biblioteca de variantes de genes. Se pueden construir bibliotecas de variantes genes por medio de PCR sujeto a error (véase, por ejemplo, Leung, 1989; Cadwell y Joyce, 1992), a partir de agrupaciones de fragmentos digeridos con DNasa I preparados con templados parentales (Stemmer, 1994a; Stemmer, 1994b; Crameri et al., 1998; Coco et al., 2001) a partir de oligonucleótidos degenerados (Ness et al., 2002, Coco, 2002) o a partir de mezclas de ambos o aún a partir de templados parentales no digeridos (Zhao et al., 1998; Eggert et al., 2005; Jézéquek et al., 2008) y habitualmente se ensamblan por medio de PCR. Las bibliotecas también se pueden obtener a partir de secuencias parentales recombinadas *in vivo* o *in vitro* por recombinación homóloga o no homóloga (Ostermeier et al., 1999; Volkov et al., 1999; Sieber et al., 2001). Las bibliotecas de variantes de genes también se pueden construir por subclonación de un gen de interés en un vector adecuado, transformar el vector en una cepa "mutator" tal como cepa XL-1 red de *E. coli* (Stratagene) y propagar la bacteria transformada por una cantidad adecuada de generaciones. Las bibliotecas de variantes de genes también se puede construir sometiendo al gen de interés a entremezclado de ADN (es decir, recombinación homóloga *in vitro* de agrupaciones de genes mutantes seleccionados por fragmentación y re-ensamblaje aleatorio) como se describe ampliamente en Harayama (1998).

35  
40 2) *Selección*: La biblioteca es evaluada por la presencia de mutantes (variantes) que posean la propiedad deseada usando pruebas o selección. El examen permite identificar y aislar mutantes de alto rendimiento manualmente, en tanto las selecciones automáticamente eliminan todos los mutantes no funcionales. Una prueba puede comprender el examen de la presencia de motivos de aminoácidos conservados conocidos. Como alternativa o además, una prueba puede comprender expresar el polinucleótido mutado en un organismo

hospedador o una parte del mismo y evaluar el nivel de actividad aciltransferasa, por ejemplo, por cuantificación del nivel del producto resultante en el lípido extraído del organismo o una parte del mismo y determinar el nivel de producto en los lípidos extraídos del organismo o una parte del mismo, con relación a un organismo correspondiente o una parte del mismo que no contiene al polinucleótido mutado y, opcionalmente, expresar el polinucleótido parental (no mutado). Como alternativa, la prueba puede comprender alimentar el organismo o una parte del mismo, con el sustrato marcado y determinar el nivel de sustrato o producto en el organismo o una parte del mismo, con relación a un organismo correspondiente o una parte del mismo que no contiene al polinucleótido mutado y, opcionalmente, expresar el polinucleótido parental (no mutado).

3) *Amplificación*: Las variantes identificadas en la selección o prueba se replican muchas veces, lo que permite que los investigadores puedan secuenciar su ADN con el fin de comprender las mutaciones que se produjeron.

Conjuntamente, estas tres etapas se conocen como una "ronda" de evolución dirigida. La mayoría de los experimentos comprenderá más de una ronda. En estos experimentos, los "ganadores" de la ronda previa se diversifican en la ronda siguiente para crear una biblioteca nueva. Al final del experimento, se caracterizan todos los mutantes de proteínas o polinucleótidos desarrollados usando métodos bioquímicos.

#### *Diseño racional*

Se puede diseñar racionalmente una proteína, sobre la base de la información conocida acerca de la estructura y plegamiento de la proteína. Esto se puede lograr por diseño desde cero (diseño *de novo*) o por rediseño basado en andamiajes nativos (véase, por ejemplo, Hallinga, 1997; y Lu y Berry, *Protein Structure Design and Engineering, Handbook of Proteins 2*, 1153-1157 (2007)). El diseño de proteínas comprende normalmente la identificación de secuencias que se pliegan en una estructura dada o blanco y se puede efectuar usando modelos computarizados. Los algoritmos de diseño de proteínas por computadora buscan el espacio de secuencia-conformación para secuencias de energía baja cuando se pliegan en la estructura blanco. Los algoritmos de diseño de proteínas computacionales emplean modelos de energía de proteínas para evaluar la manera en que las mutaciones afectarían la estructura y función de la proteína. Estas funciones de energía incluyen normalmente una combinación de mecánicas moleculares, estadísticas (es decir, basado en el conocimiento previo) y otros términos empíricos. Un software disponible adecuado incluye IPRO (Rediseño y Optimización Interactivo de Proteínas), EGAD (un Algoritmo Genético para el Diseño de Proteínas), Rosetta Design, Sharpen y Abalone.

También se incluyen en el alcance de la invención los polipéptidos definidos en el presente documento que fueron modificados diferencialmente durante o después de la síntesis, por ejemplo, por biotilación, bencilación, glicosilación, acetilación, fosforilación, amidación, derivatización con grupos protectores/bloqueantes conocidos, clivaje proteolítico, unión a una molécula de anticuerpo o a otro ligando celular, etc. Estas modificaciones pueden servir para aumentar la estabilidad y/o bioactividad del polipéptido de la invención.

#### Identificación de aciltransferasas

En el presente documento se divulga método para identificar una molécula de ácido nucleico que codifica una aciltransferasa que tiene mayor capacidad para producir MAG, DAG y/o TAG en una célula.

El método comprende obtener una célula que comprende una molécula de ácido nucleico que codifica una aciltransferasa ligada operativamente a un promotor que es activo en la célula. La molécula de ácido nucleico puede codificar una aciltransferasa natural tal como MGAT, GPAT y/o DGAT o uno o más mutantes de las mismas. Los mutantes se pueden modificar usando procedimientos estándar en la técnica (véase previamente) tal como mediante una mutagénesis aleatoria, mutagénesis dirigida o mutagénesis de saturación en genes de interés conocidos o sometiendo diferentes genes a entremezclado de ADN. Por ejemplo, se puede mutar y/o recombinar aleatoriamente un polinucleótido que comprende una secuencia seleccionada entre cualquiera de las SEQ ID NO: 1 a 44 que codifica una MGAT para crear una gran biblioteca de variantes de genes (mutantes) usando por ejemplo, PCR sujeto a error y/o entremezclado de ADN. Los mutantes se pueden seleccionar para realizar una investigación adicional sobre la base que comprenden un motivo de aminoácido conservado. Por ejemplo, en el caso de un ácido nucleico candidato que codifica una MGAT, un especialista puede determinar si comprende una secuencia provista en las SEQ ID NO: 220, 221, 222, 223 y/o 224 antes de evaluar si el ácido nucleico codifica una MGAT mutante funcional (por ejemplo, por transfección en una célula hospedadora, tal como una célula vegetal y evaluar la actividad aciltransferasa (es decir, MGAT) según se describe en el presente documento). Se puede usar secuenciación directa por PCR del ácido nucleico o un fragmento del mismo, para determinar la secuencia de nucleótidos exacta y deducir la secuencia de aminoácidos correspondiente y de esa manera identificar secuencias de aminoácidos conservadas. Se pueden usar cebadores degenerados basado en secuencias de aminoácidos conservadas para dirigir la amplificación por PCR. También se pueden usar cebadores degenerados como sondas en ensayos de hibridación de ADN. Como alternativa, la o las secuencias de aminoácidos conservadas se pueden detectar en ensayos de hibridación de proteínas que utilizan, por ejemplo, un anticuerpo que se une específicamente a las secuencias de aminoácidos conservadas o un sustrato que se une específicamente a las secuencias de aminoácidos conservadas tal como, por ejemplo, un lípido que se une a FLXLXXXN (un dominio de unión de lípidos neutro putativo; SEQ ID NO: 224).

La molécula de ácido nucleico puede comprender una secuencia de nucleótidos que codifica una MGAT. La secuencia de nucleótidos puede i) comprender una secuencia seleccionada entre cualquiera de las SEQ ID NO:1 a 44, ii) codificar un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO:45 a 82 o un fragmento biológicamente activo de la misma, iii) ser al menos un 50 % idéntica a i) o ii) o iv) hibridarse con cualquiera de i) a iii) en condiciones rigurosas. En una realización adicional, la molécula de ácido nucleico comprende una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos conservadas de DGAT2 y/o MGAT1/2 provistas en las SEQ ID NO:220, 221, 222, 223 y 224. En una realización preferida, la molécula de ácido nucleico comprende una secuencia de nucleótidos que codifica las secuencias de aminoácidos conservadas provistas en la SEQ ID NO:220 y/o en la SEQ ID NO:224.

La molécula de ácido nucleico puede comprender una secuencia de nucleótidos que codifica una GPAT, preferentemente una GPAT con actividad fosfatasa. La secuencia de nucleótidos puede i) comprender una secuencia seleccionada entre cualquiera de las SEQ ID NO:84 a 141, ii) codificar un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO:144 a 201 o un fragmento biológicamente activo de la misma, iii) ser al menos un 50 % idéntica a i) o ii) o iv) hibridarse con cualquiera de i) a iii) en condiciones rigurosas. En una realización adicional, la molécula de ácido nucleico comprende una secuencia de nucleótidos que codifica una o más secuencias de aminoácidos conservadas de GPAT provistas en las SEQ ID NO: 225, 226 y 227 o una secuencia de aminoácidos que es al menos un 50 %, preferentemente al menos un 60 %, más preferentemente al menos un 65 % idéntica a las mismas.

En una realización adicional, la molécula de ácido nucleico comprende una secuencia de nucleótidos que codifica una DGAT2. La secuencia de nucleótidos puede i) comprender una secuencia de nucleótidos seleccionada entre cualquiera de las SEQ ID NO:204 a 211, ii) codificar un polipéptido que comprende aminoácidos cuya secuencia se proporciona en cualquiera de las SEQ ID NO:212 a 219 o un fragmento biológicamente activo de la misma, iii) ser al menos un 50 % idéntica a i) o ii) o iv) hibridarse con cualquiera de i) a iii) en condiciones rigurosas. En una realización preferida adicional, la DGAT2 comprende una secuencia de nucleótidos de la SEQ ID NO: 204 y/o una secuencia de nucleótidos que codifica un polipéptido que comprende aminoácidos cuya secuencia se proporciona en la SEQ ID NO: 212.

Se puede obtener una célula que comprende una molécula de ácido nucleico que codifica una aciltransferasa ligada operativamente a un promotor que es activo en la célula usando procedimientos estándar en la técnica tal como por introducción de la molécula de ácido nucleico en una célula, por ejemplo, mediante precipitación por fosfato de calcio, tratamiento con polietilenglicol, electroporación y combinaciones de estos tratamientos. También se pueden usar otros métodos de transformación de células e incluyen, pero en un sentido no limitativo, la introducción de ADN en plantas por transferencia o inyección directa de ADN. Las células vegetales transformadas también se pueden obtener usando transferencia mediada por *Agrobacterium* y los métodos de aceleración que se describen en el presente documento.

El método comprende además determinar si el nivel de MAG, DAG y/o TAG producido en la célula está aumentado en comparación con una célula correspondiente que no contiene al ácido nucleico usando técnicas conocidas en la técnica, tales como los ejemplificados en el Ejemplo 1. Por ejemplo, los lípidos se pueden extraer en una solución de cloroformo/metanol, secar y separar por cromatografía en capa delgada (TLC). Las identidades de TAG, DAG, MAG, ácidos grasos libres y otros lípidos se pueden verificar usando estándares internos de lípidos después de tinción con vapor de yodo. Los cromatogramas resultantes se pueden analizar usando un equipo FosforImager y la cantidad de MAG, DAG y TAG se puede cuantificar sobre la base de la cantidad conocida de estándares internos o, como alternativa, las células pueden ser alimentados con *sn*-2 monooleoilglicerol[<sup>14</sup>C] o [<sup>14</sup>C]glicerol-3-fosfato y luego se puede cuantificar la radioactividad asociada mediante recuento de centelleo líquido (es decir, se cuantifica la cantidad de MAG, DAG y TAG marcados).

El método comprende además identificar una molécula de ácido nucleico que codifica una aciltransferasa con mayor capacidad para producir MAG, DAG y/o TAG en una célula. En una realización preferida, la aciltransferasa cataliza una reacción enzimática en la vía MGAT. En una realización adicional preferida, DAG aumenta debido a la vía MGAT (es decir, la acilación de MAG con acil-CoA grasa está catalizada por una MGAT para formar DAG). En otra realización o en una forma adicional, el sustrato MAG es producido por una GPAT que también tiene actividad fosfatasa y/o DAG es acilado con acil-CoA grasa por una DGAT y/o una MGAT que tiene actividad DGAT para formar TAG.

## Ejemplos

### 60 **Ejemplo 1. Materiales y métodos generales**

#### Expresión de genes en células vegetales en un sistema de expresión transitorio

Los genes se expresaron en células vegetales usando un sistema de expresión transitorio esencialmente como fue descrito por Voinnet et al., (2003) y Wood et al., (2009). Se introdujeron vectores binarios que contienen a la región de codificación a expresar por un fuerte promotor constitutivo e35S que contiene una región potenciadora duplicada

en la cepa AGL1 de *Agrobacterium tumefaciens*. Además se introdujo por separado el vector binario quimérico, 35S:p19, para la expresión del supresor de silenciamiento viral p19 en AGL1, como se describe en el documento WO2010/057246. Las células recombinantes se cultivaron hasta la fase estacionaria a 28°C en caldo LB suplementado con kanamicina 50 mg/l y rifampicina 50 mg/l. Las bacterias se agruparon luego en pélets por centrifugación a 5000 g durante 5 min a temperatura ambiente antes de resuspenderlas hasta una DO600 = 1,0 en una infiltración solución amortiguadora que contiene MES 10 mM pH 5.7, MgCl<sub>2</sub> 10 mM y acetosiringona 100 µM. Las células se incubaron luego a 28°C con agitación por 3 horas después de lo cual se midió la DO600 y se agregó un volumen de cada cultivo, incluyendo 35S:p19, necesario para alcanzar una concentración final de DO600 = 0,125 en un tubo limpio. El volumen final se conformó con la solución amortiguadora anterior. Luego se infiltraron hojas con la mezcla de cultivo y las plantas se cultivaron normalmente por otros tres a cinco días después de la infiltración antes de recuperar discos foliares para la preparación de lisados celulares purificados o para el aislamiento de lípidos totales. Las infiltraciones control eran con la cepa 35S:p19 solamente.

#### Ensayo con lisados de hojas purificados

Se molieron tejidos foliares de *Nicotiana benthamiana* infiltradas previamente como se describió previamente en una solución que contiene solución amortiguadora de fosfato de potasio 0,1 M (pH 7,2) y sacarosa 0,33 M usando un homogeneizador de vidrio. El homogenato de hojas se centrifugó a 20.000 g durante 45 minutos a 4 °C de spués de lo cual se recolectó cada sobrenadante. El contenido de proteínas en cada sobrenadante se midió de acuerdo con el método Bradford (1976) usando un contador multi-marca Wallac1420 y un reactivo colorante para ensayo de proteínas de Bio-Rad (Bio-Rad Laboratories, Hercules, CA, EE.UU.). Los ensayos con aciltransferasa emplearon 100 µg de proteína de acuerdo con Cao et al., (2007) con algunas modificaciones. El medio de reacción contenía Tris-HCl 100 mM (pH 7,0), MgCl<sub>2</sub> 5 mM, BSA 1 mg/ml (ácidos grasos, libres), sacarosa 200 mM, oleoil-CoA frío 40 mM, *sn*-2 monooleoilglicerol[<sup>14</sup>C] 16,4 µM (55 mCi/mmol, American Radiochemicals, Saint Louis, MO, EE.UU.) o [<sup>14</sup>C]glicerol-3-fosfato (G-3-P) sal disódica 6,0 µM (150 mCi/mmol, American Radiochemicals). Los ensayos se condujeron por 7,5, 15 o 30 minutos.

#### Análisis de lípidos

##### *Análisis de los lípidos de los ensayos de lisados de hojas*

Los lípidos de los ensayos con lisados se extrajeron usando cloroformo:metanol:KCl 0,1 M (2:1:1) y se recuperaron. Las diferentes clases de lípidos presentes en las muestras se separaron sobre placas de cromatografía en capa delgada (TLC) de gel de sílice 60 (MERCK, Darmstadt, Alemania) impregnadas con ácido bórico 10 %. El sistema de solventes usado para fraccionar TAG del extracto de lípidos consistía de cloroformo/acetona (90/10 v/v). Las clases individuales de lípidos se visualizaron por exposición de las placas a vapor de yodo y se identificaron corriendo estándares auténticos en paralelo sobre la misma placa de TLC. Las placas se expusieron a pruebas de obtención de imágenes con fósforo durante la noche y se analizaron con un equipo Fosforimager Fujifilm FLA-5000 antes del recuento de centelleo líquido para la cuantificación del DPM.

##### *Aislamiento y fraccionamiento de lípidos totales*

Los tejidos se secaron por congelamiento, se pesaron y los lípidos totales se extrajeron como se describe en Bligh y Dyer (1959). Las diferentes clases de lípidos presentes en cada muestra se separaron sobre placas de TLC de gel de sílice 60. El sistema de solventes usado para fraccionar los lípidos neutros (NL) y los lípidos polares (PL) consistió de hexano/éter dietílico/ácido acético (70/30/1 v/v/v). Las clases individuales de lípidos se visualizaron por exposición de la placa a vapor de yodo y se identificaron corriendo estándares auténticos en paralelo sobre la misma placa de TLC.

Para determinar la composición de ácidos grasos en las muestras de lípidos, se produjeron metil ésteres de ácidos grasos (FAMES) de NL, incluyendo fracciones TAG, DAG y/o MAG y PL eliminando las correspondientes bandas de las placas de TLC e incubando las mismas en una solución de metanol/HCl/diclorometano (10/1/1 v/v/v) durante 2 horas a 80°C junto con una cantidad conocida de ácido hexadecanoico como estándar interno. Los FAMES se extrajeron en hexano/DCM, se concentraron hasta un volumen pequeño en hexano y se inyectaron en un cromatógrafo gaseoso.

La cantidad de ácidos grasos individuales y totales (TFA) presentes en las fracciones de lípidos se cuantificó determinando el área bajo cada pico y se calculó por comparación con el área pico de la cantidad conocida de estándar interno.

##### *Cromatografía gas-líquido (GC) capilar*

Los FAME se analizaron por cromatografía de gases (GC) usando un cromatógrafo gaseoso Agilent Technologies 6890N (Palo Alto, California, EE.UU.) equipado con una columna capilar de sílice fundida Equity™-1 (15 m x 0,1 mm d.i., 0,1 µm de espesor de película), FID, un inyector *split/splitless* y un muestreador automático e inyector Agilent Technologies 7683 Series. Se usó helio como gas transportador. Las muestras se inyectaron en la modalidad



*splitless* a una temperatura de horno de 120°C. Después de la inyección, se elevó la temperatura del horno a 270°C a razón de 10°C.min<sup>-1</sup> y finalmente a 310°C a razón de 5°C.min<sup>-1</sup>. Los picos se cuantificaron con el software Agilent Technologies ChemStation (Rev B.03.01 (317), Palo Alto, California, EE.UU.).

#### 5 Ensayo con DGAT en H1246 de *Saccharomyces cerevisiae*

La cepa H1246 de *Saccharomyces cerevisiae* carece por completo de actividad DGAT y no contiene TAG ni ésteres de esteroles como resultado de mutaciones de noqueo en cuatro genes (DGA1, LRO1, ARE1, ARE2). La adición de ácidos grasos libre (por ejemplo 1 mM 18:1<sup>Δ9</sup>) al medio de crecimiento de H1246 es tóxica en ausencia de actividad DGAT. El crecimiento sobre dicho medio se puede usar entonces como un indicador o selección de la presencia de actividad DGAT en esta cepa de levadura.

Se transformó la cepa H1246 de *S. cerevisiae* con la construcción pYES2 (control negativo), una construcción que codifica la DGAT1 de *Arabidopsis thaliana* en el pYES2 o una construcción que codifica la MGAT2 de *Mus musculus* en pYES2. Los transformantes fueron alimentados con [<sup>14</sup>C]18:1<sup>Δ9</sup> ácidos grasos libres.

En un experimento separado, se transformó la cepa H1246 de *S. cerevisiae* con la construcción pYES2 (control negativo), una construcción que codifica la DGAT1 de *Bernardia pulchella* en el pYES2 o una construcción que codifica la MGAT1 de *M. musculus* en el pYES2 y fue alimentada con 18:1<sup>Δ9</sup> ácidos grasos libres. La cepa S288C de *S. cerevisiae* de tipo silvestre transformada con el pYES2 sirvió como control positivo.

Los transformantes de levadura se resuspendieron en agua mQ estéril y se diluyeron hasta una DO600 = 1. Las muestras se diluyeron luego en cuatro diluciones consecutivas, cada una 1/10. Se sembraron 2 µl de cada dilución sobre cada una de las placas (YNBD, YNBG, YNBG+FA) junto con 2 µl de agua mQ y 2 µl de una suspensión de células H1246 no transformadas (DO600 = 1). Las placas se incubaron durante 6 días a 30 °C antes de calificar el crecimiento.

*Medio de placas, 40 ml de medio por placa*

- YNBD: medio mínimo selectivo [*dropout*] sin uracilo y que contiene 2 % de glucosa, 0,01 % de NP40 y 100 µl de etanol.
- YNBG: medio mínimo selectivo [*dropout*] sin uracilo y que contiene 2 % de galactosa, 1 % de rafinosa, 0,01 % de NP40 y 100 µl de etanol.
- YNBG+FA: medio mínimo [*dropout*] sin uracilo y que contiene 2 % de galactosa, 1 % de rafinosa, 0,01 % de NP40 y C18:1<sup>Δ9</sup> 1 mM disuelto en 100 µl de etanol.

#### **Ejemplo 2. Expresión constitutiva de una monoacilglicerol aciltransferasa en células vegetales**

##### MGAT1

Se demostró la actividad enzimática de la monoacilglicerol aciltransferasa 1 (MGAT1) codificada por el gen de *M. musculus* (Yen et al., 2002) y de la diacilglicerol aciltransferasa (DGAT1) de *A. thaliana* (Bouvier-Nave et al., 2000), usada aquí para su comparación con la MGAT1, en tejido foliar de *N. benthamiana* usando un sistema de expresión transitorio según se describió en el Ejemplo 1.

Se elaboró un vector denominado 35S-pORE04 por inserción de un fragmento *Pst*I que contiene un promotor 35S en el sitio *Sfo*I del vector pORE04 después de un tratamiento con la T4 ADN polimerasa para obtener extremos adhesivos (Coutu et al., 2007). Se sintetizó un ADN quimérico que codifica la MGAT1 de *M. musculus*, de codones optimizados para *Brassica napus*, mediante Genart y se denominó 0954364\_MGAT\_pMA. Se elaboró un ADN quimérico denominado 35S:MGAT1 y que codifica la MGAT1 de *M. musculus* (Genbank, N° Acceso Q91ZV4) para expresión en células vegetales por inserción de la región de codificación completa de 0954364\_MGAT\_pMA, contenido en un fragmento *Eco*RI, en 35S-pORE04 en el sitio *Eco*RI. El vector que contiene a la construcción 35S:MGAT1 se denominó pJP3184. De manera similar, se elaboró un ADN quimérico 35S:DGAT1 que codifica la DGAT1 de *A. thaliana* (Genbank, N° Acceso AAF19262) para expresión en células vegetales, por inserción de la región de codificación completa de pXZP513E, contenido en un fragmento *Bam*HI-*Eco*RV, en 35S-pORE04 en el sitio *Bam*HI-*Eco*RV. El vector que contiene a la construcción 35S:DGAT1 se denominó pJP2078.

Los vectores quiméricos se introdujeron en la cepa AGL1 de *A. tumefaciens* y las células de los cultivos de la misma se infiltraron en tejido foliar de plantas de *N. benthamiana* en un cuarto de crecimiento a 24°C. Con el fin de permitir comparaciones directas entre muestras y para reducir la variación entre hojas, las muestras comparadas eran infiltradas de cada lado de la misma hoja. Los experimentos se condujeron por triplicado. Después de la infiltración, las plantas se cultivaron por otros tres días antes de recolectar discos foliares, secarlos por congelamiento y los lípidos que se extrajeron de las muestras se fraccionaron y cuantificaron según se describió en el Ejemplo 1. Este análisis reveló que los genes de MGAT1 y de DGAT1 funcionaban aumentando los niveles de aceite en las hojas de *N. benthamiana* de la siguiente manera.

El tejido foliar transformado con la construcción 35S:p19 solamente (control negativo) contenía un promedio de 4 µg de ácidos grasos libres (FFA) derivados de DAG/mg de peso seco de hojas y 5 µg de FFA derivados de TAG/mg de peso seco de hojas. El tejido foliar transformado con las construcciones 35S:p19 y 35S:DGAT1 (control para la expresión de DGAT1) contenía un promedio de 4 µg de FFA derivados de DAG/mg de peso seco de hojas y 22 µg de FFA derivados de TAG/mg de peso seco de hojas. El tejido foliar transformado con las construcciones 35S:p19 y 35S:MGAT1 contenía un promedio de 8 µg de FFA derivados de DAG/mg de peso seco de hojas y 44 µg de FFA derivados de TAG/mg de peso seco de hojas. El tejido foliar transformado con las construcciones 35S:p19, 35S:DGAT1 y 35S:MGAT1 no contenía niveles de DAG o de TAG mayores que los observados en la infiltración con 35S:p19 y 35S:MGAT1 (Figura 2). Además, se notó una disminución en el nivel de saturados en semillas después de la expresión de MGAT en comparación con ya sea el control p19 o las muestras DGAT1 (Tabla 2).

Los datos descritos previamente demostraban que la enzima MGAT1 era mucho más activa que la enzima DGAT1 en la promoción de la acumulación de ambos DAG y TAG en el tejido foliar. La expresión del gen de MGAT1 dio como resultado el doble de acumulación de DAG y TAG en el tejido foliar en comparación con la expresión de DGAT1. Este resultado era muy sorprendente e inesperado, considerando que la MGAT es una enzima que se expresa en intestino de ratón, un sistema biológico muy diferente de las hojas vegetales. Este estudio era la primera demostración de expresión ectópica de MGAT en una célula vegetal.

Las muestras de hoja infiltradas con MGAT1 de *M. musculus* acumularon el doble de DAG y TAG con relación al tejido foliar infiltrado solamente con la DGAT1 de *A. thaliana*. La eficiencia de la producción de TAG también fue sorprendente e inesperado dado que la MGAT de ratón sólo tiene una actividad muy baja como DGAT. El tejido foliar infiltrado con genes que codifican ambas MGAT1 y DGAT1 no acumularon significativamente más TAG que la muestra de hojas de MGAT1 solamente. La Figura 1 es una representación de diversas vías de acumulación de TAG, la mayoría de las cuales convergen en DAG, una molécula central en la síntesis de lípidos. Por ejemplo, MAG, DAG y TAG se pueden interconvertir a través de diversas actividades enzimáticas incluyendo transacilación, lipasa, MGAT, DGAT y PDAT. Se notó asimismo una disminución en el nivel de saturados después de la expresión de MGAT.

### MGAT2

Se elaboró un ADN quimérico denominado 35S:MGAT2 y que codifica la MGAT2 de *M. musculus* para expresión en células vegetales por inserción de la región de codificación completa de MGAT2, contenido en un fragmento *EcoRI*, en 35S:pORE04 en el sitio *EcoRI*. La actividad enzimática de la monoacilglicerol aciltransferasa 2 (MGAT2) codificada por el gen de *M. musculus* (Yen, 2003) (Genbank, N° Acceso Q80W94) y DGAT1 de *A. thaliana* (Bouvier-Nave et al., 2000), usada aquí para su comparación con la MGAT2, se demostró también en tejido foliar de *N. benthamiana* usando un sistema de expresión transitorio según se describió en el Ejemplo 1.

En comparación con los controles, la expresión de DGAT1 incremento el TAG en hojas en 5,9 veces, la MGAT2 en 7,3 veces y la combinación de MGAT2+DGAT1 en 9,8 veces (Figura 3). La capacidad de la MGAT2 sola para obtener tales incrementos significativos en TAG era inesperada por numerosas razones. En primer lugar, se sabe que la cantidad de sustrato MAG presente en el tejido foliar es baja y no serían esperables grandes incrementos en la acumulación de TAG de este sustrato. En segundo lugar, se esperaba que la adición de actividad MGAT solamente (es decir, adición de MGAT2 sin actividad DGAT) solamente diera como resultado DAG, no TAG, en especial en un entorno foliar donde habitualmente hay poca actividad DGAT nativa.

### Discusión

Los inventores de la presente han demostrado sorprendentemente que la expresión transgénica de un gen MGAT da como resultado incrementos significativos en el rendimiento de lípidos en células vegetales. Los inventores de la presente entienden que Tumaney et al. Han aislado una DGAT con alguna actividad MGAT y que no tuvieron éxito en sus intentos por clonar un gen que codificara una MGAT según se define en el presente documento. Tumaney et al., (2001), informaron actividad MGAT en cacahuete y aislaron una enzima responsable de esta actividad. Sin embargo, Tumaney et al. no publicaron los resultados de las pruebas de actividad DGAT y por ello parece que la enzima informada era una DGAT con alguna actividad MGAT. Aún más, en trabajos anteriores no se pudo identificar ninguna actividad MGAT en otras especies (Stobart et al., 1997). Además, era sorprendente que la enzima aislada por Tumaney et al. era una enzima soluble, citosólica, en lugar de una enzima unida a membrana. Finalmente, aunque más tarde Tumaney et al. publicaron resúmenes sosteniendo que habían aislado genes MGAT de cacahuete (Rajasekharan et al., 2006) y de *Arabidopsis* (Ghosh et al., 2006), nunca se publicaron artículos que describieran el aislamiento de estos genes.

### **Ejemplo 3. Demostración bioquímica de la actividad MGAT transgénica en extractos de hojas**

Se obtuvieron lisados celulares de tejido foliar de *N. benthamiana* que había sido infiltrado con 35S:MGAT1, 35S:MGAT2 y 35S:DGAT1, según se describió en el Ejemplo 1. Se condujeron infiltraciones separadas de hojas, cada una por triplicado, para cepas que solamente contienen la construcción 35S:p19 (control negativo), la cepa 35S:MGAT2 junto con la cepa 35S:p19 y una mezcla de las cepas 35S:MGAT2 y 35S:DGAT1 de *Agrobacterium* con

la cepa 35S:p19. Se cosecharon muestras por triplicado después de tres días y se preparó un lisado celular purificado preparado lisis mecánica del tejido y centrifugación. Las actividades MGAT de los lisados celulares purificados se compararon alimentándolos con [<sup>14</sup>C]MAG según se describió en el Ejemplo 1. Los datos se muestran en la Figura 4.

5 Se observó poca actividad MGAT en la muestra control 35S:p19, dado que la mayor parte de la radioactividad permanecía en MAG durante todo el ensayo. Por el contrario, la mayoría del MAG marcado en la muestra 35S:MGAT2 era convertido rápidamente en DAG (Figura 4, panel central), lo cual era indicativo de una fuerte actividad MGAT expresada por la construcción 35S:MGAT2. Además, también se produjo una cantidad significativa de TAG. La producción de TAG observada en la muestra 35S:MGAT2 probablemente se debía a la actividad DGAT nativa de *N. benthamiana* o era producido por otra ruta de síntesis de TAG. La cantidad de TAG producida aumentaba mucho con la adición de 35S:DGAT1 (Figura 4, panel derecho), lo cual indica que la enzima MGAT2 produjo DAG que era accesible para su conversión en TAG por la DGAT1 en tejidos vegetales vegetativos.

#### 15 **Ejemplo 4. Demostración bioquímica de la producción de MAG accesible para MGAT en extractos de hojas**

En los ensayos *in vitro* que se describen en el Ejemplo 3 usando lisados de hojas, los sustratos (*sn*-2 MAG y oleoil-CoA) eran suministrados exógenamente, en tanto la actividad MGAT *in vivo* en tejidos vegetales intactos requerirán la presencia nativa de estos sustratos. La presencia de niveles bajos de MAG en diversos tejidos vegetales ya fue descrita previamente (Hirayama y Hujii, 1965; Panekina et al., 1978; Lakshminarayana et al., 1984; Perry & Harwood, 1993). Para evaluar si la MGAT2 tenía acceso al MAG producido por las vías nativas de las plantas, se repitió el experimento anterior pero esta vez se alimentó [<sup>14</sup>C]G-3-P a los lisados. Los datos resultantes se muestran esquemáticamente en la Figura 5.

25 Se observó producción de MAG marcado en todas las muestras, lo cual indica la producción *de novo* de MAG a partir de G-3-P en los lisados de hojas. También se observaron los productos marcados, DAG y TAG, en todas las muestras aunque eran relativamente bajos en la muestra control 35S:p19, lo cual indica que la producción de estos lípidos neutros por la vía de Kennedy endógena era relativamente baja en esta muestra. Por el contrario, la mayor parte de la marca en las muestras de MGAT2 y MGAT2 + DGAT1 apareció en las agrupaciones de DAG y TAG, lo cual indica que la MGAT agregada exógenamente catalizaba la conversión del MAG producido a partir del G-3-P marcado por una vía nativa de la planta.

Los ejemplos 2 a 4 demuestran diversos puntos clave: 1) El tejido foliar puede sintetizar MAG a partir de G-3-P de manera tal que el MAG es accesible para una MGAT exógena expresada en el tejido foliar; 2) Aún una MGAT que deriva de intestino de mamífero puede funcionar en tejidos vegetales, en los cuales no habría una MGAT endógena, lo que requiere de una interacción exitosa con otros factores de plantas relacionados con la síntesis de lípidos; 3) El DAG producido por la actividad MGAT exógena se encuentra accesible para una DGAT vegetal o una DGAT exógena, para producir TAG; y 4) la expresión de una MGAT exógena puede dar niveles muy aumentados de TAG en tejidos vegetales, siendo dichos niveles al menos tan elevados como los obtenidos con la expresión de la DGAT1 exógena de *A. thaliana*.

#### 40 **Ejemplo 5. Expresión de DGAT1, MGAT1 y MGAT2 en levadura**

Se construyeron vectores de expresión quiméricos de levadura por inserción de genes que codifican la DGAT1 de *A. thaliana*, MGAT1 de *M. musculus* y MGAT2 de *M. musculus* en el vector pYES2 para obtener pYES2:DGAT1, pYES2:MGAT1 y pYES2:MGAT2. Estas construcciones fueron transformadas en la cepa H1246 de *Saccharomyces cerevisiae* que carece por completo de actividad DGAT y no contiene TAG ni ésteres de esteroles como resultado de mutaciones de noqueo en cuatro genes (DGA1, LRO1, ARE1, ARE2). La cepa H1246 de levadura tiene la capacidad para sintetizar DAG a partir de ácidos grasos agregados exógenamente, pero no puede convertir el DAG en TAG debido a las mutaciones de noqueo. Los cultivos de levadura transformados fueron alimentados con [<sup>14</sup>C]18:1<sup>Δ9</sup> antes que se extrajeran y fraccionaran los lípidos totales por TLC según se describió en el Ejemplo 1. En la Figura 6 se muestra un autorradiograma de una placa de TLC representativa.

Se observó formación de TAG, indicativo de la presencia de actividad DGAT, en las células de levadura que contienen DGAT1 (control positivo) y la MGAT1 de mamífero, pero no en las células que contienen a la MGAT2. Se concluyó que la MGAT1 de ratón también tenía actividad DGAT en células de levadura y por ello funcionaba como una enzima de función doble MGAT/DGAT, en tanto MGAT2 no presentaba una actividad DGAT detectable y por ello solamente era una MGAT.

#### 60 **Ejemplo 6. Expresión de una monoacilglicerol aciltransferasa en plantas, semillas y hongos**

##### Expresión de MGAT1 en semillas de *Arabidopsis thaliana*

Se usó un gen que codifica MGAT1 *M. musculus* y se encuentra bajo el control de un promotor específico de semillas (FP1, un promotor de napina truncado de *Brassica napus*) para generar plantas y semillas de la progenie transformadas de forma estable de *A. thaliana*. El vector denominado pJP3174 se obtuvo por inserción de un

fragmento *Sall* que contiene un sitio *EcoRI* flanqueado por el promotor FP1 y la señal de poliadenilación de lectina de *Glycine max* en el sitio *Sall*-*XhoI* del vector pCW141. El vector pCW141 también contenía un GFP dirigido por FP1, interrumpido por intrón, secretado por semillas como gen marcador rastreable. El gen quimérico denominado FP1:MGAT1-GFP se obtuvo por inserción de la región de codificación completa de la construcción 0954364\_MGAT\_pMA, contenida en un fragmento *EcoRI*, en el pJP3174 en el sitio *EcoRI*, generando así el pJP3179. Este vector quimérico se introdujo en la cepa AGL1 de *A. tumefaciens* y las células del cultivo del *Agrobacterium* transformado se usó para tratar plantas de *A. thaliana* (ecotipo Columbia) usando el método de inmersión floral para la transformación (Clough y Bent, 1998). Después de la maduración, las semillas de las plantas tratadas se visualizaron bajo un microscopio de disección Leica MZFLIII y una lámpara de mercurio 100 ebq. Se aislaron quince semillas transgénicas (fuertemente GFP-positivas) y quince semillas no transgénicas (GFP-negativas) y se agrupó cada conjunto. Las agrupaciones GFP-positivas y GFP-negativas se analizaron por su contenido total de ácidos grasos según se describió en el Ejemplo 1. Este análisis proporcionó el contenido promedio y la composición de ácidos grasos para semillas transformadas con la construcción MGAT, pero en una población que puede haber contenido semillas transformadas hemicigotas y homocigotas.

Este análisis reveló que el gen MGAT1 funcionaba aumentando los niveles de aceite de semillas en semillas de *A. thaliana*, donde las quince semillas no transgénicas (control, igual que el tipo silvestre) contienen un promedio de 69,4 µg de ácidos grasos totales en tanto las quince semillas transgénicas transformadas con el gen GFP y por ello era probable que contuvieran a la construcción genética FP1:MGAT1, contenían un promedio de 71,9 µg de ácidos grasos totales. Esto representaba un aumento del 3,5 % en el contenido de aceite con relación al control (tipo silvestre). El análisis también reveló que el gen MGAT enriquecía las semillas en ácidos grasos poliinsaturados, como se puede ver por la composición de ácidos grasos de los lípidos totales extraídos obtenidos a partir de las semillas. En particular, la cantidad de ALA presentes como un porcentaje de los ácidos grasos totales extraídos de las semillas aumentó del 16,0 % al 19,6 %. De manera similar, el porcentaje de los ácidos grasos 20:2n6 aumentó del 1,25 % al 1,90 % y los ácidos grasos 20:3n3 aumentó del 0,26 % al 0,51 % (Tabla 2).

**Tabla 2.** Efecto de la expresión de MGAT sobre la composición de ácidos grasos de semillas.

Perfil de FA (% de TFA)										
	16:00	16:01	16:3w3	C18:0	C18:1d9	C18:1d11	C18:2	C18:3		
<b>Muestra</b>										
<b>Control</b>	7,41	0,36	0,12	3,00	15,26	1,98	30,93	15,98		
<b>MGAT1</b>	7,11	0,32	0,11	2,95	13,86	1,51	28,87	19,59		
	C20:0	20:1d11	20:1iso	20:2n6	20:3n3	C22:0	C22:1	C24:0	24:1d15	
<b>Muestra</b>										
<b>Control</b>	1,86	17,95	1,74	1,25	0,26	0,57	0,98	0,20	0,17	100,00
<b>MGAT1</b>	1,90	17,22	1,71	1,90	0,51	0,57	1,52	0,19	0,17	100,00

Se condujo un experimento adicional, en donde se modificó el ADN quimérico FP1:MGAT1-GFP para eliminar el gen GFP. Esta construcción genética, denominado FP1:MGAT1, se transformó en una línea de *A. thaliana* que era mutante para FAD2. Se determinó el contenido total de ácidos grasos de semillas de T<sub>2</sub> de plantas T<sub>1</sub> resistentes a antibióticos, así como líneas parentales cultivadas junto con estas plantas, de acuerdo con el Ejemplo 1. Los datos se muestran en la Tabla 3. El promedio de ácidos grasos totales de semillas de las líneas control era de 347,9 µg/100 semillas en tanto el promedio en semillas transgénicas era de 381,0 µg/100 semillas. Cuando se excluyeron los datos para la línea control C6 para determinar el promedio, el promedio para los controles era de 370 µg/100 semillas. El contenido de aceite en las semillas transgénicas representaba un aumento de aproximadamente un 3 % en términos relativos en comparación con el contenido de aceite en las semillas no transformadas.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Arabidopsis*. Las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Tabla 3. Perfiles de ácidos grasos y cuantificación de ácidos grasos totales en semillas T<sub>2</sub> transgénicas FP1:MGAT1 y controles parentales de *Arabidopsis thaliana*.

Muestra	C16:0	C16:1	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:1d11	C22:0	C24:0	24:1d15	µg de FA/100 semillas
C7	6,2	0,5	2,5	81,3	4,2	0,6	2,7	0,7	0,7	0,3	0,2	0,1	442,5
C4	6,3	0,4	2,4	81,7	3,9	0,5	2,5	0,9	0,6	0,4	0,2	0,1	403,8
C8	6,4	0,5	2,6	81,1	4,1	0,6	2,6	0,8	0,6	0,4	0,2	0,1	403,2
C2	6,2	0,5	2,4	81,4	4,1	0,6	2,7	0,8	0,6	0,4	0,2	0,1	377,0
C1	6,4	0,5	2,4	80,6	4,1	0,7	3,3	0,8	0,6	0,4	0,2	0,1	344,8
C3	6,4	0,5	2,6	80,0	4,1	0,6	3,5	0,8	0,6	0,4	0,2	0,2	314,3
C5	6,3	0,5	2,6	80,7	4,4	0,6	2,4	0,7	0,6	0,4	0,2	0,1	310,6
C6	6,7	0,7	2,7	77,2	5,0	0,8	4,3	0,9	0,7	0,4	0,3	0,2	186,8
M23	5,9	0,4	2,0	81,4	5,0	0,8	2,4	0,7	0,7	0,5	0,2	0,2	455,7
M10	6,0	0,4	2,4	82,3	4,2	0,7	2,2	0,7	0,6	0,3	0,2	0,1	437,7
M22	5,9	0,4	2,2	81,4	4,8	0,8	2,4	0,7	0,6	0,4	0,2	0,2	425,0
M25	6,0	0,4	2,2	81,7	4,6	0,7	2,4	0,8	0,6	0,3	0,2	0,1	406,7
M8	6,0	0,4	2,2	81,6	4,5	0,8	2,5	0,7	0,6	0,3	0,2	0,2	404,5
M14	5,7	0,4	2,1	81,8	4,6	0,8	2,5	0,7	0,6	0,3	0,2	0,2	396,4
M26	6,2	0,4	2,2	81,8	4,4	0,8	2,2	0,7	0,6	0,4	0,2	0,1	393,0
M6	5,9	0,4	2,2	81,8	4,5	0,8	2,4	0,7	0,6	0,3	0,2	0,2	392,9
M5	5,9	0,5	2,2	80,9	4,8	0,9	2,6	0,7	0,7	0,5	0,2	0,2	389,7
M17	6,3	0,4	2,3	75,1	4,7	4,8	4,4	0,7	0,6	0,3	0,2	0,1	388,4
M9	6,1	0,4	2,3	81,8	4,4	0,7	2,4	0,7	0,6	0,3	0,2	0,1	388,2
M20	6,2	0,4	2,2	81,5	4,7	0,7	2,3	0,7	0,6	0,3	0,2	0,1	379,1
M12	6,2	0,4	2,2	81,6	4,4	1,0	2,2	0,7	0,6	0,3	0,2	0,1	374,7
M18	6,2	0,4	2,4	81,3	4,7	0,7	2,3	0,8	0,5	0,3	0,2	0,1	369,1
M24	6,1	0,4	2,2	81,6	4,6	0,7	2,3	0,9	0,6	0,3	0,2	0,1	361,7
M7	5,9	0,4	2,3	81,9	4,5	0,7	2,3	0,7	0,6	0,3	0,2	0,1	359,4
M4	6,2	0,4	2,4	81,3	4,4	0,7	2,5	0,8	0,5	0,3	0,2	0,2	352,3
M13	6,1	0,4	2,3	81,4	4,5	0,8	2,4	0,7	0,6	0,3	0,2	0,1	352,0
M16	6,1	0,4	2,2	81,8	4,4	0,7	2,4	0,7	0,6	0,3	0,2	0,2	340,5
M19	6,1	0,4	2,3	80,9	4,9	0,8	2,7	0,7	0,5	0,3	0,2	0,1	318,6
M3	6,0	0,5	2,3	81,1	4,7	0,9	2,7	0,7	0,6	0,3	0,2	0,2	316,6

Expresión de MGAT1 en semillas de *Brassica napus*

Se usó el vector FP1:MGAT1 usado para la expresión de MGAT1 de *M. musculus* en semillas de *Arabidopsis thaliana* para generar plantas transformadas de *B. napus*. El vector se introdujo en la cepa AGL1 de *A. tumefaciens* mediante procedimientos estándar de electroporación. Los cultivos se dejaron crecer durante la noche a 28°C en medio LB con agitación a 150 rpm. Las células bacterianas se recolectaron por centrifugación a 4000 rpm durante 5 minutos, se lavaron con medio AB de Winans (Winans, 1988) y se resuspendieron en 10 ml de medio AB de Winans (pH 5,2) y se cultivaron con kanamicina (50 mg/l) y rifampicina (25 mg/l) durante la noche con la adición de acetosiringona 100 µM. Dos horas antes de la infección de las células de *Brassica*, se agregó espermidina (120 mg/l) y se ajustó la densidad final de las bacterias hasta una DO 600nm de 0,3-0,4 con medio AB fresco. Se infectaron peciolo de cotiledones recién aislados de plántulas de 8 día de vida de *B. napus* cultivadas sobre medio 1/2 MS (Murashige-Skoog, 1962) o segmentos de hipocotilo preacondicionado por 3-4 días sobre medio MS con tiazurón (TDZ) 1 mg/l + ácido alfa-naftalenacético (NAA) 0,1 mg/l con 10 ml de cultivos de *Agrobacterium* durante 5 minutos. Los explantes (peciolo de cotiledón e hipocotilo) infectados con *Agrobacterium* fueron transferidos luego sobre papel de filtro estéril para eliminar el exceso de *Agrobacterium* y luego fueron transferidos a medio de cocultivo (medio MS con TDZ 1 mg/l + NAA 0,1 mg/l + acetosiringona 100 µM) suplementado con o sin diferentes antioxidantes (L-cisteína 50 mg/l y ácido ascórbico 15 mg/l). Todas las placas se sellaron con Parafilm y se incubaron en oscuridad a 23-24 °C durante 48 horas.

Los explantes cocultivados (peciolo de cotiledón e hipocotilo) se lavaron luego con agua destilada estéril + cefotaxime 500 mg/l + timentina 50 mg/l durante 10 minutos, se lavaron en agua destilada estéril durante 10 minutos, se secaron sobre papel de filtro estéril, se transfirieron a medio de preselección (MS + TDZ 1 mg/l + NAA 0,1 mg/l + sulfato de adenina (ADS) 20 mg/l + AgNO<sub>3</sub> 1,5 mg/l + cefotaxime 250 mg/l y timentina 50 mg/l) y se cultivaron por cinco días a 24°C con un fotoperíodo de 16 horas/8 horas. Después fueron transferidos a medio de selección (MS + TDZ 1 mg/l + NAA 0,1 mg/l + ADS 20 mg/l + AgNO<sub>3</sub> 1,5 mg/l + cefotaxime 250 mg/l y timentina 50 mg/l) con glufosinato de amonio 1,5 mg/l y se cultivaron durante 4 semanas a 24°C con un fotoperíodo de 16 horas/8 horas, con un subcultivo bisemanal sobre el mismo medio. Los explantes con callos verdes fueron transferidos a medio de inicio de brotes (MS + kinetina 1 mg/l + ADS 20 mg/l + AgNO<sub>3</sub> 1,5 mg/l + cefotaxime 250 mg/l + timentina 50 mg/l + glufosinato de amonio 1,5 mg/l) y se cultivaron durante otras 2-3 semanas. Los brotes que emergieron de los explantes resistentes fueron transferidos a medio de elongación de brote (medio MS con ácido giberélico 0,1 mg/l + ADS 20 mg/l + AgNO<sub>3</sub> 1,5 mg/l + cefotaxime 250 mg/l + glufosinato de amonio 1,5 mg/l) y se cultivaron durante otras dos semanas. Se seleccionaron brotes saludables de 2-3 cm de longitud y se transfirieron a medio de enraizamiento (1/2 MS con NAA 1 mg/l + ADS 20 mg/l + AgNO<sub>3</sub> 1,5 mg/l + cefotaxime 250 mg/l) y se cultivaron durante 2-3 semanas. Los brotes con raíces bien establecidos fueron transferidos a macetas (mezcla para plántulas) y se cultivaron en un gabinete de crecimiento durante dos semanas y a continuación fueron transferidos a invernadero. Se confirmó que dieciséis transformantes individuales eran transgénicos para la construcción FP1:MGAT1 y crecieron normalmente bajo las condiciones de invernadero. El crecimiento aparentemente era normal y las plantas eran fértiles, florecían y establecían semillas normalmente. Las plantas se cultivaron hasta la madurez, se cosecharon las semillas obtenidas de las plantas transformadas y luego eran analizadas por el contenido de aceite de semillas y la composición de ácidos grasos. La expresión específica de semillas de MGAT1 aumenta el contenido de aceite y aumenta el porcentaje de ácidos grasos poliinsaturados en el aceite de semillas de *Brassica*.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Brassica*. Las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Expresión de MGAT1 en semillas de *Gossypium hirsutum*

Se usó el mismo gen quimérico específico de semillas usado para la expresión de MGAT1 de *M. musculus* en semillas de *Arabidopsis thaliana* para generar plantas transformadas de *Gossypium hirsutum*. El vector denominado FP1:MGAT1 se introdujo en la cepa AGL1 de *A. tumefaciens* mediante procedimientos estándar de electroporación y las células del cultivo de *Agrobacterium* se usaron para introducir los ADN quiméricos en las células de *Gossypium hirsutum*, variedad Coker315. Se usaron cotiledones recortados de plántulas de algodón de 10 días de vida como explantes y se infectaron y se cocultivaron con *A. tumefaciens* por un período de dos días. Esto fue seguido por una selección de seis semanas sobre medio MS (Murashige y Skoog, 1962) que contiene 2,4-D 0,1 mg/l, kinetina 0,1 mg/l, sulfato de kanamicina 50 mg/l y cefotaxime 25 mg/l. Se transfirieron callos saludables derivados de los explantes de cotiledones a medio MS que contiene 6-(γ,γ-dimetilalilamino)purina (2ip) 5 mg/l, ácido naftalenacético (NAA) 0,1 mg/l, kanamicina 25 mg/l y cefotaxime 250 mg/l por un segundo período de seis semanas a 28 °C. Los embriones somáticos formados después de aproximadamente seis a diez semanas de incubación se hicieron germinar y se mantuvieron sobre el mismo medio, pero sin adición de fitohormona o antibióticos. Las plántulas desarrolladas a partir de los embriones somáticos fueron transferidas a tierra y se mantuvieron en invernadero una vez que se habían desarrollado hojas y raíces, a una temperatura de crecimiento de 28 °C/20 °C (día/noche). Las plantas transgénicas que contienen a la construcción FP1-MGAT1 se cultivaron en invernadero, florecieron y produjeron copos que contienen semillas. Las semillas se cosechan al alcanzar la madurez. La expresión específica de semillas de MGAT1 aumenta el contenido de aceite y aumenta el porcentaje de ácidos grasos poliinsaturados en el aceite de semillas de algodón.

Expresión de genes de MGAT1 y MGAT2 en plantas de *N. benthamiana* después de una transformación estable

Se transformó *N. benthamiana* de forma estable con la construcción 35S:MGAT1 descrita en el Ejemplo 2. Se introdujo 35S:MGAT1 en la cepa AGL1 de *A. tumefaciens* mediante un procedimiento estándar de electroporación. Las células transformadas se cultivaron sobre medio LB sólido suplementado con kanamicina (50 mg/l) y rifampicina (25 mg/l) y se incubaron a 28 °C por dos días. Se usó una sola colonia para iniciar cultivos frescos. Después de 48 horas de cultivo vigoroso, las células se recolectaron por centrifugación a 2.000x g y se eliminó el sobrenadante. Las células se resuspendieron en solución fresca que contiene 50 % de medio LB y 50 % de medio MS a una densidad  $DO_{600} = 0,5$ .

Se recortaron muestras de hojas de *N. benthamiana* cultivada *in vitro* y se cortaron en secciones cuadradas de 0,5-1 cm<sup>2</sup> aproximadamente con un bisturí filoso mientras estaban sumergidas en la solución de *A. tumefaciens*. Los trozos de hoja lesionados de *N. benthamiana* sumergidos en *A. tumefaciens* se dejaron en reposo a temperatura ambiente durante 10 minutos antes de secarlos sobre un papel de filtro estéril y luego se transfirieron sobre placas MS sin suplemento. Después de un período de co-cultivo de dos días a 24°C, los explantes se lavaron tres veces con medio MS líquido, estéril, luego se secaron con papel de filtro estéril y se colocaron sobre el agar MS selectivo suplementado con bencilaminopurina (BAP) 1,0 mg/l, ácido indolacético (IAA) 0,25 mg/l, kanamicina 50 mg/l y cefotaxime 250 mg/l. Las placas se incubaron a 24°C por dos semanas para permitir el desarrollo de brote a partir de los trozos de hoja transformadas de *N. benthamiana*.

Para establecer plantas transgénicas con raíces *in vitro*, se recortaron brotes verdes saludables y se transfirieron a recipientes para cultivo tisular de 200 ml que contienen medio de agar MS suplementado con IAA 25 µg/l, kanamicina 50 mg/l y cefotaxime 250 mg/l. Se tomaron discos foliares suficientemente grandes de los brotes transgénicos y se secaron por congelamiento para el fraccionamiento de TAG y el análisis de cuantificación según se describió en el Ejemplo 1 (Tabla 4). La mejor planta 35S:MGAT1 de *N. benthamiana* tenía un contenido de TAG de 204,85 µg/100 mg de peso seco de tejido foliar en comparación con un contenido promedio de TAG de 85,02 µg/100 mg de peso seco de tejido foliar en las líneas control, lo que representa un aumento en el contenido de TAG del 241 %.

*N. benthamiana* también se transformó de forma estable con la construcción 35S:MGAT2 que se describió en el Ejemplo 2 y un vector binario control pORE4 (Tabla 5). La mejor planta 35S:MGAT2 de *N. benthamiana* tenía un contenido de TAG de 79,0 µg/100 mg de peso seco de tejido foliar en comparación con un contenido de TAG de 9,5 µg/100 mg de peso seco de tejido foliar en la línea control, lo que representa un aumento en el contenido de TAG del 731 %. El perfil de ácidos grasos de las fracciones de TAG también estaba alterado con niveles significativamente reducidos de los ácidos grasos saturados 16:0 y 18:0 y niveles aumentados de ácidos grasos poliinsaturados, en particular 18:3ω3 (ALA) (Tabla 5). El perfil de ácidos grasos de los lípidos polares las mismas muestras de hojas no se vieron significativamente afectadas, lo cual indica que los cambios en la composición de ácidos grasos de los lípidos no polares era real. Las plantas control en este experimento eran más pequeñas y fisiológicamente diferentes que en el experimento previo con la construcción 35S:MGAT1 y esto podría explicar el contenido diferente de aceites de las plantas control entre un experimento y otro. Los experimentos para comparar directamente las construcciones 35S:MGAT1 y 35S:MGAT2 con las plantas control se condujeron usando plantas del mismo tamaño y fisiología.

Se obtuvo un nuevo juego de vectores de expresión binarios constitutivos usando un promotor 35S con una región potenciadora duplicada (e35S). Las construcciones 35S:MGAT1#2 (pJP3346), 35S:MGAT2#2 (pJP3347) y 35S:DGAT1#2 (pJP3352) se obtuvieron clonando primero el promotor e35S, contenido dentro de un fragmento *Bam*HI-*Eco*RI, en el pORE04 en los sitios *Bam*HI-*Eco*RI para obtener el pJP3343. Los plásmidos pJP3346 y pJP3347 se produjeron luego por clonación de los genes MGAT1 y MGAT2, respectivamente, en el sitio *Eco*RI del pJP3343. El pJP3352 se produjo por clonación de DGAT1 de *A. thaliana*, contenido en un sitio *Xho*I-*Asi*SI, en los sitios *Xho*I-*Asi*SI del pJP3343.

Se usaron los plásmidos pJP3346, pJP3347 y pJP3352 en la cepa AGL1 de *Agrobacterium* para transformar *N. benthamiana* como se describió previamente. Se recuperaron catorce plantas transgénicas confirmadas para pJP3346 y 22 para pJP3347. Se había generado un número de brotes transformados resistentes a kanamicina con el pJP3352. El análisis de la expresión de los transgenes se condujo en plantas transformadas con MGAT1 o MGAT2. Se seleccionaron las plantas con niveles elevados de expresión. Se efectúa un análisis de la expresión en plantas transformadas con DGAT1 de *A. thaliana*. Las plantas crecen normalmente y se dejan crecer hasta alcanzar la madurez. Las semillas se cosechan cuando están maduras. Las semillas de la progenie de gran expresión se siembran directamente en tierra para el análisis de lípidos de la población segregante T2, que incluye plantas homocigotas y heterocigotas. El contenido de aceite en las hojas de plantas que expresan niveles elevados de MGAT1 o MGAT2 es significativamente mayor en comparación con las plantas transformadas con DGAT1 de *A. thaliana* o las plantas control.

También se usó pJP3346, pJP3347 y un vector control en AGL1 para transformar *A. thaliana* como se describió previamente. Se identificaron veinticinco plantas transgénicas T2 confirmadas que comprenden al ADN-T del pJP3346 y 43 plantas transgénicas para pJP3347. El análisis de la expresión se efectúa en las plantas transgénicas.

## ES 2 640 100 T3

Se cosechan las semillas de la progenie de gran expresión y se siembran directamente en tierra. Se efectúa un análisis de lípido incluyendo contenido el de aceite de las hojas de la progenie T2 y T3, incluyendo de los segregantes que no contienen a los transgenes. Los mayores niveles de TAG se obtienen en las plantas que son homocigotas para los transgenes MGAT.



**Tabla 4.** Perfil de ácidos grasos y cuantificación de TAG en tejido foliar de *Nicotiana benthamiana* transformado de forma estable con la construcción 35S:MGAT1. Las muestras 'M' son 35S:MGAT1, en tanto las muestras 'C' son plantas control parentales.

Muestra	C16:0	16:3w3	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:3n3	C22:0	C24:0	µg/100 mg de PS
M1	38,7	0,7	5,1	8,5	0,4	7,0	34,4	1,1	0,3	0,2	0,4	204,85
M8	33,2	0,8	4,4	8,1	0,3	6,5	42,8	0,9	0,2	0,2	0,2	184,20
M3	41,1	0,6	5,3	10,4	0,4	5,5	31,8	1,0	0,4	0,2	0,2	133,62
M2	42,5	0,5	5,2	7,4	0,0	4,8	34,4	1,0	0,2	0,3	0,2	133,57
M7	35,2	0,6	4,5	8,6	0,0	4,9	41,7	1,1	0,3	0,3	0,2	128,49
M5	49,1	0,6	6,4	9,0	0,4	3,7	16,9	1,1	0,0	0,5	0,7	107,39
M4	41,9	0,4	6,0	9,6	0,0	4,2	33,0	1,1	0,2	0,4	0,2	93,71
M6	41,4	0,4	5,8	8,2	0,0	4,3	34,6	1,1	0,2	0,3	0,2	88,38
C1	40,2	0,4	6,1	8,3	0,0	7,8	31,9	1,3	0,2	0,4	0,3	81,53
C2	39,9	0,6	5,5	7,1	0,0	6,9	35,4	1,1	0,3	0,4	0,3	88,52

**Tabla 5.** Perfil de ácidos grasos y cuantificación de TAG en tejido foliar de *Nicotiana benthamiana* transformado de forma estable con la construcción 35S:MGAT2. Las muestras 'M' son 35S:MGAT2, en tanto las muestras 'C' son plantas control parentales. Se tomaron dos hojas de cada planta y se analizaron por separado.

Muestra	C16:0	16:1d7	16:1d13t	C16:1	16:3n3	C18:0	C18:1	C18:1d11	C18:2	C18:3n	C20:0	µg/100 mg de PS
C, hoja 1	34,0	2,7	0,8	0,0	0,0	17,3	6,6	0,0	15,9	18,7	0,0	12,9
C, hoja 2	35,0	1,8	0,0	0,0	1,3	25,0	3,0	0,0	13,0	17,6	1,4	6,1
M, hoja 1	14,6	0,4	1,0	0,4	7,7	5,9	4,0	0,4	16,8	47,0	0,6	97,1
M, hoja 2	18,1	0,3	1,0	0,0	6,0	8,1	2,8	0,3	14,0	46,9	1,0	60,9
C, hoja 1	13,4	0,0	3,0	0,2	7,4	2,0	2,5	0,4	8,4	61,4	0,3	2439,3
C, hoja 2	10,3	0,0	2,4	0,2	9,7	1,4	2,0	0,3	9,5	63,3	0,0	4811,5
M, hoja 1	11,6	0,0	2,4	0,2	8,7	1,9	2,4	0,3	8,7	63,0	0,0	3568,8
M, hoja 2	10,7	0,0	2,4	0,2	9,5	1,6	1,9	0,3	9,2	63,3	0,0	3571,2

Expresión de MGAT1 en plantas de *Trifolium repens* transformadas de forma estable

Se usó un gen quimérico que codifica MGAT1 de *M. musculus* para transformar *Trifolium repens*, otra planta dicotiledónea. Los vectores que contienen los genes quiméricos 35S:MGAT1 y 35S:DGAT1 se introdujeron en *A. tumefaciens* por medio de un procedimiento de electroporación estándar. Ambos vectores también contienen un gen marcador de selección 35S:BAR. Las células transformadas de *Agrobacterium* se cultivaron sobre medio LB sólido suplementado con kanamicina (50 mg/l) y rifampicina (25 mg/l) y se incubaron a 28 °C por dos días. Se usó una sola colonia para iniciar un cultivo fresco para cada construcción. Después de 48 horas de cultivo vigoroso, los cultivos de *Agrobacterium* se usaron para tratar cotiledones de *T. repens* (cv. Haifa) que habían sido disecadas de semillas embebidas como se describe en Larkin et al. (1996). Después de tres días de co-cultivo, los explantes fueron expuestos a PPT 5 mg/l para seleccionar los brotes transformados y luego fueron transferidos a medio de enraizamiento para formar raíces, antes de su transferencia a tierra. Se obtuvo una planta transformada que contiene MGAT1. El promotor 35S se expresa de manera constitutiva en las células de las plantas transformadas. El contenido de aceite está aumentado al menos en los tejidos vegetativos, tales como hojas.

Expresión de MGAT en *Hordeum vulgare* transformada de forma estable

Se usó un vector quimérico que incluye MGAT1 de *M. musculus* para producir *Hordeum vulgare*, una planta monocotiledónea, transformada de forma estable. Los vectores que contienen a los genes quiméricos Ubi:MGAT1 y Ubi:DGAT1 se construyeron por clonación por separado de las regiones de codificación completas de MGAT1 de *M. musculus* y DGAT1 de *A. thaliana* en el pWVEC8-Ubi. Los vectores que contienen los genes quiméricos Ubi:MGAT1 y Ubi:DGAT1 se introdujeron en la cepa AGL1 de *A. tumefaciens* por medio de un procedimiento de electroporación estándar. Las células transformadas de *Agrobacterium* se cultivaron sobre medio LB sólido suplementado con kanamicina (50 mg/l) y rifampicina (25 mg/l) y las placas se incubaron a 28 °C por dos días. Se usó una sola colonia de cada una para iniciar cultivos frescos.

Después de 48 horas de cultivo vigoroso, los cultivos de *Agrobacterium* se usaron para transformar células en embriones inmaduros de cebada (cv. Golden Promise) de acuerdo con los métodos publicados (Tingay et al., 1997; Bartlett et al., 2008) con algunas modificaciones. Brevemente, se aislaron embriones de entre 1,5 y 2,5 mm de longitud de cariopses inmaduros y se retiraron los ejes embrionarios. Los explantes resultantes se co-cultivaron durante 2-3 días con el *Agrobacterium* transgénico y luego se cultivaron en oscuridad durante 4-6 semanas sobre un medio que contiene timentina e higromicina para generar callos embriogénicos antes de moverlos a un medio de transición bajo condiciones de poca luz por dos semanas. Los callos fueron transferidos luego a un medio de regeneración para permitir la regeneración de brotes y raíces antes de su transferencia a tierra. Se obtuvieron plantas transformadas y fueron transferidas al invernadero. La región de codificación de MGAT1 se expresa de manera constitutiva bajo el control del promotor Ubi en las células de las plantas transformadas. El contenido de aceite aumentado al menos en los tejidos vegetativos.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Hordeum* como se describió previamente. Los tejidos vegetativos de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Expresión de MGAT en células de levadura

Se usó un vector quimérico que incluye MGAT1 de *M. musculus* para transformar levadura; en este ejemplo se trata de *Saccharomyces cerevisiae*, un microbio fúngico adecuado para la producción de aceite por fermentación. Se obtuvo una construcción genética Gal1:MGAT1 por inserción de la región de codificación completa de una construcción denominada 0954364\_MGAT\_pMA, contenida en un fragmento EcoRI, en el pYES2 en el sitio EcoRI, generando así el pJP3301. De manera similar, se obtuvo una construcción genética Gal1:DGAT1, usada aquí por comparación y que codifica separadamente la enzima DGAT1 de *A. thaliana*, por inserción de la región de codificación completa de DGAT1 de *A. thaliana* en el pYES2. Estos vectores quiméricos se introdujeron en la cepa S288C de *S. cerevisiae* por shock de calor y los transformantes se seleccionaron sobre placas con medio mínimo de levadura (YMM) que contienen 2 % de rafinosa como única fuente de carbono. Se establecieron cultivos de inóculos clonales en YMM líquido con 2 % de rafinosa como única fuente de carbono. A partir de ellos se inocularon cultivos experimentales en medio YMM que contiene 1 % de NP-40, hasta una DO600 inicial de 0,3 aproximadamente. Los cultivos se dejaron crecer a 28°C con agitación (100 rpm aproximadamente) hasta que la DO600 era de 1,0 aproximadamente. En este punto, se agregó galactosa hasta una concentración final de 2 % (p/v). Los cultivos se incubaron a 25°C con agitación por otras 48 horas antes de cosecharlos por centrifugación. Los pélets celulares se lavaron con agua antes de secarlos por congelamiento para el fraccionamiento en las clases de lípidos y el análisis de cuantificación según se describió en el Ejemplo 1. El promotor Gal se expresa de manera inducible en las células de levadura transformadas, incrementando el contenido de aceite en las células.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células de levadura, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en levadura. Las células transgénicas resultantes tienen un contenido de aceite aumentado. Los genes también se introducen en

la levadura oleaginosa, *Yarrowia lipolytica*, para aumentar el contenido de aceite.

Expresión de MGAT en células del alga *Chlamydomonas reinhardtii*

5 Se usa un vector quimérico que incluye MGAT1 de *M. musculus* para transformar de forma estable células de algas. La construcción genética denominada 35S:MGAT1 se obtiene por clonación de la región de codificación de MGAT1 en un vector de clonación que contiene un casete promotor 35S del virus en mosaico de la coliflor y un gen de resistencia a paramomicina (aminoglucósido-O-fosfotransferasa VIII) expresado por un promotor RBCS2 de *C. reinhardtii*. La construcción 35S:MGAT1 se introduce por separado en un cultivo en fase logarítmica de  $5 \times 10^7$  cc503, una cepa deficiente en pared celular de *Chlamydomonas reinhardtii* mediante un método de esferas de vidrio modificadas (Kindle, 1990). Ambos vectores también contienen el gen de resistencia a BLE como gen marcador de selección. Brevemente, se hace crecer una colonia de células no transformadas sobre una placa de agar TAP mantenida a 24°C aproximadamente hasta  $5 \times 10^6$  células/ml aproximadamente sobre un período de cuatro días, las células resultantes se agrupan en pélets a 3000 g durante 3 minutos a temperatura ambiente y se resuspendieron para producir  $5 \times 10^7$  células en 300 ul de medio TAP. Se agregan 300 ul de esferas de vidrio de 0,6 mm de diámetro, 0,6 µg de plásmido en 5 µl y 100 µl de PEG MW8000 20 % y la mezcla se agita por vortexeo a velocidad máxima por 30 segundos, luego se transfieren a 10 ml de TAP y se incuban durante 16 horas con agitación en oscuridad. Las células se agrupan en pélets, se resuspenden en 200 µl de TAP, luego se plaquean sobre placas TAP que contienen zeocina 5 mg/l y se incuban en oscuridad durante 3 semanas. Las colonias transformadas se subcultivan sobre una placa fresca de TAP + zeocina 5 mg/l después de lo cual se dejan crecer bajo condiciones de medio estándar usando selección con zeocina. Después de una cosecha por centrifugación, los pélets celulares se lavan con agua antes de secarlos por congelamiento para el fraccionamiento en clases de lípidos y el análisis de cuantificación según se describió en el Ejemplo 1. El promotor 35S:MGAT1 se expresa de manera constitutiva en las células de algas transformadas. El contenido de aceite de las células está significativamente aumentado.

25 La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en la construcción mencionada previamente y se introducen en *Chlamydomonas*. El contenido de aceite en las células transgénicas resultantes está aumentado significativamente.

30 Expresión de MGAT en *Lupinus angustifolius* transformada de forma estable

Se usa un vector quimérico que incluye MGAT1 de *M. musculus* para transformar *Lupinus angustifolius*, una planta leguminosa. Se usan los vectores quiméricos 35S:MGAT1 y 35S:DGAT1 en *Agrobacterium* para transformar *L. angustifolius* como describieran Pigeaire et al. (1997). Brevemente, se co-cultivan explantes de ápices de brotes con *Agrobacterium* transgénico antes de humedecerlos exhaustivamente con una solución de PPT (2 mg/ml) y transferirlos sobre un medio de regeneración sin PPT. Se recortan los múltiples brotes axilares que se desarrollan a partir de los ápices de brotes y se colocan sobre un medio que contiene PPT 20 mg/l y los brotes sobrevivientes se transfieren sobre medio fresco que contiene PPT 20 mg/l. A continuación se transfieren los brotes saludables a tierra. El promotor 35S se expresa de manera constitutiva en células de las plantas transformadas, incrementando el contenido de aceite en los tejidos vegetativos y en las semillas. Se usa un promotor específico de semillas para incrementar aún más el contenido de aceite en semillas transgénicas de *Lupinus*.

45 La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Lupinus*. Las semillas y los tejidos vegetativos de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Expresión de MGAT en células de *Sorghum bicolor* transformadas de forma estable

50 Se usa un vector quimérico que incluye MGAT1 de *M. musculus* para transformar de forma estable *Sorghum bicolor*. Se usan Ubi:MGAT1 y Ubi:DGAT1 en la cepa AGL1 de *A. tumefaciens* para transformar *Sorghum bicolor* como describen Gurel et al. (2009). El *Agrobacterium* se centrifuga primero a 5.000 rpm a 4 °C durante 5 minutos y se diluye hasta una DO550 = 0,4 con medio de co-cultivo líquido. Los embriones inmaduros aislados previamente se cubren luego por completo con la suspensión de *Agrobacterium* durante 15 minutos y después se cultivan, con el lado del escutelo hacia arriba, sobre medio de co-cultivo en oscuridad durante 2 días a 24°C. Los embriones inmaduros son transferidos luego a medio de inducción de callos (CIM) con carbenicilina 100 mg/l para inhibir el crecimiento de *Agrobacterium* y se deja así durante 4 semanas. Los tejidos son transferidos luego a un medio de regeneración para formar brotes y raíces. El promotor Ubi se expresa de manera constitutiva en células de las plantas transformadas, incrementando el contenido de aceite por lo menos en los tejidos vegetativos.

60 La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Sorghum*. Los tejidos vegetativos de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

65

Expresión de MGAT en plantas de *Sorghum bicolor* transformadas de forma estable

Se usa un gen quimérico que codifica MGAT1 de *M. musculus* para transformar de forma estable *Glycine max*, otra legumbre que se puede usar para la producción de aceite. Se usa 35S:MGAT1 en *Agrobacterium* para transformar *G. max* como se describe en Zhang et al. (1999). El *Agrobacterium* se co-cultiva durante tres días con explantes de cotiledones derivados de plántulas de cinco días de vida. Los explantes se cultivan luego sobre medio B5 de Gamborg suplementado con BAP 1,67 mg/l y glufosinato 5,0 mg/l por cuatro semanas después de lo cual los explantes se subcultivan en un medio que contiene sales MS mayores y menores y vitaminas B5 (MS/B5) suplementado con zeatina-ribósido 1,0 mg/l, GA3 0,5 mg/l e IAA 0,1 mg/l con 1m7 mg/l o glufosinato 2,0 mg/l. Los brotes elongados forman raíces sobre un medio de enraizamiento MS/B5 suplementado con NAA 0,5 mg/l sin selección adicional con glufosinato. El promotor 35S se expresa de manera constitutiva en células de las plantas transformadas, incrementando el contenido de aceite en los tejidos vegetativos y en las semillas.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Glycine*. Los tejidos vegetativos y las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Expresión de MGAT en *Zea mays* transformada de forma estable

Se usa un gen quimérico que codifica MGAT1 de *M. musculus* para transformar de forma estable *Zea mays*. Los vectores que comprenden 35S:MGAT1 y 35S:DGAT1 se usan para transformar *Zea mays* según se describe en Gould et al. (1991). Brevemente, se co-cultivan explantes de ápices de brotes con *Agrobacterium* transgénico durante dos días antes de transferirlos sobre un medio de sales MS que contiene kanamicina y carbenicilina. Después de varias rondas de subcultivo, se forman espontáneamente brotes y raíces transformadas y se transplantan a tierra. El promotor 35S se expresa en células de las plantas transformadas, incrementando el contenido de aceite en los tejidos vegetativos y en las semillas.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Zea mays*. Los tejidos vegetativos y las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado. Como alternativa, las regiones de codificación MGAT se expresan bajo el control de un promotor específico de endosperma, tal como el promotor de zeína o un promotor específico de embriones obtenido a partir de una planta monocotiledónea, para una mayor expresión y un mayor contenido de aceite en las semillas. Se introduce otro gen quimérico que codifica una GPAT con actividad fosfatasa, tal como GPAT4 o GPAT6 de *A. thaliana*, en *Zea mays* en combinación con la MGAT, incrementando aún más el contenido de aceite en semillas de maíz.

Expresión de MGAT en *Elaeis guineensis* transformada de forma estable (aceite de palma)

Se usa un gen quimérico que codifica MGAT1 de *M. musculus* para transformar de forma estable *Elaeis guineensis*. Se usan los vectores quiméricos denominados Ubi:MGAT1 y Ubi:DGAT1 en *Agrobacterium*. Después de 48 horas de cultivo vigoroso, las células se usan para transformar *Elaeis guineensis* como describieron Izawati et al. (2009). El promotor Ubi se expresa de manera constitutiva en células de las plantas transformadas, incrementando el contenido de aceite por lo menos en los frutos y las semillas y se pueden usar para obtener aceite.

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Elaeis*. Las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

Expresión de MGAT en *Avena sativa* (avena) transformada de forma estable

Se usa un gen quimérico que codifica MGAT1 de *M. musculus* para transformar de forma estable *Avena sativa*, otra planta monocotiledónea. Se usan los vectores quiméricos denominados Ubi:MGAT1 y Ubi:DGAT1, como se describió previamente y que contienen, ambos, un marcador de selección Ubi:BAR, para transformar *Avena sativa* como describen Zhang et al. (1999).

La región de codificación del gen MGAT2 de ratón, de codones optimizados para la expresión en células vegetales, se sustituye por la MGAT1 en las construcciones mencionadas previamente y se introducen en *Avena*. Las semillas de las plantas transgénicas resultantes presentan un contenido de aceite aumentado.

**Ejemplo 7. Manipulación mediante ingeniería genética de una MGAT con actividad DGAT**

Se puede manipular mediante ingeniería genética una MGAT con actividad DGAT comparable mediante una mutagénesis aleatoria, una mutagénesis dirigida o una mutagénesis por saturación de el o los genes MGAT de interés o sometiendo diferentes genes MGAT y/o DGAT a entremezclado de ADN. La función de DGAT se pueden seleccionar positivamente usando, por ejemplo, una cepa de levadura que tienen un requerimiento absoluto de

complementación de síntesis de TAG cuando se le alimentan ácidos grasos libres, tal como H1246, que contiene mutaciones en cuatro genes (*DGA1*, *LRO1*, *ARE1*, *ARE2*). La transformación de las variantes de MGAT en dicha cepa y el posterior suministro a la levadura transformada de una concentración de ácidos grasos libres que previene la complementación por el gen MGAT de tipo silvestre solamente permitirá el crecimiento de variantes con mayor capacidad de síntesis de TAG debido a una actividad DGAT mejorada. La actividad MGAT de estos genes mutados se puede determinar por alimentación con *sn-1* o *sn-2* MAG marcado y cuantificación de la producción del DAG marcado. Diversas rondas de evolución dirigida en combinación con un diseño racional de proteína daría como resultado la producción de un nuevo gen MGAT con actividades MGAT y DGAT similares.

#### 10 **Ejemplo 8 Expresión constitutiva de la diacilglicerol aciltransferasa 2 de *A. thaliana* en plantas**

La expresión de la DGAT2 de *A. thaliana* en levadura (Weselake et al., 2009) y células de insecto (Lardizabal et al., 2001) no demostró actividad DGAT. De manera similar, la DGAT2 no pudo complementar un noqueo de la DGAT1 de *A. thaliana* (Weselake et al., 2009). La actividad enzimática de la DGAT2 de *A. thaliana* en tejido foliar se determinó usando un sistema de expresión transitorio de *N. benthamiana* según se describió en el Ejemplo 1. La DGAT2 de *A. thaliana* (Acceso Q9ASU1) se obtuvo por PCR genómico y se clonó en un vector de expresión binario bajo el control del promotor 35S para generar 35S:DGAT2. Este vector quimérico se introdujo en la cepa AGL1 de *A. tumefaciens* y las células de cultivos de la misma se infiltraron en tejido foliar de plantas de *N. benthamiana* en un cuarto de crecimiento a 24 °C usando 35S:DGAT1 como control. Se infiltraron diversas comparaciones directas con las muestras en comparación ubicadas a cada lado de la misma hoja. Los experimentos se condujeron por triplicado. Después de la infiltración, las plantas se cultivaron por otros cinco días antes de retirar discos foliares y secarlos por congelamiento para el fraccionamiento en clases de lípidos y el análisis de cuantificación según se describió en el Ejemplo 1. Este análisis reveló que ambas DGAT1 y DGAT2 aumentaban los niveles de aceite en las hojas de *Nicotiana benthamiana* (Tabla 6).

El tejido foliar transformado con la construcción 35S:p19 (control negativo) contenía un promedio de 34,9 µg de ácidos grasos libres derivados de TAG/100 mg de peso seco de hojas. El tejido foliar transformado con las construcciones 35S:p19 y 35S:DGAT1 (control positivo) contenía un promedio de 126,7 µg de ácidos grasos libres derivados de TAG/100 mg de peso seco de hojas. El tejido foliar transformado con las construcciones 35S:p19 y 35S:DGAT2 contenía un promedio de 310,0 µg de ácidos grasos libres derivados de TAG/100 mg de peso seco de hojas.

Los datos descritos previamente demuestran que la enzima DGAT2 de *A. thaliana* es más activa que la enzima DGAT1 *A. thaliana* en promover la acumulación de TAG en tejido foliar. La expresión del gen de DGAT2 dio como resultado un 245 % de acumulación de TAG en el tejido foliar en comparación con la expresión de DGAT1.

#### 35 **Ejemplo 9. Co-expresión de MGAT y GPAT en semillas transgénicas**

Yang et al., (2010) describieron dos glicerol-3-fosfato aciltransferasas (GPAT4 y GPAT6) de *A. thaliana*, ambas con preferencia por *sn-2* (es decir, forman preferencialmente *sn-2* MAG en lugar de *sn-1/3* MAG) y actividad fosfatasa, que podían producir *sn-2* MAG a partir de G-3-P (Figura 1). Se propuso que estas enzimas formaran parte de la vía de síntesis de cutina. GPAT4 y GPAT6 no presentaron una gran expresión en tejido de semillas. La combinación de dicha GPAT/fosfatasa bifuncional con una MGAT permite obtener una nueva vía de síntesis de DAG usando G-3-P como sustrato que puede reemplazar o suplementar la típica vía de Kennedy para la síntesis de DAG en plantas, en particular en el aceite de semillas, u otras células, lo que da como resultado un mayor contenido de aceite, en particular de niveles de TAG.

Se obtuvieron los ADN quiméricos denominados pJP3382 y pJP3383, que codifican las GPAT4 y GPAT6 de *A. thaliana*, respectivamente, junto con la MGAT2 de *M. musculus* para su expresión en semillas insertando primero la región de codificación completa de MGAT2, contenida en un fragmento *Swal*, en el pJP3362 en el sitio *SmaI* para obtener el pJP3378. El plásmido pJP3362 era un vector de expresión binario que contiene los casetes de expresión vacíos FAE1 y FP1 y un gen de resistencia a kanamicina como marcador de selección. La GPAT4 de *A. thaliana* se amplificó a partir del ADNc y se clonó en el pJP3378 en el sitio *NotI* para obtener el pJP3382 en el cual se expresaba la GPAT4 por el promotor de napina truncado, FP1 y la MGAT2 era expresada por el promotor FAE1 de *A. thaliana*. De manera similar, se amplificó GPAT6 de *A. thaliana* a partir del ADNc y se clonó en el pJP3378 en el sitio *NotI* para obtener el pJP3384 en donde la GPAT6 estaba ligada operativamente al promotor de napina truncado, FP1 y la MGAT2 era expresada por el promotor FAE1 de *A. thaliana*. Los plásmidos pJP3382 y pJP3383 fueron transformados en *A. thaliana* (ecotipo Columbia) mediante el método de inmersión floral. Se plaquearon semillas de las plantas tratadas sobre un medio que contiene al antibiótico, kanamicina, para seleccionar las plantas de progenie (plantas T1) que fueron transformadas. Las plántulas transgénicas fueron transferidas a tierra y se cultivaron en invernadero. Se determinó la expresión de los transgenes en los embriones en desarrollo. Se seleccionan las plantas transgénicas con el mayor nivel de expresión y que muestran una relación 3:1 de plantas transgénicas : no transgénicas por línea, indicativa de una inserción de un solo locus de los transgenes y se dejan crecer hasta alcanzar la madurez. Se obtienen semillas a partir de estas plantas (T2) que incluyen algunas que son homocigotas para los transgenes. Se cultivaron cincuenta plantas de progenie (plantas T2) de cada línea en tierra y se determina el contenido de lípidos, el contenido de TAG y las composiciones de ácidos grasos de la semilla

resultante. El contenido de lípidos neutros, en particular el contenido de TAG, de las semillas que comprenden tanto una MGAT como una GPAT4 o GPAT6 está sustancialmente aumentado en comparación con los controles y con las semillas que comprenden solamente MGAT o solamente la DGAT1 de *A. thaliana*. También está aumentado el contenido de MAG y de DAG. La composición de ácidos grasos del lípido extraído de las semillas está modificada, en particular el contenido de ácidos grasos poliinsaturados, tal como ALA, está significativamente aumentado.

**Ejemplo 10. Evaluación del efecto de GPAT4 y GPAT6 sobre el incremento de TAG mediado por MGAT mediante silenciamiento y mutación de GPAT**

10 La familia GPAT es grande y todos los miembros conocidos contienen dos dominios conservados, un dominio plsC aciltransferasa y un dominio de la superfamilia de hidrolasas tipo HAD. Además de esto, todas las GPAT4-8 de *A. thaliana* contienen una región N-terminal homóloga de un dominio fosfoserina fosfatasa. Ambas GPAT4 y GPAT6 de *A. thaliana* contienen restos conservados conocidos por ser críticos para la actividad fosfatasa (Yang et al., 2010).

15 Se diseñaron cebadores degenerados basados en la secuencia de aminoácidos conservada GDLVICPEGTTTCREP (SEQ ID NO:228) para amplificar fragmentos en las GPAT de *N. benthamiana* expresados en tejido foliar. 3' Se conducirá un 3' RACE usando estos cebadores y cebadores de oligo-dT inversos con ARN aislado del tejido foliar de *N. benthamiana*. Las GPAT con actividad fosfatasa (es decir, tipo GPAT4/6) se identificarán por su homología con la región del dominio de fosfoserina fosfatasa N-terminal descrita previamente. Se van a generar construcciones de  
20 ARNi dirigidas por 35S para estos genes blanco y serán transformadas en la cepa AGL1 de *A. tumefaciens*. De manera similar, se va a transformar una construcción 35S:V2 que contiene a la proteína supresora de silenciamiento viral V2 en la cepa AGL1 de *A. tumefaciens*. Se sabe que V2 suprime el mecanismo de silenciamiento de la planta nativa para permitir una expresión transitoria efectiva, pero también permitir que funcione el silenciamiento de genes basado en ARNi.

25 Luego se comparará la acumulación de TAG entre muestras de hojas transformadas transitoriamente infiltradas con las siguientes mezclas de cepas: 1) 35S:V2 (control negativo); 2) 35S:V2 + 35S:MGAT2 (control positivo); 3) 35S:V2 + GPAT-RNAi; 4) 35S:V2 + GPAT-RNAi + 35S:MGAT2. Se espera que la mezcla 35S:V2 + GPAT-RNAi + 35S:MGAT2 dará como resultado menos acumulación de TAG que 35S:V2 + muestra 35S:MGAT2 debido a la  
30 interrupción de la síntesis de *sn-2* MAG debida al silenciamiento de GPAT.

Se conducirá un experimento similar usando secuencias tipo GPAT4/6 de *A. thaliana* y *N. benthamiana* que están mutadas para eliminar los restos conservados conocidos por ser críticos para la actividad fosfatasa (Yang et al., 2010). Estos genes mutados (conocidos conjuntamente como GPAT4/6-delta) se clonarán luego en vectores de expresión binarios dirigidos por 35S y transformados en *A. tumefaciens*. Luego se comparará la acumulación de TAG entre muestras de hojas transformadas transitoriamente infiltradas con las siguientes mezclas de cepas: 1) 35S:p19 (control negativo); 2) 35S:p19 + 35S:MGAT2 (control positivo); 3) 35S:p19 + GPAT4/6-delta; 4) 35S:p19+ GPAT4/6-delta + 35S:MGAT2. Se espera que la mezcla 35S:p19 + GPAT4/6-delta + 35S:MGAT2 dará como resultado menos acumulación de TAG que 35S:p19 + muestra 35S:MGAT2 debido a la interrupción de la síntesis de  
35 *sn-2* MAG debida a la mutación de GPAT. Si bien los genes tipo GPAT4/6 nativos de *N. benthamiana* estarán presentes en este experimento, se espera que el alto nivel de expresión de las construcciones GPAT4/6-delta les dotará de una mayor capacidad para acceder al sustrato G-3-P que los genes endógenos.

**Tabla 6.** Perfil de ácidos grasos y cuantificación de TAG en tejido foliar triplicado de *Nicotiana benthamiana* transformado transitoriamente con las construcciones 35S:p19, 35S:DGAT1 y 35S:DGAT2.

Muestra	C16:0	16:1w13t	C16:1d7	16:3w3	C18:0	C18:1	C18:1d11	C18:2	C18:3	C20:0	20:1d11	20:2	20:3n3	C22:0	C24:0	µg/100 mg de PS
<b>P19</b>	44.7	0.1	0.0	0.0	33.9	1.2	0.0	6.5	12.7	0.9	0.0	0.0	0.0	0.0	0.0	43.29
	44.1	1.7	0.0	0.0	15.3	2.0	0.0	15.2	19.5	2.2	0.0	0.0	0.0	0.0	0.0	23.12
	43.3	1.5	0.0	0.0	10.5	1.5	0.0	17.2	23.9	2.2	0.0	0.0	0.0	0.0	0.0	38.35
<b>P19+AtDGAT1</b>	36.3	0.5	0.1	0.4	11.6	2.3	0.3	17.8	24.5	3.6	0.0	0.0	0.2	1.5	0.2	144.77
	33.6	0.5	0.1	0.4	11.2	2.9	0.3	23.1	21.5	3.8	0.0	0.0	0.2	1.5	0.9	145.34
	36.8	0.5	0.0	0.0	12.4	2.9	0.4	21.3	19.3	3.9	0.0	0.0	0.0	1.5	1.0	90.04
<b>P19+AtDGAT2</b>	18.6	0.3	0.1	0.5	9.3	7.7	0.4	28.0	33.1	1.1	0.2	0.1	0.1	0.2	0.3	439.25
	17.5	0.3	0.1	0.3	10.2	9.9	0.5	32.7	26.5	1.2	0.1	0.0	0.1	0.2	0.4	282.50
	18.4	0.3	0.1	0.3	9.8	7.5	0.5	32.3	29.1	1.2	0.0	0.0	0.0	0.3	0.2	208.40



Los expertos en la técnica apreciarán que pueden realizarse numerosas variaciones y/o modificaciones a la invención, tal como se ilustran en las realizaciones específicas, sin apartarse del espíritu o alcance de la invención, tal como ha sido descripta ampliamente. Las realizaciones de la presente deben, por lo tanto, considerarse en todo sentido a modo ilustrativo y no restrictivo.

5 Toda descripción de documentos, actas, materiales, dispositivos, artículos o semejantes que se incluyera en el presente documento memoria descriptiva solamente sirve a los efectos de proporcionar un contexto para la presente invención. No debe considerarse como un reconocimiento que cualquiera o todos estos temas formen parte del sustento de la técnica anterior o que fueran de conocimiento general en el campo relevante de la presente invención  
10 como existía antes de la fecha de prioridad de cada reivindicación de esta solicitud.

## REFERENCIAS

- 15 Abdullah et al. (1986) *Biotech.* 4:1087.  
 Alvarez et al. (2000) *Theor Appl Genet* 100:319-327.  
 Al-Mariri et al. (2002) *Infect. Immun.* 70:1915-1923.  
 Bao and Ohlrogge (1999) *Plant Physiol.* 120:1057-62.  
 Bartlett et al. (2008) *Plant Methods* 4:22.  
 20 Baumlein et al. (1991) *Mol. Gen. Genet.* 225: 459-467.  
 Baumlein et al. (1992) *Plant J.* 2:233-239.  
 Benghezal et al. (2007) *Journal of Biological Chemistry* 282:30845-30855.  
 Bligh and Dyer (1959) *Canadian Journal of Biochemistry and Physiology* 37:911-7.  
 Bouvier-Nave et al. (2000) *European Journal of Biochemistry / FEBS* 267:85-96.  
 Bradford (1976) *Anal. Biochem.* 72:248.  
 25 Broothaerts et al. (2005) *Nature* 433:629-633.  
 Broun et al. (1998) *Plant J.* 13:201-210.  
 Cadwell and Joyce (1992) *PCR Methods Appl.* 2:28-33.  
 Cao et al. (2003) *J. Biol. Chem.* 278:13860-13866.  
 Cao et al. (2007) *J. Lipid Res.* 48:583.  
 30 Cheng et al. (1996) *Plant Cell Rep.* 15:653-657.  
 Cheng et al. (2003) *J. Biol. Chem.* 278, 13611-13614.  
 Chikwamba et al. (2003) *Proc. Natl. Acad. Sci. U.S.A.* 100:11127-11132.  
 Chung et al. (2006) *BMC Genomics* 7:120.  
 Clough and Bent (1998) *Plant J.* 16:735-743.  
 35 Coco et al. (2001) *Nature Biotechnology* 19:354-359.  
 Coco et al. (2002) *Nature Biotechnology* 20:1246-1250.  
 Comai et al. (2004) *Plant J* 37: 778-786.  
 Courvalin et al. (1995) *Life Sci.* 318:1209-1212.  
 Coutu et al. (2007) *Transgenic Res.* 16:771-781.  
 40 Cramer et al. (1998) *Nature* 391:288-291.  
 Darji et al. (1997) *Cell* 91:765-775.  
 Deshpande and Mukund (1992) *Appl. Biochem. Biotechnol.*, 36:227.  
 Dietrich et al. (1998) *Nature Biotech.* 18:181-185.  
 Eggert et al. (2005) *Chembiochem* 6:1062-1067.  
 45 Ellerstrom et al. (1996) *Plant Mol. Biol.* 32:1019-1027.  
 Fennelly et al. (1999) *J. Immunol.* 162:1603-1610.  
 Fujimura et al. (1985) *Plant Tissue Culture Lett.* 2:74.  
 Ghosal et al. (2007) *Biochimica et Biophysica Acta* 1771:1457-1463.  
 Ghosh et al. (2006) *International Symposium on Plant Lipids, Abstract.*  
 50 Glevin et al. (2003) *Microbiol. Mol. Biol. Rev.* 67:16-37.  
 Grant et al. (1995) *Plant Cell Rep.* 15:254-258.  
 Grillot-Courvalin et al. (1998) *Nature Biotech.* 16:862-866.  
 Grillot-Courvalin (1999) *Curr. Opin. Biotech.* 10:477-481.  
 Gould et al. (1991) *Plant Physiol.* 95:426-434.  
 55 Gurel et al. (2009) *Plant Cell Rep.* 28:429-444.  
 Harayama (1998) *Trends Biotechnol.* 16: 76-82.  
 Hellinga (1997) *Proc. Natl. Acad. Sci.* 94(19):10015-10017.  
 Henikoff et al. (2004) *Plant Physiol* 135: 630-636.  
 Hense et al. (2001) *Cell Microbiol.* 3:599-609.  
 60 Hinchee et al. (1988) *Biotechnology* 6:915-922.  
 Hirayama and Hujii K (1965) *Agricultural and Biological Chemistry* 29:1-6.  
 Horvath et al. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97:1914-1919.  
 Izawati et al. (2009) *J. Oil Palm Res.* 21:643-652.  
 Jako et al. (2001) *Plant Physiol.* 126:861-874.  
 65 Jézéquel et al. (2008) *Biotechniques* 45:523-532.  
 Lacroix et al. (2008) *Proc. Natl. Acad. Sci., U.S.A.* 105: 15429-15434.

- Leung et al. (1989) *Technique* 1:11-15.  
 Kindle (1990) *Proc. Nat. Acad. Sci. USA* 87: 1228-1232.  
 Koziel et al. (1996) *Plant Mol. Biol.* 32:393-405.  
 Kunik et al. (2001) *Proc. Natl. Acad. Sci. U.S.A.* 98:1871-1876.  
 5 U.S.A. 105: 15429-15434.  
 Lakshminarayana et al. (1984) *Journal of the American Oil Chemists Society* 61:1249-1253.  
 Lardizabal et al. (2001) *J. Biol. Chem.* 276:38862-38869.  
 Larkin et al. (1996) *Transgenic Res.* 5:325-335.  
 Murashige and Skoog (1962) *Physiologia Plantarum* 15:473-497.  
 10 Ness et al. (2002) *Nature Biotechnology* 20:1251–1255.  
 Niedz et al. (1995) *Plant Cell Reports* 14:403.  
 Ostermeier et al. (1999) *Nature Biotechnology* 17:1205–1209.  
 Ow et al. (1986) *Science* 234:856-859.  
 Panekina (1978) *Chemistry of Natural Compounds* 14:33–36.  
 15 Perez-Vich et al. (1998) *JAOCS* 75:547-555  
 Perrin et al. (2000) *Mol Breed* 6:345-352.  
 Perry and Harwood (1993) *Phytochemistry* 33:329–333.  
 Phillips et al. (2002) *Journal of Food Composition and Analysis* 12:123-142.  
 Pigeaire et al. (1997) *Mol. Breed.* 3:341-349.  
 20 Potenza et al. (2004) *In Vitro Cell Dev. Biol. Plant* 40:1-22.  
 Powell et al. (1996) *Vaccines* 183, Abstract.  
 Prasher et al. (1985) *Biochem. Biophys. Res. Commun.* 127:31-36.  
 Rajasekharan et al. (2006) *International Symposium on Plant Lipids, Abstract.*  
 Saha et al. (2006) *Plant Physiol.* 141:1533-1543.  
 25 Schaffner et al. (1980) *Proc. Natl. Acad. Sci. U.S.A.* 77:2163-2167.  
 Shiau et al. (2001) *Vaccine* 19:3947-3956.  
 Sieber et al. (2001) *Nature Biotechnology* 19:456–460.  
 Sizemore et al. (1995) *Science* 270:299-302.  
 Slade and Knauf (2005) *Transgenic Res.* 14: 109-115.  
 30 Stobart et al. (1997) *Planta* 203:58-66.  
 Stalker et al. 1988 *Science* 242: 419-423.  
 Stemmer (1994a) *Proc. Natl. Acad. Sci. USA* 91:10747–10751.  
 Stemmer (1994b) *Nature* 370(6488):389–391.  
 Thillet et al. (1988) *J. Biol. Chem* 263:12500-12508.  
 35 Tingay et al. (1997) *Plant J.* 11:1369-1376.  
 Toriyama et al. (1986) *Theor. Appl. Genet.* 205:34.  
 Tumaney et al. (2001) *J. Biol. Chem.* 276:10847-10852.  
 Tzfira and Citovsky (2006) *Curr. Opin. Biotech.* 17:147-154.  
 Weiss et al. (2003) *Int. J. Med. Microbiol.* 293:95:106.  
 40 Weselake et al. (2009) *Biotechnology Advances* 27:866–878.  
 Winans et al. (1988) *Journal of Bacteriology* 170:4047-4054.  
 Wood et al. (2009). *Plant Biotech. J.* 7: 914-924.  
 Yang et al. (2003) *Planta* 216:597-603.  
 Yang et al. (2010) *PNAS* 107:12040-12045.  
 45 Yen and Farese (2003) *J. Biol. Chem.* 278:18532-18537.  
 Yen et al. (2002) *Proc. Natl. Acad. Sci. U.S.A.* 99:8512-8517.  
 Yen et al. (2005) *J. Lipid Res.*, 46: 1502-1511.  
 Voinnet et al. (2003) *Plant J.* 33:949-956.  
 Volkov et al. (1999) *Nucleic acids research* 27(18):e18.  
 50 Zhang et al. (1999) *Plant Cell Reports* 18:959-966.  
 Zhang et al. (1999) *Plant Cell, Tissue and Organ Culture* 56:-46.  
 Zheng et al. (2003) *The Plant Cell* 15:1872-1887.  
 Zhao et al. (1998) *Nature Biotechnology* 16:258–261.

## 55 LISTADO DE SECUENCIAS

<110> Commonwealth Scientific and Industrial Research Organisation

<120> Métodos para producir lípidos

60

<130> 510797

<150> AU2010902841

<151> 28-06-2010

65

<150> US61/399286

ES 2 640 100 T3

<151> 09-07-2010  
 <150> US61/485349  
 <151> 12-05-2011  
 5 <160> 230  
 <170> PatentIn versión 3.5  
 10 <210> 1  
 <211> 1011  
 <212> ADN  
 <213> *Mus musculus*  
 15 <400> 1

```

atgatggttg agttcgctcc tcttaacact cctcttgcta gatgccttca aactgctgct      60
gttcttcaat gggttctctc tttcttgctc cttgttcaag tgtgcatcgg aatcatggtt      120
atgctcgttc tctacaacta ctggttcctc tacatccctt acctcgtttg gttctactac      180
gattggagaa ctcctgagca aggtggaaga agatggaact gggttcaatc ttggcctggt      240
tggaagtact tcaaagagta cttccctatc tgcctcgtta agactcagga tcttgatcct      300
ggacacaact acatcttogg attccatcct catggaatct tctgctcctg tgctttcgga      360
aacttctgta ctaagtacag cgacttcaag aagcttttcc ctggattcac ctcttacctc      420
catggttgcta agatctgggt ctgcttccca ctcttcagag aatacctcat gtctaacgga      480
cctgtgtctg tgtctaaaga atctctctct cacttctctc ccaaggatgg tggtggaaac      540
gtttcaatca tegtctcttg aggtgctaaa gaggctcttg aggctcatcc tggaaacttc      600
actttgtgta tcagacagcg taagggatc gttaagatgg ctcttactca cggtgcttct      660
cttgttccag ttttctcatt cggagagaac gatctotaca agcagatcaa caaccctaag      720
ggatcttggc ttagaactat ccaggatgct atgtacgact ctatgggagt tgctttgcoct      780
cttatctacg ctagaggaat cttccagcac tacttcggaa tcatgcctta cagaaagctc      840
atctacaccg ttgttggtag acctatccct gttcaacaga tccttaacce tacttctgag      900
cagatcgagg aacttcatca gacttacctt gaggaactca agaagttggt caacgagcac      960
aagggaaagt acggaatccc tgagcatgag actctcgttt tcaagtgatg a              1011
    
```

20 <210> 2  
 <211> 1008  
 <212> ADN  
 <213> *Mus musculus*  
 25 <400> 2

ES 2 640 100 T3

atggtogagt tgcctccact tottgctcct tgggagagaa gattgcagac ttctgctggt 60  
cttcagtggg tttctcatt ccttgctcct gctcaactct goactcgttat ctctgctgga 120  
ctcttggtca ccaggctctg gctttctct gttctttaag ctactcgtg gtacttggtat 180  
tgggataagc ctagacaagg tggtagacct atccaattct tcagaagatt ggccatctgg 240  
aagtacatga aggattactt ccctgtgtct ctctgtaaga ctgctgagct tgatcctagc 300  
cgtaactaca ttgctggatt ccctcctcat ggtgttcttg ctgctgggtc tttcctaac 360  
ctctgtactg agtctactgg attcacctct ttgtccctg gaatcagatc ctacctatg 420  
atgcttaacc ttggttcag ggctccattc ttcagagatt acatcatgct tgggtgactc 480  
gtgtctagcg agaaagttc tgctgatcac atccttagca gaaaggggtg tggaaacctt 540  
cttctctca tctgtggag tgctcaagag gctcttgatg ctgacctgg tgcttacaga 600  
cttctctga agaaccgtaa gggattcctc agacttgctc ttatgcatgg tgctgctctt 660  
gttctctct totcattcgg agagaacaac ctcttcaacc aggttgagaa cactcctgga 720  
acttgctta gatggatcca gaacagactc caaaagatta tgggaatcag ccttccactt 780  
ttccatgaa ggggagttt ccaatactct ttoggactca tgctttcag acagcctatc 840  
actactatcg ttgaaagcc tatcgaggtt caaatgactc ctgagcctc tagagaggaa 900  
gttgatagac ttcaccagcg ttacatcaa gagttgtgca agcttttcga ggaacacaag 960  
ctcaagtca acgttccaga ggatcagcat cttgagttct gctaatga 1008

<210> 3  
<211> 1008  
<212> ADN  
<213> *Ciona intestinalis*

5

<400> 3

atgtgggagt tgcgccctct tagaatcct atgatcagaa gattgcagac ccttgcctgtt 60  
tctcaactcg tgtctgttt cttggctcct gctgctactt gcactcttct tactatctat 120  
atcaacctct tcaccagcta ctggcctatc actactgctt acctcgtttt cttgttctc 180  
gatcgtaaca ctctgagtc tgggtgtaga agatctgagt gggtagaaa ctggtctttg 240  
tggaaagtga tgtctgatta ctcccttgt acccttcaaca agaccgttga tcttgatcct 300  
aagaggaact acatcttcgg aatccatcct catggtgttc tctgtatcgg atctttcact 360  
cacttctcta ccaacggatc tggattctct catgttttcc ctggatcacc ctctcactt 420  
actatgcttc cattctggtt caagatgcct ttcttcaggg attacgtgat gtctggtgga 480

10

ES 2 640 100 T3

cttactcctg ctactaggaa ggctatcaag cacactatca cttagacctg tggtggacat 540  
 atctgttgta ttatccctgg tggtgctcct gagtctotta acgctagacc aggtgatggt 600  
 gttcttttgc ttaagcagag actcggatc ctttaagcttg ctatcactaa cgggtttcct 660  
 cttgttcctg tgttctcttt cgggtgatcat gctccttggg agcaaaagcc taaccctcct 720  
 ggatctotta tcagaagatt ccaggactct agccaaaagt ggatgcaagt tgctctccct 780  
 gttttccatg cttagaggaat cttccagtac aacttcggac tcatccctta cagaagatct 840  
 gttcacactg ttgttgagaga gcctatcgaa gttcctcaaa actctaacc tacctctgag 900  
 gatctcatgt ctcttcaaga ggactacatc aacagactca gggctatott cgatgagcac 960  
 aagtctaagt acctccctga ggattgcaag ctcatcatca actgatga 1008

5 <210> 4  
 <211> 1020  
 <212> ADN  
 <213> *Tribolium castaneum*

<400> 4

atgaagatcc tcggaatcaa gttcgcctcct cttcatatcc ctcttgagag aagattgcag 60  
 actcttgctg ctggatgttg gttcactact cttgctttcg gaactttcat cggaaccttc 120  
 atctggggtt acgtgttctt caccaggttc tggtacttgt ctggtcttta cgcaccatc 180  
 atctacctog agaagtctaa gtgtgagaag ggtggttagac ctatcgagtg gatcagacat 240  
 tggggatggt ggtactacct caagaactac ttcccttgca agctcgattt cgttcctgga 300  
 cttactttog atcetaagag gaactacctc ttgccttgtt accctcatgg aatctcctc 360  
 gctggacctt tcaaacctat cggatctcca tactctgagt tcagcaagct tttcccaaag 420  
 ttcagagtga gactcgttat ccttcaccag catttcttca tccattcct tagagagatc 480  
 gcttacggaa ctggtggaat ctctgcttct gctaagtctc ttaaccacgt tctcggatct 540  
 ccagagggag gttacattgc tgttcttatg cctgggtggtg ctggtgagggc ttacaactct 600  
 agacctggac agtacagaat catcctcaag aacagaaagg gattcgtaa gctcgcctc 660  
 agaaacggat ctctcttgt tctgttatc tctttcggag agcctgagct tttcgatcaa 720  
 gttgagggaa gaacctcag aaagatccaa gagtctatca gaaagtacct tggactcgt 780  
 cctgttatct tctctggaag aggattcttc cagtactctt tcggagttat ccctcaaaga 840  
 aggocctatca ctactgttgt tggacacct atogaggta caaagatcga gaagcctacc 900  
 aacgaggaag ttgacgagct tcacaaaaag tacatgcaag agcttgagaa cttgttcgaa 960  
 10 gagtacaagt tcaagtacct cgagaacct aaggatatcc atcttgagtt cgagtgatga 1020

15 <210> 5  
 <211> 867  
 <212> ADN  
 <213> *Danio rerio*

<400> 5

ES 2 640 100 T3

atgctcggag ccctttacgc tggatggctt taccttgata gagatacacc ttcttgcgga 60  
 ggtagaagat ctcaatgggt tagatcttgg aggatctgga tgcacttcag agattacttc 120  
 cctatcaccc tcgttaagac cgttgatctt gatcctaggc acaactacct tctcggattc 180  
 catcctcatg gtgttcttgt tgetgggga ttcgaaaact tctgtactga ggcttctgga 240  
 ttctctcaaa tgttccttgg acttaccctt tacctcttta tgettccatt ctgggttcagg 300  
 gtgccattct tccgtgagta catcatgtgc ggaggacttg tttcttctga gaaggcttct 360  
 gcttcttacc ttcttggaca tccctgggtg ggacaagctg ctgttatcgc tgttggaggt 420  
 gctcctgagt ctcttgaggc tagacctggt gctcttactc ttcaacttct tcagaggaag 480  
 ggattcatca agcttgctct taagcacggt gcttggcttg ttccagtttt ctcatcggga 540  
 gagaacgagc ttttcgacca aatggaaaac cctgctggat ctgctctcag aagaatgcaa 600  
 gaaagactcc agaggattat gggagttgct ctcccacttt tccatgctag gggagttttc 660  
 caatactctt tcggacttct ccacatacaga aagcctatcc atacogttgt tggtagacct 720  
 atccctgttt ctcaaactcc ttgcccttct aaagaggata tcgacgctct ccacactctc 780  
 tacatgcaag gacttaccca ggtgttcgag gaaaacaaga agcactacgg aatogctgat 840  
 gataagcacc ttaagttcac ctgatga 867

5 <210> 6  
 <211> 762  
 <212> ADN  
 <213> *Danio rerio*  
 <400> 6

atgagggact acttccctat cagacttacc aagaccgctg atctcgacac cagaagaac 60  
 tacgttatgg gattccatcc tcacggaatc cttgttgctg gtgcttttac taacttctgc 120  
 actgaggcta ccggattctc taagtgttc cctggaatca agagcaacct tcttatgctt 180  
 cctctttggt tccgtgctcc attcttcaga gactacatca tgtctgctgg actcgttctc 240  
 tctgacaaag agtctgcttc ttacctcctt agaagaaag gtgggtgaaa cgctgttgtt 300  
 atogctgttg gaggtgctcc tgaggctctt gatgctcatc ctgggtgatta cactgttcac 360  
 cttgctaaca agaagggatt catcaagctt gctatcgagc atggtgctga tcttgttcct 420  
 atctactctt tcggagagaa cgaggttttc gatcaagttc aaaaccctag aggaacctgg 480  
 cttcgttaca tccaagaaag actccagagg attatgggag tttctcttc acttttccac 540  
 gctaggggag ttttcagta cactttcgga cttatgcctt accgtaagcc tatcaacacc 600  
 gttgttgga gacctatccc tgttgagaag aacgagaagc ctactgctga ggaacttgat 660  
 gtttaccatc agcgttacat ggatgagctt gctagacttt tcgaggatca caagggaaac 720

tacgggtgttc ctgaggatac tcatctcgtt ttccagtgat ga 762

10 <210> 7  
 <211> 1008

ES 2 640 100 T3

<212> ADN  
<213> *Mus musculus*

<400> 7

5

```

atgatggtcg agttcgggcc actcaacacc ccgctggcac ggtgcctaca gaccgctgcg      60
gtgctgcagt gggctcctgtc cttcctcctg ctctgtcagg tgtgcattgg aattatggtg     120
atgctgggtcc tgtacaacta ttggttcctt tacatcccat atctggtctg gttttactat     180
gactggagaa ccccagagca aggaggcaga agatggaact gggtccaaag ctggcctgtg     240
tggaagtatt ttaaggagta tttccaatc tgtcttgtca aaacgcagga tttgatccg      300
ggtcacaatt atatatttgg gtttcacctt catggaatat tctgtcctgg agcctttgga     360
aatttttgta caaaatactc ggacttcaag aagctatttc ctggctttac atcgtatctc     420
cacgtggcca agatctggtt ctgtttcccg ttgttccgag aatatctgat gagtaacggg     480
ccggtttcag tgtctaagga gagtttgtct catgtgctga gcaaggatgg aggtggcaat     540
gtctcaatca ttgtcctcgg agtgcaaaag gaggcctgg aggtcacc caggacattc     600
accctgtgca tccgccagcg caaagggttt gttaagatgg ccttgacca tgggtgccagt     660
ttggttccag tattttcttt tggtgaaaat gatctatata agcaaattaa caaccccaaa     720
ggctcctggc tacgaactat acaagacgca atgtatgatt caatgggagt agccttgcca     780
ctgatatatg ccagaggaat tttccagcac tactttggca taatgcccta tcggaagctg     840
atctacactg ttgttggccg ccctatccct gttcagcaga ttctgaacc caccctcagag     900
cagattgaag agctgcatca gacataccta gaggagctaa agaaactatt caatgaacac     960
aaagggaaat atgggattcc ggagcacgaa actctggtat ttaaataa                    1008

```

<210> 8  
<211> 1093  
<212> ADN  
<213> *Homo sapiens*

10

<400> 8

```

cgtgggtgca ggctgcagtg gctggcgccg tectcgcccg gccaggccat gaaggtagag      60
tttgcaaccgc tcaacatcca gctggcgccg cggctgcaga cgggtggccgt gctgcagtgg     120
gtcctttott ttcttacagg gccgatgtcc attggaatca ctgtgatgct gatcatacac     180
aactatttgt tcctttacat cccttatttg atgtggcttt actttgactg gcatacccca     240
gagcgaggag gcaggagatc cagctggatc aaaaattgga ctctttggaa acactttaag     300
gactattttc caattcatct tatcaaaact caagatttgg atccaagtca caactatata     360

```

15

ES 2 640 100 T3

```

tttgggtttc acccccatgg aataatggca gttggagcct ttgggaattt ttctgtaa420
tattctgact tcaaggacct gtttctggc tttacttcat atcttcacgt gctgccactt480
tggttctggt gtcctgtctt tcgagaatat gtgatgagtg ttgggctggt ttcagtttcc540
aagaaaagtg tgtcctacat ggtaagcaag gagggagggtg gaaacatctc tgtcattgtc600
cttgggggtg caaaagaatc actggatgct catcctggaa agttcactct gttcatccgc660
cagcggaaag gatttggtta aattgctttg acccatggcg cctctctggt occagtggtt720
tcttttgggtg aaaatgaact gtttaacaa actgacaacc ctgaaggatc atggattaga780
actgttcaga ataaactgca gaagatcatg gggtttgctt tgcccctggt tcatgccagg840
ggagtttttc agtacaattt tggcctaatag acctatagga aagccatcca cactgttggt900
ggccgoccca tcctgttctg tcagactctg aaccogacc aggagcagat tgaggagtta960
catcagacct atatggagga acttaggaaa ttgtttgagg aacacaaagg aaagtatggc1020
attocagagc acgagactct tgttttaaaa tgacttgact ataaaaaaaa attaaaaaat1080
aaaaataaat gac1093

```

<210> 9  
<211> 1093  
<212> ADN  
<213> *Pan troglodytes*

5

<400> 9



ES 2 640 100 T3

```

cgtgggtgca ggctgcagtg gctggcgccg tcctcgcccg gccaggccat gaaggtagag      60
tttgcaccgc tcaacatcca gccggcgccg cggctgcaga cggtggccgt gctgcagtg      120
gtcctttcct ttcttacagg gccgatgtcc attggaatca ctgtgatgct gatcatacac      180
agctattcgt tcctttacat cccttatttg atgtggcttt actttgactg gcatacccca      240
gagcgaggag gcaggagatc cagctggatc aaaaattgga ctctttggaa acactttaag      300
gactattttc caattcatct catcaaaact caagatttgg atccaagtca caactatata      360
tttgggtttc acccccatgg aataatggca gttggagcct ttgggaattt ttctgtaaat      420
tattctgact tcaaggacct gtttcctggc tttacttcat atcttcacgt gctgccactt      480
tggttctggt gtccctgtct tcgagaatat gtgctgagtg ttggcctggt ttcagtttcc      540
aagaaaagtg tgtcctacat ggtaagcaag gagggaggtg gaaacatctc tgtcattgtc      600
cttgggggtg caaaagaatc actggatgct catcctggaa agttcactct gttcatccac      660
cagcgaaag gatttgtaa aactgctttg acccatggcg cctctctggt cccagtggtt      720
tcttttggtg aaaatgaact gtttaacaa actgacaacc ctgaaggatc atggattaga      780
actgttcaga ataaactaca gaagatcatg gggtttgctt tgcccctgtt tcatgccagg      840
ggagtttttc agtacaattt tggcctaatag ccctatagga aagccatcca cactgttgtt      900

ggccgcccga tcctgttctg tcagactctg aaccggaccg aggagcagat tgaggagtta      960
catcagacct atatggagga acttaggaaa ttgtttgagg agcacaagg aaagtatggc     1020
attccagagc acgagactct tgttttaaaa tgacttgact ataaaaaaaa atttttaaat     1080
aaaaataaat gac                                                                1093

```

<210> 10  
 <211> 1008  
 <212> ADN  
 <213> *Pan troglodytes*  
  
 <400> 10

5

ES 2 640 100 T3

```

atgaaggtag agtttgacc gctcaacatc cagccggcgc ggcggctgca gacgggtggc      60
gtgctgcagt gggctctgaa atacctgctg ctogggccga tgtccattgg aatcactgtg      120
atgctgatca tacacagcta ttcgttcctt tacatccctt attgatgtg gctttacttt      180
gactggcata cccagagcgc aggaggcagg agatccagct ggatcaaaaa ttggactett      240
tggaaacact ttaaggacta tttccaatt catctcatca aaactcaaga tttggatcca      300
agtcacaact atatatttgg gtttcacccc catggaataa tggcagttgg agcctttggg      360
aatttttctg taaattattc tgacttcaag gacctgttcc ctggctttac ttcatatctt      420
cacgtgctgc cactttgggt ctggtgtcct gtctttogag aatatgtgct gagtgttggc      480
ctggtttcag tttccaagaa aagtgtgtcc tacatggtaa gcaaggaggg aggtggaaac      540
atctctgtca ttgtccttgg gggtgcaaaa gaatcactgg atgctcatcc tggaaagtcc      600
actctgttca tccaccagcg gaaaggattt gttaaaactg ctttgacca tgggcctctc      660
ctggtcccag tggtttcttt tggtgaaaat gaactgttta aacaaactga caaccctgaa      720
ggatcatgga ttagaactgt tcagaataaa ctacagaaga tcatgggggt tgctttgccc      780
ctgtttcatg ccaggggagt ttttcagtac aattttggcc taatgcccta taggaaagcc      840
atccacactg ttgttggccg ccogatccct gttcgtcaga ctctgaacct gaccagagg      900
cagattgagg agttacatca gacctatatg gaggaactta ggaaattggt tgaggagcac      960
aaaggaaagt atggcattcc agagcacgag actcttgttt taaaatga      1008

```

<210> 11  
 <211> 1008  
 <212> ADN  
 <213> *Canis familiaris*

5

<400> 11

```

atgaaggtcg agtttgccgc gctccacatc ccgctggcgc ggcggctgca gacggccgcg      60
gtgctgcagt gggctctgtc cttcctgctg ctccgcgagg tgtgcacgg aatcatcgtg      120
gtactgatcg tgcacaacta ttggttcctc tacgtccctt atctgacatg gctttgcctt      180
gactggcggc cccagagca gggaggcagg agatccaact gggtcagaag ctggaccggt      240

```

10

ES 2 640 100 T3

tggaggtatt ttaaggacta tttccaatt cacctcatca aaacttggga tttgatcca	300
agtcacaact atatatggg gtttcacccc cacggagtgc ttgttgctgg agcctttgga	360
aatTTTTgta cgaatcattc agacttcgag gagctgttcc ctggcttcac tgcgtatctt	420
cacgtgctcc cattttgggt cgggtgtcct ctcttcgag aatatctgat gactagcggg	480
tcagtctcag tttctaagaa aagtgtgtcc tatgtgtgga gcaaggagg agtggaaac	540
atttcagtca ttgttcttgg ggggtcagaa gaatcactgg atgccatcc tggaaaatc	600
actctgttca tccgccagcg gaaaggattt gttaaagtgg ctttgacca tgggcctcc	660
ttgtcccgg tgtttcttt tggagaaaat gaactattta agcaagtaa caaccctgaa	720
ggctcatggc ttcgaactgt gcaggagaag ctacagaaga tcatgggatt tgctttgct	780
ctgttcatg ccagagggat tttcaatac aatTTTggcc tgatgccta taggaaacct	840
atccacactg tggttggtcg ccaatcctt gttcatcggc ctccgcacc aagcccggga	900
cagattgagg agttgcatca gacctatag gaggagctaa ggaaattatt tgaagcacac	960
aaaaggaagt acggtattcc cgagcatgaa actctcattt ttaaataa	1008

<210> 12  
 <211> 1118  
 <212> ADN  
 <213> *Bos taurus*  
  
 <400> 12

5

ES 2 640 100 T3

```

atgaaggtcg agtttgcgcc actcaacatc ccgctggcgc gggggctgca gacggctgcg      60
gtgctgcatt ggttgctgtc cttcctgctg ttgcacagc tatgcctcgg aattatcgtg      120
ttcctgatca tatacaacta ctggttcctc taccttcctt acttgacatg gctttacttt      180
gactggcaaa cccagagca aggaggcaga agatctgaat gggtcagaaa ctgggccatt      240
tggaggtact ttaaggacta tttccaatt catctcatca aaacctggga tttggatccg      300
agtcacaact acatatttgg gtttcacccc catggagtgc ttgtggttgg agcctttgga      360
aacttctgta caaattattc ggccttcaag gagctgttcc ccggctttac ctcctatctt      420
cacgtgctgc cgtattggtt ccggtgtccg ctcttcgggg aatatctgat gacgagtggg      480
ccagtctcag tttccaagaa aagtgtgtgc cacgtgttaa gcaaagaggg aggtggaaac      540
atctcagtca ttgtgctcgg ggtgacagag gaatcactgg atgccatcc tggaaagtcc      600
actctgttca tccgccagcg gaagggattc gtgaaaattg ctttgacca tgggtgcttat      660
ttggtgccag tgttttcttt tggtgaaaat gaactattta aacaagttag caatcccgaa      720
ggatcatggc ttcgaaatgt gcaggagaaa ctacagaaga tcatgggatt tgctttgcca      780
ctgttocacg ccagaggaat ttttcaatac aattttggcc taatccccta taggaagccc      840
attcacactg ttgttgccg cccaatccct gttcgtcaga cctgaaccc aacctcagag      900

cagattgagg agttgcatca gacctacatg gaggagctga gaaaattatt tgaagaacac      960
aaaggaagt atggtattcc agaaaatgaa actctcattt ttagataact gttgtcttag     1020
aagtcatgca agagagcgag agaacatccc actgaaagaa aagatattct aaataaattt     1080
ttttctttga caatggaatc aaatttgcaa ctgtaaaa                                1118

```

<210> 13  
 <211> 749  
 <212> ADN  
 <213> *Rattus norvegicus*

5

<400> 13

ES 2 640 100 T3

ttgactcact aagggccagt ttcagtttct aagaggagtt tgtctcatgt gatgagcaag 60  
 gaaggagggtg ggaacatctc agtcattgtc atcgggggtg cgaaagagct actggagtcc 120  
 ttcoctggaa gatattctct gtgccttctt cagcgcgaaag gatttgtaa gatcgccttg 180  
 acccatgggtg cccatttgggt tocagtgttt tcttttgggtg aaaatgaact atatagccaa 240  
 gtcgacaacc ctaaaggctc atggcttcgg actgcccaag acaaagtga taacctaacg 300  
 gggttggcct tgccactatt ctatgccaga ggaattttcc agaacagctt tggcctaatag 360  
 ccctatcgga aattgatcta cactgttgct gcggttgca cagcagctcg tttgactcag 420  
 aagtcaaacc atcagaacga cagctcagtg tctgctgaga agggaaactc ttttatttgt 480  
 tgtcaagtgg gatggtgtga ccgagatagg cagccattgg ccgccccatc cctgttcaga 540  
 agactctgca ccgacccca gagcagattg atgagctgca tcagacgtac atggaggagc 600  
 taaggaaact cttcaatgaa cataaagga aatacggat tccggagcat gaaactctgg 660  
 tttttaata accgtcttcc tccaagccat actgaaggaa gacaagaaa cattgttgag 720  
 aaaataaata ttttaataa cttctctgt 749

<210> 14  
 <211> 1547  
 <212> ADN  
 <213> *Danio rerio*

5

<400> 14

tttttttttt ttttttctg tttgagctgt ctttattatc gtcctcgcca ttgagttcgc 60  
 tctctgogc gtgcctctgc gcaggcgcac tcagacggcc gcggtggccc agtgggtctt 120  
 cagcttcttg gctctggctc aggtgtgtct ggccgcgttc gtgctgctgt gtctcagtga 180  
 ttggtgatg ctcggagcgc tgtatgctgg ttggctgtat ctggaccggg acacgcctc 240  
 atgtggaggt cggaggtcac agtgggtcag aagctggagg atctggatgc atttcaggga 300  
 ttatttcca atcacactgg tcaagacggg ggatctggac ccgaggcaca attacctgct 360  
 gggtttccat ccgcacgggt ttctggtggc oggtggattc ggaacttct gcacggaggc 420

10

ES 2 640 100 T3

gtccggattc tcccagatgt tcccgggact cacgccgtac ctgctgatgc tgccgttctg	480
gttcaggggtg cccttcttca gagaatacat catgtgcgga gggctgggtg ccagtgagaa	540
ggccagtgcg agttacctgc tgggtcatcc aggcgggggt caagcggccg tcattgcagt	600
ggcggagct cggagtctc tggagggcgc gcctggagca ctgacactgc aactgctgca	660
gcgcaaaggc ttcatacaac tggccctcaa acacggagcg tggctggttc cagtgttctc	720
gtttggggag aacgagctct tcgaccagat ggagaatccc gcaggctccg ccctccgccc	780
catgcaggaa cgcctccaga ggataatggg cgtggccttg cctctcttcc atgctagagg	840
agttttccag tacagcttcg ggctgctgcc gtacaggaaa cccatacaca ccgtcgttgg	900
acgtcccatt cctgtaagtc agactccgtg tcccagtaaa gaagacatcg acgctcttca	960
cacactctac atgcagggcc tgacacaggt gtttgaggag aacaagaaac attacggcat	1020
cgcggacgac aaacatctga agttcacctg aactgatctg tgtgtgtgtg tgccgctgcg	1080
tatgtgtttg tttgtgtgtg tgtgtgtgtg tgtgtgtgtg cgttgcatat aatctagaat	1140
gattatctac catctagtta aaggctcttg tacgaggggt tgtgtagtgtg tgtgtcatg	1200
tgatatgacg atattcagaa taaccacagc cttgagctga agtacggta ctacacagtg	1260
tataaaacat taaggagatc cttttactgt taccctcctt cactacagta tggtaaatat	1320
tgttactgta ttttatgact tcattaaaac actatgtgac cgccattaag attattaaat	1380
gttcagatta aacaggtcat gaacactctg ggttttacag atcttcagga ttccctoctg	1440
tgcttaaadc agatttgctc ttaatcagca aatgtgtgtt gatgtgtatg acaacagaaa	1500
cacaaataaa gcggtgctgt ttattatctt aaaaaaaaaa aaaaaaa	1547

<210> 15  
 <211> 1279  
 <212> ADN  
 <213> *Caenorhabditis elegans*

5

<400> 15

ES 2 640 100 T3

attcttcatt ggacctctac gaccaccaca cagtcagtca tgtagatcg aacagaaagt	60
gaagataaga tgccacatct actaggagtt gagtgggctc cgctcaatat tccgttggct	120
cgcgctcttc aaactttggg agcacttcat ttctttctca tcaactctctt cacaccagta	180
ctcgtttctca ccgttccatt ctacatgta tataccgtac tctggccttt gatctttctg	240
tatgggcttt ggatgattta cgattggaat tcaccaaga agggagccta tatgagcaat	300
tggttccaga gacaaagaat tcattogtgg tatgccaact attttocagt caaattgcac	360
acaacatctg acatgccaga agaacataac tatttgattg ggtaccatcc gcatggaata	420
atttcaatgg ccgcattcat caactttgca acaaatggaa ctggaattct cgatactctt	480
ccacgaattc gtttccattt gtgcacactt gtgggtcaat tctggactcc gtggagacgt	540
gagtggggat tggtgcacgg aatgatagac tgcagtcgag aaagcatcaa gcacgttttg	600
gagcatgaaa agaaggaaa agcagttgta ttgggtggtg gtggagctga agaagcactt	660
gatgcacatc caggatgcca tattttgact ttgaaaaaa ggaaaggatt cgtgaaaatt	720
gccctgcaaa ctggagctca actggttcca tgctattcat tcggtgaaaa tgatattttc	780
aatcaagccg aaaatccaaa gggatcaaca attcgacagt tccaaacgat aatgaaaaga	840
gtcttgggat tctcccctcc agcattctat gggagaggag tattcaacta tacatttgg	900
cttcttccat tcagaaaacc tatcaacact gttctcggcg ctccaatttc agtgacaaag	960
acagtgaatc caactcaaga acaaatcgac acacttcac agacatacat ggaccgtctt	1020
catgagcttt tcgaggagca caagacaaaa tacgatgtct ctccaactac acaacttgtt	1080
atcaattaa tattatctat ttaatcactt ttgatgcact tttctcatat ttacatctct	1140
attatacctc attctgcaat ttagttcata ttttcagcca atttttcaat attacgagtt	1200
ctcatgtcaa atatocccat cagattcagt tactagttca gttattttta atcgatctct	1260
attataagtt tcgacacag	1279

<210> 16  
 <211> 1089  
 <212> ADN  
 <213> *Caenorhabditis elegans*

5

<400> 16

ES 2 640 100 T3

attcttcatt ggacctctac gaccaccaca cagtcagtca tgtagatcg aacagaaagt	60
gaagataaga tgccacatct actaggagtt gagtgggctc cgctcaatat tccgttggct	120
cgcggtcttc aaactttggg agcacttcat ttcttcttca tcaactcttt cacaccagta	180
ctcgttctca ccgttccatt ctacatgta tataccgtac tctggccttt gatctttctg	240
tatgggcttt ggatgattta cgattggaat tcaccaaaga agggagccta tatgagcaat	300
tggtccaga gacaaagaat tcattcgtgg tatgccaaact atttccagt caaattgcac	360
acaacatctg acatgccaga agaacataac tatttgattg ggtaccatcc gcatggaata	420
atctcaatgg ccgattcat caactttgca acaaatggaa ctggaattct cgatactctt	480
ccacgaattc gtttccattt gtgcacactt gttggtcaat tctggactcc gtggagacgt	540
gagtggggat tgttgacgg aatgatagac tgcagtcgag aaagcatcaa gcacgttttg	600
gagcatgaaa agaaaggaaa agcagttgta ttggtggtg gtggagctga agaagcactt	660
gatgcacatc caggatgcca tttttgact ttgaaaaaa ggaaaggatt cgtgaaaatt	720
gccctgcaaa ctggagctca actggtcca tgctattcat tcggtgaaa tgatattttc	780
aatcaagccg aaaatccaaa gggatcaaca attcgacagt tccaaacgat aatgaaaaga	840
gtcttgggat tctccctcc agcattctat gggagaggag tattcaacta tacatttgg	900
cttcttccat tcagaaaacc tatcaaacact gttctcggcg ctccaatttc agtgacaaag	960
acagtgaatc caactcaaga acaaatcgac acacttcatc agacatacat ggaccgtctt	1020
catgagcttt tcgaggagca caagacaaaa tacgatgtct ctccaactac acaacttgtt	1080
atcaattaa	1089

<210> 17  
 <211> 1149  
 <212> ADN  
 <213> *Caenorhabditis elegans*

5

<400> 17



ES 2 640 100 T3

tcagctttoc actgaaaatg cgcgaatttc toggaataga atgggtggat ctattctcat 60  
caattcaacg taaaaagaca tacttgggag ttgtctatca ctttatgctc acttaccac 120  
ttgcactttt cgtcaccatt ttgccatttt tcttgctttt cactttcca tggcacattt 180  
tggctctcta cgcttgctgg tacttctacg atatggattc tccgaggaga ggtggatatt 240  
ccagtgattg ggtcagaaaa tggcgtgtca acgactgggt cgcccaatac ttccaatca 300  
acttgcacaa aactgccgaa ctctccacgg acaagaacta tctggttggga atccatcctc 360  
atggtatcat ctccatggct gcatggctca attttgccac caatggaacg ggaatctatg 420  
agaaattccc tggaaattcg tggaaattgt gcactttagc acttcagttc agaatggcca 480  
tccgtcgtga gcttctcctt ctcaaccggc taatcgactg ctccagagaa tccattgaat 540  
atgtgcttga caaatgtggc cagaagggac gtgcagtggg attggtgatt ggaggagccg 600  
aagaagcttt agatgctcat ccaggctacc acacactcac tttggcatca agaaagggat 660  
ttgttcggga ggctttgatt actggtgcct acttggctcc ggtgtattct tttggagaga 720  
atgacgtttt tgaacagatg gagaatccag ttggctcaog gctccgcaac ttccaagaat 780  
ggtgcaaaaag cattttcggc atctcatatc caattttoca tggctcgtggc ttcttccaac 840  
taactttcgg atatcttcca ttcogtaaac caatcgatac cgtagtogga gccccaattc 900  
ccgttgaaaa agtggaaaat ccgacgaaag agcaaatcga cgagcttcac accatttatt 960  
gtcaaaaagt gacggagctg ttcgatgagc acaaggaaaa atatggagtg gagaaggacg 1020  
tgccgttggg gctccggtag agcggaaaat ttgaaatggt ctgtgaaaag tgccagtttt 1080  
tgatgagatt tctgtatttt atcgtgccat aatttttggt tttcgagcct tgttttaagt 1140  
tgttacggt 1149

<210> 18  
<211> 1136  
5 <212> ADN  
<213> *Caenorhabditis elegans*

<400> 18

atgctaaact accaaattca caaaaagctc accgacatca agtgggtgaa catcttctcc 60

10

ES 2 640 100 T3

ccatgggatc gccagcgtgc ctacttcgcc ttggctgtct ggttcgggct catctacca	120
ttctgctgcc tgtgccaggt ggctccgttt gtgctctttt tcaccggcca gtggattatt	180
ttgggtctct acgcagtttg gtacctttac gatcgagaat ctccgagaag aggaggatat	240
cgggataatt ggttcagaaa tttgtcgtcg cacaaagtgt tcgccgagta ttttctgtt	300
aaacttcaca aaactgcgga gttggatcca aaccaaatt atttattcgg atatcatcct	360
catggaattc tcggtgctcg agcgtggtct tgttttggat ttgatgcgtg caatgtgaag	420
caagtgttca aaggcatccg cttcaacatc tgcaccttgc ccggcaactt caccgcaatg	480
ttccgcgcg agatctcct cagcatcggg atgatcgaga gctccaaaga atccatcgag	540
cacgtgctca actccgagga aaagggccgt gccgtttaa ttgtcgtggg tggagccgct	600
gaagctcttg aagctcacc agggaagcat actctaacac tggcaaatcg caaaggtttc	660
gtgagagaag ccgtgaagac cggagctcat ctgggtgccag tttatgcgtt tggagagaat	720
gacatatata agcaaattga caaccggaa ggctcgaat tacggaaaat tcaagaatgg	780
ggaaagaaga aatgggaat ttcactgcca ctaatctacg gaagaggata ttttcaatg	840
gctcttgggc ttcttccaat gagcogagct gtgaatgtag ttgtcggagc gcctattcaa	900
gtggaaaaag agctcgatcc ttctaaggaa gtcattgatg aaattcatgg agtttatatg	960
gaaaagctcg ccgagttatt tgaagagcac aaggcaaagt tcggagtttc caaggacact	1020
cggtcgtttt ttcagtgaaa taacttttat ttagattat ttgtttgatg tttgatgtct	1080
gttatgttat cttttttat caccaaactg tgtaattttg tttacgggtt cattac	1136

<210> 19

<211> 1176

<212> ADN

<213> *Caenorhabditis elegans*

5

<400> 19

ES 2 640 100 T3

atgagactcc ggctgagctc gatatctgga aaggogaagc ttcccgataa agaaatatgc	60
tcatcagttt cgagaatatt ggcaccattg cttgttccat ggaagcgacg actcgagact	120
cttgccgtga tgggtttcat tttcatgtgg gtaatcctac caatcatgga cctctgggtt	180
ccattccacg tcttgttcaa tactcgatgg tggttccttg ttccactcta cgctgtctgg	240
ttctactatg attttgatac accgaaaaa gcttcaagaa gatggaattg ggccagaaga	300
cacgtagcct ggaagtactt tgccagctac ttoccattga gattgatcaa gactgctgac	360
cttccggcgg atcgttaatta catcattggc tctcatcctc atggaatgtt ctcggttgg	420
ggttttactg caatgagcac caacgcgacc ggatttgaag acaagttccc gggaataaaa	480
tctcacatca tgacgctaaa tgggcaattt tatttcccat ttcgtcgaga attcgggaata	540
atgctcggtg gaatcgaagt ttcgaaagaa tcaattgaat acaacttaac taaatgtgga	600
aaaggacgag catgcgcaat tgtcattggc ggagcctcgg aggctcttga agctcatccc	660
aataaaaata cattgacggt gatcaatcga cgtggtttct gcaaatatgc tctgaagttt	720
ggagcagacc tcgtaccaat gtacaatttc ggagagaatg atttatacga gcagtatgaa	780
aaccogaagg gatctagatt gcgagaagtt caggagaaaa tcaaggacat gttcggattg	840
tgtcccccot tgctccggcg tcgatcgttg ttcaaccaat accttatcgg attgctgccg	900
ttccgaaaac cagttacaac agtcatggga aggccaattc gagtcaccca aaccgacgag	960
ccaaccgttg agcagattga tgagctgcat gcaaaatatt gtgatgctct ctacaatctg	1020
ttcgaggagt acaagcatct tcaactccatt cctcccgaca ctcatctcat cttccagtga	1080
ctttttcaca aattttgtat attttgacaca cagattgttt tategccaag ttaatgtata	1140
gtaatcttct tgtaaataaa tcagataaac attaaa	1176

<210> 20

<211> 1239

<212> ADN

<213> *Kluyveromyces lactis*

<400> 20

5

ES 2 640 100 T3

```

atgggtaagg aacaaatcag aaagagacaa actaaaaagg tacgcaaggg ttcacaaccc      60
gtggcaaatg gtaaactctc tgtagtaggg gtctccaaag atgggaaacc agaattttgt      120
gctattgata cgccataga aaggagactt caaactctca gcgtagcatg gtatgtcgaa      180
tgtattccgt taatgatcat cataatggtg tttgtttggg ttaaccogtt gatgtggtcg      240
tttgaatac catacaccat ttattatctc atagaccgga ctgcctccaa tgggaacgct      300
gtgaagagac actccaagtg gttccgatct ctaaaggctc ggttttatct ccgtgattac      360
ttcctataa gtatgcacaa aagcaactgaa ctogagccaa catttacaag tattgattcc      420
acagagctgg aaaatgatgc gagtgagcca ggctacctgg atagctcaca gcccgtttta      480
ccagataaat ggtggaatcc tttcagagaa aaggatgaaa ctgtagacc aacaggaccg      540
aggtatattt ttggctacca tccacatggt atagcagcat tcggagcgtt tggagcattc      600
gctaccogag catgtaattg gagtaaagtt ttcccaggca ttccagtgtg tttgcttact      660
ttggtaaatc aatttcaaat accagtgtac cgtgattatt tattggcctt aggcataaca      720
tcggtcgcta gaaaaaatgc tatgaaagta cttgagaaaa attattcaat agccatcgta      780
attggaggtg ccagtgaatc ccttcttact aaccttgggt catctgatat cattttaagc      840
aagcggaaag gttttgtgaa gttggctttg caaactggga atgtgagtct tgtacctgtg      900
tatggatttg gtgaaactga tacttacaag attctgaagc ttaaaaatga ctctataata      960
ggcaggatac agatatggct taaagagaac tacagcttca ctgtccctct atttttcgct     1020
cgtggtgtat tcaattatga ttttggttta ctaccgttta gacatccagt caatgttgtt     1080
gttgaaacc ctattcatat taaagaaaag atagatcatc caacaattga ggagattgat     1140
cattaccatt ccttatacat tgaagaattg aaacgtcttt acgatgataa caaggcaaaa     1200
tttaactatt cagaaaaaac tttaaatata gtggaatga                               1239

```

<210> 21  
 <211> 1386  
 <212> ADN  
 <213> *Ashbya gossypii*

5

<400> 21

ES 2 640 100 T3

```

atgcaggaca gcatggatga cagcttgagg gaggtgagg gcagacagga cgactcggag      60
gtcagcagcg ggacgacact cggaagcagc accccggagg acagcggggt gacggcgaag     120
ctgcgcaaga agtaccagat ggcgtcggca ctgctgcggc gggagctgga ggagctgagc     180
gtgtacgacg cgaagacggc gggggtaagc ggccgcagca gcggcagcgg cagcggcggg     240
ctggcgctgc tgggcgggcg attccacgtg gcgccgctgc ggatccccgc gcggcggcgg     300
ctgcagaagc tgggtgtggc gtggcacacg agctcgttca tctacatgac ggtgctggtg     360
ctgtttctgg cggcgaacct gctgatgtgg tggttcatgg tgcogtatat ggtgtactac     420
gtgtggaacc ggtcgcggc gaacggcggg gtggtgcggc ggtactcgcc gcgctgcgg     480
tcgctggcgc tgtggcggtg ctactgcgag tactaccoga tctcactgca caagagcgag     540
gacctggcgc cgacgttcgt gccggaccgc cggggggcgg agccgcgcga gtggaagctg     600
cggctgtggc tgtggccgac gcgcgtggag ctgctgaact tgacgctcca gtggacgcgc     660
gcgoggcgc aggttgcgac agggccgcgc tacatcttcg ggtaccacct gcacggcgtg     720
ggcgcaactg gcgccttcgg cgcgatcgcg accgagggct gcaactggag caaggtcttc     780
gcgggcatcc cggcctgcct gtgcacgctg gtgaaccagt tccagatccc catctaccgc     840
gactacctgc tgggcctcgg gtgcacctcc gtggcccgca agaactgct caaggtgctc     900
gagcagaact actcggctcg catcgtcgtg gggggcgccc aggaagccct gctcagcgt     960
gtgggtcca ccgagctcgt gctgaacaag cgcaagggt tcatcaagct cgcgctggag    1020
acgggcaacg tcaacctggt gcctatctac gccttcggcg agacagactg cttcaactg     1080
ctcgacacgg gcaacgagag ctacctacgc aagttccagc tgtggatcaa gaagacctac     1140
ggcttcacca tcccgttctt ttttgcgcgc ggtgtcttca actatgactt cggcttctg     1200
ccctcagga atccgatcaa cgtcgtcgtg ggcaagcctg tgtacgtaga caagcgcaga     1260
acgaacccca ccatggagga aatcgaccac taccacgacc tgtacgtgca ggagctgcgc     1320
aatgtgttcg acaagaacaa acacaagttc ggctacgccg ggaaggaact taaaatagta     1380
gagtag                                           1386

```

<210> 22  
 <211> 1653  
 <212> ADN  
 <213> *Magnaporthe oryzae*  
  
 <400> 22

5

ES 2 640 100 T3

atggccacgg tagatgatac ccagaacccat gtccaactga atgcagagtc cagcctaaac 60  
 gaatttcogt accaagcaag cgcggacaag gaggagtatc caagccttcc ttctgccccg 120  
 ctccctgacg tggacggcag ctcccaaat gggcatacag ccgccaacgc tgaccttccg 180  
 ccctcggagt cgaccgaaca gccatggaac aaggtagcca gtggcgactt catccatgaa 240  
 accaagatgc ccagtgaaga ccgggacaac cttgccaggc ccggcagtct accaactccc 300  
 gagccgagcc gtagaggcag caacacgctg gtcggcagcc cgcgtcaggg cgacggattg 360  
 caaatgctca tggacgaatt cgcagacgat gaatcctcag atacctcgtg ggacattcac 420  
 gagaccatca ggagtcgcac ccaggccttg gctgagaagg accaccaggg gatattctca 480  
 aatcagcggc gaggcaacca ggccgacgcg ccagacacgg taaccagtgc caacgcttgg 540  
 ttgccggctg gcttcgaaga cgacctggac gaagtgaccg gtgctgacga cgagttcgag 600  
 ccaaccogtg ggagtcgctg gagagtgagg ctggcgcccc tggatacccc ttttcatcgc 660  
 cggatgcaga ccctggtggt tctgctacac gtcctcggca tgggaatcac tttcagcttt 720  
 ttctgcttcc tatgcaogct cccoctgttt tggcctatca tcatcgccta cctcgtattc 780  
 atcagactgt cgagagctgg atccgatggg aagacaaacc gtcgcataga atggctccgc 840  
 agggctcgca tctggaagta ctttgccgat tacttccctg tcaagotaca caagacagcc 900  
 gacctgcccc ccacccgcaa gtacatTTTT gcgattcatc cacacggaat catctcgcac 960  
 ggtgcctttg cctccttgcg caaggaggcc ctccgcttct ctgagaagtt tcccggcacc 1020  
 acaaacagtg tgtgcaactc tcacggcaac ttcaagaccg cgttctaccg cgagtatctg 1080  
 atggccatgg gcctgatctc ggtttcaaag gagtccatta ttaacacttt gatgaccggc 1140  
 gggatcaacg gtgaaggcat gggcagggcc gtctccataa ttgtcgggtg ggogcgagag 1200  
 gcactcgagg cctgccccaa gacgatgocg ttgatactga aacggggggg tttttgtcgc 1260  
 atggcactac gcaccggtgc agaccttggt ccgggtgctct gctttggtga gaacgacctc 1320  
 taccaacagt gggggccgca ggatcatccg cgcttccgtc gcctgcaaat gaaggctctc 1380  
 gcgtacctag gctttgcggt ccccgttctg aggggaaggg gcgtcttcaa ctatgacttt 1440  
 ggagtccttc ctcagcggcg gccgatcaac gttggtggtg gtgagcctat caagggtgag 1500  
 cagtttcggg gtagcggtaa cattgagcca cgcgtcgagg agctgtggaa gctatattgc 1560  
 cagaagcttc aggaactgta cgaccagaac aagcacgtct atttcaagga tcgtcaggag 1620  
 gatatgaggt ttggtgactc ggaggagttg taa 1653

<210> 23  
 <211> 1410  
 <212> ADN  
 <213> *Ciona intestinalis*  
 <400> 23

5

ES 2 640 100 T3

gccataaaac gttgagaaaa tgttttgtgt tacgtcaagt tactcattaa ctcaatgtaa 60  
 agattgatat agccagagca aattgtagtt aagtagcgca tatacccaac ttggaaactg 120  
 ctatatagtt ggtaaagcgg attaggaacc atgtgggaat togcaccgct tcgtattcct 180  
 atgataagaa gacttcaaac gttggctggt totcattttg tgttctgctt tottgcccta 240  
 gctgccacat gcactttatt gactatctac attactttgt tcacttcata ctggccgatc 300  
 accaccgctt accttgtggt ottattcttg gaccgaaaca ccccggaatc tgggtggaaga 360  
 agaagcgagt ggggtgagaaa ttggagtctc tggaaatgga tgagcgatta ctttccttgt 420  
 acgttgacaa aaaccgttga cctcgatcca aaaagaaact acatatttgg catccatcca 480  
 catggggttt tgtgcatcgg atctttcact cacttctcca cgaatggctc cggattttcg 540  
 cacgtttttc cgggatttac atcttatctt acgatgttgc ctttctggtt caaaatgcc 600  
 tttttctggt actatgttat gagcggcgggt ttgactccag ccacaaggaa agccatcaaa 660  
 cacacaataa caaggccagg tgggtggcat atatgctgca ttatacctgg tgggtgctcca 720  
 gaatctctaa acgcacggcc tggagatggt gtattgttgc tgaaacaaag acttgattt 780  
 ctcaagctgg caataacgaa cggagttcct ttggttcccg tgttttcatt tggcgatcac 840  
 gcgttgtggg agcagaaaacc gaatccccc ggttccctca tccgtaggtt ccaagatagc 900  
 tctcagaagt ggatgcaagt ogctcttccc gttttccacg cacggggtat ctttcagtac 960  
 aattttggac taattccgta caggcgtagc gtccataccg tagtgggcca gcccatagaa 1020  
 gtgccacaga acagtaaccg gacatcggag gaccttatga gtttgcaaga ggattatatt 1080  
 aatcgattga gagcaatttt tgacgagcac aagagcaaat atctccaga agattgcaaa 1140  
 ctgatcatca actaaactaa actaaactat gcaaccaaag ctttttaaat tgagttagtt 1200  
 atagtagtgt gaagtaagat agagcacctt tagcaataat gctgtccaaa tttcctgggt 1260  
 caaacaatta ataacggacg ggagtcgtga tgaaacgggt ttagaatact ttaagtgttt 1320  
 ttgtttacta ccaaatgaaa agataaggta gaataaaagg gtgtctcatc tccccacc 1380  
 gtactatata tatacacgat cgaatcattg 1410

<210> 24  
 <211> 1728  
 <212> ADN  
 <213> *Mus musculus*

<400> 24

agcatggtgg agttcgcgcc cctggtggta ccatgggagc gcaggctaca gaccttcgcg 60

ES 2 640 100 T3

gtccttcagt gggctcttctc cttcctggcc ttggcccagc totgcatcgt catcttcgta 120  
 ggctctctat tcacaagggt ctggetcttc tctgtcctgt atgccacctg gtggtacctg 180  
 gactgggaca agccgcggca gggaggccgg cccatccagt tottcagacg cttggccata 240  
 tggaagtaca tgaaggatta tttccctgtc tctttggtca agacagctga gctggacct 300  
 tcccggaaact acatcgcggg cttccacccc catggagtcc tagcggctgg agcctttctt 360  
 aacctgtgca ctgaaagcac gggctttacc tcgcttttcc cgggcatccg ctccctatctg 420  
 atgatgctga ctgtgtggtt ccgggcccc ttcttcogag attacatcat gtctgggggg 480  
 ctggtctcat cagaaaagggt gagtgccgat cacattctgt ccaggaaggg cggcgggaac 540  
 ttgcttgcca tcatcgttgg gggcgcgcag gaggcactgg acgccaggcc tggagcctac 600  
 aggctgctgc tgaagaatcg caaggcttc atcaggctcg ccotgatgca tggggcagct 660  
 cttgtgocaa tcttctcctt tggagaaaac aacctgttca accaggttga gaacacctt 720  
 ggtacctggc tgcgctggat ccagaaccgg ctacagaaga tcatgggcat ctccctccct 780  
 ctcttcacg gcagaggtgt cttccagtac agctttggcc tcatgccctt ccgccagccc 840  
 atcaccacca tagtggggaa gcccatcgag gtgcagatga caccacagcc ctcaaggagg 900  
 gaggtggacc ggcttacca gcgctatac aaggagctct gcaagctctt tgaggagcac 960  
 aaactcaagt tcaacgtccc tgaggaccag catctggagt tctgctaagt gtctccagcc 1020  
 ggaagacagc tgcatctgag cgctgcagg agtgtgggat tagggggact tccacagcca 1080  
 ccagacactc ctacaaacct agccacaact gccaatgag aagagggggc agctcctaact 1140  
 cctgggattt gaacctgcag ccaaagctct gaggtctccc tgtccttggc ctgtctgcac 1200  
 atctgtagaa tgggggaaaa gcaggcagag agaaattcct gaggtctctt cccacagttg 1260  
 taatgtcatt caaacatgac caaaggacaa acagggagaa agagaacaaa actgtttctt 1320  
 atctaccctt gagggacagt gcaagagaag ccagcaccoc aggcctccct gtgcatgctc 1380  
 cctgatgctg cttcttccct ctgaggcaga gacggggagc caagtctgcc ctggcaccta 1440  
 ctctatgttt cttcagatto tgggtcctct gagctatgat accaaaggag ccacagaaggc 1500  
 agataaggag ggacggggtc actgactatg accgagggtg ggtctccttc ccatatcctg 1560  
 agcctcagtt tccccagcct taatgacctg ggagcggcac actgctcacc acagaggctc 1620  
 caccagagag cctcttactc atgctttcta gtgaactcca gcctctgtct tggcactgaa 1680  
 gggcagcact gtacatgtta cctcaataaa tgaaggagt ctgtctta 1728

<210> 25  
 <211> 1005  
 <212> ADN  
 <213> *Homo sapiens*  
 <400> 25

5



ES 2 640 100 T3

```

atggtagagt tgcgcocctt gtttatgccg tgggagcgca ggctgcagac acttgctgtc      60
ctacagtttg tcttctcctt cttggcactg gccagatctt gcaactgtggg ottcatagcc      120
ctcctgttta caagattctg gctcctcact gtctctgatg cggcctgggtg gtatctggac      180
cgagacaagc cacggcaggg gggccggcac atccaggcca tcaggtgctg gactatatgg      240
aagtacatga aggactatct ccccatctcg ctgggtcaaga ctgctgagct ggaccctctt      300
cggaactaca ttgcgggctt ccacccccat ggagtcctgg cagtcggagc ctttgccaac      360
ctgtgcactg agagcacagg cttctcttcg atcttccccg gtatccgccc ccatctgatg      420
atgctgaocct tgtggttccg ggcoccttc ttcagagatt acatcatgtc tgcagggttg      480
gtcacatcag aaaaggagag tgctgctcac attctgaaca ggaagggttg cggaaacttg      540
ctgggcatca ttgtaggggg tgcccaggag gccctggatg ccaggcctgg atccttcacg      600
ctgttactgc ggaaccgaaa gggcttctgc aggctcgccc tgacacacgg ggcaccctg      660
gtgccaatct tctccttcgg ggagaatgac ctatttgacc agattcccaa ctcttctggc      720
tcctggttac gctatatcca gaatcggttg cagaagatca tgggcatctc cctcccactc      780
tttcatggcc gtggtgtctt ccagtacagc tttggtttaa taccctaccg ccggcccatc      840
accactgtgg tggggaagcc catcgaggta cagaagacgc tgcacccctc ggaggaggag      900
gtgaaccagc tgcaccagcg ttatatcaaa gagctgtgca acctcttcga ggcccacaaa      960
cttaagttca acatccctgc tgaccagcac ttggagttct gctga                          1005

```

5

<210> 26  
 <211> 1128  
 <212> ADN  
 <213> *Pan troglodytes*  
  
 <400> 26

ES 2 640 100 T3

aaaaacctgt ggggtgcctca gaccacagca gagctcacag aacctgcggg agccaggctg	60
accogccagc atggtagagt togcaccctt gtttgtgccg tgggagcgca ggctgcagac	120
acttgctgtc ctacagtttg tottctcctt cttggcactg gctgagatct gcactgtggg	180
cttcatagcc ctctgttta caagattctg gctcctcact gtcctgtatg oggcctggtg	240
gtatctggac cgagacaagc cacggcaggg gggccggcac atccaggcca tcagggtctg	300
gactatatgg aagtacatga aggactattt ccccatctcg ctggtcaaga ctgctgagct	360
ggacccctct cggaactaca ttgcgggctt ccacccccat ggagtcctgg cagtcggagc	420
ctttgccaac ctgtgcaactg agagcacagg tttctcttctg atcttccccg gtatccgtcc	480
ccatctgatg atgctgacct tgtggttccg gggccccctc tcagagatt acatcatgtc	540
tgcagggttg gtcacatcag aaaaggagag tgctgctcac attctgaaca ggaagggtgg	600
cggaaacttg ctgggcatca ttgtaggggg tgcccaggag gccttgatg ccaggcctgg	660
atccttcacg ctgttactgc ggaaccgaaa gggcttcgtc aggetcgcgc tgacacacgg	720
ggcaccctct gtgccaatct tctccttcgg ggagaatgac ctatttgacc agattcccaa	780
ctcttctggc tcttggttac gctatatcca gaatcggttg cagaagatca tgggcatctc	840
cctcccactc tttcatggcc gtggtgtctt ccagtacagc tttggtttaa taccctaccg	900
ccggcccatc accactgtgg tggggaagcc catcgaggta cagaagacgc tgcacccctc	960
ggaggaggag gtgaaccagc tgcaccagcg ttatatcaaa gagctgtgca acctcttcca	1020
ggcccacaaa ctttaagtca acatccctgc tgaccagcac ttggagttct gctgagccca	1080
aagggcaggg ccaacattag ggagcccagc aggaggtgct gtgctgag	1128

<210> 27  
 <211> 1005  
 <212> ADN  
 <213> *Canis familiaris*

5

<400> 27

ES 2 640 100 T3

```

atggtgaagt ttgcaccact gttcgtgccca tgggagcgca ggctgcagac cttcatggtc      60
ctgcagtggg tcttctcctt cctggccctg gccagatct gcaactgtgtt cttcgtaggc      120
ctcctgttca caaggttctg ggtggtcagc gtccctgatg ccgcctgggtg gtaccgggac      180
cgggacacgc cgaggcaggg gggcaggccc gtccaagccc tgaggcgctg cttcttgtgg      240
aagtacatga gggactactt ccccgtcacg ctgggtcaaga ctgctgagct ggaccctcc      300
cggaactacc ttgctggctt ccacccccac ggggtcctgg cgcaggagc ctttgtcaac      360
ctgtgcacgg agagcacggg cttccctttg ctcttccctg gcatccgctc ccatctgatg      420
atgctgacct tgtggttccg ggtcccttcc ttcagggatt acatcatgtc cgggggcctg      480
gtcacttcgg acaaggagag cgctgctcac atcctgagca ggaaggaggg tggcaatctg      540
ctggccatca ttgtaggggg tgcccaggag gcgctgaacg ccaggcctgg agactccacg      600
cttctgctcc ggaaccgcaa gggcttcacg agacttgccc tgatgcacgg ggcagctctg      660
gtgccgatct tttcctttgg ggagaatgag ctatttgacc aggttgagaa ctctcctggc      720
tcctggttgc gtaggattca gaaccggctg cagaagatca tgggtatctc cctcccgtc      780
ttccatggcc gtggtgtctt ccagtacagc tttggcttta taccctaccg ccagcccac      840
accaccgtgg tgggaaagcc catcgaggtg cagaagacac tgtgtccctc aaaggaggag      900
gtggacaagc tgcacacagc ctatgtcaag gaactgtgtg aactcttcga aaccacaag      960
ctcaagtaca atgtccctgt ggaccaacat ttggagttct gctga      1005

```

<210> 28  
 <211> 1820  
 <212> ADN  
 <213> *Bos taurus*  
  
 <400> 28

5

ES 2 640 100 T3

gaggagccag catggtggag tttgcacctt tgtttgtgcc actggagcgc agactgcaga 60  
cctttgcagt cctgtactgg atcttctgct tcatggcttt acccccttta tgccttgtgg 120  
tcttcatagg cctcctgttc acaagattct ggctcttcaa tatoctgtat gtgatctggg 180  
tgtacttggg ccaaaatagg ccacggcagg ggggcagtca caataagttc ctaaagcgct 240  
gggtcctgtg gaagtacatg aaggactatt tcccatcac gctagtcaag accactgagt 300  
tggaccctc ccggaactac cttgctgcct tccaccctca tggactcctg tcagtcggag 360  
cttttaccaa catgtgcaca gacagcacgg gcttctctc acttttccct ggcacccgcc 420  
cacatctgac aacgataaac atatatttca ggatcccact cttcagagat tacatcatgc 480  
aaggggggct ggtctcatca gacaaggaga gtattgctta cattctgagc aggaaaggag 540  
gcggaacct agtggccatc actgttgggg gcatcagga ggcactgaaa accaggcctg 600  
gagccaacaa gctggtgctg cggaaccgca agggcttcat caggctcgc ctgatgcatg 660  
gagcagttct ggtgccgatc ttctccttg gggacaatga attatatgct aagacttctc 720  
ctggcttttg gtggaaatgg ttcagagatc aactgtacaa gaagacacgt cttgccatcc 780  
cgttcttcta tggcogtggg gtcttccagt acagttttgg ttttatgccc tacogccgac 840  
ccatcaccac cgtggtgggg aagcccatcg aggtgccgaa gataccacat cctcccacag 900  
aggaggtgga caggctgcac cagcattacc tgaaggagct gagcaacctc ttcgaaacct 960  
acaagctcaa gtacaacatc cccaaagacc aacacttggg gttctgctga gtacctcccc 1020  
agaggagggg cttgccccaa agggcagggc tggcattagg aaggagggtg ctgtgatccc 1080  
agaaggtttc ctggaggtat ctggtgaaca catctccaga gccttactg actcctgcat 1140  
ctccacccca tgccaggccc caggtcacat ctggctcata gaagaggaga ccagactctg 1200  
gcatcaatga tttgctctcc tccactaact cctctgtaac gggcctctc tttggttctg 1260  
cccactgag aatgggaga gatactggga gaaaatgctg tctgtgtgtc cgtcctctgt 1320  
ggctctgtoa ctcaataagc aggacaaaca gcagctccag tgagcagggt tcccttctc 1380  
tgccttgagt gacagctgga agcagcccc agctgcagcc ctttctccta caagctccct 1440  
tatoctacat cctgcctgtc ccatcttctc cctccctgag ccggagagga gggggagggc 1500  
acaacctcct ggctcatcct ggcacctctg agtttctgag ccccatatgg cagccagaag 1560  
ggcctggagg tggggaagac ctgagctctg tccactggta aatgactgta caccaacctc 1620  
ttcaacctcc ctctgagtct cagtttccc aaggactaaa atgacctctg acctoggctc 1680  
taatcacaca taggctgaaa gctgcttatg gtaactccta ctotatcctt tgtaattcag 1740  
cctgacctt gggtccata tatgtcaatt catcacatga aagattcgat tttaaaaaaa 1800  
aaaaaaaaa aaaaaaaaaa 1820

<210> 29  
<211> 1774  
<212> ADN  
<213> *Rattus norvegicus*

ES 2 640 100 T3

<400> 29

```

gtgtgtggag cctcctgagc cagctcagca tgggtgaatt cgccccctg ttggtacct      60
gggagcgcag gctacagacc ttcgcggtcc ttcagtgggt cttctccttc ctggccttgg      120
cccagctctg tatcttcata ttcatoggcc tcctattcac gaggttctgg ctcttctcag      180
tcctatatgc cacctggtgg tacctggact gggacaggcc gaggcagga ggccggccca      240
tccaattctt cagacgcgat gccatagga agtacatgaa ggatttcttc cctgtctctt      300
tgggtcaagac agctgagctg gacccttccc ggaactacat cgcaggcttt ccccccatg      360
gagtcctggc ggccggagcc tttcttaacc tgtgcaactga aagcacaggc tttacctcgc      420
tcttccaggg cattcgctcc tatctgatga tgctgactgt gtggttccgg gcccccctat      480
tccgagatta catcatgtct ggggggctgg tctcgtcaga aaaggtgagt gctgatcaca      540
ttctgagcag aaagggcggc gggaaactgc tagccatcat cgttgggggc gcacaggaag      600
cactggatgc caggcccgga gcctacaggc tgctgctgaa gaatcgcaag ggcttcatca      660
ggctcgcctt gacacatggg gcagctcttg tgccgatctt ctcccttggg gagaacaacc      720
tgtttaacca agttgagaac acccctggca cctggctgog ctggatccag aactggctgc      780
agaagatcat gggcatctcc cttccgctct tccatggcag aggtgtcttc cagtacagct      840
ttggcctcgt gcccttccgc cagcccatca ccaccgtagt ggggaagccc atcgaggtgc      900
agatgatacc acatccctcg gaggaggagg tgaaccggct tcaccagctc tatatcaagg      960
agctatgcaa gctctttgag gagcacaaac tcaagttcaa cgtccctgag gaccagcacc      1020
tggagttctg ctaagtgtct ccagccgga gacagctgca tctgagcggc tgcaggagtg      1080
ctgggatgag agggatttcc acagccacca ggaactccta ggagcccagc cacagctggc      1140
aagatggagg cagttcctaa ccctgggatt cgtccctgca gccaaagctc tgagctctcc      1200
ctgtccttgg tctgcacatc tgtagaatgg gggggaaagc aggcagagag caattcctga      1260
ggtctctctc cacagttgta atgtcattca aacatgacca aaggacagcc atggagaaag      1320
agaacaaaat tgttctttac ctacccttga gggacagtgc aagaggagcc agcaccaccag      1380
gcctctctgt acatgcccc tggtgcctc ttcttacct ctgagctgga gaggggaagc      1440
catatatgcc ctggcaacca ctctatgttt cttccagatt ctgagtcctc tgagctatga      1500
catgaaggag cccagagaca gttaaggagg acagagctca ctgactatga ccgaggttag      1560
atctccttcc catgtcctga gcctcggttt cctcagcctt aatgaactgg gagccacacc      1620
tctcacatca cagaggctcc cccagggagc ctcttactca tgctttctag tgaactctag      1680

cctctgtctt ggcacaaaag ggcagtgtg tacatgttgc ttcaataaat aaaaggagtc      1740
tgtcgtaaaa aaaaaaaaaa aaaaaaaaaa aaaa      1774

```

5

<210> 30

<211> 1056

ES 2 640 100 T3

<212> ADN  
<213> *Gallus gallus*

<400> 30

5  
atggagcgtg cctataaaat ctgtgatgac accacgtcca aaagactgaa aaacctcttc 60  
tgggctgtgc agaacgggtc ccccaggtg agccgttcct ctcccaccac cctcaccgct 120  
caccgcctgg cttttctccc agcacagtgc tgcaccgctg ccttcacggc cctcctcttc 180  
accgccttct ggctgctcag tgtcctctat gccgcctggt ggttcgtgga caggaagca 240  
ccgctcaggg gtggcgacg gatccacatg gtgcggaaca gcgcagtctg gagacacatg 300  
agggacttct tccccgtcac gctggtgaaa acagcagagt tggaccccag gcagaactac 360  
ttggtgggat ttcaccccca cggggtcctg gcagtgggag ccttcaccaa ttttggaaaca 420  
gagcgacag gcttttccac gatcctcca gccatcacc cacacctgat gatgctctcc 480  
ttatggtttc gtgtcccctt cttaaggac tacctgatga gggaggcct tgtgtcctct 540  
gataaggaga gtgcgtacca tgtgctgcag aggccagagg gggggaacct gttggccatc 600  
atcgttggtg gggctcagga ggctctggat gcccgccag gatcctgcac cttgctgctg 660  
aagaatagga agggatttgt ccgcgtggc atcgagcagg gcaccccact ggtccctgcc 720  
tttcctttg gggagaatga gctcttcgac caggtgagca accccaaggg ctcgtggctg 780  
cggtgatcc agcaccggct ccagcagatc atgggcatct cctacccct ctccacgca 840  
agaggtatct tccagtacag cttcgggtg gtgccctacc gccggcctat caacaccgtg 900  
attggaagc caattcccgt gctgaagaag cacaagccaa cggaggaaga ggttgatcga 960  
gtccataaga aatacttaga agagctgagc aagctctttg aggagcaca agctaaatac 1020  
aacgtcccag aagacagtca cctcgaattt atatag 1056

<210> 31  
<211> 1287  
<212> ADN  
<213> *Danio rerio*

<400> 31

10  
15  
gtgtttggtt cagaatgacc cagtaaaactt gttcatttgt ggagcctgtg ggaataagta 60  
aacaaatcat acctggatct ttgagttctc cagagctgga ttttactcgc cctgctgtat 120  
tgtcctgttc ttcacctcgc tcttcaccog tttctggctc atcagtgtgc tttatgccac 180  
ctggtggtac atcgactggg acacaccttc acgaggagg agaatagtgc catatctctg 240  
caaactgcgt gtgtgggagt acatgagaga ttatttcccc atcaggctga taaagacggc 300

ES 2 640 100 T3

tgacttggac actaggaaga actatgtgat gggttttcat cctcatggta ttctcgtggc	360
tggtgctttt acaaatttct gcacggaggc gactggattt agcaaaactct tcctcgggat	420
aaagagcaac cttcttatgc tgcctctctg gttcagagcc ccttcttca gagattacat	480
catgtcagca gggctggttc catcagacaa ggagagtgcc agttatctgc tccggaggaa	540
gggcggaggg aatgcggtgg tcattgcagt cggtaggagct cctgaagctt tggatgcaca	600
tcctggtgac tacacogttc acctggccaa taagaaaggc tttattaagc ttgcaataga	660
gcacggtgca gatttggttc ctatatattc gtttgagag aacgaggtgt ttgaccaggt	720
gcagaatcct cgtggcacat ggctgogata cattcaagag cgtctgcagc gcatcatggg	780
tgtttctctt cctctctttc atgogogagg agttttccag tacaccttcg gctgatgcc	840
ttacagaaag cccatcaaca ctgttggtgg aaggccaatt ccagtggaga aaaatgagaa	900
accaacagca gaagaattgg acgtctatca tcagcgctat atggatgaac tggcacggct	960
gttcgaggac cacaagggaa attatggagt acctgaagac actcatctgg tgttccagtg	1020
agatgatcac ttatggtgac gatgagtttg tgtgtttgta cttgtattt cttacattgt	1080
aggaacccaaa tgtcttcaca tatatagaca taccactaca tttgaccttg tggggacatt	1140
tgtttggttc taatgttcaa aatggctcat gaatcatgct gaatttaata tttagaaaat	1200
gtaaaaatgc cgattgtttt ctgtgagggg tgggtttagg ggtagaaaaa aaacccaaaa	1260
aaaaaaaaa aaaaaaaaaa aaaaaaa	1287

5 <210> 32  
 <211> 1268  
 <212> ADN  
 <213> *Drosophila melanogaster*  
 <400> 32

ES 2 640 100 T3

```

ccgagagccg ataagataa gtataagtct agtacagggc aatttgacca togctggctt      60
ccatatcatt cagtcgtggg agaaccgccg aaagagcagc tgggtggagt tttttgggca      120
tttttgtaac acagcagaca tgaaaatcga gtgggcacca ctgctgggtc ctctggaacg      180
ccgactgcag atactggtca cggccttttt cacctccatg ctgctgatac tattgtcagt      240
ttccttcctt ttggtagctg gatcactgat ctacggaggt cttttggtgc gtagtctgat      300
ggtaacttac ttggcctaag tctttgtgca ccacaagaaa acccaatccg ttgtggatgg      360
caatggctgg atgataaac gcaccaacct tttgcatcgc cactatcgtg attactttcc      420
cgtggagctg gtgaaaacag ccgaactgcc agctactaag aactacatct tggccagctt      480
tccccacgga attctgggca caggcattgg cattaacatg ggcttgaaa tctccaagtg      540
gctggagcta tcccccaag tgcgtcccaa actgggcaact ctggatcagc atttccatgt      600
tccgttcatg cgtgaggtcc tccgctgctg gggctctggtg tcagtgtcca aagaggcgct      660

gatccgtatg ctcagcaagt caaatgatcc caagcacaag gataatcggg atggtttcac      720
ctccaatgog gtggccattc tggttggcgg tgcccagaaa gccatggact ctcatcctgg      780
gcagtacatt ttaacottga agaataggaa aggcttcgtg cgaatggcca ttagaacggg      840
ctcatcgatt gttccttcat tttcctttgg agaggtggac attttcgatc aggtggcaaa      900
tcccccaaac tcgctgctcc gacggtttca ggactttgtc aagaagctca cgggagtctc      960
tccgctgatt cctgtgggcc gggattctt caactacacc tttggttcc tcccattccg     1020
acgaocgatt gtccaagttg ttggtgctcc catcgatgtt gttaagaacg agcaccaga     1080
ctcggagtat gtggataaag tgcatggaca ggtcattgag tcgctggaga agttattcga     1140
tcagtacaaa gacaagtact tggagaattc gaagagtgcc actctagttg tacactagtt     1200
ttaaatcatg ttgatacaaa atattcctat gtattgaaat agaaacgagc tttttaattg     1260
aatttaaa                                     1268

```

<210> 33  
 <211> 1269  
 <212> ADN  
 <213> *Drosophila melanogaster*  
 <400> 33

5



ES 2 640 100 T3

```

agttaaatgt tgtacgctcg catcgctcgt tctgctgctc aaacgcgata acgatccatt      60
actcgccgga ggaccgtgct accaagtagt tcaactttcc gagttcgaga gctctgacat      120
catgaaaaac gagtgggctc ccaaaggagt tcccatggaa cggcgtcgcc agacgtttgc      180
catggccttc ctaatcctat cctttatgat actctccttt ggatcctatt tcttctgtgc      240
tgcogtctg ttctatggaa gccttttctg gcgcaccatt atggtcactc atttggctca      300
cgtctacgcg aatcacaaga gaaccactc cattatggat ggcaatggct ggaaaattaa      360
ccgcaacaat tggctgttcc ggcattatcg ggactacttt cccgtacagc tggtaagac      420
cgcagagctg ccgccgaata agaactacat tttggccagt tttcgcacg gaattttggg      480
gactggcatt tccattaata tgggtctaga catctccaag tggctgcaac tgtcccaca      540
ggtgaggccc aagggtggcca ctttggatca aaacttctct acgcccactc tgcgtggtct      600
cctgagatct tggggcctgg tgtcggtttc caaggaggcc ttagtatatc tgctaaccaa      660
atcaaacgat cccaagcaca aggacaatcg tgatgggttc acctccaacg cggtggtat      720
tctggtgggt ggtgcccagc aggcactaga ctgcaccct ggcaaatata tattgacact      780
gaagaaccgc aagggtctcg ttaaaatggc cattaggaca ggttcttcca ttgtcccac      840
gttttccttt ggcgaggtgg acattctgga tcaggttgcc aatccaccga actctcgggt      900
tcgtcgtctt caagactttg taaagaggat aacggggata tctcogctga ttccogttgg      960
ccggggcctc ttcaactact cctttggctt tctgcccac cgtcgacgca ttgtccaagt     1020

tgttggcgt cccatcgacg tggttcagag cgatcaacca gacgcagcct atgtggataa     1080
gatacacaaa caggttattg acgacctgga gaagatgttc gccaaagtaca aggatcagta     1140
cataccgaac tccaagcagg acaagctgat tatacactag tcgagaagtt attgtctaaa     1200
caatgtgatc aatgagtgc tttataagca tttaatataa aaattacgaa ttttatcagg     1260
caaataaaa                                     1269

```

<210> 34  
 <211> 1139  
 <212> ADN  
 <213> *Drosophila melanogaster*  
  
 <400> 34

5

ES 2 640 100 T3

atgacaatcg aatgggctcc tctccgggtt cggctggaac ggcggcttca gaccctggtt	60
acgtcattct tcaacctatac cttctttaog ctgcccatct cgtcctgcct cgcagtagct	120
atcctgctgt actacggcga aatgtttggt cgtagctctgc ttcttatcta ttttgtgaaa	180
athtatttgg attataaaaag aaattacggc attatggaag gtaatggttg gttattctac	240
cgcagcaatt ggagatatcg aaactacttc cccgtcgaac ttgtaaaaac cgcagtagctg	300
cctoccaaaca ggaactatat cgtagcaagc tttccacacg gaatacttgg tactggaact	360
tgcattaaca tgagcctgga catcgacaat tggctatccc tttaccogca cgttcgacct	420
aagatcgcca ccttgatca tcaactcaag agcctttcc tgcgtgacat attgcgctgg	480
tggggcatgg tatcggttcc caaggaatcc ctgtcctatt tactcagcaa gtcaaatgat	540
ccaatgcaca aggataatcg ggatggtttc acatccaatg cggtgccctg acttgttggg	600
ggcggcaagg aagccatgga ttgcgcatcc ggacagtaca ttttgacct caagataga	660
aagggtttcg tcaaaatggc tgttcgaacc ggatcgtcga ttgtgccctc gttgtccttt	720
ggcgaggtag acatatttga tcaggtggct aatcctcgg actcctcgt cgcgccttc	780
cagaatgtgg tcaagaaatt cacaggcatt tcaccgctcc ttccaaggg tcgcggtatc	840
ttcaactaca actatggcat tctcccacat cgacgacgta ttgtccaagt tgttggttcc	900
cccatgatg tggaaagatg cgagactccc gatcccgaa atgtggacaa gatccatggg	960
caggtgatcg acgcgctggc gaggatgttc gatgaatata aagagaagta cactocaaac	1020
tcaaagcata tcaaacttat aatacagtga ctcagtagag tcccttagat tgtgattttt	1080
tatagaatta aggaaacttt tgaaacgcaa cacagtttta gactaattta aaaaaaaaa	1139

<210> 35  
 <211> 1971  
 <212> ADN  
 <213> *Anopheles gambiae*

5

<400> 35

ES 2 640 100 T3

cgcatgagga gaaaagttgg ccactcagtt attatacggg acaagctccc agtataagcg 60  
 tcattgtag atcgattct acaatgacaa acggaccoga catcattagt catgtttggc 120  
 cggtaggtcag aagaaagcac agagggtgga agaaaagctg attcattggg tgaggacggt 180  
 agtctgoggg cgtgttttta tegtctcagt ttgcatcatg ccccagtga agtgataccg 240  
 acacgacatt tggatgaagcc attcggttca agattaaggg acaggctaaa ccagtacaag 300  
 tgagtattcg ttgattcatt cgtaaaccga acgttgtttg ctcagcaaga agcaatgact 360  
 gataagtgat aacgtaattg catcatcggc atcaatgctg ataagatctc gagctgoggg 420  
 aaggaaaacg catcaagtgg tacatcgttt gtgggtgat agattttggg agggcgagtc 480  
 gtagtggtgg aacatcccc ggtaccggtt ctagacgggc tcacccaaac ctaagatggg 540  
 taaaatcgaa tgggccccac tgaatgtgcc gatgagggcg cggatcgaaa ccctgtccac 600  
 tgcctgtgg atgtggtga tactgttcgg cgagctgggc atgctgatat cgtacttcc 660  
 gctgctgac tatggtaact tgtttattaa atogctgtgc gtcactacg gatacttcat 720  
 ctacactgat cgaagtgga ctttgaacgg gggccgtgga cagggtgtaa catggtggcg 780  
 agagctgttc tggatgaagc tctaccagag ctacttccc gcaagctgc acaaacgg 840  
 tgatctcgac cctaaccgga actacctgtt tggccgctc cgcacggcg tgcctggcct 900  
 cggtgccctc atcaactttg ccaccaatgc gaccggctt cagacaagt ttcccaaat 960  
 ccgttccgc ccggtgacgc taaacttoca ctttgtgata ccgttcttc gogagctgct 1020  
 gctcagctgg ggcctcgtgt cggccaacc gaacagcatc ctgagcctgc tgaaggcgcc 1080  
 gaacaaaccg gaccaccgc tgaacgacga cggctacacg gccaacggg tctgatag 1140  
 ggtgggoggt cgggcccgaat cgctgcaactg ccggccgaac aactacacgc tgggtgctgcg 1200  
 caagcggaag ggcttctgca agctggcgat caaggcggg accccgctgg tgcgggtgat 1260  
 gagcttccgc gaggtcgacc tgttcgatca gccgcccgaat ccgcccggct cggggttgcg 1320  
 ccggttccag gagtttgtga aaaacacgac cggcatcggc ccggccgct tctcggggcg 1380  
 gggcttcttc cagtacagct acgggttgat cccgcccgg aagccactga acacagctcgt 1440  
 cggtgctccg gtggaggtga cgcaaataga caaccgacg caggaacagg tggatgaggt 1500  
 gcacgagcgg ttctgcccgc cgctggacaa tctgttcgaa acgaacaaat cgcgcttcat 1560  
 tgccgactac aagaatgtga agcttgtcat ggagtaatag gtaaatgca gatacttccc 1620  
 gtaactccc ccacttttac acacaaacga tcgggtgata gcgggtggaa aacggggcaa 1680  
 aatcacgcac aagagacaca gaaacacact atcatcacga taccctacgc aaccogtgg 1740  
 aattgtgta taattcgcat cgtaaggatt gagactgtac caaaaaaaaa aaaaaaacga 1800  
 acgatcaaaa ccaactctt tcatacctaa gatccaacca ttgtgoggtg ccaacaggtg 1860  
 atattgttat acagctctgc gagagggagt caattactag ctgcgtaaga gaggacacac 1920  
 tacattggga gtgttagca tttaaattct cgtgtaagt taattgtgta a 1971

5 <210> 36  
 <211> 1612  
 <212> ADN

ES 2 640 100 T3

<213> *Anopheles gambiae*

<400> 36

```

gtcacattct aaccttacgc ccgtgctgaa cccagggga gaccagtgcg cgctccactg      60
tgaactccct agtgtgtagt gtattctttt cctgttgttt tagatgtatt aagggcgagt      120
cgtagtggtg gaacatcccc cggtagccgtg tctagacgcg ctcacaaaaa cctaagatgg      180
gtaaaatcga atgggcccc actgaatgtgc cgatgaggcg gcggatcgaa accctgtcca      240
ctgccctgtg gatgtggctg atactgttcg gcgagctggg catgctgata tcgtacttcc      300
tgctgctgat ctatggtaac ttgtttatta aatcgctgtg cgtcactctac ggatacttca      360
tctacactga tcgaaaagtg actttgaacg ggggccgtgg acagggtgta acatggtggc      420
gagagctggt ctggtggaag ctctaccaga gctacttccc ggcgaaactg cacaaaacgg      480
ttgatctcga ccctaaccgg aactacctgt ttgccgcctt cccgcaacggc gtgcttgccc      540
tcggtgcctt catcaacttt gccaccaatg cgaccgctt tcacgacaag ttcccaaaa      600
tccgttcccg ccoggtgacg ctaaacttcc actttgtgat accgttcttc cgcgagctgc      660
tgctcagctg gggcctcgtg toggccaacc cgaacagcat cctgagcctg ctgaaggcgc      720
cgaacaaaacc ggaccaccog ctgaacgacg acggctacac ggccaacgcg gtcgtgatag      780
tggtggggcg tgcggccgaa tcgctgcaact gccggccgaa caactacacg ctggtgctgc      840
gcaagcggaa gggcttctgc aagctggcga tcaaggcggg caccocgctg gtgccgggta      900
tgacgttcgg cgaggtcgac ctgttcgatc agccgccgaa tccgcccggc tcgcggttgc      960
gccggttcca ggagtttgtg aaaaacacga ccggcatcgc gccggccgcc ttcgtcgggc     1020
ggggcttctt ccagtacagc tacgggttga tcccgccgog gaagccactg aacacagtog     1080
tcggtgctcc ggtggagggt acgcaaatac acaaccogac gcaggaacag gtggatgagg     1140
tgcacgagcg gttctgcgcg gcgctggaca atctgttcga aacgaacaaa tcgcgcttca     1200
ttgccgacta caagaatgtg aagcttgtca tggagtaata ggtaaaatgc agatacttcc     1260
cgtacttccc gccatcttta cacacaaacg atcggtgtat agcggtgga aaacggggca     1320
aatcacgca caagagacac agaaacacac tatcatcaog atatcctacg caaccogtgg     1380
aaattgtgtt ataattogca tcgtaaggat tgagactgta ccaaaaaaaaa aaaaaaacg     1440
aacgatcaaa accaactctt ttcataccta agatccaacc attgtgcggg gccaacaggt     1500
gatattgta tacagctctg cgagagggag tcaattacta gctgcgtaag agaggacaca     1560
ctacattggg agtgtttagc atttaaatc tcgtgtaagt gtaattgtgt aa              1612

```

5

<210> 37

<211> 1017

<212> ADN

10 <213> *Tribolium castaneum*

<400> 37

ES 2 640 100 T3

atgaagattt tgggatcaa gttcgccct ttgcacatcc ccctcgagcg ccgcttgcaa 60  
 accottgccg ccggatgctg gttcacaacc ctggcctttg gtaccttcat cggtaacctc 120  
 atctgggtgt acgtcttctt cacccgatcc tggtaacctg ccgtccttta cgccaccata 180  
 atatacctcg agaaaagcaa gtgtgagaag ggcggccgcc ccatcgaatg gatcagacat 240  
 tggggatggt ggtactactt gaaaaactac ttcccctgta agcttgactt tgtcccgggc 300  
 ctaacctttg atccgaagag gaattatcta tttgcctggt acccccatgg gatcctcccg 360  
 gcgggccctt tcaacacccat tggtagccct tacagogaat ttagcaagtt gtttccaaag 420  
 tttcgggtga ggttggatgat tctgcaccaa catttcttca ttccgtttct gggggagatt 480  
 gcttatggga caggggggat ttcggcctcg gccaaagtcgt tgaatcatgt tttggggagt 540  
 ccggaagggg ggtatattgc ggtggtgatg ccggggggcg cagttgaggc ctataacagc 600  
 cggcctggac agtacagaat catcctgaag aacaggaaag gcttcgtaa actggccttg 660  
 agaatgggt caccactagt accagttatc togttgggg agccggaatt gttgatcaa 720  
 gttgaaggga ggactttgag gaagattcag gattccatca ggaagtacct cgggctggca 780  
 cctgtgattt tctcaggag agggttcttc cagtactcgt ttggcgttat tccacaaagg 840  
 aggcctatta cgaccgttgt cgggcacccg attgaagtta ccaaatcga aaaaccaaca 900  
 aacgaggaag tggacgaact acataaaaag tacatgcaag agctggagaa cttgttcgaa 960  
 gagtacaaat tcaaatattt ggagaatccg aaggacatac acttgaggt tgaataa 1017

<210> 38  
 <211> 1420  
 <212> ADN  
 <213> *Homo sapiens*

5

<400> 38

cactcacaca cctacggaca cacgctactc tgggaggtga tttgcgactt agccaggccc 60  
 ccaaagctgg gctcctgtag ggagaaagtc tgcccaggtc cacatccaag ccttcatcgt 120  
 ttgtcctccg ggttctggga tctgtctgga agaggggagc ttctgcaatg ggagttgcca 180  
 caacctgca gcccccaacc acttccaaaa ccttgcagaa gcagcatcta gaagcagtg 240  
 gcgcctacca atatgtgctc actttctctt tcatgggccc tttcttctcc cttcttctct 300  
 ttgtctctct cttoacgtca ctctggocct tctctgtttt ttacttggtg tggctctatg 360  
 tggactggga cacaccaac caaggtggaa ggcgttcgga gtggataag aaccgggcaa 420

10

ES 2 640 100 T3

tttggagaca actaagggat tattatcctg tcaagctggt gaaaacagca gagctgcccc 480  
 cggatcggaa ctacgtgctg ggcgcccacc ctcatgggat catgtgtaca ggcttctct 540  
 gtaatttctc caccgagagc aatggcttct cccagctctt cccggggctc cggccctggt 600  
 tagccgtgct ggctggcctc ttctacctcc cgggtctatcg cgactacatc atgtcctttg 660  
 gactctgtcc ggtgagccgc cagagcctgg acttcatcct gtcccagccc cagctcgggc 720  
 aggcctggt catcatggtg gggggtgctc acgaggccct gtattcagtc cccggggagc 780  
 actgccttac gctccagaag cgcaaaggct tcgtgctcct ggcgctgagg cacggggcgt 840  
 ccctggtgcc cgtgtactcc tttggggaga atgacatctt tagacttaag gcttttgcca 900  
 caggctcctg gcagcattgg tgccagctca ccttcaagaa gctcatgggc ttctctcctt 960  
 gcatottctg gggctcgggt ctcttctcag ccacctcctg gggcctgctg ccotttctgtg 1020  
 tgccatcac cactgtggtg ggccgcccac tcccgtccc ccagcgcctc caccaccgcg 1080  
 aggaggaagt caatcactat cacgcccctc acatgacggc cctggagcag ctcttcgagg 1140  
 agcacaagga aagctgtggg gtcccgcctt ccacctgcct caccttcac taggcctggc 1200  
 cgcgcccttt cgctgagccc ctgagcccaa ggcactgaga cctccacca ctgtggactc 1260  
 catgcctcca ataaaaggta gttctgggcc cagcgcagtg cctcgtgcct gtgatcccag 1320  
 cactttggga ggccagggtg ggaggatcgt ttgagcccag gagtgaaga ccagcctggg 1380  
 caacacagtg agacttcatt tctacaaaaa aaaaaaaaaa 1420

<210> 39  
 <211> 1934  
 <212> ADN  
 <213> *Pan troglodytes*

5

<400> 39

ccaactcacac acctacggac acacgctact ctgggaggtg atttgcgact tagccaggcc 60  
 cccaaagctg ggctcctgta gggagaaagt ctgccaggt ccacatccaa gccttcatcg 120  
 tttgtcctct gggttctggg atcttgctgg aagaggggag cttctgcaat gggagttgcc 180  
 acaaccctgc tgcccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg 240  
 ggcgcctacc aatatgtgct cactttctc ttcatgggc cttctctctc ccttcttctc 300  
 tttgtcctcc tcttcacgtc actctggccc ttctctgttt tttacttgggt gtggctctat 360  
 gtggactggg acacacccea ccaagggtgga aggcgttcgg agtggataag gaactgggca 420  
 atttgagac aactaaggga ttattatcct gtcaagctgg tgaaaacggc agagctgccc 480  
 cggatcggga actacgtgct ggcgcccac cctcatggga tcatgtgtac aggttctctc 540  
 tgtaatttct ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg 600  
 ttagccgtgc tggctggcct cttctacctc ccggctctac gcgactacat catgtccttt 660

10

ES 2 640 100 T3

ggactctgtc ccgtgagccg ccagagcctg gacttcatcc tgtcccagcc ccagctcggg 720  
 caggccgtgg tcatcatggt tgggggtgcg catgaggccc tgtattcagt ccccggggag 780  
 cactgcctaa cgctccagaa gcgcaaaggc ttcgtgcgcc tggcgctgag gcacgggtggg 840  
 ccgccccatc cctgtcccc agcgcctoca cccaccgag gaggaagtca atcactatca 900  
 cgcoctctac atgatggccc tgcagcagct cttcgaggag cacaaggaaa gctgtggggg 960  
 ccccgcttcc acctgcctca ccttcatota ggottggcgg cggcctttcg ctgagccoct 1020  
 gagcccaagg cactgagacc tccaccact gtggactcca tgcctccaat aaaaggtagt 1080  
 tctgggceca gcgcagtgcc tcacgcctgt aatcccagca ctttgggagg ccaagggtggg 1140  
 aggatcgttt gagcccagga gttgaagacc agcctgggca acatagtgag acttcatttc 1200  
 tacaaaaatt aaaaaataat gttgttaatt agccgggcat ggtggcatgt gcctgtaatc 1260  
 ccagctattht gggaggctga ggcaggagac ttgcttgaac ccaggaggcg gaggttgcag 1320  
 tgagctgaga tcacaccaca gtactccggc ctgggcaaca cagcgagact caatctcaaa 1380  
 aaaaaaaaaa aaaaaaaaaa aagatggtta aagctatagt aggctgggcg tgggtgctca 1440  
 tgcccataat cccaacactt tgggaagcca aggcagggtg atcacctgag gtcaggagtt 1500  
 caagaccagc ctggccaaaa tggcaaaacc ccctctctac taaaactaca aaaattagcc 1560  
 gggcatggtg gctcatgcct ataatcccag gtactccgga ggctgaggca ggagaatcgc 1620  
 ttgaaccoag gaggcagagg ttgcagtgag ccgagatcac accactgcac tccagtctgg 1680  
 gcaacagagc gagattccat ctcaaaaaaa aaaaaaaat gttgttaaag ctataatcat 1740  
 tctggaagtg aacatggcca ggcattgacct gccccatcc tgatgtgcta caaagctgct 1800  
 gccttagatt gggatcagg accctgaact gtcctcccag tgagagttca gctctgaggca 1860  
 tgacctgga caagcctctt ccttttgcct gtttacttct ttgtctgtaa ataaaggcat 1920  
 ctgcttagca aaga 1934

<210> 40  
 <211> 2223  
 <212> ADN  
 <213> *Pan troglodytes*

5

<400> 40

ccactcacac acctacggac acacgtact ctgggagggtg atttgcgact tagccaggcc 60  
 cccaaagctg ggctcctgta gggagaaagt ctgccagggt ccacatccaa gccttcatcg 120  
 tttgtcctct gggttctggg atcttgctgg aagaggggag cttctgcaat gggagttgcc 180  
 acaaccctgc tgccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg 240  
 ggcgcctacc aatatgtgct cactttcctc ttoatgggoc ctttctctc ccttctgtc 300  
 tttgtcctcc tcttcacgtc actctggccc ttctctgttt tttacttggg gtggctctat 360

10

ES 2 640 100 T3

gtggactggg acacacccaa ccaaggtgga aggcgttcgg agtggataag gaactgggca 420  
 atttggagac aactaagga ttattatcct gtcaagctgg tgaaaacggc agagctgccc 480  
 ccggatcggg actacgtgct gggcgccac cctcatggga tcatgtgtac aggcttcctc 540  
 tgtaatttct ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg 600  
 ttagccgtgc tggctggcct ettctacctc ccggtctatc gcgactacat catgtccttt 660  
 ggactctgtc ccgtgagccg ccagagcctg gacttcatcc tgtcccagcc ccagctcggg 720  
 caggccgtgg tcatcatggt tgggggtgcg catgaggccc tgtattcagt ccccggggag 780  
 cactgcctaa cgctccagaa gcgcaaaggc ttcgtgcgcc tggcgctgag gcacggggcg 840  
 tccttggctc ccgtgtactc ctttggggag aatgacatct ttagacttaa ggcttttgcc 900  
 acaggctcct ggcagcattg gtgccagctc accttcaaga agttcatggg cttctctcct 960  
 tgcattctct ggggtcgtgg tctcttctca gccacctcct ggggcctgct gcccttctct 1020  
 gtgcccatca ccactgtggg tgagtgccca cctccggggg gacggccacc agcagctgcg 1080  
 tgggcatcag ggatccctcg cccacctgtc tccctctccc tgcagtgggc cgccccatcc 1140  
 ctgtcccca ggcctccac cccaccgagg aggaagtcaa tcaactcac gccctctaca 1200  
 tgatggccct gcagcagctc ttcgaggagc acaaggaaag ctgtggggtc cccgcttcca 1260  
 cctgcctcac cttcatctag gcttggccgc ggcctttcgc tgagcccctg agcccaaggc 1320  
 actgagacct ccaccactg tggactccat gcctccaata aaaggtagtt ctgggcccag 1380  
 cgcagtgcct cagcctgta atcccagcac tttgggaggg caaggtggga ggatcgtttg 1440  
 agcccaggag ttgaagacca gcctgggcaa catagtgaga cttcatttct acaaaaatta 1500  
 aaaaataatg ttgttaatta gccgggcatg gtggcatgtg cctgtaatcc cagctatttg 1560  
 ggaggctgag gcaggagact tgcttgaacc caggaggcgg aggttgacgt gagctgagat 1620  
 cacaccacag tactccggcc tgggcaacac agcgagactc aatctcaaaa aaaaaaaaaa 1680  
 aaaaaaaaaa agatgggtaa agctatagta ggctgggcgt ggtggctcat gcccataatc 1740  
 ccaacacttt ggggaagcaa ggcaggtgga tcacctgagg tcaggagttc aagaccagcc 1800  
 tggccaaaat ggcaaaacc cctctctact aaaactaaa aaattagccg ggcattggtg 1860  
 ctcatgccta taatcccagg tactccggag gctgaggcag gagaatcgct tgaaccagg 1920  
 aggcagaggt tgcagtgagc cgagatcaca ccactgcact ccagctcggg caacagagcg 1980  
 agattccatc tcaaaaaaaaa aaaaaaaaaatg ttgttaaagc tataatcatt ctggaagtga 2040  
 acatggccag gcatgacctg cccccatcct gatgtgctac aaagctgctg ccttagattg 2100  
 ggtatcagga cctgaactg tctcccagc gagagttcag tctgtggcat gacctggac 2160  
 aagcctcttc cttttgectg tttacttctt tgtctgtaaa taaaggcacc tgcttagcaa 2220  
 aga 2223

<210> 41  
 <211> 2137



ES 2 640 100 T3

<212> ADN  
 <213> *Pan troglodytes*

<400> 41

5

```

ccactcacac acctacggac acaogctact ctgggaggtg atttgcgact tagccaggcc      60
cccaaagctg ggctcctgta gggagaaagt ctgccaggt ccacatcca gccttcatcg      120
tttgtcctct gggttctggg atcttgctgg aagaggggag cttctgcaat gggagttgcc      180
acaaccctgc tgcccccaac cacttccaaa accctgcaga agcagcatct agaagcagtg      240
ggcgcctacc aatatgtgct cactttcctc ttcattgggc ctttcttctc ccttcttgtc      300
tttgtcctcc tcttcacgtc actotggccc ttctctgttt tttacttggg gtggctctat      360
gtggactggg acacacccaa ccaaggtgga aggcgttcgg agtggataag gaactgggca      420
atctggagac aactaaggga ttattatcct gtcaagctgg tgaaaacggc agagctgcc      480
ccggatcgga actacgtgct gggcgcccac cctcatggga tcatgtgtac aggcttcctc      540
tgtaatttct ccaccgagag caatggcttc tcccagctct tcccggggct ccggccctgg      600
ttagccgtgc tggctggcct cttctacctc ccggtctatc gcgactacat catgtccttt      660
ggactctgtc ccgtgagcgg ccagagcctg gacttcatcc tgtcccagcc ccagctcggg      720
caggccgtgg tcatcatggt tgggggtgcg catgaggccc tgtattcagt ccccggggag      780
cactgcctaa cgctccagaa gcgcaaaggc ttogtgcgcc tggcgtgag gcacggggcg      840
tccttggtgc ccgtgtactc ctttggggag aatgacatct ttagacttaa ggcttttgcc      900
acaggctcct ggcagcattg gtgccagctc accttcaaga agttcatggg cttctctcct      960
tgcattctct ggggtcgtgg tctcttctca gccacctcct ggggectgct gccctttgct     1020
gtgoccatca ccactgtggt gggcgcgcc atccctgtcc ccagcgcct ccaccccacc     1080
gaggaggaag tcaatcacta tcacgcctc tacatgatgg cctgcagca gctcttcgag     1140
gagcacaagg aaagctgtgg ggtccccgct tccacctgcc tcaccttcat ctaggcttgg     1200
ccgcgccctt tcgctgagcc cctgagccca aggcactgag acctccaccc actgtggact     1260
ccatgcctcc aataaaagg agttctgggc ccagcgcagt gcctcacgcc tgtaatccca     1320
gcactttggg aggccaagg gggaggatcg tttgagccca ggagttgaag accagcctgg     1380
gcaacatagt gagacttcat ttctacaaaa attaaaaaat aatggtgta attagccggg     1440
catggtggca tgtgcctgta atccagcta tttgggaggc tgaggcagga gacttgcttg     1500
aaccagagag gcggagggtg cagtgagctg agatcacacc acagtactcc ggocctgggca     1560
acacagcgag actcaatctc aaaaaaaaa aaaaaaaaa aaaaagatgg ttaaagctat     1620
agtaggctgg gcgtggtggc tcatgccc atcccaaca ctttgggaag ccaaggcagg     1680
    
```

ES 2 640 100 T3

tggatcacct gaggtcagga gttcaagacc agcctggcca aaatggcaaa accccctctc	1740
tactaaaact acaaaaatta gcogggcatg gtggetcatg cctataatcc caggtactcc	1800
ggaggctgag gcaggagaat cgcttgaacc caggaggcag aggttgcagt gagccgagat	1860
cacaccactg cactccagtc tgggcaacag agcgagattc catctcaaaa aaaaaaaaaa	1920
aatgttgta aagctataat cattctggaa gtgaacatgg ccaggcatga cctgccccca	1980
tcctgatgtg ctacaaagct gctgccttag attgggtatc aggaccctga actgtcctcc	2040
cagtgagagt tcagtctgtg gcatgacott ggacaagcct cttccttttg cotgtttact	2100
tctttgtctg taaataaagg catctgotta gcaaaga	2137

<210> 42

<211> 1749

5

<212> ADN

<213> *Canis familiaris*

<400> 42

ES 2 640 100 T3

atgacgttcc ctgaccccaa gacgggacag tgctgcctct ctccggtcct gtccgggcct 60  
 gggggcggcc ggggtggcct gctggccctg ggagggcccc tggaagcgct ggaagcaaaa 120  
 cccggcgcgc tgttttgcgg atccggaatc agaagagatt tgtccagctg gcaccgggac 180  
 acggagcgag ctcttcagc cgttcgogaa ccctccggac ccttgggcct ccttctgccc 240  
 ttccgggagc gcccatccac ccggtccggg agtcccagcc tctgccccct cggctcccct 300  
 gcgcgcacca gctccgcagc gctcgggcca gcgcgctccc cgcgccggg cgggaacccc 360  
 gcgctgctgc ccacgcctgg gcacaaccggc tcacctctt tgacactccc agtccctgac 420  
 tctgtcgaga cggctgcgcg tggcaatgaa gaggccgggc ggtgccgggg atcccgttct 480  
 gcgcgcatgg cgcggctcac gcgctgttc cgggagcgca aggcgcgagc cggcctcccg 540  
 ggccaccgac acctcccgcc cgtgtccctg aggtgccgac actcccggg acctgcacac 600  
 ggagccgggg cgcgcaccag gggcctgccc ggccacgctg ggccccggag gcgtggagcg 660  
 tttggtcgct ccagcaagaa agccttcaact ccgtgcccaa accttcgtcg tttgcgctcc 720  
 aggccttgga acccactgg cggaggggag cgcgccacca tgagaaccac gcagaaacag 780  
 tggctagaag ccttaagcgt ctctattac gtgttcaact tcctcttcat gggccttttc 840  
 ttctcccttc ttgtctctt cctctcttc acttcattct ggtccatctc tgttctctac 900  
 ttggtgtggc tcttctgga ctgggacaca cccaaccaag gtgggaggtg ctttgagtgc 960  
 aacaggagct gtaccatttg gaaacacttg aaggattatt ttcctattaa gctggggaag 1020  
 acagcagagc ttccccctga ccggaactat gttgtgggtt gtcaccaca tgggatcatg 1080  
 tgcattggaa ccttotgcaa tttcttcagc gaggccaaca acttctcaa gcagttcccg 1140  
 gggattcaga catcaccgct aacactggcc tttctcttgc acctccagc ctatcgtgac 1200  
 tacctcatgt accttgact gtgttctgtg aaccgcgga gcctggactt tattttgtcc 1260  
 aagccccagc ccgggcagc ggtagccatc gtagttgggg gtgcccaaga gtccctgttt 1320  
 gcgatcccag ggatgcactg cctcgttctc cggaaacgca aaggcttctg gcgcctggca 1380  
 ctgaggcacg gagcttccct ggtacctgtg tactcctttg gggagaatga tatcttcaat 1440  
 ttttaaggctt ttcccacaaa ctctggcag tatctgtgcc agatcaccat caagaagatc 1500  
 atgaaatctt ctcttgcct cttctggggc cgcggactct tctcagctga ctctgggggt 1560  
 ttactgcctc tcgccaagcc tatcaacct gtggtgggcc gcccatccc tgtgccacag 1620  
 cgcctcaacc cactgagga ggaagttgac cactatcaca tgctctacat ggaggctcta 1680  
 gagcaactgt ttgaggaaca caaggaaagc tgtggtgtcc cagcctctac tcacctcctc 1740  
 tttaaatag 1749

<210> 43  
 <211> 1221  
 <212> ADN  
 <213> *Bos taurus*

ES 2 640 100 T3

<400> 43

```

atgataagtg agagcagact ggagacctct ggggctctga tgagagagtc tccccacgca      60
cacagccccc tacaagcgat cctgggaggc tcccgaagcgc ttggccactc agaccagaaa      120
gatttcctgc gctccaacca aagccttcat cccatTTTTc ctggttcgga agccctgctg      180
gctgaagggg agcacctggg agtctccacc accccaccac ccacccccag catgaaaacc      240
ttaaagaaac agtggttaga agtactgagc acctaccaat atgtgctctg tttcctcttc      300
ctgggcccctt tcttttccct tgttggcttc ttctcctct tcaacctgct ctggtatctc      360
tctgttctct acttggtatg gttattcctg gactgggaca caccocagca aggtggaagg      420
cgtaatcagt ggttgaagaa ctgcaactgtg tggaaacacc tgagcgatta tttccccatt      480
aagctggtga aaacagtaga gctgccccca gacaggaact acgtgctatt gtcccacca      540
catgggatca tgggctttgg aaccgtctgt aacttcagca cggagggcac tggctgctcg      600
cagctgttcc cagggttog gttctcaactg gccgtgttga attgtctctt gtacgtgcoa      660
ggctgtcgag agtacattat gtccctgtgga acatgttctg tgaaccgtca gagcctggat      720
tatgttctat ctcaaccaca gctcggcagg gctgtggtca tcatggtcgg aggggccaat      780
gaggccctgc ataccgtccc aggggagcac tgctcactc tccggaatcg taaaggcttc      840
gtccgcctgg cactgaggca cggtgccctc ctggtgcccg tgtactcctt tggggagaa      900
gacgtcttca gagttaaggc ttttgcccca gactcctggc agcgtctggt ccaggtcacc      960
atcaagaggc tcctgagctt ctctccttgc atcttctggg gccgtggtct cttctcagcc     1020
aagtctggg gccctgatgc cctcggcaga cccatcacca ctgtggtggg ccgccccatc     1080
ccggtgcccc agtgtccaca gccacccgag gagcaggtgg accactatca caggctctac     1140
atgaaggctc tggagcaact gtttgaggag cacaagaaga gctcgggcct cccggttct     1200
actcacctca ccttcatcta g                                             1221

```

5 <210> 44  
 <211> 1872  
 <212> ADN  
 <213> *Danio rerio*

10 <400> 44

ES 2 640 100 T3

acacacagtc gctccaccct catacagacg tcgcacgagc gaggagcata cattcaactca 60  
 aactcaaggg atatttaaaa gatatttccc gacttaatga cttttctag ctcaagcagg 120  
 tattcgcggc gtcttctagt caccagtcg tgacagcaaa gactcgaatt ctttctctgt 180  
 ttacaaatga ttcacgttaa agcaaagcgt cgtttacagg gggttcagtc atgcggacgc 240  
 ggattcaaaa gtggctggag ataaacctca gatagccagt ccagaaacta cccacataca 300  
 tttgcaatga cgcgataaga caacgaatcg ttatctgttt gaccaattga accgatcaaa 360  
 gtgcgagcca gacaaataca agacttcggg cagtgaaggt gaggaagagg attcgtttgc 420  
 cacaggatca ctaaacttgg ccactttaat gggaacagat aacgtttcag aggtaaagga 480  
 gaagggggac aggtcaccat ggaagatat aatcgaggat ataagtgttt tgcaattggt 540  
 gttgtcattt ctttttttag gagtggcttg tctactattg atgatctacc tgatgttcac 600  
 ctctctgtgg attttcccc ctctgtactt cacctggcag atctatgatt ggcacacacc 660  
 tgagagaggt ggcagaagaa caaagtttgt gagaggctgg gaagtatgga aacacttgcg 720  
 agactathtt cctgtgaagc tggtaaagac tgctgaatta aatccaata agaactatat 780  
 tatgggctgt caccgcacg gcatcatgtg ttttgagcc ttctcctgct tcagcacaga 840  
 ccgcaatgga tttgcggaaa ctttctctgg aatagcgtcc actcttgcaa tcttagccgg 900  
 acttttccgc cttctctct tcaggaata cattttagct gcaggattgc tgctgtcag 960  
 taaggcaagt ttggattatt tgctgagcca aactgggtgtg ggtaacgctg tggatcat 1020  
 cattgggtgga gcagaggaat cactgacctc ctctacagga gtaaacactg tggtcataaa 1080  
 acacaggaaa ggctttgtac gactggctct tgaacacggg gcagacctgg ttccagtgt 1140  
 ctcttttggga gagaacgagc tgttcccgca ggtggttctg tctgagggca gtgtaggtcg 1200  
 acggctgcag gttcttttta agcagattat gggttttgct ccatgcatct tcacaggagg 1260  
 ccgggtggctt cttctgccgt acaaacttcc agtcaccact gtgggtgggt gcccataaa 1320  
 cgtccctctg gtgaaaatc caactcagga acaggtggat cattaccacg gtttatacat 1380  
 ggcatctctg gctgatcttt tccacaagca taagaccagc tatggactgg cagagacgca 1440  
 cgagctgcac ttcatttagc agtgcctgaa agcacaagc catttaaaa gctgactggg 1500  
 atataatgtg gggctgatat aggaaaagga caatgcactt atttctgtac tgttccatgt 1560  
 gaatgtttta acttcttggg tctctgaaat gattcacaaa gcacaaaatg aaccaaatga 1620  
 atacggagat gggcatttgt ctgtaacaaa aactcatcc ttaagaaagg caaaaggttt 1680  
 gcttacgggtt atctcagctt tgttctgtaa atttgcactg aaatttttac aacatgttag 1740  
 tgtttattat aaaggttata caaagacttt tgaagaaaat catttttaa aattggttta 1800  
 atataaatga tttacatatg tattttgtta tgtctattca tgtgaaagat tataaaacca 1860  
 gcacattttt ac 1872

- 5 <210> 45
- <211> 335
- <212> PRT
- <213> *Mus musculus*
- <400> 45

ES 2 640 100 T3

Met Met Val Glu Phe Ala Pro Leu Asn Thr Pro Leu Ala Arg Cys Leu  
 1 5 10 15

Gln Thr Ala Ala Val Leu Gln Trp Val Leu Ser Phe Leu Leu Leu Val  
 20 25 30

Gln Val Cys Ile Gly Ile Met Val Met Leu Val Leu Tyr Asn Tyr Trp  
 35 40 45

Phe Leu Tyr Ile Pro Tyr Leu Val Trp Phe Tyr Tyr Asp Trp Arg Thr  
 50 55 60

Pro Glu Gln Gly Gly Arg Arg Trp Asn Trp Val Gln Ser Trp Pro Val  
 65 70 75 80

Trp Lys Tyr Phe Lys Glu Tyr Phe Pro Ile Cys Leu Val Lys Thr Gln  
 85 90 95

Asp Leu Asp Pro Gly His Asn Tyr Ile Phe Gly Phe His Pro His Gly  
 100 105 110

Ile Phe Val Pro Gly Ala Phe Gly Asn Phe Cys Thr Lys Tyr Ser Asp  
 115 120 125

Phe Lys Lys Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Ala Lys  
 130 135 140

Ile Trp Phe Cys Phe Pro Leu Phe Arg Glu Tyr Leu Met Ser Asn Gly  
 145 150 155 160

ES 2 640 100 T3

Pro Val Ser Val Ser Lys Glu Ser Leu Ser His Val Leu Ser Lys Asp  
 165 170 175

Gly Gly Gly Asn Val Ser Ile Ile Val Leu Gly Gly Ala Lys Glu Ala  
 180 185 190

Leu Glu Ala His Pro Gly Thr Phe Thr Leu Cys Ile Arg Gln Arg Lys  
 195 200 205

Gly Phe Val Lys Met Ala Leu Thr His Gly Ala Ser Leu Val Pro Val  
 210 215 220

Phe Ser Phe Gly Glu Asn Asp Leu Tyr Lys Gln Ile Asn Asn Pro Lys  
 225 230 235 240

Gly Ser Trp Leu Arg Thr Ile Gln Asp Ala Met Tyr Asp Ser Met Gly  
 245 250 255

Val Ala Leu Pro Leu Ile Tyr Ala Arg Gly Ile Phe Gln His Tyr Phe  
 260 265 270

Gly Ile Met Pro Tyr Arg Lys Leu Ile Tyr Thr Val Val Gly Arg Pro  
 275 280 285

Ile Pro Val Gln Gln Ile Leu Asn Pro Thr Ser Glu Gln Ile Glu Glu  
 290 295 300

Leu His Gln Thr Tyr Leu Glu Glu Leu Lys Lys Leu Phe Asn Glu His  
 305 310 315 320

Lys Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Phe Lys  
 325 330 335

<210> 46  
 <211> 334  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 46

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Leu Ala Arg Arg Leu  
 1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Ser Phe Leu Thr Gly Pro  
 20 25 30

Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Asn Tyr Leu Phe  
 35 40 45

10

ES 2 640 100 T3

Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr Pro  
 50 55 60

Glu Arg Gly Gly Arg Arg Ser Ser Trp Ile Lys Asn Trp Thr Leu Trp  
 65 70 75 80

Lys His Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Gln Asp  
 85 90 95

Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly Ile  
 100 105 110

Met Ala Val Gly Ala Phe Gly Asn Phe Ser Val Asn Tyr Ser Asp Phe  
 115 120 125

Lys Asp Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro Leu  
 130 135 140

Trp Phe Trp Cys Pro Val Phe Arg Glu Tyr Val Met Ser Val Gly Leu  
 145 150 155 160

Val Ser Val Ser Lys Lys Ser Val Ser Tyr Met Val Ser Lys Glu Gly  
 165 170 175

Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Lys Glu Ser Leu  
 180 185 190

Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile Arg Gln Arg Lys Gly  
 195 200 205

Phe Val Lys Ile Ala Leu Thr His Gly Ala Ser Leu Val Pro Val Val  
 210 215 220

Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Thr Asp Asn Pro Glu Gly  
 225 230 235 240

Ser Trp Ile Arg Thr Val Gln Asn Lys Leu Gln Lys Ile Met Gly Phe  
 245 250 255

Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Asn Phe Gly  
 260 265 270

Leu Met Thr Tyr Arg Lys Ala Ile His Thr Val Val Gly Arg Pro Ile  
 275 280 285

Pro Val Arg Gln Thr Leu Asn Pro Thr Gln Glu Gln Ile Glu Glu Leu  
 290 295 300



ES 2 640 100 T3

His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His Lys  
 305 310 315 320

Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Leu Lys  
 325 330

5  
 <210> 47  
 <211> 334  
 <212> PRT  
 <213> *Pan troglodytes*  
 <400> 47

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Pro Ala Arg Arg Leu  
 1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Ser Phe Leu Thr Gly Pro  
 20 25 30

Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Ser Tyr Ser Phe  
 35 40 45

Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr Pro  
 50 55 60

Glu Arg Gly Gly Arg Arg Ser Ser Trp Ile Lys Asn Trp Thr Leu Trp  
 65 70 75 80

Lys His Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Gln Asp  
 85 90 95

Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly Ile  
 100 105 110

Met Ala Val Gly Ala Phe Gly Asn Phe Ser Val Asn Tyr Ser Asp Phe  
 115 120 125

Lys Asp Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro Leu  
 130 135 140

Trp Phe Trp Cys Pro Val Phe Arg Glu Tyr Val Leu Ser Val Gly Leu  
 145 150 155 160

Val Ser Val Ser Lys Lys Ser Val Ser Tyr Met Val Ser Lys Glu Gly  
 165 170 175

Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Lys Glu Ser Leu  
 180 185 190

10

ES 2 640 100 T3

Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile His Gln Arg Lys Gly  
 195 200 205

Phe Val Lys Thr Ala Leu Thr His Gly Ala Ser Leu Val Pro Val Val  
 210 215 220

Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Thr Asp Asn Pro Glu Gly  
 225 230 235 240

Ser Trp Ile Arg Thr Val Gln Asn Lys Leu Gln Lys Ile Met Gly Phe  
 245 250 255

Ala Leu Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Asn Phe Gly  
 260 265 270

Leu Met Pro Tyr Arg Lys Ala Ile His Thr Val Val Gly Arg Pro Ile  
 275 280 285

Pro Val Arg Gln Thr Leu Asn Pro Thr Gln Glu Gln Ile Glu Glu Leu  
 290 295 300

His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His Lys  
 305 310 315 320

Gly Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Val Leu Lys  
 325 330

<210> 48  
 <211> 335  
 <212> PRT  
 <213> *Pan troglodytes*  
 <400> 48

5

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Gln Pro Ala Arg Arg Leu  
 1 5 10 15

Gln Thr Val Ala Val Leu Gln Trp Val Leu Lys Tyr Leu Leu Leu Gly  
 20 25 30

Pro Met Ser Ile Gly Ile Thr Val Met Leu Ile Ile His Ser Tyr Ser  
 35 40 45

Phe Leu Tyr Ile Pro Tyr Leu Met Trp Leu Tyr Phe Asp Trp His Thr  
 50 55 60

Pro Glu Arg Gly Gly Arg Arg Ser Ser Trp Ile Lys Asn Trp Thr Leu  
 65 70 75 80

10

ES 2 640 100 T3

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

ES 2 640 100 T3

<210> 49  
 <211> 335  
 <212> PRT  
 <213> *Canis familiaris*

5

<400> 49

Met Lys Val Glu Phe Ala Pro Leu His Ile Pro Leu Ala Arg Arg Leu  
 1 5 10 15

Gln Thr Ala Ala Val Leu Gln Trp Val Leu Ser Phe Leu Leu Leu Ala  
 20 25 30

Gln Val Cys Ile Gly Ile Ile Val Val Leu Ile Val His Asn Tyr Trp  
 35 40 45

Phe Leu Tyr Val Pro Tyr Leu Thr Trp Leu Cys Leu Asp Trp Arg Thr  
 50 55 60

Pro Glu Gln Gly Gly Arg Arg Ser Asn Trp Val Arg Ser Trp Thr Val  
 65 70 75 80

Trp Arg Tyr Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Trp  
 85 90 95

Asp Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly  
 100 105 110

Val Leu Val Ala Gly Ala Phe Gly Asn Phe Cys Thr Asn His Ser Asp  
 115 120 125

Phe Glu Glu Leu Phe Pro Gly Phe Thr Ala Tyr Leu His Val Leu Pro  
 130 135 140

Phe Trp Phe Arg Cys Pro Leu Phe Arg Glu Tyr Leu Met Thr Ser Gly  
 145 150 155 160

Ser Val Ser Val Ser Lys Lys Ser Val Ser Tyr Val Leu Ser Lys Glu  
 165 170 175

Gly Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Glu Glu Ser  
 180 185 190

Leu Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile Arg Gln Arg Lys  
 195 200 205

Gly Phe Val Lys Val Ala Leu Thr His Gly Ala Ser Leu Val Pro Val  
 210 215 220

ES 2 640 100 T3

Phe Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Val Asn Asn Pro Glu  
 225 230 235 240

Gly Ser Trp Leu Arg Thr Val Gln Glu Lys Leu Gln Lys Ile Met Gly  
 245 250 255

Phe Ala Leu Pro Leu Phe His Ala Arg Gly Ile Phe Gln Tyr Asn Phe  
 260 265 270

Gly Leu Met Pro Tyr Arg Lys Pro Ile His Thr Val Val Gly Arg Pro  
 275 280 285

Ile Pro Val His Arg Thr Pro His Pro Ser Pro Gly Gln Ile Glu Glu  
 290 295 300

Leu His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Ala His  
 305 310 315 320

Lys Arg Lys Tyr Gly Ile Pro Glu His Glu Thr Leu Ile Phe Lys  
 325 330 335

<210> 50  
 <211> 335  
 <212> PRT  
 <213> *Bos taurus*  
 <400> 50

5

Met Lys Val Glu Phe Ala Pro Leu Asn Ile Pro Leu Ala Arg Arg Leu  
 1 5 10 15

Gln Thr Ala Ala Val Leu His Trp Leu Leu Ser Phe Leu Leu Phe Ala  
 20 25 30

Gln Val Cys Leu Gly Ile Ile Val Phe Leu Ile Ile Tyr Asn Tyr Trp  
 35 40 45

Phe Leu Tyr Leu Pro Tyr Leu Thr Trp Leu Tyr Phe Asp Trp Gln Thr  
 50 55 60

Pro Glu Gln Gly Gly Arg Arg Ser Glu Trp Val Arg Asn Trp Ala Ile  
 65 70 75 80

Trp Arg Tyr Phe Lys Asp Tyr Phe Pro Ile His Leu Ile Lys Thr Trp  
 85 90 95

Asp Leu Asp Pro Ser His Asn Tyr Ile Phe Gly Phe His Pro His Gly  
 100 105 110

10

ES 2 640 100 T3

Val Leu Val Val Gly Ala Phe Gly Asn Phe Cys Thr Asn Tyr Ser Ala  
 115 120 125

Phe Lys Glu Leu Phe Pro Gly Phe Thr Ser Tyr Leu His Val Leu Pro  
 130 135 140

Tyr Trp Phe Arg Cys Pro Leu Phe Arg Glu Tyr Leu Met Ser Ser Gly  
 145 150 155 160

Pro Val Ser Val Ser Lys Lys Ser Val Cys His Val Leu Ser Lys Glu  
 165 170 175

Gly Gly Gly Asn Ile Ser Val Ile Val Leu Gly Gly Ala Glu Glu Ser  
 180 185 190

Leu Asp Ala His Pro Gly Lys Phe Thr Leu Phe Ile Arg Gln Arg Lys  
 195 200 205

Gly Phe Val Lys Ile Ala Leu Thr His Gly Ala Tyr Leu Val Pro Val  
 210 215 220

Phe Ser Phe Gly Glu Asn Glu Leu Phe Lys Gln Val Ser Asn Pro Glu  
 225 230 235 240

Gly Ser Trp Leu Arg Asn Val Gln Glu Lys Leu Gln Lys Ile Met Gly  
 245 250 255

Phe Ala Leu Pro Leu Phe His Ala Arg Gly Ile Phe Gln Tyr Asn Phe  
 260 265 270

Gly Leu Ile Pro Tyr Arg Lys Pro Ile His Thr Val Val Gly Arg Pro  
 275 280 285

Ile Pro Val Arg Gln Thr Leu Asn Pro Thr Ser Glu Gln Ile Glu Glu  
 290 295 300

Leu His Gln Thr Tyr Met Glu Glu Leu Arg Lys Leu Phe Glu Glu His  
 305 310 315 320

Lys Gly Lys Tyr Gly Ile Pro Glu Asn Glu Thr Leu Ile Phe Arg  
 325 330 335

<210> 51  
 <211> 183  
 <212> PRT  
 <213> *Rattus norvegicus*

5

<400> 51

ES 2 640 100 T3

Met Ser Lys Glu Gly Gly Gly Asn Ile Ser Val Ile Val Ile Gly Gly  
 1 5 10 15

Ala Lys Glu Leu Leu Glu Ser Phe Pro Gly Arg Tyr Ser Leu Cys Leu  
 20 25 30

Leu Gln Arg Lys Gly Phe Val Lys Ile Ala Leu Thr His Gly Ala His  
 35 40 45

Leu Val Pro Val Phe Ser Phe Gly Glu Asn Glu Leu Tyr Ser Gln Val  
 50 55 60

Asp Asn Pro Lys Gly Ser Trp Leu Arg Thr Ala Gln Asp Lys Val Tyr  
 65 70 75 80

Asn Leu Thr Gly Leu Ala Leu Pro Leu Phe Tyr Ala Arg Gly Ile Phe  
 85 90 95

Gln Asn Ser Phe Gly Leu Met Pro Tyr Arg Lys Leu Ile Tyr Thr Val  
 100 105 110

Ala Ala Val Gly Thr Ala Ala Arg Leu Thr Gln Lys Ser Asn His Gln  
 115 120 125

Asn Asp Ser Ser Val Ser Ala Glu Lys Gly Asn Ser Phe Ile Cys Cys  
 130 135 140

Gln Val Gly Trp Cys Asp Arg Asp Arg Gln Pro Leu Ala Ala Pro Ser  
 145 150 155 160

Leu Phe Arg Arg Leu Cys Thr Arg Pro Gln Ser Arg Leu Met Ser Cys  
 165 170 175

Ile Arg Arg Thr Trp Arg Ser  
 180

<210> 52  
 <211> 287  
 <212> PRT  
 <213> *Danio rerio*

<400> 52

Met Leu Gly Ala Leu Tyr Ala Gly Trp Leu Tyr Leu Asp Arg Asp Thr  
 1 5 10 15

Pro Ser Cys Gly Gly Arg Arg Ser Gln Trp Val Arg Ser Trp Arg Ile  
 20 25 30

ES 2 640 100 T3

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



# ES 2 640 100 T3

**275**

**280**

**285**

5  
<210> 53  
<211> 349  
<212> PRT  
<213> *Caenorhabditis elegans*  
  
<400> 53

ES 2 640 100 T3

Met Leu Asp Arg Thr Glu Ser Glu Asp Lys Met Pro His Leu Leu Gly  
 1 5 10 15

Val Glu Trp Ala Pro Leu Asn Ile Pro Leu Ala Arg Arg Leu Gln Thr  
 20 25 30

Leu Gly Ala Leu His Phe Phe Phe Ile Thr Leu Phe Thr Pro Val Leu  
 35 40 45

Val Leu Thr Val Pro Phe Tyr Met Leu Tyr Thr Val Leu Trp Pro Leu  
 50 55 60

Ile Phe Leu Tyr Gly Leu Trp Met Ile Tyr Asp Trp Asn Ser Pro Lys  
 65 70 75 80

Lys Gly Ala Tyr Met Ser Asn Trp Phe Gln Arg Gln Arg Ile His Ser  
 85 90 95

Trp Tyr Ala Asn Tyr Phe Pro Val Lys Leu His Thr Thr Ser Asp Met  
 100 105 110

Pro Glu Glu His Asn Tyr Leu Ile Gly Tyr His Pro His Gly Ile Ile  
 115 120 125

Ser Met Ala Ala Phe Ile Asn Phe Ala Thr Asn Gly Thr Gly Ile Leu  
 130 135 140

Asp Thr Leu Pro Arg Ile Arg Phe His Leu Cys Thr Leu Val Gly Gln  
 145 150 155 160

Phe Trp Thr Pro Trp Arg Arg Glu Trp Gly Leu Leu His Gly Met Ile  
 165 170 175

Asp Cys Ser Arg Glu Ser Ile Lys His Val Leu Glu His Glu Lys Lys  
 180 185 190

Gly Lys Ala Val Val Leu Val Val Gly Gly Ala Glu Glu Ala Leu Asp  
 195 200 205

Ala His Pro Gly Cys His Ile Leu Thr Leu Lys Lys Arg Lys Gly Phe

ES 2 640 100 T3

210	215	220																	
Val	Lys	Ile	Ala	Leu	Gln	Thr	Gly	Ala	Gln	Leu	Val	Pro	Cys	Tyr	Ser				
225					230					235					240				
Phe	Gly	Glu	Asn	Asp	Ile	Phe	Asn	Gln	Ala	Glu	Asn	Pro	Lys	Gly	Ser				
				245					250					255					
Thr	Ile	Arg	Gln	Phe	Gln	Thr	Ile	Met	Lys	Arg	Val	Leu	Gly	Phe	Ser				
			260					265					270						
Pro	Pro	Ala	Phe	Tyr	Gly	Arg	Gly	Val	Phe	Asn	Tyr	Thr	Phe	Gly	Leu				
		275					280					285							
Leu	Pro	Phe	Arg	Lys	Pro	Ile	Asn	Thr	Val	Leu	Gly	Ala	Pro	Ile	Ser				
290						295					300								
Val	Thr	Lys	Thr	Val	Asn	Pro	Thr	Gln	Glu	Gln	Ile	Asp	Thr	Leu	His				
305					310					315					320				
Gln	Thr	Tyr	Met	Asp	Arg	Leu	His	Glu	Leu	Phe	Glu	Glu	His	Lys	Thr				
				325					330					335					
Lys	Tyr	Asp	Val	Ser	Pro	Thr	Thr	Gln	Leu	Val	Ile	Asn							
			340					345											

<210> 54  
 <211> 339  
 <212> PRT  
 <213> *Caenorhabditis elegans*  
 <400> 54

5

ES 2 640 100 T3

Met Pro His Leu Leu Gly Val Glu Trp Ala Pro Leu Asn Ile Pro Leu  
 1                                      5                                      10                                      15

Ala Arg Arg Leu Gln Thr Leu Gly Ala Leu His Phe Phe Phe Ile Thr  
                                     20                                      25                                      30

Leu Phe Thr Pro Val Leu Val Leu Thr Val Pro Phe Tyr Met Leu Tyr  
                                     35                                      40                                      45

Thr Val Leu Trp Pro Leu Ile Phe Leu Tyr Gly Leu Trp Met Ile Tyr  
                                     50                                      55                                      60

Asp Trp Asn Ser Pro Lys Lys Gly Ala Tyr Met Ser Asn Trp Phe Gln  
 65                                      70                                      75                                      80

Arg Gln Arg Ile His Ser Trp Tyr Ala Asn Tyr Phe Pro Val Lys Leu

## ES 2 640 100 T3

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

Val Ile Asn

5  
 <210> 55  
 <211> 340  
 <212> PRT  
 <213> *Caenorhabditis elegans*  
 <400> 55

```

Met Pro Gln Phe Leu Gly Ile Glu Trp Val Asp Leu Phe Ser Ser Ile
1           5           10           15

Gln Arg Lys Lys Thr Tyr Leu Gly Val Val Tyr His Phe Met Leu Thr
           20           25           30

Tyr Pro Leu Ala Leu Phe Val Thr Ile Leu Pro Phe Phe Leu Leu Phe
           35           40           45

Thr Phe Gln Trp His Ile Leu Ala Leu Tyr Ala Cys Trp Tyr Phe Tyr
50           55           60

Asp Met Asp Ser Pro Arg Arg Gly Gly Tyr Ser Ser Asp Trp Val Arg
65           70           75           80

Lys Trp Arg Val Asn Asp Trp Phe Ala Gln Tyr Phe Pro Ile Asn Leu
           85           90           95

His Lys Thr Ala Glu Leu Ser Thr Asp Lys Asn Tyr Leu Val Gly Ile
100          105          110

His Pro His Gly Ile Ile Ser Met Ala Ala Trp Ser Asn Phe Ala Thr
115          120          125

Asn Gly Thr Gly Ile Tyr Glu Lys Phe Pro Gly Ile Arg Trp Asn Leu
130          135          140

Cys Thr Leu Ala Leu Gln Phe Arg Met Ala Ile Arg Arg Glu Leu Leu
145          150          155          160

Leu Leu Thr Gly Leu Ile Asp Cys Ser Arg Glu Ser Ile Glu Tyr Val
165          170          175

Leu Asp Lys Cys Gly Gln Lys Gly Arg Ala Val Val Leu Val Ile Gly
180          185          190

Gly Ala Glu Glu Ala Leu Asp Ala His Pro Gly Tyr His Thr Leu Thr
195          200          205
    
```

10

ES 2 640 100 T3

Leu Ala Ser Arg Lys Gly Phe Val Arg Glu Ala Leu Ile Thr Gly Ala  
 210 215 220

Tyr Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Val Phe Glu Gln  
 225 230 235 240

Met Glu Asn Pro Val Gly Ser Arg Leu Arg Asn Phe Gln Glu Trp Cys  
 245 250 255

Lys Ser Ile Phe Gly Ile Ser Tyr Pro Ile Phe His Gly Arg Gly Phe  
 260 265 270

Phe Gln Leu Thr Phe Gly Tyr Leu Pro Phe Arg Lys Pro Ile Asp Thr  
 275 280 285

Val Val Gly Ala Pro Ile Pro Val Glu Lys Val Glu Asn Pro Thr Lys  
 290 295 300

Glu Gln Ile Asp Glu Leu His Thr Ile Tyr Cys Gln Lys Leu Thr Glu  
 305 310 315 320

Leu Phe Asp Glu His Lys Glu Lys Tyr Gly Val Glu Lys Asp Val Pro  
 325 330 335

Leu Val Leu Arg  
 340

<210> 56  
 <211> 345  
 <212> PRT  
 <213> *Caenorhabditis elegans*

5

<400> 56

Met Leu Asn Tyr Gln Ile His Lys Lys Leu Thr Asp Ile Lys Trp Val  
 1 5 10 15

Asn Ile Phe Ser Pro Trp Asp Arg Gln Arg Ala Tyr Phe Ala Leu Val  
 20 25 30

Val Trp Phe Gly Leu Ile Tyr Pro Phe Cys Cys Leu Cys Gln Val Ala  
 35 40 45

Pro Phe Val Leu Phe Phe Thr Gly Gln Trp Ile Ile Leu Gly Leu Tyr  
 50 55 60

Ala Val Trp Tyr Leu Tyr Asp Arg Glu Ser Pro Arg Arg Gly Gly Tyr  
 65 70 75 80

10

ES 2 640 100 T3

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



Ser Lys Asp Thr Arg Leu Val Phe Gln  
 340 345

5 <210> 57  
 <211> 359  
 <212> PRT  
 <213> *Caenorhabditis elegans*  
 <400> 57

Met Arg Leu Arg Leu Ser Ser Ile Ser Gly Lys Ala Lys Leu Pro Asp  
 1 5 10 15

Lys Glu Ile Cys Ser Ser Val Ser Arg Ile Leu Ala Pro Leu Leu Val  
 20 25 30

Pro Trp Lys Arg Arg Leu Glu Thr Leu Ala Val Met Gly Phe Ile Phe  
 35 40 45

Met Trp Val Ile Leu Pro Ile Met Asp Leu Trp Val Pro Phe His Val  
 50 55 60

Leu Phe Asn Thr Arg Trp Trp Phe Leu Val Pro Leu Tyr Ala Val Trp  
 65 70 75 80

Phe Tyr Tyr Asp Phe Asp Thr Pro Lys Lys Ala Ser Arg Arg Trp Asn  
 85 90 95

Trp Ala Arg Arg His Val Ala Trp Lys Tyr Phe Ala Ser Tyr Phe Pro  
 100 105 110

Leu Arg Leu Ile Lys Thr Ala Asp Leu Pro Ala Asp Arg Asn Tyr Ile  
 115 120 125

Ile Gly Ser His Pro His Gly Met Phe Ser Val Gly Gly Phe Thr Ala  
 130 135 140

Met Ser Thr Asn Ala Thr Gly Phe Glu Asp Lys Phe Pro Gly Ile Lys  
 145 150 155 160

Ser His Ile Met Thr Leu Asn Gly Gln Phe Tyr Phe Pro Phe Arg Arg  
 165 170 175

Glu Phe Gly Ile Met Leu Gly Gly Ile Glu Val Ser Lys Glu Ser Leu  
 180 185 190

Glu Tyr Thr Leu Thr Lys Cys Gly Lys Gly Arg Ala Cys Ala Ile Val  
 195 200 205

10

ES 2 640 100 T3

Ile Gly Gly Ala Ser Glu Ala Leu Glu Ala His Pro Asn Lys Asn Thr  
 210 215 220

Leu Thr Leu Ile Asn Arg Arg Gly Phe Cys Lys Tyr Ala Leu Lys Phe  
 225 230 235 240

Gly Ala Asp Leu Val Pro Met Tyr Asn Phe Gly Glu Asn Asp Leu Tyr  
 245 250 255

Glu Gln Tyr Glu Asn Pro Lys Gly Ser Arg Leu Arg Glu Val Gln Glu  
 260 265 270

Lys Ile Lys Asp Met Phe Gly Leu Cys Pro Pro Leu Leu Arg Gly Arg  
 275 280 285

Ser Leu Phe Asn Gln Tyr Leu Ile Gly Leu Leu Pro Phe Arg Lys Pro  
 290 295 300

Val Thr Thr Val Met Gly Arg Pro Ile Arg Val Thr Gln Thr Asp Glu  
 305 310 315 320

Pro Thr Val Glu Gln Ile Asp Glu Leu His Ala Lys Tyr Cys Asp Ala  
 325 330 335

Leu Tyr Asn Leu Phe Glu Glu Tyr Lys His Leu His Ser Ile Pro Pro  
 340 345 350

Asp Thr His Leu Ile Phe Gln  
 355

<210> 58  
 <211> 412  
 <212> PRT  
 <213> *Kluyveromyces lactis*

<400> 58

Met Gly Lys Glu Gln Ile Arg Lys Arg Gln Thr Lys Lys Val Arg Lys  
 1 5 10 15

Gly Ser Gln Pro Val Ala Asn Gly Lys Ser Ser Val Val Gly Val Ser  
 20 25 30

Lys Asp Gly Lys Pro Glu Phe Cys Ala Ile Asp Thr Pro Ile Glu Arg  
 35 40 45

Arg Leu Gln Thr Leu Ser Val Ala Trp Tyr Val Glu Cys Ile Pro Leu  
 50 55 60

5

10

ES 2 640 100 T3

Met Ile Ile Ile Met Leu Phe Val Trp Val Asn Pro Leu Met Trp Ser  
65 70 75 80

Phe Val Ile Pro Tyr Thr Ile Tyr Tyr Phe Ile Asp Arg Thr Ala Ser  
85 90 95

Asn Gly Asn Ala Val Lys Arg His Ser Lys Trp Phe Arg Ser Leu Lys  
100 105 110

Val Trp Phe Tyr Phe Arg Asp Tyr Phe Pro Ile Ser Met His Lys Ser  
115 120 125

Thr Glu Leu Glu Pro Thr Phe Thr Ser Ile Asp Ser Thr Glu Leu Glu  
130 135 140

Asn Asp Ala Ser Glu Pro Gly Tyr Leu Asp Ser Ser Gln Pro Val Leu  
145 150 155 160

Pro Asp Lys Trp Trp Asn Pro Phe Arg Glu Lys Asp Glu Thr Val Arg  
165 170 175

Pro Thr Gly Pro Arg Tyr Ile Phe Gly Tyr His Pro His Gly Ile Ala  
180 185 190

Ala Phe Gly Ala Phe Gly Ala Phe Ala Thr Glu Ala Cys Asn Trp Ser  
195 200 205

Lys Val Phe Pro Gly Ile Pro Val Cys Leu Leu Thr Leu Val Asn Gln  
210 215 220

Phe Gln Ile Pro Val Tyr Arg Asp Tyr Leu Leu Ala Leu Gly Ile Thr  
225 230 235 240

Ser Val Ala Arg Lys Asn Ala Met Lys Val Leu Glu Lys Asn Tyr Ser  
245 250 255

Ile Ala Ile Val Ile Gly Gly Ala Ser Glu Ser Leu Leu Thr Asn Leu  
260 265 270

Gly Ser Ser Asp Ile Ile Leu Ser Lys Arg Lys Gly Phe Val Lys Leu  
275 280 285

Ala Leu Gln Thr Gly Asn Val Ser Leu Val Pro Val Tyr Gly Phe Gly  
290 295 300

Glu Thr Asp Thr Tyr Lys Ile Leu Lys Leu Lys Asn Asp Ser Ile Ile  
305 310 315 320

Gly Arg Ile Gln Ile Trp Leu Lys Glu Asn Tyr Ser Phe Thr Val Pro  
 325 330 335

Leu Phe Phe Ala Arg Gly Val Phe Asn Tyr Asp Phe Gly Leu Leu Pro  
 340 345 350

Phe Arg His Pro Val Asn Val Val Val Gly Asn Pro Ile His Ile Lys  
 355 360 365

Glu Lys Ile Asp His Pro Thr Ile Glu Glu Ile Asp His Tyr His Ser  
 370 375 380

Leu Tyr Ile Glu Glu Leu Lys Arg Leu Tyr Asp Asp Asn Lys Ala Lys  
 385 390 395 400

Phe Asn Tyr Ser Glu Lys Thr Leu Asn Ile Val Glu  
 405 410

<210> 59  
 <211> 461  
 <212> PRT  
 <213> *Ashbya gossypii*

5

<400> 59

Met Gln Asp Ser Met Asp Asp Ser Leu Arg Glu Ala Glu Gly Arg Gln  
 1 5 10 15

Asp Asp Ser Glu Val Ser Ser Gly Thr Thr Leu Gly Ser Ser Thr Pro  
 20 25 30

Glu Asp Ser Gly Val Thr Ala Lys Leu Arg Lys Lys Tyr Gln Met Ala  
 35 40 45

Ser Ala Leu Leu Arg Arg Glu Leu Glu Glu Leu Ser Val Tyr Asp Ala  
 50 55 60

Lys Thr Ala Gly Val Ser Gly Arg Ser Ser Gly Ser Gly Ser Gly Gly  
 65 70 75 80

Leu Ala Leu Leu Gly Gly Arg Phe His Val Ala Pro Leu Arg Ile Pro  
 85 90 95

Ala Arg Arg Arg Leu Gln Thr Leu Val Val Ala Trp His Thr Ser Ser  
 100 105 110

Phe Ile Tyr Met Thr Val Leu Val Leu Phe Leu Ala Ala Asn Pro Leu  
 115 120 125

10

ES 2 640 100 T3

Met Trp Trp Phe Met Val Pro Tyr Met Val Tyr Tyr Val Trp Asn Arg  
130 135 140

Ser Pro Ala Asn Gly Gly Val Val Arg Arg Tyr Ser Pro Arg Leu Arg  
145 150 155 160

Ser Leu Ala Leu Trp Arg Tyr Tyr Cys Glu Tyr Tyr Pro Ile Ser Leu  
165 170 175

His Lys Ser Glu Asp Leu Ala Pro Thr Phe Val Pro Asp Pro Arg Gly  
180 185 190

Ala Glu Pro Arg Glu Trp Lys Leu Arg Leu Trp Leu Trp Pro Thr Arg  
195 200 205

Val Glu Leu Leu Asn Leu Thr Leu Gln Trp Thr Arg Ala Arg Pro Gln  
210 215 220

Val Ala Thr Gly Pro Arg Tyr Ile Phe Gly Tyr His Pro His Gly Val  
225 230 235 240

Gly Ala Leu Gly Ala Phe Gly Ala Ile Ala Thr Glu Gly Cys Asn Trp  
245 250 255

Ser Lys Val Phe Ala Gly Ile Pro Ala Cys Leu Cys Thr Leu Val Asn  
260 265 270

Gln Phe Gln Ile Pro Ile Tyr Arg Asp Tyr Leu Leu Gly Leu Gly Cys  
275 280 285

Thr Ser Val Ala Arg Lys Asn Val Leu Lys Val Leu Glu Gln Asn Tyr  
290 295 300

Ser Val Cys Ile Val Val Gly Gly Ala Gln Glu Ala Leu Leu Ser Arg  
305 310 315 320

Val Gly Ser Thr Glu Leu Val Leu Asn Lys Arg Lys Gly Phe Ile Lys  
325 330 335

Leu Ala Leu Glu Thr Gly Asn Val Asn Leu Val Pro Ile Tyr Ala Phe  
340 345 350

Gly Glu Thr Asp Cys Phe Asn Val Leu Asp Thr Gly Asn Glu Ser Tyr  
355 360 365

Leu Arg Lys Phe Gln Leu Trp Ile Lys Lys Thr Tyr Gly Phe Thr Ile

ES 2 640 100 T3

```
370              375              380

Pro Phe Phe Phe Ala Arg Gly Val Phe Asn Tyr Asp Phe Gly Phe Leu
385              390              395              400

Pro Phe Arg Asn Pro Ile Asn Val Val Val Gly Lys Pro Val Tyr Val
              405              410              415

Asp Lys Arg Arg Thr Asn Pro Thr Met Glu Glu Ile Asp His Tyr His
              420              425              430

Asp Leu Tyr Val Gln Glu Leu Arg Asn Val Phe Asp Lys Asn Lys His
              435              440              445

Lys Phe Gly Tyr Ala Gly Lys Glu Leu Lys Ile Val Glu
              450              455              460
```

- 5
- <210> 60
  - <211> 550
  - <212> PRT
  - <213> *Magnaporthe oryzae*
  - <400> 60

ES 2 640 100 T3

Met Ala Thr Val Asp Asp Thr Gln Asn His Val Gln Leu Asn Ala Glu  
 1 5 10 15

Ser Ser Leu Asn Glu Phe Pro Tyr Gln Ala Ser Ala Asp Lys Glu Glu  
 20 25 30

Tyr Pro Ser Leu Pro Ser Ala Pro Leu Pro Asp Val Asp Gly Ser Ser  
 35 40 45

Gln Asn Gly His Thr Ala Ala Asn Ala Asp Leu Pro Pro Ser Glu Ser  
 50 55 60

Thr Glu Gln Pro Trp Asn Lys Val Ala Ser Gly Asp Phe Ile His Glu  
 65 70 75 80

Thr Lys Met Pro Ser Glu Asp Arg Asp Asn Leu Ala Arg Pro Gly Ser  
 85 90 95

Leu Pro Thr Pro Glu Pro Ser Arg Arg Gly Ser Asn Thr Leu Val Gly  
 100 105 110

Ser Pro Arg Gln Gly Asp Gly Leu Gln Met Leu Met Asp Glu Phe Ala  
 115 120 125

Asp Asp Glu Ser Ser Asp Thr Ser Trp Asp Ile His Glu Thr Ile Arg

ES 2 640 100 T3

130	135	140																			
Ser 145	Arg	Thr	Gln	Ala	Leu	Ala	Glu	Lys	Asp	His	Gln	Gly	Ile	Phe	Ser 160						
					150						155										
Asn	Gln	Arg	Arg	Gly	Asn	Gln	Ala	Asp	Ala	Pro	Asp	Thr	Val	Thr	Ser 175						
				165						170											
Ala	Asn	Ala	Trp	Leu	Pro	Ala	Gly	Phe	Glu	Asp	Asp	Leu	Asp	Glu	Val						
			180						185												
Thr	Gly	Ala	Asp	Asp	Glu	Phe	Glu	Pro	Thr	Arg	Gly	Ser	Arg	Trp	Arg						
		195						200						205							
Val	Arg	Leu	Ala	Pro	Leu	Asp	Thr	Pro	Phe	His	Arg	Arg	Met	Gln	Thr						
		210						215						220							
Leu	Val	Val	Leu	Leu	His	Val	Leu	Gly	Met	Gly	Ile	Thr	Phe	Ser	Phe 240						
				225						230						235					
Phe	Cys	Phe	Leu	Cys	Thr	Leu	Pro	Leu	Phe	Trp	Pro	Ile	Ile	Ile	Ala 255						
			245						250												
Tyr	Leu	Val	Phe	Ile	Arg	Leu	Ser	Arg	Ala	Gly	Ser	Asp	Gly	Lys	Thr						
			260						265						270						
Asn	Arg	Arg	Ile	Glu	Trp	Leu	Arg	Arg	Ala	Arg	Ile	Trp	Lys	Tyr	Phe						
		275						280						285							
Ala	Asp	Tyr	Phe	Pro	Val	Lys	Leu	His	Lys	Thr	Ala	Asp	Leu	Pro	Pro						
		290						295						300							
Thr	Arg	Lys	Tyr	Ile	Phe	Ala	Ile	His	Pro	His	Gly	Ile	Ile	Ser	His 320						
				305						310						315					
Gly	Ala	Phe	Ala	Ser	Phe	Ala	Thr	Glu	Ala	Leu	Gly	Phe	Ser	Glu	Lys						
				325						330											
Phe	Pro	Gly	Ile	Thr	Asn	Ser	Val	Cys	Thr	Leu	His	Gly	Asn	Phe	Lys						
			340						345						350						
Thr	Pro	Phe	Tyr	Arg	Glu	Tyr	Leu	Met	Ala	Met	Gly	Leu	Ile	Ser	Val						
		355						360						365							
Ser	Lys	Glu	Ser	Ile	Ile	Asn	Thr	Leu	Met	Thr	Gly	Gly	Ile	Asn	Gly						
		370						375						380							



ES 2 640 100 T3

Glu Gly Met Gly Arg Ala Val Ser Ile Ile Val Gly Gly Ala Arg Glu  
385 390 395 400

Ala Leu Glu Ala Cys Pro Lys Thr Met Arg Leu Ile Leu Lys Arg Arg  
405 410 415

Gly Phe Cys Arg Met Ala Leu Arg Thr Gly Ala Asp Leu Val Pro Val  
420 425 430

Leu Cys Phe Gly Glu Asn Asp Leu Tyr Gln Gln Trp Gly Pro Gln Asp  
435 440 445

His Pro Arg Phe Arg Arg Leu Gln Met Lys Ala Leu Ala Tyr Leu Gly  
450 455 460

Phe Ala Val Pro Val Leu Arg Gly Arg Gly Val Phe Asn Tyr Asp Phe  
465 470 475 480

Gly Val Leu Pro Gln Arg Arg Pro Ile Asn Val Val Val Gly Glu Pro  
485 490 495

Ile Lys Val Glu Gln Phe Arg Gly Ser Gly Asn Ile Glu Pro Arg Val  
500 505 510

Glu Glu Leu Trp Lys Leu Tyr Cys Gln Lys Leu Gln Glu Leu Tyr Asp  
515 520 525

Gln Asn Lys His Val Tyr Phe Lys Asp Arg Gln Glu Asp Met Arg Phe  
530 535 540

Val Asp Ser Glu Glu Leu  
545 550

<210> 61  
<211> 334  
<212> PRT  
<213> *Ciona intestinalis*

5

<400> 61

Met Trp Glu Phe Ala Pro Leu Arg Ile Pro Met Ile Arg Arg Leu Gln  
1 5 10 15

Thr Leu Ala Val Ser His Phe Val Phe Cys Phe Leu Ala Leu Ala Ala  
20 25 30

Thr Cys Thr Leu Leu Thr Ile Tyr Ile Thr Leu Phe Thr Ser Tyr Trp  
35 40 45

10

ES 2 640 100 T3

Pro Ile Thr Thr Ala Tyr Leu Val Phe Leu Phe Leu Asp Arg Asn Thr  
 50 55 60

Pro Glu Ser Gly Gly Arg Arg Ser Glu Trp Val Arg Asn Trp Ser Leu  
 65 70 75 80

Trp Lys Trp Met Ser Asp Tyr Phe Pro Cys Thr Leu His Lys Thr Val  
 85 90 95

Asp Leu Asp Pro Lys Arg Asn Tyr Ile Phe Gly Ile His Pro His Gly  
 100 105 110

Val Leu Cys Ile Gly Ser Phe Thr His Phe Ser Thr Asn Gly Ser Gly  
 115 120 125

Phe Ser His Val Phe Pro Gly Phe Thr Ser Tyr Leu Thr Met Leu Pro  
 130 135 140

Phe Trp Phe Lys Met Pro Phe Phe Arg Asp Tyr Val Met Ser Gly Gly  
 145 150 155 160

Leu Thr Pro Ala Thr Arg Lys Ala Ile Lys His Thr Ile Thr Arg Pro  
 165 170 175

Gly Gly Gly His Ile Cys Cys Ile Ile Pro Gly Gly Ala Pro Glu Ser  
 180 185 190

Leu Asn Ala Arg Pro Gly Asp Val Val Leu Leu Leu Lys Gln Arg Leu  
 195 200 205

Gly Phe Leu Lys Leu Ala Ile Thr Asn Gly Val Pro Leu Val Pro Val  
 210 215 220

Phe Ser Phe Gly Asp His Ala Leu Trp Glu Gln Lys Pro Asn Pro Pro  
 225 230 235 240

Gly Ser Leu Ile Arg Arg Phe Gln Asp Ser Ser Gln Lys Trp Met Gln  
 245 250 255

Val Ala Leu Pro Val Phe His Ala Arg Gly Ile Phe Gln Tyr Asn Phe  
 260 265 270

Gly Leu Ile Pro Tyr Arg Arg Ser Val His Thr Val Val Gly Glu Pro  
 275 280 285

Ile Glu Val Pro Gln Asn Ser Asn Pro Thr Ser Glu Asp Leu Met Ser  
 290 295 300

ES 2 640 100 T3

Leu Gln Glu Asp Tyr Ile Asn Arg Leu Arg Ala Ile Phe Asp Glu His  
 305 310 315 320

Lys Ser Lys Tyr Leu Pro Glu Asp Cys Lys Leu Ile Ile Asn  
 325 330

<210> 62  
 <211> 334  
 <212> PRT  
 <213> *Mus musculus*  
 <400> 62

5

Met Val Glu Phe Ala Pro Leu Leu Val Pro Trp Glu Arg Arg Leu Gln  
 1 5 10 15

Thr Phe Ala Val Leu Gln Trp Val Phe Ser Phe Leu Ala Leu Ala Gln  
 20 25 30

Leu Cys Ile Val Ile Phe Val Gly Leu Leu Phe Thr Arg Phe Trp Leu  
 35 40 45

Phe Ser Val Leu Tyr Ala Thr Trp Trp Tyr Leu Asp Trp Asp Lys Pro  
 50 55 60

Arg Gln Gly Gly Arg Pro Ile Gln Phe Phe Arg Arg Leu Ala Ile Trp  
 65 70 75 80

Lys Tyr Met Lys Asp Tyr Phe Pro Val Ser Leu Val Lys Thr Ala Glu  
 85 90 95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val  
 100 105 110

Leu Ala Ala Gly Ala Phe Leu Asn Leu Cys Thr Glu Ser Thr Gly Phe  
 115 120 125

Thr Ser Leu Phe Pro Gly Ile Arg Ser Tyr Leu Met Met Leu Thr Val  
 130 135 140

Trp Phe Arg Ala Pro Phe Phe Arg Asp Tyr Ile Met Ser Gly Gly Leu  
 145 150 155 160

Val Ser Ser Glu Lys Val Ser Ala Asp His Ile Leu Ser Arg Lys Gly  
 165 170 175

Gly Gly Asn Leu Leu Ala Ile Ile Val Gly Gly Ala Gln Glu Ala Leu  
 180 185 190

10

ES 2 640 100 T3

Asp Ala Arg Pro Gly Ala Tyr Arg Leu Leu Leu Lys Asn Arg Lys Gly  
 195 200 205

Phe Ile Arg Leu Ala Leu Met His Gly Ala Ala Leu Val Pro Ile Phe  
 210 215 220

Ser Phe Gly Glu Asn Asn Leu Phe Asn Gln Val Glu Asn Thr Pro Gly  
 225 230 235 240

Thr Trp Leu Arg Trp Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile  
 245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly  
 260 265 270

Leu Met Pro Phe Arg Gln Pro Ile Thr Thr Ile Val Gly Lys Pro Ile  
 275 280 285

Glu Val Gln Met Thr Pro Gln Pro Ser Arg Glu Glu Val Asp Arg Leu  
 290 295 300

His Gln Arg Tyr Ile Lys Glu Leu Cys Lys Leu Phe Glu Glu His Lys  
 305 310 315 320

Leu Lys Phe Asn Val Pro Glu Asp Gln His Leu Glu Phe Cys  
 325 330

<210> 63  
 <211> 334  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 63

5

Met Val Glu Phe Ala Pro Leu Phe Met Pro Trp Glu Arg Arg Leu Gln  
 1 5 10 15

Thr Leu Ala Val Leu Gln Phe Val Phe Ser Phe Leu Ala Leu Ala Glu  
 20 25 30

Ile Cys Thr Val Gly Phe Ile Ala Leu Leu Phe Thr Arg Phe Trp Leu  
 35 40 45

Leu Thr Val Leu Tyr Ala Ala Trp Trp Tyr Leu Asp Arg Asp Lys Pro  
 50 55 60

Arg Gln Gly Gly Arg His Ile Gln Ala Ile Arg Cys Trp Thr Ile Trp  
 65 70 75 80

10

ES 2 640 100 T3

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

ES 2 640 100 T3

<210> 64  
 <211> 334  
 <212> PRT  
 <213> *Pan troglodytes*

5

<400> 64

Met Val Glu Phe Ala Pro Leu Phe Val Pro Trp Glu Arg Arg Leu Gln  
 1 5 10 15

Thr Leu Ala Val Leu Gln Phe Val Phe Ser Phe Leu Ala Leu Ala Glu  
 20 25 30

Ile Cys Thr Val Gly Phe Ile Ala Leu Leu Phe Thr Arg Phe Trp Leu  
 35 40 45

Leu Thr Val Leu Tyr Ala Ala Trp Trp Tyr Leu Asp Arg Asp Lys Pro  
 50 55 60

Arg Gln Gly Gly Arg His Ile Gln Ala Ile Arg Cys Trp Thr Ile Trp  
 65 70 75 80

Lys Tyr Met Lys Asp Tyr Phe Pro Ile Ser Leu Val Lys Thr Ala Glu  
 85 90 95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val  
 100 105 110

Leu Ala Val Gly Ala Phe Ala Asn Leu Cys Thr Glu Ser Thr Gly Phe  
 115 120 125

Ser Ser Ile Phe Pro Gly Ile Arg Pro His Leu Met Met Leu Thr Leu  
 130 135 140

Trp Phe Arg Ala Pro Phe Phe Arg Asp Tyr Ile Met Ser Ala Gly Leu  
 145 150 155 160

Val Thr Ser Glu Lys Glu Ser Ala Ala His Ile Leu Asn Arg Lys Gly  
 165 170 175

Gly Gly Asn Leu Leu Gly Ile Ile Val Gly Gly Ala Gln Glu Ala Leu  
 180 185 190

Asp Ala Arg Pro Gly Ser Phe Thr Leu Leu Leu Arg Asn Arg Lys Gly  
 195 200 205

Phe Val Arg Leu Ala Leu Thr His Gly Ala Pro Leu Val Pro Ile Phe  
 210 215 220

ES 2 640 100 T3

Ser Phe Gly Glu Asn Asp Leu Phe Asp Gln Ile Pro Asn Ser Ser Gly  
 225 230 235 240

Ser Trp Leu Arg Tyr Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile  
 245 250 255

Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly  
 260 265 270

Leu Ile Pro Tyr Arg Arg Pro Ile Thr Thr Val Val Gly Lys Pro Ile  
 275 280 285

Glu Val Gln Lys Thr Leu His Pro Ser Glu Glu Glu Val Asn Gln Leu  
 290 295 300

His Gln Arg Tyr Ile Lys Glu Leu Cys Asn Leu Phe Glu Ala His Lys  
 305 310 315 320

Leu Lys Phe Asn Ile Pro Ala Asp Gln His Leu Glu Phe Cys  
 325 330

- <210> 65
- <211> 334
- <212> PRT
- <213> *Canis familiaris*

5

<400> 65

Met Val Lys Phe Ala Pro Leu Phe Val Pro Trp Glu Arg Arg Leu Gln  
 1 5 10 15

Thr Phe Met Val Leu Gln Trp Val Phe Ser Phe Leu Ala Leu Ala Gln  
 20 25 30

Ile Cys Thr Val Val Phe Val Gly Leu Leu Phe Thr Arg Phe Trp Val  
 35 40 45

Val Ser Val Leu Tyr Ala Ala Trp Trp Tyr Arg Asp Arg Asp Thr Pro  
 50 55 60

Arg Gln Gly Gly Arg Pro Val Gln Ala Leu Arg Arg Cys Phe Leu Trp  
 65 70 75 80

Lys Tyr Met Arg Asp Tyr Phe Pro Val Thr Leu Val Lys Thr Ala Glu  
 85 90 95

Leu Asp Pro Ser Arg Asn Tyr Leu Ala Gly Phe His Pro His Gly Val  
 100 105 110

10

ES 2 640 100 T3

Leu Ala Ala Gly Ala Phe Val Asn Leu Cys Thr Glu Ser Thr Gly Phe  
 115 120 125  
 Pro Leu Leu Phe Pro Gly Ile Arg Ser His Leu Met Met Leu Thr Leu  
 130 135 140  
 Trp Phe Arg Val Pro Phe Phe Arg Asp Tyr Ile Met Ser Gly Gly Leu  
 145 150 155 160  
 Val Thr Ser Asp Lys Glu Ser Ala Ala His Ile Leu Ser Arg Lys Glu  
 165 170 175  
 Gly Gly Asn Leu Leu Ala Ile Ile Val Gly Gly Ala Gln Glu Ala Leu  
 180 185 190  
 Asn Ala Arg Pro Gly Asp Ser Thr Leu Leu Leu Arg Asn Arg Lys Gly  
 195 200 205  
 Phe Ile Arg Leu Ala Leu Met His Gly Ala Ala Leu Val Pro Ile Phe  
 210 215 220  
 Ser Phe Gly Glu Asn Glu Leu Phe Asp Gln Val Glu Asn Ser Pro Gly  
 225 230 235 240  
 Ser Trp Leu Arg Arg Ile Gln Asn Arg Leu Gln Lys Ile Met Gly Ile  
 245 250 255  
 Ser Leu Pro Leu Phe His Gly Arg Gly Val Phe Gln Tyr Ser Phe Gly  
 260 265 270  
 Phe Ile Pro Tyr Arg Gln Pro Ile Thr Thr Val Val Gly Lys Pro Ile  
 275 280 285  
 Glu Val Gln Lys Thr Leu Cys Pro Ser Lys Glu Glu Val Asp Lys Leu  
 290 295 300  
 His Gln Arg Tyr Val Lys Glu Leu Cys Glu Leu Phe Glu Thr His Lys  
 305 310 315 320  
 Leu Lys Tyr Asn Val Pro Val Asp Gln His Leu Glu Phe Cys  
 325 330

<210> 66  
 <211> 332  
 <212> PRT  
 <213> *Bos taurus*

5

<400> 66



ES 2 640 100 T3

Met Val Glu Phe Ala Pro Leu Phe Val Pro Leu Glu Arg Arg Leu Gln  
1 5 10 15

Thr Phe Ala Val Leu Tyr Trp Ile Phe Cys Phe Met Val Leu Pro Pro  
20 25 30

Leu Cys Leu Val Val Phe Ile Gly Leu Leu Phe Thr Arg Phe Trp Leu  
35 40 45

Phe Asn Ile Leu Tyr Val Ile Trp Leu Tyr Leu Asp Gln Asn Arg Pro  
50 55 60

Arg Gln Gly Gly Ser His Asn Lys Phe Leu Lys Arg Trp Val Leu Trp  
65 70 75 80

Lys Tyr Met Lys Asp Tyr Phe Pro Ile Thr Leu Val Lys Thr Thr Glu  
85 90 95

Leu Asp Pro Ser Arg Asn Tyr Leu Ala Ala Phe His Pro His Gly Leu  
100 105 110

Leu Ser Val Gly Ala Phe Thr Asn Met Cys Thr Asp Ser Thr Gly Phe  
115 120 125

Ser Ser Leu Phe Pro Gly Ile Arg Pro His Leu Thr Thr Ile Asn Ile  
130 135 140

Tyr Phe Arg Ile Pro Leu Phe Arg Asp Tyr Ile Met Gln Gly Gly Leu  
145 150 155 160

Val Ser Ser Asp Lys Glu Ser Ile Ala Tyr Ile Leu Ser Arg Lys Gly  
165 170 175

Gly Gly Asn Leu Val Ala Ile Thr Val Gly Gly Ile Arg Glu Ala Leu  
180 185 190

Lys Thr Arg Pro Gly Ala Asn Lys Leu Val Leu Arg Asn Arg Lys Gly  
195 200 205

Phe Ile Arg Leu Ala Leu Met His Gly Ala Val Leu Val Pro Ile Phe  
210 215 220

Ser Phe Gly Asp Asn Glu Leu Tyr Ala Lys Thr Ser Pro Gly Phe Trp  
225 230 235 240

Trp Lys Trp Phe Arg Asp Gln Leu Tyr Lys Lys Thr Arg Leu Ala Ile

ES 2 640 100 T3

				245						250					255
Pro	Phe	Phe	Tyr	Gly	Arg	Gly	Val	Phe	Gln	Tyr	Ser	Phe	Gly	Phe	Met
			260					265					270		
Pro	Tyr	Arg	Arg	Pro	Ile	Thr	Thr	Val	Val	Gly	Lys	Pro	Ile	Glu	Val
		275					280					285			
Pro	Lys	Ile	Pro	His	Pro	Ser	Gln	Glu	Glu	Val	Asp	Arg	Leu	His	Gln
	290					295					300				
His	Tyr	Leu	Lys	Glu	Leu	Ser	Asn	Leu	Phe	Glu	Thr	His	Lys	Leu	Lys
305					310					315					320
Tyr	Asn	Ile	Pro	Lys	Asp	Gln	His	Leu	Glu	Phe	Cys				
				325					330						

5

<210> 67  
 <211> 334  
 <212> PRT  
 <213> *Rattus norvegicus*  
 <400> 67

ES 2 640 100 T3

Met Val Glu Phe Ala Pro Leu Leu Val Pro Trp Glu Arg Arg Leu Gln  
 1                   5                   10                   15

Thr Phe Ala Val Leu Gln Trp Val Phe Ser Phe Leu Ala Leu Ala Gln  
                  20                   25                   30

Leu Cys Ile Phe Ile Phe Ile Gly Leu Leu Phe Thr Arg Phe Trp Leu  
                  35                   40                   45

Phe Ser Val Leu Tyr Ala Thr Trp Trp Tyr Leu Asp Trp Asp Arg Pro  
                  50                   55                   60

Arg Gln Gly Gly Arg Pro Ile Gln Phe Phe Arg Arg Met Ala Ile Trp  
                  65                   70                   75                   80

Lys Tyr Met Lys Asp Phe Phe Pro Val Ser Leu Val Lys Thr Ala Glu  
                  85                   90                   95

Leu Asp Pro Ser Arg Asn Tyr Ile Ala Gly Phe His Pro His Gly Val  
                  100                   105                   110

Leu Ala Ala Gly Ala Phe Leu Asn Leu Cys Thr Glu Ser Thr Gly Phe  
                  115                   120                   125

Thr Ser Leu Phe Pro Gly Ile Arg Ser Tyr Leu Met Met Leu Thr Val

ES 2 640 100 T3

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

<210> 68  
 <211> 351  
 <212> PRT  
 <213> *Gallus gallus*  
 <400> 68

5

ES 2 640 100 T3

Met Glu Arg Ala Tyr Lys Ile Cys Asp Asp Thr Thr Ser Lys Arg Leu  
1                   5                   10                   15

Lys Asn Leu Phe Trp Ala Val Gln Asn Gly Ser Pro Gln Val Ser Arg

ES 2 640 100 T3

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

ES 2 640 100 T3

Ile Ser Leu Pro Leu Phe His Ala Arg Gly Ile Phe Gln Tyr Ser Phe  
 275 280 285

Gly Leu Val Pro Tyr Arg Arg Pro Ile Asn Thr Val Ile Gly Lys Pro  
 290 295 300

Ile Pro Val Leu Lys Lys His Lys Pro Thr Glu Glu Glu Val Asp Arg  
 305 310 315 320

Val His Lys Lys Tyr Leu Glu Glu Leu Ser Lys Leu Phe Glu Glu His  
 325 330 335

Lys Ala Lys Tyr Asn Val Pro Glu Asp Ser His Leu Glu Phe Ile  
 340 345 350

<210> 69  
 <211> 252  
 <212> PRT  
 <213> *Danio rerio*

5

<400> 69

Met Arg Asp Tyr Phe Pro Ile Arg Leu Ile Lys Thr Ala Asp Leu Asp  
 1 5 10 15

Thr Arg Lys Asn Tyr Val Met Gly Phe His Pro His Gly Ile Leu Val  
 20 25 30

Ala Gly Ala Phe Thr Asn Phe Cys Thr Glu Ala Thr Gly Phe Ser Lys  
 35 40 45

Leu Phe Pro Gly Ile Lys Ser Asn Leu Leu Met Leu Pro Leu Trp Phe  
 50 55 60

Arg Ala Pro Phe Phe Arg Asp Tyr Ile Met Ser Ala Gly Leu Val Pro  
 65 70 75 80

Ser Asp Lys Glu Ser Ala Ser Tyr Leu Leu Arg Arg Lys Gly Gly Gly  
 85 90 95

Asn Ala Val Val Ile Ala Val Gly Gly Ala Pro Glu Ala Leu Asp Ala  
 100 105 110

His Pro Gly Asp Tyr Thr Val His Leu Ala Asn Lys Lys Gly Phe Ile  
 115 120 125

Lys Leu Ala Ile Glu His Gly Ala Asp Leu Val Pro Ile Tyr Ser Phe  
 130 135 140

10

ES 2 640 100 T3

Gly Glu Asn Glu Val Phe Asp Gln Val Gln Asn Pro Arg Gly Thr Trp  
145 150 155 160

Leu Arg Tyr Ile Gln Glu Arg Leu Gln Arg Ile Met Gly Val Ser Leu  
165 170 175

Pro Leu Phe His Ala Arg Gly Val Phe Gln Tyr Thr Phe Gly Leu Met  
180 185 190

Pro Tyr Arg Lys Pro Ile Asn Thr Val Val Gly Arg Pro Ile Pro Val  
195 200 205

Glu Lys Asn Glu Lys Pro Thr Ala Glu Glu Leu Asp Val Tyr His Gln  
210 215 220

Arg Tyr Met Asp Glu Leu Ala Arg Leu Phe Glu Asp His Lys Gly Asn  
225 230 235 240

Tyr Gly Val Pro Glu Asp Thr His Leu Val Phe Gln  
245 250

<210> 70  
<211> 352  
<212> PRT  
<213> *Drosophila melanogaster*

5

<400> 70

Met Lys Ile Glu Trp Ala Pro Leu Arg Val Pro Leu Glu Arg Arg Leu  
1 5 10 15

Gln Ile Leu Val Thr Ala Phe Phe Thr Ser Met Leu Leu Ile Leu Leu  
20 25 30

Ser Val Ser Phe Leu Leu Val Ala Gly Ser Leu Ile Tyr Gly Gly Leu  
35 40 45

Leu Val Arg Ser Leu Met Val Thr Tyr Leu Ala Tyr Val Phe Val His  
50 55 60

His Lys Lys Thr Gln Ser Val Val Asp Gly Asn Gly Trp Met Ile Thr  
65 70 75 80

Arg Thr Asn Leu Leu His Arg His Tyr Arg Asp Tyr Phe Pro Val Glu  
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Ala Thr Lys Asn Tyr Ile Leu Ala  
100 105 110

10



ES 2 640 100 T3

Ser Phe Pro His Gly Ile Leu Gly Thr Gly Ile Gly Ile Asn Met Gly  
 115 120 125

Leu Glu Ile Ser Lys Trp Leu Glu Leu Phe Pro Gln Val Arg Pro Lys  
 130 135 140

Leu Gly Thr Leu Asp Gln His Phe His Val Pro Phe Met Arg Glu Val  
 145 150 155 160

Leu Arg Cys Trp Gly Leu Val Ser Val Ser Lys Glu Ala Leu Ile Arg  
 165 170 175

Met Leu Ser Lys Ser Asn Asp Pro Lys His Lys Asp Asn Arg Asp Gly  
 180 185 190

Phe Thr Ser Asn Ala Val Ala Ile Leu Val Gly Gly Ala Gln Glu Ala  
 195 200 205

Met Asp Ser His Pro Gly Gln Tyr Ile Leu Thr Leu Lys Asn Arg Lys  
 210 215 220

Gly Phe Val Arg Met Ala Ile Arg Thr Gly Ser Ser Ile Val Pro Ser  
 225 230 235 240

Phe Ser Phe Gly Glu Val Asp Ile Phe Asp Gln Val Ala Asn Pro Pro  
 245 250 255

Asn Ser Leu Leu Arg Arg Phe Gln Asp Phe Val Lys Lys Leu Thr Gly  
 260 265 270

Val Ser Pro Leu Ile Pro Val Gly Arg Gly Phe Phe Asn Tyr Thr Phe  
 275 280 285

Gly Phe Leu Pro Phe Arg Arg Arg Ile Val Gln Val Val Gly Ala Pro  
 290 295 300

Ile Asp Val Val Lys Asn Glu His Pro Asp Ser Glu Tyr Val Asp Lys  
 305 310 315 320

Val His Gly Gln Val Ile Glu Ser Leu Glu Lys Leu Phe Asp Gln Tyr  
 325 330 335

Lys Asp Lys Tyr Leu Glu Asn Ser Lys Ser Ala Thr Leu Val Val His  
 340 345 350

<210> 71  
 <211> 352

<212> PRT  
<213> *Drosophila melanogaster*

<400> 71

5

ES 2 640 100 T3

Met Lys Ile Glu Trp Ala Pro Lys Gly Val Pro Met Glu Arg Arg Arg  
 1 5 10 15

Gln Thr Phe Ala Met Ala Phe Leu Ile Leu Ser Phe Met Ile Leu Ser  
 20 25 30

Phe Gly Ser Tyr Phe Phe Val Ala Ala Val Leu Phe Tyr Gly Ser Leu  
 35 40 45

Leu Trp Arg Thr Ile Met Val Ile Tyr Leu Val Tyr Val Tyr Ala Asn  
 50 55 60

His Lys Arg Thr His Ser Ile Met Asp Gly Asn Gly Trp Lys Ile Asn  
 65 70 75 80

Arg Asn Asn Trp Leu Phe Arg His Tyr Arg Asp Tyr Phe Pro Val Gln  
 85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asn Lys Asn Tyr Ile Leu Ala  
 100 105 110

Ser Phe Pro His Gly Ile Leu Gly Thr Gly Ile Ser Ile Asn Met Gly  
 115 120 125

Leu Asp Ile Ser Lys Trp Leu Gln Leu Phe Pro Gln Val Arg Pro Lys  
 130 135 140

Val Ala Thr Leu Asp Gln Asn Phe Leu Thr Pro Ile Val Arg Gly Leu  
 145 150 155 160

Leu Arg Ser Trp Gly Leu Val Ser Val Ser Lys Glu Ala Leu Val Tyr  
 165 170 175

Leu Leu Thr Lys Ser Asn Asp Pro Lys His Lys Asp Asn Arg Asp Gly  
 180 185 190

Phe Thr Ser Asn Ala Val Ala Ile Leu Val Gly Gly Ala Gln Glu Ala  
 195 200 205

Leu Asp Ser His Pro Gly Lys Tyr Ile Leu Thr Leu Lys Asn Arg Lys  
 210 215 220

Gly Phe Val Lys Met Ala Ile Arg Thr Gly Ser Ser Ile Val Pro Thr  
 225 230 235 240

ES 2 640 100 T3

Phe Ser Phe Gly Glu Val Asp Ile Leu Asp Gln Val Ala Asn Pro Pro  
 245 250 255

Asn Ser Arg Val Arg Arg Phe Gln Asp Phe Val Lys Arg Ile Thr Gly  
 260 265 270

Ile Ser Pro Leu Ile Pro Val Gly Arg Gly Ile Phe Asn Tyr Ser Phe  
 275 280 285

Gly Phe Leu Pro Asn Arg Arg Arg Ile Val Gln Val Val Gly Ala Pro  
 290 295 300

Ile Asp Val Val Gln Ser Asp Gln Pro Asp Ala Ala Tyr Val Asp Lys  
 305 310 315 320

Ile His Lys Gln Val Ile Asp Asp Leu Glu Lys Met Phe Ala Lys Tyr  
 325 330 335

Lys Asp Gln Tyr Ile Pro Asn Ser Lys Gln Asp Lys Leu Ile Ile His  
 340 345 350

<210> 72  
 <211> 349  
 <212> PRT  
 <213> *Drosophila melanogaster*  
 <400> 72

5

Met Thr Ile Glu Trp Ala Pro Leu Arg Val Pro Leu Glu Arg Arg Leu  
 1 5 10 15

Gln Thr Leu Val Thr Ser Phe Phe Thr Tyr Thr Phe Phe Thr Leu Pro  
 20 25 30

Ile Ser Ser Cys Leu Ala Val Ala Ile Leu Leu Tyr Tyr Gly Glu Met  
 35 40 45

Phe Val Arg Ser Leu Leu Leu Ile Tyr Phe Val Lys Ile Tyr Leu Asp  
 50 55 60

Tyr Lys Arg Asn Tyr Gly Ile Met Glu Gly Asn Gly Trp Leu Phe Tyr  
 65 70 75 80

Arg Ser Asn Trp Arg Tyr Arg Asn Tyr Phe Pro Val Glu Leu Val Lys  
 85 90 95

Thr Ala Glu Leu Pro Pro Asn Arg Asn Tyr Ile Val Ala Ser Phe Pro  
 100 105 110

10

ES 2 640 100 T3

His Gly Ile Leu Gly Thr Gly Thr Cys Ile Asn Met Ser Leu Asp Ile  
 115 120 125

Asp Asn Trp Leu Ser Leu Tyr Pro His Val Arg Pro Lys Ile Ala Thr  
 130 135 140

Leu Asp His His Phe Lys Thr Pro Phe Leu Arg Asp Ile Leu Arg Trp  
 145 150 155 160

Trp Gly Met Val Ser Val Ser Lys Glu Ser Leu Ser Tyr Leu Leu Ser  
 165 170 175

Lys Ser Asn Asp Pro Met His Lys Asp Asn Arg Asp Gly Phe Thr Ser  
 180 185 190

Asn Ala Val Ala Val Leu Val Gly Gly Ala Lys Glu Ala Met Asp Ser  
 195 200 205

His Pro Gly Gln Tyr Ile Leu Thr Leu Lys Asp Arg Lys Gly Phe Val  
 210 215 220

Lys Met Ala Val Arg Thr Gly Ser Ser Ile Val Pro Ser Leu Ser Phe  
 225 230 235 240

Gly Glu Val Asp Ile Phe Asp Gln Val Ala Asn Pro Pro Asp Ser Ser  
 245 250 255

Leu Arg Arg Phe Gln Asn Val Val Lys Lys Phe Thr Gly Ile Ser Pro  
 260 265 270

Leu Leu Pro Lys Gly Arg Gly Ile Phe Asn Tyr Asn Tyr Gly Ile Leu  
 275 280 285

Pro His Arg Arg Arg Ile Val Gln Val Val Gly Ser Pro Ile Asp Val  
 290 295 300

Glu Arg Cys Glu Thr Pro Asp Pro Glu Tyr Val Asp Lys Ile His Gly  
 305 310 315 320

Gln Val Ile Asp Ala Leu Ala Arg Met Phe Asp Glu Tyr Lys Glu Lys  
 325 330 335

Tyr Thr Pro Asn Ser Lys His Ile Lys Leu Ile Ile Gln  
 340 345

<210> 73  
 <211> 353

<212> PRT  
<213> *Ashbya gossypii*

<400> 73

5

ES 2 640 100 T3

Met Gly Lys Ile Glu Trp Ala Pro Leu Asn Val Pro Met Arg Arg Arg  
1 5 10 15

Ile Glu Thr Leu Ser Thr Ala Leu Trp Met Trp Leu Ile Leu Phe Gly  
20 25 30

Glu Leu Gly Met Leu Ile Ser Tyr Phe Leu Leu Leu Ile Tyr Gly Asn  
35 40 45

Leu Phe Ile Lys Ser Leu Cys Val Ile Tyr Gly Tyr Phe Ile Tyr Thr  
50 55 60

Asp Arg Lys Val Thr Leu Asn Gly Gly Arg Gly Gln Gly Val Thr Trp  
65 70 75 80

Trp Arg Glu Leu Phe Trp Trp Lys Leu Tyr Gln Ser Tyr Phe Pro Ala  
85 90 95

Lys Leu His Lys Thr Val Asp Leu Asp Pro Asn Arg Asn Tyr Leu Phe  
100 105 110

Ala Ala Phe Pro His Gly Val Leu Gly Leu Gly Ala Phe Ile Asn Phe  
115 120 125

Ala Thr Asn Ala Thr Gly Phe His Asp Lys Phe Pro Lys Ile Arg Ser  
130 135 140

Arg Pro Val Thr Leu Asn Phe His Phe Val Ile Pro Phe Phe Arg Glu  
145 150 155 160

Leu Leu Leu Ser Trp Gly Leu Val Ser Ala Asn Pro Asn Ser Ile Leu  
165 170 175

Ser Leu Leu Lys Ala Pro Asn Lys Pro Asp His Pro Leu Asn Asp Asp  
180 185 190

Gly Tyr Thr Ala Asn Ala Val Val Ile Val Val Gly Gly Ala Ala Glu  
195 200 205

Ser Leu His Cys Arg Pro Asn Asn Tyr Thr Leu Val Leu Arg Lys Arg  
210 215 220

Lys Gly Phe Cys Lys Leu Ala Ile Lys Ala Gly Thr Pro Leu Val Pro  
225 230 235 240

Val Met Thr Phe Gly Glu Val Asp Leu Phe Asp Gln Pro Pro Asn Pro  
 245 250 255

Pro Gly Ser Arg Leu Arg Arg Phe Gln Glu Phe Val Lys Asn Thr Thr  
 260 265 270

Gly Ile Ala Pro Ala Ala Phe Val Gly Arg Gly Phe Phe Gln Tyr Ser  
 275 280 285

Tyr Gly Leu Ile Pro Arg Arg Lys Pro Leu Asn Thr Val Val Gly Ala  
 290 295 300

Pro Val Glu Val Thr Gln Ile Asp Asn Pro Thr Gln Glu Gln Val Asp  
 305 310 315 320

Glu Val His Glu Arg Phe Cys Arg Ala Leu Asp Asn Leu Phe Glu Thr  
 325 330 335

Asn Lys Ser Arg Phe Ile Ala Asp Tyr Lys Asn Val Lys Leu Val Met  
 340 345 350

Glu

<210> 74  
 <211> 353  
 <212> PRT  
 <213> *Drosophila melanogaster*

5

<400> 74

Met Gly Lys Ile Glu Trp Ala Pro Leu Asn Val Pro Met Arg Arg Arg  
 1 5 10 15

Ile Glu Thr Leu Ser Thr Ala Leu Trp Met Trp Leu Ile Leu Phe Gly  
 20 25 30

Glu Leu Gly Met Leu Ile Ser Tyr Phe Leu Leu Leu Ile Tyr Gly Asn  
 35 40 45

Leu Phe Ile Lys Ser Leu Cys Val Ile Tyr Gly Tyr Phe Ile Tyr Thr  
 50 55 60

Asp Arg Lys Val Thr Leu Asn Gly Gly Arg Gly Gln Gly Val Thr Trp  
 65 70 75 80

Trp Arg Glu Leu Phe Trp Trp Lys Leu Tyr Gln Ser Tyr Phe Pro Ala  
 85 90 95

10



ES 2 640 100 T3

Lys Leu His Lys Thr Val Asp Leu Asp Pro Asn Arg Asn Tyr Leu Phe  
 100 105 110  
 Ala Ala Phe Pro His Gly Val Leu Gly Leu Gly Ala Phe Ile Asn Phe  
 115 120 125  
 Ala Thr Asn Ala Thr Gly Phe His Asp Lys Phe Pro Lys Ile Arg Ser  
 130 135 140  
 Arg Pro Val Thr Leu Asn Phe His Phe Val Ile Pro Phe Phe Arg Glu  
 145 150 155 160  
 Leu Leu Leu Ser Trp Gly Leu Val Ser Ala Asn Pro Asn Ser Ile Leu  
 165 170 175  
 Ser Leu Leu Lys Ala Pro Asn Lys Pro Asp His Pro Leu Asn Asp Asp  
 180 185 190  
 Gly Tyr Thr Ala Asn Ala Val Val Ile Val Val Gly Gly Ala Ala Glu  
 195 200 205  
 Ser Leu His Cys Arg Pro Asn Asn Tyr Thr Leu Val Leu Arg Lys Arg  
 210 215 220  
 Lys Gly Phe Cys Lys Leu Ala Ile Lys Ala Gly Thr Pro Leu Val Pro  
 225 230 235 240  
 Val Met Thr Phe Gly Glu Val Asp Leu Phe Asp Gln Pro Pro Asn Pro  
 245 250 255  
 Pro Gly Ser Arg Leu Arg Arg Phe Gln Glu Phe Val Lys Asn Thr Thr  
 260 265 270  
 Gly Ile Ala Pro Ala Ala Phe Val Gly Arg Gly Phe Phe Gln Tyr Ser  
 275 280 285  
 Tyr Gly Leu Ile Pro Arg Arg Lys Pro Leu Asn Thr Val Val Gly Ala  
 290 295 300  
 Pro Val Glu Val Thr Gln Ile Asp Asn Pro Thr Gln Glu Gln Val Asp  
 305 310 315 320  
 Glu Val His Glu Arg Phe Cys Arg Ala Leu Asp Asn Leu Phe Glu Thr  
 325 330 335  
 Asn Lys Ser Arg Phe Ile Ala Asp Tyr Lys Asn Val Lys Leu Val Met

340

345

350

Glu

5

<210> 75  
 <211> 338  
 <212> PRT  
 <213> *Tribolium castaneum*  
 <400> 75

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

10

			195				200				205				
Leu	Lys	Asn	Arg	Lys	Gly	Phe	Val	Lys	Leu	Ala	Leu	Arg	Asn	Gly	Ser
	210					215					220				
Pro	Leu	Val	Pro	Val	Ile	Ser	Phe	Gly	Glu	Pro	Glu	Leu	Phe	Asp	Gln
	225				230					235					240
Val	Glu	Gly	Arg	Thr	Leu	Arg	Lys	Ile	Gln	Glu	Ser	Ile	Arg	Lys	Tyr
				245					250					255	
Leu	Gly	Leu	Ala	Pro	Val	Ile	Phe	Ser	Gly	Arg	Gly	Phe	Phe	Gln	Tyr
			260					265						270	
Ser	Phe	Gly	Val	Ile	Pro	Gln	Arg	Arg	Pro	Ile	Thr	Thr	Val	Val	Gly
		275					280						285		
His	Pro	Ile	Glu	Val	Thr	Lys	Ile	Glu	Lys	Pro	Thr	Asn	Glu	Glu	Val
	290					295					300				
Asp	Glu	Leu	His	Lys	Lys	Tyr	Met	Gln	Glu	Leu	Glu	Asn	Leu	Phe	Glu
	305				310						315				320
Glu	Tyr	Lys	Phe	Lys	Tyr	Leu	Glu	Asn	Pro	Lys	Asp	Ile	His	Leu	Glu
				325					330					335	
Phe Glu															

<210> 76  
 <211> 341  
 <212> PRT  
 <213> *Homo sapiens*

5

<400> 76

ES 2 640 100 T3

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

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr  
20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu  
35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr  
50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile

ES 2 640 100 T3

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

ES 2 640 100 T3

Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser Thr  
 325 330 335

Cys Leu Thr Phe Ile  
 340

5

<210> 77  
 <211> 281  
 <212> PRT  
 <213> *Pan troglodytes*  
 <400> 77

Met Gly Val Ala Thr Thr Leu Leu Pro Pro Thr Thr Ser Lys Thr Leu  
 1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr  
 20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu  
 35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr  
 50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile  
 65 70 75 80

Arg Asn Trp Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys  
 85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly  
 100 105 110

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser  
 115 120 125

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp  
 130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr  
 145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe  
 165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly  
 180 185 190

10

ES 2 640 100 T3

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr  
 195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Gly  
 210 215 220

Pro Pro His Pro Cys Pro Pro Ala Pro Pro Pro His Arg Gly Gly Ser  
 225 230 235 240

Gln Ser Leu Ser Arg Pro Leu His Asp Gly Pro Ala Ala Ala Leu Arg  
 245 250 255

Gly Ala Gln Gly Lys Leu Trp Gly Pro Arg Phe His Leu Pro His Leu  
 260 265 270

His Leu Gly Leu Ala Ala Ala Phe Arg  
 275 280

<210> 78  
 <211> 344  
 <212> PRT  
 <213> *Pan troglodytes*  
 <400> 78

5

Met Gly Val Ala Thr Thr Leu Leu Pro Pro Thr Thr Ser Lys Thr Leu  
 1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr  
 20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu  
 35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr  
 50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile  
 65 70 75 80

Arg Asn Trp Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys  
 85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly  
 100 105 110

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser  
 115 120 125

10

ES 2 640 100 T3

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp  
 130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr  
 145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe  
 165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly  
 180 185 190

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr  
 195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Ala  
 210 215 220

Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Arg Leu  
 225 230 235 240

Lys Ala Phe Ala Thr Gly Ser Trp Gln His Trp Cys Gln Leu Thr Phe  
 245 250 255

Lys Lys Phe Met Gly Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly Leu  
 260 265 270

Phe Ser Ala Thr Ser Trp Gly Leu Leu Pro Phe Ala Val Pro Ile Thr  
 275 280 285

Thr Val Gly Glu Cys Pro Pro Pro Gly Gly Arg Pro Pro Ala Ala Ala  
 290 295 300

Trp Ala Ser Gly Ile Pro Arg Pro Pro Val Ser Leu Ser Leu Gln Trp  
 305 310 315 320

Ala Ala Pro Ser Leu Ser Pro Ser Ala Ser Thr Pro Pro Arg Arg Lys  
 325 330 335

Ser Ile Thr Ile Thr Pro Ser Thr  
 340

<210> 79  
 <211> 341  
 <212> PRT  
 <213> *Pan troglodytes*  
 <400> 79



ES 2 640 100 T3

Met Gly Val Ala Thr Thr Leu Leu Pro Pro Thr Thr Ser Lys Thr Leu  
1 5 10 15

Gln Lys Gln His Leu Glu Ala Val Gly Ala Tyr Gln Tyr Val Leu Thr  
20 25 30

Phe Leu Phe Met Gly Pro Phe Phe Ser Leu Leu Val Phe Val Leu Leu  
35 40 45

Phe Thr Ser Leu Trp Pro Phe Ser Val Phe Tyr Leu Val Trp Leu Tyr  
50 55 60

Val Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Arg Ser Glu Trp Ile  
65 70 75 80

Arg Asn Trp Ala Ile Trp Arg Gln Leu Arg Asp Tyr Tyr Pro Val Lys  
85 90 95

Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu Gly  
100 105 110

Ala His Pro His Gly Ile Met Cys Thr Gly Phe Leu Cys Asn Phe Ser  
115 120 125

Thr Glu Ser Asn Gly Phe Ser Gln Leu Phe Pro Gly Leu Arg Pro Trp  
130 135 140

Leu Ala Val Leu Ala Gly Leu Phe Tyr Leu Pro Val Tyr Arg Asp Tyr  
145 150 155 160

Ile Met Ser Phe Gly Leu Cys Pro Val Ser Arg Gln Ser Leu Asp Phe  
165 170 175

Ile Leu Ser Gln Pro Gln Leu Gly Gln Ala Val Val Ile Met Val Gly  
180 185 190

Gly Ala His Glu Ala Leu Tyr Ser Val Pro Gly Glu His Cys Leu Thr  
195 200 205

Leu Gln Lys Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly Ala  
210 215 220

Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Arg Leu  
225 230 235 240

Lys Ala Phe Ala Thr Gly Ser Trp Gln His Trp Cys Gln Leu Thr Phe  
245 250 255

Lys Lys Phe Met Gly Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly Leu  
 260 265 270

Phe Ser Ala Thr Ser Trp Gly Leu Leu Pro Phe Ala Val Pro Ile Thr  
 275 280 285

Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Arg Leu His Pro Thr  
 290 295 300

Glu Glu Glu Val Asn His Tyr His Ala Leu Tyr Met Met Ala Leu Gln  
 305 310 315 320

Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser Thr  
 325 330 335

Cys Leu Thr Phe Ile  
 340

<210> 80  
 <211> 582  
 <212> PRT  
 <213> *Canis familiaris*

5

<400> 80

Met Thr Phe Pro Asp Pro Lys Thr Gly Gln Cys Cys Leu Ser Pro Val  
 1 5 10 15

Leu Ser Gly Pro Gly Gly Gly Arg Val Ala Leu Leu Ala Leu Gly Gly  
 20 25 30

Pro Leu Glu Ala Leu Glu Ala Lys Pro Gly Ala Leu Phe Cys Gly Ser  
 35 40 45

Gly Ile Arg Arg Asp Leu Ser Ser Trp His Arg Asp Thr Glu Arg Ala  
 50 55 60

Leu Pro Ala Val Arg Glu Pro Ser Gly Pro Leu Gly Leu Leu Leu Pro  
 65 70 75 80

Phe Arg Ala Arg Pro Ser Thr Arg Ser Gly Ser Pro Ser Leu Cys Pro  
 85 90 95

Leu Gly Ser Pro Ala Arg Thr Ser Ser Ala Ala Leu Gly Pro Ala Arg  
 100 105 110

Ser Pro Arg Arg Gly Gly Asn Pro Ala Leu Leu Pro Thr Pro Gly His  
 115 120 125

10

ES 2 640 100 T3

Thr Gly Ser Pro Leu Leu Thr Leu Pro Val Pro Asp Ser Val Glu Thr  
 130 135 140

Val Ala Arg Gly Asn Glu Glu Ala Gly Arg Cys Arg Gly Ser Arg Ser  
 145 150 155 160

Ala Arg Met Ala Arg Leu Thr Arg Leu Phe Arg Glu Arg Lys Ala Arg  
 165 170 175

Ala Gly Leu Pro Gly His Pro His Leu Pro Ala Val Ser Leu Arg Cys  
 180 185 190

Arg His Ser Arg Gly Pro Ala His Gly Ala Gly Ala Arg Thr Arg Gly  
 195 200 205

Leu Pro Gly His Ala Gly Pro Arg Arg Arg Gly Ala Phe Gly Arg Ser  
 210 215 220

Ser Lys Lys Ala Phe Thr Pro Cys Pro Asn Leu Arg Arg Leu Arg Ser  
 225 230 235 240

Arg Pro Trp Asn Pro Thr Gly Gly Gly Glu Arg Pro Thr Met Arg Thr  
 245 250 255

Thr Gln Lys Gln Trp Leu Glu Ala Leu Ser Val Ser Tyr Tyr Val Phe  
 260 265 270

Thr Phe Leu Phe Met Gly Leu Phe Phe Ser Leu Leu Val Leu Phe Leu  
 275 280 285

Leu Phe Thr Ser Phe Trp Ser Ile Ser Val Leu Tyr Leu Val Trp Leu  
 290 295 300

Phe Leu Asp Trp Asp Thr Pro Asn Gln Gly Gly Arg Cys Phe Glu Cys  
 305 310 315 320

Asn Arg Ser Cys Thr Ile Trp Lys His Leu Lys Asp Tyr Phe Pro Ile  
 325 330 335

Lys Leu Val Lys Thr Ala Glu Leu Pro Pro Asp Arg Asn Tyr Val Val  
 340 345 350

Gly Cys His Pro His Gly Ile Met Cys Met Gly Thr Phe Cys Asn Phe  
 355 360 365

Phe Thr Glu Ala Asn Asn Phe Ser Lys Gln Phe Pro Gly Ile Gln Thr  
 370 375 380

ES 2 640 100 T3

Ser Pro Val Thr Leu Ala Phe Leu Leu His Leu Pro Val Tyr Arg Asp  
 385 390 395 400

Tyr Leu Met Tyr Leu Gly Leu Cys Ser Val Asn Arg Arg Ser Leu Asp  
 405 410 415

Phe Ile Leu Ser Lys Pro Gln Pro Gly Gln Ala Val Ala Ile Val Val  
 420 425 430

Gly Gly Ala His Glu Ser Leu Phe Ala Ile Pro Gly Met His Cys Leu  
 435 440 445

Val Leu Arg Asn Arg Lys Gly Phe Val Arg Leu Ala Leu Arg His Gly  
 450 455 460

Ala Ser Leu Val Pro Val Tyr Ser Phe Gly Glu Asn Asp Ile Phe Asn  
 465 470 475 480

Phe Lys Ala Phe Pro Thr Asn Ser Trp Gln Tyr Leu Cys Gln Ile Thr  
 485 490 495

Ile Lys Lys Ile Met Lys Phe Ser Pro Cys Ile Phe Trp Gly Arg Gly  
 500 505 510

Leu Phe Ser Ala Asp Ser Trp Gly Leu Leu Pro Phe Ala Lys Pro Ile  
 515 520 525

Thr Thr Val Val Gly Arg Pro Ile Pro Val Pro Gln Arg Leu Asn Pro  
 530 535 540

Thr Glu Glu Glu Val Asp His Tyr His Met Leu Tyr Met Glu Ala Leu  
 545 550 555 560

Glu Gln Leu Phe Glu Glu His Lys Glu Ser Cys Gly Val Pro Ala Ser  
 565 570 575

Thr His Leu Ile Phe Lys  
 580

<210> 81  
 <211> 406  
 <212> PRT  
 <213> *Bos taurus*

5

<400> 81

Met Ile Ser Glu Ser Arg Leu Glu Thr Ser Gly Ala Leu Met Arg Glu  
 1 5 10 15

10

ES 2 640 100 T3

Ser Pro His Ala His Ser Pro Leu Gln Ala Ile Leu Gly Gly Ser Gln  
 20 25 30

Ala Leu Gly His Ser Asp Gln Lys Asp Phe Leu Arg Ser Asn Gln Ser  
 35 40 45

Leu His Pro Ile Phe Pro Gly Ser Glu Ala Leu Leu Ala Glu Gly Glu  
 50 55 60

His Leu Gly Val Ser Thr Thr Pro Pro Pro Thr Pro Ser Met Lys Thr  
 65 70 75 80

Leu Lys Lys Gln Trp Leu Glu Val Leu Ser Thr Tyr Gln Tyr Val Leu  
 85 90 95

Cys Phe Leu Phe Leu Gly Pro Phe Phe Ser Leu Val Gly Phe Phe Leu  
 100 105 110

Leu Phe Thr Ser Leu Trp Tyr Leu Ser Val Leu Tyr Leu Val Trp Leu  
 115 120 125

Phe Leu Asp Trp Asp Thr Pro Gln Gln Gly Gly Arg Arg Asn Gln Trp  
 130 135 140

Leu Lys Asn Cys Thr Val Trp Lys His Leu Ser Asp Tyr Phe Pro Ile  
 145 150 155 160

Lys Leu Val Lys Thr Val Glu Leu Pro Pro Asp Arg Asn Tyr Val Leu  
 165 170 175

Leu Ser His Pro His Gly Ile Met Gly Phe Gly Thr Val Cys Asn Phe  
 180 185 190

Ser Thr Glu Gly Thr Gly Cys Ser Gln Leu Phe Pro Gly Leu Arg Phe  
 195 200 205

Ser Leu Ala Val Leu Asn Cys Leu Leu Tyr Val Pro Gly Cys Arg Glu  
 210 215 220

Tyr Ile Met Ser Cys Gly Thr Cys Ser Val Asn Arg Gln Ser Leu Asp  
 225 230 235 240

Tyr Val Leu Ser Gln Pro Gln Leu Gly Arg Ala Val Val Ile Met Val  
 245 250 255

Gly Gly Ala Asn Glu Ala Leu His Thr Val Pro Gly Glu His Cys Leu

ES 2 640 100 T3

					260						265						270
Thr	Leu	Arg	Asn	Arg	Lys	Gly	Phe	Val	Arg	Leu	Ala	Leu	Arg	His	Gly		
		275					280					285					
Ala	Ser	Leu	Val	Pro	Val	Tyr	Ser	Phe	Gly	Glu	Asn	Asp	Val	Phe	Arg		
	290					295					300						
Val	Lys	Ala	Phe	Ala	Pro	Asp	Ser	Trp	Gln	Arg	Leu	Phe	Gln	Val	Thr		
305					310					315					320		
Ile	Lys	Arg	Leu	Leu	Ser	Phe	Ser	Pro	Cys	Ile	Phe	Trp	Gly	Arg	Gly		
				325					330					335			
Leu	Phe	Ser	Ala	Lys	Ser	Trp	Gly	Leu	Met	Pro	Leu	Ala	Arg	Pro	Ile		
			340					345					350				
Thr	Thr	Val	Val	Gly	Arg	Pro	Ile	Pro	Val	Pro	Gln	Cys	Pro	Gln	Pro		
		355					360					365					
Thr	Glu	Glu	Gln	Val	Asp	His	Tyr	His	Arg	Leu	Tyr	Met	Lys	Ala	Leu		
	370					375					380						
Glu	Gln	Leu	Phe	Glu	Glu	His	Lys	Lys	Ser	Cys	Gly	Leu	Pro	Ala	Ser		
385					390					395					400		
Thr	His	Leu	Thr	Phe	Ile												
				405													

<210> 82  
 <211> 336  
 <212> PRT  
 <213> *Danio rerio*

5

<400> 82

ES 2 640 100 T3

Met Gly Thr Asp Asn Val Ser Glu Val Lys Glu Lys Gly Asp Arg Ser  
1 5 10 15

Pro Trp Lys Asp Ile Ile Glu Asp Ile Ser Val Leu Gln Leu Val Leu  
20 25 30

Ser Phe Leu Phe Leu Gly Val Ala Cys Leu Leu Leu Met Ile Tyr Leu  
35 40 45

Met Phe Thr Ser Leu Trp Ile Phe Pro Thr Leu Tyr Phe Thr Trp Gln  
50 55 60

Ile Tyr Asp Trp His Thr Pro Glu Arg Gly Gly Arg Arg Thr Lys Phe

ES 2 640 100 T3

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



ES 2 640 100 T3

His Lys Thr Ser Tyr Gly Leu Ala Glu Thr His Glu Leu His Phe Ile  
 325 330 335

<210> 83  
 <211> 520  
 5 <212> PRT  
 <213> *Arabidopsis thaliana*  
 <400> 83

Met Ala Ile Leu Asp Ser Ala Gly Val Thr Thr Val Thr Glu Asn Gly  
 1 5 10 15

Gly Gly Glu Phe Val Asp Leu Asp Arg Leu Arg Arg Arg Lys Ser Arg  
 20 25 30

Ser Asp Ser Ser Asn Gly Leu Leu Leu Ser Gly Ser Asp Asn Asn Ser  
 35 40 45

Pro Ser Asp Asp Val Gly Ala Pro Ala Asp Val Arg Asp Arg Ile Asp  
 50 55 60

Ser Val Val Asn Asp Asp Ala Gln Gly Thr Ala Asn Leu Ala Gly Asp  
 65 70 75 80

Asn Asn Gly Gly Gly Asp Asn Asn Gly Gly Gly Arg Gly Gly Gly Glu  
 85 90 95

Gly Arg Gly Asn Ala Asp Ala Thr Phe Thr Tyr Arg Pro Ser Val Pro  
 100 105 110

Ala His Arg Arg Ala Arg Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe  
 115 120 125

Lys Gln Ser His Ala Gly Leu Phe Asn Leu Cys Val Val Val Leu Ile  
 130 135 140

Ala Val Asn Ser Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Trp  
 145 150 155 160

Leu Ile Arg Thr Asp Phe Trp Phe Ser Ser Arg Ser Leu Arg Asp Trp  
 165 170 175

Pro Leu Phe Met Cys Cys Ile Ser Leu Ser Ile Phe Pro Leu Ala Ala  
 180 185 190

Phe Thr Val Glu Lys Leu Val Leu Gln Lys Tyr Ile Ser Glu Pro Val  
 195 200 205

10

ES 2 640 100 T3

Val Ile Phe Leu His Ile Ile Ile Thr Met Thr Glu Val Leu Tyr Pro  
 210 215 220

Val Tyr Val Thr Leu Arg Cys Asp Ser Ala Phe Leu Ser Gly Val Thr  
 225 230 235 240

Leu Met Leu Leu Thr Cys Ile Val Trp Leu Lys Leu Val Ser Tyr Ala  
 245 250 255

His Thr Ser Tyr Asp Ile Arg Ser Leu Ala Asn Ala Ala Asp Lys Ala  
 260 265 270

Asn Pro Glu Val Ser Tyr Tyr Val Ser Leu Lys Ser Leu Ala Tyr Phe  
 275 280 285

Met Val Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro Arg Ser Ala  
 290 295 300

Cys Ile Arg Lys Gly Trp Val Ala Arg Gln Phe Ala Lys Leu Val Ile  
 305 310 315 320

Phe Thr Gly Phe Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn Pro Ile  
 325 330 335

Val Arg Asn Ser Lys His Pro Leu Lys Gly Asp Leu Leu Tyr Ala Ile  
 340 345 350

Glu Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp Leu Cys  
 355 360 365

Met Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu  
 370 375 380

Leu Cys Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Ala Lys  
 385 390 395 400

Ser Val Gly Asp Tyr Trp Arg Met Trp Asn Met Pro Val His Lys Trp  
 405 410 415

Met Val Arg His Ile Tyr Phe Pro Cys Leu Arg Ser Lys Ile Pro Lys  
 420 425 430

Thr Leu Ala Ile Ile Ile Ala Phe Leu Val Ser Ala Val Phe His Glu  
 435 440 445

Leu Cys Ile Ala Val Pro Cys Arg Leu Phe Lys Leu Trp Ala Phe Leu  
 450 455 460

ES 2 640 100 T3

Gly Ile Met Phe Gln Val Pro Leu Val Phe Ile Thr Asn Tyr Leu Gln  
465 470 475 480

Glu Arg Phe Gly Ser Thr Val Gly Asn Met Ile Phe Trp Phe Ile Phe  
485 490 495

Cys Ile Phe Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu  
500 505 510

Met Asn Arg Lys Gly Ser Met Ser  
515 520

- 5  
<210> 84  
<211> 1856  
<212> ADN  
<213> *Arabidopsis thaliana*  
  
<400> 84

ES 2 640 100 T3

aaggaccttc cctctctttc tttocaaatc tctctttatc tctctttctc tgtttactaa 60  
atcgccagta ttcgocgcaa gaaagatgtc tccggcgaag aagagccgaa gttttcctcc 120  
gataagcgag tgtaaaagca gagagtatga ttcgatcgcg gcggatctcg acgggactct 180  
gcttctgtca agaagctctt tcccttactt catgctggtg gccattgaag caggtagtct 240  
cttccgtgga ttgatccttc ttctttctct tcccatcgtc atcatcgtt atctcttctg 300  
ctcogaatct ctoggaatcc aaatccttat cttcatctcc ttcgocggtc tcaagatcaa 360  
gaacatcgaa ctgctctctc gcgocgttct tacacggttt tacgocgocg atgtgaggaa 420  
agatagtttt gaggtgtttg ataaatgtaa gaagaggaaa gtagtggtga ctgcgaatcc 480  
gatagtgatg gttgagccat tcgtcaaaga ctatttgga ggagataaag ttttggaac 540  
agagattgaa gtaaacccca aaacgatgaa agctactggt tttgtgaaa agcctggtgt 600  
tcttgttggg gatctcaaga gattagccat cttgaaagag ttcggogatg actcacogga 660  
tcttggcctc ggtgaccgca cctocgatca cgatttcatg tccatttgca aggaaggtta 720  
catggtgcat gagaccaa atcagccacaac agtccctata gagagtctca agaaccgcat 780  
aatcttccat gatggocgtc ttgtccaacg tccgaccca ttaaaccgct taatcattta 840  
cctttggctt cctttogget tcatgctctc tgtcttccgc gtctacttca acctcccttt 900  
accggaacgc ttcgtccgct acacttacga gatcctcggc attcacctca caatccgtgg 960  
ccaccgtcct ccacctcctt cccocgaaa acctgaaaac ctctaogttc tcaaccaccg 1020  
cacogccctt gacccatta tcatcgccat tgctctcggc cgtaagatca catgtgtcac 1080  
ttatagtgtc tctcgcctct cctgatgct ctoaccatt cctgctgttg ctctgaccog 1140  
agaccgtgtc gctgacgcag cccgatgag gcaactctc gagaaagtg atttgtgat 1200  
ctgtcctgaa ggcacaacgt gtagagaacc atatctacta aggttcagtg ctctatttgc 1260  
agagctgagc gaccggatcg tgctgtggc catgaactgc aagcaaggga tgttcaacgg 1320  
gacgacagtt agaggtgtga aattctggga tccgtacttc ttcttcatga acccccgaac 1380  
tagctacgag gccactttct tggaccgttt gccogaagaa atgacggtaa acggtggtgg 1440  
caagactcct tttgaggtgg ctaattacgt tcagaagtg atcgggtggag ttttggggtt 1500  
cgaatgtact gaacttaca ggaaggataa gtatctgttg cttggaggca acgacggaaa 1560  
ggttgagtct atcaataaga ccaagtccat ggagtaataa tccagctagt ctggttctg 1620  
ctattttctg tgggagtga gttaggtctt tcgggttttc ccaattactt tttttttgt 1680  
tttttttcat gtattgttgc aggaaaatgt tttgttgaat atgcaactgga taatgaaact 1740  
acatttgaat caaagtacgt acatgtagt taccatttcc ccttgcctta actatgaatg 1800  
atttgtaaac tcaccttcta ttataagata tcatagagaa aaacacattc tcgact 1856

<210> 85  
<211> 1845  
<212> ADN  
<213> *Arabidopsis thaliana*

5

<400> 85

ES 2 640 100 T3

atatccattc tcaccttcca aatctcccca ctctctctct ctctctctct ctctctctct 60  
 ctctctctct ctctatctct ctctatctct ctctatcttc tcttttctcc tctctccaac 120  
 ttcttcaacta atgggagctc aggagaaacg ggcgcgttcc gagcagatat caaagtgcga 180  
 tgtaaggac cggccaacc ataccgtggc cgctgatcta gacggaacac tactaatctc 240  
 tegtagcgcc ttccttact atttctctgt agccctcgag gcagggagct tgctccgagc 300  
 gttgatccta cttgtgtccg taccattcgt ttatcttacg tacttgacca tctccgagac 360  
 tttagccatc aacgtatttg tcttcatcac gttcgcgggt ctcaagatcc gagacgttga 420  
 gctagtggtc cgttcogtcc tcccgagggt ctatgocggag gacgtgaggc ccgatacctg 480  
 gcgtatcttc aacacgttcg gaaaacggta cataataact gcgagccctc gaattatggt 540  
 cgagccattc gtgaaaacat tccctaggagt tgataaagtt cttggaacag agctagagggt 600  
 ctccaaatcg ggtcgggcaa ccgggttcac cagaaaacca ggtattctcg tcggtcagta 660  
 caaacgtgac gtcgttttga gagagtttgg tggcctagcg tctgatttac ctgatttggg 720  
 gctcggcgat agcaagacgg accacgactt catgtccatc tgcaaggaag gttacatggt 780  
 gccacgtaag aatgogaac cattaccaag aaacaaactc ttaagccca taatattcca 840  
 cgagggcaga ttagtccaac gcccaacgcc gttagttgct ctgttaactt tccctctggct 900  
 tcccgtcggg ttctctctct ctatcatcog cgtctacacg aatattcogt tacoggaacg 960  
 tatcgccogt tacaactaca agcttactgg catcaagcta gtcgtcaacg gccaccctcc 1020  
 tccgccgcca aaacctggcc agccaggcca tcttttggtc tgcaaccacc gcaccgttct 1080  
 cgatcctgtg gtcacagctg tcgcactcgg ccggaaaatc agctgcgtca cttacagcat 1140  
 cagcaagttc tctgagctaa tctaccaat caaagccgtt gcggtgactc gtcaacgtga 1200  
 gaaagacgca gcgaacatca agcgtctttt ggaggaaggc gatctcgtga tatgtcccga 1260  
 gggaaaccacg tgcggtgagc ctttctctct ccggtttagt gctcttttcg ctgagctcac 1320  
 ggaccggatc gttcccgtgg cgatcaacac aaagcagagc atggtcaatg gtaccaccac 1380  
 acgtggatac aagcttcttg atccttactt tgcgttcatg aacccgaggc cgacgtatga 1440  
 gatcacgttc ctcaaacaga ttccagctga gctgacgtgt aaaggaggca aatctccgat 1500  
 agaggttgcg aattacatac agagggtttt gggaggaacc ttaggttttg agtgcaccaa 1560  
 tttcacaaga aaggataagt acgcaatgct tgctggtact gacggtaggg ttccggtgaa 1620  
 gaaggagaag acgtgattat tgatcggagg atggagaaaa caaagaaata atacaaagat 1680  
 tggaatcttt ttcatttggg tggttaatta tcaatactta aggggcaaag aattgataag 1740  
 tttgtttttg ttgttacctt tatcttctct gcttcggtga atgtttcttc aagttttaca 1800  
 tttagtacaa taacggttta agaaaagaaa gaaaaaagta tcggt 1845

5 <210> 86  
 <211> 1216  
 <212> ADN  
 <213> *Arabidopsis thaliana*

10 <400> 86

ES 2 640 100 T3

tcacttcttg gtggtgttga tagactccac cttgccgtca ttacctcaa gcaaaagata	60
tttatccttg cgagtaagtt cgggtgcattc gaagcccaaa accgogccga taactttctg	120
gacgtaatta gccacctota taggagtctt gccaccaccg ttgacagtca tttcttcagg	180
caaacgatcc aagaaagtgg cttcatagct tggctctggg ttcatagaaga agaagtaagg	240
gtcccaaaac ttcacacccc taactgtggt cccgttgaac attccttgtt tacagttcat	300
cgctactggc acaatccggg cgcttagctc tgcgaataga gcgctaaatc tcagtagata	360
ctcttctcta cacgtcgtgc cttccggaca tatcaccaag tcgcctgtta tacaccaaaa	420
acatatcatt agctcaacaa ctccattgat ttttaatttc tcaactaaag attaagtga	480
atgtaagtca atattcaaaa gaagtaaaca atattatfff atgttcattt atataaattt	540
ggttagaaga tgcgttaatt aggtacccaa gtaaacatca aagaccatac caaagtaata	600
tttactatat tgagatttgc atgggctact atatatgtat tcatcgaaa ataaaaagaa	660
agaaatagaa actcctattg accccctgg acagtgaac ttgttatagg tatgctctag	720
taatatatga tgagaagatg tgtaaagtac ctttctcgag aagttttctc atgttggcag	780
catcggtggc acggtcacgg gtgagggcaa cagcaggaat aggagaaagc ataagggaga	840
gacgagagac actgtaagtg acgcaacaga tcttacgtcc aagagcaata gcaacgatga	900
tgggatcaag cgcggtacgg tggtaagga catagaggtt gccaaagatt ccaggggaag	960
gaggtggagg acgatgacca cgaatggta agtggatccc gagcatctcg taagtgtaac	1020
ggacaaatct ttcaggtaaa gggagggtga agtagacgcg aatgatggag aggatgaaac	1080
caaaaggaag ccataggtat gtgataatgg cgtttaaagg agttggacgt tgcgctaaac	1140
gccatcatg gaagactatg cggttcttta agcgttcttt tggaatcgtt gtggctgact	1200
tggtcgcattg aacct	1216

- 5 <210> 87
- <211> 1736
- <212> ADN
- <213> *Arabidopsis thaliana*
- 10 <400> 87

ES 2 640 100 T3

atctctcttt ctctgtttac taaatcgcca gtattcgccg caagaaagat gtctccggcg 60  
 aagaagagcc gaagttttcc tccgataagc gagggtaaaa gcagagagta tgattcgatc 120  
 gctgcggatc tcgaocggac tctgcttctg tcaagaagct ctttccctta cttcatgctg 180  
 gtggccattg aagcaggtag tctcttccgt ggattgatcc ttcttctttc tcttccatc 240  
 gtcacatcgc cttatctctt cgtctccgaa tctctcggaa tccaaatcct tatcttcatc 300  
 tccttcgccg gtatcaagat caagaacatc gaactcgtct ctgcgcgccg tcttacacgg 360  
 ttttacgccg cggatgtgag gaaagatagt tttgaggtgt ttgataaatg taagaagagg 420  
 aaagtagtgg tgactcggaa tccgatagtg atgggtgagc cattcgtcaa agactatttg 480  
 ggaggagata aagttttggg aacagagatt gaagttaacc ccaaacgat gaaagctact 540  
 ggttttgtga aaaagcctgg tgttcttgtt ggtgatctca agagattagc catcttgaaa 600  
 gagttcggcg atgactcacc ggatcttggc ctccgtgacc gcacctccga tcacgatttc 660  
 aagtcattt gcaaggaagg ttacatggtg catgagacca aatcagccac aacagtcct 720  
 atagagagtc tcaagaaccg cataatcttc catgatggcc gtcttgtcca acgtccgacc 780  
 ccattaaacg ccttaatcat ttaccttgg cttcctttcg gcttcatgct ctctgtcttc 840  
 cgcgtctact tcaacctccc tttaccggaa ogcttcgtcc gotacactta cgagatcctc 900  
 ggcatcacc tcacaatcog tggccaccgt cctccacctc cttccccog aaaacctgga 960  
 aacctctacg ttctcaacca ccgcaccgcc cttgacocca ttatcatogc cattgctctc 1020  
 ggccgtaaga tcacatgtgt cacttatagt gtctctcgcc tctccctgat gctctcacc 1080  
 attcctgctg ttgctctgac ccgagaccgt gtcgctgacg cagcccgcac gaggcaactc 1140  
 ctcgagaaag gtgatttggg gatctgtcct gaaggcacia cgtgtagaga accatatcta 1200  
 ctaaggttca gtgctctatt tgcagagctg agcgaccgga togtgcctgt ggccatgaac 1260  
  
 tgcaagcaag ggatgttcaa cgggacgaca gttagaggtg tgaattctg ggatccgtac 1320  
 ttcttcttca tgaacccccg acctagctac gaggccactt tcttggaccg tttgcccga 1380  
 gaaatgacgg taaacggtgg tggcaagact ccttttgagg tggctaatta cgttcagaag 1440  
 gtgatcggtg gaggtttggg gttcgaatgt actgaactta caaggaagga taagtatctg 1500  
 ttgcttggag gcaacgacgg aaaggtttag tctatcaata agaccaagtc catggagtaa 1560  
 taatccagct agtctggtg ctgctatctt gcttgggagt gaagttaggt ctttcgggtt 1620  
 ttccaatta cttttttttt tgtttttttt catgtattgt tcgaggaaaa tgttttgtt 1680  
 aatatgcact ggataatgaa actacatttg aatcaaagta ogtacatgta gtgtac 1736

<210> 88  
 <211> 1817  
 <212> ADN  
 <213> *Oryza sativa*

5

<400> 88

ES 2 640 100 T3

ttaacatccc atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcgtt 60  
 gccggcgagc gccctgtact tgtccttcct ggtgaagctg gtgcactcgt aggagagcgt 120  
 ggaggcgatg agcctctgga tgtagtggc cacctcgtgg ctgctccggc cgccgccgcc 180  
 gccgccaccg ccgttgcagg tgagctgcc ggggagcttg ctgaggaagg tgaccacgta 240  
 tccggggcta gggttcatga agaagtagaa ggggtccagc gccttcacc ccctcgcgt 300  
 cgtgccgtgg aacatgctca tctggttctc catcgccacc ggcacgatct cgtcgggtgag 360  
 ctccgggaac aggccgaga accggaggag gaacggctcc cggcacgtcg tcccctccgg 420  
 gcagatcacc aggtcgccct ccgccagcag ccgccggatc atcgccgcgt ccgccgccct 480  
 gtcccgcgtc agccgcaccg tcctgatcgg cgacaatc tccgacagcc gcgacacctg 540  
 ttcagacacg caccaagatt tttattttta aaattttgaa caacttaaat tttaaaattt 600  
 aaaaataaac ctttctaaga tttattatca acatctaaac ttttttgacc atgttagccg 660  
 agttggcatc atacaggtct ggctagccta acagcgttgt cctgccacac caccctaggt 720  
 aacgtagcag acaaaaactg tcacgottga cagcagctg tcattttgcc atgttagaca 780  
 caggctggcg tgattaaaat gttcaacttt ctaaataaag ctatagaata gtttgttttt 840  
 aatataaaaa aagaaaaatt caaaagtaa tataaattcc atgcaccggc gagttagctg 900  
 gtgagctagc tgctcgccgg agttggtggt ggtgtgatca tgtgtactac cgagtaggtg 960  
 acggcgggtg tggggcggcc gagggcggtg gagaggaaga tggggtcgag gagggtcggg 1020  
 tgggagcaga tgaagagcac gccggactgg ccggtctcgc ggctggccgg tggcgggggg 1080  
 gtgccottga cggtgacggc cacgccgagg gcgcggaagg cgtggtacac cacgcgcctc 1140  
 gggaggagcg cgccggcggc gatgoggagg caggcgagca cgaagccgat ggggatccag 1200  
  
 agcacggtga gcagcgcgag ogccggcgac ggcttctgga ccaggcggcc gtcgtggaac 1260  
 accaccggtc tcgggagctc ctcccgggc acgggcctcg gcgtcggcga caccggcacc 1320  
 acgtaccct ccttgcatag cctcatgaac gggtagtccg tcctccggtc gccgatgccg 1380  
 atctccggcg aggcgtcgcc gaacgcctcc cggagcgcgg ccgccttgtt ctgcgccagc 1440  
 agcacgccgg ggctccggac gagccccgtg gcgcggccac gccagaccac gagctccgtg 1500  
 cccacgacga tgtcggcgcc gaggtactcc ttcaagaacg cctccaccat catcctcggg 1560  
 ttccgggtga gcacgcaccg gcgcccgcac gccgagaaca cgcgccacga ctccgggtgg 1620  
 atgtcggagc agtagaactt gggcagcacg gcgcgcgcca ccgcctcaac gtcggcgacc 1680  
 ctgcgcggcg ccatggacgc gaagatgagc acctggatgc cggccgactc ggacaccagg 1740  
 tagtagagca cgccggcgag cggcgccagc aggatcagcg ccagcagccg gaggacgccc 1800  
 cccgtctcga acgcat 1817

5 <210> 89  
 <211> 1924  
 <212> ADN  
 <213> *Picea sitchensis*



ES 2 640 100 T3

<400> 89

ccccccggtg	cgttttgttt	gagaaggaat	atggcgaagg	gccagagtgg	cctcgactgt	60
caagggctct	ttgatgtgat	tagcaaatgc	agctccaagg	gccgagaaaa	tcagaccgtg	120
gctgcggtac	ttgatggcac	catgctcaga	tccaggagct	cgttccotta	tttcatgctc	180
gtggccatgg	agggcgggag	cctgttacgc	ggcatggtgc	tctctgcctg	cgctcccatg	240
gcctggtttc	tctacaactt	catctccgaa	gctgccccga	ttcagcttct	catctttatt	300
tccttcgctg	gactcaagat	ccgggatatc	gagctcgttt	cccgcgccgt	gttgcccaag	360
ttctatgccg	aggacgtgca	cccggagtca	tggagggtct	tcagcgcctt	tgggaagcgg	420
tacatogtga	ctgcgaatcc	aagaatcatg	gtggagccct	tctggaaggg	ctatttgggg	480
gctgacaagg	ttcttggaac	ggagctgcag	gtttccagga	gcgggcgggc	taccgggttc	540
gtgaagaagc	ctggtgtgct	cgttggggat	ttgaagaagg	cagctgtgga	agcggaattc	600
ggtgataaat	tgctgagct	tggccttggg	gatcgagaga	cagatcatcc	tttcatgtcc	660
ctctgcaagg	aggggtacat	ggtgcccaag	atgaagggtg	acgaagtgcc	cacaaacaaa	720
ctaatgagcc	caatcgtctt	ccacgacggg	cgtttggtgc	agagaccgaa	ccccggggcc	780
gcccttctca	cctttctatg	gctgcccata	ggcttctttt	tggctcagtt	tcgagtcttc	840
ggcaacattc	ccatcccaga	gaaatacgtg	aaaatctcct	acaagataat	gggaatcaag	900
ctggtogtca	agggaaagcc	tccgccggcc	ccgaagaaga	aaggcgagcg	ggcgtgctg	960
ttcgtatgcg	atcacagaac	gttgctggat	cccgtgatcg	tggccgtggc	cctgggaagg	1020

ES 2 640 100 T3

aaagttagcg ccgtcacgta cagcatcagt cggttctcgg agattatctc gccattaaa	1080
accgtccggc tcaactcggga ccgcgagagg gacgccgcta acatcaagcg cctgctggag	1140
gagggcgacc tggatgatctg cccggagggc accacgtgca gggagccctt cctcctgagg	1200
ttcagtgtctc tattcgggga gctcaaggac aggatagtgc cagtggccat ctgcaacaag	1260
atgagcatgt tccacggaac aacggtgctc gggtggaagg gcttcgatcc ttttttctt	1320
ttcatgaatc cgggtgccac ctacgaggtg acgttcttgg atcagttgcc tacggagctc	1380
acgtgctcgg gaggcaaate tcccatcgaa gtcgctaatt acattcagag ggtgctggcg	1440
gccactcttg gctacgagtg caccaatttc accaggaagg acaagtaccg gatgctcgca	1500
ggcaacgatg gcattgttcc ccttaataag actgctactg ccgcagcaa caaccacaaa	1560
cagaattgat ccaccgaatt ccaccgaatt ccaggccggg tattcttaat catgataaat	1620
tttaataatc tcttattaat ctatgattcc aattattaca tactgcaaat ttgaacgcgt	1680
agttaagcag aatcttccca cgaatgaatg tatacacctc ataattggta tatctgtatt	1740
gcaaattaga tcagatcgta gttaaagaag aatccacta atggttgctc cataattggc	1800
acatctggtc attatattgc aaattattcc agtaatttgc ctggactgta ttatgctcaa	1860
ggatcagaaa taatataatg cccaaggatt tttttgtcca aaaacaagaa aaaaaaaaaa	1920
aaaa	1924

5

<210> 90  
 <211> 1765  
 <212> ADN  
 <213> *Zea mays*  
  
 <400> 90

ES 2 640 100 T3

cgcacgcagc agtacgacgt ctctcctctg ggtctggggc cgagacaccg agcacgtact 60  
 accagcaaga tgggtggcgtc tcccagattc aagcccatcg aggagtgctg ctccggagggg 120  
 cggtcggagc agacggtggc cgccgacctg gacggcacgc tgctcatctc caggagcgcg 180  
 ttcccctact acctcctcgt ggctctcgag gccggcagcg tccctccgcg cgcgctgctg 240  
 ctctgtcog tgccgttcgt ctacgtcacc tacgccttct tctccgagtc gctggccatc 300  
 agcacgctgg tgtacatctc cgtggcgggg ctcaaggtgc gcagcatcga gatggtggcg 360  
 cgctccgtgc tcccaggtt ctacgcgggc gacgtgcacc cggagagctg gagggtgttc 420  
 agctccttcg gcaggcgta cgtcgtcacg gccagcccca gggcatggtt cgagcccttc 480  
 gccagggcct tcctcggcgc cgacaaggtc gtccggaccg agctggaggt cggcagggac 540  
 ggcaaggcca cgggcttcat ggccaggcca ggcgtgctcg tcggcgacca caagaagaag 600  
 gccgtcgtca aggagctcgg cgacgcgctg cccgacgtcg gcatggggga tagggagacc 660  
 gacttcgact tcatgtccat ctgcaaggag gcctacctgg tgacgtcaag gaagtacagc 720  
 ccggtgcca ggaaccagct gctgagcccg ctgatcgtgc acgacggccg cctcgtgcag 780  
 cgcccgacgc cgtcgtcgc gctcgtcacc ttctctgga tgccgttcgg cttcgcgctg 840  
 gcgtcatgc gcgtgtacat caacctgcog ctgcccgagc gcacgtcta ctacacctac 900  
 aagctcatgg gcatcaggot cgtcgtcaag ggcaccccg cgcgcgcc caagaagggc 960  
 caccocggcg tcctcttcgt ctgcaaccac cgcaccgtgc togaccccg ctaggtggcc 1020  
 gtggcgctgc gccgcaaggt cagctcgtc acctacagca tctccaagtt ctccgagctc 1080  
 atctcgcca tcaaggcctg cgcgctcgt cgggagcgcg acaaggacgc cgagaacatc 1140  
 cgccgcctgc tggaggaggg cgacctggtc atctgccccg agggcaccac ctgccgcgag 1200  
 cccttcctgc tgcgcttcag cgcgctcttc gccgagctca ccgaccgcat cgtgccctg 1260  
 gccatcaaca ccaaggagac catggtccac ggctccaccg tgccgggctt caagotcatg 1320  
 gatocctact tcttcttcat gaaccocgoc ccgacgtacg agatcacggt cctcaccag 1380  
 ctccccaaag acctcacgtg cagcggcggc aagtgcocca togaggtggc caactatctc 1440  
 cagaagacgc tcagcggaca gcttggttc gagtgcacgt ccatcacgcg caaggagaag 1500  
 tacggcatgc tcgctggcac cgacggccgt gtcccgtcca agaacaagga gaaggagaag 1560  
 gagaaggata agaactaacg gccaacgtcg ctctcttctg gctgccccagc agactgccgt 1620  
 ttaattactg ctcaaattta ttccaattta ttgttccttg tcgacaagtt ggccagtggt 1680  
 atgtgggccc tgtactagta caagcttgat gatgattagt atatctaaa ttcttccttt 1740  
 ccggaagaag ggcggaagag ttggtt 1765

<210> 91  
 <211> 1778  
 <212> ADN  
 <213> *Arabidopsis thaliana*

5

<400> 91

ES 2 640 100 T3

agagacacac acacagcgag tgaaaatgtc tccggagaag aagagtcaaa acttccctcc 60  
 aataacggaa tgcagagacg gagagtacga ttgatagcc gccgatcttg acgggactct 120  
 gottctctca agaagctcct toccttactt catgctogtc gctggtgaag ctggaagcct 180  
 tttacgtgga ctaatcctcc ttctctcggt accattcgtc attatctctt acctattogt 240  
 atccgaatct cttggtatcc agatcctcat cttcatctca ttcgctggtc tcaaaaatccg 300  
 cgatatcgaa cttgtctctc gcgcagttct tccacggttt tacgcggcgg atgtgaggaa 360  
 agacagtttt gaggtgtttg ataagtgtaa gagaaaagtg gtagtgacgg cgaatccgat 420  
 tgtgatgggt gaggcgtttg tgaaggatta tcttgaggt gataaagttt tgggaacaga 480  
 gattgaagtt aaccctaata ccaatagac cactggattt gtgaagaagc ctggtgttct 540  
 tgttggtgat cttaaagagt tagccatttt aaaagagttt ggtaacgaat cacctgatct 600  
 cggcctcggg gatcgaaacct ctgatcatga tttcatgtct ctctgcaaga aaggttacat 660  
 ggttcatgcg accaagtcag ccacaacgat tccaaaagaa cgcttaaaga accgcatagt 720  
 cttccatgat gggcgtttag cgcaacgtcc aactcogtta aacgccatta tcacatacct 780  
 atggcttctt tttggtttca tctctccat cattcogctc tacttcaacc tccctttacc 840  
 tgaagattt gtcggttaca cttacgagat gctcgggatc cacttaacca ttcgtggtca 900  
 tcgtcctcca ctccttccc ctggaactct tggcaacctc tatgtcctta accaccgtac 960  
 cgcgcttgat cccatcatcg ttgctattgc tcttgacgt aagatctgtt gcgtcactta 1020  
 cagtgtctct cgtctctccc ttatgcttcc tctattcct gctggtgcc tcaccctgta 1080  
 ccgtgccacc gatgctgcca acatgagaaa acttctcgag aaaggcgact tggtgatatg 1140  
 tccggaaggc acgacgtgta gagaagagta tctactgaga tctagcgtc tattcgcaga 1200  
 gctaagcgac cggattgtgc cagtagcgat gaactgtaaa caaggaatgt tcaacgggac 1260  
 cacagttagg ggtgtgaagt tttgggacct ttacttctc ttcattgaacc caagaccaag 1320  
 ctatgaagcc actttcttgg atcgtttgcc tgaagaaatg actgtcaacg gtggtggcaa 1380  
 gactcctata gaggtggcta attacgtcca gaaagttatc ggcgcggtt tgggcttoga 1440  
 atgcaccgaa cttactcgca aggataaata tcttttgctt ggaggtaatg acggcaaggt 1500  
 ggagtctatc aacaacacca agaagtgaat atottattct tgttataact tggcattgac 1560  
 ttattgcaag taaggttgga tcatatacga attttcatgt gtgtaattgc ttgagctctt 1620  
 gtcacatta tggtagtcta ttgttatggt tattctattt tgttcttaat gtcgaacggg 1680  
 taactatata cttaaaaact attttgtgat cccggttcta tacggatgtg tatagagact 1740  
 ttctatttaa agtgcaaaat tttattggag aaaaaaaa 1778

5 <210> 92  
 <211> 1832  
 <212> ADN  
 <213> *Oryza sativa*

10 <400> 92

ES 2 640 100 T3

ggctagctgc tccgtgtctc cctcgttgca tggtaagcta gcatgtogcc gttcaagccg 60  
 atcgagcagt gctcgacgga ggggcggtcg cagcagacgg tggcctcoga cctcgacggc 120  
 acgctgctcc tgtcccgcag cgccttcccg tactacctcc tcgtcgcgct cgaggccggc 180  
 ggccccctcc gcgcggtcgc gctgctcatg tccgtgccct tcgtctacct cacgtacgtc 240  
 accatctccg agtcctcgc cgtgcgcgcg ctgctgtaca tcgccgtggc gggctctcgag 300  
 gtgagggacg tcgagtcggt ggcgcggtcg gtgctgcca ggttctacgc cggcgacgtc 360  
 caccgggagg ggtggcgggt gttcagctcg ttcggcagga ggtgcgtcgt gacggcgagc 420  
 cccagggatga tgggtggagcc gttcgcgagg gcgttcctcg gcgcggacag ggtcatcggc 480  
 acggagctgg aggtcggcga ggacggcagg gccaccggct tcgtcgccaa gcccgccgtg 540  
 ctcatccgcg agcacaagag gaacgcgggtg gtgagggagt tcggcgacgc gctgccggac 600  
 gtcggcatgg gggacagga gagcgacttc gacttcatgg ccatctgcaa ggacgcgtac 660  
 gtcgtgacca cctcgcgaa gcaccggccg gtgccggaga gccagctgct gagaaccgtc 720  
 gtcctccacg acggccgcct cgcaccagcgc ccgacggcga tcaacacgct gctcgtcttc 780  
 ctctggatgc cggtoeggct cgcgctcggc ctctccggcg cgtgcctcag cctgctctc 840  
 cccgagcggg tcctctccta cgcctacaag ctacacggcg tcgggctcgt ggtgcgcggc 900  
 cgcgccccgc cgcgggacgg gagccccggc gtgctcttcg tgtgcaacca ccgcaccgtg 960  
 ctcgaccagg tcgccgtcgc cgcgcgctg gcccgcaagg tgatctcgt cacgtacagc 1020  
 gtcctcggga aaacgtacgg catgagctcc aggtcccgg aggcgctcac ggctcgcog 1080  
 gtcaagcccg ccgtcgcgct gtgcaggag cgcgacaggg acgcggaccg tgtccggcgg 1140  
 ctgctggagg agggcgtcga catcgtcgc ttccccgagg gcaccacctg ccgaggggog 1200  
 ttctgctgc ggttcagctc gctgttcgcc gagctcaccg accgcctcgt gccgggtggc 1260  
 atcgccacca gggagaccat gttccacggc tccacggcgc gcgggttaa gggcatggac 1320  
 ccttacttot tcttcatgaa cccgcggcgc gcgtacgagg tcacgttctt gagccagctg 1380  
 ccgagtgagc tcacctccgg cggcggcggc aagtgcggcg tcgaggtggc caactacgtg 1440  
 cagaagggc tcgcgggca gctggggtcc gagcacatcg gcatcacgcy caaggagaag 1500  
 tagcgaatgc ttgcggcac cgcggccgt gtggagtcga atgaggacta gaaccagaag 1560  
 ccacacccc actattattg gttccaatat atatggccag ggcctcgtt tcttcgttat 1620  
 ttctgtcaca taacaagctt caagtgtta tacgagaact gaagatgcaa catagtagaa 1680  
 ctcatgttgt tcgttgtgga cttgtggtat aacgactttg atcatttctt tgcctttta 1740  
 tttttacaaa aatccaagaa atatagtcat ataagttca tactgatgta aattctttg 1800  
 ctatataata gatttgcta tgtttttttt tc 1832

5 <210> 93  
 <211> 2529  
 <212> ADN  
 <213> *Oryza sativa*

10 <400> 93

# ES 2 640 100 T3

tcaattggcc ttgggcttgt cggcggcggc ggcgaccgg ccgctggtgc cggcgagctt	60
catgtacttg tccttcctgg tgagcgtggt gcattggaac cctagctoct tggcgatcac	120
cctctgcacg tggttcgcca cctocaccgc gtcctcccg cggcgaagc acgtctctc	180
cggcctcagc gccggcagga acgtcacctc gtacccggc cgcgggttca tgtagaagaa	240
gtaagggtcc aggaacttcc accccctcgc cgtcgacca tagtacgtcc cctgcctcgc	300

ES 2 640 100 T3

ctccaccggc accggcacga tcctcgccgt cagctccggg aacagcgccg agaaccggag 360  
caggtacggc tcccggcacg tcgtcccctc cgggcacacc accacgtccc cctcctccag 420  
cagcgcggcg atgcgcggcg cgtecgccgc gcgggtccgg gtcagcggcg ccgcgcggat 480  
cggcgagatc gccgtcgaca gccggctaac gctgtacgtc acgcacgtca cggggcggcc 540  
cagcgcgatg gacacgatga tggggtcacg cgcggtgcgg tggttgcaca ccaggagggg 600  
gcccgggctc cccggggcgcg gcggcgccgg cggcgccccc cgcacggcga ggcggtatgc 660  
ggtgaggcgg taggtgtggc ggacgaggcg cgcggggacg gggaggttga ggaacacggc 720  
gaggagcggc accggcgaag ccacggggag gtaggccacg gcgaacagcg cgctcgctgg 780  
ctccggggcg cggacgaggc ggccgtcgtg gaagatggcg cgcgacagca gctcgtcgcc 840  
ggcggcgccg ggcgcgcgct tgttcttggg caccatgtag gcttcctggc cattagcaat 900  
ttggcgccaa aacagagtta gcacactaag gctgtgttta gttcacatc aaaattgaaa 960  
gtttggagaa attagaacgg tgtaacagaa aagttggaag tttgtgtgtg tagaaaaatt 1020  
cgatgtgaca gaaaagttaa aagtttaaag aaaaaagta aatctaaaca aggcctaatt 1080  
tttacaataa tctggctttg gagtatattc atatgtacat tttgaacagt tgtcttcttt 1140  
ttttttgtta aataaaaaa aatattgatg atggactagg ttgcattctg aatgtaagca 1200  
tgaaatttta gtttatgata gtctaaccaa ttaaaaattg ctagtgttca aagttgaaac 1260  
agtagagtca atattcagag taaatgcaa tggagagagt actaaatcca cattaattga 1320  
tgcaaagaaa acaatactat atgagtaaag tgatcaatga taatgatatt actataacaa 1380  
gtcagcatga gcaacaata tttgaatata ggagtacata ctaatttggg gttttggttt 1440  
taaaaaggac gaaataaata ggaagcaact gctaactgcc tactaaagga ggtgatggta 1500  
ccaacagtat tgagtttttg ccacaagttt tgcataacca taatgtaaat tctagagcct 1560  
tottgatctt tccatgaaa agaggggaaa aaaaaggcaa ttcaacaaaa agaaaaggca 1620  
aaaagctaag aagcgtacat attcctgcta ctcgtgctac cacaaatgag ccaatcggat 1680  
cgaagccatg aacctcga gaacaatga gcgtgcaatt cctaccaca ttaatttcac 1740  
aatattttct tttttatth ttggcgttct aatgaattg ttaaattatt agtggccaca 1800  
tgatgcaaaa cttcaattht gttcatgtgt gattgattca attcaagaaa tgatgaaat 1860  
ttcgagctth ggatcgtagg ccatgaagtg taccttgacg atggccatga agtcgtggtc 1920  
gctctcggg tcgcccggc cgacgtcggg catgtcggcg ccggcgaaga gcctctgac 1980  
gacctcggc ttcttctcgc cgacgaggac ggcccttgat cggccgggtga accgcttgc 2040  
ggaggcgaac gtctcgagct cggtgccggc cacctcggcg ccgaggaact cgcgcacgaa 2100  
ctccccacc atgacggccg gcgacgggtg gaccacgacg cggcgtcgg cgcagccgtg 2160  
gaacaccggc cacgtgtcgg cgcgcacccc ggccggcgtg tggcggggca gcacggcgcg 2220

ES 2 640 100 T3

cgccacggcc tcgacgtcgc ggacgogcag ccogggcaac gtgatgaaca cgagcagggc	2280
gatggcggcg gcctcggaga cgccgacgta gagcgcgagg agccacggcg cggcggagcag	2340
cagcgccagc gcgcggaggt agctgccggc ctogagagcc acgaggaagt agtacgggaa	2400
cgccgacgac gacaccagca gcgtcccgtc gaggtccgcc gccaccgtcc gcctcgcgcc	2460
cgccgcgcgc tcgtaogccg tcaccggcgg aaagaacctc ctccctcccc cgccgcgcgc	2520
cgccgccat	2529

5  
 <210> 94  
 <211> 3320  
 <212> ADN  
 <213> *Oryza sativa*  
 <400> 94

atgtcgcctg tcaagccgat caagcagtc tcgacggagg ggcggtcgca gcagacgggtg	60
gcctcogacc tcgacggcac gctgctcctg tcccgagcg ccttcccgtc ctacctctc	120
gtcgcgctcg aggcggcgcg cccctccgc gcggtcgcgc tgcctcatgtc cgtgcccttc	180
gtctacctca cgtaagtcac catctccgag tccctcgcgc tgcgcgcgct gctgtacatc	240
gccgtggcgg gtctcaggt gagggacgtc gagtcggtag cgcggtcggt gctgccagg	300
ttctacgccc gcgacgtcca cccggagggg tggcgggtgt tcagctcgtt cggcaggagg	360
tgcgtcgtga cggcgagccc cagggtgatg gtggagccgt tcgogagggc gttcctcggc	420
gcggacaggg tcatcggcac ggagctggag gtccggcagag acggcagggc caccggcttc	480
gtcgccaagc cgggcgtgct catccgcgag cacaagagga acgcgggtggt gagggagttc	540
ggcgacgcgc tgccggacgt cggcatgggg gacagggaga gcgacttcga ctctcatggc	600
atctgcaagg tacggacaaa gaatattgcc tattccctta actacttact actacctaa	660
ttccaaaaat agtttgcttt tacacatgga caactgctg tctaaatgta gaaaaatgag	720
ttctttttta aaaagtagag gtaggattgt actagcacac atgttcatat tacgtgtgtc	780
ctagtatttg gtacagtaca tttataacgg caagtagctt atagtgaac ctgcaggcag	840
catatctgca tatggaattt tgogatgtgg ccgcccaatt atgtacttag tatttttttt	900
aaaagagatg tacttagtat ttogcataca attttgctat atatttgacg acaaatttta	960
ccacaaaaat ttgacagata taattacagt acaatcgtag tgtaattaca ctgtaacttg	1020
catgtaatta cactgtaact atagtgaac ttgtatgtaa ttttcaaaa tatctccgta	1080
acatgttatt tcgataaaat gaggttgtgg gaataaatc tttcatatat gtgtgctatg	1140
atgtgttttc ttccctacca aaacaaatct tgtaatatag ctaacgattc aaaattatgt	1200
gaaacttaca tataagttac actgtagtta catgcaagtt acagtgtaat tacatataag	1260
ttacagtgta attacactac cactgtacta taattacatc tgtcaaattt ttaggagaaa	1320

10



ES 2 640 100 T3

atttgtcgac aaatgtatag gtggtccctt tggcatatat aagtaggta attgcaacag 1380  
 tacaacaat atattattaa ttccttgtag tatgtaatac actgaatagt acaacgagta 1440  
 tgcattttgt tcggtgccat ttatgctgag gataattgta gtatggcatt aatcgatctg 1500  
 gaccaatctc tocacgtag gtgtctgcat tcaccgggta gatgatcagt tgagaggtag 1560  
 ttgtgccatt gatatcgtcc actgctttct cgttatcctt tctctgtagg tgactctgaa 1620  
 tctagctctt ttcacggatt gcatcaactg atatacgtcg tattcttccc aagttccctg 1680  
 cattcatgtc ttatatatgc taactggaca attaatggct catagcaggg cagctgcctg 1740  
 tcagtgatca ctocagcttg gttgaatggg tgagttagcc tgtctgtttc atgatatctg 1800  
 gtacaggtcc cactcaatac cgtttgtag gaatttatac ttgtgatgta cttatgtcct 1860  
 caaataagt cagggagaat atgaacgtaa gcaaatttgt gattgctaga cagacatgaa 1920  
 aaaccaaagt ctagccatgc atgatctaca gatttccttg tttcttccct ccaactttgg 1980  
 gaagtttgca ctgaattaat acggttttgc ttcattgtag gccattatat ttgcaagtaa 2040  
 tggctctact agtataactg gatcccaaat ttccgatgga gcatcacaat cacgatcata 2100  
 tgcaagatc gatattatgc tgtttggggc aaccgggcaa acatagcacc gggtaaactt 2160  
 tggttcaact tacagttaat cagtagtaa ccatgctgat atgataataa gctgattagt 2220  
 ccgtgtaagt taattaatta acctaccggg catgcatgca agttttggat ggatttacet 2280  
 aatggataag ctggtactct gtttghaat tctaactctac tagcgtcctc aattttttta 2340  
 cttctccatt taaccagcga cccatgcatt attctgtatt tccgttgcatt acttgcattt 2400  
 ccatgcacgt gaactatgca tgacatgatg catcttgatc aacgtggcac gtgcacgcat 2460  
 gcatgcagga cgcgtacgtc gtgaccacct cgcggaagca ccggccgggtg ccggagagcc 2520  
 agctgctgag aaccgtcgtc ctccacgacg gccgcctcgc ccagcgcccg accgggatca 2580  
 acacgctgct cgtcttctc tggtatgccg tcggtctcgc gctcgcgctc ctccgcgct 2640  
 gcctcagcct gctcctccc gagcgcgtcc tctcctacgc ctacaagctc accggcgtcg 2700  
 ggctcgtggt ggcgggccc cgcgcgccgc oggacgggag ccccgcgctg ctcttcgtgt 2760  
 gcaaccaccg caccgtgctc gaccoggtcg ccgtcgccgc cgcgctgggc cgcaaggtga 2820  
 tctgcgtcac gtacagcgtc cctcggaaaa cgtacggcat gagctccagg ctcccggagg 2880  
 cgtcacggc ctgcgggtc aaggccggcg tcgcgctgtg cagggagcgc gacagggacg 2940  
 cggaccgtgt cggggggctg ctggaggagg gcgtcgacat cgtcgcttc cccgagggca 3000  
 ccacctgccg aggggcgttc ctgctgcggg tcagctcgtt gttcgcgag ctccaccgacc 3060  
 gcatcgtgcc ggtggccatc gccaccaggg agaccatgtt ccacggctcc accggcgcgcg 3120  
 ggtttaaggg catggaccct tactttctt tcatgaacce gcggccggcg tacgaggtca 3180

ES 2 640 100 T3

cgttcctgag ccagctgccg agtgagctca cctccggcgg cggcggcaag tcgccggctc 3240  
 aggtggccaa ctacgtgcag aaggcgctcg ccgggcagct ggggtccgag cacatcggca 3300  
 tcacgcgcaa ggagaagtag 3320

<210> 95  
 <211> 2532  
 <212> ADN  
 <213> *Oryza sativa*

5

<400> 95

tcaattggcc ttgggcttgt cggcggcggc ggcgacccgg ccgtcgttgc cggcgagctt 60  
 catgtacttg tccttctcgg tgagcgtggt gcattggaac octagctcct tggcgatcac 120  
 cctctgcacg tggttcgcca cctccaccgc gctcctcccg ccggcgacgc acgtctcctc 180  
 cggcctcagc gccggcagga acgtcacctc gtaccccggc cgcgggttca tgtagaagaa 240  
 gtaagggtcc aggaacttcc acccctcgc cgtcgcacca tagtacgtcc cctgcctcgc 300  
 ctccaccgcc accggcacga tcctcgcggt cagctccgcg aacagcgcgg agaaccgcag 360  
 caggtacggc tcccgcaacg tcgtccctc cgggcacacc accacgtccc cctcctccag 420  
 cagcgcggcg atgcgcggcg cgtccgccc gcgggtcccgc gtcagcggcg ccgcgcggat 480  
 cggcgagatc gccgtcgaca gcggggctaa cgctgtacgt aacgcacgta acggggcggc 540  
 caacgccgat ggacacgatg aaggggtccg aggcgggtcc gtggttgac aacaggaggg 600  
 agccccggcg tccccgggc gcagcggcg ccgccgcgcc ccgcgcacgg cgaggcggat 660  
 gccggtgagg cggtaggtgt ggcggacgag gcgcgcggg acggggaggt tgaggaacac 720  
 gcggaggagc gccacggoga agcccacggg gaggtaggcc agcgogaaca gcgcgctcgc 780  
 tggctccggg cggcggacga ggcggcggc gtggaagatg gcgcgcgaca gcagctcgtc 840  
 ggcgcgggcg cgcggcggc gcttgttctt gggcaccatg taggcttctt gcgcattggc 900  
 aatttgggc caaacagag ttagcacact aaggctgtgt ttagtttcac atcaaaattg 960  
 aaagtccgga gaaattgaa cgggtgtaaca gaaaagtgg aagtttgtgt atgtagaaaa 1020  
 attcgatgtg acagaaaagt taaaagtta aagaaaaag ttaaatctaa acaaggccta 1080  
 atttttacaa atatctggct ttggagtata ttcatatgta cattttgaac acttgtcttc 1140  
 ttttttttc gttaaataaa aacaaatatt gatgatgggc taggttgcac tctgaatgta 1200  
 agcatgaaat ttagtttat gatagtctaa ccaattaaaa attgctagtg gtcaaagttg 1260  
 aaacagtaga gtcaatattc agagtaaatg caaatggaga gactactaaa tccacattaa 1320  
 ttgatgcaaa gaaaacaata ctatatgagt aaagtgatca atgataatga tattactata 1380  
 acaagtcagc atgagcaaac aatattgaa tataggagta catactaatt tggggttttg 1440  
 gttttaaaaa ggacgaaata aataggaagc aactgctaac tgccactaa aggaggtgat 1500

10

ES 2 640 100 T3

ggtaccaaca gtattgagtt ttggccacaa gttttgcata accataatgt aaattctaga 1560  
 gccttcttga tctttcctat gaaaagaggg gaaaaaaagg caattcaaca aaaaaaaaaag 1620  
 gcaaaaagct aagaagcgt catattcttg ctactcgtgc taccacaaat gagccaatcg 1680  
 gatcgaagcc atgaacactc gaagaacaat cgagcgtgca attoctacca caattaattt 1740  
 cacaaatattt tcttttttta tttttggogt tctaaatgaa ttgttaaatt attagtggcc 1800  
 acatgatgca aaacttcaat tttgttcagtg tgtgattgat tcaattcaag aatgatgaa 1860  
 aatttcgagc ttggatcgt aggccatgaa gtgtacctg cagatggcca tgaagtcgtg 1920  
 gtcgctctcg cggtcgcca ggccgacgtc gggcatgtcg ccgccggcaa aaagcctctc 1980  
 gagcacctcg cgttcttct cgccgacgag gacggccttg atgcggccgg taaaccgctt 2040  
 gccggaggca aacgtctcga gctcggctcc ggccacctcg gcgccgagga actcgcgcac 2100  
 gaactcccc accatgacgg ccggcaacgc ggtgacaaga acgcgcgctt tggcccagcc 2160  
 gtggaaaacg gccaggttt cggcgcgaac cccggcggcg tagtggcgcg gcagcacggc 2220  
 gcgcgccacg gcctcgacgt cggcgacgcg cagcccggcg aacgtgatga acacgagcag 2280  
 ggcgatggcg gcggcctcgg agacgcgcg gtagagcgcg aggagccacg gcgcggcgag 2340  
 cagcagcgcc agcgcgcgga ggtagctgcc ggccctcgaga gccacaagga agtagtacgg 2400  
 gaacgccgac gacgacacca gcagcgtccc gtogaggtcc gccgccaccg tccgcctcgc 2460  
 gcccgccgcc gcgtcgtacg ccgtcacggc cggaaagaac ctctctctcc ccgcgcgcgc 2520  
 cgccgcgcgc at 2532

<210> 96  
 <211> 2941  
 <212> ADN  
 <213> *Oryza sativa*

5

<400> 96

ctagttcttt tccttctcct tgttcttggg agggacgcgg ccgtcggctc cggcgagtat 60  
 gctgtacttc tccttgcgcg ttatggcggg gcaactcgaag ccgagctggc cgtgagcgt 120  
 cttctggatg tagttggcca cctcgatggg cgacttgccg ccgctgcaag tgagctcctt 180  
 gggcagctgg ttcaggaagg tgatctcgta cgtcggccgc gggttcatga agaagaagta 240  
 ggggtccatg agcttgaagc cccgcacggt ggagccgtgg aacatgctct ccttgggtgtt 300  
 gatogccacc ggcacgatgc ggtcgggtgag ctoggcgaag agcgcgctga accgcagcag 360  
 gaacggctcg cggcaggtgg tgccctcggg gcagatcacc aggtcgcctt cctccagcag 420  
 ccgccggatg ttctcggcgt ccttctcgcg ctcccgcgac agcgcgacgg ccttgatcgg 480  
 cgagatgagc tccgagaact tggagatgct gtatgtgacg cagctgacct tgcggcgag 540  
 cgccacggcg acctcaaccg ggtcagacac ggtcgggtgg ttgcagacga agaggacgcc 600

10

ES 2 640 100 T3

aggatgtccc ttcttggggg gagggggcgg gttgcccttc acgatgaggc ggatgcccot 660  
 gagctttagt gtgtagaaga cgatccgctc cgggagtggc aggttgacgt acacgaggag 720  
 gagcgcgagc acgaagccga acggcatcca caggaaggtg acgagcgcca ccagcggcgt 780  
 cggacgctgc acgagggcggc cgtcgtggag gatgagtggg ctcagcagct ggttcttggg 840  
 caccgcgctg tacttccttg atgtcacgag gtaggcctcc tgcacagaca cgcagcatgc 900  
 atgttacaaa cacttcactg atgcaaaaat cacagcatca ccatcacttc ttattttatc 960  
 gaatatataa atgcttgctt atctcgccat gtacgccaaa cctacatcct accacgatta 1020  
 caattcagta cagtccacta ctgatggctt cacagatgac agatcttcac atgcttaatc 1080  
 aaaattatgt tttttcggag ttaaactggt ggatttaagt tgacgtgtat cagccaatca 1140  
 atagggcaga gtagaccaga gcagtggtaa agaagtacag cgattactgt accgcaacaa 1200  
 ctgcaagact ggcgccata aatgggtgat ttcccctcca cgagagaaat tcggtatacc 1260  
 cccggggcag tgacagaaaag acgggtccca tggagaaggg aaagaagggg caagaccatg 1320  
 aatggcagta tttatttcag tggttctctg cattaccggc tggaaagaaag gaagaaagaa 1380  
 agggaaagca agggcagccg tggatctggc aggcacagcc agcaaccagct ctgatgtctc 1440  
 acgcgccagc cggccgatct ccttcaatgc ggccgggaca aggcttgaat tctcacaaa 1500  
 caaccgtaaa taatggagcc agccagccgg ccggacggac tgactcgacc gtggcatact 1560  
 ggcgatgagt aatccacgca ttcaatccgg cgggcccagc gcgagtaaaa accatcggct 1620  
 tccttccatg ctctggctct gccgcctcc agattccttt totgatcatt tgctcgagct 1680  
 ttcgagttgt taagcttcag gtgaggatct agtatgtgcg acgtagacgt tttgtgagtt 1740  
 ctgataatga ccttgatcct tattattagt ttactgatta aggatgttgg tggaaacagag 1800  
 cgaccggctc gtgcaggaag cttaaactgg gtttcgcata atgggctgag caacatctac 1860  
 agatcagtgt cataaatggc ctgacaatga ctatacacc ctcgtgttcg cgtgctcgac 1920  
 cactgtggca tactotactc tgatctgaaa ccaactctact ctgatctgaa acctctgcat 1980  
 gtttgtacag attgcagaat ttatagtagg cgggtgtcat agcttcgtca tagcaaaaat 2040  
 caattgtaac tgatgtgcta attccacgct atgaaagtaa tatgtgtgat ggagactata 2100  
 cttaattcca attttcagt ttcttgtttg gaatgcagta attccgaaat aacacagagca 2160  
 ttctatacgg aatttatgta caagatggtc taaccagttc gtagttgccca gtaattggca 2220  
 gtgctactac taattcgtcg caggacttga tcaactgcaat taacaatcaa cacgacagta 2280  
 ctattactat gagcaaaaaa tgagtttgag ggggcgtcaa atttaccttg catatggaca 2340  
 tgaagtcgaa atccgtctcc ctgtcgccca agccgacgct gggcaccgcg tcgctgagct 2400  
 ccttgacgac ggctgctc ttgtggctcg ccacgagcac tccgggcttc accatgaacc 2460  
 ccgtggcctt gccgttcttg ccgaacctca gctccgtccc gaccaccttg tcggcgccga 2520

ES 2 640 100 T3

ggaacgtcct ggogaagtgc tocaccatga tcctggggct cgccgtgatg atgtaccgct 2580  
 tgccgaagga gttgaacacc ctccagctct ccgggtgcac gtccctccgcg tagaacttgg 2640  
 gcagcaccga ccgcgccacc atctcgatgt tcctcacctt gagccccgcc acggagatgt 2700  
 acaccagcgt gctgatggcc agcgactcgg agaagaaaat gtaggtcacg tagacgaacg 2760  
 gcacggacag cagcagcacg acggcgcgga gcacgctgcc ggcctcagc gcgacgagga 2820  
 ggtagtacgg gaagggcctc cggacactga ccagcgtgcc gtcgaagtcg ggggccaccg 2880  
 tctgctccga cctcccatcc gagctgcact cctccacggg cttgaacctc cggacacca 2940  
 t 2941

<210> 97  
 <211> 1930  
 <212> ADN  
 <213> *Oryza sativa*

5

<400> 97

ttaacatccc atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcgtt 60  
 gccggcgagc gccctgtact tgtccttct ggtgaagctg gtgcaactgt aggagagcgt 120  
 ggaggcgatg agcctctgga tgtagtggc cacctcgtgg ctgctccggc cgcgcggcc 180  
 gttgcaggtg agctcgcggg ggagcttgc gaggaaggtg accacgtatc cggggctagg 240  
 gttcatgaag aagtagaagg ggtccagcgc ctccacccc ctgcgccgtg tgccgtggaa 300  
 catgctcacc tggttctcca tcgccaccgg cacgatctcg tcgggtgagct ctgccaacag 360  
 cgccgagaac cggaggagga acggctcccg gcacgtcgtc ccctccgggc agatcaccag 420  
 gtcgccctcc gccagcagcc gccggatcat cgcgcgtcc gccgccctgt ccgcgctcag 480  
 ccgcaccgtc ctgatcggcg acaatatctc cgacagccgc gacacctgt cagacacgca 540  
 ccaagatddd tatttttaaa attttgaaca acttaaatd taaaattta aaataaacct 600  
 ttctaagatt tattatcaac atctaaactt tttgacaatg ttagccgagt tggcatcata 660  
 caggtctggc tagcctaaca gcgttgcct gccacaccac tctaggtaac gtggcagaac 720  
 aaaactgtca cgcttgacac gacagtgtca ttttgccatg ttagacacag gctggcgtaa 780  
 ttaaaatggt caactttcta aaataagcta tagaatagtt tgtttttaat ataaaaaaag 840  
 aaaaattcaa aagttaatat aaattccatg caccggcgag ttagctgggt agctagttgc 900  
 tcgccggagt tgggtgggt gtgatcatgt gtactaccga gtaggtgacg gcggtgatgg 960  
 ggcggccgag ggcgggtggag aggaagatgg ggtcgaggag ggtgcggtgg gagcagatga 1020  
 agagcacgcc ggactggccg gtctcgcggc tggccgggtg cggcgggggt cccttgacgg 1080  
 tgacgcgcac gccgagggcg cgggaaggcgt ggtacaccac gcgcatcggg aggagcgcgc 1140  
 cggcggcgat gccgaggcag gcgagcacga agccgatggg gatccagagc acggtgagca 1200

10

ES 2 640 100 T3

gcgcgagcgc cggcgacggc ttctggacca ggcggccgtc gtggaacacc accggcttcg 1260  
 ggagctctc cggcgacag ggctcggcg tcggcgacac cggcaccacg taccctctc 1320  
 tgcatagcct catgaacggg tagtccgtcc tccggtcggc gatgccgatc tccggcgagg 1380  
 cgtcgcgaa cgcctcccgg agcgcggcgg ccttgttctc gccgacgagc acgccggggc 1440  
 tccggacgag ccccgtggcg cggccacgcc agaccacgag ctccgtgcc acgacgatgt 1500  
 cggcgcgag gtactcctc aagaacgcct ccaccatcat cctcgggttc gcggtgagca 1560  
 cgcaccggcg ccgcacgcc gagaacacgc gccacgactc cgggtggatg tcggagcagt 1620  
 agaacttggg cagcacggcg cgcgccaccg cctcaacgtc ggcgaccctc gcgccggcca 1680  
 tggacgcaa gatgagcacc tggatgccgg ccgactcggg caccaggtag tagagcacgc 1740  
 cggcgagcgg cgcacgagg atcagcgcca gcagccggag gacgcccccc gtctcgaacg 1800  
 ccatgtgcgc gaagtacggg aacgagctcc ggcgcacag cagcgtgccg tcgaggtcgg 1860  
 cgacgatcgt gtgcgcccg cggtccttgg acgaacactt ctccaccgcc gggaaacggcg 1920  
 acgcgatcat 1930

<210> 98  
 <211> 1960  
 <212> ADN  
 <213> *Oryza sativa*

5

<400> 98

ttaacatccc atggccttct tctcgtcgat gttgggcttg gacacgacgg tgccgtcgtt 60  
 gccggcgagc gccctgtact tgccttctc ggtgaagctg gtgcactcgt aggagagcgt 120  
 ggaggcgatg agcctctgga tgtagttggc cacctcgtgg ctgctccggc cgcgccgcc 180  
 gccgccaccg ccgttgacgg tgagctcggc ggggagcttg ctgaggaagg tgaccacgta 240  
 tccggggcta ggttcatga agaagtagaa ggggtccagc gccttccacc ccctcgcgt 300  
 cgtgccgtgg aacatgctca tctggttctc catcgccacc ggcacgatct cgtcgtgag 360  
 ctccgcgaac agcgcgaga accggaggag gaacggctcc cggcacgtcg tccctccgg 420  
 gcagatcacc aggtgcctc ccgccagcag ccgccggatc atcgccgcgt ccgccgcct 480  
 gtcccgctc agccgcaccg tctgatcgg cgacaatc tccgacagcc gcgacacctg 540  
 ttcagacacg caccaagatt tttatttta aaattttgaa caacttaaat tttaaaattt 600  
 aaaaataaac ctttctaaga tttattatca acatctaac ttttttgacc atgttagcgg 660  
 agttggcatc atacaggtct ggctagccta acagcgttgt cctgccacac caccctaggt 720  
 aacgtagcag aacaaaactg tcacgcttga cacgacagtg tcattttgcc atgttagaca 780  
 caggctggcg tgattaaat gttcaacttt ctaaataag ctatagaata gtttgtttt 840  
 aatataaaaa aagaaaatt caaaagtaa tataaattcc atgcaccggc gagttagctg 900

10

ES 2 640 100 T3

gtgagctagc tgctcgccgg agttggtggt ggtgtgatca tgtgtactac cgagtaggtg 960  
 acggcgggtga tggggcggcc gagggcgggt gagaggaaga tggggtcgag gagggtgcgg 1020  
 tgggagcaga tgaagagcac gccggactgg ccggtctcgc ggctggccgg tggcggcggg 1080  
 gtgcccttga cggtgacgcg cacgccgagg gcgcggaagg cgtggtacac cacgcgcac 1140  
 gggaggagcg cgcggcgccg gatgcccagg caggcgagca cgaagccgat ggggatccag 1200  
 agcacgggtga gcagcgcgag cgcggcgac ggcttctgga ccaggcggcc gtcgtggaac 1260  
 accaccggct tcgggagctc ctcccggcg acgggcctcg gcgtcggcga caccggcacc 1320  
 acgtaccct ccttgcatag cctcatgaac gggtagtccg tcctccggtc gccgatgccg 1380  
 atctccggcg aggcgtcgcc gaacgcctcc cggagcgcgg ccgccttgtt ctgcgcgacg 1440  
 agcacgccgg ggctccggac gagccccgtg gcgcggccac gccagaccac gagctccgtg 1500  
 cccacgacga tgcggcgcc gaggtactcc ttcaagaacg cctccaccat catcctcggg 1560  
 ttcccggtta acaaggcacc ggggccggc aagccgaaaa aaaggccaa gaatccgggt 1620  
 ggatttcgga gcattagaac ttgggcagca cggcgggggg caccgcctta aagtccggga 1680  
 cccttgggcc ggccatggac gcgaagatga gcacctggat gccggccgaa tcgaaacca 1740  
 ggtagtagag caagccggca aacggcgcca gcaggatcag gccagcagc cggagagcgc 1800  
 ccccgtttt gaacgccatg tgcccgaagt acgggaacga gctccggccg cacagcagcg 1860  
 tgccgtcgag gtcggggacg atcgtgtcgc cgcgcgggtc cttggacgaa aactttcca 1920  
 ccgccgggaa cggcgacgcg atcatcgggg ggtccccat 1960

<210> 99  
 <211> 2311  
 <212> ADN  
 <213> *Selaginella moellendorffii*

5

<400> 99

cttctgagcc agcagtacag gggctggacg gacttgaaag gcatttttct ccagttccat 60  
 ttcaagtgag cacgttgga gcatattatc tcctctgtgg cgttctttct aactcggagt 120  
 tgtcccagga ttcgacatgg ccgacttaa gtactcgcgg ctcgtccagt tcggcctcgt 180  
 ggagaattgc tcatctgaag gacgtggaat gcacacggtt gcggcggatt tggatggtac 240  
 actgctcaga cgcagaagct cgtttocata ctctctgctg gtggctcttg aaggaggtag 300  
 cctcttcaga gcagttgtc tcggactttt ggcccctatc gcctggctgc tttaccactt 360  
 cgtctcggag gcagcagcaa ttcaacttct catctatgct tctttttcgg ggcagaaggt 420  
 gtccgagatt gagaaggtgg caacggcgggt tctgccgaag ttttactccc aggacgtcca 480  
 cagtgatgct ttcagagttt tctcgtcctg cgggaagaag tatgtcatca ctgccaatcc 540  
 toggatcatg gtggagtatt tctgcaaggc ctctcttggc gcggataaag ttattgggac 600

10

ES 2 640 100 T3

tgaaatcgaa gtggacgatg atggacgtgc tacaggcttc gtcaagttcc caggtgtgct 660  
 cgtgagtcaa aacaagagga cgccttgaa gctcgagttc gtggacgggg agcttcccga 720  
 gatcggactc ggcgatcgag aaaccgactt cgccttcatg tcgctttgca aggtaagcgc 780  
 actctttcca tatcatatac atattttgca tttcgagcat ataataacta gctctgcctc 840  
 tttcgtgctg ctttaactgt tgaggggtgc caggatttgc gtgacactcc agaatcagtt 900  
 acgagaagcg agttttctaa ctctggttgt aacogtgtga tacaggaagc ttacgttgtt 960  
 cccaagagga aggttgacgc agtggacggc agcgotctgg cgcagcgggt gatcttccat 1020  
 gatggttaggc tgggtccagct cccaacgccg ttogtggcgc tcgtcacttt cctctggctc 1080  
 cctatoggtc tcgcgctcgc agtgctccgg atogggcctca cctccogttg cccocgaaag 1140  
 tacatgggaa tgatctacgc catcctcggc gtccgcccga tcaccaaggg caaagttcct 1200  
 cgcaaacaag acggcagggg gtcctcttc gtgtgcaacc acaggactct ggtggtatccc 1260  
 atcttctcgg gattggccgc cgagagggac gtgaccgggc tcacttacag catcagccgg 1320  
 gtgtccgagg tgctatgcc catcgtcacg gtccggctga cgagagacaa ggagcgcgac 1380  
 tttacgaaga tgaagtgcct gctggagagg gggagggacc tgtgcctgtg tccggagggg 1440  
 acgacttgc gtgagccgtt cttgctgcga ttcagcgtc tcttcaccca gctcaagcac 1500  
 aggatcgtcc cggcggccgt gaagacgagg cagagcatgt ttaacgggac ggcggtgagg 1560  
 gggtggaagg ggatggatcc attcttcttc ttcattgaatc cattcccgac gtacgaggtg 1620  
 gagtttttg aggagctccc cggggagatg agcgtccaaa agggcggcaa gagctcgttc 1680  
 gagacggcca atttcatcca gaagatgctg ggcgacaagc tgggatacga gtgcacgacc 1740  
 ttcacccgca aggacaagta cctgatgctc aatggcagcg atggaacagt cccacgaaa 1800  
 taggaagggg ttggaaaaga aaacaagatg gaacogttgg ctggtgttga cttgaattga 1860  
 ctagagttga cggttgaata cttttttttt agtgttggct ggctatctag cttgcaactag 1920  
 taaatgctaa aggttatagg caggcgattt taatccttcc acatggctgg gacaggaaat 1980  
 attgacttgc ttcaattaaa gggggaaagg ttttagggca ggggtgttct agggcacggc 2040  
 agcgagctgg catcctcaca ccggcagcat ggcgtccggc cggcggtagc gcgctcgtgt 2100  
 ctgcattgc cagagaaatg ccaactccaa cctccgccac cttttctctc tcggccacct 2160  
 ttgcgaacta aacaaatgtg ttagaggaat caaagatcca acacaccgaa acttcgtttt 2220  
 tgttttcttc ttgcctggtc tttgtggagg cgaagaaaaa aaggaaacga gagagtcttt 2280  
 gtacctttgg gaatcaaact gctattgcgt a 2311

<210> 100  
 <211> 1989  
 <212> ADN  
 <213> *Selaginella moellendorffii*

5

<400> 100



ES 2 640 100 T3

atgtcgtctg acggcggcaa ggtcacggat tctttcctgg tagatgctcc attcgagccg 60  
 gtggagcgat gctcgtcggc ggggcgtgag aagcagacgt gcgtggcggg tttcgacggg 120  
 actctcatcc gcggccggag ctcttcccg tatttcatgc tggggcgct agagggagga 180  
 ggctcatcc ggagcgtgat tcttggectc ttcgcgcgcg tcgcctggat cctctacat 240  
 ttcgtctcgg aatccgcccg catccgectc ctcatcttca tctccttcgc cgggatgaag 300  
 gccaaagcaga tcgagtcggt gtcgcgcgcc gtgctgccca aattctactc ggaggacgtc 360  
 cacagcgagg octaccgagt gttctcctcg tggggcaagc gtgtggtggt gaccgccaat 420  
 ccccgatca tgggtggagca cttcgccaag acgtgcctgg gagcggacaa ggtgctgggg 480  
 accgagattg aggtgacgag gagtggctac gcgacgggat tcctcaagcg gccgggagtt 540  
 ctcgtggggg tcaacaagcg caaggctgtg aggcaggaat tcggcgagca tttgccccat 600  
 gtcggcattg gcgatcgagc cacggacttt gctttcatgg cgctctgcaa ggtacttacc 660  
 tcgaatccaa tcccgctctg cttgggtgggt ggagtggttg cttagttctc gatcgagctc 720  
 gatgcattag atcaacctc tgggcattgt tttcgatcga tgatccatcc ctgcaattgc 780  
 tgcatgagtg gatgattctc ttgtttcttt tgtttatctt ttttccttgt ctcatgcaca 840  
 cattcttaca agcgcactag octaaaagaa tctcttccaa gcagcagtag aaacttcttt 900  
 ccagtcaact aacgaacca tttgctgacc actatcagga agcatacgtc gtcccacct 960  
 ccaaagtgga ggcgctgagc cgagacaagc tcgtcaagcc cgttgtcttc cacgacggcc 1020  
 gcctcgtcca gcggcccacg cccttgaccg ccctcgtcac cttcctgtgg ctgcccgtca 1080  
 gcttcgtgct ggcggtgttc cggatcatgg tcacggtagc ctgcccgcgc gactacgtga 1140  
 ccatcgtota caagctcgtc ggcacccgac tcacgtcaa ggggcccgatc ccgcccacca 1200  
 agaagcgggg cgagtcgggc gtgctcttcg tctccagcca caggacactg tgcgatccgg 1260  
 tgttcgtggg ggtcggggcc cggcgggagg tgacggcgct cacctacagc atcagcaggg 1320  
 tgteggagtt cctgtccccg atcaagacgg tgggcctgtc cagggaccgc gagagggagc 1380  
 cggccaagat caaggcgtg ctccagaag gggacctgt catctgtccc gagggaaacca 1440  
 cctgcgcgga gccattcctt ttgcatcga gcgcaactgt cggggagctg actgacaaga 1500  
 tcgtgcccgt ggcgctgtgc acgaagggga gcactttcca cgggacaacg gtgcgcgggt 1560  
 ggaaggggtt ggatccattc ttcttcttca tgaatccct cccgacgtac gaggtgacgt 1620  
 ttctccagca gctgcgcggc gagctgacgg tccagcaggg cggcaagagc gcgatcgagg 1680  
 tggcgaatca catccagcgg gtgatcgcgg acacgctggg attcgagtgc acgaacttca 1740  
 cccgcaagga caagtacggg atgctggcgg ggaacgacgg caccgttccg gagaggcgga 1800  
 ggagcggcgg cggaggtggc ggcgcgcgca aggatctccc gaaggacgag caagcggccc 1860  
 aggggcaaga ccaggtaaat ggctgtagct aaaattgatt tgcgaggatc gtggatccat 1920  
 gccatgggtc ggtctcacc cgctccttcg ttataagtga atgtaagcta agctactggc 1980  
 ttgctctct 1989

ES 2 640 100 T3

<210> 101  
 <211> 1554  
 <212> ADN  
 <213> *Selaginella moellendorffii*

5

<400> 101

```

atggcagagt ttccacccat cgaaaagtgt ggcaccacaa tcagccgcga gagcgaatcg      60
gtggcctccg atctggatgg aacgctgctg cgcgggagga gttccttccc atactttttc      120
ctcaccactt tcgagacgtc ggggatcttg cgctcctgyc tgctcctcct cgtctcgcct      180
ctcgtctggc ttctctatca ctctcatctcc gaggccgctg gaatcaagct cctcatcttc      240
gcttccctgg cgggtgttcc agtctctcag atogaatgca togcccgag tggttctccg      300
cggttttatc tccaggatct caaccccgag acgtggaagg tcttctcgtc ctttggaag      360
aagattggtg tcacagcgaa tcccaggatc ttggtggagt cgctctgcaa ggaattccta      420
ggtgctgacg aggtgattgg aaccgagctg gaggtggata accgaggacg agcaaccggt      480
ttcgtcaaga gcccoggagt tctcgttggg gagaacaaag ccaaggcctt gctccaggcc      540
tgtgatcccg agcagctccc agacgttggg cttggcgatc ggcagaccga tttcccgttc      600
atgaagcttt gcaaggtacg taacaaaaga tcgtccctcc tcccgcgac ttagcagttc      660
tcacgctggc ttctctcaat ccaggaagga tacgtcgtcc cccaagcaa caacgtggag      720
ccagttcaa agtcccagct tetgaagccc atcatctttc acgatggteg tctcgtccag      780
cttccgacgc cactcgtggc tctcgtcacc ctctctgga tcccgcgctg ctctcgttcta      840
gootgcctcc gaatcgccgc aggtcgtctg ctcccatgt caatcgtgta ccatgccttc      900
tggctgctgg gagtccgagt gaccgtgaaa ggaacgccac ctccgagagc caccaagaac      960
tcccgaggcg tcctcttcat ctgctgcac cggacgctgc tagatccaat ctctctgtcg     1020
tgtgcgctgg gccgaggcgt ctccgcctg acttactcgg tgtcccgaact gtccgagatc     1080
ctcagcccga tcaaacaggt ggctctctcc aggaaccgct acaaggacgc ggctcagatc     1140
cgggagctcc tgaaagaagg cgacctggta atctgccccg agggaacgac gtgtcgggag     1200
ccattcctgc tgcgtttcag cgcactgttc gcggagctgg ccgacgatat cgtcccgcctc     1260
gccatggaca accggatgag catgttccac ggcaccgctg ctcgtgcctg gaaggggatg     1320
gaccattct acttcttcat gaatcccagc cccagctacg aggtaacgct cctccagcag     1380
ctcccggaag agctcacttg caagggcggc aagagctccc acgaagtcgc aaaccatata     1440

cagcgagctc tggcgggagt gctagatttc gagtgtacca atctgacgag gaaggacaag     1500
tacaaggcgc tcgtcggcac ggacggatcg gtggagacca agaagtcaac gtga          1554
    
```

10

<210> 102  
 <211> 1554  
 <212> ADN  
 <213> *Selaginella moellendorffii*

15

ES 2 640 100 T3

<400> 102

tcacggtgac	ttcttgact	ccaccgatcc	gtccgtgccg	acgagcgcct	tgtacttgtc	60
cttcctcgtc	agattggtag	actcgaaatc	cagcaactccc	gccagagctc	gctggatatg	120
gtttgcgact	tcgtgggagc	tcttgccgcc	cttgcaagtg	agctcttccg	ggagctgctg	180
gaggaacggt	acctcgtagc	tggggctggg	attcatgaag	aagtagaatg	ggtccatccc	240
cttcocagca	cgagcagcgg	tgccgtggaa	catgotcatc	cggttgtcca	tggcgaocggg	300
gacgatatcg	tcggccagct	ccgogaacag	tgcgotgaaa	cgcagcagga	atggctcccg	360
gcacgctcgt	ccctcggggc	agattaccag	gtcgccttct	ttcaggagct	cccggatctg	420
agccgcgtcc	ttgtcgcggt	tcctggagag	agccaactgtt	ttgatcgggc	tgaggatctc	480
ggacagtcgg	gacaccgagt	aagtcacggc	ggagacgcct	cggcccagcg	cacacgacag	540
gaagattgga	tctagcagcg	tcoggtgcga	gcagatgaag	aggacgcctc	gggagttctt	600
ggtggctctc	ggaggtggcg	ttcctttcac	ggtcactcgg	actcccagca	gccagaagggc	660
atggtacacg	attgacatgg	gaagcagcga	gcctgcggcg	attcggagggc	aggctagaac	720
gaagccgacg	gggatccaga	ggaggggtgac	gagagccacg	agtggcgtcg	gaagctggac	780
gagacgacca	tcgtgaaaga	tgatgggctt	cagaagctgg	gactttggaa	ctggctccac	840
gttgttgctt	ggggggacga	cgtatccttc	ctggattgag	agaagccagc	gtgagaattg	900
ctaagtgcg	ggaaggagg	acgatctttt	gttacgtacc	ttgcaaagct	tcatgaacgg	960
gaaatcgtc	tgccgatcgc	caagtccaac	gtctgggagc	tgctcgggat	cacaggcctg	1020
gagcaagccc	ttggctttgt	tototccaac	tagaactccg	gggctottga	cgaaaccggg	1080
tgctcgtcct	oggctatcca	totocagctc	ggttccaatc	acctcgtcag	cacctaggaa	1140
ttccttgacg	agcgactcca	ccaagatcct	gggattcgtc	gtgacaacaa	tcttcttccc	1200
aaaggacgag	aagaccttcc	acgtctcggg	gttgagatcc	tggagataaa	accgcggaag	1260
aaactcctcg	gcgatgcatt	cgatctgaga	gactggaaca	cccgccaggg	aagcgaagat	1320
gaggagcttg	attccagcgg	cctcggagat	gaagtgatag	agaagccaga	cgagagggca	1380
gacgaggagg	agcagccag	agcgaagat	ccccgacgtc	tcgaaagtgg	tgaggaaaaa	1440
gtatgggaag	gaactcctcc	cgcgcagcag	cgttocatcc	agatcggagg	ccaccgatcc	1500
gctctcgcgg	ctgattgtgc	tgtacactt	ttcgatgggt	ggaaactctg	ccat	1554

5 <210> 103  
 <211> 1757  
 <212> ADN  
 <213> *Selaginella moellendorffii*

10 <400> 103

ES 2 640 100 T3

atggagcacc cgtatgacgt gatcgacaat tgctcggcgg acaaccgcga gaagcagagc 60  
 gtggtagccg acttgatgg ctctgcgtt cgcagccgga gctccttccc ttacttcatg 120  
 ctggtggctc tggaaggagg cagcctcctg agaagcgtgc tctcatgggt catggcccc 180  
 gtcgcgtggc tgctctacca ctctcgtctg gaggcagcag gcatcaaagt cctcatcttc 240  
 ttctcctgcg ccggccagag agctcgggac atcgaagcgg tggtctcgcgc agttctcccc 300  
 aagttctact ccgaggacgt ccacagctcc acctggagag tcttctcgtc gtgcggcaag 360  
 cgctacattc tcacggccag cccgagactg atggtggagc actttgccac gacatacctg 420  
 ggcgctgaca aggtgatagg gacggagctg caagtgacaa ggagcgggta cgctacgggg 480  
 ttctcaaga gccccggcgt tttggttggc gtcaacaaga ggcgggcatt gaagcgcgtg 540  
 tttggtgagg aggtgccaga cctggggatt ggcgatcgca gctctgatta tccgttcatg 600  
 gctttttgca aggttggtae ttctcgttct ctctgagctc tcttctttt cgcctgtagt 660  
 cagcaggagc atatgcatat attgcagtgc atgctatgct agtctgtgat ctttggagca 720  
 actcgtcggc tggctttctt tttctttgat tttggtgat ggttgttggg gtcaagtccg 780  
 cgttgtcttc gttgtctccg tcgtcttctg gctttatttc atccacggat gtggaacttc 840  
 cgagtctcac cagaactaat aaagtcgctt ctcttttttc cttttcggct ttgcaggaag 900  
 cttacattgt tcctagccag cctgccgacg ccgttccagc tcacaaactc gcgcgcaaga 960  
 ttatctttca cgacggcagg ctggtccaga ggcgcagccc gttcgtggcg ctggtcaact 1020  
 ttctgtggct acccgccggc ttctcctcgt ccttcttccg catcctgggt gcgcttccat 1080  
 gcccggtcga gtacgtcccg ctgctctaca aggtgatcgg cgtccggctc atcgtcaagg 1140  
 gcaaagtccc gccaccttc aagaacggga tcctcttctg ctgcacgcac cggacgctca 1200  
 tcgacccgat aatggcctcg gcggcctcgg ggagcgacgt gaccgcgctc acctacagca 1260  
 ttagccgggt gtcggaggtg ctctcgccca tcaagacggt gccgctgtcg cgggaccggg 1320  
 agagggacgc ctccaacatc aagcggctgc tgctgaagga gaacttggcg atctgtcccg 1380  
 agggaacgac gtgccgggag cctttcctgc tgccggtcag cgcctgttc gcggagctgt 1440  
 ccgagaacat cgtcccgtt gcgacgtcga ccaaaacgag catcttccac gggacgacgg 1500  
 tgccgggctg gaaggggatg gaccggttct tcttcttcat gaaccggtc ccggtgtacg 1560  
 aggtgacgtt cctcagcggc ctgccgcacg atctgagcgt ccaggggggc aagtcggcga 1620  
 tcgaggtggc caactacatc cagaggtgc tggcggacac gctgggattc gagtgcacga 1680  
 actttacggc caaggataag tatgggatgc tcctcggcag cgatggaacg gtgcctgcca 1740  
 ggaagaagtc cttttga 1757

<210> 104  
 <211> 2256  
 <212> ADN  
 <213> *Selaginella moellendorffii*

ES 2 640 100 T3

<400> 104

agcttctgag	ccagcagtac	aggggctgga	cggacttgaa	aggcattttt	ctccagttcc	60
atttcaagtg	agcacgttgg	aagcatttat	tctcctctgt	ggcgttcttt	ctaactcgga	120
gttgtcccag	gattogacat	ggccgactta	aagtactcgc	ggctcttccg	gttcggcctc	180
gtggagaatt	gctcatctga	aggacgtgga	atgcacacag	ttgcggcggga	tttggatggt	240
acactgctca	gaggcagaag	ctcgtttcca	tacttcctgc	tggtggctct	tgaaggaggt	300
agcctcttca	gagcagttgt	tctcggactt	ttggccccta	tcgcctggct	gctttaccac	360
ttcgtctcgg	aggcagcagc	aattcaactt	ctcatctatg	tctctttttc	ggggcagaag	420
gtgtccgaga	ttgagaaggt	ggcaacggcg	gttctgcca	agttttactc	ccaggacgtc	480
cacagtgatg	ctttcagagt	tttctcgtcc	tgcggaaga	agtatgtcat	cactgccaat	540
cctoggatca	tggtggagta	tttctgcaag	gccttccttg	gcgcggacaa	agttattggg	600
acggaaatcg	aagtggacga	tgatggacgt	gctacaggct	tcgtcaagtt	cccaggtgtg	660
ctcgtgagtc	aaaacaagag	gaccgccttg	aagctcaggt	tcgtggacgg	ggagcttccc	720
gagatcggac	tcggcgatcg	agaaaccgac	ttcgccttca	tgtcgctttg	caaggtaagc	780
gcactctttc	catatcatat	acatattttg	catttcgagc	atataatact	tagctctgcc	840
tctttcgtgc	tgctttaact	gttgaggggt	gccaggattt	gcgtgacact	ccagaatcag	900
ttacogagaag	cgagttttct	aactctgggt	gtaaccgtgt	gatacaggaa	gcttacgttg	960
ttcccaagag	gaaggttgac	gcagtggacg	gcagcgtct	ggcgcagcgc	gtgatcttcc	1020
atgatggtag	gctggtccag	ctcccaacgc	cgttcgtggc	gctcgtcact	ttcctctggc	1080
tcctatcgg	cttcgcgctc	gcagtgctcc	ggatcggcct	cacctcccgt	tgcccccgaa	1140
agtacatggg	aatgatctac	gccatcctcg	gcgtccgcog	catcaccaa	ggcaaagttc	1200
ctcgcaaaca	agacggcagg	gggtcctct	tcgtgtgcaa	ccacaggact	ctggtggatc	1260
ccatcttctg	gggattggcc	gocgagaggg	acgtgaccgg	gctcacttac	agcatcagcc	1320
gggtgtccga	ggtgctatcg	cccatcgtca	cggtccggct	gacgagagac	aagcagcgcg	1380
actttacgaa	gatgaagtcg	ctgctggaga	gggggagggg	cctgtgcctg	tgtccggagg	1440
ggacgacttg	ccgggagcct	ttcttgctgc	gattcagcgc	tctcttcacc	gagctcacgc	1500
acaggatcgt	cccggcggcc	gtgaagacga	ggcagagcat	gtttaacggg	acggcgggtgc	1560

ES 2 640 100 T3

gggggtggaa ggggatggat ccattcttct tcttcatgaa tccattcccg acgtacgagg 1620  
 tggagttttt ggaggagctc cccgtggaga tgagcgtcca aaagggcggc aagagctcgt 1680  
 tcgagacggc caatttcate cagaagatgc tgggcgacaa gctgggatac gagtgcacga 1740  
 ccttcacccg caaggacaag tacctgatgc tcaatggcag cgatggaaca gtccccacga 1800  
 aataggaagg ggttgaaaaa gaaaacaaga tggaaccggt gactggtggt gacttgaatt 1860  
 gactagagt gacggttgaa tttttttttt ttggtggtg ctggctatct agcttgcaact 1920  
 agtaaagct aaagttttt caagcaggca ggcgattta atccttccac atgggtggga 1980  
 cagaaaatat tgacttcgtt catttaaagg gggcaaagg tttagggcac ggaggggtg 2040  
 tctagggcac ggaagcgagc tggcatcctc acacaggcag catggcgtcc gcccgcggt 2100  
 acggcgctcg tgtctcgcat tgcacgcaa atgccactcc caacctcggc caocttttct 2160  
 ctctcctgga gttctctgc caccttccg aactgaacaa atgtgttaga ggaatcaaag 2220  
 atccaacaca ccgaaacttc gtttttgttt tcttct 2256

<210> 105  
 <211> 2044  
 <212> ADN  
 <213> *Selaginella moellendorffii*  
 <400> 105

5

atgtcgtcgc acggcggcaa ggtcacggat tctttcctgg tagatgctcc attcgagccg 60  
 gtggagcgat gctcgtcggc ggggcgtgag aagcagacgt gcgtggcgga tttcgacggt 120  
 actctcatcc gcggcggag ctcttcccg tatttcatgc tgggtggcgt agagggagga 180  
 ggcctcatcc ggagcgtgat tcttggcctc ttgcgcggc togcctggat cctctacat 240  
 ttcgtctcgg aatccgcccg catccgcctc ctcatcttca totccttgc ogggatgaag 300  
 gocaagcaga tcgagtcggt gtcgcgcgcc gtgctgocca aattctactc ggaggacgtc 360  
 cacagcgagg cctaccgagt gttctcctcg tgcggcaagc gtgtggtggt gaccgccaat 420  
 ccccgatca tggtgagca cttcgccaag acgtgcctgg gggcggacaa ggtgctgggg 480  
 accgagattg aggtgacgag gagtggctac ggcacgggat tcctcaagcg gccgggagtt 540  
 ctcgtggggg tcaacaagcg caaggctgtg aggcaggaat toggcgagca tttgccccat 600  
 gtcggcattg gcgatcgagc cacggacttt gctttcatgg cgctctgcaa ggtacttatc 660  
 togaatccaa tcccgtctcg cttggtgggt ggagtgggtg cttagttctc gatcgagctc 720  
 gatgcattag atcaaacctc tgggcattgt tttcgatoga tgatccatcc ctgcaattgc 780  
 tgcagtagtg gatgattctc ttgtttcttt tgtttatctt ttttcttgt ctcatgcaca 840  
 cattcttaca agcgcactag cactagccta aaagaatctc ttccaagcag cagtagaac 900  
 ttctttccag tcaactaacg aaccatttg ctgaccacta tcaggaagca taogtcgtcc 960

10

ES 2 640 100 T3

caccctccaa agtggaggcg ctgagccgag acaagctcgt caagcccgtc gtcttccaag 1020  
 acggccgcct cgtccagcgg cccacgcctt tgaccgccct cgtcaccttc ctgtggetgc 1080  
 cegtcagctt cgtgctggcg gtcttccgga tcatggtcac ggtaccctgc ccgcgcgagt 1140  
 acgtgaccat cgtctacaag ctcgctcgca tccggctcat cgtcaagggg ccgatcccgc 1200  
 cgcccaagaa gcggggcgag tccggcgtgc tcttcgtctc cagccacagg acaactgtgcg 1260  
 atccggtggt cgtgggggtc ggggcccggc gggagggtac ggcgctcacc tacagcatca 1320  
 gcaggggtgc ggagttcctg tccccgatta agacgggtgg cctgtccagg gaccgcgaga 1380  
 gggacgcggc caagatcaag gcgctgctcc agaaggggga cctgtgcata tgtcccgagg 1440  
 gaaccacctg ccgcgagcca ttccttttgc gattcagcgc actgttcgcg gagctgactg 1500  
 acaagatcgt gccggtggcg ctgtgcacga aggggagcac tttccacggg acaacggtgc 1560  
 gcgggtggaa ggggttgat ccattcttct tcttcatgaa tccttcccg acgtaogagg 1620  
 tgacgttctt ccagcagctc ccgcgggagt tgacggcca gcagggcggc aagagcgcga 1680  
 tccaggtggc gaatcacatc cagcgggtga tccgcggcac gctgggattc gagtgcacga 1740  
 acttcacccg caaggacaag tacgggatgc tggcggggaa cgacggcacc gtgccggaga 1800  
 ggcggaggag cggcggcgga ggtggcggcg ccgccaaagga tctcccggag gacgagcaag 1860  
 cggcccaggg gcaagaccag gtaaattggct gtagctaaga ttgatttgcg aggatcgtgg 1920  
 atccatgcca tgggtcgggt ctcacgcgct ccttcgttat aagtgaatgt aagctaagct 1980  
 actggcttgc tctctcgatt gactcatcag gaatagatcg atcgggtgga gagcatcgtg 2040  
 atcg 2044

<210> 106  
 <211> 1715  
 <212> ADN  
 <213> *Oryza sativa*  
 <400> 106

5

ES 2 640 100 T3

actcccacac catcagcgag cacctagcgg togcacgct agctgccatc tcttgggtgag 60  
atccatcgag acgcccggca cgacgacgat ggtgtcgcgg aggttcaagc cagtggagga 120  
gtgcagctcg gatgggaggt cggagcagac ggtggccggc gacttcgacg gcacgctggt 180  
caggtcgggg agcgccttcc cgtactacct cctcgtcggc ctcgaggccg gcagcgtgct 240  
ccgcgccgctc gtgctgctgc tgtccgtgcc gttcgtctac gtgacctaca ttttcttctc 300  
cgagtcgctg gccatcagca cgctgggtga catctccgtg gcggggctca aggtgaggaa 360  
catcgagatg gtggcgcggt cgggtgctgcc caagttctac gcggaggacg tgcacccgga 420  
gagctggagg gtgttcaact ccttcggcaa gcgggtacatc atcacggcga gcccaggat 480  
catggtggag cacttcgcca agacgttccg cggcgccgac aaggtggtcg ggacggagct 540  
  
ggaggtcggc aagaacggca aggccacggg gttcatggtg aagcccggag tgctcgtggg 600  
cgaccacaag aggcaggccg tcgtcaagga gctacgcgac gcggtgcccg acgtcggctt 660  
gggcgacagg gagacggatt tcgacttcat gtccatatgc aaggaggcct acctcgtgac 720  
atcaaggaag tacagcgcgg tgcccaagaa ccagctgctg agcccactca tctccacga 780  
cggccgcctc gtgcagcgtc cgacgccgct ggtggcgctc gtcaccttcc tgtggatgcc 840  
gttcggcttc gcgctcgcgc tctcgcgctg gtacgtcaac ctgccactcc cggagcggat 900  
cgtcttctac acctacaagc tcatgggcat ccgcctcatc gtgaaggga acccgcgcc 960  
tcccccaag aaggacatc ctggcgtcct cttcgtctgc aaccaccgca cagtgtctga 1020  
cccggttgag gtcgccgtgg cgctgcgccg caaggtcagc tgcgtcacgt acagcatctc 1080  
caagttctcg gagctcatct cgcgatcaa ggccgtcggc ctgtcgcggg agcgcgagaa 1140  
ggacgccgag aacatccggc ggctgctgga ggagggcgac ctggtgatct gcccggagg 1200  
caccacctgc cgcgagccgt tctgctgctg gttcagcggc ctcttcgccg agctcaccga 1260  
ccgcatcgtg ccggtggcga tcaacaccaa ggagagcatg ttccacggct ccaccgtgctg 1320  
gggcttcaag ctcatggacc cctacttctt cttcatgaac ccgcggccga cgtacgagat 1380  
caccttctctg aaccagctgc ocaaggagct cacttgacgc ggcggcaagt cgcctatcga 1440  
ggtggccaac tacatccaga agacgctcag cggccagctc ggcttcgagt gcaccgccat 1500  
aacgcgcaag gagaagtaca gcatactcgc cgggaccgac ggccgcgtcc cttccaagaa 1560  
caaggagaag gaaaagaact agcgagtcac agcctcctat cgatcggagt actcgtatt 1620  
cgcccagccg actagtgtca ctcatggctt ctatctattg tttacgattt attgttgttt 1680  
ttcaagaagt tggtaaatta catattgttt ccgag 1715

<210> 107  
<211> 1521  
<212> ADN  
<213> *Oryza sativa*

5

<400> 107



ES 2 640 100 T3

atggcggcgg cggcggcggc ggggaggagg aggttctttc cgccgggtgac ggcgtacgac 60  
gcgcgggcgg gcgcgaggcg gacggtggcg gcggacctcg acgggacgct gctgggtgctg 120  
tcgtcggcgt tcccgtaacta cttccttgtg gctctcgagg ccggcagcta cctccgcgcg 180  
ctggcgctgc tgctcgccgc gcogtggctc ctcgcgctct acgtcggcgt ctccgaggcc 240  
gccgccatcg ccctgctcgt gttcatcacg ttcgccgggc tgcgcgtccg cgacgtcgag 300  
gccgtggcgc gcgcogtget gccgcgccac tacgcccgcg gggtgccgcg cgacacgtgg 360  
gcggtgttcc acggctgcgc cgagcgcgcg gtcgtggtca ccgctcgcc gcccgcatg 420  
gtgggggagt tcgtgcgcga gttcctcggc gccgaggtgg ccggcaccga gctcgagacg 480  
ttcgcctcgg gcaagcgggt caccggccgc atcaaggcgg tcctcgtcgg cgagaagaag 540  
cgcgaggctg tcgagaggct cttcgccggc ggcgacatgc ccgacgtcgg cctcggcgac 600  
cgcgagagcg accacgactt catggccatc tgcaaggaag cctacatggt gcccaagaac 660  
aagcgcgcgc cgcgcgcgcg ccgcgacgag ctgctgtcgc gcgccatctt ccacgaecgc 720  
cgctcgtcc gccgcccggg gccagcgagc gcgctgttcg cgctggccta cctcccogtg 780  
ggcttcgccc tggcgtcct ccgctgttc ctcaacctcc ccgtcccggc gcgcctcgtc 840  
cgccacacct accgcctcac cggcatccgc ctcgccgtgc gcggggcgcc gccgcgcgcg 900  
ccgcgcccgg ggacgcccgg ctccctcctg gtgtgcaacc accgcaccgc gctggacccc 960  
atcatcgtgt ccatcgcgct gggccgccc gtgacgtgcg tgacgtacag cgttagccgg 1020  
ctgtcgacgg cgatctcgcc gatccgcgcg gcggcgctga cgcgggaccg cgcggcggac 1080  
gcggcgcgca tcgcggcgct gctggaggag ggggacgtgg tgggtgtccc ggaggggacg 1140  
acgtgccggg agccgtacct gctgcggttc tcggcgctgt tcgcgagct gacggcgagg 1200  
atcgtgccgg tggcgggtgga ggcgaggcag gggacgtact atgggtcgac ggcgaggggg 1260  
tggaagtcc tggaacctta cttctctac atgaaccgc gcccggggta cgaggtgacg 1320  
ttcctgccgg cgctgaggcc ggaggagacg tgcgtcgccg gcgggaggag cgcggtggag 1380  
gtggcgaacc acgtgcagag ggtgatcgcc aaggagctag ggttccaatg caccacgctc 1440  
accaggaagg acaagtacat gaagctcgcc ggcaacgacg gccgggtcgc cgccgcccgc 1500  
gacaagccca aggccaattg a 1521

<210> 108  
<211> 1755  
<212> ADN  
<213> *Zea mays*

5

<400> 108

ES 2 640 100 T3

aaccoagaca ggcaccgtac caccgtgttc cccgtcgcca tggccgcgag gttccccccg 60  
gtgtogtcct acgacgcttc ggcgcgggtc cgcgcacggy ccgcggcgga cctggacggc 120  
acgctgctgg cgtcgctctc ggcgttcccc tactacttcc tggcggcgct ggagccggc 180  
tcctgcgcgc gcgcgctgct gctcctcctg gccgcgccgc tgctgctggc gctctacacg 240  
ctggctctccg aggcggcgcc catcgcgctg ctggcgctcg tcacgttcgc ggggctccgg 300  
gtgcgcgacg tcgaggccgt cgcgccgggc gtcctcccgc gccactacgc cgcaggcgtg 360  
cgcgcgcaca cctgggccgt gttccgcgac tgcggcgccg ccgcgagggt cgtcgtcacc 420  
gcgtcccccg ccgtcatggt cgcgcacttc gtcgcgaggt tcctcgccgc cgagctcgcc 480  
gccaccgagc tcgagacctg ccgcgcgctc ggcgacgcct ggttcacggg caggatcaag 540  
gccgtgctcg tcggggagag gaaggccgag gtcgtgccc ggctcttcgc cgcggcgac 600  
ctgcccagc tcgggctcgg ggaaccgcgag agcgaccacg acttcatggc catctgcaag 660  
gaggcttaca tggtgcctcc ggaccggcgc gcgcgcgag ccgcccgcga ccgcctgctg 720  
tcccgcgcgc tcttcacga cggccgcctc gtcggcgccc cggaccggc gcacgcgctc 780  
ttcgcgctgg cctacctccc tctcggett cgtctggcgc tgttccgcgt gctcttcaac 840  
ctcgccatgc cggcgcgect ggtccgccac acgtaccgcc tgaccgggat ccggtccgc 900  
gtgcgcggga ctccgccgcc ggcgccggcg ccgggcgcgc cggggtcgt gctggtgtgc 960  
aaccaccgca cggcgcctga ccccatcctc gtgtcggtgg ccgtggggcg gcccggtggc 1020  
tgctgacct acagcgtgag ccggctgtcg accggcgtgt ccgcggctcc cgcggtggcg 1080  
ctgtcgcggg accgggcggc cgcgcgcggc cgtatagcgg ccgtgctggc ggagggcgac 1140  
gtggtggtgt gcccgcaggg gaccacgtgc cgggagccct gcctgctgcg cttctcggcg 1200  
ctgttcgcgg agctcacgga ccggatcgtg cccgtggcgc tggacgcgcg gcaggggacc 1260  
tactacgggt ccacggcgag ggggtggaag tggctcgacc ctacttctt ctacatgaac 1320  
ccgcgccccg ggtacgacgt caccttcctg ccgcgcgctc ggcgcgagga gacgtgcggc 1380  
gccggcggga ggagcgccgt cgacgtggcc aaccacgtgc agacgctcat ccgaaggag 1440  
ctcgggttcc ggtgcaccaa gctcaccag aaggacaagt acatgaagct ccgggcaac 1500  
gacggcacgg tcggggcgag gcctgagaat gagaaggcgc ccgcgatga tagtctgcc 1560  
acggccacca ccaagaagct tgtataaaga gactgtattt atagctagct ggtcgactgt 1620  
atcggagcag catccattat gtatttatgt acgtatgcta atgacgaca cattagtgtg 1680  
atgtgatgga ttcgtgaaca tagaagatta caagcttaac ttcatgttta ccatgaccaa 1740  
aaaaaaaa aaaaa 1755

5 <210> 109  
<211> 1852  
<212> ADN  
<213> *Zea mays*

10 <400> 109

ES 2 640 100 T3

ggcggcgtat attaactcca cctcgccgcc ccactccact gccagcaac tcgcacgcag 60  
 cagtacgacg tctctctctt gggctctgggg ccgagacacc gagcacgtac taccagcaag 120  
 atggtggcgt ctcccagatt caagcccatc gaggagtgtt gctcggaggg gcggctggag 180  
 cagacgggtg ccgcccacct ggacggcacg ctgctcatct ccaggagcgc gttcccctac 240  
 tacctcctcg tggctctcga ggcgggcagc gtctcctcgg ccgctgtgtt gctcctgtcc 300  
 gtgcccgttc tctacgtcac ctaagccttc ttctccgagt cgtggccat cagcacgtg 360  
 gtgtacatct ccgtggcggg gctcaagggt cgcacatct agatgggtgg gcgctccgtg 420  
 ctccccaggt tctacgcggg cgacgtgcac ccggagagct ggagggtgtt cagctccttc 480  
  
 ggcaggcgtc acgtcgtcac ggccagcccc agggatcatg tcgagccctt cggcagggcc 540  
 ttctcggcgg ccgacaaggt cgtcgggacc gagctggagg tcggcaggga cggcaaggcc 600  
 acgggcttca tggccaggcc aggcgtgttc gtcggcgacc acaagaaga gccgctcgtc 660  
 aaggagctcg gcgacgcgct gcccgacgtc ggcatggggg ataggagac cgacttcgac 720  
 ttcattgtca tctgcaagga ggcctacctg gtgacgtcaa ggaagtacag ccggtgccc 780  
 aggaaccagc tgtgagccc gctgatcgtg cacgacggcc gcctcgtgca gcgcccagc 840  
 ccgctcgtcg cgtcgtcac cttcctctgg atgccgttcg gcttcgctgt ggcgtcatg 900  
 cgcgtgtaca tcaacctgcc gctgccgag cgcatcgtct actaaccta caagctcatg 960  
 ggcacagggc tcgtcgtcaa gggcaccgcc ccgcccgcgc ccaagaaggg ccaccggggc 1020  
 gtcctcttcg tctgcaacca ccgaccgtg ctgaccccg togagggtgg cgtggcgtg 1080  
 cgcggcaagg tcagctcgt cacctacagc atctccaagt tctccgagct catctcggcc 1140  
 atcaaggccg tcgctcgtc gcgggagcgc gacaaggacg ccgagaacat ccgcccctg 1200  
 ctggaggagg gcgacctggt catctgcccc gagggcacca cctgcccgga gcccttctg 1260  
 ctgctgttca gcgctctctt cgcgagctc accgaccgca tcgtgcccgt ggccatcaac 1320  
 accaaggaga ccatgttcca cggtccacc gtgcgagggt tcaagctcat ggatccctac 1380  
 ttcttttca tgaaccgcg ccgacgtac gagatcacgt tctcaccga gctcccaaaa 1440  
 gacctcagct gcagggcgg caagtgcgc atcgagggtg ccaactatat ccagaagacg 1500  
 ctgagcggac agcttggtt cgagtgcacg tccatcacgc gcaaggagaa gtacggcatg 1560  
 ctgctggca ccgacggccc tgtcccgtcc aagaacaagg agaaggagaa ggataagaac 1620  
 taacggccaa cgtcgtctt cctgcttgc cagaagactg ccgtttaatt actgtcaaaa 1680  
 ttatttcaa ttattgttc cttgtcgaag agttggccag tgtgatgtgg gccgtgtact 1740  
 agtacaagct tgatgatcat tagtatgtct acaattcttc ctttcgggaa gaaggagga 1800  
 agagttgata gagccgtctt aatacaaaagt ttgcgcaaaa aaaaaaaaaa aa 1852

5 <210> 110  
 <211> 1704  
 <212> ADN  
 <213> *Zea mays*

ES 2 640 100 T3

<400> 110

```

caccaatata atcaccattt ctgttcaogc gatcgagagg caaaattaag catggtgtcc      60
cgaaggttca agccgatcga gctgtgcgac tccgaggggc ggtcgcgccca gacggtcgcc      120
gcogacctcg acggcacgct cctcctgtcc cgcagcgcgt tcccgtacta cctcctcate      180
gcgctggagg ccggtagcct ccttcgcgcg gtggcgcttc tcttgccgtg gccccctcgtg      240
tacctcacct acgtgaccgt gtccgagacc ctggccgtgc ggcggttcat ctacgtcgcc      300

atggccgggc tggaggogag ggacatcgag gccgtcgcgc ggtccgtgct cccagcttc      360
tacgccggcg acgtocaccc ggaggctgg cgcgtgttcc gctcgttcgg caggaggtgc      420
gtcgtcaccg ccagcccgcg cctcatggtg gagccgttcg ccaaggcgtt tcttggcgcg      480
gacgtggtga tcgggacgga ggtggaggtg gccagtcgc ggaaggccac ggggttcggt      540
gccgggcccg gggtgctcgt cggcgagcac aagaggcggg cggtggcgag ggagttcggc      600
gactcgtcgc cggacgtcgg catgggggat cgcgagagcg acttcgactt catgtccate      660
tgcaaggaag cgtacaccgt aacacgacag aagtaccgcg cgctgccccg ggatcagctg      720
cagggccggg ccatoctcca cgacggccgc ctggcccgcg gcccgacggc gaccaacgcg      780
cttctcactt tcctgtggat gccgctcggc ttcgcgctcg cgctggtgcg cgcgcatctg      840
cacctgctcc tcccggtgcg cgcctcgcg tacgctaca agcttacggg cgtcaagctg      900
gtggtgcgcg gcaacaggcc gccgccgtcc aagaagaaga agggcgacca gctaggggtt      960
ctcttcgtgt gcaaccaccg tagcacgctc gacgccgtcg cggtgggcgt cgcgctgggc      1020
cgcaaggtga gatgggtcgt cacagacggc ggtgcgggtg cctccaggtt ctoggagccc      1080
gtagtgtcgc ctatcatgac aggogtgccg ctgcccgtgc ccaccagcct cgagggcgac      1140
gccgacgcgg acgccgcccc ccgcatcggg cggctgctcg aggagggaga cgacgtcgtc      1200
gtcttccccg agggcgccat ctgccgcgag ccgttctcgc tgcggttcgg cgcgctcttc      1260
gctgaggtca ccgaccgcat cgtgcccgtg gccatcgaac ctagggagag catgtttcac      1320
gggtccacgg cgcgcgggct cagagggatg gatecctact tctttttcat gaaccacgg      1380
ccgacgtacg aggtcacggt cctgaaccag cttccagggg agctcacctg cggcggcggg      1440
aggtcgccgg tcgaggtggc cagctacgtc caggaggtcc tggccgcgca gcttggattc      1500
gagtgcacct gatgaagcag caagtgcaaa agtggcatgg tgacggttgc gtggcgttca      1560
aaggtggact agggaacaca gagcatgcaa ccaactgtca attgccgtgt gttgagagtg      1620
gtgatcagtt catggtttca cgtttgctgt tgatgtctgt gataagaacc atatgtatct      1680
ctacaactat taagagcgtt gtgt      1704

```

5

<210> 111  
 <211> 1845  
 <212> ADN  
 <213> *Arabidopsis thaliana*

<400> 111

atattccattc	tcaccttcca	aatctcccca	ctctctctct	ctctctctct	ctctctctct	60
ctctctctct	ctctatctct	ctctatctct	ctctatcttc	tcttttctcc	tctctccaac	120
ttcttcaacta	atgggagctc	aggagaaacg	gcgcggttc	gagcagatat	caaagtgcga	180
tgtaaggac	cggccaacc	ataccgtggc	cgctgatcta	gacggaacac	tactaatctc	240
tcgtagcgcc	ttccttact	atttctcgt	agccctcgag	gcagggagct	tgctccgagc	300
gttgatccta	cttgtgtccg	taccattcgt	ttatcttaag	tacttgacca	tctccgagac	360
tttagccatc	aacgtatttg	tcttcatcac	gttcgcgggt	ctcaagatcc	gagacgttga	420
gctagtggtc	cgttccgtcc	tcccagaggt	ctatgcggag	gacgtgaggc	ccgataacctg	480
gcgtatcttc	aacacgttcg	ggaaacggta	cataataact	gcgagccctc	gaattatggt	540
cgagccattc	gtgaaaacat	tcctaggagt	tgataaagtt	cttggaacag	agctagaggt	600
ctccaaatcg	ggtcgggcaa	cgggttcac	cagaaaacca	ggtattctcg	tcggtcagta	660
caaacgtgac	gtcgttttga	gagagtttgg	tgccctagcg	tctgatttac	ctgatttggg	720
gctcggcgat	agcaagacgg	accacgactt	catgtccatc	tgcaaggaag	gttacatggt	780
gccacgtacg	aatgcgaac	cattaccaag	aaacaaactc	ttaagcccca	taatattcca	840
cgagggcaga	ttagtccaac	gcccaacgcc	gtagttgct	ctgttaactt	tcctctggct	900
tcccgtcggc	ttcgtctctc	ctatcatcog	cgtctacacg	aatattccgt	taccggaacg	960
tatcgccogt	tacaactaca	agcttactgg	catcaagcta	gtcgtcaacg	gccaccctcc	1020
tccgcgcgca	aaacctggcc	agccaggcca	tcttttggtc	tgcaaccacc	gcaccgttct	1080
cgatcctgtg	gtcacagctg	tcgcactcgg	ccggaaaatc	agctgcgtca	cttacagcat	1140
cagcaagttc	tctgagctaa	tctcaccaat	caaagccgtt	gcgttgactc	gtcaacgtga	1200
gaaagacgca	gcgaacatca	agcgtctttt	ggaggaaggc	gatctcgtga	tatgtcccga	1260
gggaaccacg	tgccgtgagc	ctttccttct	ccggtttagt	gctcttttcg	ctgagctcac	1320
ggaccggatc	gttcccgtgg	cgatcaacac	aaagcagagc	atgttcaatg	gtaccaccac	1380
acgtggatac	aagcttcttg	atccttactt	tgcgttcatg	aaccggaggc	cgacgtatga	1440
gatcacgttc	ctcaaacaga	ttccagctga	gctgacgtgt	aaaggaggca	aatctccgat	1500
agaggttgcg	aattacatac	agagggtttt	gggaggaacc	ttaggttttg	agtgcaccaa	1560
tttcacaaga	aaggataagt	acgcaatgct	tgctggtact	gacggtaggg	ttccggtgaa	1620
gaaggagaag	acgtgattat	tgatcggagg	atggagaaaa	caaagaaata	atacaaagat	1680
tggaatcttt	ttcatttggg	tggttaatta	tcaatactta	aggggcaaag	aattgataag	1740
tttgtttttg	ttgttacctt	tatcttctct	gcttcggtga	atgtttcttc	aagttttaca	1800
tttagtacia	taacggttta	agaaaagaaa	gaaaaaagta	tcggt		1845

ES 2 640 100 T3

<210> 112  
<211> 1872  
<212> ADN  
<213> *Arabidopsis thaliana*

5

<400> 112

caatataggt gcatttaaaa cataacgttc cctctcttgt cttcttattt ttgctctctg

60

ES 2 640 100 T3

ttcttttagaa ctgccattaa cgagagtgag agagagagac acacacacag cgagtgaaaa 120  
 tgtctocgga gaagaagagt caaaacttcc ctccaataac ggaatgcaga gacggagagt 180  
 acgattcgat agccgcogat cttgacggga ctctgcttct ctcaagaagc tccctccctt 240  
 acttcatgct cgtcgtgtt gaagctgga gccttttacg tggactaatc ctccottctct 300  
 cgttaccatt cgtcattatc tcttacctat tcgtatocga atctcttggg atccagatcc 360  
 tcatcttcat ctcatcogct ggtctcaaaa tccgcgatat cgaacttgc tctcgcgcag 420  
 ttcttccacg gttttacgcg gcggatgtga ggaaagacag ttttgagggtg tttgataagt 480  
 gtaagagaaa agtggtagtg acggcgaatc cgattgtgat ggtggaggcg tttgtgaagg 540  
 attatcttgg aggtgataaa gttttgggaa cagagattga agttaaccct aaaaccaata 600  
 gagccactgg atttgtgaag aagcctgggtg ttcttgttgg tgatcttaag aggttagcca 660  
 ttttaaaaga gtttggtaac gaatcacctg atctcggcct cggatgatcga acctctgatc 720  
 atgatttcat gtctctctgc aagaaagggt acatggttca tgcgaccaag tcagccacaa 780  
 cgattccaaa agaacgctta aagaaccgca tagtcttcca tgatgggctg ttagcgcaac 840  
 gtccaactcc gttaaacgcc attatcacat acctatggct tcccttttggg ttcatectct 900  
 ccatcattcg cgtctaactc aacctccctt tacctgaaag atttgtccgt tacacttacg 960  
 agatgctcgg gatecactta accttcgtg gtcctctcc tccacctcct tcccctggaa 1020  
 ctcttggcaa cctctatgtc cttaacacc gtaccgogct tgatcccatc atcgttgcta 1080  
 ttgctcttgg acgtaagatc tgttgcgta cttacagtgt ctctcgtctc tccottatgc 1140  
 tttctcctat tctgtgtgtt gccctcacc gtgaccgtgc caccgatgct gccaacatga 1200  
 gaaaacttct cgagaaaggc gacttgggtga tatgtccgga aggcacgacg tgtagagaag 1260  
 agtatctact gagatttagc gctctattcg cagagctaag cgaccggatt gtgocagtag 1320  
 cgatgaactg taaacaagga atgttcaacg ggaccacagt taggggtgtg aagttttggg 1380  
 accttactt cttcttcatg aaccaagac caagctatga agccacttc ttggatcgtt 1440  
 tgctgaaga aatgactgtc aacggtggtg gcaagactcc tatagagggt gctaattacg 1500  
 tccagaaagt taccggcgcg gttttgggct tcgaatgcac cgaacttact cgcaaggata 1560  
 aatatctttt gcttgagggt aatgacggca aggtggagt tatcaacaac accaagaagt 1620  
 gaatatctta ttcttgttat aacttggcat tgacttattg caagtaagggt tggatcatat 1680  
 acgaattttc atgtgtgtaa ttgcttgagc tcttgcctac attatggtac gttattgtta 1740  
 tgtttattct attttgttct taatgtcgaa cgggtaacta tacccttaa aactattttg 1800  
 tgatcccggt tctatacgga tgtgtataga gactttctat ttaaagtgca aaattttatt 1860  
 ggaggaagta aa 1872

<210> 113  
 <211> 1545  
 <212> ADN  
 <213> *Physcomitrella patens*

ES 2 640 100 T3

<400> 113

atgcagcaac	tgcaccatt	gactagtttt	cactgcagtg	aaagcttcga	aggagtggag	60
acctgcagag	tggaagggcg	tgogaaccag	agcattgttt	cagatttggga	cggaacgctt	120
ctgcggtcaa	gaagctcttt	tcctactttt	atggtgctcg	ctttcgaagc	agggagccca	180
ctccggtctg	togttcttct	gcttatttct	cccgtagtat	ggcttgtcta	caacttogtc	240
tccgaggcag	ctgggattaa	gatgctcatt	ttogtctcgc	tggcgggact	caaagtgtcc	300
gccgtcgagt	cagtggcgcg	gggtgttctg	occaaattct	acttgggaaga	catgcactcc	360
atcagctacc	gagtcttcac	ctcctgcggg	aagcgctacg	tctgacggc	gaatccaagg	420
atcatagtgg	agccgtttct	gaaagagtat	cttgacgtag	aagctgtaat	gggcactgaa	480
ctacaaatct	cgagcggggg	atatgtcacc	ggttttgtaa	cgggccccgg	tgtactcgtt	540
gggtgctgtca	aagagcgagc	tgtgaagaaa	tactttggac	cagaccctcc	ggacgtcggc	600
ctaggagaca	ggcagagcga	cttcttattc	atggatctct	gcaaggaatc	ttacattgta	660
cacagcgaca	aggatgtgcc	cgccgtgtcg	aaggaagact	tcttaaaacc	tcttatcttc	720
cacgacggta	gactgggtgtg	ccgtcctacg	ccgctgatgt	cgcttatcgt	catgctatgg	780
tgtcctattg	gctttgtcct	tgccgttggt	cggatgctca	tccgcattgc	gttgccaatg	840
tgggtgggcgc	ttcctcttga	ggcaatgctc	ggcgtcaccg	tccgtgtgaa	aggaattccc	900
ccgacgtgcc	ccctccgaaa	caagaagcgg	gggtgtgctct	togtgtgctc	tcatcgcact	960
cttcttgacc	ctatcttctt	gtccattgcc	gttcgaagga	aagtcaactgc	agtaacatac	1020
agcatcagtc	ggctctccga	agtgctgtcc	ccaatccgca	ctgtgogcct	cacgagggac	1080
aggaacacgg	acgccaacac	catggcaagt	ctattggagg	aaggtgacct	tgtgggtgtgc	1140
cccgagggca	ccacttgccg	ggagccttat	ctggtgcggg	tcagttcctt	gttcgctgag	1200
ctgacggatc	aaattgtccc	cgtgacgatg	aatatcaaaa	ttaccatggt	tcatgggacg	1260
agcgtctgag	gctggaaagg	catggacccc	ttcttcttct	tcatgaaacc	ttgccctaag	1320
tatgaggtaa	catttctgga	ccaactcccg	catcagctga	catgcaatgg	cggaaagacc	1380
agccacgagg	tggcgaacta	cattcagcgc	gtgttagcga	actctottgg	atttcagtgc	1440
accaatctca	cgcgcaagga	caagtatact	gtattggccg	ggaacgatgg	gattgtgccg	1500
gatcacactc	tgcggcagcg	agctggcaga	atgttcggct	tgtag		1545

5

<210> 114  
 <211> 1506  
 <212> ADN  
 <213> *Physcomitrella patens*

10

<400> 114



ES 2 640 100 T3

atgaagagag atcccttoga taccatogac aagtgcagtg acaaaggcag aagcaagcag 60  
 acggttgtct cggatctoga cggaaactctt ctctgcagcc gaagctcgtt ccctacttc 120  
 atgctggtgg cattogaagc tggaggagtg gcccgggcct ttgtgctcct cttgtgctct 180  
 ccgctgtgtt ggcttctgta ccattgcgtc tcagaatcca tcggcattcg cctcctgatc 240  
 tttgtcacct tcgcgggcct aaaaatcggc ggtatcgaag ccatcgcacg cggaacgttg 300  
 ccaaagtttt acgcggaaga cgtgcatccg gacacatggc gaatcttttc ttcgtttggg 360  
 gagcgttaca tacttacggc gacgccgagg atcatggcgg agacottcgc caagacgtat 420  
 ctgggagtgg acggtgtgtt ggggacggag ctgcatttca cttctggcgg catcgcgact 480  
 ggtttgttga tgaaccagg ggttctaacc gggagaaaca aagaaattgt gctgaggcaa 540  
 gaattccaag gcttgaacct tcctgacgtg ggtctcggag accgtccgtc cgaccacaac 600  
 ttcattgcca tttgcaaaga aggatacatc gttccgccga gcaacacaat actggcagcg 660  
 tcgaaagaaa gtctaataa cctgctggtg tttcacgacg gccgctgat ccaaaggccc 720  
 acagcaggca tagcgtcat catactgctg tggtaaccga ttggcgctgt gttggccgtg 780  
 cttcggatcc tggctggcat ccttcttccc ttccacttgt tgaaattggt atacaagttt 840  
 cttgggggtg ggggtgtgtt tcgtgggacc ccacccccag agcccacaaga cggccccgga 900  
 cgaggaggct acctgtacgt gtgctcgcac aggacgctac tagaccccgt gatggtgggc 960  
 gtggcggtga agcggcgcgct gacggccgtc acgtacagca taagccggct gtcggaggtg 1020  
 ttgtccocaa tcaaaacagt ggcaactgaag aggaatcggg agaaagacgc ggcgaagatc 1080  
 aggtctctgc tgcgagaagg agacottgog atctgccccg agggcaccac ctgtcgggag 1140  
 ccgtatctgc tgcgattcag cgcgctgttc gcggagctgt cgaaccagtt ggtacctgtg 1200  
 gcgatgaaca cgcgcatgag catgttccac ggcaccaccg cgcaggggtg gaaatgcttg 1260  
 gaccctttct acttcttcat gaatcccaac cctatctatg aggttacgtt cctcaacgaa 1320  
 ttaccggttg agcttacatg tgccggtggg aaatcttctg atgaagttgc gaaccacatc 1380  
 caacaacttt tgtogcaaac gttgggcttt gaatgcacaa attatacacg aaaagacaag 1440  
 tatggggtgc tctgtggaaa cgacggttca attcctttaa agtctcaaga ctcctttggt 1500  
 tcgtag 1506

<210> 115  
 <211> 2100  
 <212> ADN  
 <213> *Physcomitrella patens*

<400> 115

atggagacgc aagaaccgca ccctgtcaca agctttcact gcagcgaaaa ctatgacgag 60

ES 2 640 100 T3

gtggagaact gcaggggtgga gggtcgggca aatcagagca ttgtatcaga tttggacgga 120  
 acgcttcttc ggtcgcgaag ctccctcccc tatttcatgc tgctagcctt cgaagccggt 180  
 agtccattac ggttcatcat tcttcttctt gtctctccca tcgtttgggt cgtctacaac 240  
 ttcattctctg aggtagtcgg tattaagatg ctaaatattca tatcgcttgc ggggctcaaa 300  
 gtgtcagcca ttgagtcagt ggcgcgtggg gtactgccga aatttttctt ggaggacatg 360  
 cactccatca gctacagagt gtttatttcc tgcgagaaac gctacgtcgt cacagccaat 420  
 ccaaggataa tagtggagcc attccttaag gagtatttag atgtagacgc tgtgatgggc 480  
 accgagcttc aaatttcgag tgggggatat gctaccggct tcgtcaccgg ccctgggtgc 540  
 ttagtgggga cagcgaagca acgagctgtg aagaaatatt ttgggtccga ccagccggac 600  
 ctgggcctag gggaccgaac gagcgacttc gcattcatgg acctctgcaa agaggcttac 660  
 atagtgccta gctataagga ggtgccgtcc gtgacgaaga aagactactt gaagccggtg 720  
 attttccatg acggaaggct ggtgtgtcgt cctacgccgc tgatggcgct ggcggtgaca 780  
 ctttggtgcc cgattggctt cgtcttggcc atcatccgaa tgttcattgg cataatgta 840  
 ccaatgtggt gggctctccc tctagaagcc atgcttggcg tcacggtcog tataaaggga 900  
 attccccga catgtccctt tcggagcaag aaacgcggcg tgctgttcgt ttgttcacat 960  
 cgcactctcc tggaccccat tttcttgtct atagccgtgc gaaggaaagt cacggcagtg 1020  
 acgtacagca tcagtcgggt gtccgaagtc ctggctccga tccgaactgt acgtctcacg 1080  
 agagacagga acacagatgc gagcactatg gctagtctct tggacgaggg tgacctcgta 1140  
 gtgtgtcctg agggcaccac ttgccgcgag ccttatcttc tgcggttcag ttccttgttt 1200  
 gcagaattga ctgaccaaat cgtccccgta accatgaaca tcaagatttc aatgtttcat 1260  
 gggacgagtg ctcgaggctg gaagggcatg gatccgttct tttctttat gaacccccgc 1320  
 cctaagtaag aggtaacatt cttagatcaa ctcccgcatc agttgacttg caatggcggga 1380  
 aaaactagcc acgaggttgc gaattacatc cagcgtgtgt tggcgaactc tctcggattt 1440  
 caatgcacca atcttacgcg caaagataag tacaggggtg tggcggggaa cgatggaatt 1500  
 gtaccagatc gaagtctgog gcagagggcc agcagaatat ttggtttgca gtgacattag 1560  
 ggtttacttg taacgccaca tgtcaaactc cagcatcgga tggtcctgta cccaaataat 1620  
 ttcttaggtg gcggaggacc gatttttttc ctcaaattgt gcggagttag ggaaaacctt 1680  
 cggccccag ctatgacca catgtgctgc gttgtctgca tccagccccg cctcaactgt 1740  
 gtcgagcttc tgctgttgat tgattgacac cgaactagaaa ccaaatagta caattttaag 1800  
 aggataaaca tcgtgttata atttggaaag ttgaaccata caattccttt ctatcagtta 1860  
 cagcgtttga ctaactacag aaacaggaag acaactgttg acattgctct atctgcggcg 1920  
 gtaaaaagga cactggctgc ttttcagaga cgacataaa cttcggtta tgactgagtt 1980

ES 2 640 100 T3

cttgttgctc ctcaaacctt gaactgtagg actgtagcta gtgtctctgc tgtctcgcct 2040  
gtaagcagaa gcattaggcg tttgtattgt gtattattcc taccattcat tgaaggacat 2100

5

<210> 116  
<211> 1676  
<212> ADN  
<213> *Physcomitrella patens*  
<400> 116

ES 2 640 100 T3

atggaggtga cgaagttgaa agaggccggg aactacagct tcgcggaggt gaaatcgtgc 60  
agctcagga atgatcgag gagccagacc gtggtggcgg acttggacgg caccctcatt 120  
cgggggcggt ccgctttccc gtatttcttc ctggtggcgt tcgaggccgg gagttacttc 180  
agggcgctgc tgctgtcgct catggcgccg ctcatgggt ttcttacta ctttgtggac 240  
gaggcgccg ggattcacct catgatcttt ctctcattcg cggggctgaa gattaaggac 300  
attgagggtg tggcgcgcgc tgtgcttacc aagttctacg ccgatgacct tcaccctgag 360  
acttggcggg tgttctcgag ctttggcaag cgtgtgggtg tgacggcaaa cccgcgcgtg 420  
atggtggagc cgtttctgaa ggaggtcctc ggcgcggacg aggttatggg gacggagatt 480  
gaggtgaaca aaaagggccg cgctacaggc aggcttctga agccgggtgt tctgggtggg 540  
acacacaagg aggaagctct aaagcgggtg aacattggag gtgagcgccc gcacgtcggg 600  
ctaggcgacc gcgttacgga tttccctttc atggcgctact gcaaggaagg ctatgtagta 660  
cccaaaacga aagtgccagc cgttaagaag caagacatgc cgaagcagct aatctttcat 720  
gacggtcggc ttgtgcaact cccaacgccc cgtaatgctt tcaaagtgtt ggcatggatg 780  
cctatcggcc tgcttttgtc catcgctccg gtcacgatgg gtgtatgggt ccccatagc 840  
ttgatgccac tgtgttacia actcacaggc atcaacctcg ttgttaaggg aaacatccct 900  
gagaagccca agaatggaga acccgggaga ctcttcgtct gcaaccaccg gactcttctg 960  
gaccccgctc tcatcgctct ggcgcttggg cgcgccgtcc ctgocgtgac ttactccata 1020  
tccaaagtct cagagttcct gtccccatg ccaacaatcg gcctgtgccg agatagggag 1080  
aaagactcgg ccaatatgag aagggtgctg aaagaggggtg agctcactct ttgtcctgag 1140  
ggcactactt gcagggagcc tttcttctt cgattctcag ctctcttcgc cgagttgagt 1200  
gacaggattg ttctgtggc cgtaaaaatt aacatgaata tgtttcacgg cacaactgct 1260  
cgcgtaaca aagccatgga ccccttcttt gcttacatga accctcggcc gaccatcgaa 1320  
ctcaaattcc tagacgagat tccaaagaac atgacatgtg gtagcggaaa atcgtccatt 1380  
gaagtggcga actacataca acatctcctt gccggagaac tcggatatga ttgcacagat 1440  
ttcactcgca aggataagta ccggttctc gcaggaaacg atggcattgt acctgttaag 1500  
gagaaaacga ggtaatgaat cgacctgcaa gcacattacc acagaacaag acattgaaag 1560  
tgcttacgat tcagtatcac tagatggagg ggtaacccc gactcagtac tataagtgta 1620  
gaattcagtg ccaaaggcaa taaactgtga ttgattgtta cataaactcg atgcaa 1676

<210> 117  
<211> 1590  
<212> ADN  
<213> *Physcomitrella patens*

5

<400> 117

ES 2 640 100 T3

atggaaattc caggagtgtt tgatgctaatt ctctttctgcg ctgagtcatt tcttgaagtc 60  
 gaaacttgta aagtggaagg tcgtgaaaaa cagacgatta tttcagactt agatgggacg 120  
 ctctcagat caagaagctc gttccottac ttcattgctca ttgcttoga cgttgaagc 180  
 cctctceget tcatcgteet cttgctggcc tctccaataa tttggctggg ctacaacttc 240  
 ttctccgagg ctgccggtac caaactgctc attttcattg catgtgctgg cttgaaagcg 300  
 tccgacatcg agtctgtagc ggcggcgctg ttgccaaaat tttacttggg agacatgcac 360  
 tccgtcagct acagcatctt cgtctcctgc ggtaagcggg atgtggtgac tgcgaatccg 420  
 cgcattcatg tagagtcatt tctgaaagag tacatggggc togaggcagt cataggctact 480  
 gagcttcata tcacgaagag cggttacgcc acggggttgg taatgggtgc cgggtgtgatt 540  
 gtgggagcta acaaaacgcg cgtctggaag aaatactttg gcgatgacct tccagacatc 600  
 gctctcggag accgagccag cgactttcct tttatggctc tttgcaagga ggcatactta 660  
 gtaccgagct tcaagcctgt cgagccagtt gccaaagctg actacctgaa gccgctcatt 720  
 ttccatgacg gcagattagt atgtcgacca acacctctca tgtcaactcg aattattctc 780  
 tggctgccga tcggtcttat tctagcactc attcggatgg tgtttggcac cgtgttacct 840  
 atgtggtatg cgttccgct ggaggccatg cttggcgtta gcattccgct gaaaggaatc 900  
 cccccagct gcccccaaac cagcggcaga cgcgggtgtgc tcttctctg ctctcaccgc 960  
 actttgctcg accccatctt cctgtcaata gcctgtcggc gacaagttac agcagtaacc 1020  
 tacagcatca gcaaggtctc ggagatactg tcgccgatcc ccacagtcg actcacaagg 1080  
 tgtcggaaaa ccgatgcgga gaccatgaca aagcttctca acgagggaga tcttgcagtg 1140  
 tgcccagaag gcaagcgtg ccgagagcca taccttcttc ggttcagctc cttgttcgcg 1200  
 gagttagccg accaaatcat tcccgctcag atgaacgtca aaacgacct gttccacggc 1260  
 acgacagccc ggggctggaa ggcaactggac ccctttatct ttctgatgaa tcccagccct 1320  
 aggtatgaat tggagtttct agatcagctt ccgcaagaga tgacctgcag cggcgggaag 1380  
 agtagtcacg aggttgcgaa ctacattcaa cgagtcctcg cggcatctct gggtttcaag 1440  
 tgcaagaaac tgacgcggcg agacaagtat cgaatcttag cagggaaacga tggcgtgggt 1500  
 cctgcaagcc attttttgaa acaagaagc aggatcttcg catgcttgc ttcaggaat 1560

agtgagagag gccagcgtc agaccgtaa 1590

5

<210> 118  
 <211> 1512  
 <212> ADN  
 <213> *Vitis vinifera*  
 <400> 118

ES 2 640 100 T3

atggctgtaa ctaccttcc aactgtggac cagtgtcaat ccattggccg agaaaagcac 60  
 actgtggttg cggacatgga tggaaacctg ottattggca ggagctcctt tccttacttt 120  
 gctttggtgg cttttgaggg gggaggtgtt otaaggctgc tottcctgct ottggcttca 180  
 ccactogctg gacttcttta ctattttgtg tccgagctcg caggaatcca gattctcatc 240  
 tttgtaacct tcgctgggat gaaggtgtca gacattgagt cagttgcacg tgctgtgctg 300  
 cccaaatfff atttgagcga tctccatccc gagtcatggc ggggtgttctc atcatgtgga 360  
 aagcgtgtg ttcttaccgc aaaccccagg ataatggttg aggcatttct gaaagatfff 420  
 ttgggagctg atttagtttt ggggactgag atagccactt ataagggtag ggcaactggg 480  
 tttgtttgtg atccaggtgt acttgtggc aagaacaagg catatgctct caacaagact 540  
 tttggagaga ccagcctga gattgggctg ggtgatcgtc aactgattt tcccttcatg 600  
 gcttcatgca aggagggcta tattgtggca gccaaaccag aggttaaggc ggtaacaagt 660  
 gacaagcttc ccaagcccat catcttccac gacagccgcc tcgtccaaa gccaaactcct 720  
 ctcaocggcac tcctcaccct tctctggatc cccataggct tcgcaactcg atgcttgoga 780  
 atcgcogctg gctcactcct ccccatggc attgtctacc acgctttctg ggcactgggt 840  
 gtccgagtct ccgtcaaggg ttccccacct octgttgcca aaaagtcaat aggccagtca 900  
 ggcgtgctat tcatctgctc ccaccgaacc ctocctgacc caatattcct ctccgcagcg 960  
 ctcgccgctc ccattcctgc agtcacgtac tccgtctccc gcctctccga gttcatctcc 1020  
 cctatcaaaa ccattcgcct aagccgagat cgagtcaaag acgcagccat gataaagaaa 1080  
 ctgcttgaag aaggtgacct agctatattgc cccgaaggaa ccacttgccg tgagcctttc 1140  
 cttctcaggt tctcagcatt gtttgctgaa ctgactgacc agcttgttcc agtggctatg 1200  
 atgaaccgca tgagcatggt ccatggaaca acagcaagag ggtggaaagg gatggacca 1260  
 tttttcttct tcatgaatcc aagcccagcg tatgaggtaa cgttcttgaa caagctgcc 1320  
 ctggaactta cttgcagttc ggggaaatct agccatgag tggccaacta catgcagagg 1380  
 gtgattgctg cactctttc ctatgaatgc actagcttta ccaggaaaga taagtaccga 1440  
 gcaactcgtg gaaatgatgg gaccgtgcc aaaaaacctc tgctatatcc cgacaaagtc 1500  
 atgggatgct ga 1512

<210> 119  
 <211> 1784  
 <212> ADN  
 <213> *Vitis vinifera*  
 <400> 119

5

ES 2 640 100 T3

gtttcttttc gttgtccact gtgctgcctg aaaggctgaa tccggttcta ccggccgaga 60  
 catgtctccg ccaccgaaac ggcgccgaaa atttccctcc atcaactacct acgcccggcg 120  
 cgaccatogc tctattgcgg ccgacctoga tggcactctc ctctctctct gctcctcctt 180  
 cccttacttc atgcttgtcg ccgtcgaagc cggcagcctc ctacgcggtc tcttccctct 240  
 gctctctctt cctatagtga tcgttgccca cctcttctac tcagaagaaa tcggaattca 300  
 gattctctac ttcatttcat tttccggcct caagatccga gacatcgagc tcgctgcgca 360  
 ggcggtgctg ccgcggttct acgccaacga cgtgaggcga gagagctggg aggtgttcga 420  
 gaagtgcgaa cgggaaggtg tggtgacggc caaccgcagc ttgatggttg agccgttcgt 480  
 gagagatttt cttggcggaa cgaaggtgtt agggacggag attgaggtga atccgaagac 540  
 gaagaaggcc actggtttcg tgaagaagcc aggggtgttg gttggagaca gaaagagatt 600  
 ggcgctgttg aaggagtttg gtgatgaatt gcctgatatc gggattggtg accgcgagtc 660  
 tgatcatgat ttcattgtcca tttgcaagga gggatacatg gtgcttccga gtaaatcagc 720  
 tacgcggtg ccaccaaac gcctgaaaac cccaataata ttccacgatg gtccggttcgt 780  
 ccaaccccc acgccactca cggccctcat catctacctc tggcttccat ttgggtttgc 840  
 actctctatc ttccgagtct acttcaacct cccctccca gagcgcatcg tccgctacac 900  
 ctaccgatg ctgggcatca accttgttat ccgaggcaac cctcctccgc cccctcccc 960  
 aggttcccc ggcaacctct acgtctgcaa ccaccgcacc gccttgacc ccatcgctat 1020  
 cgccatcgcc ctccgccgca aggtctcctg cgtcacctac agcgtcagcc gcctctctcg 1080  
 ctctctctcc cccatccccg ccgtcgcctc caaccgcgac cgcgcgcgag acgctgcccg 1140  
 catctctctc attctccaga agggcgacct cgtggtctgc ccagagggga ccacctgccc 1200  
 cgagccctac ctctccggt ttagcgcact gttcgcggaa ctccagcagc gcctcgtgcc 1260  
 ggtggcggtg aacgtgaagc agaacatggt ccaccggcacc accgtccgag ggtggaagtt 1320  
 ctgggacgcc tatttttact tcatgaacct tcggcccacc tatgagataa cgtttctgga 1380  
 tcggttgccg gaggagatga cgtgtaaggc cggaggggag tcggccatag aagtggcgaa 1440  
 tcacgtgcaa aaggtgttg gtggagtgtc ggggtttgag tgcaactggac tgactagaaa 1500  
 ggataagtac atgttgctgg gaggaatga tggcaaggtg gaatccatgt acaatgocaa 1560  
 gaagtgaaag gctgggatog gagaagaagg ccacattgga attgtatggg ttttagaaaa 1620  
 attacgaggg aggtatatcg gttaatgtag aattacaact tctatggaaa ttcccttagt 1680  
 atgattgata attgttgatc attttctac tggattatc atattatttg gacattcaag 1740  
 ttttagagtta tgaagagtgt cttctgattt ctgtggttga tttt 1784

- 5 <210> 120
- <211> 1533
- <212> ADN
- <213> *Vitis vinifera*
- <400> 120

ES 2 640 100 T3

atgaccaa at caaacctcca ctctcctccc acctccctaa ccatcgatca atgcacctcc 60  
 atcggccggg aaaaggagac ggttgtggcg gacatggacg gaactcttct cagagggcgc 120  
 agttcgttcc cctactttgc acttgttgcc ttcgaggttg gtgggatttt taggcttctt 180  
 ctgttgttgc tatgttctcc attggctgga gttctctact acttcatatc ggagccagct 240  
 ggcatcagc tgctcatatt cgcgacctt gtggggatga aggtttcggg tatcagatcg 300  
 gtggcgcggg cagttttgcc caagttttac gcgagcgacc tgcacccgga ggcgtggagg 360  
 gtgttctcgt cgtgtgggaa gaggtgtgtg ctcaccgcga atccgaggat tatggtggag 420  
 gccttcctga aggatttttt ggggtgcagac atgggtgttg ggacggagggt ggacacgtat 480  
 aaagggagag ctactgggct tgtccgtggt ggtggggttc ttgttgggaa gagaaaggca 540  
 gaggccctcc gaaatgcctt cagcgagacg tctttaccgg agattggact cggatgatcg 600  
 gagacggatt atccatttat gagtctctgc aaggagagtt acatagttcc agogaagcct 660  
 gaggtggagg ctgttagcca tgacaaattg ccaaagccca ttgtattcca cgcgcggcgc 720  
 ctctccaaa aaccaacccc tttcatggcg cttctcacca tctctggat cccagtaggc 780  
 ttctcttg cctgcttgcg catcgccgct ggtgcgctcc ttccaatgcc gcttgtctac 840  
 tacgcctttt gggcactcgg tgtcagagtt tatatcaagg gcaaccccc tccaccagcc 900  
 aaaaaatcaa taggccaatc tggcgtcctc ttcatttgc ttcacagAAC cctcctcgac 960  
 cctatcttcc tctccgcgcg attaggcgcg cctatccctg cagtcacgta ctccctttct 1020  
 cgtctctcgg aattcatctc ccccatcaag acagtcgggc tcagtcgcca tcgagccacc 1080  
 gacgcagcaa tgataaagaa gctcttagag gcaggtgacc tagccatag tcccgaagga 1140  
 accacttgcg gggagccatt tctgctaaga ttttctgcct tgttttctga gctaaccgac 1200  
 gagctggtgc cgggtggctat ggcaaaccgc atgaccatgt ttcattgggac aacagctaga 1260  
 ggttgaagg ggatggacc cttttactac ttcattgaacc cttagccctgc ctatgaggta 1320  
 aacttcttga ataagcttcc acatgagctc acttgcgggtg ctggaaaatc gagccacgac 1380  
 gtcgcgaact acatacagag gatgattgcc gcaagtctat catatgagtg caccaacttc 1440  
 actaggaagg ataagtatcg ggcactggcc ggaacgatg gaactgtagt ggaaaagccc 1500  
 aaggttgatg ctaataaagt catgggctgc tag 1533

5 <210> 121  
 <211> 1726  
 <212> ADN  
 <213> *Vitis vinifera*  
 10 <400> 121



ES 2 640 100 T3

atgggggctc aacgccattt ccaaccaata tcgaaatgca gctccgacgg tcgagccaat 60  
cagacagtgg cggcagactt agatggcaog cttcttgtgt cccgaagcgc atttccctac 120  
ttcatgcttg tggcacttga agctggtagc cttgtgaggg caatcctcct cctagcatcg 180  
gttccactag tgtatttcat ttacttgttc atttccgagt cgctagccat taaaatcttt 240  
atctatatgg cttttgcggg actcaaagtt agagagattg agcttgtttc gaggtcagtc 300  
ctgcccaagt tctatgctga ggatgtgcac cccgagagtt ggagagtgtt taattccttt 360  
gggaagaggt atatcattac tgccaatcct agaattatgg tggaaacctt tgtgaagacc 420  
tatatgggtg ccgataaggt tcttgggacc gagttggaag tatcgaactc cggaagggcg 480  
actgggtttg tgaagaaacc aggagttttg gttggggaac ataagaaagc ggcagtgcaag 540  
caagaatttg gcatgaattt gccagatgtg ggcctgggag atagagatac ggatcatgat 600  
ttcatgtcct tatgcaagga aggatacatg gtgccaaagga caaaatgtga gccactacca 660  
aggaacaagc ttctgagtcc agtcattttt caccgaggac gtttggttca gagaccacc 720  
cctgtggctg ccctgttgac tttcttatgg atgccaatg ggatcattct ctccatcctt 780  
agggtctacc tcaacattcc cttgccagag cgtattgtga ggtacaatta caagatcctg 840  
ggaatcaagc ttattgttaa ggttacacot cctccccgc caaagggagg ccaaagtgga 900  
gtcctcttcg tttgcaacca taggacagtt ttggaccccg tggtcacggc tgtggettta 960  
ggcagaaaaa tcagctgtgt cacttacagt attagcaaat tctcggagct aatttctcca 1020  
atcaaagccg ttgcactatc cagggagaga gaaaaggatg cagcaaatat caagcgetta 1080  
cttgaagaag gtgatttagt aatttgtccg gagggaacta catgtagaga gccatttctg 1140  
ttgagattta ggcgcccttt tgctgagctc actgatagaa ttgtgcctgt tgccattaat 1200  
accaagcaga gcgtgtttta tggcacgtca acccggggtt ataagttggt ggatccttac 1260  
ttcgtgttca tgaatccaat gccacctac gagatcacct tcttgaatca gctgcccgcc 1320  
gagctcactt gcaaaggtgg gaaatcccca atcgaagttg caaattacat ccagagggtg 1380  
ctggctggaa cccttggtt tgagtgcact aacttgacca ggaaggacaa atacgccaca 1440  
cttgccggaa cagacggcag tgttccgttc aagaacgaaa aggaaaagga aaaggcttga 1500  
gaactgaaag tgattcaaag taaattaatt agtctcatga ttttctattg ttttcatat 1560  
attatatgcc ggaatcatta tcatgacaat gaactccgat tgtgatgata ataaattggt 1620  
gtgattataa aaataatagt ttcttatatt gtcgtggaga caagtcaacg gtcatttttt 1680  
ttaagttttt gaaaataaat tttatctgga caagtttcac cttaga 1726

<210> 122  
<211> 1504  
<212> ADN  
<213> *Populus trichocarpa*  
<400> 122

ES 2 640 100 T3

tcctctctca	tggaatggg	ggctcatgc	cattttgaac	caatagcaa	atgtagcaca	60
gagggacgat	caaatcagac	tgtggctgcc	gacctggatg	gaacacttct	tgtatcaaga	120
agtgcctttc	catacttctt	gcttggtgct	attgaagctg	ggagtctctt	aagaggacta	180
attctcttag	catcagtccc	atttgatatac	tttacgtact	tatttatatc	agaggcaatg	240
gcaatcaaga	cttcoactctt	cattgccttt	gcgggactta	aaataagaga	cattgagctt	300
gtttcaaggt	ctgtccttgcc	aaagttttat	gctgatgatg	ttcatccaga	aacttggagg	360
gtgttcaatt	cttttgaaa	aagatacatt	gttactgcta	atcctaggat	tatggtcgag	420
ccatttgtca	agacctttt	agggctgac	aaggttcttg	gcaactgaatt	agaggctaca	480
aaatctggaa	gagcaactgg	gtttatcaag	aaacctggag	ttcttggttg	agatcataaa	540
agagatgctc	tcttaaaaga	gtttggcaca	aatttgcctg	atntaggoct	gggagaccgc	600
gaaactgacc	atgacttcat	gtccatttgc	aaggaaggat	atatggtgcc	aggaaccaag	660
tgtgagcctc	tagcaagaaa	caagcttcta	agccctgtca	tattccacga	tggccggtta	720
gttcaaaggc	ctactccttt	ggctgccctt	ttgaccttct	tatggatgcc	aattggtatt	780
attctctcca	tacttagagt	ctaccttaac	atcccttgc	ccgagagact	tgccctggtac	840
acttataagc	tactaggaat	tagagttatt	gtcaagggtg	ccctccacc	ccctccagga	900
aaaggccaca	gtggagtcct	ctttgtttgc	aatcatogta	ctgtcttaga	cccagttgtc	960
actgctggtg	cactaggaag	aaaaattagt	tgtgtcacct	atagcataag	caagttcact	1020
gaaattattt	caccattaa	agctggtgct	ttatcaaggg	agagagacaa	agatgctgct	1080
aacattaagc	gactgcttga	agaagggtgac	ctggtcattt	gccctgaagg	aaccacatgt	1140
agagagccat	ttctcttgag	atthagtgct	ctttttgctg	agcttactga	taggattgtg	1200
ccagttgcta	tcaatacaaa	aaaagtgtg	tttcatggca	cgacggttcg	agggcacaaa	1260
ctgttggacc	cttattttgt	gttcatgaat	ccaatgcaa	catatgaggt	caccttcttg	1320
aatcagctgc	ccaaagagct	tacttgcaaa	ggaggcaaat	cagocattga	agttgcaaat	1380
tacatacaaa	gggtgctggc	tgggacactt	ggatttgagt	gcactaactt	gaccaggaag	1440
gacaagtatg	ctaagcttgc	aggaacagac	ggccgtgttc	tatctaagaa	ggaaaatgct	1500
tgag						1504

<210> 123  
 <211> 1506  
 <212> ADN  
 <213> *Populus trichocarpa*  
 <400> 123

5

ES 2 640 100 T3

atggttagct tcccaactgt tgacaaatgt gcatccatcg gcagagaaaa gcacagtgt 60  
 gttgctgaca tggatgggac cttgcttaga ggocgtagct cattccotta ctttgctcta 120  
 cttgcottcg aagctggtgg gatttttagg ctgcttttct tgctcttgaa ctcaccacta 180  
 gcaggacttc tgtattactt tgtgtccgag tctgctggta ttaaagttct tatctttgct 240  
 acatgtgccg gaatgaagtt atcagatata gagtctgtag cacgtgctgt gctgccaaag 300  
 ttttactcta gtgaccttca ttctgagtcg tggcgtgtgt tctcttcttg cggaaaacgt 360  
 tgcgttotta ctgcaaatcc aagaattatg gtggaagcat ttttgaaaga tttcttggga 420  
 gctgatttgg ttttggggac tgagatatta acttataaag gtagagcaac agggtttgtt 480  
 cagagcccgg gactacttgt ggggaagaac aaggcagatg cccttaaaaa ggcttttggg 540  
 aagacacaac cagagattgg acttggagat cgtcatacag atgcaccctt catggctcta 600  
 tgcaaggagg gctacatagt gccacctaag ccagaggttg aagctgtaac tactgacaag 660  
 ctcccaaac cagtcatctt ccatgatggc cggctagtcc aaaaaccaac accattatca 720  
 gcactactca taattctctg gatccctata ggcttcattt tagcctgctt gaggattgca 780  
 gcagggtcac tcctgcctat gccaatggtc tactatgctt tctgggcact tgggtgcccgt 840  
 gtcaacataa agggaacccc acctcccca gccaaaaaat caattggtca atctggtgtc 900  
 ctctttgttt gctcacacag aacottgott gatcccatat ttctctccac tggccttggg 960  
 cgcctattc cagcagttac atactcactt tctcgtctct ctgaaattat ctcaccaatc 1020  
 aaaactgtta gactcagccg tgaccgagct gcagacgcat caatgatcaa gaagttatta 1080  
 gaagaagggtg acctagctat atgtcccag cgaactactt gccgcgaacc gtttctttta 1140  
 aggttctcgg ctttgtttgc tgaattaaca gacgaactcg ttccagtggc tatggcgaat 1200  
 cgcattgagta tgtttcacgg aaccactgca agaggctgga aagggatgga tccattttac 1260  
 ttcttcatga acctagccc agcatatgaa gtaactttcc tgaacaagtt gccactagag 1320  
 ctaacattaa gtggtgggaa atctagccat gaggtggcaa actacataca aagagtgatt 1380  
 gctgctactc tctcctatga atgcacttcc tttactagga gagacaagta cagggcactt 1440  
 gctgggaatg atgggaatgt ggttgaaaag acgaaggtcc aagccaacaa agtaatgggg 1500  
 tgttaa 1506

<210> 124  
 <211> 1506  
 <212> ADN  
 <213> *Populus trichocarpa*  
  
 <400> 124

5

ES 2 640 100 T3

atggatagct ttccaactgt tgacaaatgt gcatccatag gtagagaaaa gcacactgtg 60  
 gtagctgaca tggatggaac cctgcttaga ggtcgtagct cattccctta ctttgcctta 120  
 cttgcctttg aagctgggtg gatttttagg ctgattttct tgctattggc ctcaccgcta 180  
 gcaggacttc tatactactt tgtgtccgag tctgctggta ttaaggttct tatatttgc 240  
 acatgtgccg gaatgaaggt ttccgatatt gagtctgtgg ctcgtgctgt actgcctaag 300  
 ttttactcta gcgatcttca ttctgagtcg tggcgtgtgt tttottcttg tggaaaaaga 360  
 tgtgttctta ctgcaaatcc aagaattatg gtggaagcat ttttgaaga tttcttggga 420  
 gctgatttgg ttttggggac agagatgtct acttataaag gtagagcaac agggtttgtt 480  
 cagagcccgg gagtacttgt ggggaagaac aaggcagatg cccttaaaaa ggcttttggg 540  
 atgacacaac cagagattgg acttgagat cgtcatacag atgcaccctt catggctcta 600  
 tgcaaggagg gctacatagt gccacctaag ccagagggtg aagctgtaac tactgacaag 660  
 ctcccaaac cagtcactctt coatgatggc cggctagtcc aaaaaccaac accattatca 720  
 gcactactca caattctctg gatcccaata ggcttcattt tagcctgctt gagaattgca 780  
 gccgggtcac tcctgcctat gccaatggtc tactatgctt tctgggcact tgggtgccgt 840  
 gtcacataa agggaacccc acctcccccg gccaaagaaat caatcgggtca atctggtgtc 900  
 ctctttattt gctctcacog aaccttgott gatccaatat ttctctccac tgcctcggc 960  
 cgtcctatcc cagcagtcac gtactcagtt tctcgtctct ctgaaattat ctcaccatc 1020  
 aaaactgtta gactcagccg tgaccgtgct acagacgcat ccatgatcaa gaagttgtta 1080  
 gaagaagggtg acctagctat atgccctgag ggaactactt gcogcgaacc ctttctttta 1140  
 aggttctcag ctttgtttgc tgaattaaca gaccaacttg ttccagtggc tgtggtgaat 1200  
 cgcagagta tgtttcatgg gacaacagca agagggtgga aagggatgga tccgttttac 1260  
 ttcttcatga acctagccc agcatatgaa gtaactttct tgaacagggtt accacaggag 1320  
 ctaacctgca ctggtgggaa atctagtcac gaggtggcaa actacataca aagagtgatt 1380  
 gctgccaccc tgtcctatga atgcacttcc ttacttagga gagataagta cagggcactt 1440  
 gctgggaacg atggtactgt ggttgaaaaa acgaagctcc aagccaaca agtaatgggg 1500  
 tgttga 1506

<210> 125  
 <211> 1797  
 <212> ADN  
 <213> *Populus trichocarpa*

5

<400> 125

ES 2 640 100 T3

acaacataac tcacacttcc taagaactct caaaataaac actatctatc cctctccctc 60  
 tccctctcta tccccctctg tgatggcatc gggggctcat cgccatthttg aaccaatagc 120  
 aaaatgtagc acagaggggc gatcgaatca gaccgtggct gcagacctag atggaacact 180  
 tcttgatca agaagtgctt ttccttactt cttgcttggt gctattgaag ctgggagtct 240  
 cctgagagga ctaattcttt tatcatcagt cccatthttg tactttacgt acttaattat 300  
 atcagaggaa atggcaatca agaccttcat cttcattgcc ttcgctggac ttaaaataag 360  
 agacattgag cttgtttcaa ggtctgtctt gccaaagttt tatgctgatg atgtccatcc 420  
 agaaacttgg aggggtgttca attcctthttg aaagagatac attgttacag caaatcctag 480  
 gattatggta gagccatthg ccaagaactt tttaggtgct gacaagttc ttggcactga 540  
 aatacaggct acaaaatctg gaagggcaac tgggtthttg aagaaacctg gagttcttgt 600  
 tggagatcat aagaggaatg ctctcttaa ggagthttgt acaagthttc ctgatttagg 660  
 cctgggagac cgcgaaactg accatgactt catgtccatc tgcaaggaag gatatattgt 720  
 gccaaagatcc aagtgcgagc cactaccaag aaacaagctt ctaagccctg tcatattcca 780  
 tgagggccgg ctggttcaaa gacctactcc cttggttgcc cttctgacgt tcttatggat 840  
 gccaatthgt attattctct ccgtccttag ggtctactt aacatccctt tgctgagag 900  
 aattgctggt tacaattata agctgctagg aattagagta gttgttaagg gtaatcctcc 960  
 accccctcca agaaaaggcc aaagtggagt cctctthttg tgcaatcacc gtactgtctt 1020  
 agaccctgth gtcactgctg tgcocctaag aagaaaaatt agctgtgtca catacagcat 1080  
 aagcaagthc actgagatta tttcacctat caaggctgth gctthtcaa gggagagaga 1140  
 cagagatgct gcaaatataa agcgtctact tgaagaaggt gacttggtaa tttgcccgga 1200  
 aggaaccact tgtagagagc cattctctct gagatthagt gctctthttg cagaacttac 1260  
 agataggatt gtgccagtcg ctatcaatac aaaacaatct gtcttcaatg gcacgacggt 1320  
 tcgagggcac aaactgthgg acccttattt tgtgttcatg aatccaatgc caacatatga 1380  
 gatcaccttc ttgaatcagc tgcccacaga gcttacttgc aaaggaggca aatcatccat 1440  
 tgaagthtga aattacatac aaaggggtgct ggctgggaca cttggatthg agtgcactaa 1500  
 tthgactagg aaggacaagt acgctatcct tgcaggaaca gacggccgtg ttgcaactaa 1560  
 gaaggaaaag gcttcagaaa attaatctct ccaaaaggcg gcataaaaga aatccatga 1620  
 tthaatthct ctattgthtt tthttatata taaaatattg ctataacaat aaaatthgac 1680  
 aatcgtatt gtgattaatt actataaaat aactthtagag tatatthgt tthtttattt 1740  
 tthaatgtag tataccgaca tcaaatthct ctaatgcaat gthttatthaa tcttcaa 1797

<210> 126  
 <211> 1575  
 <212> ADN  
 <213> *Sorghum bicolor*

ES 2 640 100 T3

<400> 126

atggagcgtc ccagcccaac gtgcttagcg gcgaacggcc cacgtagaaa agcccgaaca 60  
atcccgttcc cgccgceccc ggctgatctt cgtacgggga gggtgcggtc gcggcagacg 120  
gtogccgcg acctcgacgg cacgctgctc ctgtcccga ggcggttccc gtactacctc 180  
ctcgtggcgc tggaggccgg cagcctcctc cgcgccgtgg cgcttctcat gtccgtgccc 240  
ttcgtgtacc tcacctacgt cacogtgtcc gagaccctgg ccgtgcgcgc gttcctctac 300  
gtcgccgtgg cggggctcga cgtgagcgac atcgaggccg tcgcccggac cgtgctcccc 360  
aggttctacg ccggcgacgt ccacccggag ggctggcgcg tcttccgctc gttcggggag 420  
agggtgcgtc tcaccgccag ccgcgggtg atgggtggag cgttcgccag ggcgttcctt 480  
ggcgccgacg tggtcacgga gacggagatg gaggtggcg cgctccggaa ggccacgggg 540  
ttcgtgcgcg ggcccggggt gctcgtcggc gaacacaaga ggcgggcggt ggtgagggag 600  
ttcggcgacg cgctgccgga cgtcggcatg ggggatcgcg agagcgactt cgacttcatg 660  
gccatctgca aggaagcgta catggtaaaca cggcagaagt accgcgcgtt gcccggggag 720  
cagctgcaga gccgggtcat cctccacgac ggccgcctgg cccggcgacc gacggcgacc 780  
aacacgcttc tcaactctct gtggatgcog ctcggtctcg cgctcgcgct ggcgcgcgtg 840  
catctccacc tgctcctccc ggcgcgcgcc ctctcctacg cctacaagct tatgggcgtc 900  
aagctgggtg tgcgcgga caagcccgcc ccgcccgcgc cgtccaagaa ggcggcgggc 960  
cccggggttc tcttctgtg caaccacgcg accacgctcg accctgtcgc ggtggcgcctc 1020  
gogctgggcc gcaaggtgag atgggtcgtc acggacggcg cctccagcag gttctcggag 1080  
gccgtgtcgc ctgtcatgac aggcgtggcg ctgcccgtgc ccgtgccag ccgcgagagc 1140  
gacgacgacg cggacgcgcg cgcccgcac cggcggtcgc tcgaggaggg agacgacgtc 1200  
gtcatcttcc ccgagggcac catctgcgcg gagcccttcc tgctgcggtt cggcgcgctc 1260  
ttcgtgagg tcaccgaccg catcgtgccc gtggccatcg gcgctagga gggcatgttc 1320  
cacgggtcca cggcgcgcgg gctcaggaga atggatccct acttctttt catgaacca 1380  
cggccgacgt acgaggtgac gttcctgaac cagcttcaa gggagctcac ctgcggcggc 1440  
ggaaggtcgc cggtcgaggt ggccaactat gtccaggagg tcctggccgc gcagcttggg 1500  
ttcgactgca cttgcaccag caagcaggca aaaacaggat ggtctccggc ggccccgacg 1560  
gttgcgtggc gttga 1575

5

<210> 127  
<211> 1967  
<212> ADN  
<213> *Sorghum bicolor*

10

<400> 127

gcctcccata tccatactcc ccgcgcgccg tcaccaccaa cctcccgctc cgtatattat 60

ES 2 640 100 T3

aactacaact cgccgcccc ctctccactg ccaactcgca gcagcagtct ctctctgct 120  
cccacacact ccttccccgg ggccctggcag tgagatcgag cctcgatcca tcgagcacta 180  
tagtagcaaa gacacaagat ggtggcgctg tatcgcgatc gcaggttcaa gcccatcgag 240  
gagtgctgct cggagggggc gtccggagcag acgggtggccg ccgacctgga cggcacgctg 300  
ctcatctccc ggagcgcggt cccctactac ctccctcgtag cgctcgaggc cgggagcgtc 360  
ctccgcgcgc tgetgctgct cctgtccgtg cctttcgtct acgtcaceta catcttcttc 420  
tccgagtcgc tggccatcag cacgctcgtc tacatctccg tcgcggggct caaggtgcgc 480  
agcatcgaga tggtagcgcg ctccgtgctc cccaggttct acgttggcga cgtgcacccg 540  
gagagctgga ggggtttcaa ctccctcggc aagcgctaca tcatcaccgc cagccccagg 600  
atcatggtcg agcccttcgc cagggccttc ctccggcgccg acaaggtcgt cgggaccgag 660  
ctcgaggtcg gcaagaacgg caaggccacg gggttcatgg tcaagcccg cgtgctcgtc 720  
ggcgaccaca agaagaaggc cgtcgtcaag gagctggggc acgccgtgcc cgacgtcggc 780  
atgggggata gggagaocga cttcgacttc atgtocatct gcaaggaggc gtacctggtg 840  
acgtcgagga agtacagccc ggtgggcaag aaccagctgc tgagcccgt gatcctacac 900  
gacggtcgcc tggtagcagc cccgacgccg ctccgtcgcgc tcgtcacctt cctctggatg 960  
ccgttcggct tcgogctcgc gctcatgccc gtgtacatca acctgccgct gcccgagcgc 1020  
atcgtctact acacctaaa gctcatgggc atcaggctga tcgtcaaggc caaccgcgcg 1080  
ccgccgcccc agaagggcc cccgggcgct ctcttcgtct gcaaccaccg caccgtgctc 1140  
gaccccgctc aggtcgcgct ggcgctgccc cgcgaaggta gctgcgtgac ctacagcatc 1200  
tccaagttct ccgagctcat ctcccccacc aaggccgtcg cgctgtcgcg ggagcgagag 1260  
aaggacgccc agaacatccg ccgcctgctg gaggaaggcg acctggtgat ctgccccgag 1320  
ggcaccacct gccgcgagcc cttcctgctg cggttcagcg cgctcttcgc cgagctcacc 1380  
gaccgcatcg tgccggtggc gatcaacacc aaggagagca tgttccacgg ctccaccgtg 1440  
cgcgggttca agctcatgga ccttacttc ttcttcatga acccgcgccc gacgtacgag 1500  
atcacgttcc tgaccagct ccccaaggac ctccatgca gggcgggcaa gtcgcccacc 1560  
gaggtggcca actacatcca gaagacgctc agcggacagc ttggcttcga gtgcaccacc 1620  
atcacgcgca aggagaagta cggcatactc gctggaacgg acggccgtgt cccgtccaag 1680  
aacaaggaca aggagaagga gaagaactaa cggccaacgt cgtcatcgt gctgcccgcc 1740  
cagcacactg ccgttacgag tactgctgaa atttattcca attaattggt ccttgtogag 1800  
aagttggcca attactcatt gtttttaaaa ggaccatgtg atgtgggccc tgtactagta 1860  
caagcttgat gatcatataa ttattatcta caagttcttc ctcttcggaa gaagggcaaa 1920  
agtgttgta ggcocgtctt aatacaaagt ttcggtgcat agtttcc 1967

<210> 128  
<211> 1518  
<212> ADN  
<213> *Sorghum bicolor*

ES 2 640 100 T3

<400> 128

```

atggctgcct cgccgttccc cacggtggag aagtgtcct ccaccgaccg gtccggcgac      60
acggtggtgg cggacctgga cggcacgctg ctgtgcggcc ggagctcctt cccgtacttc      120
gcgcacatgg cgttcgagac cggcggcgtg ctgocgctgc tgctgctcat cgcgctggcg      180
ccgctcgcgg gcctcctcta ctacttcgtg tccgagccgg cgggcatcca ggtgctcatc      240
ttcgcgtcca tggcgggggc cagggtggcc gacatcgagg ccgtggcgcg ggcggtgctg      300
cccaagttct actgcgccga cctccaccgc gagtcgtggc gcgtcttctc ggcgtgcggc      360
cgccggtgcg tgctcaccgc gaaccocggg atcatggtgg aggcgttcct caaggagtac      420
gtcggcacgg acgtcgtcgt cggcacggag ctcgtcgtgt ggcgaggccg cgcgacgggg      480
ctggtgcgct ccccgggcgt cctggtcggc gagcagaagg cggacgcgct ccggaggacg      540
ttcggcgacg acgtcgcgcc cgaggtcggc ctcgcgaca ggaagacgga ctaccogttc      600
atgaggctgt gcaaggaggg ctacgtcgtg ccggcgacgc ccaagctgaa gcccgtgccg      660
cgcgagaacc tgccgaagcc ggtggtcttc cagcaggggc ggctcgtcca gaagccgtcc      720
ccggcgctcg cgtgctcac cgtgctctgg atcccgatcg ggttctctct cgcgtgcctg      780
cgcacgcgcy cgggcgcgct cctgcogatg cgcacgtgt accacgcgtt ccgcgccctc      840
ggcgtccgtg tcacatcag gggaaacct ccgcgcggc caagccgcga gacgggccag      900
accggcgtgc tttcatctg ctcccaccgc acctcctcg accccatctt cctctccacc      960
gccctgggoc gccocatcac cgcctcacc tactcgtct cgcggctgtc ggaaattctg     1020
tcgccaatcc gcacggtgcy cctgaccoga gaccgcggc cggacgcggc catgatccgg     1080
cggctcctga ccgagggaga cctggtgatc tgcccggagg ggacgacgtg ccgggagccg     1140
ttctgctgc gtttctcggc gctgttcgcy gagctgacgg acgagatcgt gccggtggcg     1200
atggagaacc agatgagcat gttccacggg acgacggcgc gagggtggaa agggctggac     1260
cctttctact tttcatgaa cccgagcccc gggtacgtgg tcacgttcct caacaagctc     1320
ccagcagagc tcacctgcaa tggcggcggc aagacagcc acgaggtggc caactacatc     1380
cagcggctca tcgctccac gctgtcctac gactgcacca acttcacacg gaaggacaag     1440
tacaaggcgc tcgctggcaa cgacggcacc gtcgtgtcca agcccaacat tgacaagaac     1500
aaggtcatgg gctgctag                                     1518

```

5

<210> 129  
 <211> 1776  
 <212> ADN  
 <213> *Sorghum bicolor*

10

<400> 129



ES 2 640 100 T3

gcagtgcgca gtgcgagaga gagacagagt gagtgcacagc tccgatccgc cgcacacgog 60  
cgtgccacag cagatagcag cgagccatcc atccccggcg cccgcatgga ggaggcggtg 120  
gaggctgtgg ttgccgcccg agtggagccc ttccccgacg tggacaagtg cgacgcctcg 180  
ggcctcggcg cgcactcctc cgtggtggcg gaccttgacg gcacgctgct cccgtcccgc 240  
agcgcgttcc cgtactacgc gctggtggcc ttcgagacog gcggcgtgcc acgcctgctc 300  
ctcctcctgc tcctctcccc gctggcccgc ttgctgogct tccctggcgtc ggagtccgog 360  
tgctgcgog tgctggtgtt cggtgccacg gcgggcccgc gtgtccgoga cgtcgagtcc 420  
gcggcgcgog cctgctccc caggttctac gccgcccacg tgcaccoggc ggctggcgc 480  
gtcttctccg cgtgctccc gaggcgogtc gtgctcacgg ccacgcccag gctcttcgcc 540  
gagccgttcc tcagggactg cctcggcgcg gacgcccgtc cgggcaccga gctcggccag 600  
tggcgcggac gcgccacggg cttggtggac gcgcgcaggg gcggcgtcct cgtcggcagg 660  
agcaaggccc aggcgctgog ggagatcttg gccgcccacg gcggcgcagc gcccgacgtc 720  
ggactcgggg acagcccctc cgactacccc ttcatgagca tgtgcaagga ggcgtacatc 780  
gtaccgcgog cgcgggtgga gccggcattg cccatggacc agctgcctcg tccggctatc 840  
ttccacgacg ggcgctggt gcgccggccg acgcccgtgg tggcgtggt ggtggtgctg 900  
tggttccccg tggggcttgc gctggcctgc ctccgcatcg cggcgggggc gctgctgcc 960  
atgccgctgg tgtactacgc cttctggggc ctgggcgtgc gcgtcgtggt gcggggcogog 1020  
ccgcccgcgc cgcgcgccga gcgcgccacg gggcggaggg gcgtcctctt cgctgctog 1080  
caccggaogc tgctggacc catcttctg tccacggcgc tgggcogccc cgtcggccgc 1140  
gtcacctact cgctctccag gctctccogag ttctgtcgc ccatccggac ggtgcgctg 1200  
tcgcccggacc gcgccacgga cgcgccatg atcagggacc tgctggcggga gggagacctg 1260  
gtcatctgcc ccgaggggac cacgtgccgg gageccttcc tgctgcgctt ctcggectc 1320  
ttcgcogagc tcaccagcga ggtggtgccc gtgccatgg agaacaggat gagcatgttc 1380  
cacggcacca ccgccagggg gtggaaggg atggaccctt totacttctt catgaaccog 1440  
agcccggcct acgtcgtcac ctctctcaac aagctcccgc cggagctcac ctgcccggc 1500  
ggcaggacca gccacgaggt ggccaactac atccagagge tcatcggcgc cacgctctcc 1560  
tacgagtgca caagcctcac caggaaggac aagtaccggg cgcttgccagg caacgacgga 1620  
gtcgtcgcca ccccaaagcc gccgccaaog gttgctagct agctagctaa ggttattatt 1680  
attattattt tttttaaaaa aaaagagaga gaaagaaaat tgggttatta tattgtggtg 1740  
atgatgctac tgaattcatt gtgattgtta tgatgc 1776

<210> 130  
<211> 1744  
<212> ADN  
<213> *Ricinus communis*

ES 2 640 100 T3

<400> 130

catttataac attttcagag agagagagag agagagtaat ataaatatgc cggcgatgaa	60
acaggctcaa aagtttccgt caataacatc atgcagcggc acctcatacg aatcaatagc	120
agctgatctg gatggtactc tgctcgtctc cagtagctcc tttccttact tcatgctcgt	180
ggcctggaa gccggtagtc tgttgctggg tctagtggtg ctgctttcgc tgccttttat	240
catcatctct tatttcttca tatcagaagc catcgggac caaatcctaa tcttcatctc	300
tttgcgtggt ctcaaaatcc gtgacatcga gcttgtttcc cgcgctgttc tgcctagatt	360
ttatgctcgc gatgtgagga aagaaagcta cgaggtgttc gatagtgca agaggaaagt	420
ggtggtgacg gcgaatccga cgattatggt ggagccgttt gtgaaggatt ttcttgccgg	480
agacaaagt ttgggaacgg agattgaagt aaatcctaag acaaagaggg ccaactggtt	540
cgtgaagaag cctggcgtgt tggttggcaa atggaagaag ttggccattt tgaagagtt	600
tggagaggac gctcctgac ttgggatcgg agaccgcaag actgatcatg acttcatgtc	660
catttgcaag gagggctaca tggtttatca cagcaaaaaa gcagctactc cactgccacg	720
ggatcgcctg aaaagtctta taatcttoca cgatggcctt tttgttcaac gaccggacc	780
actcaatgcc ctgcctacct atctttggtt gccatttggg ttcattgctc ccatctttcg	840
ggtatacttc aatctaccac tccccgaacg catcgtacgg tacacatacg agatgtagg	900
catccacctt gtgatccgtg ggtaccacc tccagcgcca tcacgtggaa cccctggaaa	960
tctatatgtg tgcaaccacc gtacagcttt ggatcctatc gtgattgcaa tgcactcgg	1020
acgcaaagt tcgtgtgtaa catacagcgt tagccgtctc tcgagattcc tatcaccaat	1080
cccagccatc gctttgacac gtgatcgtgc ggcgtacgca gaacgcatca cagctctact	1140
ccaaaagggt gatctttagt tgtgtcctga gggaaactaca tgtcgcgagc agtttttgtt	1200
gagattcagt gcactgtttg ctgaaatgag cgataggatt gtgcctgtgg cggttaattg	1260
caagcaaagt atgttctatg gcaccactgt acgtggggtt aaattttggg acccttattt	1320
tttctttatg aatccaaggc caacttatga ggtgacattt cttgatcggg taccgaaga	1380
gatgacgggt aaagcgggag ggaaatcatc gattgaggtg gctaattatg tgcagaaggt	1440
tttgggtgat gtgttgggtt ttcagtgtac tggattaacg aggaaggata agtacttgtt	1500
gcttgagggg aatgatggca aggtggagtc tatgtacaac tccaaaagt aagcatttgg	1560
tttttcgttt ctataagatt tttagttacg gatgtttgta aggccagaca tatcaaatat	1620
tttagacatt gatgtcgtcg tgetgaattt gcgcacagct ttatatgcaa gactctttaa	1680
agaatttggg tgtccgattt catggacgtg actgggttaag atttcatgat gttataaggg	1740
aatt	1744

5

<210> 131  
 <211> 1539  
 <212> ADN  
 <213> *Ricinus communis*

10

ES 2 640 100 T3

<400> 131

atggcaaacc ctaaagaaac atatgctcct tttagctcat tcccaaatat agaaaagtgc	60
acatccgtag gtcgacaaaa cgacactgta gtgaccgata tggatggaac cctactttgt	120
ggacgtagcc cgtttcctta ctttgccctta gttgcctatg aagctagtgg aatattaagg	180
cttctattct tgctattagg ttgcctatt gctggaattc tccgtaactt catctccgaa	240
tctgctggaa tccgagtact tgtatttgca acctttgcag ggatgogagt gtctgatata	300
gaatcgggtg caagagctgt tttgcccaag ttttattcta gtgacttgca cccagagact	360
tggcgggtgt tctcttcgtg tgggaagaag tgcgtaatca cggaaaatcc taggattatg	420
gtagaggcat tcttgaaaga gtttgttggg gctgatatgg ttattggcac tgagattggt	480
gtgcataaag ggagagcgcac tgggtttatt aggagtcccg gtgttcttgt tggacagaac	540
aaggctgatg cgctaaaaat ggctttctat gatgcgcggg tgccaaatat tgggatcgggt	600
gacaggagga ctgatcgtcc attcatgaaa ttgtgcaagg agagctacat agtccaagc	660
aattctcaag tagagccagt tggtcctaaa aagttaccaa aaccocatagt tttccatgac	720
gggcggcttg tccaaaaacc aacacattta atcgcgcttg ctactcttct ttggattcca	780
gtaggcttta ttctgcctg tctacgctta gccgcgggtg cgctccttcc aatgccatta	840
gtgtactaag ccttgtaggc acttggtggt cgtattcaca taaaaggcac accaccgccg	900
ccagccaaga aatccactgg ccaagctgga gtocctcttcg tatgttccca tagaactctt	960
cttgaccogg tcttctctc catcgttcta ggacgtccaa tcccagcogt cacttaactcc	1020
ctttcgcggc tatcggagtt catttcaccc atcaagactg tcaggctcac tcgagaccgc	1080
gttactgatg caaacatgat caaggaacta ttacaagaag gtgacttagt catatgtcct	1140
gaagggacta cttgtaggga accatttttg cttcggtttt cagctttggt cgcggagtta	1200
accgatgaac ttgtccctgt cgccatggct aataagatga ctatgtttca tggacaaca	1260
gctagaggat ggaaagggat ggacccttc tatttcttga tgaacctag cctgcatat	1320
gaagtgacat tttgaaaca gttgccttac aatttaacat gcggagcagg aaaatctagc	1380
catagtgttg ctaattatat acaaaggacg attgctgcaa ctttgtctta ccaatgcacc	1440
agctttacta ggaaggataa atatagagca cttgctggca atgacggaac tgtcgtggag	1500
gagcccaagc aggctcgtga tgaagtcatg gattcctag	1539

5

<210> 132  
 <211> 1515  
 <212> ADN  
 <213> *Ricinus communis*

10

<400> 132

ES 2 640 100 T3

atggttatgg tggacagctt cccagaagtt agcaaatgtg cgtccaaagg ccgagaaaag 60  
cacacggtgg tcgctgatat ggatggcacc ctgtttattg gtcgtagctc cttcccttat 120  
ttcgctctaa ttgccttcga ggctggtggg attttttaggc ttcttttctt gctcttggca 180  
tcaccacttg oggcacttct ttattacttt gtttcagagt ctgctggtat tcaagttctt 240  
atctttgcat cctttgctgg tatgaaagta tcggatattg aatctgtagc acgcgcagtg 300  
ctgccaaagt tctactccag cgatctccat cctgagtcac ggcgtgtgtt ctcttcttgt 360  
gggaaacgct gcgttcttac agcgaacca agaattatgg tggaagcatt tctgaaggat 420  
ttcttgggag ctgatttagt ttttggtaact gaaatatcaa cttataaagg tagagctaca 480  
ggttttgttc gtaccocctgg tgtacttcta ggcaagaaca aggcaaagtc tcttaaaaag 540  
gcttttgag acgcaaggcc agaagttggg cttggagata ggcgcactga tgctcccttt 600  
atggctttgt gcaaggaggg atacctggtg ccaccaagc cggaagtacg ggcagtaaca 660  
ggtgacaagc tcccaaaacc catcgtcttc catgatggca ggctagtcca gaaaccaaca 720  
ccattaatgg cacttctcat cattctctgg ataccaatag gcttcattct agcttgcttg 780  
agaattgcag caggatcact cctgcctatg ccaatggtct actatgcctt cttagcactc 840  
gggtgocctg tcacagtaaa aggcaacca cctcctccag ccaagaaatc aactggtcaa 900  
tcaggtgtcc tctttgtttg ctcccacaga actctccttg accocatatt cctctccact 960  
gcccttgccc gcocatttgc cgcagtcaca tactcagtat ctgcctctc ggaaattata 1020  
tcacccatca aaacagttag gctcagccgc gatagagcaa cagatgcagc aatgatcaag 1080  
aagttgtag tagaaggtga cttagcgata tgtcctgaag ggaccacttg cagggaaacc 1140  
tttcttttaa ggttttcagc tttgttcgca gaattaaccg atcaaattgt gccggtagct 1200  
atggtgaaca gaatgagcat gttocatgga actaccgcta gaggggtggaa agggatggat 1260  
ccattttact tcttcatgaa ccctagtcca gcttatgaag taactttctt gagtaagttg 1320  
ccccagaat taacgtgcag ttcaggcaag tctagccacg aagtggctaa ttacatacaa 1380  
agggatgatt ctgctactct ttctatgaa tgcactggat ttactagaaa agataagtac 1440  
agagcacttg ctggcaacga tggcagtggt gttgagaaac ctaagctcga agccaacaaa 1500  
gtaatgggtt gctaa 1515

<210> 133  
<211> 1804  
<212> ADN  
<213> *Arabidopsis lyrata*  
<400> 133

5

ES 2 640 100 T3

tttgtgctct ctgttcttta gagcggccat taacgagaga gagagagaga gagagagact 60  
 tggcgagtta taatgtctoc cgcaagaag agtggaaact ttctccgat aacggaatgc 120  
 agagacggag agtacgattc gatagccgcc gatcttgacg ggactctgct tctctcaaga 180  
 agctcgttcc cttacttcat gctggtcgcc gttgaagctg gaagcctttt ccgtggacta 240  
 atactgcttc tctcattacc aatcgtcatt atttcttacc tattcgtatc cgaatctctt 300  
 ggaatccaga tctcatctt catctccttc gctggctca aaatccgga tatcgaactc 360  
 gtctctcgcg cagtcttcc acggttttat gcgccgatg tgaggaaaga cagttttgat 420  
 gtgtttgata agtgaagag gaaagtggta gtgacggcga atccgatagt gatggtggag 480  
 gcgtttgtga aggattatct tggagtgat aaagttttgg gaacagagat tgaagttaac 540  
 cctaaatoca atagagccac tggatttggt aagaagcctg gtgttcttgt tggatgattc 600  
 aagaggttag ccattttgaa agagtttggg gacgaatcac ctgatctcgg cctcggatg 660  
 cgaacctctg atcacgattt catgtccctc tgcaagaaag gtacatggt tcatgcgtgc 720  
 aagtcagcca caacaattcc aaaagaacgt ttaaagaacc gcatagtctt ccacgacggg 780  
 cgtttagcgc aacgtccaac gcccttaaac gccattgtca catacctatg gcttcccttt 840  
 ggtttcatac tctccatcat tcgctctac ttcaacctcc ctttaacctga acgctttgtc 900  
 cgttacacct acgagattct cggaattcac ttaaccattc gtggccatcg tcctccacct 960  
 ccttcccccg gaactcttgg caacctctat gtccttaacc accgtaccgc gcttgatccc 1020  
 atcatcgttg ctattgctct tggacgcaag atctgttctg ttacttacag tgtctctcgt 1080  
 ctctccctta tgctttctcc tattectgct gttgcctca cccgtgaccg tgtcaccgat 1140  
 gctgccaaca tgagaaaact tctcgagaaa ggcgacttgg tgatatgtcc ggaaggcaca 1200  
 acgtgcagag aagagtatct actgagattt agcgtctat tctcagagct aagtgaccgg 1260  
 attgtgccag tagcgatgaa ctgtaaacaa ggaatgttca acgggaccac agtgaggggt 1320  
 gtgaagttct gggaccctta cttctttttc atgaacccaa gaccaagcta tgaagccact 1380  
 ttcttgatc gttgcctga agaatgact gtcaacgggtg gcggcaagac tcctatagag 1440  
 gtggccaatt acgtccagaa agttatcggc gcggctcttg gcttcgaatg caccgaactt 1500  
 actcgcaagg ataaatatct tttgcttggg ggtaatgacg gcaaggtgga gtctatcatg 1560  
 ttgggataaa cttgaaataa ctttaagagat aagttactaa aatgtaaaag agataaagct 1620  
 gaatattcta ggagttatct agaataaatg tattaggagt ttttaattgat gatgtattgg 1680  
 taogttattg ttatgttaat tctatttttc cctaaaacta ttttgtgatc acatttctat 1740  
 acgcatgtgc ttaagggata ataagaactt tatatttaaa gagcaaaatt ttattggagg 1800  
 aagt 1804

5 <210> 134  
 <211> 1746  
 <212> ADN

ES 2 640 100 T3

<213> *Arabidopsis lyrata*

<400> 134

```

atatccattc tcacctcca aatctctctc ttctcttctc tttcoggctt cttcactaat      60
gggagotcag gagaaacgac gtcgtttoga gcctatatca aagtgogacg ttaaggacog      120
gtccaacat  accgtggcog ctgatctaga cgggacacta ctgatctcac gtagcgcctt      180
ccottactat ttctctgtag ccctcagaggc agggggcttg ctccgagcat tgatcctact      240
cgtgtccgta ccgttctct atcttaogta cttgaccgtc tcggagactt tggccatcaa      300
cgtttttgtc ttcacacgt ttgcgggtct caagatccga gacgttgagc tagtgggccg      360
ttccgtctc  ccgaggttct atgcccagga cgtgaggccc gatacgtggc gtatctttaa      420
cacgtttggg aaacggtaca tagtaacggc gagtctcoga atcatggtcg agccgttcgt      480
gaagacatac ctaggagttg ataaagttct cggaacagaa ctagaggtct ccaaatcggg      540
tcgggcaacc gggttcacca gaaaaccogg tattctcgtg ggtcagcaca aacgtgaagt      600
cgttttgaga gagtttggtg gccttgcgtc tgatttacct gatttggggc tcggcgatag      660
caagacagac tacgacttca tgtccatctg caaggaaggt tacatggtgc cacgtacgaa      720
atgcgaaccc ttaccaagaa acaaactctt aagccccata atattccacg agggcagatt      780
agtccaacgc ccaacgcogt tagttgctct gttaactctc ctttggcttc ccatcggtt      840
cctctctct  cttatccggg totacacgaa tattccgta ccggaacgta tcgcccgta      900
caactacaag ctactggca tcaagtaat cgtcaacggt caccctctc cgccgcaaaa      960
acctggccaa ccaggccatc tcttggtttg taaccaccgc accgtccttg acccggtggt      1020
gacagccgtc gcgctcggcc ggaaaatcag ctgcgtcact tatagcatta gcaagttctc      1080
tgagctaatc tcaccaatca aagccggtgc gttgactcgt caacgtgaga aagacgcagc      1140
gaacatcaag cgcttttg aagaaggoga tcttgtgatc tgtccgagg gaactacgtg      1200
ccgtgagcct ttccttctc ggtttagtgc cctcttccgt gagctcacgg acccgattgt      1260
tcccgtagcg atcaacacga agcagagcat gttcaatggt actaccacac gtggctacaa      1320
gcttcttgat ccttacttct cgttcatgaa cccgaggccg acgtacgaga tcacgttctc      1380
caaacagatt cccgctgagc tgacgtgtaa aggaggcaaa tcgccgatag aggtggcgaa      1440
ttacatacaa agggttttg gaggaacctt aggttttgag tgcaccaatt tcacaagaaa      1500
ggataagtac gcaatgcttg ctggtaccga cggtagggtt ccggtgaaga aggagaagac      1560
gtgattattg atcggagaat ggagaaaacg aagaaataat acaaagattg gaatctttt      1620
tcatttggat ggttaattat caatatttaa ggggcaaaga attgataagt ttatttttt      1680
gttgtttctt ttatcttctc tgcttccatg aatgtttcaa gatttacttt tagtacaata      1740
acggtt                                           1746

```

ES 2 640 100 T3

<210> 135  
 <211> 1509  
 <212> ADN  
 <213> *Vernicia fordii*

5

<400> 135

```

atgtcgcaaa cgaaaccggc tccaaaattc ccttcaataa catcatgtac cggtcggcg      60
tatcagtcaa tagcggctga tcttgatggc acgctcctcg tgtccagtag ctccctccct      120
tacttcatgt tagtggctgt cgaagctggc agtctcctcc gtggtctggt gtcgcttctg      180
tctctacctc ttgtgatcat ctccatattt ttcatatctg aagccgtcgg aatccaaatc      240
ctaactata tctcctttgc tgggctcaag atccgcgaca tcgagctggc tccccgct      300
gttttgcta ggttttatgc tgctgatgtg aggaaggaga gttttgaggt ttttgacaaa      360
tgcaagagga aggtgggtgg gacggcaaat ccgacgatta tgggtggagcc atttgtgaag      420
gattttcttg gcggagacaa ggtgttaggt acggagattg aagtaaacc aaagactaag      480
agggccaccg gattcgtgaa gaatcctggc gtgttagtgg gaaatggaa gaaattgtcc      540
atthtgaaag agtttgaga ggaatcaccg gatcttggga tcggagaccg caagacggat      600
catgatttca tgtccatttg caaggagggc tacatggttc aacgcagcaa atcagctact      660
ccaataccac tggatcgtct caaaagtcgt ataattttcc atgatggccg gttcgttcag      720
cgacctgacc cactgaacgc tcttgtcacc tatctttggc tgccatttgg attcatcctc      780
tctattattc gcgataactt caatctccca ctcccagaac gcatcgtacg gtacacatac      840
gagatgctag gcatccacct tgtgatccgt gggaaccgcg ctccctgccc atcacctgga      900
acccagaa atctctatgt ttgcaaccac cgctcagctt tggatcctat tgtcattgca      960
atagcacttg gacgcaaagt ttctgtgta acatacagcg taagccgtct ctgaggttc      1020
ctatctccga ttccagctat cgctttgact cgtgatcgtg ccgctgatgc agcacgcatc      1080
tcagagctac tccaaaaggg tgatcttgta gtgtgtccag aggggactac gtgtcgcgag      1140
ccgttcctgt tgcgattcag tgcattgttc gcagaaatga gcgataggat tgtgcgggtg      1200
gcggttaact gtaagcaaaa tatgttctat ggtaccaccg tacgtggtgt caagtthtg      1260
gatccttatt acttcttcat gaaccaaga ccaacctatg aggtgacatt ccttgatcgg      1320
ttgccggaag agatgacagc taaggcggga gggaaatcgt cgatcgaggt ggctaattat      1380
gtgcagaagg tgttggtgta tgtgttgga ttcgaatgca ctggattaac tagaaaggat      1440
aagtatatgt tgctaggagg aatgatggc aaggtggagt caatgtacaa caccaaaaaa      1500
    
```

gccggttaa

1509

10

<210> 136  
 <211> 1545  
 <212> ADN  
 <213> *Oryza sativa*

ES 2 640 100 T3

<400> 136

atgggggacg	ccacgatgat	cgcgtcgccg	ttcccggcgg	tggagaagtg	ttcgtccaag	60
gaccgcggcg	gcgacacgat	cgtcgcccgc	ctcgacggca	cgctgctgtg	eggccggagc	120
tcgttcccg	acttcgcgca	catggcggtc	gagacggggg	gcgtcctccg	gctgctggcg	180
ctgatcctgc	tggcgccgct	cgccggcggtg	ctctactacc	tgggtgtccga	gtcggccggc	240
atccagggtgc	tcatcttcgc	gtccatggcc	ggcgcgaggg	tcgccgacgt	tgaggcggtg	300
gcgcgcccgc	tgtctcccaa	gttctactgc	tccgacatcc	accgggagtc	gtggcgcggtg	360
ttctcggcgt	gcgggcgccc	gtgcgtgctc	accgcgaacc	cgaggatgat	ggtggaggcg	420
ttcttgaagg	agtacctcgg	cgccgacatc	gtcgtgggca	cggagctcgt	ggtctggcgt	480
ggccgcgccca	cggggtcgtg	ccggagcccc	ggcgtgctcg	tcggcgagaa	caaggcgccc	540
gcgctccggg	aggcgttcgg	cgacgcctcg	ccggagatcg	gcatcggcga	ccggaggacg	600
gactaccctg	tcatgaggct	atgcaaggag	gggtacgtgg	tgccgggtgc	gccgacgccg	660
aggcccggtgc	cgcgggagga	gctcccgaag	ccgggtggtg	tcacgacgg	ccgcctggtc	720
cagaagccgt	cgccggcgct	cgcgctgctc	accgtgctct	ggatccccat	cggttctgtg	780
ctcgccctgcc	tcgcatcgc	cgccggcgcg	ctcctcccca	tgcgctggtg	gtaccacgcc	840
ttccgcgcc	tggcgtgcg	cgtcaccgtc	aagggcacc	cgccgccacc	ggccagccgc	900
gagaccggcc	agtcggcgtg	gctcttcatc	tgctccacc	gcacctcct	cgaccccatc	960
ttctctcca	cgccctcgg	cgccccatc	accgccgca	cctactcgg	gtcgcggctg	1020
tcggagatat	tgtcgcgat	caggacggtg	cggtgacgc	gggacaggg	ggcggacgg	1080
gcgatgatcc	ggcgctgct	ggcggagggc	gacctggtga	tctgcccgga	ggggacgacg	1140
tgccgggagc	cgttcctcct	ccggttctcg	gcgctgttcg	oggagctcac	cgacgagatc	1200
gtgccgggtg	cgatggagaa	ccagatgagc	atgttcacgc	gcacgacggc	gagggggtgg	1260
aaggcgctgg	acccttcta	cttcttcatg	aacctagcc	ccggatacgt	ggtcaccttc	1320
ctcagcaagc	tccccggcga	gctcacctgc	aacggcggtg	gcggcggcgg	cgccggccgg	1380
agcagccacg	aggtggccaa	ctacatccag	aggctcatcg	cctccacgct	ctcctacgag	1440
tgcaccagct	tcaccaggaa	ggacaagtac	agggcgctcg	ccggcaacga	cgccaccgtc	1500
gtgtccaagc	ccaacatoga	cgagaagaag	gccatgggat	gttaa		1545

5

<210> 137  
 <211> 1856  
 <212> ADN  
 <213> *Arabidopsis thaliana*

10

<400> 137



ES 2 640 100 T3

aaggaccttc cctctcttcc tttccaaatc tctctttatc tctctttctc tgtttactaa 60  
atcgccagta ttcgccgcaa gaaagatgtc tccggcgaag aagagccgaa gttttcctcc 120  
gataagcgag tgtaaaagca gagagtatga ttgatogct ggggatctcg acgggactct 180  
gcttctgtca agaagctctt tcccttactt catgctgggt gccattgaag caggtagtct 240  
cttccgtgga ttgatccttc ttctttctct tcccatcgtc atcategctt atctcttctg 300  
ctccgaatct ctcggaatcc aaatccttat ctccatctcc ttcgccgta tcaagatcaa 360  
gaacatcgaa ctgctctctc ggcgcttct tacacggtt tacgcggcgg atgtgaggaa 420  
agatagtttt gaggtgttg ataatgtaa gaagaggaaa gtagtgtgga ctgcgaatcc 480  
gatagtgatg gttgagccat tcgtcaaaga ctatttggga ggagataaag ttttgggaac 540  
agagattgaa gtttaaccca aaacgatgaa agctactggt tttgtgaaa agcctggtgt 600  
tcttgttggg gatctcaaga gattagccat ctgaaagag ttggcgatg actcaccgga 660  
tcttggcctc ggtgaccgca cctccgatca cgatttcattg tccatttga aggaaggta 720  
catggtgcat gagaccaa atcagccacaac agtccctata gagagtctca agaaccgcat 780  
aatcttccat gatggcgtc ttgtccaacg tccgacccca ttaaacgcct taatcattta 840  
cctttggctt cctttcggct tcatgctctc tgtcttccgc gtctacttca acctcccttt 900  
acctgaacgc ttcgctcgt acacttacga gatcctcggc attcacctca caatccgtgg 960  
ccaccgtcct ccacctcct cccccgaaa acctggaaac ctctacgttc tcaaccaccg 1020  
caccgcctt gacccatta tcatcgccat tgctctcggc cgtaagatca catgtgtcac 1080  
ttatagtgtc tctcgctct cctgatgct ctacccatt cctgctgttg ctctgaccgg 1140  
agaccgtgtc gctgacgcag cccgatgag gcaactctc gagaaagtg atttggtgat 1200  
ctgtcctgaa ggcacaacgt gtagagaacc atatctacta aggttcagt ctctatttgc 1260  
agagctgagc gaccggatcg tgcctgtggc catgaactgc aagcaaggga tgttcaacgg 1320  
gacgacagtt agaggtgtga aattctggga tccgtacttc ttcttcatga acccccgacc 1380  
tagctacgag gccactttct tggaccgtt gccgaagaa atgacggtaa acgggtgtgg 1440  
caagactcct tttgaggtgg ctaattacgt tcagaagtg atcgggtggag ttttggggtt 1500  
cgaatgtact gaacttaca ggaaggataa gtatctgttg cttggaggca acgacggaaa 1560  
ggttgagtct atcaataaga ccaagtccat ggagtaataa tocagctagt ctggttctg 1620  
ctattttctg tgggagtgaa gttaggtctt tccgggtttc ccaattactt tttttttgt 1680  
ttttttcat gtattgtctg aggaaaatgt tttgttgaat atgcactgga taatgaaact 1740  
acatttgaat caaagtacgt acatgtagt taccatttcc ccttgcctta actatgaatg 1800  
atgtgaaac tcacctgta ttataagata toatagagaa aaacacattc togact 1856

- 5 <210> 138
- <211> 1503
- <212> ADN
- <213> *Populus trichocarpa*

ES 2 640 100 T3

<400> 138

```

atgtcgccaa cgaaacccgc aaaaaaattt ccgccgataa cggcctgtaa cggcacgacg      60
caccaatota tagcagcggg tcttgatggc acactccttg tgtcgagtag ctcgtttcca      120
tactttatga ttgttgctgt tgaagctgga agcctctttc gtggtcttgt tttacttctt      180
tcaactgcta tcgtcatcgt gtcttacctc ttcatatctg aagccttggg gatacaaatg      240
ttaatcttca tttctatgtc tgggctcaag atccgtgata ttgaacttgt ttcacgcgct      300
gtattgecta ggttttatgc tgcagatggt agatcggaga gttttgaggt gtttgatagg      360
tgcaagagaa aggtggtggt gacagcaaat ccaactataa tggttgagcc ttttgtgaag      420
gattttctcg gtggagataa ggttttgggc acggagattg aagtgaacco gaagaccaag      480
agggctacag gattgttaa gaagcctggt gttctggtgg gcaagtggaa agaattggct      540
gttttaaagg agtttgaga ggaagcaccg gatcttggtg ttggtgaccg caagactgat      600
catgatttca tgtccctttg caaggagggc tacatggtgc atcgacgcaa atcagcgcact      660
ccactaccgc gggatcgcct caaaaatcgt atcatcttcc atgatggcog tcttgttcaa      720
cgtccagacc cactcaatgc cttgatcact tacatttggg tgcggtttgg attcatcctc      780
tocatcattc gcgtctactt caacctcca ctcccagaac gcatcgtaac gtacacatac      840
gaaatgcttg gcattcacct ggtcattcgt gggacccac ctctgcccc atcacctggg      900
accccaggaa atctctatgt ttgcaatcac cgtacagccc tggatccaat tgtcatagcc      960
atagcaacttg gacgcaaagt ctcgtgtggt acgtacagcg taagccgtct ctcgaggttt     1020
ctttcaccaa tcccagccat cgctttgact cgtgatcgcg cggtgatgc tgcaagaata     1080
tcatccatac ttcaaaaggg tgatctagtg gtgtgtccgg agggaaccac ttgtcgtgag     1140
gagttcttgc tgcgatttag tgcattgttt gcagagttga gtgataggat tgtgcccggt     1200
gcagtcaatt gtaaacaaaa catgttttat ggcaocgacc tgcgtggtgt caagttctgg     1260
gatccttatt ttttctcat gaacccgagg ccaacttacg aagtacatt ccttgatcgc     1320
ttgccagagg aatgacggt caaggctggt ggcaaatctt cgatcgaggt ggcaactat     1380
gtgcagaaag tgttggtgga agtgtttaga ttcgagaaca ctgggttgac caggaaggac     1440
aagtacctgt tgcttgggtg gaatgatggt aaggctcaggt caatgcacaa tgccaaaaaa     1500
taa

```

5 <210> 139  
 <211> 2117  
 <212> ADN  
 <213> *Sorghum bicolor*

10 <400> 139

ES 2 640 100 T3

gcogagctcg actcgggaatc cgatocccca cccaactcca tgctccgatt cctoctcoct 60  
 tctctgctcc gactccgacc accaccacct ctgctgccag ccccgtocta catataaatc 120  
 cctccgtcat catcgtcaac cgcagcgttg ggcgctcaa caaacgcagc caccgcgcgc 180  
 gccgccacca cgacaccaac catataatcg cctgcccaag tcccaacctt tcccgacca 240  
 gccagacacc accaccaccg togtcttctt cgccatggcc ccccgggcga gcgcgaggtt 300  
 cccaccggtg tcatcctacg acgcgtccgc acgcaaccgc cgaaccgccg ccgcggaact 360  
 ggacggcacg ctgctcgggt cctcctoggc gttccoctac tacttctctg tggcgtcga 420  
 ggcggtctcg tgccctccgc cggcggcgtt cctcctcgtg gctccgctcc tgctagcgtt 480  
 ctacacgctg gtgtccgagg cggcggccat cgcgctgctg gcgttcgtca cgttcgcggt 540  
 gctgcgcgtg cgcgacgtcg aggcctcgc ccgcggcgtc ctgccgcggc actacgcggc 600  
 cggcgtgcgc gccgacacct ggaccgtgtt ccgcggctgc ggcgagggga ggagggtcgt 660  
 cgtcacgcgc tccccgcgcg tcatggtggg cgagttcgtc cgcgagttcc ttggtgctga 720  
 ggtcgtggc accgagctcg agacctctc cgcgttcggc gccgcaggt ttacggggag 780  
 gatcaaggcc gtgctcgtcg gggagaggaa ggcggaggtc gtcggcggcg tcttcgcccg 840  
 tggggagatg cccgacgtcg ggcttgggga ccgcgagagt gaccatgact tcatggccat 900  
 ctgcaaggaa gcttacatgg tgccctccga caagcgcgcg gcgcgtgcgg ccgccgacac 960  
 gctgctgtcc cgtccgtct tccacgacgg ccgcctcgtc cagcgcgccg acccggcgca 1020  
 ggcgctgttc gcgctggcct acctccgct gggcttctg ctggcgtgt tccgctctt 1080  
 cttcaacctc atgatgccg tacgcctggt ccgccacacg taccgcctga cgggatccg 1140  
 cctccggtc cggggcacgc cgcgcgcgc cccggcgccg ggcgccccg gttcgtctgt 1200  
 ggtctgcaac caccgcacgg cgtgggacct gatcctctg tccgtggcgc tgggtcgccc 1260  
 cgtgtcgtgc gtcacgtaca gcgcgagccg cctctcgacg gcgatctcgc cgatccgggc 1320  
 cgtcgcgctg tccccggacc gcgccaccga cgcggcgcgc atggcggcgc tgcctggcga 1380  
 gggcgaagtc gtcgtgtgcc ccgaggggac cacgtgccgg gagcctgccc tgcctccggtt 1440  
 ctggcgtctg ttcgcgagc tcacggaccg gatcgtgccc gtggccatgg aggcgcggca 1500  
 gggcacctac tacgggtcga cggcgagggg gtggaagtgg ctgcacccct acttcttcta 1560  
 catgaaccgc cggccaggct acgacgtcac gttcctgccg ccgctgcggc cggaggagac 1620  
 gtgcggcgc ccggggagga gcgccgtcga cgtggctaac cacgtgcaga gggctatcgc 1680

ES 2 640 100 T3

```
caaggagctc gggttccagt gcaccacgct caccaggaag gacaagtaca tgaagctcgc 1740
cggcaacgac ggctcggtcg ccgccagggc taaaaggac gccgccgatg ataacgctgc 1800
cacggtcacc accaccaaga agtttgtatg aagagatoga ttatatatat tggcaaccta 1860
gctogactat atcggagcag cattgtactg acgtaaaaaa aaaaaaaca tgcattgcaca 1920
aaaggtctag ttaggcttat gtatattgga gtatcgtggt cttttattta tactacgtat 1980
acgtagaaaag aatggacgag tgaatggagt ccaattttaa aggcgcaatt tatatatgat 2040
tatgattatt ttgggtagtg tatttgctta aaaatctaaa cacatactaa attgcattgc 2100
aacaatttaa atttttt 2117
```

5

```
<210> 140
<211> 1547
<212> ADN
<213> Ricinus communis

<400> 140
```

ES 2 640 100 T3

acacatatatt cccactgcta ctagcttccc aaaactcaca octtaatcta tcaaattggt 60  
 catgggggct catcggcatt tcgagtcaat caagaaatgc agcaccgagg gacgggtcgaa 120  
 tcagacggtg gctgcagacc tggacggcac acttcttgtt tcatcaagtg cttttccata 180  
 cttcttgctc gttgcaactg aagctggaag cgtgttaaga gctatactcc ttttagcatc 240  
 tgttcctttc gtgtatttta cgtacctggt tttatcagag goaatagcta tcaaaacctt 300  
 catctttatt gcctttgcag gacttaaaat aaggagcgtt gagcttgttt cggggtctgt 360  
 cttgcctaga ttctacgagg atgatgtcca tcttgaaact tggagagtgt tcaattcttt 420  
 tgggaaaaga tatattgtca ccgccagtc taggattatg gttgaacctt ttgttaagac 480  
 ctttttaggt gctgacaaag ttcttggcac tgaactagag gttacaaaat cgggaaaggc 540  
 aactgggttc gtcaagaaac ctggaattct tgttgctgaa cataaaagag atgccatttt 600  
 aaaagaactg ggogctgatt tgccagattt aggcctgtgg agaccgagaa actgaccaca 660  
 atttcatgtc tatttgcaag gaaggatata tgggtgccaag aacaaagtgt gagccactac 720  
 caagaacaa gcttctaagc cctgtaatat tccatgaagg acggttagtt caaaggccaa 780  
 cacctctagt tgccctcttg acctttttat ggatgccaat cgggtataatt ctctcaatcc 840  
 ttagggata ccttaatatc cctttacctg aaagaattgc ctattacaat tacaaaattc 900  
 ttggaattaa agttatagtt aagggcaccc ctctctctgc tocaagaaaa ggccaaagtg 960  
 gagtctttt tgtatgcaac catcgtactg ttttagaccc agtagtgaca gcagttgcct 1020  
 taggcagaaa gatcagttgt gttacatata gcataagcaa attcacagaa atcatatcgc 1080  
 caatcaaagc agtcgcatta tcaagagaaa gagagaaaaga tgctgctaata attaagcgtt 1140  
 tattagaaga aggtgacttg gtcatttgcc ctgaaggaac aacttgtaga gagccatttc 1200  
 ttttgagatt tagtgccctt tttgctgaac ttaccgacag aattgtgcct gttgctatca 1260  
 ataccaagca aacagttttc cacggcacga ctgtccgagg gcataaatta ttggaccctt 1320  
 attttggtt catgaatcca atgccaacct atgagatcac attcttgaat cagttgccta 1380  
 tagagcttac ttgcaaaggt ggtaaactgt cgatcgaagt tgccaattac attcagaggg 1440  
 ttcttgctgg tacattgggg tttgaatgta ctaatttgac aaggaaggac aagtatgcta 1500  
 tccttgctgg tacagacggt cgtgttccat ctaagaagga aaagtaa 1547

<210> 141  
 <211> 1549  
 <212> ADN  
 <213> *Arabidopsis lyrata*

5

<400> 141

ES 2 640 100 T3

aagattcgcc gcaagaaaac atgtctccgg cgaagaagag ccgaagtttt cctccgataa 60  
gcgagtgtaa aagcagagag tatgattoga tcgctgcgga tctcgacggc actctgcttc 120  
tctcaagaag ctctttccct tacttcatgc tcgtcgccat tgaagcaggt agtctcttcc 180  
gtggattgat ccttcttctt tctcttccca tcgtcatcat cgcttatctc ttcgtctccg 240  
aatctctcgg aatccaaatc cttatctaca tctccttcgc cggtatcaaa atcaagaaca 300  
tcgaactcgt ctctcgcgcc gttcttccac ggttttacgc ggcggatgtg aggaaagata 360  
gttttgaggt gtttgataaa tgcaagagga aagtgggtgt tactgcgaat ccgatagtga 420  
tggttgagcc attcgtcaaa gattatttgg gaggagaaaa agttttggga acagagattg 480  
aagttaacc caaaacgatg aaagctactg gttttgtaa aaagcctggt gttcttgttg 540  
gtgatttgaa gagattagcc atcttgaaag agttcggcga agaatcaccg gatattggac 600  
tcggtgaccg gacctccgat caagatttca tgtccatttg caaggaaggt tacatggtgc 660  
atgagaccaa atcagccaca acagtcctta tagagcgtct caagaaccgc ataatcttcc 720  
acgatggccg tcttgtccaa cgtccgacgc cgttgaacgc cttaatcatt tacctttggc 780  
ttcctttcgg ctttatgctc totatcttcc gcgtctactt caacctccct ttaccogaac 840  
gcttcggtcg ctacacttac gagatcctcg gcattcacct cacaatccgt ggccaccgtc 900  
ctccccctcc ttccccggc aaacctggaa acctttacgt ccttaaccac cgaaccgccc 960  
ttgaccccat aattatcgcc attgctctcg gccgtaagat cacatgtgtc acttatagtg 1020  
tctctcgccct ctccctgatg ctctcaccca ttctgctgtg tgctctgacc cgagaccgtg 1080  
tcgctgatgc agcccgcatg agacaactcc tcgagaaagg tgatttgggt atctgtcctg 1140  
aaggcacaac gtgtagagaa coatatctac taagattcag tgctctattt gcggagctga 1200  
gcgaccggat cgtgcctgtg gccatgaact gcaagcaagg gatgttcaac gggacgacag 1260  
ttagagggtg gaaattctgg gatccgtact tcttcttcat gaacctcga cctagctacg 1320  
aggccacttt cttggaccgt ttgcccgaag aaatgacggt aaacgggtgt ggcaagactc 1380  
cttttgaggt ggctaattac gttcagaagg tgatcgggtg agttttggg ttcgaatgca 1440  
ctgaacttac aaggaaggat aagtatctgt tgcttgagg caacgacgga aaggttgagt 1500  
ctatcaataa gaccaagtcc atggactaat aatctagcta gttactggt 1549

<210> 142  
<211> 1995  
<212> ADN  
<213> *Arabidopsis thaliana*  
<400> 142

5

ES 2 640 100 T3

ccaacacatt tacttttctt cttcttctct ctctacgcca tagctatggt tttaccagag 60  
 cttctogtga ttctagcaga atgggttctt taccgtcttc ttgctaaatc atggttacaga 120  
 gctgctagaa agctcagagg ctatggttcc cagttaaaaa atctcctaag cctctccaaa 180  
 acacaatctc tacacaacaa ttctcagcat catcttcata accatcatca acaaaatcac 240  
 cctaaccaaa cccttcaaga ttctctagac cctctgttcc caagccttac aaaataccaa 300  
 gaacttcttc tagacaaaaa cagagcatgc tctgtttcgt ctgatcatta tcgtgatacc 360  
 ttttctgog atatcgatgg tgttttatta agacaacatt ogagcaaaca cttccacacg 420  
 ttcttccott acttcatgct cgtggctttt gaaggaggat cgatcatcag agcgatcctt 480  
 ctcttgctct cctgttcttt cctttggact ctgcaacaag aaaccaaact cagagttctc 540  
 tctttcataa ccttctctgg tcttagagtc aaagacatgg acaatgttcc tagatcggtt 600  
 ctacaaaaat tcttcctoga gaacttgaat atccaagttt acgacatctg ggctcgaact 660  
 gagtactcca aagttgtggt cagcagcttg ccgcaagtgc tgggtggagag gttcctcaga 720  
 gaacatctca acgctgatga tgtgattggg actaaactgc aagaaataaa agtcatggga 780  
 aggaaattct acacaggcct agcttctggt tctggtttcg tcttgaaca taaatcagct 840  
 gaagactact tttttgatag caagaagaag cctgcacttg gaattggaag ctottcaagc 900  
 ccacaagatc acatcttcat ctccatttgc aaagaagctt atttttgaa tgaagaagag 960  
 tccatgagca agaacaacgc attgcctaga gaaagatacc caaagccact gatattccac 1020  
 gacgggagat tggctttctt gccaacacca ttagctacat tagccatggt catatggctc 1080  
 cctatcggat tttactcgc tgtcttcagg atctcggtag gagtctttct tcttaccat 1140  
 gtagccaatt tcttggttc tatgtccggt gttagatca ctttcaagac ccataatttg 1200  
 aataacggtc gtccagagaa aggtaatagt ggagtcttgt atgtatgcaa ccataggacc 1260  
 ctctcagacc cagtttctc caccacttcg ctgggcaagc ccttgaccgc ggttacttac 1320  
 agcctaagca agttctccga gttcatagca cctctcaaga cggttagcct taaaagagat 1380  
 aggaagaaag acgggtgagc aatgcagaga ctattaagca aaggagatct agtggtttgt 1440  
 ccggagggaa ccacatgtag agagccttat ctactaaggt tcagtccttt gttcgtgag 1500  
 ttaactgaag acattgttcc cgtggctgta gatgctcgag tgagcatggt ctatgggaca 1560  
 actgcgagcg ggctcaagtg ttagaccct atcttcttcc tcatgaacce taggccggtt 1620  
 tattgtcttg agattttaa gaagttgcca aaagagatga cttgtgcagg aggcaagtca 1680  
 agctttgaag ttgcaattt catacagga gaacttgcta gggttttggg atttgaatgc 1740  
 actaatctca caaggagga taagtacttg gttcttgctg gaaatgaagg cattgtcaga 1800  
 taaactatc ogaatttca ttagttttt tgtttgttg aaaagctaaa tattattgtt 1860  
 atttcaggac ttagatggtt aattctatat ggttgattca ttggcatcat ttctatttct 1920  
 tctactcgg gaccattgat atgtattgaa aatgtttctt ttatatgatt ttgtttctgg 1980  
 atttgtatag tccat 1995

<210> 143  
<211> 1828  
<212> ADN  
<213> *Arabidopsis thaliana*

5

<400> 143



ES 2 640 100 T3

gtctgaaacc aaacaaagaa aatacacgaa ggaaggaaaa gaaaagtaa aagaaaagct 60  
 ttctgcaaaa tccaaacatg tccgccaaga tttcaatatt ccaagctctt gtctttctat 120  
 tctaccggtt tatcctccgg cgatatcgga actctaaacc aaaataccaa aatggccctt 180  
 cttctctcct ccaatccgac ctatcacgcc acacattgat cttcaacgta gaaggagctc 240  
 ttctcaaatc cgactctctc ttoccttact tcatgttagt agcatttgag gggggaggcg 300  
 taataaggtc atttctcctc ttcattctct atccattgat aagcttgatg agccatgaga 360  
 tgggtgtcaa agtgatggta atggtgagct tcttcgggat caaaaaagaa ggttttcgag 420  
 cggggagagc ggttttgcct aaatactttc tagaagatgt cggactcgag atgttcgaag 480  
 tgttgaagag aggagggag aaaatcggag tgagtgatga tcttcctcaa gttatgatcg 540  
 aagggttctt gagagattac ttggagattg acgttgtggt cgggagagaa atgaaagtcg 600  
 ttggaggtta ttatctaggt atcatggagg ataaaaccaa acatgatctt gtctttgatg 660  
 agttagttcg taaagagaga ctaaacaccg gtcgtgttat tggcatcact tccttcaata 720  
 catctcttca ccgatatcta ttctctcagt tttgccagga aatttatttc gtgaagaaat 780  
 cagacaagcg aagctggcaa accctaccac gaagccagta ccctaaacca ttgattttcc 840  
 atgatggccg tctcgcgatc aaaccaacce taatgaacac tttggtcttg ttcatgtggg 900  
 gtcctttcgc agccgcagcc gcagcagcca gactcttcgt ctctctttgc atcccttact 960  
 ctttatcaat ccgatacctc gccttttccg gttgcagact aaccgtcaact aacgactacg 1020  
 tttcatctca aaaacaaaaa ccaagtcaac gcaaaggttg tctctttggt tgtaaccata 1080  
 ggactttatt ggaccctctc tatggttgcac tcgctttgag aaagaaaaac atcaaaactg 1140  
 taacgtatag tttgagtagg gtatctgaga ttttggtcc gatcaagacg gtgagactga 1200  
 cccgtgatcg ggtgagcgac ggtcaagcca tggagaaatt gttaaccgaa ggagatctcg 1260  
 ttgtttgtcc tgaaggaacc acttgtagag aacottaact gcttaggttt agccctttgt 1320  
 tcaccgaggt tagtgatgtc atcgttcccg tggotgtgac ggtacacgtg accttcttct 1380  
 acggtacaac ggcgagtggc ctttaaggcac ttgaccogct tttcttctc ttggatcctt 1440  
 atcctaccta caccatcaa tttctcgacc ctgtctccgg tggcacgtgc caagatcctg 1500  
 atggaaagtt gaagtttgag gtggccaaca atgttcagag tgatattggg aaggcgttg 1560  
 atttcgagtg cacaagtctc actagaaaag acaagtattt gatcttggcc ggtaataatg 1620  
 gagtagttaa gaaaaattaa aatttgttt aggatgatga gtgagtgatg agttcggttc 1680  
 tttgggtatg taagtaacca atcgattatt tcatttgttt attaataatg tctacgtaaa 1740  
 tgtaactaca atcttttttt cttggtaatg tacacatag ttctatatca tatccaaaat 1800  
 ctttgtcaga caattttgcc tttcatgc 1828

<210> 144  
 <211> 503  
 <212> PRT

ES 2 640 100 T3

<213> *Arabidopsis thaliana*

<400> 144

```

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys
1          5          10          15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
20          25          30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu
35          40          45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Ile
50          55          60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
65          70          75

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

Val Ser Arg Ala Val Leu Thr Arg Phe Tyr Ala Ala Asp Val Arg Lys
100         105         110

```

5

ES 2 640 100 T3

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Lys Arg Lys Val Val Val  
 115 120 125  
 Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu  
 130 135 140  
 Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr  
 145 150 155 160  
 Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp  
 165 170 175  
 Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Asp Ser Pro Asp  
 180 185 190  
 Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys  
 195 200 205  
 Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro  
 210 215 220  
 Ile Glu Ser Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val  
 225 230 235 240  
 Gln Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro  
 245 250 255  
 Phe Gly Phe Met Leu Ser Val Phe Arg Val Tyr Phe Asn Leu Pro Leu  
 260 265 270  
 Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu  
 275 280 285  
 Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly  
 290 295 300  
 Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile  
 305 310 315 320  
 Ala Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser  
 325 330 335  
 Arg Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg  
 340 345 350  
 Asp Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly  
 355 360 365

ES 2 640 100 T3

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu  
 370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro  
 385 390 395 400

Val Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg  
 405 410 415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro  
 420 425 430

Ser Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val  
 435 440 445

Asn Gly Gly Gly Lys Thr Pro Phe Glu Val Ala Asn Tyr Val Gln Lys  
 450 455 460

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys  
 465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile  
 485 490 495

Asn Lys Thr Lys Ser Met Glu  
 500

<210> 145  
 <211> 501  
 <212> PRT  
 <213> *Arabidopsis thaliana*  
 <400> 145

5

Met Gly Ala Gln Glu Lys Arg Arg Arg Phe Glu Gln Ile Ser Lys Cys  
 1 5 10 15

Asp Val Lys Asp Arg Ser Asn His Thr Val Ala Ala Asp Leu Asp Gly  
 20 25 30

Thr Leu Leu Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala  
 35 40 45

Leu Glu Ala Gly Ser Leu Leu Arg Ala Leu Ile Leu Leu Val Ser Val  
 50 55 60

Pro Phe Val Tyr Leu Thr Tyr Leu Thr Ile Ser Glu Thr Leu Ala Ile  
 65 70 75 80

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Ile Ser Lys Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu  
 340 345 350

Thr Arg Gln Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu  
 355 360 365

Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 370 375 380

Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile  
 385 390 395 400

Val Pro Val Ala Ile Asn Thr Lys Gln Ser Met Phe Asn Gly Thr Thr  
 405 410 415

Thr Arg Gly Tyr Lys Leu Leu Asp Pro Tyr Phe Ala Phe Met Asn Pro  
 420 425 430

Arg Pro Thr Tyr Glu Ile Thr Phe Leu Lys Gln Ile Pro Ala Glu Leu  
 435 440 445

Thr Cys Lys Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln  
 450 455 460

Arg Val Leu Gly Gly Thr Leu Gly Phe Glu Cys Thr Asn Phe Thr Arg  
 465 470 475 480

Lys Asp Lys Tyr Ala Met Leu Ala Gly Thr Asp Gly Arg Val Pro Val  
 485 490 495

Lys Lys Glu Lys Thr  
 500

<210> 146  
 <211> 289  
 <212> PRT  
 <213> *Arabidopsis thaliana*  
 <400> 146

Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys Glu Arg Leu  
 1 5 10 15

Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln Arg Pro Thr  
 20 25 30

Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe Gly Phe Ile  
 35 40 45

5

10

ES 2 640 100 T3

Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro Glu Arg Phe  
50 55 60

Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr Ile Arg Gly  
65 70 75 80

His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn Leu Tyr Val  
85 90 95

Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala Ile Ala Leu  
100 105 110

Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg Leu Ser Leu  
115 120 125

Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp Arg Ala Thr  
130 135 140

Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp Leu Val Ile  
145 150 155 160

Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu Arg Phe Ser  
165 170 175

Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val Ala Met Asn  
180 185 190

Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly Val Lys Phe  
195 200 205

Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser Tyr Glu Ala  
210 215 220

Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn Gly Gly Gly  
225 230 235 240

Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val Ile Gly Ala  
245 250 255

Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp Lys Tyr Leu  
260 265 270

Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn Asn Thr Lys  
275 280 285

Lys

ES 2 640 100 T3

<210> 147  
 <211> 503  
 <212> PRT  
 <213> *Arabidopsis thaliana*

5

<400> 147

```

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys
 1          5          10          15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu
          20          25          30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu
          35          40          45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Ser Leu Pro Ile
          50          55          60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile
 65          70          75          80

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

Val Ser Arg Ala Val Leu Thr Arg Phe Tyr Ala Ala Asp Val Arg Lys
          100          105          110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Lys Arg Lys Val Val Val
          115          120          125

Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu
          130          135          140

Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr
 145          150          155          160

Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp
          165          170          175

Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Asp Ser Pro Asp
          180          185          190

Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Lys Ser Ile Cys
          195          200          205

Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro
    
```



ES 2 640 100 T3

210		215		220											
Ile	Glu	Ser	Leu	Lys	Asn	Arg	Ile	Ile	Phe	His	Asp	Gly	Arg	Leu	Val
225					230					235					240
Gln	Arg	Pro	Thr	Pro	Leu	Asn	Ala	Leu	Ile	Ile	Tyr	Leu	Trp	Leu	Pro
				245					250					255	
Phe	Gly	Phe	Met	Leu	Ser	Val	Phe	Arg	Val	Tyr	Phe	Asn	Leu	Pro	Leu
			260					265					270		
Pro	Glu	Arg	Phe	Val	Arg	Tyr	Thr	Tyr	Glu	Ile	Leu	Gly	Ile	His	Leu
		275					280					285			
Thr	Ile	Arg	Gly	His	Arg	Pro	Pro	Pro	Pro	Ser	Pro	Gly	Lys	Pro	Gly
290						295					300				
Asn	Leu	Tyr	Val	Leu	Asn	His	Arg	Thr	Ala	Leu	Asp	Pro	Ile	Ile	Ile
305					310					315					320
Ala	Ile	Ala	Leu	Gly	Arg	Lys	Ile	Thr	Cys	Val	Thr	Tyr	Ser	Val	Ser
				325					330					335	
Arg	Leu	Ser	Leu	Met	Leu	Ser	Pro	Ile	Pro	Ala	Val	Ala	Leu	Thr	Arg
			340					345					350		
Asp	Arg	Val	Ala	Asp	Ala	Ala	Arg	Met	Arg	Gln	Leu	Leu	Glu	Lys	Gly
		355					360					365			
Asp	Leu	Val	Ile	Cys	Pro	Glu	Gly	Thr	Thr	Cys	Arg	Glu	Pro	Tyr	Leu
370						375					380				
Leu	Arg	Phe	Ser	Ala	Leu	Phe	Ala	Glu	Leu	Ser	Asp	Arg	Ile	Val	Pro
385					390					395					400
Val	Ala	Met	Asn	Cys	Lys	Gln	Gly	Met	Phe	Asn	Gly	Thr	Thr	Val	Arg
				405					410					415	
Gly	Val	Lys	Phe	Trp	Asp	Pro	Tyr	Phe	Phe	Phe	Met	Asn	Pro	Arg	Pro
			420					425					430		
Ser	Tyr	Glu	Ala	Thr	Phe	Leu	Asp	Arg	Leu	Pro	Glu	Glu	Met	Thr	Val
		435					440					445			
Asn	Gly	Gly	Gly	Lys	Thr	Pro	Phe	Glu	Val	Ala	Asn	Tyr	Val	Gln	Lys
450						455						460			

ES 2 640 100 T3

Val Ile Gly Gly Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys  
465 470 475 480

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile  
485 490 495

Asn Lys Thr Lys Ser Met Glu  
500

5  
<210> 148  
<211> 467  
<212> PRT  
<213> *Oryza sativa*  
  
<400> 148

Met Ala Phe Glu Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu  
1 5 10 15

Leu Ala Pro Leu Ala Gly Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala  
20 25 30

Gly Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Ala Arg Val Ala  
35 40 45

Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser  
50 55 60

Asp Ile His Pro Glu Ser Trp Arg Val Phe Ser Ala Cys Gly Arg Arg  
65 70 75 80

Cys Val Leu Thr Ala Asn Pro Arg Met Met Val Glu Ala Phe Leu Lys  
85 90 95

Glu Tyr Leu Gly Ala Asp Ile Val Val Gly Thr Glu Leu Val Val Trp  
100 105 110

Arg Gly Arg Ala Thr Gly Leu Val Arg Ser Pro Gly Val Leu Val Gly  
115 120 125

Glu Asn Lys Ala Ala Ala Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro  
130 135 140

Glu Ile Gly Ile Gly Asp Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu  
145 150 155 160

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

ES 2 640 100 T3

Pro Arg Glu Glu Leu Pro Lys Pro Val Val Phe His Asp Gly Arg Leu  
180 185 190

Val Gln Lys Pro Ser Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile  
195 200 205

Pro Ile Gly Phe Val Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu  
210 215 220

Leu Pro Met Arg Val Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg  
225 230 235 240

Val Thr Val Lys Gly Thr Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly  
245 250 255

Gln Ser Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro  
260 265 270

Ile Phe Leu Ser Thr Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr  
275 280 285

Ser Val Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg  
290 295 300

Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu  
305 310 315 320

Ala Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu  
325 330 335

Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu  
340 345 350

Ile Val Pro Val Ala Met Glu Asn Gln Met Ser Met Phe His Gly Thr  
355 360 365

Thr Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn  
370 375 380

Pro Ser Pro Gly Tyr Val Val Thr Phe Leu Ser Lys Leu Pro Gly Glu  
385 390 395 400

Leu Thr Cys Asn Gly Gly Gly Gly Gly Gly Gly Arg Ser Ser His  
405 410 415

Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile Ala Ser Thr Leu Ser Tyr  
420 425 430

ES 2 640 100 T3

Glu Cys Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly  
 435 440 445

Asn Asp Gly Thr Val Val Ser Lys Pro Asn Ile Asp Glu Lys Lys Ala  
 450 455 460

Met Gly Cys  
 465

<210> 149  
 <211> 512  
 <212> PRT  
 <213> *Picea sitchensis*  
 <400> 149

5

Met Ala Lys Gly Gln Ser Gly Leu Asp Cys Gln Gly Leu Phe Asp Val  
 1 5 10 15

Ile Ser Lys Cys Ser Ser Lys Gly Arg Glu Asn Gln Thr Val Ala Ala  
 20 25 30

Asp Leu Asp Gly Thr Met Leu Arg Ser Arg Ser Ser Phe Pro Tyr Phe  
 35 40 45

Met Leu Val Ala Met Glu Gly Gly Ser Leu Leu Arg Gly Met Val Leu  
 50 55 60

Leu Ala Cys Ala Pro Met Ala Trp Phe Leu Tyr Asn Phe Ile Ser Glu  
 65 70 75 80

Ala Ala Gly Ile Gln Leu Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys  
 85 90 95

Ile Arg Asp Ile Glu Leu Val Ser Arg Ala Val Leu Pro Lys Phe Tyr  
 100 105 110

Ala Glu Asp Val His Pro Glu Ser Trp Arg Val Phe Ser Ala Phe Gly  
 115 120 125

Lys Arg Tyr Ile Val Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe  
 130 135 140

Val Lys Gly Tyr Leu Gly Ala Asp Lys Val Leu Gly Thr Glu Leu Gln  
 145 150 155 160

Val Ser Arg Ser Gly Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val  
 165 170 175

10

ES 2 640 100 T3

Leu Val Gly Asp Leu Lys Lys Ala Ala Val Glu Ala Glu Phe Gly Asp  
 180 185 190

Lys Leu Pro Glu Leu Gly Leu Gly Asp Arg Glu Thr Asp His Pro Phe  
 195 200 205

Met Ser Leu Cys Lys Glu Gly Tyr Met Val Pro Lys Met Lys Val Asp  
 210 215 220

Glu Val Pro Thr Asn Lys Leu Met Ser Pro Ile Val Phe His Asp Gly  
 225 230 235 240

Arg Leu Val Gln Arg Pro Asn Pro Gly Ala Ala Leu Leu Thr Phe Leu  
 245 250 255

Trp Leu Pro Ile Gly Phe Phe Leu Ala Gln Phe Arg Val Phe Gly Asn  
 260 265 270

Ile Pro Ile Pro Glu Lys Tyr Val Lys Ile Ser Tyr Lys Ile Met Gly  
 275 280 285

Ile Lys Leu Val Val Lys Gly Thr Pro Pro Pro Ala Pro Lys Lys Lys  
 290 295 300

Gly Glu Arg Gly Val Leu Phe Val Cys Asp His Arg Thr Leu Leu Asp  
 305 310 315 320

Pro Val Ile Val Ala Val Ala Leu Gly Arg Lys Val Ser Ala Val Thr  
 325 330 335

Tyr Ser Ile Ser Arg Phe Ser Glu Ile Ile Ser Pro Ile Lys Thr Val  
 340 345 350

Arg Leu Thr Arg Asp Arg Glu Arg Asp Ala Ala Asn Ile Lys Arg Leu  
 355 360 365

Leu Glu Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg  
 370 375 380

Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp  
 385 390 395 400

Arg Ile Val Pro Val Ala Ile Cys Asn Lys Met Ser Met Phe His Gly  
 405 410 415

Thr Thr Val Arg Gly Trp Lys Gly Phe Asp Pro Phe Phe Phe Phe Met  
 420 425 430

ES 2 640 100 T3

Asn Pro Val Pro Thr Tyr Glu Val Thr Phe Leu Asp Gln Leu Pro Thr  
 435 440 445

Glu Leu Thr Cys Ala Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr  
 450 455 460

Ile Gln Arg Val Leu Ala Ala Thr Leu Gly Tyr Glu Cys Thr Asn Phe  
 465 470 475 480

Thr Arg Lys Asp Lys Tyr Arg Met Leu Ala Gly Asn Asp Gly Ile Val  
 485 490 495

Ser Leu Asn Lys Thr Ala Thr Ala Arg Ser Asn Asn His Lys Gln Asn  
 500 505 510

<210> 150  
 <211> 502  
 <212> PRT  
 <213> *Zea mays*  
 <400> 150

5

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

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

Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala  
 35 40 45

Gly Ser Val Leu Arg Ala Ala Leu Leu Leu Leu Ser Val Pro Phe Val  
 50 55 60

Tyr Val Thr Tyr Ala Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu  
 65 70 75 80

Val Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Ser Ile Glu Met Val  
 85 90 95

Ala Arg Ser Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu  
 100 105 110

Ser Trp Arg Val Phe Ser Ser Phe Gly Arg Arg Tyr Val Val Thr Ala  
 115 120 125

Ser Pro Arg Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala  
 130 135 140

10

ES 2 640 100 T3

Asp Lys Val Val Gly Thr Glu Leu Glu Val Gly Arg Asp Gly Lys Ala  
 145 150 155 160  
  
 Thr Gly Phe Met Ala Arg Pro Gly Val Leu Val Gly Asp His Lys Lys  
 165 170 175  
  
 Lys Ala Val Val Lys Glu Leu Gly Asp Ala Leu Pro Asp Val Gly Met  
 180 185 190  
  
 Gly Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala  
 195 200 205  
  
 Tyr Leu Val Thr Ser Arg Lys Tyr Ser Pro Val Pro Arg Asn Gln Leu  
 210 215 220  
  
 Leu Ser Pro Leu Ile Val His Asp Gly Arg Leu Val Gln Arg Pro Thr  
 225 230 235 240  
  
 Pro Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala  
 245 250 255  
  
 Leu Ala Leu Met Arg Val Tyr Ile Asn Leu Pro Leu Pro Glu Arg Ile  
 260 265 270  
  
 Val Tyr Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Val Val Lys Gly  
 275 280 285  
  
 Thr Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val  
 290 295 300  
  
 Cys Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu  
 305 310 315 320  
  
 Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu  
 325 330 335  
  
 Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Asp Lys  
 340 345 350  
  
 Asp Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile  
 355 360 365  
  
 Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser  
 370 375 380  
  
 Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn

ES 2 640 100 T3

```

385                390                395                400

Thr Lys Glu Thr Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu
                405                410                415

Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile
                420                425                430

Thr Phe Leu Thr Gln Leu Pro Lys Asp Leu Thr Cys Ser Gly Gly Lys
                435                440                445

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

Leu Gly Phe Glu Cys Thr Ser Ile Thr Arg Lys Glu Lys Tyr Gly Met
465                470                475                480

Leu Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu
                485                490                495

Lys Glu Lys Asp Lys Asn
                500

```

5

```

<210> 151
<211> 500
<212> PRT
<213> Arabidopsis thaliana

<400> 151

```



ES 2 640 100 T3

Met Ser Pro Glu Lys Lys Ser Gln Asn Phe Pro Pro Ile Thr Glu Cys  
1                   5                   10                   15

Arg Asp Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
20                   25                   30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu  
35                   40                   45

Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Phe  
50                   55                   60

Val Ile Ile Ser Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile  
65                   70                   75                   80

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu  
85                   90                   95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys

ES 2 640 100 T3

100 105 110  
 Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr  
 115 120 125  
 Ala Asn Pro Ile Val Met Val Glu Ala Phe Val Lys Asp Tyr Leu Gly  
 130 135 140  
 Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Asn  
 145 150 155 160  
 Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp Leu  
 165 170 175  
 Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asn Glu Ser Pro Asp Leu  
 180 185 190  
 Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Leu Cys Lys  
 195 200 205  
 Lys Gly Tyr Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys  
 210 215 220  
 Glu Arg Leu Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln  
 225 230 235 240  
 Arg Pro Thr Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe  
 245 250 255  
 Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro  
 260 265 270  
 Glu Arg Phe Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr  
 275 280 285  
 Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn  
 290 295 300  
 Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala  
 305 310 315 320  
 Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg  
 325 330 335  
 Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp  
 340 345 350

ES 2 640 100 T3

Arg Ala Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp  
 355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu  
 370 375 380

Arg Ser Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser  
 420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn  
 435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val  
 450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp  
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn  
 485 490 495

Asn Thr Lys Lys  
 500

<210> 152  
 <211> 486  
 <212> PRT  
 <213> *Oryza sativa*  
 <400> 152

5

Met Ser Pro Phe Lys Pro Ile Glu Gln Cys Ser Thr Glu Gly Arg Ser  
 1 5 10 15

Gln Gln Thr Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Leu Ser Arg  
 20 25 30

Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly Gly Pro  
 35 40 45

Leu Arg Ala Val Ala Leu Leu Met Ser Val Pro Phe Val Tyr Leu Thr  
 50 55 60

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Ile Cys Val Thr Tyr Ser Val Pro Arg Lys Thr Tyr Gly Met Ser Ser  
 325 330 335

Arg Leu Pro Glu Ala Leu Thr Ala Ser Pro Val Lys Ala Ala Val Ala  
 340 345 350

Leu Cys Arg Glu Arg Asp Arg Asp Ala Asp Arg Val Arg Arg Leu Leu  
 355 360 365

Glu Glu Gly Val Asp Ile Val Ala Phe Pro Glu Gly Thr Thr Cys Arg  
 370 375 380

Gly Ala Phe Leu Leu Arg Phe Ser Ser Leu Phe Ala Glu Leu Thr Asp  
 385 390 395 400

Arg Ile Val Pro Val Ala Ile Ala Thr Arg Glu Thr Met Phe His Gly  
 405 410 415

Ser Thr Ala Arg Gly Phe Lys Gly Met Asp Pro Tyr Phe Phe Phe Met  
 420 425 430

Asn Pro Arg Pro Ala Tyr Glu Val Thr Phe Leu Ser Gln Leu Pro Ser  
 435 440 445

Glu Leu Thr Ser Gly Gly Gly Gly Lys Ser Pro Val Glu Val Ala Asn  
 450 455 460

Tyr Val Gln Lys Ala Leu Ala Gly Gln Leu Gly Ser Glu His Ile Gly  
 465 470 475 480

Ile Thr Arg Lys Glu Lys  
 485

<210> 153  
 <211> 506  
 <212> PRT  
 <213> *Oryza sativa*

5

<400> 153

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val  
 1 5 10 15

Thr Ala Tyr Asp Ala Ala Ala Gly Ala Arg Arg Thr Val Ala Ala Asp  
 20 25 30

Leu Asp Gly Thr Leu Leu Val Ser Ser Ser Ala Phe Pro Tyr Tyr Phe  
 35 40 45

10

ES 2 640 100 T3

Leu Val Ala Leu Glu Ala Gly Ser Tyr Leu Arg Ala Leu Ala Leu Leu  
 50 55 60

Leu Ala Ala Pro Trp Leu Leu Ala Leu Tyr Val Gly Val Ser Glu Ala  
 65 70 75 80

Ala Ala Ile Ala Leu Leu Val Phe Ile Thr Phe Ala Gly Leu Arg Val  
 85 90 95

Arg Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Arg His Tyr Ala  
 100 105 110

Ala Gly Val Arg Ala Asp Thr Trp Ala Val Phe His Gly Cys Ala Glu  
 115 120 125

Arg Arg Val Val Val Thr Ala Ser Pro Ala Val Met Val Gly Glu Phe  
 130 135 140

Val Arg Glu Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr  
 145 150 155 160

Phe Ala Ser Gly Lys Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val  
 165 170 175

Gly Glu Lys Lys Arg Glu Val Val Glu Arg Leu Phe Ala Gly Gly Asp  
 180 185 190

Met Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met  
 195 200 205

Ala Ile Cys Lys Glu Ala Tyr Met Val Pro Lys Asn Lys Arg Ala Pro  
 210 215 220

Arg Ala Ala Ala Asp Glu Leu Leu Ser Arg Ala Ile Phe His Asp Gly  
 225 230 235 240

Arg Leu Val Arg Arg Pro Glu Pro Ala Ser Ala Leu Phe Ala Leu Ala  
 245 250 255

Tyr Leu Pro Val Gly Phe Ala Val Ala Leu Leu Arg Val Phe Leu Asn  
 260 265 270

Leu Pro Val Pro Ala Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly  
 275 280 285

Ile Arg Leu Ala Val Arg Gly Ala Pro Pro Pro Pro Pro Arg Pro Gly  
 290 295 300

Thr Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro  
 305 310 315 320

Ile Ile Val Ser Ile Ala Leu Gly Arg Pro Val Thr Cys Val Thr Tyr  
 325 330 335

Ser Val Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala Ala  
 340 345 350

Leu Thr Pro Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu  
 355 360 365

Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu  
 370 375 380

Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala Arg  
 385 390 395 400

Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser  
 405 410 415

Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met Asn  
 420 425 430

Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro Glu  
 435 440 445

Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn His  
 450 455 460

Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu  
 465 470 475 480

Thr Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Arg Val  
 485 490 495

Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn  
 500 505

<210> 154  
 <211> 486  
 <212> PRT  
 <213> *Oryza sativa*

5

<400> 154

Met Ser Pro Phe Lys Pro Ile Lys Gln Cys Ser Thr Glu Gly Arg Ser  
 1 5 10 15

10

ES 2 640 100 T3

Gln Gln Thr Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Leu Ser Arg  
 20 25 30

Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly Gly Pro  
 35 40 45

Leu Arg Ala Val Ala Leu Leu Met Ser Val Pro Phe Val Tyr Leu Thr  
 50 55 60

Tyr Val Thr Ile Ser Glu Ser Leu Ala Val Arg Ala Leu Leu Tyr Ile  
 65 70 75 80

Ala Val Ala Gly Leu Glu Val Arg Asp Val Glu Ser Val Ala Arg Ser  
 85 90 95

Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu Gly Trp Arg  
 100 105 110

Val Phe Ser Ser Phe Gly Arg Arg Cys Val Val Thr Ala Ser Pro Arg  
 115 120 125

Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala Asp Arg Val  
 130 135 140

Ile Gly Thr Glu Leu Glu Val Gly Glu Asp Gly Arg Ala Thr Gly Phe  
 145 150 155 160

Val Ala Lys Pro Gly Val Leu Ile Arg Glu His Lys Arg Asn Ala Val  
 165 170 175

Val Arg Glu Phe Gly Asp Ala Leu Pro Asp Val Gly Met Gly Asp Arg  
 180 185 190

Glu Ser Asp Phe Asp Phe Met Ala Ile Cys Lys Asp Ala Tyr Val Val  
 195 200 205

Thr Thr Ser Arg Lys His Arg Pro Val Pro Glu Ser Gln Leu Leu Arg  
 210 215 220

Thr Val Val Leu His Asp Gly Arg Leu Ala Gln Arg Pro Thr Ala Ile  
 225 230 235 240

Asn Thr Leu Leu Val Phe Leu Trp Met Pro Val Gly Phe Ala Leu Ala  
 245 250 255

Leu Leu Arg Ala Cys Leu Ser Leu Leu Leu Pro Glu Arg Val Leu Ser



ES 2 640 100 T3

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

<210> 155  
 <211> 507  
 <212> PRT  
 <213> *Oryza sativa*

ES 2 640 100 T3

<400> 155

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val  
 1 5 10 15

Thr Ala Tyr Asp Ala Ala Ala Gly Ala Arg Arg Thr Val Ala Ala Asp  
 20 25 30

Leu Asp Gly Thr Leu Leu Val Ser Ser Ser Ala Phe Pro Tyr Tyr Phe  
 35 40 45

Leu Val Ala Leu Glu Ala Gly Ser Tyr Leu Arg Ala Leu Ala Leu Leu  
 50 55 60

Leu Ala Ala Pro Trp Leu Leu Ala Leu Tyr Val Gly Val Ser Glu Ala  
 65 70 75 80

Ala Ala Ile Ala Leu Leu Val Phe Ile Thr Phe Ala Gly Leu Arg Val  
 85 90 95

Arg Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Arg His Tyr Ala  
 100 105 110

Ala Gly Val Arg Ala Glu Thr Trp Ala Val Phe His Gly Trp Ala Lys  
 115 120 125

Arg Arg Val Leu Val Thr Ala Leu Pro Ala Val Met Val Gly Glu Phe  
 130 135 140

Val Arg Glu Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr  
 145 150 155 160

Phe Ala Ser Gly Lys Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val  
 165 170 175

Gly Glu Lys Lys Arg Glu Val Val Glu Arg Leu Phe Ala Gly Gly Asp  
 180 185 190

Met Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met  
 195 200 205

Ala Ile Cys Lys Glu Ala Tyr Met Val Pro Lys Asn Lys Arg Ala Pro  
 210 215 220

Arg Ala Ala Ala Asp Glu Leu Leu Ser Arg Ala Ile Phe His Asp Gly  
 225 230 235 240

ES 2 640 100 T3

Arg Leu Val Arg Arg Pro Glu Pro Ala Ser Ala Leu Phe Ala Leu Ala  
 245 250 255

Tyr Leu Pro Val Gly Phe Ala Val Ala Leu Leu Arg Val Phe Leu Asn  
 260 265 270

Leu Pro Val Pro Ala Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly  
 275 280 285

Ile Arg Leu Ala Val Arg Gly Ala Arg Arg Arg Pro Leu Arg Pro Gly  
 290 295 300

Asp Ala Gly Ala Pro Ser Cys Cys Ala Thr Thr Glu Pro Pro Arg Thr  
 305 310 315 320

Pro Ser Ser Cys Pro Ser Ala Leu Ala Ala Pro Leu Arg Ala Leu Arg  
 325 330 335

Thr Ala Leu Ala Pro Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala  
 340 345 350

Ala Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu  
 355 360 365

Leu Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg  
 370 375 380

Glu Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala  
 385 390 395 400

Arg Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly  
 405 410 415

Ser Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met  
 420 425 430

Asn Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro  
 435 440 445

Glu Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn  
 450 455 460

His Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr  
 465 470 475 480

Leu Thr Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Arg  
 485 490 495

ES 2 640 100 T3

Val Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn  
 500 505

5  
 <210> 156  
 <211> 497  
 <212> PRT  
 <213> *Oryza sativa*  
 <400> 156

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

10

ES 2 640 100 T3

Leu Val Thr Ser Arg Lys Tyr Ser Ala Val Pro Lys Asn Gln Leu Leu  
 210 215 220  
 Ser Pro Leu Ile Leu His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro  
 225 230 235 240  
 Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Val Leu  
 245 250 255  
 Ala Leu Leu Arg Val Tyr Val Asn Leu Pro Leu Pro Glu Arg Ile Val  
 260 265 270  
 Phe Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Ile Val Lys Gly Asn  
 275 280 285  
 Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val Cys  
 290 295 300  
 Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu Arg  
 305 310 315 320  
 Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu Leu  
 325 330 335  
 Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys Asp  
 340 345 350  
 Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile Cys  
 355 360 365  
 Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala  
 370 375 380  
 Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn Thr  
 385 390 395 400  
 Lys Glu Ser Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu Met  
 405 410 415  
 Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile Thr  
 420 425 430  
 Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Ser Gly Gly Lys Ser  
 435 440 445  
 Pro Ile Glu Val Ala Asn Tyr Ile Gln Lys Thr Leu Ser Gly Gln Leu  
 450 455 460

Gly Phe Glu Cys Thr Ala Ile Thr Arg Lys Glu Lys Tyr Ser Ile Leu  
 465 470 475 480

Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu Lys  
 485 490 495

Asn

<210> 157  
 <211> 505  
 <212> PRT  
 <213> *Oryza sativa*

5

<400> 157

Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys Cys Ser Ser Lys Asp  
 1 5 10 15

Arg Gly Gly Asp Thr Ile Val Ala Asp Leu Asp Gly Thr Leu Leu Cys  
 20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met Ala Phe Glu Thr Gly  
 35 40 45

Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu Ala Pro Leu Ala Gly  
 50 55 60

Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala Gly Ile Gln Val Leu Ile  
 65 70 75 80

Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp Val Glu Ala Val Ala  
 85 90 95

Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser Asp Ile His Pro Glu Ser  
 100 105 110

Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys Val Leu Thr Ala Asn  
 115 120 125

Pro Arg Met Met Val Glu Ala Phe Leu Lys Glu Tyr Leu Gly Ala Asp  
 130 135 140

Ile Val Val Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly  
 145 150 155 160

Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Ala Ala  
 165 170 175

10

ES 2 640 100 T3

Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro Glu Ile Gly Ile Gly Asp  
 180 185 190

Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu Cys Lys Glu Gly Tyr Val  
 195 200 205

Val Pro Val Ser Pro Thr Pro Arg Pro Val Pro Arg Glu Glu Leu Pro  
 210 215 220

Lys Pro Val Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Ser Pro  
 225 230 235 240

Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro Ile Gly Phe Val Leu  
 245 250 255

Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Arg Val Val  
 260 265 270

Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val Thr Val Lys Gly Thr  
 275 280 285

Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln Ser Gly Val Leu Phe  
 290 295 300

Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala  
 305 310 315 320

Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser Val Ser Arg Leu Ser  
 325 330 335

Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg Ala  
 340 345 350

Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Ala Glu Gly Asp Leu Val  
 355 360 365

Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe  
 370 375 380

Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile Val Pro Val Ala Met  
 385 390 395 400

Glu Asn Gln Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys  
 405 410 415

Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr Val

ES 2 640 100 T3

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

<210> 158  
 <211> 474  
 <212> PRT  
 <213> *Oryza sativa*  
 <400> 158

5



ES 2 640 100 T3

Met Gly Asp Pro Pro Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys  
 1 5 10 15

Phe Ser Ser Lys Asp Arg Gly Gly Asp Thr Ile Val Pro Asp Leu Asp  
 20 25 30

Gly Thr Leu Leu Cys Gly Arg Ser Ser Phe Pro Tyr Phe Gly His Met  
 35 40 45

Ala Phe Lys Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu  
 50 55 60

Ala Pro Phe Ala Gly Leu Leu Tyr Tyr Leu Val Ser Asp Ser Ala Gly  
 65 70 75 80

Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Pro Arg Val Pro Asp  
 85 90 95

Phe Lys Ala Ala Phe Leu Lys Glu Tyr Leu Gly Ala Asp Ile Val Val  
 100 105 110

Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly Leu Val Arg  
 115 120 125

Ser Pro Gly Val Leu Val Gly Glu Asn Lys Ala Ala Ala Leu Arg Glu



ES 2 640 100 T3

Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr Val Val Thr Phe  
385 390 395 400

Leu Ser Lys Leu Pro Gly Glu Leu Thr Cys Asn Gly Gly Gly Gly Gly  
405 410 415

Gly Gly Gly Arg Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu  
420 425 430

Ile Ala Ser Thr Leu Ser Tyr Glu Cys Thr Ser Phe Thr Arg Lys Asp  
435 440 445

Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly Thr Val Val Ser Lys Pro  
450 455 460

Asn Ile Asp Glu Lys Lys Ala Met Gly Cys  
465 470

<210> 159  
<211> 497  
<212> PRT  
<213> *Selaginella moellendorffii*

5

<400> 159

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

Asn Cys Ser Ser Glu Gly Arg Gly Met His Thr Val Ala Ala Asp Leu  
20 25 30

Asp Gly Thr Leu Leu Arg Arg Arg Ser Ser Phe Pro Tyr Phe Leu Leu  
35 40 45

Val Ala Leu Glu Gly Gly Ser Leu Phe Arg Ala Val Val Leu Gly Leu  
50 55 60

Leu Ala Pro Ile Ala Trp Leu Leu Tyr His Phe Val Ser Glu Ala Ala  
65 70 75 80

Ala Ile Gln Leu Leu Ile Tyr Val Ser Phe Ser Gly Gln Lys Val Ser  
85 90 95

Glu Ile Glu Lys Val Ala Thr Ala Val Leu Pro Lys Phe Tyr Ser Gln  
100 105 110

Asp Val His Ser Asp Ala Phe Arg Val Phe Ser Ser Cys Gly Lys Lys  
115 120 125

10

ES 2 640 100 T3

Tyr Val Ile Thr Ala Asn Pro Arg Ile Met Val Glu Tyr Phe Cys Lys  
 130 135 140

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

Asp Asp Gly Arg Ala Thr Gly Phe Val Lys Phe Pro Gly Val Leu Val  
 165 170 175

Ser Gln Asn Lys Arg Thr Ala Leu Lys Leu Glu Phe Val Asp Gly Glu  
 180 185 190

Leu Pro Glu Ile Gly Leu Gly Asp Arg Glu Thr Asp Phe Ala Phe Met  
 195 200 205

Ser Leu Cys Lys Glu Ala Tyr Val Val Pro Lys Arg Lys Val Asp Ala  
 210 215 220

Val Asp Gly Ser Ala Leu Ala Gln Arg Val Ile Phe His Asp Gly Arg  
 225 230 235 240

Leu Val Gln Leu Pro Thr Pro Phe Val Ala Leu Val Thr Phe Leu Trp  
 245 250 255

Leu Pro Ile Gly Phe Ala Leu Ala Val Leu Arg Ile Gly Leu Thr Ser  
 260 265 270

Arg Cys Pro Arg Lys Tyr Met Gly Met Ile Tyr Ala Ile Leu Gly Val  
 275 280 285

Arg Arg Ile Thr Lys Gly Lys Val Pro Arg Lys Gln Asp Gly Arg Gly  
 290 295 300

Leu Leu Phe Val Cys Asn His Arg Thr Leu Val Asp Pro Ile Phe Leu  
 305 310 315 320

Gly Leu Ala Ala Glu Arg Asp Val Thr Gly Leu Thr Tyr Ser Ile Ser  
 325 330 335

Arg Val Ser Glu Val Leu Ser Pro Ile Val Thr Val Arg Leu Thr Arg  
 340 345 350

Asp Lys Glu Arg Asp Phe Thr Lys Met Lys Ser Leu Leu Glu Arg Gly  
 355 360 365

Arg Asp Leu Cys Leu Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe  
 370 375 380

Leu Leu Arg Phe Ser Ala Leu Phe Thr Gln Leu Thr His Arg Ile Val  
 385 390 395 400

Pro Ala Ala Val Lys Thr Arg Gln Ser Met Phe Asn Gly Thr Ala Val  
 405 410 415

Arg Gly Trp Lys Gly Met Asp Pro Phe Phe Phe Phe Met Asn Pro Phe  
 420 425 430

Pro Thr Tyr Glu Val Glu Phe Leu Glu Glu Leu Pro Ala Glu Met Ser  
 435 440 445

Val Gln Lys Gly Gly Lys Ser Ser Phe Glu Thr Ala Asn Phe Ile Gln  
 450 455 460

Lys Met Leu Gly Asp Lys Leu Gly Tyr Glu Cys Thr Thr Phe Thr Arg  
 465 470 475 480

Lys Asp Lys Tyr Leu Met Leu Asn Gly Ser Asp Gly Thr Val Pro Thr  
 485 490 495

Lys

- <210> 160
- <211> 534
- <212> PRT
- <213> *Selaginella moellendorffii*
- <400> 160

5

Met Ser Leu Asp Gly Gly Lys Val Thr Asp Ser Phe Leu Val Asp Ala  
 1 5 10 15

Pro Phe Glu Pro Val Glu Arg Cys Ser Ser Ala Gly Arg Glu Lys Gln  
 20 25 30

Thr Cys Val Ala Asp Phe Asp Gly Thr Leu Ile Arg Gly Arg Ser Ser  
 35 40 45

Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Gly Gly Gly Leu Ile Arg  
 50 55 60

Ser Val Ile Leu Gly Leu Phe Ala Pro Leu Ala Trp Ile Leu Tyr His  
 65 70 75 80

Phe Val Ser Glu Ser Ala Gly Ile Arg Leu Leu Ile Phe Ile Ser Phe  
 85 90 95

10

ES 2 640 100 T3

Ala Gly Met Lys Ala Lys Gln Ile Glu Ser Val Ser Arg Ala Val Leu  
 100 105 110

Pro Lys Phe Tyr Ser Glu Asp Val His Ser Glu Ala Tyr Arg Val Phe  
 115 120 125

Ser Ser Cys Gly Lys Arg Val Val Val Thr Ala Asn Pro Arg Ile Met  
 130 135 140

Val Glu His Phe Ala Lys Thr Cys Leu Gly Ala Asp Lys Val Leu Gly  
 145 150 155 160

Thr Glu Ile Glu Val Thr Arg Ser Gly Tyr Ala Thr Gly Phe Leu Lys  
 165 170 175

Arg Pro Gly Val Leu Val Gly Val Asn Lys Arg Lys Ala Val Arg Gln  
 180 185 190

Glu Phe Gly Glu His Leu Pro His Val Gly Ile Gly Asp Arg Ala Thr  
 195 200 205

Asp Phe Ala Phe Met Ala Leu Cys Lys Glu Ala Tyr Val Val Pro Pro  
 210 215 220

Ser Lys Val Glu Ala Leu Ser Arg Asp Lys Leu Val Lys Pro Val Val  
 225 230 235 240

Phe His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro Leu Thr Ala Leu  
 245 250 255

Val Thr Phe Leu Trp Leu Pro Val Ser Phe Val Leu Ala Val Phe Arg  
 260 265 270

Ile Met Val Thr Val Pro Cys Pro Arg Glu Tyr Val Thr Ile Val Tyr  
 275 280 285

Lys Leu Val Gly Ile Arg Leu Ile Val Lys Gly Pro Ile Pro Pro Pro  
 290 295 300

Lys Lys Arg Gly Glu Ser Gly Val Leu Phe Val Ser Ser His Arg Thr  
 305 310 315 320

Leu Cys Asp Pro Val Phe Val Gly Val Gly Ala Arg Arg Glu Val Thr  
 325 330 335

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

ES 2 640 100 T3

Lys Thr Val Gly Leu Ser Arg Asp Arg Glu Arg Asp Ala Ala Lys Ile  
 355 360 365

Lys Ala Leu Leu Gln Lys Gly Asp Leu Cys Ile Cys Pro Glu Gly Thr  
 370 375 380

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu  
 385 390 395 400

Leu Thr Asp Lys Ile Val Pro Val Ala Leu Cys Thr Lys Gly Ser Thr  
 405 410 415

Phe His Gly Thr Thr Val Arg Gly Trp Lys Gly Leu Asp Pro Phe Phe  
 420 425 430

Phe Phe Met Asn Pro Phe Pro Thr Tyr Glu Val Thr Phe Leu Gln Gln  
 435 440 445

Leu Pro Pro Glu Leu Thr Val Gln Gln Gly Gly Lys Ser Ala Ile Glu  
 450 455 460

Val Ala Asn His Ile Gln Arg Val Ile Ala Asp Thr Leu Gly Phe Glu  
 465 470 475 480

Cys Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ala Gly Asn  
 485 490 495

Asp Gly Thr Val Pro Glu Arg Arg Arg Ser Gly Gly Gly Gly Gly Gly  
 500 505 510

Ala Ala Lys Asp Leu Pro Lys Asp Glu Gln Ala Ala Gln Gly Gln Asp  
 515 520 525

Gln Val Asn Gly Cys Ser  
 530

- <210> 161
- <211> 494
- <212> PRT
- <213> *Selaginella moellendorffii*

5

<400> 161

Met Ala Glu Phe Pro Pro Ile Glu Lys Cys Gly Thr Thr Ile Ser Arg  
 1 5 10 15

Glu Ser Glu Ser Val Ala Ser Asp Leu Asp Gly Thr Leu Leu Arg Gly  
 20 25 30

10

ES 2 640 100 T3

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





ES 2 640 100 T3

Met Ala Glu Phe Pro Pro Ile Glu Lys Cys Ser Ser Thr Ile Ser Arg

ES 2 640 100 T3

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

ES 2 640 100 T3

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

<210> 163  
<211> 490

<212> PRT  
<213> *Selaginella moellendorffii*

<400> 163

5

ES 2 640 100 T3

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

ES 2 640 100 T3

Val Ala Leu Val Thr Phe Leu Trp Leu Pro Ala Gly Phe Leu Leu Ala  
 245 250 255

Phe Phe Arg Ile Leu Val Ala Leu Pro Cys Pro Val Glu Tyr Val Pro  
 260 265 270

Leu Leu Tyr Lys Val Ile Gly Val Arg Leu Ile Val Lys Gly Lys Val  
 275 280 285

Pro Pro Pro Ser Lys Asn Gly Ile Leu Phe Val Cys Thr His Arg Thr  
 290 295 300

Leu Ile Asp Pro Ile Met Ala Ser Ala Ala Ser Gly Ser Asp Val Thr  
 305 310 315

Ala Leu Thr Tyr Ser Ile Ser Arg Val Ser Glu Val Leu Ser Pro Ile  
 325 330 335

Lys Thr Val Pro Leu Ser Arg Asp Arg Glu Arg Asp Ala Ser Asn Ile  
 340 345 350

Lys Arg Leu Leu Leu Lys Glu Asn Leu Ala Ile Cys Pro Glu Gly Thr  
 355 360 365

Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu  
 370 375 380

Leu Ser Glu Asn Ile Val Pro Val Ala Thr Ser Thr Lys Thr Ser Ile  
 385 390 395

Phe His Gly Thr Thr Val Arg Gly Trp Lys Gly Met Asp Pro Phe Phe  
 405 410 415

Phe Phe Met Asn Pro Phe Pro Val Tyr Glu Val Thr Phe Leu Glu Arg  
 420 425 430

Leu Pro His Asp Leu Ser Val Gln Gly Gly Lys Ser Ala Ile Glu Val  
 435 440 445

Ala Asn Tyr Ile Gln Arg Val Leu Ala Asp Thr Leu Gly Phe Glu Cys  
 450 455 460

Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ile Gly Ser Asp  
 465 470 475 480

Gly Thr Val Pro Ala Arg Lys Lys Ser Phe  
 485 490

## ES 2 640 100 T3

<210> 164  
<211> 497  
<212> PRT  
5 <213> *Selaginella moellendorffii*  
  
<400> 164



ES 2 640 100 T3

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

Asn Cys Ser Ser Glu Gly Arg Gly Met His Thr Val Ala Ala Asp Leu  
20 25 30

Asp Gly Thr Leu Leu Arg Gly Arg Ser Ser Phe Pro Tyr Phe Leu Leu  
35 40 45

Val Ala Leu Glu Gly Gly Ser Leu Phe Arg Ala Val Val Leu Gly Leu  
50 55 60

Leu Ala Pro Ile Ala Trp Leu Leu Tyr His Phe Val Ser Glu Ala Ala  
65 70 75 80

Ala Ile Gln Leu Leu Ile Tyr Val Ser Phe Ser Gly Gln Lys Val Ser  
85 90 95

Glu Ile Glu Lys Val Ala Thr Ala Val Leu Pro Lys Phe Tyr Ser Gln  
100 105 110

Asp Val His Ser Asp Ala Phe Arg Val Phe Ser Ser Cys Gly Lys Lys  
115 120 125

Tyr Val Ile Thr Ala Asn Pro Arg Ile Met Val Glu Tyr Phe Cys Lys  
130 135 140

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

Asp Asp Gly Arg Ala Thr Gly Phe Val Lys Phe Pro Gly Val Leu Val  
165 170 175

Ser Gln Asn Lys Arg Thr Ala Leu Lys Leu Glu Phe Val Asp Gly Glu  
180 185 190

Leu Pro Glu Ile Gly Leu Gly Asp Arg Glu Thr Asp Phe Ala Phe Met  
195 200 205

Ser Leu Cys Lys Glu Ala Tyr Val Val Pro Lys Arg Lys Val Asp Ala  
210 215 220

ES 2 640 100 T3

Val Asp Gly Ser Ala Leu Ala Gln Arg Val Ile Phe His Asp Gly Arg  
 225 230 235 240

Leu Val Gln Leu Pro Thr Pro Phe Val Ala Leu Val Thr Phe Leu Trp  
 245 250 255

Leu Pro Ile Gly Phe Ala Leu Ala Val Leu Arg Ile Gly Leu Thr Ser  
 260 265 270

Arg Cys Pro Arg Lys Tyr Met Gly Met Ile Tyr Ala Ile Leu Gly Val  
 275 280 285

Arg Arg Ile Thr Lys Gly Lys Val Pro Arg Lys Gln Asp Gly Arg Gly  
 290 295 300

Leu Leu Phe Val Cys Asn His Arg Thr Leu Val Asp Pro Ile Phe Val  
 305 310 315 320

Gly Leu Ala Ala Glu Arg Asp Val Thr Gly Leu Thr Tyr Ser Ile Ser  
 325 330 335

Arg Val Ser Glu Val Leu Ser Pro Ile Val Thr Val Arg Leu Thr Arg  
 340 345 350

Asp Lys Gln Arg Asp Phe Thr Lys Met Lys Ser Leu Leu Glu Arg Gly  
 355 360 365

Arg Asp Leu Cys Leu Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe  
 370 375 380

Leu Leu Arg Phe Ser Ala Leu Phe Thr Glu Leu Thr His Arg Ile Val  
 385 390 395 400

Pro Ala Ala Val Lys Thr Arg Gln Ser Met Phe Asn Gly Thr Ala Val  
 405 410 415

Arg Gly Trp Lys Gly Met Asp Pro Phe Phe Phe Phe Met Asn Pro Phe  
 420 425 430

Pro Thr Tyr Glu Val Glu Phe Leu Glu Glu Leu Pro Val Glu Met Ser  
 435 440 445

Val Gln Lys Gly Gly Lys Ser Ser Phe Glu Thr Ala Asn Phe Ile Gln  
 450 455 460

Lys Met Leu Gly Asp Lys Leu Gly Tyr Glu Cys Thr Thr Phe Thr Arg

ES 2 640 100 T3

465

470

475

480

Lys Asp Lys Tyr Leu Met Leu Asn Gly Ser Asp Gly Thr Val Pro Thr  
485 490 495

Lys

<210> 165

<211> 534

<212> PRT

<213> *Selaginella moellendorffii*

<400> 165

5

ES 2 640 100 T3

Met Ser Leu Asp Gly Gly Lys Val Thr Asp Ser Phe Leu Val Asp Ala  
 1 5 10 15

Pro Phe Glu Pro Val Glu Arg Cys Ser Ser Ala Gly Arg Glu Lys Gln  
 20 25 30

Thr Cys Val Ala Asp Phe Asp Gly Thr Leu Ile Arg Gly Arg Ser Ser  
 35 40 45

Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Gly Gly Gly Leu Ile Arg  
 50 55 60

Ser Val Ile Leu Gly Leu Phe Ala Pro Leu Ala Trp Ile Leu Tyr His  
 65 70 75 80

Phe Val Ser Glu Ser Ala Gly Ile Arg Leu Leu Ile Phe Ile Ser Phe  
 85 90 95

Ala Gly Met Lys Ala Lys Gln Ile Glu Ser Val Ser Arg Ala Val Leu  
 100 105 110

Pro Lys Phe Tyr Ser Glu Asp Val His Ser Glu Ala Tyr Arg Val Phe  
 115 120 125

Ser Ser Cys Gly Lys Arg Val Val Val Thr Ala Asn Pro Arg Ile Met  
 130 135 140

Val Glu His Phe Ala Lys Thr Cys Leu Gly Ala Asp Lys Val Leu Gly  
 145 150 155 160

Thr Glu Ile Glu Val Thr Arg Ser Gly Tyr Ala Thr Gly Phe Leu Lys  
 165 170 175

Arg Pro Gly Val Leu Val Gly Val Asn Lys Arg Lys Ala Val Arg Gln

ES 2 640 100 T3

				180						185					190				
Glu	Phe	Gly	Glu	His	Leu	Pro	His	Val	Gly	Ile	Gly	Asp	Arg	Ala	Thr				
		195					200					205							
Asp	Phe	Ala	Phe	Met	Ala	Leu	Cys	Lys	Glu	Ala	Tyr	Val	Val	Pro	Pro				
	210					215					220								
Ser	Lys	Val	Glu	Ala	Leu	Ser	Arg	Asp	Lys	Leu	Val	Lys	Pro	Val	Val				
225					230					235									
Phe	His	Asp	Gly	Arg	Leu	Val	Gln	Arg	Pro	Thr	Pro	Leu	Thr	Ala	Leu				
				245					250					255					
Val	Thr	Phe	Leu	Trp	Leu	Pro	Val	Ser	Phe	Val	Leu	Ala	Val	Phe	Arg				
			260					265					270						
Ile	Met	Val	Thr	Val	Pro	Cys	Pro	Arg	Glu	Tyr	Val	Thr	Ile	Val	Tyr				
		275					280						285						
Lys	Leu	Val	Gly	Ile	Arg	Leu	Ile	Val	Lys	Gly	Pro	Ile	Pro	Pro	Pro				
	290					295					300								
Lys	Lys	Arg	Gly	Glu	Ser	Gly	Val	Leu	Phe	Val	Ser	Ser	His	Arg	Thr				
305					310					315					320				
Leu	Cys	Asp	Pro	Val	Phe	Val	Gly	Val	Gly	Ala	Arg	Arg	Glu	Val	Thr				
				325					330					335					
Ala	Leu	Thr	Tyr	Ser	Ile	Ser	Arg	Val	Ser	Glu	Phe	Leu	Ser	Pro	Ile				
			340					345					350						
Lys	Thr	Val	Gly	Leu	Ser	Arg	Asp	Arg	Glu	Arg	Asp	Ala	Ala	Lys	Ile				
		355					360					365							
Lys	Ala	Leu	Leu	Gln	Lys	Gly	Asp	Leu	Cys	Ile	Cys	Pro	Glu	Gly	Thr				
	370					375					380								
Thr	Cys	Arg	Glu	Pro	Phe	Leu	Leu	Arg	Phe	Ser	Ala	Leu	Phe	Ala	Glu				
385					390					395					400				
Leu	Thr	Asp	Lys	Ile	Val	Pro	Val	Ala	Leu	Cys	Thr	Lys	Gly	Ser	Thr				
				405					410					415					
Phe	His	Gly	Thr	Thr	Val	Arg	Gly	Trp	Lys	Gly	Leu	Asp	Pro	Phe	Phe				
			420					425					430						

ES 2 640 100 T3

Phe Phe Met Asn Pro Phe Pro Thr Tyr Glu Val Thr Phe Leu Gln Gln  
 435 440 445

Leu Pro Pro Glu Leu Thr Val Gln Gln Gly Gly Lys Ser Ala Ile Glu  
 450 455 460

Val Ala Asn His Ile Gln Arg Val Ile Ala Asp Thr Leu Gly Phe Glu  
 465 470 475 480

Cys Thr Asn Phe Thr Arg Lys Asp Lys Tyr Gly Met Leu Ala Gly Asn  
 485 490 495

Asp Gly Thr Val Pro Glu Arg Arg Arg Ser Gly Gly Gly Gly Gly Gly  
 500 505 510

Ala Ala Lys Asp Leu Pro Glu Asp Glu Gln Ala Ala Gln Gly Gln Asp  
 515 520 525

Gln Val Asn Gly Cys Ser  
 530

<210> 166  
 <211> 497  
 <212> PRT  
 <213> *Oryza sativa*  
 <400> 166

5

Met Val Ser Arg Arg Phe Lys Pro Val Glu Glu Cys Ser Ser Asp Gly  
 1 5 10 15

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

Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala Gly  
 35 40 45

Ser Val Leu Arg Ala Val Val Leu Leu Leu Ser Val Pro Phe Val Tyr  
 50 55 60

Val Thr Tyr Ile Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu Val  
 65 70 75 80

Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Asn Ile Glu Met Val Ala  
 85 90 95

Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His Pro Glu Ser  
 100 105 110

10

ES 2 640 100 T3

Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Ile Thr Ala Ser  
 115 120 125

Pro Arg Ile Met Val Glu His Phe Ala Lys Thr Phe Leu Gly Ala Asp  
 130 135 140

Lys Val Val Gly Thr Glu Leu Glu Val Gly Lys Asn Gly Lys Ala Thr  
 145 150 155 160

Gly Phe Met Val Lys Pro Gly Val Leu Val Gly Asp His Lys Arg Gln  
 165 170 175

Ala Val Val Lys Glu Leu Arg Asp Ala Val Pro Asp Val Gly Leu Gly  
 180 185 190

Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala Tyr  
 195 200 205

Leu Val Thr Ser Arg Lys Tyr Ser Ala Val Pro Lys Asn Gln Leu Leu  
 210 215 220

Ser Pro Leu Ile Leu His Asp Gly Arg Leu Val Gln Arg Pro Thr Pro  
 225 230 235 240

Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala Leu  
 245 250 255

Ala Leu Leu Arg Val Tyr Val Asn Leu Pro Leu Pro Glu Arg Ile Val  
 260 265 270

Phe Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Ile Val Lys Gly Asn  
 275 280 285

Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val Cys  
 290 295 300

Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu Arg  
 305 310 315 320

Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu Leu  
 325 330 335

Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys Asp  
 340 345 350

Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile Cys  
 355 360 365

ES 2 640 100 T3

Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala  
 370 375 380

Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn Thr  
 385 390 395 400

Lys Glu Ser Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu Met  
 405 410 415

Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile Thr  
 420 425 430

Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Ser Gly Gly Lys Ser  
 435 440 445

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

Gly Phe Glu Cys Thr Ala Ile Thr Arg Lys Glu Lys Tyr Ser Ile Leu  
 465 470 475 480

Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu Lys  
 485 490 495

Asn

<210> 167  
 <211> 506  
 <212> PRT  
 <213> *Oryza sativa*  
 <400> 167

5

Met Ala Ala Ala Ala Ala Ala Gly Arg Arg Arg Phe Phe Pro Pro Val  
 1 5 10 15

Thr Ala Tyr Asp Ala Ala Ala Gly Ala Arg Arg Thr Val Ala Ala Asp  
 20 25 30

Leu Asp Gly Thr Leu Leu Val Ser Ser Ser Ala Phe Pro Tyr Tyr Phe  
 35 40 45

Leu Val Ala Leu Glu Ala Gly Ser Tyr Leu Arg Ala Leu Ala Leu Leu  
 50 55 60

Leu Ala Ala Pro Trp Leu Leu Ala Leu Tyr Val Gly Val Ser Glu Ala  
 65 70 75 80

10



ES 2 640 100 T3

Ala Ala Ile Ala Leu Leu Val Phe Ile Thr Phe Ala Gly Leu Arg Val  
85 90 95

Arg Asp Val Glu Ala Val Ala Arg Ala Val Leu Pro Arg His Tyr Ala  
100 105 110

Ala Gly Val Arg Ala Asp Thr Trp Ala Val Phe His Gly Cys Ala Glu  
115 120 125

Arg Arg Val Val Val Thr Ala Ser Pro Ala Val Met Val Gly Glu Phe  
130 135 140

Val Arg Glu Phe Leu Gly Ala Glu Val Ala Gly Thr Glu Leu Glu Thr  
145 150 155 160

Phe Ala Ser Gly Lys Arg Phe Thr Gly Arg Ile Lys Ala Val Leu Val  
165 170 175

Gly Glu Lys Lys Arg Glu Val Val Glu Arg Leu Phe Ala Gly Gly Asp  
180 185 190

Met Pro Asp Val Gly Leu Gly Asp Arg Glu Ser Asp His Asp Phe Met  
195 200 205

Ala Ile Cys Lys Glu Ala Tyr Met Val Pro Lys Asn Lys Arg Ala Pro  
210 215 220

Arg Ala Ala Ala Asp Glu Leu Leu Ser Arg Ala Ile Phe His Asp Gly  
225 230 235 240

Arg Leu Val Arg Arg Pro Glu Pro Ala Ser Ala Leu Phe Ala Leu Ala  
245 250 255

Tyr Leu Pro Val Gly Phe Ala Val Ala Leu Leu Arg Val Phe Leu Asn  
260 265 270

Leu Pro Val Pro Ala Arg Leu Val Arg His Thr Tyr Arg Leu Thr Gly  
275 280 285

Ile Arg Leu Ala Val Arg Gly Ala Pro Pro Pro Pro Pro Arg Pro Gly  
290 295 300

Thr Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro  
305 310 315 320

Ile Ile Val Ser Ile Ala Leu Gly Arg Pro Val Thr Cys Val Thr Tyr  
325 330 335

ES 2 640 100 T3

Ser Val Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Ala Ala  
 340 345 350

Leu Thr Arg Asp Arg Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu  
 355 360 365

Glu Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu  
 370 375 380

Pro Tyr Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ala Arg  
 385 390 395 400

Ile Val Pro Val Ala Val Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser  
 405 410 415

Thr Ala Arg Gly Trp Lys Phe Leu Asp Pro Tyr Phe Phe Tyr Met Asn  
 420 425 430

Pro Arg Pro Gly Tyr Glu Val Thr Phe Leu Pro Ala Leu Arg Pro Glu  
 435 440 445

Glu Thr Cys Val Ala Gly Gly Arg Ser Ala Val Glu Val Ala Asn His  
 450 455 460

Val Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu  
 465 470 475 480

Thr Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Arg Val  
 485 490 495

Ala Ala Ala Ala Asp Lys Pro Lys Ala Asn  
 500 505

<210> 168  
 <211> 515  
 <212> PRT  
 <213> *Zea mays*

5

<400> 168

Met Ala Ala Arg Phe Pro Pro Val Ser Ser Tyr Asp Ala Ser Ala Arg  
 1 5 10 15

Val Arg Arg Thr Ala Ala Ala Asp Leu Asp Gly Thr Leu Leu Ala Ser  
 20 25 30

Ser Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala Leu Glu Ala Gly Ser  
 35 40 45

10

ES 2 640 100 T3

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

ES 2 640 100 T3

290 295 300

Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ser Val  
 305 310 315 320

Ala Leu Gly Arg Pro Val Ala Cys Val Thr Tyr Ser Val Ser Arg Leu  
 325 330 335

Ser Thr Ala Val Ser Pro Val Arg Ala Val Ala Leu Ser Arg Asp Arg  
 340 345 350

Ala Ala Asp Ala Ala Arg Ile Ala Ala Leu Leu Ala Glu Gly Asp Val  
 355 360 365

Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Cys Leu Leu Arg  
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala  
 385 390 395 400

Leu Asp Ala Arg Gln Gly Thr Tyr Tyr Gly Ser Thr Ala Arg Gly Trp  
 405 410 415

Lys Trp Leu Asp Pro Tyr Phe Phe Tyr Met Asn Pro Arg Pro Gly Tyr  
 420 425 430

Asp Val Thr Phe Leu Pro Pro Leu Arg Pro Glu Glu Thr Cys Gly Ala  
 435 440 445

Gly Gly Arg Ser Ala Val Asp Val Ala Asn His Val Gln Thr Leu Ile  
 450 455 460

Ala Lys Glu Leu Gly Phe Arg Cys Thr Lys Leu Thr Arg Lys Asp Lys  
 465 470 475 480

Tyr Met Lys Leu Ala Gly Asn Asp Gly Thr Val Gly Gly Arg Pro Glu  
 485 490 495

Asn Glu Lys Gly Ala Ala Asp Asp Ser Ala Ala Thr Ala Thr Thr Lys  
 500 505 510

Lys Leu Val  
 515

<210> 169  
 <211> 500  
 <212> PRT

<213> *Zea mays*

<220>

<221> misc\_feature

5

<222> (92)..(92)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 169

ES 2 640 100 T3

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

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

Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu Glu Ala  
35 40 45

Gly Ser Val Leu Arg Ala Ala Leu Leu Leu Leu Ser Val Pro Phe Val  
50 55 60

Tyr Val Thr Tyr Ala Phe Phe Ser Glu Ser Leu Ala Ile Ser Thr Leu  
65 70 75 80

Val Tyr Ile Ser Val Ala Gly Leu Lys Val Arg Xaa Ile Glu Met Val  
85 90 95

Ala Arg Ser Val Leu Pro Arg Phe Tyr Ala Gly Asp Val His Pro Glu  
100 105 110

Ser Trp Arg Val Phe Ser Ser Phe Gly Arg Arg Tyr Val Val Thr Ala  
115 120 125

Ser Pro Arg Val Met Val Glu Pro Phe Ala Arg Ala Phe Leu Gly Ala  
130 135 140

Asp Lys Val Val Gly Thr Glu Leu Glu Val Gly Arg Asp Gly Lys Ala  
145 150 155 160

Thr Gly Phe Met Ala Arg Pro Gly Val Leu Val Gly Asp His Lys Lys  
165 170 175

Lys Ala Val Val Lys Glu Leu Gly Asp Ala Leu Pro Asp Val Gly Met  
180 185 190

Gly Asp Arg Glu Thr Asp Phe Asp Phe Met Ser Ile Cys Lys Glu Ala  
195 200 205

Tyr Leu Val Thr Ser Arg Lys Tyr Ser Pro Val Pro Arg Asn Gln Leu  
210 215 220

ES 2 640 100 T3

Leu Ser Pro Leu Ile Val His Asp Gly Arg Leu Val Gln Arg Pro Thr  
 225 230 235 240

Pro Leu Val Ala Leu Val Thr Phe Leu Trp Met Pro Phe Gly Phe Ala  
 245 250 255

Leu Ala Leu Met Arg Val Tyr Ile Asn Leu Pro Leu Pro Glu Arg Ile  
 260 265 270

Val Tyr Tyr Thr Tyr Lys Leu Met Gly Ile Arg Leu Val Val Lys Gly  
 275 280 285

Thr Pro Pro Pro Pro Pro Lys Lys Gly His Pro Gly Val Leu Phe Val  
 290 295 300

Cys Asn His Arg Thr Val Leu Asp Pro Val Glu Val Ala Val Ala Leu  
 305 310 315 320

Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu  
 325 330 335

Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Asp Lys  
 340 345 350

Asp Ala Glu Asn Ile Arg Arg Leu Leu Glu Glu Gly Asp Leu Val Ile  
 355 360 365

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser  
 370 375 380

Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn  
 385 390 395 400

Thr Lys Glu Thr Met Phe His Gly Ser Thr Val Arg Gly Phe Lys Leu  
 405 410 415

Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Ile  
 420 425 430

Thr Phe Leu Thr Gln Leu Pro Lys Asp Leu Thr Cys Ser Gly Gly Lys  
 435 440 445

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

Leu Gly Phe Glu Cys Thr Ser Ile Thr Arg Lys Glu Lys Tyr Gly Met

ES 2 640 100 T3

465

470

475

480

Leu Ala Gly Thr Asp Gly Arg Val Pro Ser Lys Asn Lys Glu Lys Glu  
485 490 495

Lys Asp Lys Asn  
500

<210> 170

<211> 486

<212> PRT

<213> *Zea mays*

<400> 170

5



ES 2 640 100 T3

Met Val Ser Arg Arg Phe Lys Pro Ile Glu Leu Cys Asp Ser Glu Gly  
 1 5 10 15

Arg Ser Arg Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu Leu  
 20 25 30

Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Ile Ala Leu Glu Ala Gly  
 35 40 45

Ser Leu Leu Arg Ala Val Ala Leu Leu Leu Ser Val Pro Leu Val Tyr  
 50 55 60

Leu Thr Tyr Val Thr Val Ser Glu Thr Leu Ala Val Arg Ala Phe Ile  
 65 70 75 80

Tyr Val Ala Met Ala Gly Leu Glu Ala Arg Asp Ile Glu Ala Val Ala  
 85 90 95

Arg Ser Val Leu Pro Ser Phe Tyr Ala Gly Asp Val His Pro Glu Gly  
 100 105 110

Trp Arg Val Phe Arg Ser Phe Gly Arg Arg Cys Val Val Thr Ala Ser  
 115 120 125

Pro Arg Leu Met Val Glu Pro Phe Ala Lys Ala Phe Leu Gly Ala Asp  
 130 135 140

Val Val Ile Gly Thr Glu Val Glu Val Ala Gln Ser Gly Lys Ala Thr  
 145 150 155 160

Gly Phe Val Ala Gly Pro Gly Val Leu Val Gly Glu His Lys Arg Arg  
 165 170 175

Ala Val Ala Arg Glu Phe Gly Asp Ser Leu Pro Asp Val Gly Met Gly

ES 2 640 100 T3

			180					185					190			
Asp	Arg	Glu	Ser	Asp	Phe	Asp	Phe	Met	Ser	Ile	Cys	Lys	Glu	Ala	Tyr	
		195					200					205				
Thr	Val	Thr	Arg	Gln	Lys	Tyr	Arg	Ala	Leu	Pro	Arg	Asp	Gln	Leu	Gln	
	210					215					220					
Gly	Arg	Ala	Ile	Leu	His	Asp	Gly	Arg	Leu	Ala	Arg	Arg	Pro	Thr	Ala	
225					230					235					240	
Thr	Asn	Ala	Leu	Leu	Thr	Phe	Leu	Trp	Met	Pro	Leu	Gly	Phe	Ala	Leu	
				245					250					255		
Ala	Leu	Val	Arg	Ala	His	Leu	His	Leu	Leu	Leu	Pro	Val	Arg	Ala	Leu	
			260					265					270			
Ala	Tyr	Ala	Tyr	Lys	Leu	Thr	Gly	Val	Lys	Leu	Val	Val	Arg	Gly	Asn	
		275					280					285				
Arg	Pro	Pro	Pro	Ser	Lys	Lys	Lys	Lys	Gly	Asp	Gln	Leu	Gly	Val	Leu	
	290					295					300					
Phe	Val	Cys	Asn	His	Arg	Ser	Thr	Leu	Asp	Ala	Val	Ala	Val	Gly	Val	
305					310					315					320	
Ala	Leu	Gly	Arg	Lys	Val	Arg	Trp	Val	Val	Thr	Asp	Gly	Gly	Ala	Gly	
				325					330					335		
Ala	Ser	Arg	Phe	Ser	Glu	Pro	Val	Val	Ser	Pro	Ile	Met	Thr	Gly	Val	
			340					345						350		
Pro	Leu	Pro	Val	Pro	Thr	Ser	Leu	Glu	Gly	Asp	Ala	Asp	Ala	Asp	Ala	
		355					360					365				
Ala	Pro	Arg	Ile	Arg	Arg	Leu	Leu	Glu	Glu	Gly	Asp	Asp	Val	Val	Val	
	370					375					380					
Phe	Pro	Glu	Gly	Ala	Ile	Cys	Arg	Glu	Pro	Phe	Leu	Leu	Arg	Phe	Gly	
385					390					395					400	
Ala	Leu	Phe	Ala	Glu	Val	Thr	Asp	Arg	Ile	Val	Pro	Val	Ala	Ile	Glu	
				405					410					415		
Pro	Arg	Glu	Ser	Met	Phe	His	Gly	Ser	Thr	Ala	Arg	Gly	Leu	Arg	Gly	
			420					425					430			

ES 2 640 100 T3

Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr Glu Val  
 435 440 445

Thr Phe Leu Asn Gln Leu Pro Gly Glu Leu Thr Cys Gly Gly Gly Arg  
 450 455 460

Ser Pro Val Glu Val Ala Ser Tyr Val Gln Glu Val Leu Ala Ala Gln  
 465 470 475 480

Leu Gly Phe Glu Cys Thr  
 485

<210> 171

<211> 501

5 <212> PRT

<213> *Arabidopsis thaliana*

<400> 171

Met Gly Ala Gln Glu Lys Arg Arg Arg Phe Glu Gln Ile Ser Lys Cys  
 1 5 10 15

Asp Val Lys Asp Arg Ser Asn His Thr Val Ala Ala Asp Leu Asp Gly  
 20 25 30

Thr Leu Leu Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala  
 35 40 45

Leu Glu Ala Gly Ser Leu Leu Arg Ala Leu Ile Leu Leu Val Ser Val  
 50 55 60

Pro Phe Val Tyr Leu Thr Tyr Leu Thr Ile Ser Glu Thr Leu Ala Ile  
 65 70 75 80

Asn Val Phe Val Phe Ile Thr Phe Ala Gly Leu Lys Ile Arg Asp Val  
 85 90 95

Glu Leu Val Val Arg Ser Val Leu Pro Arg Phe Tyr Ala Glu Asp Val  
 100 105 110

Arg Pro Asp Thr Trp Arg Ile Phe Asn Thr Phe Gly Lys Arg Tyr Ile  
 115 120 125

Ile Thr Ala Ser Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Phe  
 130 135 140

Leu Gly Val Asp Lys Val Leu Gly Thr Glu Leu Glu Val Ser Lys Ser  
 145 150 155 160

10

ES 2 640 100 T3

Gly Arg Ala Thr Gly Phe Thr Arg Lys Pro Gly Ile Leu Val Gly Gln  
 165 170 175

Tyr Lys Arg Asp Val Val Leu Arg Glu Phe Gly Gly Leu Ala Ser Asp  
 180 185 190

Leu Pro Asp Leu Gly Leu Gly Asp Ser Lys Thr Asp His Asp Phe Met  
 195 200 205

Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu Pro  
 210 215 220

Leu Pro Arg Asn Lys Leu Leu Ser Pro Ile Ile Phe His Glu Gly Arg  
 225 230 235 240

Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu Trp  
 245 250 255

Leu Pro Val Gly Phe Val Leu Ser Ile Ile Arg Val Tyr Thr Asn Ile  
 260 265 270

Pro Leu Pro Glu Arg Ile Ala Arg Tyr Asn Tyr Lys Leu Thr Gly Ile  
 275 280 285

Lys Leu Val Val Asn Gly His Pro Pro Pro Pro Pro Lys Pro Gly Gln  
 290 295 300

Pro Gly His Leu Leu Val Cys Asn His Arg Thr Val Leu Asp Pro Val  
 305 310 315 320

Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser  
 325 330 335

Ile Ser Lys Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu  
 340 345 350

Thr Arg Gln Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu  
 355 360 365

Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 370 375 380

Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile  
 385 390 395 400

Val Pro Val Ala Ile Asn Thr Lys Gln Ser Met Phe Asn Gly Thr Thr  
 405 410 415

Thr Arg Gly Tyr Lys Leu Leu Asp Pro Tyr Phe Ala Phe Met Asn Pro  
 420 425 430

Arg Pro Thr Tyr Glu Ile Thr Phe Leu Lys Gln Ile Pro Ala Glu Leu  
 435 440 445

Thr Cys Lys Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln  
 450 455 460

Arg Val Leu Gly Gly Thr Leu Gly Phe Glu Cys Thr Asn Phe Thr Arg  
 465 470 475 480

Lys Asp Lys Tyr Ala Met Leu Ala Gly Thr Asp Gly Arg Val Pro Val  
 485 490 495

Lys Lys Glu Lys Thr  
 500

<210> 172

<211> 500

5 <212> PRT

<213> *Arabidopsis thaliana*

<400> 172

Met Ser Pro Glu Lys Lys Ser Gln Asn Phe Pro Pro Ile Thr Glu Cys  
 1 5 10 15

Arg Asp Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
 20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu  
 35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Phe  
 50 55 60

Val Ile Ile Ser Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile  
 65 70 75 80

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu  
 85 90 95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys  
 100 105 110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr  
 115 120 125

10

ES 2 640 100 T3

Ala Asn Pro Ile Val Met Val Glu Ala Phe Val Lys Asp Tyr Leu Gly  
 130 135 140

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Asn  
 145 150 155 160

Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp Leu  
 165 170 175

Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asn Glu Ser Pro Asp Leu  
 180 185 190

Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Leu Cys Lys  
 195 200 205

Lys Gly Tyr Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys  
 210 215 220

Glu Arg Leu Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln  
 225 230 235 240

Arg Pro Thr Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe  
 245 250 255

Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro  
 260 265 270

Glu Arg Phe Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr  
 275 280 285

Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn  
 290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala  
 305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg  
 325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp  
 340 345 350

Arg Ala Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp  
 355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu  
 370 375 380

ES 2 640 100 T3

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser  
 420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn  
 435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val  
 450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp  
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn  
 485 490 495

Asn Thr Lys Lys  
 500

<210> 173  
 <211> 514  
 <212> PRT  
 <213> *Physcomitrella patens*  
 <400> 173

5

Met Gln Gln Leu Asp Pro Leu Thr Ser Phe His Cys Ser Glu Ser Phe  
 1 5 10 15

Glu Gly Val Glu Thr Cys Arg Val Glu Gly Arg Ala Asn Gln Ser Ile  
 20 25 30

Val Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser Phe Pro  
 35 40 45

Tyr Phe Met Leu Leu Ala Phe Glu Ala Gly Ser Pro Leu Arg Ser Val  
 50 55 60

Val Leu Leu Leu Ile Ser Pro Val Val Trp Leu Val Tyr Asn Phe Val  
 65 70 75 80

Ser Glu Ala Ala Gly Ile Lys Met Leu Ile Phe Val Ser Leu Ala Gly  
 85 90 95

10

ES 2 640 100 T3

Leu Lys Val Ser Ala Val Glu Ser Val Ala Arg Gly Val Leu Pro Lys  
 100 105 110

Phe Tyr Leu Glu Asp Met His Ser Ile Ser Tyr Arg Val Phe Thr Ser  
 115 120 125

Cys Gly Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Ile Val Glu  
 130 135 140

Pro Phe Leu Lys Glu Tyr Leu Asp Val Glu Ala Val Met Gly Thr Glu  
 145 150 155 160

Leu Gln Ile Ser Ser Gly Gly Tyr Val Thr Gly Phe Val Thr Gly Pro  
 165 170 175

Gly Val Leu Val Gly Ala Val Lys Glu Arg Ala Val Lys Lys Tyr Phe  
 180 185 190

Gly Pro Asp Pro Pro Asp Val Gly Leu Gly Asp Arg Gln Ser Asp Phe  
 195 200 205

Leu Phe Met Asp Leu Cys Lys Glu Ser Tyr Ile Val His Ser Asp Lys  
 210 215 220

Asp Val Pro Ala Val Ser Lys Glu Asp Phe Leu Lys Pro Leu Ile Phe  
 225 230 235 240

His Asp Gly Arg Leu Val Cys Arg Pro Thr Pro Leu Met Ser Leu Ile  
 245 250 255

Val Met Leu Trp Cys Pro Ile Gly Phe Val Leu Ala Val Val Arg Met  
 260 265 270

Leu Ile Gly Ile Ala Leu Pro Met Trp Trp Ala Leu Pro Leu Glu Ala  
 275 280 285

Met Leu Gly Val Thr Val Arg Val Lys Gly Ile Pro Pro Thr Cys Pro  
 290 295 300

Leu Arg Asn Lys Lys Arg Gly Val Leu Phe Val Cys Ser His Arg Thr  
 305 310 315 320

Leu Leu Asp Pro Ile Phe Leu Ser Ile Ala Val Arg Arg Lys Val Thr  
 325 330 335

Ala Val Thr Tyr Ser Ile Ser Arg Leu Ser Glu Val Leu Ser Pro Ile



ES 2 640 100 T3

	340		345		350														
Arg	Thr	Val	Arg	Leu	Thr	Arg	Asp	Arg	Asn	Thr	Asp	Ala	Asn	Thr	Met				
		355					360					365							
Ala	Ser	Leu	Leu	Glu	Glu	Gly	Asp	Leu	Val	Val	Cys	Pro	Glu	Gly	Thr				
		370					375					380							
Thr	Cys	Arg	Glu	Pro	Tyr	Leu	Leu	Arg	Phe	Ser	Ser	Leu	Phe	Ala	Glu				
					390					395					400				
Leu	Thr	Asp	Gln	Ile	Val	Pro	Val	Thr	Met	Asn	Ile	Lys	Ile	Thr	Met				
				405					410						415				
Phe	His	Gly	Thr	Ser	Ala	Arg	Gly	Trp	Lys	Gly	Met	Asp	Pro	Phe	Phe				
			420					425					430						
Phe	Phe	Met	Asn	Pro	Cys	Pro	Lys	Tyr	Glu	Val	Thr	Phe	Leu	Asp	Gln				
		435					440					445							
Leu	Pro	His	Gln	Leu	Thr	Cys	Asn	Gly	Gly	Lys	Thr	Ser	His	Glu	Val				
		450				455					460								
Ala	Asn	Tyr	Ile	Gln	Arg	Val	Leu	Ala	Asn	Ser	Leu	Gly	Phe	Gln	Cys				
		465			470					475					480				
Thr	Asn	Leu	Thr	Arg	Lys	Asp	Lys	Tyr	Thr	Val	Leu	Ala	Gly	Asn	Asp				
				485					490					495					
Gly	Ile	Val	Pro	Asp	His	Thr	Leu	Arg	Gln	Arg	Ala	Gly	Arg	Met	Phe				
			500					505					510						

Gly Leu

<210> 174  
 <211> 501  
 <212> PRT  
 <213> *Physcomitrella patens*

5

<400> 174

ES 2 640 100 T3

Met Lys Arg Asp Pro Phe Asp Thr Ile Asp Lys Cys Ser Asp Lys Gly  
1                                   5                                   10                                   15

Arg Ser Lys Gln Thr Val Val Ser Asp Leu Asp Gly Thr Leu Leu Arg  
                                  20                                   25                                   30

Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Phe Glu Ala Gly

ES 2 640 100 T3

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

ES 2 640 100 T3

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

<210> 175  
 <211> 517  
 <212> PRT  
 <213> *Physcomitrella patens*

5

<400> 175

ES 2 640 100 T3

Met Glu Thr Gln Glu Pro Asp Pro Val Thr Ser Phe His Cys Ser Glu  
 1 5 10 15

Asn Tyr Asp Glu Val Glu Asn Cys Arg Val Glu Gly Arg Ala Asn Gln  
 20 25 30

Ser Ile Val Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser  
 35 40 45

Phe Pro Tyr Phe Met Leu Leu Ala Phe Glu Ala Gly Ser Pro Leu Arg  
 50 55 60

Phe Ile Ile Leu Leu Leu Val Ser Pro Ile Val Trp Phe Val Tyr Asn  
 65 70 75 80

Phe Ile Ser Glu Val Val Gly Ile Lys Met Leu Ile Phe Ile Ser Leu  
 85 90 95

Ala Gly Leu Lys Val Ser Ala Ile Glu Ser Val Ala Arg Gly Val Leu  
 100 105 110

Pro Lys Phe Phe Leu Glu Asp Met His Ser Ile Ser Tyr Arg Val Phe  
 115 120 125

Ile Ser Cys Glu Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Ile  
 130 135 140

Val Glu Pro Phe Leu Lys Glu Tyr Leu Asp Val Asp Ala Val Met Gly  
 145 150 155 160

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

Gly Pro Gly Val Leu Val Gly Thr Ala Lys Gln Arg Ala Val Lys Lys  
 180 185 190

Tyr Phe Gly Ser Asp Gln Pro Asp Leu Gly Leu Gly Asp Arg Thr Ser  
 195 200 205

Asp Phe Ala Phe Met Asp Leu Cys Lys Glu Ala Tyr Ile Val Pro Ser  
 210 215 220

Tyr Lys Glu Val Pro Ser Val Thr Lys Lys Asp Tyr Leu Lys Pro Leu  
 225 230 235 240

Ile Phe His Asp Gly Arg Leu Val Cys Arg Pro Thr Pro Leu Met Ala  
 245 250 255

ES 2 640 100 T3

Leu Ala Val Thr Leu Trp Cys Pro Ile Gly Phe Val Leu Ala Ile Ile  
 260 265 270

Arg Met Phe Ile Gly Ile Met Leu Pro Met Trp Trp Ala Leu Pro Leu  
 275 280 285

Glu Ala Met Leu Gly Val Thr Val Arg Ile Lys Gly Ile Pro Pro Thr  
 290 295 300

Cys Pro Leu Arg Ser Lys Lys Arg Gly Val Leu Phe Val Cys Ser His  
 305 310 315 320

Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Ile Ala Val Arg Arg Lys  
 325 330 335

Val Thr Ala Val Thr Tyr Ser Ile Ser Arg Leu Ser Glu Val Leu Ala  
 340 345 350

Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg Asn Thr Asp Ala Ser  
 355 360 365

Thr Met Ala Ser Leu Leu Asp Glu Gly Asp Leu Val Val Cys Pro Glu  
 370 375 380

Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Ser Leu Phe  
 385 390 395 400

Ala Glu Leu Thr Asp Gln Ile Val Pro Val Thr Met Asn Ile Lys Ile  
 405 410 415

Ser Met Phe His Gly Thr Ser Ala Arg Gly Trp Lys Gly Met Asp Pro  
 420 425 430

Phe Phe Phe Phe Met Asn Pro Arg Pro Lys Tyr Glu Val Thr Phe Leu  
 435 440 445

Asp Gln Leu Pro His Gln Leu Thr Cys Asn Gly Gly Lys Thr Ser His  
 450 455 460

Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Asn Ser Leu Gly Phe  
 465 470 475 480

Gln Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Arg Val Leu Ala Gly  
 485 490 495

Asn Asp Gly Ile Val Pro Asp Arg Ser Leu Arg Gln Arg Ala Ser Arg  
 500 505 510

ES 2 640 100 T3

Ile Phe Gly Leu Gln  
515

5  
<210> 176  
<211> 504  
<212> PRT  
<213> *Physcomitrella patens*  
  
<400> 176

Met Glu Val Thr Lys Leu Lys Glu Ala Gly Asn Tyr Ser Phe Ala Glu  
1 5 10 15

Val Lys Ser Cys Ser Ser Gly Asn Asp Arg Arg Ser Gln Thr Val Val  
20 25 30

Ala Asp Leu Asp Gly Thr Leu Ile Arg Gly Arg Ser Ala Phe Pro Tyr  
35 40 45

Phe Phe Leu Val Ala Phe Glu Ala Gly Ser Tyr Phe Arg Ala Leu Leu  
50 55 60

Leu Ser Leu Met Ala Pro Leu Ile Gly Phe Leu Tyr Tyr Phe Val Asp  
65 70 75 80

Glu Ala Ser Gly Ile His Leu Met Ile Phe Leu Ser Phe Ala Gly Leu  
85 90 95

Lys Ile Lys Asp Ile Glu Gly Val Ala Arg Ala Val Leu Thr Lys Phe  
100 105 110

Tyr Ala Asp Asp Leu His Pro Glu Thr Trp Arg Val Phe Ser Ser Phe  
115 120 125

Gly Lys Arg Val Val Leu Thr Ala Asn Pro Arg Val Met Val Glu Pro  
130 135 140

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

Glu Val Asn Lys Lys Gly Arg Ala Thr Gly Arg Leu Leu Lys Pro Gly  
165 170 175

Val Leu Val Gly Thr His Lys Glu Glu Ala Leu Lys Arg Cys Asn Ile  
180 185 190

Gly Gly Glu Arg Pro His Val Gly Leu Gly Asp Arg Val Thr Asp Phe  
195 200 205

10

ES 2 640 100 T3

Pro Phe Met Ala Tyr Cys Lys Glu Gly Tyr Val Val Pro Lys Thr Lys  
 210 215 220

Val Pro Ala Val Lys Lys Gln Asp Met Pro Lys Gln Leu Ile Phe His  
 225 230 235 240

Asp Gly Arg Leu Val Gln Leu Pro Thr Pro Arg Asn Ala Phe Lys Val  
 245 250 255

Leu Ala Trp Met Pro Ile Gly Leu Leu Leu Ser Ile Val Arg Val Thr  
 260 265 270

Met Gly Val Trp Val Pro Ile Arg Leu Met Pro Leu Cys Tyr Lys Leu  
 275 280 285

Thr Gly Ile Asn Leu Val Val Lys Gly Asn Ile Pro Glu Lys Pro Lys  
 290 295 300

Asn Gly Glu Pro Gly Arg Leu Phe Val Cys Asn His Arg Thr Leu Leu  
 305 310 315 320

Asp Pro Val Ile Ile Ala Leu Ala Leu Gly Arg Pro Val Pro Ala Val  
 325 330 335

Thr Tyr Ser Ile Ser Lys Val Ser Glu Phe Leu Ser Pro Met Pro Thr  
 340 345 350

Ile Gly Leu Cys Arg Asp Arg Glu Lys Asp Ser Ala Asn Met Arg Arg  
 355 360 365

Val Leu Lys Glu Gly Glu Leu Thr Leu Cys Pro Glu Gly Thr Thr Cys  
 370 375 380

Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser  
 385 390 395 400

Asp Arg Ile Val Pro Val Ala Val Lys Ile Asn Met Asn Met Phe His  
 405 410 415

Gly Thr Thr Ala Arg Gly Asn Lys Ala Met Asp Pro Phe Phe Ala Tyr  
 420 425 430

Met Asn Pro Arg Pro Thr Ile Glu Leu Lys Phe Leu Asp Glu Ile Pro  
 435 440 445

Lys Asn Met Thr Cys Gly Ser Gly Lys Ser Ser Ile Glu Val Ala Asn



ES 2 640 100 T3

450 455 460

Tyr Ile Gln His Leu Leu Ala Gly Glu Leu Gly Tyr Asp Cys Thr Asp  
465 470 475 480

Phe Thr Arg Lys Asp Lys Tyr Arg Leu Leu Ala Gly Asn Asp Gly Ile  
485 490 495

Val Pro Val Lys Glu Lys Thr Arg  
500

5 <210> 177  
<211> 529  
<212> PRT  
<213> *Physcomitrella patens*  
  
<400> 177

ES 2 640 100 T3

Met Glu Ile Pro Gly Val Phe Asp Ala Asn Leu Phe Cys Ala Glu Ser  
 1 5 10 15

Phe Pro Glu Val Glu Thr Cys Lys Val Glu Gly Arg Glu Lys Gln Thr  
 20 25 30

Ile Ile Ser Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ser Phe  
 35 40 45

Pro Tyr Phe Met Leu Ile Ala Phe Asp Ala Gly Ser Pro Leu Arg Phe  
 50 55 60

Ile Val Leu Leu Leu Ala Ser Pro Ile Ile Trp Leu Val Tyr Asn Phe  
 65 70 75 80

Phe Ser Glu Ala Ala Gly Thr Lys Leu Leu Ile Phe Ile Ala Cys Ala  
 85 90 95

Gly Leu Lys Ala Ser Asp Ile Glu Ser Val Ala Arg Gly Val Leu Pro  
 100 105 110

Lys Phe Tyr Leu Glu Asp Met His Ser Val Ser Tyr Ser Ile Phe Val  
 115 120 125

Ser Cys Gly Lys Arg Tyr Val Val Thr Ala Asn Pro Arg Ile Met Val  
 130 135 140

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

Glu Leu His Ile Thr Lys Ser Gly Tyr Ala Thr Gly Leu Val Met Gly

ES 2 640 100 T3

				165					170					175				
Ala	Gly	Val	Ile	Val	Gly	Ala	Asn	Lys	Thr	Arg	Ala	Val	Lys	Lys	Tyr			
			180					185					190					
Phe	Gly	Asp	Asp	Leu	Pro	Asp	Ile	Ala	Leu	Gly	Asp	Arg	Ala	Ser	Asp			
		195					200					205						
Phe	Pro	Phe	Met	Ala	Leu	Cys	Lys	Glu	Ala	Tyr	Leu	Val	Pro	Ser	Phe			
	210					215					220							
Lys	Pro	Val	Glu	Pro	Val	Ala	Lys	Ala	Asp	Tyr	Leu	Lys	Pro	Leu	Ile			
225					230					235					240			
Phe	His	Asp	Gly	Arg	Leu	Val	Cys	Arg	Pro	Thr	Pro	Leu	Met	Ser	Leu			
				245					250					255				
Val	Ile	Ile	Leu	Trp	Ser	Pro	Ile	Gly	Leu	Ile	Leu	Ala	Leu	Ile	Arg			
			260					265					270					
Met	Val	Phe	Gly	Thr	Val	Leu	Pro	Met	Trp	Tyr	Ala	Leu	Pro	Leu	Glu			
		275					280					285						
Ala	Met	Leu	Gly	Val	Ser	Ile	Arg	Val	Lys	Gly	Ile	Pro	Pro	Ser	Cys			
	290					295					300							
Pro	Gln	Thr	Ser	Gly	Arg	Arg	Gly	Val	Leu	Phe	Val	Cys	Ser	His	Arg			
305					310					315					320			
Thr	Leu	Leu	Asp	Pro	Ile	Phe	Leu	Ser	Ile	Ala	Cys	Arg	Arg	Gln	Val			
				325					330					335				
Thr	Ala	Val	Thr	Tyr	Ser	Ile	Ser	Lys	Val	Ser	Glu	Ile	Leu	Ser	Pro			
			340					345					350					
Ile	Pro	Thr	Val	Arg	Leu	Thr	Arg	Cys	Arg	Lys	Thr	Asp	Ala	Glu	Thr			
		355					360					365						
Met	Thr	Lys	Leu	Leu	Asn	Glu	Gly	Asp	Leu	Ala	Val	Cys	Pro	Glu	Gly			
	370					375					380							
Thr	Thr	Cys	Arg	Glu	Pro	Tyr	Leu	Leu	Arg	Phe	Ser	Ser	Leu	Phe	Ala			
385					390					395					400			
Glu	Leu	Ala	Asp	Gln	Ile	Ile	Pro	Val	Thr	Met	Asn	Val	Lys	Thr	Thr			
				405					410					415				

ES 2 640 100 T3

Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe  
 420 425 430

Ile Phe Leu Met Asn Pro Ser Pro Arg Tyr Glu Leu Glu Phe Leu Asp  
 435 440 445

Gln Leu Pro His Glu Met Thr Cys Ser Gly Gly Lys Ser Ser His Glu  
 450 455 460

Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Ala Ser Leu Gly Phe Lys  
 465 470 475 480

Cys Thr Asn Leu Thr Arg Arg Asp Lys Tyr Arg Ile Leu Ala Gly Asn  
 485 490 495

Asp Gly Val Val Pro Ala Ser His Phe Leu Lys Gln Arg Ser Arg Ile  
 500 505 510

Phe Ala Cys Leu Pro Ser Arg Asn Ser Glu Arg Gly Gln Arg Ser Asp  
 515 520 525

Pro

- <210> 178
- <211> 503
- <212> PRT
- <213> *Vitis vinifera*

5

<400> 178

Met Ala Val Thr Thr Phe Pro Thr Val Asp Gln Cys Gln Ser Ile Gly  
 1 5 10 15

Arg Glu Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Leu Ile  
 20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu Val Ala Phe Glu Gly Gly  
 35 40 45

Gly Val Leu Arg Leu Leu Phe Leu Leu Leu Ala Ser Pro Leu Ala Gly  
 50 55 60

Leu Leu Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Gln Ile Leu Ile  
 65 70 75 80

Phe Val Thr Phe Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val Ala  
 85 90 95

10

ES 2 640 100 T3

Arg Ala Val Leu Pro Lys Phe Tyr Leu Ser Asp Leu His Pro Glu Ser  
 100 105 110

Trp Arg Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn  
 115 120 125

Pro Arg Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp  
 130 135 140

Leu Val Leu Gly Thr Glu Ile Ala Thr Tyr Lys Gly Arg Ala Thr Gly  
 145 150 155 160

Phe Val Cys Asp Pro Gly Val Leu Val Gly Lys Asn Lys Ala Tyr Ala  
 165 170 175

Leu Asn Lys Thr Phe Gly Glu Thr Gln Pro Glu Ile Gly Leu Gly Asp  
 180 185 190

Arg His Thr Asp Phe Pro Phe Met Ala Ser Cys Lys Glu Gly Tyr Ile  
 195 200 205

Val Ala Ala Lys Pro Glu Val Lys Ala Val Thr Ser Asp Lys Leu Pro  
 210 215 220

Lys Pro Ile Ile Phe His Asp Ser Arg Leu Val Gln Lys Pro Thr Pro  
 225 230 235 240

Leu Thr Ala Leu Leu Thr Leu Leu Trp Ile Pro Ile Gly Phe Ala Leu  
 245 250 255

Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Arg Ile Val  
 260 265 270

Tyr His Ala Phe Trp Ala Leu Gly Val Arg Val Ser Val Lys Gly Ser  
 275 280 285

Pro Pro Pro Val Ala Lys Lys Ser Ile Gly Gln Ser Gly Val Leu Phe  
 290 295 300

Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Ala Ala  
 305 310 315 320

Leu Gly Arg Pro Ile Pro Ala Val Thr Tyr Ser Val Ser Arg Leu Ser  
 325 330 335

Glu Phe Ile Ser Pro Ile Lys Thr Ile Arg Leu Ser Arg Asp Arg Val  
 340 345 350

Lys Asp Ala Ala Met Ile Lys Lys Leu Leu Glu Glu Gly Asp Leu Ala  
 355 360 365

Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe  
 370 375 380

Ser Ala Leu Phe Ala Glu Leu Thr Asp Gln Leu Val Pro Val Ala Met  
 385 390 395 400

Met Asn Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys  
 405 410 415

Gly Met Asp Pro Phe Phe Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu  
 420 425 430

Val Thr Phe Leu Asn Lys Leu Pro Leu Glu Leu Thr Cys Ser Ser Gly  
 435 440 445

Lys Ser Ser His Glu Val Ala Asn Tyr Met Gln Arg Val Ile Ala Ala  
 450 455 460

Thr Leu Ser Tyr Glu Cys Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg  
 465 470 475 480

Ala Leu Ala Gly Asn Asp Gly Thr Val Pro Lys Lys Pro Leu Leu Tyr  
 485 490 495

Pro Asp Lys Val Met Gly Cys  
 500

<210> 179  
 <211> 501  
 <212> PRT  
 <213> *Vitis vinifera*  
 <400> 179

Met Ser Pro Pro Pro Lys Arg Ala Arg Lys Phe Pro Ser Ile Thr Thr  
 1 5 10 15

Tyr Ala Gly Gly Asp His Arg Ser Ile Ala Ala Asp Leu Asp Gly Thr  
 20 25 30

Leu Leu Val Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val  
 35 40 45

Glu Ala Gly Ser Leu Leu Arg Gly Leu Phe Leu Leu Leu Ser Leu Pro  
 50 55 60

5

10

ES 2 640 100 T3

Ile Val Ile Val Ala Tyr Leu Phe Ile Ser Glu Glu Ile Gly Ile Gln  
65 70 75 80

Ile Leu Ile Phe Ile Ser Phe Ser Gly Leu Lys Ile Arg Asp Ile Glu  
85 90 95

Leu Ala Ser Gln Ala Val Leu Pro Arg Phe Tyr Ala Asn Asp Val Arg  
100 105 110

Arg Glu Ser Trp Glu Val Phe Glu Lys Cys Glu Arg Lys Val Val Val  
115 120 125

Thr Ala Asn Pro Thr Leu Met Val Glu Pro Phe Val Arg Asp Phe Leu  
130 135 140

Gly Gly Thr Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr  
145 150 155 160

Lys Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp  
165 170 175

Arg Lys Arg Leu Ala Leu Leu Lys Glu Phe Gly Asp Glu Leu Pro Asp  
180 185 190

Ile Gly Ile Gly Asp Arg Glu Ser Asp His Asp Phe Met Ser Ile Cys  
195 200 205

Lys Glu Gly Tyr Met Val Leu Pro Ser Lys Ser Ala Thr Pro Val Pro  
210 215 220

Pro Asn Arg Leu Lys Thr Pro Ile Ile Phe His Asp Gly Arg Phe Val  
225 230 235 240

Gln Pro Pro Thr Pro Leu Thr Ala Leu Ile Ile Tyr Leu Trp Leu Pro  
245 250 255

Phe Gly Phe Ala Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu  
260 265 270

Pro Glu Arg Ile Val Arg Tyr Thr Tyr Pro Met Leu Gly Ile Asn Leu  
275 280 285

Val Ile Arg Gly Asn Pro Pro Pro Pro Ser Pro Gly Ser Pro Gly  
290 295 300

Asn Leu Tyr Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Val Ile  
305 310 315 320

Ala Ile Ala Leu Arg Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser  
 325 330 335

Arg Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg  
 340 345 350

Asp Arg Ala Ala Asp Ala Ala Arg Ile Ser Ser Ile Leu Gln Lys Gly  
 355 360 365

Asp Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu  
 370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro  
 385 390 395 400

Val Ala Val Asn Val Lys Gln Asn Met Phe His Gly Thr Thr Val Arg  
 405 410 415

Gly Val Lys Phe Trp Asp Ala Tyr Phe Tyr Phe Met Asn Pro Arg Pro  
 420 425 430

Thr Tyr Glu Ile Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Cys  
 435 440 445

Lys Ala Gly Gly Lys Ser Ala Ile Glu Val Ala Asn His Val Gln Lys  
 450 455 460

Val Leu Gly Gly Val Leu Gly Phe Glu Cys Thr Gly Leu Thr Arg Lys  
 465 470 475 480

Asp Lys Tyr Met Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met  
 485 490 495

Tyr Asn Ala Lys Lys  
 500

<210> 180  
 <211> 510  
 <212> PRT  
 <213> *Vitis vinifera*

5

<400> 180

Met Thr Lys Ser Asn Leu His Ser Pro Pro Thr Ser Leu Thr Ile Asp  
 1 5 10 15

Gln Cys Thr Ser Ile Gly Arg Glu Lys Glu Thr Val Val Ala Asp Met  
 20 25 30

10



ES 2 640 100 T3

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

ES 2 640 100 T3

275		280		285											
Arg	Val	Tyr	Ile	Lys	Gly	Asn	Pro	Pro	Pro	Pro	Ala	Lys	Lys	Ser	Ile
	290					295					300				
Gly	Gln	Ser	Gly	Val	Leu	Phe	Ile	Cys	Ser	His	Arg	Thr	Leu	Leu	Asp
305					310					315					320
Pro	Ile	Phe	Leu	Ser	Ala	Ala	Leu	Gly	Arg	Pro	Ile	Pro	Ala	Val	Thr
				325					330					335	
Tyr	Ser	Leu	Ser	Arg	Leu	Ser	Glu	Phe	Ile	Ser	Pro	Ile	Lys	Thr	Val
			340					345					350		
Arg	Leu	Ser	Arg	Asp	Arg	Ala	Thr	Asp	Ala	Ala	Met	Ile	Lys	Lys	Leu
		355					360					365			
Leu	Glu	Ala	Gly	Asp	Leu	Ala	Ile	Cys	Pro	Glu	Gly	Thr	Thr	Cys	Arg
	370					375					380				
Glu	Pro	Phe	Leu	Leu	Arg	Phe	Ser	Ala	Leu	Phe	Ser	Glu	Leu	Thr	Asp
385					390					395					400
Glu	Leu	Val	Pro	Val	Ala	Met	Ala	Asn	Arg	Met	Thr	Met	Phe	His	Gly
				405					410					415	
Thr	Thr	Ala	Arg	Gly	Trp	Lys	Gly	Met	Asp	Pro	Phe	Tyr	Tyr	Phe	Met
			420					425					430		
Asn	Pro	Ser	Pro	Ala	Tyr	Glu	Val	Asn	Phe	Leu	Asn	Lys	Leu	Pro	His
		435					440					445			
Glu	Leu	Thr	Cys	Gly	Ala	Gly	Lys	Ser	Ser	His	Asp	Val	Ala	Asn	Tyr
	450					455					460				
Ile	Gln	Arg	Met	Ile	Ala	Ala	Ser	Leu	Ser	Tyr	Glu	Cys	Thr	Asn	Phe
465					470					475					480
Thr	Arg	Lys	Asp	Lys	Tyr	Arg	Ala	Leu	Ala	Gly	Asn	Asp	Gly	Thr	Val
				485					490					495	
Val	Glu	Lys	Pro	Lys	Val	Asp	Ala	Asn	Lys	Val	Met	Gly	Cys		
			500					505					510		

<210> 181  
 <211> 499  
 <212> PRT  
 <213> *Vitis vinifera*

ES 2 640 100 T3

<400> 181

Met Gly Ala Gln Arg His Phe Gln Pro Ile Ser Lys Cys Ser Ser Asp  
 1 5 10 15

Gly Arg Ala Asn Gln Thr Val Ala Ala Asp Leu Asp Gly Thr Leu Leu  
 20 25 30

Val Ser Arg Ser Ala Phe Pro Tyr Phe Met Leu Val Ala Leu Glu Ala  
 35 40 45

Gly Ser Leu Val Arg Ala Ile Leu Leu Leu Ala Ser Val Pro Leu Val  
 50 55 60

Tyr Phe Ile Tyr Leu Phe Ile Ser Glu Ser Leu Ala Ile Lys Ile Phe  
 65 70 75 80

Ile Tyr Met Ala Phe Ala Gly Leu Lys Val Arg Glu Ile Glu Leu Val  
 85 90 95

Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Glu Asp Val His Pro Glu  
 100 105 110

Ser Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Ile Thr Ala  
 115 120 125

Asn Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Tyr Met Gly Ala  
 130 135 140

Asp Lys Val Leu Gly Thr Glu Leu Glu Val Ser Asn Ser Gly Arg Ala  
 145 150 155 160

Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Glu His Lys Lys  
 165 170 175

Ala Ala Val Gln Gln Glu Phe Gly Met Asn Leu Pro Asp Val Gly Leu  
 180 185 190

Gly Asp Arg Asp Thr Asp His Asp Phe Met Ser Leu Cys Lys Glu Gly  
 195 200 205

Tyr Met Val Pro Arg Thr Lys Cys Glu Pro Leu Pro Arg Asn Lys Leu  
 210 215 220

Leu Ser Pro Val Ile Phe His Glu Gly Arg Leu Val Gln Arg Pro Thr  
 225 230 235 240

ES 2 640 100 T3

Pro Val Ala Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly Ile Ile  
 245 250 255  
 Leu Ser Ile Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu Arg Ile  
 260 265 270  
 Val Arg Tyr Asn Tyr Lys Ile Leu Gly Ile Lys Leu Ile Val Lys Gly  
 275 280 285  
 Thr Pro Pro Pro Pro Pro Lys Gly Gly Gln Ser Gly Val Leu Phe Val  
 290 295 300  
 Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val Ala Leu  
 305 310 315 320  
 Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe Ser Glu  
 325 330 335  
 Leu Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg Glu Lys  
 340 345 350  
 Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu Val Ile  
 355 360 365  
 Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser  
 370 375 380  
 Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala Ile Asn  
 385 390 395 400  
 Thr Lys Gln Ser Val Phe Tyr Gly Thr Ser Thr Arg Gly Tyr Lys Leu  
 405 410 415  
 Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr Glu Ile  
 420 425 430  
 Thr Phe Leu Asn Gln Leu Pro Ala Glu Leu Thr Cys Lys Gly Gly Lys  
 435 440 445  
 Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln Arg Val Leu Ala Gly Thr  
 450 455 460  
 Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr Ala Thr  
 465 470 475 480  
 Leu Ala Gly Thr Asp Gly Ser Val Pro Phe Lys Asn Glu Lys Glu Lys  
 485 490 495

Glu Lys Ala

5  
 <210> 182  
 <211> 497  
 <212> PRT  
 <213> *Populus trichocarpa*  
 <400> 182

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

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

Leu Leu Val Ser Arg Ser Ala Phe Pro Tyr Phe Leu Leu Val Ala Ile  
 35 40 45

Glu Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Ala Ser Val Pro  
 50 55 60

Phe Val Tyr Phe Thr Tyr Leu Phe Ile Ser Glu Ala Met Ala Ile Lys  
 65 70 75 80

Thr Phe Ile Phe Ile Ala Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu  
 85 90 95

Leu Val Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Asp Asp Val His  
 100 105 110

Pro Glu Thr Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Val  
 115 120 125

Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Phe Leu  
 130 135 140

Gly Ala Asp Lys Val Leu Gly Thr Glu Leu Glu Ala Thr Lys Ser Gly  
 145 150 155 160

Arg Ala Thr Gly Phe Ile Lys Lys Pro Gly Val Leu Val Gly Asp His  
 165 170 175

Lys Arg Asp Ala Leu Leu Lys Glu Phe Gly Thr Asn Leu Pro Asp Leu  
 180 185 190

Gly Leu Gly Asp Arg Glu Thr Asp His Asp Phe Met Ser Ile Cys Lys  
 195 200 205

10

ES 2 640 100 T3

Glu Gly Tyr Met Val Pro Gly Thr Lys Cys Glu Pro Leu Ala Arg Asn  
 210 215 220

Lys Leu Leu Ser Pro Val Ile Phe His Asp Gly Arg Leu Val Gln Arg  
 225 230 235 240

Pro Thr Pro Leu Ala Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly  
 245 250 255

Ile Ile Leu Ser Ile Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu  
 260 265 270

Arg Leu Ala Trp Tyr Thr Tyr Lys Leu Leu Gly Ile Arg Val Ile Val  
 275 280 285

Lys Gly Thr Pro Pro Pro Pro Pro Gly Lys Gly His Ser Gly Val Leu  
 290 295 300

Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val  
 305 310 315 320

Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe  
 325 330 335

Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg  
 340 345 350

Asp Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu  
 355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg  
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala  
 385 390 395 400

Ile Asn Thr Lys Gln Ser Val Phe His Gly Thr Thr Val Arg Gly His  
 405 410 415

Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr  
 420 425 430

Glu Val Thr Phe Leu Asn Gln Leu Pro Lys Glu Leu Thr Cys Lys Gly  
 435 440 445

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

Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr  
 465 470 475 480

Ala Lys Leu Ala Gly Thr Asp Gly Arg Val Leu Ser Lys Lys Glu Asn  
 485 490 495

Ala

<210> 183

<211> 501

5 <212> PRT

<213> *Populus trichocarpa*

<400> 183

Met Val Ser Phe Pro Thr Val Asp Lys Cys Ala Ser Ile Gly Arg Glu  
 1 5 10 15

Lys His Ser Val Val Ala Asp Met Asp Gly Thr Leu Leu Arg Gly Arg  
 20 25 30

Ser Ser Phe Pro Tyr Phe Ala Leu Leu Ala Phe Glu Ala Gly Gly Ile  
 35 40 45

Phe Arg Leu Leu Phe Leu Leu Leu Asn Ser Pro Leu Ala Gly Leu Leu  
 50 55 60

Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Lys Val Leu Ile Phe Ala  
 65 70 75 80

Thr Cys Ala Gly Met Lys Leu Ser Asp Ile Glu Ser Val Ala Arg Ala  
 85 90 95

Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Ser Glu Ser Trp Arg  
 100 105 110

Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn Pro Arg  
 115 120 125

Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp Leu Val  
 130 135 140

Leu Gly Thr Glu Ile Leu Thr Tyr Lys Gly Arg Ala Thr Gly Phe Val  
 145 150 155 160

Gln Ser Pro Gly Val Leu Val Gly Lys Asn Lys Ala Asp Ala Leu Lys  
 165 170 175

10

ES 2 640 100 T3

Lys Ala Phe Gly Lys Thr Gln Pro Glu Ile Gly Leu Gly Asp Arg His  
180 185 190

Thr Asp Ala Pro Phe Met Ala Leu Cys Lys Glu Gly Tyr Ile Val Pro  
195 200 205

Pro Lys Pro Glu Val Glu Ala Val Thr Thr Asp Lys Leu Pro Lys Pro  
210 215 220

Val Ile Phe His Asp Gly Arg Leu Val Gln Lys Pro Thr Pro Leu Ser  
225 230 235 240

Ala Leu Leu Ile Ile Leu Trp Ile Pro Ile Gly Phe Ile Leu Ala Cys  
245 250 255

Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Pro Met Val Tyr Tyr  
260 265 270

Ala Phe Trp Ala Leu Gly Val Arg Val Asn Ile Lys Gly Thr Pro Pro  
275 280 285

Pro Pro Ala Lys Lys Ser Ile Gly Gln Ser Gly Val Leu Phe Val Cys  
290 295 300

Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr Ala Leu Gly  
305 310 315 320

Arg Pro Ile Pro Ala Val Thr Tyr Ser Leu Ser Arg Leu Ser Glu Ile  
325 330 335

Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg Ala Ala Asp  
340 345 350

Ala Ser Met Ile Lys Lys Leu Leu Glu Glu Gly Asp Leu Ala Ile Cys  
355 360 365

Pro Glu Arg Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala  
370 375 380

Leu Phe Ala Glu Leu Thr Asp Glu Leu Val Pro Val Ala Met Ala Asn  
385 390 395 400

Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met  
405 410 415

Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu Val Thr



ES 2 640 100 T3

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

5 <210> 184  
 <211> 501  
 <212> PRT  
 <213> *Populus trichocarpa*  
 <400> 184

ES 2 640 100 T3

Met Asp Ser Phe Pro Thr Val Asp Lys Cys Ala Ser Ile Gly Arg Glu  
 1 5 10 15

Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Leu Arg Gly Arg  
 20 25 30

Ser Ser Phe Pro Tyr Phe Ala Leu Leu Ala Phe Glu Ala Gly Gly Ile  
 35 40 45

Phe Arg Leu Ile Phe Leu Leu Leu Ala Ser Pro Leu Ala Gly Leu Leu  
 50 55 60

Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Lys Val Leu Ile Phe Ala  
 65 70 75 80

Thr Cys Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val Ala Arg Ala  
 85 90 95

Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Ser Glu Ser Trp Arg  
 100 105 110

Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala Asn Pro Arg  
 115 120 125

Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala Asp Leu Val

ES 2 640 100 T3

130						135										140
Leu	Gly	Thr	Glu	Met	Ser	Thr	Tyr	Lys	Gly	Arg	Ala	Thr	Gly	Phe	Val	
145					150					155					160	
Gln	Ser	Pro	Gly	Val	Leu	Val	Gly	Lys	Asn	Lys	Ala	Asp	Ala	Leu	Lys	
				165					170					175		
Lys	Ala	Phe	Gly	Met	Thr	Gln	Pro	Glu	Ile	Gly	Leu	Gly	Asp	Arg	His	
			180					185					190			
Thr	Asp	Ala	Pro	Phe	Met	Ala	Leu	Cys	Lys	Glu	Gly	Tyr	Ile	Val	Pro	
		195					200						205			
Pro	Lys	Pro	Glu	Val	Glu	Ala	Val	Thr	Thr	Asp	Lys	Leu	Pro	Lys	Pro	
	210					215						220				
Val	Ile	Phe	His	Asp	Gly	Arg	Leu	Val	Gln	Lys	Pro	Thr	Pro	Leu	Ser	
225					230					235					240	
Ala	Leu	Leu	Thr	Ile	Leu	Trp	Ile	Pro	Ile	Gly	Phe	Ile	Leu	Ala	Cys	
				245					250					255		
Leu	Arg	Ile	Ala	Ala	Gly	Ser	Leu	Leu	Pro	Met	Pro	Met	Val	Tyr	Tyr	
			260					265					270			
Ala	Phe	Trp	Ala	Leu	Gly	Val	Arg	Val	Ile	Ile	Lys	Gly	Thr	Pro	Pro	
		275					280					285				
Pro	Pro	Ala	Lys	Lys	Ser	Ile	Gly	Gln	Ser	Gly	Val	Leu	Phe	Ile	Cys	
	290					295					300					
Ser	His	Arg	Thr	Leu	Leu	Asp	Pro	Ile	Phe	Leu	Ser	Thr	Ala	Leu	Gly	
305					310					315					320	
Arg	Pro	Ile	Pro	Ala	Val	Thr	Tyr	Ser	Val	Ser	Arg	Leu	Ser	Glu	Ile	
				325					330					335		
Ile	Ser	Pro	Ile	Lys	Thr	Val	Arg	Leu	Ser	Arg	Asp	Arg	Ala	Thr	Asp	
			340					345					350			
Ala	Ser	Met	Ile	Lys	Lys	Leu	Leu	Glu	Glu	Gly	Asp	Leu	Ala	Ile	Cys	
		355					360					365				
Pro	Glu	Gly	Thr	Thr	Cys	Arg	Glu	Pro	Phe	Leu	Leu	Arg	Phe	Ser	Ala	
	370					375						380				

ES 2 640 100 T3

Leu Phe Ala Glu Leu Thr Asp Gln Leu Val Pro Val Ala Val Val Asn  
 385 390 395 400

Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met  
 405 410 415

Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr Glu Val Thr  
 420 425 430

Phe Leu Asn Arg Leu Pro Gln Glu Leu Thr Cys Thr Gly Gly Lys Ser  
 435 440 445

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

Ser Tyr Glu Cys Thr Ser Phe Thr Arg Arg Asp Lys Tyr Arg Ala Leu  
 465 470 475 480

Ala Gly Asn Asp Gly Thr Val Val Glu Lys Thr Lys Leu Gln Ala Asn  
 485 490 495

Lys Val Met Gly Cys  
 500

<210> 185  
 <211> 500  
 <212> PRT  
 <213> *Populus trichocarpa*

5

<400> 185

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

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

Leu Leu Val Ser Arg Ser Ala Phe Pro Tyr Phe Leu Leu Val Ala Ile  
 35 40 45

Glu Ala Gly Ser Leu Leu Arg Gly Leu Ile Leu Leu Ser Ser Val Pro  
 50 55 60

Phe Val Tyr Phe Thr Tyr Leu Ile Ile Ser Glu Glu Met Ala Ile Lys  
 65 70 75 80

Thr Phe Ile Phe Ile Ala Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu  
 85 90 95

10

ES 2 640 100 T3

Leu Val Ser Arg Ser Val Leu Pro Lys Phe Tyr Ala Asp Asp Val His  
 100 105 110  
 Pro Glu Thr Trp Arg Val Phe Asn Ser Phe Gly Lys Arg Tyr Ile Val  
 115 120 125  
 Thr Ala Asn Pro Arg Ile Met Val Glu Pro Phe Ala Lys Asn Phe Leu  
 130 135 140  
 Gly Ala Asp Lys Val Leu Gly Thr Glu Ile Gln Ala Thr Lys Ser Gly  
 145 150 155 160  
 Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp His  
 165 170 175  
 Lys Arg Asn Ala Leu Leu Lys Glu Phe Gly Thr Ser Leu Pro Asp Leu  
 180 185 190  
 Gly Leu Gly Asp Arg Glu Thr Asp His Asp Phe Met Ser Ile Cys Lys  
 195 200 205  
 Glu Gly Tyr Ile Val Pro Arg Ser Lys Cys Glu Pro Leu Pro Arg Asn  
 210 215 220  
 Lys Leu Leu Ser Pro Val Ile Phe His Glu Gly Arg Leu Val Gln Arg  
 225 230 235 240  
 Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu Trp Met Pro Ile Gly  
 245 250 255  
 Ile Ile Leu Ser Val Leu Arg Val Tyr Leu Asn Ile Pro Leu Pro Glu  
 260 265 270  
 Arg Ile Ala Trp Tyr Asn Tyr Lys Leu Leu Gly Ile Arg Val Val Val  
 275 280 285  
 Lys Gly Asn Pro Pro Pro Pro Pro Arg Lys Gly Gln Ser Gly Val Leu  
 290 295 300  
 Phe Val Cys Asn His Arg Thr Val Leu Asp Pro Val Val Thr Ala Val  
 305 310 315 320  
 Ala Leu Arg Arg Lys Ile Ser Cys Val Thr Tyr Ser Ile Ser Lys Phe  
 325 330 335  
 Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala Leu Ser Arg Glu Arg  
 340 345 350

ES 2 640 100 T3

Asp Arg Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu Glu Gly Asp Leu  
 355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg  
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile Val Pro Val Ala  
 385 390 395 400

Ile Asn Thr Lys Gln Ser Val Phe Asn Gly Thr Thr Val Arg Gly His  
 405 410 415

Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn Pro Met Pro Thr Tyr  
 420 425 430

Glu Ile Thr Phe Leu Asn Gln Leu Pro Thr Glu Leu Thr Cys Lys Gly  
 435 440 445

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

Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Lys Asp Lys Tyr  
 465 470 475 480

Ala Ile Leu Ala Gly Thr Asp Gly Arg Val Ala Thr Lys Lys Glu Lys  
 485 490 495

Ala Ser Glu Asn  
 500

- <210> 186
- <211> 524
- <212> PRT
- <213> *Sorghum bicolor*

5

<400> 186

Met Glu Arg Pro Ser Pro Thr Cys Leu Ala Ala Asn Gly Pro Arg Arg  
 1 5 10 15

Lys Ala Arg Thr Ile Pro Phe Pro Pro Pro Pro Ala Asp Leu Arg Thr  
 20 25 30

Gly Arg Leu Arg Ser Arg Gln Thr Val Ala Ala Asp Leu Asp Gly Thr  
 35 40 45

Leu Leu Leu Ser Arg Ser Ala Phe Pro Tyr Tyr Leu Leu Val Ala Leu  
 50 55 60

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Pro Gly Val Leu Phe Val Cys Asn His Arg Thr Thr Leu Asp Pro Val  
 325 330 335

Ala Val Ala Val Ala Leu Gly Arg Lys Val Arg Trp Val Val Thr Asp  
 340 345 350

Gly Ala Ser Ser Arg Phe Ser Glu Ala Val Ser Pro Val Met Thr Gly  
 355 360 365

Val Ala Leu Pro Val Pro Val Pro Ser Arg Glu Ser Asp Asp Asp Ala  
 370 375 380

Asp Ala Ala Ala Arg Ile Arg Arg Leu Leu Glu Glu Gly Asp Asp Val  
 385 390 395 400

Val Ile Phe Pro Glu Gly Thr Ile Cys Arg Glu Pro Phe Leu Leu Arg  
 405 410 415

Phe Gly Ala Leu Phe Ala Glu Val Thr Asp Arg Ile Val Pro Val Ala  
 420 425 430

Ile Gly Ala Arg Glu Gly Met Phe His Gly Ser Thr Ala Arg Gly Leu  
 435 440 445

Arg Arg Met Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr Tyr  
 450 455 460

Glu Val Thr Phe Leu Asn Gln Leu Pro Arg Glu Leu Thr Cys Gly Gly  
 465 470 475 480

Gly Arg Ser Pro Val Glu Val Ala Asn Tyr Val Gln Glu Val Leu Ala  
 485 490 495

Ala Gln Leu Gly Phe Asp Cys Thr Cys Thr Ser Lys Gln Ala Lys Thr  
 500 505 510

Gly Trp Ser Pro Ala Ala Pro Thr Val Ala Trp Arg  
 515 520

<210> 187  
 <211> 503  
 <212> PRT  
 <213> *Sorghum bicolor*

<400> 187

Met Val Ala Ser Tyr Arg Ser Arg Arg Phe Lys Pro Ile Glu Glu Cys  
 1 5 10 15

5

10



ES 2 640 100 T3

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

			260					265				270			
Glu	Arg	Ile	Val	Tyr	Tyr	Thr	Tyr	Lys	Leu	Met	Gly	Ile	Arg	Leu	Ile
		275					280					285			
Val	Lys	Gly	Asn	Pro	Pro	Pro	Pro	Pro	Lys	Lys	Gly	His	Pro	Gly	Val
	290						295				300				
Leu	Phe	Val	Cys	Asn	His	Arg	Thr	Val	Leu	Asp	Pro	Val	Glu	Val	Ala
305					310					315					320
Val	Ala	Leu	Arg	Arg	Lys	Val	Ser	Cys	Val	Thr	Tyr	Ser	Ile	Ser	Lys
				325					330					335	
Phe	Ser	Glu	Leu	Ile	Ser	Pro	Ile	Lys	Ala	Val	Ala	Leu	Ser	Arg	Glu
			340					345						350	
Arg	Glu	Lys	Asp	Ala	Glu	Asn	Ile	Arg	Arg	Leu	Leu	Glu	Glu	Gly	Asp
		355					360					365			
Leu	Val	Ile	Cys	Pro	Glu	Gly	Thr	Thr	Cys	Arg	Glu	Pro	Phe	Leu	Leu
	370					375					380				
Arg	Phe	Ser	Ala	Leu	Phe	Ala	Glu	Leu	Thr	Asp	Arg	Ile	Val	Pro	Val
385					390					395					400
Ala	Ile	Asn	Thr	Lys	Glu	Ser	Met	Phe	His	Gly	Ser	Thr	Val	Arg	Gly
				405					410					415	
Phe	Lys	Leu	Met	Asp	Pro	Tyr	Phe	Phe	Phe	Met	Asn	Pro	Arg	Pro	Thr
			420					425					430		
Tyr	Glu	Ile	Thr	Phe	Leu	Thr	Gln	Leu	Pro	Lys	Asp	Leu	Thr	Cys	Ser
		435					440					445			
Gly	Gly	Lys	Ser	Pro	Ile	Glu	Val	Ala	Asn	Tyr	Ile	Gln	Lys	Thr	Leu
	450					455					460				
Ser	Gly	Gln	Leu	Gly	Phe	Glu	Cys	Thr	Thr	Ile	Thr	Arg	Lys	Glu	Lys
465					470					475					480
Tyr	Gly	Ile	Leu	Ala	Gly	Thr	Asp	Gly	Arg	Val	Pro	Ser	Lys	Asn	Lys
				485					490					495	
Asp	Lys	Glu	Lys	Glu	Lys	Asn									
			500												

ES 2 640 100 T3

<210> 188  
 <211> 505  
 <212> PRT  
 <213> *Sorghum bicolor*

5

<400> 188

Met Ala Ala Ser Pro Phe Pro Thr Val Glu Lys Cys Ser Ser Thr Asp  
 1 5 10 15

Arg Ser Gly Asp Thr Val Val Ala Asp Leu Asp Gly Thr Leu Leu Cys  
 20 25 30

Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met Ala Phe Glu Thr Gly  
 35 40 45

Gly Val Leu Arg Leu Leu Leu Leu Ile Ala Leu Ala Pro Leu Ala Gly  
 50 55 60

Leu Leu Tyr Tyr Phe Val Ser Glu Pro Ala Gly Ile Gln Val Leu Ile  
 65 70 75 80

Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp Ile Glu Ala Val Ala  
 85 90 95

Arg Ala Val Leu Pro Lys Phe Tyr Cys Ala Asp Leu His Pro Glu Ser  
 100 105 110

Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys Val Leu Thr Ala Asn  
 115 120 125

Pro Arg Ile Met Val Glu Ala Phe Leu Lys Glu Tyr Val Gly Thr Asp  
 130 135 140

Val Val Val Gly Thr Glu Leu Val Val Trp Arg Gly Arg Ala Thr Gly  
 145 150 155 160

Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu Gln Lys Ala Asp Ala  
 165 170 175

Leu Arg Arg Thr Phe Gly Asp Asp Val Ala Pro Glu Val Gly Leu Gly  
 180 185 190

Asp Arg Lys Thr Asp Tyr Pro Phe Met Arg Leu Cys Lys Glu Gly Tyr  
 195 200 205

Val Val Pro Ala Thr Pro Lys Leu Lys Pro Val Pro Arg Glu Asn Leu  
 210 215 220

ES 2 640 100 T3

Pro Lys Pro Val Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Ser  
 225 230 235 240

Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro Ile Gly Phe Leu  
 245 250 255

Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Arg Met  
 260 265 270

Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val Thr Ile Arg Gly  
 275 280 285

Asn Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln Thr Gly Val Leu  
 290 295 300

Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr  
 305 310 315 320

Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser Val Ser Arg Leu  
 325 330 335

Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu Thr Arg Asp Arg  
 340 345 350

Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Thr Glu Gly Asp Leu  
 355 360 365

Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg  
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile Val Pro Val Ala  
 385 390 395 400

Met Glu Asn Gln Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp  
 405 410 415

Lys Gly Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Gly Tyr  
 420 425 430

Val Val Thr Phe Leu Asn Lys Leu Pro Ala Glu Leu Thr Cys Asn Gly  
 435 440 445

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

Ala Ser Thr Leu Ser Tyr Glu Cys Thr Asn Phe Thr Arg Lys Asp Lys  
 465 470 475 480

ES 2 640 100 T3

Tyr Lys Ala Leu Ala Gly Asn Asp Gly Thr Val Val Ser Lys Pro Asn  
 485 490 495

Ile Asp Lys Asn Lys Val Met Gly Cys  
 500 505

5 <210> 189  
 <211> 518  
 <212> PRT  
 <213> *Sorghum bicolor*

<400> 189

Met Glu Glu Ala Val Glu Ala Val Val Ala Ala Gly Val Glu Pro Phe  
 1 5 10 15

Pro Thr Val Asp Lys Cys Asp Ala Ser Gly Leu Gly Ala His Ser Ser  
 20 25 30

Val Val Ala Asp Leu Asp Gly Thr Leu Leu Arg Ser Arg Ser Ala Phe  
 35 40 45

Pro Tyr Tyr Ala Leu Val Ala Phe Glu Thr Gly Gly Val Pro Arg Leu  
 50 55 60

Leu Leu Leu Leu Leu Leu Ser Pro Leu Ala Ala Leu Leu Arg Phe Leu  
 65 70 75 80

Ala Ser Glu Ser Ala Cys Leu Arg Val Leu Val Phe Gly Ala Thr Ala  
 85 90 95

Gly Ala Arg Val Arg Asp Val Glu Ser Ala Ala Arg Ala Val Leu Pro  
 100 105 110

Arg Phe Tyr Ala Ala Asp Val His Pro Ala Ala Trp Arg Val Phe Ser  
 115 120 125

Ala Cys Ser Arg Arg Arg Val Val Leu Thr Ala Thr Pro Arg Leu Phe  
 130 135 140

Ala Glu Pro Phe Leu Arg Asp Cys Leu Gly Ala Asp Ala Val Ala Gly  
 145 150 155 160

Thr Glu Leu Ala Thr Trp Arg Gly Arg Ala Thr Gly Leu Val Asp Ala  
 165 170 175

Arg Arg Gly Gly Val Leu Val Gly Arg Ser Lys Ala Gln Ala Leu Arg  
 180 185 190

10

ES 2 640 100 T3

Glu Ile Leu Ala Ala Asp Gly Gly Asp Ala Pro Asp Val Gly Leu Gly  
 195 200 205  
 Asp Ser Arg Ser Asp Tyr Pro Phe Met Ser Met Cys Lys Glu Ala Tyr  
 210 215 220  
 Ile Val Pro Arg Ala Pro Val Glu Pro Ala Leu Pro Met Asp Gln Leu  
 225 230 235 240  
 Pro Arg Pro Val Ile Phe His Asp Gly Arg Leu Val Arg Arg Pro Thr  
 245 250 255  
 Pro Leu Val Ala Leu Val Val Val Leu Trp Phe Pro Val Gly Leu Ala  
 260 265 270  
 Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu Pro Met Pro Leu  
 275 280 285  
 Val Tyr Tyr Ala Phe Trp Ala Leu Gly Val Arg Val Val Val Arg Gly  
 290 295 300  
 Ala Pro Pro Pro Pro Arg Ala Glu Arg Ala Thr Gly Arg Arg Gly Val  
 305 310 315 320  
 Leu Phe Ala Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser  
 325 330 335  
 Thr Ala Leu Gly Arg Pro Val Ala Ala Val Thr Tyr Ser Leu Ser Arg  
 340 345 350  
 Leu Ser Glu Phe Leu Ser Pro Ile Arg Thr Val Arg Leu Ser Arg Asp  
 355 360 365  
 Arg Ala Thr Asp Ala Ala Met Ile Arg Asp Leu Leu Ala Glu Gly Asp  
 370 375 380  
 Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu  
 385 390 395 400  
 Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Ser Glu Val Val Pro Val  
 405 410 415  
 Ala Met Glu Asn Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly  
 420 425 430  
 Trp Lys Gly Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala  
 435 440 445

ES 2 640 100 T3

Tyr Val Val Thr Phe Leu Asn Lys Leu Pro Pro Glu Leu Thr Cys Ala  
 450 455 460

Gly Gly Arg Thr Ser His Glu Val Ala Asn Tyr Ile Gln Arg Leu Ile  
 465 470 475 480

Ala Ala Thr Leu Ser Tyr Glu Cys Thr Ser Leu Thr Arg Lys Asp Lys  
 485 490 495

Tyr Arg Ala Leu Ala Gly Asn Asp Gly Val Val Ala Thr Pro Lys Pro  
 500 505 510

Pro Pro Thr Val Ala Ser  
 515

<210> 190  
 <211> 501  
 <212> PRT  
 <213> *Ricinus communis*

5

<400> 190

Met Pro Ala Met Lys Gln Ala Gln Lys Phe Pro Ser Ile Thr Ser Cys  
 1 5 10 15

Ser Gly Thr Ser Tyr Glu Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
 20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu  
 35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Val Leu Leu Leu Ser Leu Pro Phe  
 50 55 60

Ile Ile Ile Ser Tyr Phe Phe Ile Ser Glu Ala Ile Gly Ile Gln Ile  
 65 70 75 80

Leu Ile Phe Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu  
 85 90 95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys  
 100 105 110

Glu Ser Tyr Glu Val Phe Asp Arg Cys Lys Arg Lys Val Val Val Thr  
 115 120 125

Ala Asn Pro Thr Ile Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly  
 130 135 140

10

ES 2 640 100 T3

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys  
145 150 155 160

Arg Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Lys Trp  
165 170 175

Lys Lys Leu Ala Ile Leu Lys Glu Phe Gly Glu Asp Ala Pro Asp Leu  
180 185 190

Gly Ile Gly Asp Arg Lys Thr Asp His Asp Phe Met Ser Ile Cys Lys  
195 200 205

Glu Gly Tyr Met Val Tyr His Ser Lys Lys Ala Ala Thr Pro Leu Pro  
210 215 220

Arg Asp Arg Leu Lys Ser Pro Ile Ile Phe His Asp Gly Arg Phe Val  
225 230 235 240

Gln Arg Pro Asp Pro Leu Asn Ala Leu Ala Thr Tyr Leu Trp Leu Pro  
245 250 255

Phe Gly Phe Met Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu  
260 265 270

Pro Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu  
275 280 285

Val Ile Arg Gly Tyr Pro Pro Pro Ala Pro Ser Arg Gly Thr Pro Gly  
290 295 300

Asn Leu Tyr Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile Val Ile  
305 310 315 320

Ala Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser  
325 330 335

Arg Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg  
340 345 350

Asp Arg Ala Ala Asp Ala Glu Arg Ile Thr Ala Leu Leu Gln Lys Gly  
355 360 365

Asp Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Gln Phe Leu  
370 375 380

Leu Arg Phe Ser Ala Leu Phe Ala Glu Met Ser Asp Arg Ile Val Pro



ES 2 640 100 T3

385                                    390                                    395                                    400

Val Ala Val Asn Cys Lys Gln Ser Met Phe Tyr Gly Thr Thr Val Arg  
                                  405                                    410                                    415

Gly Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro  
                                  420                                    425                                    430

Thr Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val  
                                  435                                    440                                    445

Lys Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys  
                                  450                                    455                                    460

Val Leu Gly Asp Val Leu Gly Phe Gln Cys Thr Gly Leu Thr Arg Lys  
465                                    470                                    475

Asp Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met  
                                  485                                    490                                    495

Tyr Asn Ser Lys Lys  
                                  500

5

<210> 191  
<211> 512  
<212> PRT  
<213> *Ricinus communis*  
  
<400> 191

ES 2 640 100 T3

Met Ala Asn Pro Lys Glu Thr Tyr Ala Pro Phe Ser Ser Phe Pro Asn  
 1 5 10 15

Ile Glu Lys Cys Thr Ser Val Gly Arg Gln Asn Asp Thr Val Val Thr  
 20 25 30

Asp Met Asp Gly Thr Leu Leu Cys Gly Arg Ser Pro Phe Pro Tyr Phe  
 35 40 45

Ala Leu Val Ala Tyr Glu Ala Ser Gly Ile Leu Arg Leu Leu Phe Leu  
 50 55 60

Leu Leu Gly Ser Pro Ile Ala Gly Ile Leu Arg Tyr Phe Ile Ser Glu  
 65 70 75 80

Ser Ala Gly Ile Arg Val Leu Val Phe Ala Thr Phe Ala Gly Met Arg  
 85 90 95

Val Ser Asp Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr

ES 2 640 100 T3

			100					105					110				
Ser	Ser	Asp	Leu	His	Pro	Glu	Thr	Trp	Arg	Val	Phe	Ser	Ser	Cys	Gly		
		115					120					125					
Lys	Lys	Cys	Val	Ile	Thr	Glu	Asn	Pro	Arg	Ile	Met	Val	Glu	Ala	Phe		
	130					135					140						
Leu	Lys	Glu	Phe	Val	Gly	Ala	Asp	Met	Val	Ile	Gly	Thr	Glu	Ile	Val		
145					150					155					160		
Val	His	Lys	Gly	Arg	Ala	Thr	Gly	Phe	Ile	Arg	Ser	Pro	Gly	Val	Leu		
				165					170					175			
Val	Gly	Gln	Asn	Lys	Ala	Asp	Ala	Leu	Lys	Met	Ala	Phe	Tyr	Asp	Ala		
			180					185					190				
Pro	Val	Pro	Asn	Ile	Gly	Ile	Gly	Asp	Arg	Arg	Thr	Asp	Arg	Pro	Phe		
		195					200					205					
Met	Lys	Leu	Cys	Lys	Glu	Ser	Tyr	Ile	Val	Pro	Ser	Asn	Ser	Gln	Val		
	210					215						220					
Glu	Pro	Val	Gly	Pro	Lys	Lys	Leu	Pro	Lys	Pro	Ile	Val	Phe	His	Asp		
225					230					235					240		
Gly	Arg	Leu	Val	Gln	Lys	Pro	Thr	His	Leu	Ile	Ala	Leu	Ala	Thr	Leu		
				245					250					255			
Leu	Trp	Ile	Pro	Val	Gly	Phe	Ile	Leu	Ala	Cys	Leu	Arg	Leu	Ala	Ala		
			260					265					270				
Gly	Ala	Leu	Leu	Pro	Met	Pro	Leu	Val	Tyr	Tyr	Ala	Leu	Trp	Ala	Leu		
		275					280					285					
Gly	Val	Arg	Ile	His	Ile	Lys	Gly	Thr	Pro	Pro	Pro	Pro	Ala	Lys	Lys		
	290					295						300					
Ser	Thr	Gly	Gln	Ala	Gly	Val	Leu	Phe	Val	Cys	Ser	His	Arg	Thr	Leu		
305					310					315					320		
Leu	Asp	Pro	Val	Phe	Leu	Ser	Ile	Ala	Leu	Gly	Arg	Pro	Ile	Pro	Ala		
				325					330					335			
Val	Thr	Tyr	Ser	Leu	Ser	Arg	Leu	Ser	Glu	Phe	Ile	Ser	Pro	Ile	Lys		
			340					345						350			

Thr Val Arg Leu Thr Arg Asp Arg Val Thr Asp Ala Asn Met Ile Lys  
 355 360 365

Glu Leu Leu Gln Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr  
 370 375 380

Cys Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu  
 385 390 395 400

Thr Asp Glu Leu Val Pro Val Ala Met Ala Asn Lys Met Thr Met Phe  
 405 410 415

His Gly Thr Thr Ala Arg Gly Trp Lys Gly Met Asp Pro Phe Tyr Phe  
 420 425 430

Leu Met Asn Pro Ser Pro Ala Tyr Glu Val Thr Phe Leu Asn Lys Leu  
 435 440 445

Pro Tyr Asn Leu Thr Cys Gly Ala Gly Lys Ser Ser His Ser Val Ala  
 450 455 460

Asn Tyr Ile Gln Arg Thr Ile Ala Ala Thr Leu Ser Tyr Gln Cys Thr  
 465 470 475 480

Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn Asp Gly  
 485 490 495

Thr Val Val Glu Glu Pro Lys Gln Ala Arg Asp Glu Val Met Asp Ser  
 500 505 510

<210> 192  
 <211> 504  
 <212> PRT  
 <213> *Ricinus communis*

5

<400> 192

Met Val Met Val Asp Ser Phe Pro Glu Val Ser Lys Cys Ala Ser Lys  
 1 5 10 15

Gly Arg Glu Lys His Thr Val Val Ala Asp Met Asp Gly Thr Leu Phe  
 20 25 30

Ile Gly Arg Ser Ser Phe Pro Tyr Phe Ala Leu Ile Ala Phe Glu Ala  
 35 40 45

Gly Gly Ile Phe Arg Leu Leu Phe Leu Leu Leu Ala Ser Pro Leu Ala  
 50 55 60

10

ES 2 640 100 T3

Ala Leu Leu Tyr Tyr Phe Val Ser Glu Ser Ala Gly Ile Gln Val Leu  
65 70 75 80

Ile Phe Ala Ser Phe Ala Gly Met Lys Val Ser Asp Ile Glu Ser Val  
85 90 95

Ala Arg Ala Val Leu Pro Lys Phe Tyr Ser Ser Asp Leu His Pro Glu  
100 105 110

Ser Trp Arg Val Phe Ser Ser Cys Gly Lys Arg Cys Val Leu Thr Ala  
115 120 125

Asn Pro Arg Ile Met Val Glu Ala Phe Leu Lys Asp Phe Leu Gly Ala  
130 135 140

Asp Leu Val Phe Gly Thr Glu Ile Ser Thr Tyr Lys Gly Arg Ala Thr  
145 150 155 160

Gly Phe Val Arg Thr Pro Gly Val Leu Val Gly Lys Asn Lys Ala Asn  
165 170 175

Ala Leu Lys Lys Ala Phe Gly Asp Ala Arg Pro Glu Val Gly Leu Gly  
180 185 190

Asp Arg Arg Thr Asp Ala Pro Phe Met Ala Leu Cys Lys Glu Gly Tyr  
195 200 205

Leu Val Pro Pro Lys Pro Glu Val Arg Ala Val Thr Gly Asp Lys Leu  
210 215 220

Pro Lys Pro Ile Val Phe His Asp Gly Arg Leu Val Gln Lys Pro Thr  
225 230 235 240

Pro Leu Met Ala Leu Leu Ile Ile Leu Trp Ile Pro Ile Gly Phe Ile  
245 250 255

Leu Ala Cys Leu Arg Ile Ala Ala Gly Ser Leu Leu Pro Met Pro Met  
260 265 270

Val Tyr Tyr Ala Phe Leu Ala Leu Gly Val Arg Val Thr Val Lys Gly  
275 280 285

Asn Pro Pro Pro Pro Ala Lys Lys Ser Thr Gly Gln Ser Gly Val Leu  
290 295 300

Phe Val Cys Ser His Arg Thr Leu Leu Asp Pro Ile Phe Leu Ser Thr  
305 310 315 320

ES 2 640 100 T3

Ala Leu Gly Arg Pro Ile Ala Ala Val Thr Tyr Ser Val Ser Arg Leu  
 325 330 335

Ser Glu Ile Ile Ser Pro Ile Lys Thr Val Arg Leu Ser Arg Asp Arg  
 340 345 350

Ala Thr Asp Ala Ala Met Ile Lys Lys Leu Leu Val Glu Gly Asp Leu  
 355 360 365

Ala Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu Arg  
 370 375 380

Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Gln Ile Val Pro Val Ala  
 385 390 395 400

Met Val Asn Arg Met Ser Met Phe His Gly Thr Thr Ala Arg Gly Trp  
 405 410 415

Lys Gly Met Asp Pro Phe Tyr Phe Phe Met Asn Pro Ser Pro Ala Tyr  
 420 425 430

Glu Val Thr Phe Leu Ser Lys Leu Pro Gln Glu Leu Thr Cys Ser Ser  
 435 440 445

Gly Lys Ser Ser His Glu Val Ala Asn Tyr Ile Gln Arg Val Ile Ala  
 450 455 460

Ala Thr Leu Ser Tyr Glu Cys Thr Gly Phe Thr Arg Lys Asp Lys Tyr  
 465 470 475 480

Arg Ala Leu Ala Gly Asn Asp Gly Ser Val Val Glu Lys Pro Lys Leu  
 485 490 495

Glu Ala Asn Lys Val Met Gly Cys  
 500

<210> 193  
 <211> 498  
 <212> PRT  
 <213> *Arabidopsis lyrata*

5

<400> 193

Met Ser Pro Ala Lys Lys Ser Gly Asn Phe Pro Pro Ile Thr Glu Cys  
 1 5 10 15

Arg Asp Gly Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
 20 25 30

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn  
 290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala  
 305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg  
 325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp  
 340 345 350

Arg Val Thr Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp  
 355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu  
 370 375 380

Arg Phe Ser Ala Leu Phe Ser Glu Leu Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser  
 420 425 430

Tyr Glu Ala Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn  
 435 440 445

Gly Gly Gly Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val  
 450 455 460

Ile Gly Ala Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp  
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Met  
 485 490 495

Leu Gly

- <210> 194
- <211> 501
- <212> PRT
- <213> *Arabidopsis lyrata*

5

<400> 194



ES 2 640 100 T3

Met Gly Ala Gln Glu Lys Arg Arg Arg Phe Glu Pro Ile Ser Lys Cys  
1 5 10 15

Asp Val Lys Asp Arg Ser Asn His Thr Val Ala Ala Asp Leu Asp Gly  
20 25 30

Thr Leu Leu Ile Ser Arg Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala  
35 40 45

Leu Glu Ala Gly Gly Leu Leu Arg Ala Leu Ile Leu Leu Val Ser Val  
50 55 60

Pro Phe Val Tyr Leu Thr Tyr Leu Thr Val Ser Glu Thr Leu Ala Ile  
65 70 75 80

Asn Val Phe Val Phe Ile Thr Phe Ala Gly Leu Lys Ile Arg Asp Val  
85 90 95

Glu Leu Val Val Arg Ser Val Leu Pro Arg Phe Tyr Ala Glu Asp Val  
100 105 110

Arg Pro Asp Thr Trp Arg Ile Phe Asn Thr Phe Gly Lys Arg Tyr Ile  
115 120 125

Val Thr Ala Ser Pro Arg Ile Met Val Glu Pro Phe Val Lys Thr Tyr  
130 135 140

Leu Gly Val Asp Lys Val Leu Gly Thr Glu Leu Glu Val Ser Lys Ser  
145 150 155 160

Gly Arg Ala Thr Gly Phe Thr Arg Lys Pro Gly Ile Leu Val Gly Gln  
165 170 175

His Lys Arg Glu Val Val Leu Arg Glu Phe Gly Ser Leu Ala Ser Asp  
180 185 190

Leu Pro Asp Leu Gly Leu Gly Asp Ser Lys Thr Asp Tyr Asp Phe Met  
195 200 205

Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu Pro  
210 215 220

Leu Pro Arg Asn Lys Leu Leu Ser Pro Ile Ile Phe His Glu Gly Arg  
225 230 235 240

Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Leu Leu Trp

ES 2 640 100 T3

				245					250					255			
Leu	Pro	Ile	Gly	Phe	Leu	Leu	Ser	Leu	Ile	Arg	Val	Tyr	Thr	Asn	Ile		
			260					265					270				
Pro	Leu	Pro	Glu	Arg	Ile	Ala	Arg	Tyr	Asn	Tyr	Lys	Leu	Thr	Gly	Ile		
		275					280					285					
Lys	Leu	Ile	Val	Asn	Gly	His	Pro	Pro	Pro	Pro	Pro	Lys	Pro	Gly	Gln		
	290					295						300					
Pro	Gly	His	Leu	Leu	Val	Cys	Asn	His	Arg	Thr	Val	Leu	Asp	Pro	Val		
305					310					315					320		
Val	Thr	Ala	Val	Ala	Leu	Gly	Arg	Lys	Ile	Ser	Cys	Val	Thr	Tyr	Ser		
				325					330						335		
Ile	Ser	Lys	Phe	Ser	Glu	Leu	Ile	Ser	Pro	Ile	Lys	Ala	Val	Ala	Leu		
			340					345					350				
Thr	Arg	Gln	Arg	Glu	Lys	Asp	Ala	Ala	Asn	Ile	Lys	Arg	Leu	Leu	Glu		
		355					360					365					
Glu	Gly	Asp	Leu	Val	Ile	Cys	Pro	Glu	Gly	Thr	Thr	Cys	Arg	Glu	Pro		
	370					375						380					
Phe	Leu	Leu	Arg	Phe	Ser	Ala	Leu	Phe	Ala	Glu	Leu	Thr	Asp	Arg	Ile		
385					390					395					400		
Val	Pro	Val	Ala	Ile	Asn	Thr	Lys	Gln	Ser	Met	Phe	Asn	Gly	Thr	Thr		
				405					410						415		
Thr	Arg	Gly	Tyr	Lys	Leu	Leu	Asp	Pro	Tyr	Phe	Ala	Phe	Met	Asn	Pro		
			420					425						430			
Arg	Pro	Thr	Tyr	Glu	Ile	Thr	Phe	Leu	Lys	Gln	Ile	Pro	Ala	Glu	Leu		
		435					440						445				
Thr	Cys	Lys	Gly	Gly	Lys	Ser	Pro	Ile	Glu	Val	Ala	Asn	Tyr	Ile	Gln		
	450					455						460					
Arg	Val	Leu	Gly	Gly	Thr	Leu	Gly	Phe	Glu	Cys	Thr	Asn	Phe	Thr	Arg		
465					470					475					480		
Lys	Asp	Lys	Tyr	Ala	Met	Leu	Ala	Gly	Thr	Asp	Gly	Arg	Val	Pro	Val		
				485					490						495		

ES 2 640 100 T3

Lys Lys Glu Lys Thr  
500

5

<210> 195  
<211> 502  
<212> PRT  
<213> *Vernicia fordii*  
  
<400> 195

Met Ser Gln Thr Lys Pro Ala Pro Lys Phe Pro Ser Ile Thr Ser Cys  
1 5 10 15

Thr Gly Ser Ala Tyr Gln Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Val Glu  
35 40 45

Ala Gly Ser Leu Leu Arg Gly Leu Val Ser Leu Leu Ser Leu Pro Leu  
50 55 60

Val Ile Ile Ser Tyr Phe Phe Ile Ser Glu Ala Val Gly Ile Gln Ile  
65 70 75 80

Leu Ile Tyr Ile Ser Phe Ala Gly Leu Lys Ile Arg Asp Ile Glu Leu  
85 90 95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys  
100 105 110

Glu Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr  
115 120 125

Ala Asn Pro Thr Ile Met Val Glu Pro Phe Val Lys Asp Phe Leu Gly  
130 135 140

Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Lys  
145 150 155 160

Arg Ala Thr Gly Phe Val Lys Asn Pro Gly Val Leu Val Gly Lys Trp  
165 170 175

Lys Lys Leu Ser Ile Leu Lys Glu Phe Gly Glu Glu Ser Pro Asp Leu  
180 185 190

Gly Ile Gly Asp Arg Lys Thr Asp His Asp Phe Met Ser Ile Cys Lys  
195 200 205

10

ES 2 640 100 T3

Glu Gly Tyr Met Val Gln Arg Ser Lys Ser Ala Thr Pro Ile Pro Leu  
 210 215 220

Asp Arg Leu Lys Ser Arg Ile Ile Phe His Asp Gly Arg Phe Val Gln  
 225 230 235 240

Arg Pro Asp Pro Leu Asn Ala Leu Val Thr Tyr Leu Trp Leu Pro Phe  
 245 250 255

Gly Phe Ile Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro  
 260 265 270

Glu Arg Ile Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Val  
 275 280 285

Ile Arg Gly Asn Pro Pro Pro Ala Pro Ser Pro Gly Thr Pro Gly Asn  
 290 295 300

Leu Tyr Val Cys Asn His Arg Ser Ala Leu Asp Pro Ile Val Ile Ala  
 305 310 315 320

Ile Ala Leu Gly Arg Lys Val Ser Cys Val Thr Tyr Ser Val Ser Arg  
 325 330 335

Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg Asp  
 340 345 350

Arg Ala Ala Asp Ala Ala Arg Ile Ser Glu Leu Leu Gln Lys Gly Asp  
 355 360 365

Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Phe Leu Leu  
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Met Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Val Asn Cys Lys Gln Asn Met Phe Tyr Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Tyr Phe Phe Met Asn Pro Arg Pro Thr  
 420 425 430

Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Ala Lys  
 435 440 445

Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys Val  
 450 455 460

Leu Gly Asp Val Leu Gly Phe Glu Cys Thr Gly Leu Thr Arg Lys Asp  
 465 470 475 480

Lys Tyr Met Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met Tyr  
 485 490 495

Asn Thr Lys Lys Ala Gly  
 500

<210> 196  
 <211> 514  
 <212> PRT  
 <213> *Oryza sativa*

5

<400> 196

Met Gly Asp Ala Thr Met Ile Ala Ser Pro Phe Pro Ala Val Glu Lys  
 1 5 10 15

Cys Ser Ser Lys Asp Arg Gly Gly Asp Thr Ile Val Ala Asp Leu Asp  
 20 25 30

Gly Thr Leu Leu Cys Gly Arg Ser Ser Phe Pro Tyr Phe Ala His Met  
 35 40 45

Ala Phe Glu Thr Gly Gly Val Leu Arg Leu Leu Ala Leu Ile Leu Leu  
 50 55 60

Ala Pro Leu Ala Gly Val Leu Tyr Tyr Leu Val Ser Glu Ser Ala Gly  
 65 70 75 80

Ile Gln Val Leu Ile Phe Ala Ser Met Ala Gly Ala Arg Val Ala Asp  
 85 90 95

Val Glu Ala Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Cys Ser Asp  
 100 105 110

Ile His Pro Glu Ser Trp Arg Val Phe Ser Ala Cys Gly Arg Arg Cys  
 115 120 125

Val Leu Thr Ala Asn Pro Arg Met Met Val Glu Ala Phe Leu Lys Glu  
 130 135 140

Tyr Leu Gly Ala Asp Ile Val Val Gly Thr Glu Leu Val Val Trp Arg  
 145 150 155 160

Gly Arg Ala Thr Gly Leu Val Arg Ser Pro Gly Val Leu Val Gly Glu  
 165 170 175

10

ES 2 640 100 T3

Asn Lys Ala Ala Ala Leu Arg Glu Ala Phe Gly Asp Ala Ser Pro Glu  
 180 185 190  
 Ile Gly Ile Gly Asp Arg Arg Thr Asp Tyr Pro Phe Met Arg Leu Cys  
 195 200 205  
 Lys Glu Gly Tyr Val Val Pro Val Ser Pro Thr Pro Arg Pro Val Pro  
 210 215 220  
 Arg Glu Glu Leu Pro Lys Pro Val Val Phe His Asp Gly Arg Leu Val  
 225 230 235 240  
 Gln Lys Pro Ser Pro Ala Leu Ala Leu Leu Thr Val Leu Trp Ile Pro  
 245 250 255  
 Ile Gly Phe Val Leu Ala Cys Leu Arg Ile Ala Ala Gly Ala Leu Leu  
 260 265 270  
 Pro Met Arg Val Val Tyr His Ala Phe Arg Ala Leu Gly Val Arg Val  
 275 280 285  
 Thr Val Lys Gly Thr Pro Pro Pro Pro Ala Ser Arg Glu Thr Gly Gln  
 290 295 300  
 Ser Gly Val Leu Phe Ile Cys Ser His Arg Thr Leu Leu Asp Pro Ile  
 305 310 315 320  
 Phe Leu Ser Thr Ala Leu Gly Arg Pro Ile Thr Ala Val Thr Tyr Ser  
 325 330 335  
 Val Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile Arg Thr Val Arg Leu  
 340 345 350  
 Thr Arg Asp Arg Ala Ala Asp Ala Ala Met Ile Arg Arg Leu Leu Ala  
 355 360 365  
 Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 370 375 380  
 Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Glu Ile  
 385 390 395 400  
 Val Pro Val Ala Met Glu Asn Gln Met Ser Met Phe His Gly Thr Thr  
 405 410 415  
 Ala Arg Gly Trp Lys Ala Leu Asp Pro Phe Tyr Phe Phe Met Asn Pro  
 420 425 430

ES 2 640 100 T3

Ser Pro Gly Tyr Val Val Thr Phe Leu Ser Lys Leu Pro Gly Glu Leu  
 435 440 445

Thr Cys Asn Gly Gly Gly Gly Gly Gly Gly Arg Ser Ser His Glu  
 450 455 460

Val Ala Asn Tyr Ile Gln Arg Leu Ile Ala Ser Thr Leu Ser Tyr Glu  
 465 470 475 480

Cys Thr Ser Phe Thr Arg Lys Asp Lys Tyr Arg Ala Leu Ala Gly Asn  
 485 490 495

Asp Gly Thr Val Val Ser Lys Pro Asn Ile Asp Glu Lys Lys Ala Met  
 500 505 510

Gly Cys

- <210> 197
- <211> 503
- <212> PRT
- <213> *Arabidopsis thaliana*
- <400> 197

5

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys  
 1 5 10 15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
 20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu  
 35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Leu Ser Leu Pro Ile  
 50 55 60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile  
 65 70 75 80

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

Val Ser Arg Ala Val Leu Thr Arg Phe Tyr Ala Ala Asp Val Arg Lys  
 100 105 110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Lys Arg Lys Val Val Val  
 115 120 125

10

ES 2 640 100 T3

Thr Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu  
130 135 140

Gly Gly Asp Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr  
145 150 155 160

Met Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp  
165 170 175

Leu Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Asp Asp Ser Pro Asp  
180 185 190

Leu Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys  
195 200 205

Lys Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro  
210 215 220

Ile Glu Ser Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val  
225 230 235 240

Gln Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro  
245 250 255

Phe Gly Phe Met Leu Ser Val Phe Arg Val Tyr Phe Asn Leu Pro Leu  
260 265 270

Pro Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu  
275 280 285

Thr Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly  
290 295 300

Asn Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile  
305 310 315 320

Ala Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser  
325 330 335

Arg Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg  
340 345 350

Asp Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly  
355 360 365

Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu



ES 2 640 100 T3

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

<210> 198  
 <211> 500  
 <212> PRT  
 <213> *Populus trichocarpa*  
 <400> 198

5

ES 2 640 100 T3

Met Ser Pro Thr Lys Pro Ala Lys Lys Phe Pro Pro Ile Thr Ala Cys  
1 5 10 15

Asn Gly Thr Thr His Gln Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
20 25 30

Leu Val Ser Ser Ser Ser Phe Pro Tyr Phe Met Ile Val Ala Val Glu  
35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Val Leu Leu Leu Ser Leu Pro Ile  
50 55 60

Val Ile Val Ser Tyr Leu Phe Ile Ser Glu Ala Leu Gly Ile Gln Met  
65 70 75 80

Leu Ile Phe Ile Ser Met Ser Gly Leu Lys Ile Arg Asp Ile Glu Leu

ES 2 640 100 T3

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

ES 2 640 100 T3

Leu Ser Arg Phe Leu Ser Pro Ile Pro Ala Ile Ala Leu Thr Arg Asp  
 340 345 350

Arg Ala Ala Asp Ala Ala Arg Ile Ser Ser Ile Leu Gln Lys Gly Asp  
 355 360 365

Leu Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Phe Leu Leu  
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Val Asn Cys Lys Gln Asn Met Phe Tyr Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Thr  
 420 425 430

Tyr Glu Val Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Lys  
 435 440 445

Ala Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Val Gln Lys Val  
 450 455 460

Leu Gly Glu Val Leu Gly Phe Glu Asn Thr Gly Leu Thr Arg Lys Asp  
 465 470 475 480

Lys Tyr Leu Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Met His  
 485 490 495

Asn Ala Lys Lys  
 500

<210> 199  
 <211> 518  
 <212> PRT  
 <213> *Sorghum bicolor*  
 <400> 199

5

Met Ala Pro Ala Ala Ser Ala Arg Phe Pro Pro Val Ser Ser Tyr Asp  
 1 5 10 15

Ala Ser Ala Arg Thr Arg Arg Thr Ala Ala Ala Asp Leu Asp Gly Thr  
 20 25 30

Leu Leu Ala Ser Ser Ser Ala Phe Pro Tyr Tyr Phe Leu Val Ala Leu  
 35 40 45

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Pro Gly Ser Leu Leu Val Cys Asn His Arg Thr Ala Leu Asp Pro Ile  
 305 310 315 320

Ile Leu Ser Val Ala Leu Gly Arg Pro Val Ser Cys Val Thr Tyr Ser  
 325 330 335

Ala Ser Arg Leu Ser Thr Ala Ile Ser Pro Ile Arg Ala Val Ala Leu  
 340 345 350

Ser Arg Asp Arg Ala Thr Asp Ala Ala Arg Met Ala Ala Leu Leu Ala  
 355 360 365

Glu Gly Asp Val Val Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 370 375 380

Cys Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile  
 385 390 395 400

Val Pro Val Ala Met Glu Ala Arg Gln Gly Thr Tyr Tyr Gly Ser Thr  
 405 410 415

Ala Arg Gly Trp Lys Trp Leu Asp Pro Tyr Phe Phe Tyr Met Asn Pro  
 420 425 430

Arg Pro Gly Tyr Asp Val Thr Phe Leu Pro Pro Leu Arg Pro Glu Glu  
 435 440 445

Thr Cys Gly Ala Gly Gly Arg Ser Ala Val Asp Val Ala Asn His Val  
 450 455 460

Gln Arg Val Ile Ala Lys Glu Leu Gly Phe Gln Cys Thr Thr Leu Thr  
 465 470 475 480

Arg Lys Asp Lys Tyr Met Lys Leu Ala Gly Asn Asp Gly Ser Val Ala  
 485 490 495

Ala Arg Ala Lys Lys Asp Ala Ala Asp Asp Asn Ala Ala Thr Val Thr  
 500 505 510

Thr Thr Lys Lys Phe Val  
 515

<210> 200  
 <211> 293  
 <212> PRT  
 <213> *Ricinus communis*

5

<400> 200

ES 2 640 100 T3

Met Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu  
1 5 10 15

Pro Leu Pro Arg Asn Lys Leu Leu Ser Pro Val Ile Phe His Glu Gly  
20 25 30

Arg Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu  
35 40 45

Trp Met Pro Ile Gly Ile Ile Leu Ser Ile Leu Arg Val Tyr Leu Asn  
50 55 60

Ile Pro Leu Pro Glu Arg Ile Ala Tyr Tyr Asn Tyr Lys Ile Leu Gly  
65 70 75 80

Ile Lys Val Ile Val Lys Gly Thr Pro Pro Pro Ala Pro Arg Lys Gly  
85 90 95

Gln Ser Gly Val Leu Phe Val Cys Asn His Arg Thr Val Leu Asp Pro  
100 105 110

Val Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr  
115 120 125

Ser Ile Ser Lys Phe Thr Glu Ile Ile Ser Pro Ile Lys Ala Val Ala  
130 135 140

Leu Ser Arg Glu Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu  
145 150 155 160

Glu Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu  
165 170 175

Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg  
180 185 190

Ile Val Pro Val Ala Ile Asn Thr Lys Gln Thr Val Phe His Gly Thr  
195 200 205

Thr Val Arg Gly His Lys Leu Leu Asp Pro Tyr Phe Val Phe Met Asn  
210 215 220

Pro Met Pro Thr Tyr Glu Ile Thr Phe Leu Asn Gln Leu Pro Ile Glu  
225 230 235 240

Leu Thr Cys Lys Gly Gly Lys Ser Ser Ile Glu Val Ala Asn Tyr Ile  
245 250 255

ES 2 640 100 T3

Gln Arg Val Leu Ala Gly Thr Leu Gly Phe Glu Cys Thr Asn Leu Thr  
 260 265 270

Arg Lys Asp Lys Tyr Ala Ile Leu Ala Gly Thr Asp Gly Arg Val Pro  
 275 280 285

Ser Lys Lys Glu Lys  
 290

<210> 201

<211> 502

5 <212> PRT

<213> *Arabidopsis lyrata*

<400> 201

Met Ser Pro Ala Lys Lys Ser Arg Ser Phe Pro Pro Ile Ser Glu Cys  
 1 5 10 15

Lys Ser Arg Glu Tyr Asp Ser Ile Ala Ala Asp Leu Asp Gly Thr Leu  
 20 25 30

Leu Leu Ser Arg Ser Ser Phe Pro Tyr Phe Met Leu Val Ala Ile Glu  
 35 40 45

Ala Gly Ser Leu Phe Arg Gly Leu Ile Leu Leu Ser Leu Pro Ile  
 50 55 60

Val Ile Ile Ala Tyr Leu Phe Val Ser Glu Ser Leu Gly Ile Gln Ile  
 65 70 75 80

Leu Ile Tyr Ile Ser Phe Ala Gly Ile Lys Ile Lys Asn Ile Glu Leu  
 85 90 95

Val Ser Arg Ala Val Leu Pro Arg Phe Tyr Ala Ala Asp Val Arg Lys  
 100 105 110

Asp Ser Phe Glu Val Phe Asp Lys Cys Lys Arg Lys Val Val Val Thr  
 115 120 125

Ala Asn Pro Ile Val Met Val Glu Pro Phe Val Lys Asp Tyr Leu Gly  
 130 135 140

Gly Glu Lys Val Leu Gly Thr Glu Ile Glu Val Asn Pro Lys Thr Met  
 145 150 155 160

Lys Ala Thr Gly Phe Val Lys Lys Pro Gly Val Leu Val Gly Asp Leu  
 165 170 175



ES 2 640 100 T3

Lys Arg Leu Ala Ile Leu Lys Glu Phe Gly Glu Glu Ser Pro Asp Ile  
 180 185 190

Gly Leu Gly Asp Arg Thr Ser Asp His Asp Phe Met Ser Ile Cys Lys  
 195 200 205

Glu Gly Tyr Met Val His Glu Thr Lys Ser Ala Thr Thr Val Pro Ile  
 210 215 220

Glu Arg Leu Lys Asn Arg Ile Ile Phe His Asp Gly Arg Leu Val Gln  
 225 230 235 240

Arg Pro Thr Pro Leu Asn Ala Leu Ile Ile Tyr Leu Trp Leu Pro Phe  
 245 250 255

Gly Phe Met Leu Ser Ile Phe Arg Val Tyr Phe Asn Leu Pro Leu Pro  
 260 265 270

Glu Arg Phe Val Arg Tyr Thr Tyr Glu Ile Leu Gly Ile His Leu Thr  
 275 280 285

Ile Arg Gly His Arg Pro Pro Pro Pro Ser Pro Gly Lys Pro Gly Asn  
 290 295 300

Leu Tyr Val Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Ile Ala  
 305 310 315 320

Ile Ala Leu Gly Arg Lys Ile Thr Cys Val Thr Tyr Ser Val Ser Arg  
 325 330 335

Leu Ser Leu Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp  
 340 345 350

Arg Val Ala Asp Ala Ala Arg Met Arg Gln Leu Leu Glu Lys Gly Asp  
 355 360 365

Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu  
 370 375 380

Arg Phe Ser Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val  
 385 390 395 400

Ala Met Asn Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly  
 405 410 415

Val Lys Phe Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser

# ES 2 640 100 T3


<210> 202  
 <211> 585  
 <212> PRT  
 <213> *Arabidopsis thaliana*  
  
 <400> 202

5

ES 2 640 100 T3

Met Val Leu Pro Glu Leu Leu Val Ile Leu Ala Glu Trp Val Leu Tyr  
 1 5 10 15

Arg Leu Leu Ala Lys Ser Cys Tyr Arg Ala Ala Arg Lys Leu Arg Gly  
 20 25 30

Tyr Gly Phe Gln Leu Lys Asn Leu Leu Ser Leu Ser Lys Thr Gln Ser  
 35 40 45

Leu His Asn Asn Ser Gln His His Leu His Asn His His Gln Gln Asn  
 50 55 60

His Pro Asn Gln Thr Leu Gln Asp Ser Leu Asp Pro Leu Phe Pro Ser  
 65 70 75 80

Leu Thr Lys Tyr Gln Glu Leu Leu Leu Asp Lys Asn Arg Ala Cys Ser  
 85 90 95

Val Ser Ser Asp His Tyr Arg Asp Thr Phe Phe Cys Asp Ile Asp Gly  
 100 105 110

Val Leu Leu Arg Gln His Ser Ser Lys His Phe His Thr Phe Phe Pro  
 115 120 125

Tyr Phe Met Leu Val Ala Phe Glu Gly Gly Ser Ile Ile Arg Ala Ile

ES 2 640 100 T3

130		135		140											
Leu 145	Leu	Leu	Leu	Ser	Cys 150	Ser	Phe	Leu	Trp	Thr 155	Leu	Gln	Gln	Glu	Thr 160
Lys	Leu	Arg	Val	Leu 165	Ser	Phe	Ile	Thr	Phe 170	Ser	Gly	Leu	Arg	Val 175	Lys
Asp	Met	Asp	Asn 180	Val	Ser	Arg	Ser	Val 185	Leu	Pro	Lys	Phe	Phe 190	Leu	Glu
Asn	Leu	Asn 195	Ile	Gln	Val	Tyr	Asp 200	Ile	Trp	Ala	Arg	Thr 205	Glu	Tyr	Ser
Lys	Val 210	Val	Phe	Thr	Ser	Leu 215	Pro	Gln	Val	Leu	Val 220	Glu	Arg	Phe	Leu
Arg 225	Glu	His	Leu	Asn	Ala 230	Asp	Asp	Val	Ile	Gly 235	Thr	Lys	Leu	Gln	Glu 240
Ile	Lys	Val	Met	Gly 245	Arg	Lys	Phe	Tyr	Thr 250	Gly	Leu	Ala	Ser	Gly 255	Ser
Gly	Phe	Val	Leu 260	Lys	His	Lys	Ser	Ala 265	Glu	Asp	Tyr	Phe	Phe 270	Asp	Ser
Lys	Lys	Lys 275	Pro	Ala	Leu	Gly	Ile 280	Gly	Ser	Ser	Ser	Ser	Pro	Gln	Asp
His 290	Ile	Phe	Ile	Ser	Ile	Cys 295	Lys	Glu	Ala	Tyr	Phe 300	Trp	Asn	Glu	Glu
Glu 305	Ser	Met	Ser	Lys	Asn 310	Asn	Ala	Leu	Pro	Arg 315	Glu	Arg	Tyr	Pro	Lys 320
Pro	Leu	Ile	Phe	His 325	Asp	Gly	Arg	Leu	Ala 330	Phe	Leu	Pro	Thr 335	Pro	Leu
Ala	Thr	Leu	Ala 340	Met	Phe	Ile	Trp	Leu 345	Pro	Ile	Gly	Phe	Leu 350	Leu	Ala
Val	Phe	Arg 355	Ile	Ser	Val	Gly	Val 360	Phe	Leu	Pro	Tyr	His 365	Val	Ala	Asn
Phe 370	Leu	Ala	Ser	Met	Ser	Gly 375	Val	Arg	Ile	Thr	Phe 380	Lys	Thr	His	Asn

ES 2 640 100 T3

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

Cys Asn His Arg Thr Leu Leu Asp Pro Val Phe Leu Thr Thr Ser Leu  
405 410 415

Gly Lys Pro Leu Thr Ala Val Thr Tyr Ser Leu Ser Lys Phe Ser Glu  
420 425 430

Phe Ile Ala Pro Leu Lys Thr Val Ser Leu Lys Arg Asp Arg Lys Lys  
435 440 445

Asp Gly Glu Ala Met Gln Arg Leu Leu Ser Lys Gly Asp Leu Val Val  
450 455 460

Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser  
465 470 475 480

Pro Leu Phe Ala Glu Leu Thr Glu Asp Ile Val Pro Val Ala Val Asp  
485 490 495

Ala Arg Val Ser Met Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Cys  
500 505 510

Leu Asp Pro Ile Phe Phe Leu Met Asn Pro Arg Pro Val Tyr Cys Leu  
515 520 525

Glu Ile Leu Lys Lys Leu Pro Lys Glu Met Thr Cys Ala Gly Gly Lys  
530 535 540

Ser Ser Phe Glu Val Ala Asn Phe Ile Gln Gly Glu Leu Ala Arg Val  
545 550 555 560

Leu Gly Phe Glu Cys Thr Asn Leu Thr Arg Arg Asp Lys Tyr Leu Val  
565 570 575

Leu Ala Gly Asn Glu Gly Ile Val Arg  
580 585

<210> 203  
<211> 520  
<212> PRT  
<213> *Arabidopsis thaliana*

<400> 203

Met Ser Ala Lys Ile Ser Ile Phe Gln Ala Leu Val Phe Leu Phe Tyr  
1 5 10 15

5

10

ES 2 640 100 T3

Arg Phe Ile Leu Arg Arg Tyr Arg Asn Ser Lys Pro Lys Tyr Gln Asn  
 20 25 30

Gly Pro Ser Ser Leu Leu Gln Ser Asp Leu Ser Arg His Thr Leu Ile  
 35 40 45

Phe Asn Val Glu Gly Ala Leu Leu Lys Ser Asp Ser Leu Phe Pro Tyr  
 50 55 60

Phe Met Leu Val Ala Phe Glu Ala Gly Gly Val Ile Arg Ser Phe Leu  
 65 70 75 80

Leu Phe Ile Leu Tyr Pro Leu Ile Ser Leu Met Ser His Glu Met Gly  
 85 90 95

Val Lys Val Met Val Met Val Ser Phe Phe Gly Ile Lys Lys Glu Gly  
 100 105 110

Phe Arg Ala Gly Arg Ala Val Leu Pro Lys Tyr Phe Leu Glu Asp Val  
 115 120 125

Gly Leu Glu Met Phe Glu Val Leu Lys Arg Gly Gly Lys Lys Ile Gly  
 130 135 140

Val Ser Asp Asp Leu Pro Gln Val Met Ile Glu Gly Phe Leu Arg Asp  
 145 150 155 160

Tyr Leu Glu Ile Asp Val Val Val Gly Arg Glu Met Lys Val Val Gly  
 165 170 175

Gly Tyr Tyr Leu Gly Ile Met Glu Asp Lys Thr Lys His Asp Leu Val  
 180 185 190

Phe Asp Glu Leu Val Arg Lys Glu Arg Leu Asn Thr Gly Arg Val Ile  
 195 200 205

Gly Ile Thr Ser Phe Asn Thr Ser Leu His Arg Tyr Leu Phe Ser Gln  
 210 215 220

Phe Cys Gln Glu Ile Tyr Phe Val Lys Lys Ser Asp Lys Arg Ser Trp  
 225 230 235 240

Gln Thr Leu Pro Arg Ser Gln Tyr Pro Lys Pro Leu Ile Phe His Asp  
 245 250 255

Gly Arg Leu Ala Ile Lys Pro Thr Leu Met Asn Thr Leu Val Leu Phe  
 260 265 270

ES 2 640 100 T3

Met Trp Gly Pro Phe Ala Ala Ala Ala Ala Ala Ala Arg Leu Phe Val  
 275 280 285

Ser Leu Cys Ile Pro Tyr Ser Leu Ser Ile Pro Ile Leu Ala Phe Ser  
 290 295 300

Gly Cys Arg Leu Thr Val Thr Asn Asp Tyr Val Ser Ser Gln Lys Gln  
 305 310 315 320

Lys Pro Ser Gln Arg Lys Gly Cys Leu Phe Val Cys Asn His Arg Thr  
 325 330 335

Leu Leu Asp Pro Leu Tyr Val Ala Phe Ala Leu Arg Lys Lys Asn Ile  
 340 345 350

Lys Thr Val Thr Tyr Ser Leu Ser Arg Val Ser Glu Ile Leu Ala Pro  
 355 360 365

Ile Lys Thr Val Arg Leu Thr Arg Asp Arg Val Ser Asp Gly Gln Ala  
 370 375 380

Met Glu Lys Leu Leu Thr Glu Gly Asp Leu Val Val Cys Pro Glu Gly  
 385 390 395 400

Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Pro Leu Phe Thr  
 405 410 415

Glu Val Ser Asp Val Ile Val Pro Val Ala Val Thr Val His Val Thr  
 420 425 430

Phe Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Ala Leu Asp Pro Leu  
 435 440 445

Phe Phe Leu Leu Asp Pro Tyr Pro Thr Tyr Thr Ile Gln Phe Leu Asp  
 450 455 460

Pro Val Ser Gly Ala Thr Cys Gln Asp Pro Asp Gly Lys Leu Lys Phe  
 465 470 475 480

Glu Val Ala Asn Asn Val Gln Ser Asp Ile Gly Lys Ala Leu Asp Phe  
 485 490 495

Glu Cys Thr Ser Leu Thr Arg Lys Asp Lys Tyr Leu Ile Leu Ala Gly  
 500 505 510

Asn Asn Gly Val Val Lys Lys Asn  
 515 520

ES 2 640 100 T3

<210> 204  
 <211> 1330  
 <212> ADN  
 <213> *Arabidopsis thaliana*

5

<400> 204

```

    agttaaaga ttggttattt gggctctgca ctcaagtgag agagaagata gatagatctg      60
    agtagaatct tcgattcatt attcgttgtc gtcgttcatc tgtgagaagc ggacaaacca      120
    aagaatccac cggagctagt gatatgggtg gttccagaga gttccgagct gaggaacatt      180
    caaatcaatt ccaactctatc atcgccatgg ccatctggct tggcgccatt cacttcaacg      240
    tcgctcttgt tctctgttct ctcatthttcc ttctctcttc tctatctctc atggtcttgg      300
    gcttgctctc tctgthttatc thttatcccaa tcgatcatcg tagcaaatat ggtcgtaaagc      360
    tcgctaggta catatgcaag caocgctgta attatthccc cgtctctctg tacgctgagg      420
    attacgaagc thttccagcct aatcgtgcct atgtctthgg ttatgaaacca cattcgggtgc      480
    taccgattgg agthtggthgt cthttgtgac tcacagggtt tatgcctatt cctaacatta      540
    aagthcttgc aagtagtgcct atattctaca ctccccttct aaggcatata tggacatggt      600
    tagggctcac cgtgcttctc aggaagaatt tcacttccct thttggattct ggctacagtt      660
    gtgttcttgt acctggthgt gtgcaggaga cthttcatat gcaacatgat gctgagaatg      720
    thttcctthc aaggagaaga gthattgtgc gcatagccat ggaacagggg agcccctctgg      780
    thccagtatt ctgctthtgg caggcacgcg tgtacaaatg gtggaagccg gattgtgac      840
    thtatcttaa actatctaga gcaatcagat tcaccccgat ctgcttctgg ggagthtttg      900
    gatcaaccatt accgtgtcga cagcctatgc atgtggtcgt tggtaaacca atagaagtca      960
    caaaaactct gaagccaact gacgaagaga thgctaagtt tcatggccag tatgtggaag     1020
    cgcttaggga thtgtthtgg aggcacaagt cccgagtcgg ctatgatctt gagctgaaaa     1080
    thctthtgaac aaaatctcca atggaaataa thactthtgt gtatccttca thaatgtta     1140
    cctthggagct ggattthgac thaatataaa tgactacatc atgtagtcta catgtattgc     1200
    atgtctthtag catcgactgt tgaagtaatg gaatacgtth ataaagcctg thaatthacat     1260
    gtcgtcttgc acaagagtat gtgghaataa thaacattga cccaaaaata atactagthta     1320
    aaththttct
    
```

10 <210> 205  
 <211> 1381  
 <212> ADN  
 <213> *Ricinus communis*

15 <220>  
 <221> misc\_feature  
 <222> (1347)..(1347)  
 <223> n es a, c, g o t

20 <400> 205



ES 2 640 100 T3

actgggacac gaggacactg acatggactg aaggagtaga aaagggacgc aaaaaaagat 60  
 ttacatggc ggagaggtcg atgctgcttt attttcatta atacagaaag agtttgagca 120  
 ctgagcactg aatagcaaga agaagaaga gaagaaatgg ggaagaagc gaatcataat 180  
 aataataata ataatatcaa tagtaatgat gagaagaatg aagagaaatc aaattataca 240  
 gttgtaaatt cgagagaact atacccaacg aacatatttc acgcaactgtt agcgttgagc 300  
 atatggattg gttoaatoca tttcaatctc ttcttactct tcatctotta tctcttcott 360  
 tcttttccca cattcctoct gattgttga ttttttgtgg tgtaaatggt cattccgato 420  
 gacgaacaca gtaagttggg ccgtcgtttg tgcaggtatg tatgcagaca tgcgtgcagt 480  
 ctttttccgg taactctcca tgttgaagac atgaatgctt ttcattctga tctgtcttac 540  
 gtttttgggt atgagccaca ttcagtattt ccccttgggt tttctgtact atcagatcac 600  
 tttgctgtcc tgccccttcc taaaatgaag gtccttgcaa gtaacgctgt gtttcggaca 660  
 ccagttttaa ggcatatatg gacatggtgt ggtcttacat cagcaacaaa gaaaaatttc 720  
 actgccctcc tagcatctgg ttatagttgc attgtgattc ccggtggagt tcaagagaca 780  
 ttttatatga agcatggctc tgagattgct ttccttaagg cgagaagagg gtttgcgca 840  
 gtagctatgg agatgggtaa acccttggtt ccagttttct gctttggta atcgaacgtg 900  
 tacaagtggg gaaacctga tggcgagta tttatgaaa ttgctagagc tattaagttc 960  
 agoccaattg tcttttgggg agttctcgggt totcatttac cgtacaacg tccaatgcat 1020  
 gttgtcgtcg gtaaaccgat tgaggtgaag caaatccac agcctacagt ggaagaggtc 1080  
 tcagaagtac agggtcagtt tgttgcggca cttaaagatc tttttgaaag gcataaagca 1140  
 cgggttggct atgcagacct tacactgaa attccttgat agaagtcagt acttgttta 1200  
 gttcagttgg atgcctcttt cagtttttgg cgataatggt tgtattcatt agataattag 1260  
 gaaaatgaaa tgtaaacgg tgaataaagc ttttcaatat gcgggcatca caaagttggt 1320  
 ctagatatcc aaaaaaaaaa aaaaacnctg tatgccgtta cgtaccgtat cgttgacagc 1380  
 a 1381

<210> 206  
 <211> 1257  
 <212> ADN  
 <213> *Vernicia fordii*

5

<400> 206

ES 2 640 100 T3

```

agaagggcga tgcacatgat actctcctct taataagagt attgatacc agagaagcca      60
gaagcaagaa acaacaact tttgcaagaa ggagattaaa atggggatgg tggaagttaa      120

gaatgaagaa gaggttacca tattcaaatc tggagaaata taccoaaca acatttttca      180
atcagtgтта gctttggcta tatggcttgg ctctttocat ttcatoctct tcttagtctc      240
ttcttcaatc ttccttctct tctccaaatt tctcctcgtt attggggttc ttttgttttt      300
tatggtgatt cctatcaacg atcgcagtaa gcttggccaa tgcttgttca gctatatatc      360
taggcattgt tgcagctatt ttcccatcac tcttcatggt gaagatataa atgcctttcg      420
ttctgatcgt gcttacgtgt ttggttatga gccacattca gtttttccca ttggtgtgat      480
gattctttct ttgggtttaa ttcctcttcc taatataaaa ttcottgcaa gtagtgctgt      540
gttctacact ccatttttaa gacatatatg gtcattggtt ggtcttacac cagcaacaag      600
gaaaaatfff gtttccctcc tgtcttctgg ttatagttgc attctggtgc ctggtggagt      660
ccaagagaca ttttatatga agcaggactc tgagattgct ttccttaagg caagaagagg      720
atattattcga atagctatgc agacgggac acccttggtt cctgttttct gctttggtca      780
aatgcacaca ttcaaatggt ggaaacctga tggagagttg ttcattgaaa ttgctagagc      840
tattaaattc accccaacta tcttttgggg agttctcgga actcctttac ccttcaaaaa      900
tccaatgcat gtcgtggtgg gtagaccat tgaggtgaag caaaatccac agoctaccgc      960
tgaagaggtt gcggaagtac agagagagtt tattgcatca cttaaaaatc tctttgaaag     1020
gcacaaagca cgggtcggct attcagacct taaacttgaa attttttgac cagaaatcag     1080
aacttaaaag tagttttgca tgctttctca ttttcattta tttagttttt ttttaaacct     1140
tttttttcc cttttgttaa tttgtgttct tagataatga tgaaaatgag aaaaaaaaaa     1200
aatttgtaaa acagctaata agccttaatt ttgtggcaaa aaaaaaaaaa aaaaaaa      1257

```

- 5 <210> 207
- <211> 1280
- <212> ADN
- <213> *Mortierella ramanniana*
  
- 10 <220>
- <221> misc\_feature
- <222> (1265)..(1266)
- <223> n es a, c, g o t
  
- 15 <220>
- <221> misc\_feature
- <222> (1269)..(1269)
- <223> n es a, c, g o t
  
- 20 <220>
- <221> misc\_feature
- <222> (1275)..(1275)
- <223> n es a, c, g o t
  
- 25 <220>
- <221> misc\_feature
- <222> (1277)..(1277)

ES 2 640 100 T3

<223> n es a, c, g o t

<220>

<221> misc\_feature

5 <222> (1279)..(1279)

<223> n es a, c, g o t

<400> 207

tgcaaagtca cactttcagt cagccatggc cagcaaggat caacatttac agcagaaggt	60
caagcatacg ctagaagcta tcccatcccc togetatgct ccattgogag tgccattaag	120
acggagatta caaacattgg cagttttatt atgggtttcc atgatgtcaa tatgcatggt	180
catattcttc tttttatgct ccattcctgt tctcctttgg ttccccatta tcctttattt	240
gacctggatc ttggtgtggg ataaggcgcc agagaacggg ggaagaccta ttcgctggct	300
gcggaatgct gcttgggtgga agctgtttgc agggatattt cccgcacatg tcatcaagga	360
agccgattta gatccatoca agaactacat ctttggttat caccoccatg gaatcatatc	420
catgggctog ttctgtactt ttagtaccaa tgctactggc tttgatgact tgttcccagg	480
catccggcca tcgcttttga cattaacatc taattttaat atcccacttt atcgtgatta	540
tttgatggcg tgcggacttt gctccgtctc caaaacatoc tgtcaaaata ttttaaccaa	600
aggtggtccg ggccgtcca ttgccattgt cgtgggaggt gcttccgagt ctctcaatgc	660
tagaccgggt gtcatggacc ttgtgtgaa gagacgcttt ggttttatca agattgctgt	720
tcaaaccgggt gcaagtctag tgcccactat cagttttggt gaaaatgagc tgtacgaaca	780
gattgaaagc aatgaaaact caaagttgca tagatggcaa aagaagattc aacatgcoct	840
tggttttact atgccgctct ttcattggacg cgggtgtattc aattatgact ttggtttgc	900
ccccatgoc catcctatct acaogattgt tggaaagccc atccccgtcc ctagcatcaa	960
gtatggacag acaaaggatg agattataag agaactacat gactcgtaca tgcatgccgt	1020
gcaggatctc tatgatcggt acaaggatat ctatgcaaag gatcgggtaa aagaactaga	1080
attcgtogaa tagttgggtt gttaccattt aacgttttct tcaaattttt cattgattat	1140
acogttcat ccgacctgg tttttattg aatgggttct tcatctocag ttcaattccc	1200
actcctgcat gtatctatta tcatcatcct ttctaataat gaaaaccttt tgaacgtttc	1260
ataannaana gaaananana	1280

10

<210> 208

<211> 2485

<212> ADN

15 <213> *Homo sapiens*

<400> 208

tgccccgttg tgaggtgata aagtgttgcg ctocgggacg ccagcgccgc ggctgccgc	60
tctgctgggg tctaggctgt ttctctcgcg ccaccactgg ccgcccggcg cagctccagg	120

ES 2 640 100 T3

tgtcctagcc gccagcctc gacgccgtcc cgggaccctc gtgctctgcg cgaagccctg 180  
 gccccggggg ccggggcatg ggccaggggc gcggggtgaa gcggcttccc gcggggccgt 240  
 gactgggvcg gcttcagcca tgaagaccct catagccgcc tactccgggg tcttgvcvcg 300  
 cgagvcgtcag gccgaggctg accggagcca gcgctctcac ggaggacctg cvgtgtvcgvc 360  
 cgaggggtct gggagatggg gcactggatc cagcctctcc tccgccctcc aggacctctt 420  
 ctctgtcacc tggctcaata ggtccaaggt ggaaaagcag ctacaggtca tctcagtgct 480  
 ccagtgvggtc ctgtccttcc ttgtactggg agtggcctgc agtgccatcc tcatgtacat 540  
 attctgcact gattgctggc tcatvcgtgt gctctacttc acttgvcgtg tgtttgactg 600  
 gaacacaccc aagaaaggtg gcaggaggtc acagtgvggtc cgaaactggg ctgtgtvgvcg 660  
 ctactttcga gactactttc ccatccagct ggtgaagaca cacaacctgc tgaccaccag 720  
 gaactatata tttggatacc acccccatgg tatcatgggc ctgggtgcct tctgcaactt 780  
 cagcacagag gccacagaag tgagcaagaa gttcccaggc atacggcctt acctggctac 840  
 actggcagvc aacttccgaa tgctgtgtt gagggagtac ctgatgtctg gaggtatctg 900  
 ccctgtcagc cgggacacca tagactatth gctttcaaag aatgggagtg gcaatgctat 960  
 catcatvcgtg gtcgggggtg cvgctgagtc tctgagctcc atgcctggca agaatgvcgt 1020  
 caccctvcgvc aaccvcagvc gctttgtgaa actggccctg cvtcatggag ctgacctggt 1080  
 tcccatctac tctttggag agaatgaagt gtacaagcag gtgatcttvc aggagggctc 1140  
 ctggggccga tgggtccaga agaagttcca gaaatacatt ggtttcgcct catgcatctt 1200  
 ccatggctga gvcctcttct cctccgacac ctggggvcgtg gtgccctact ccaagcccat 1260  
 caccactgth gtgggagagc ccatcaccat cccaagctg gagcacccaa ccagcaaga 1320  
 catvcacctg taccacacca tgtacatgga ggcctvggtg aagctcttvc acaagcacia 1380  
 gaccaagttc gvcctccvcg agactgaggt cctggaggtg aactgagcca gccttccggg 1440  
 ccaattccct ggaggaacca gctgcaaatc acttttttgc tctgtaaatt tggaagtgtc 1500  
 atgggtgtct gtgggttatt taaaagaaat tataacaatt ttgctaaacc attacaatgt 1560  
 taggtctttt ttaagaagga aaaagtcagth atttcaagth ctttcacttc cagcttgcct 1620  
 tgttctaggt ggtggctaaa tctgggccta atctgggtgvc ctvcagctaac ctctcttctt 1680  
 cccttctga agtgacaaag gaaactcagth cttcttgggg aagaaggatt gccattagtg 1740  
 acttgvacca gttagatgat tcactttttg ccctagggga tgagagvcga aagccacttc 1800  
 tcatacaagc ccctttattg ccactacccc acgctvcgtct agtvcctgaaa ctvcagggacc 1860  
 agtttctctg ccaaggggag gagttggaga gcacagttgc cccgttgtgt gagggcagta 1920  
 gtaggcactct ggaatgctcc agtttgatct cccttctgccc accvcctacct caccvcctagth 1980  
 cactvcatct gcgvcctgga ctggvcctcca ggatgaggtat gggggtggca atgacacct 2040

ES 2 640 100 T3

```
gcaggggaaa ggactgcccc ccatgcacca ttgcaggag gatgccgcca ccatgagcta 2100
gggtggagtaa ctggtttttc ttgggtggct gatgacatgg atgcagcaca gactcagcct 2160
tggcctggag cacatgctta ctggtggcct cagtttacct tccccagatc ctagattctg 2220
gatgtgagga agagatocct cttcagaagg ggocctggcct tctgagcagc agattagttc 2280
caaagcaggt ggcccccgaa cccaagcctc acttttctgt gccttctga gggggttggg 2340
coggggagga aaccaaaccc tctcctgtgt gttctgttat ctcttgatga gatcattgca 2400
ccatgtcaga cttttgtata tgccttgaaa ataaatgaaa gtgagaatcc tcaaaaaaaaa 2460
aaaaaaaaaa aaaaaaaaaa aaaaa 2485
```

<210> 209  
<211> 1411  
<212> ADN  
<213> *Homo sapiens*  
  
<400> 209

5

ES 2 640 100 T3

actgtttctga gatctttgcc tocctcaggc tcccgagaat catggctcat tccaagcagc	60
ctagtcactt ccagagtctg atgcttctgc agtggccttt gagctacctt gccatctttt	120
ggatcttgca gccattgttc gtctacctgc tgtttacatc cttgtggccg ctaccagtgc	180
tttactttgc ctggttggtc ctggactgga agaccccaga gcgaggtggc aggcgttcgg	240
cctgggtaag gaactggtgt gtctggacct acatcagga ctatttcccc attacgatcc	300
tgaagacaaa ggacctatca cctgagcaca actacctcat gggggttcac ccccatggcc	360
tcoctgacct tggcgccttc tgcaacttct gcaactgagc cacaggcttc tcgaagacct	420
tcccaggoat cactcctcac ttggccaagc tgtcctgggt cttcaagatc ccctttgtta	480
ggagtaoct catggccaaa ggtgtgtgct ctgtgagcca gccagccatc aactatctgc	540
tgagccatgg cactggcaac ctcggtggca ttgtagtggg aggtgtgggt gaggccctgc	600
aaagtgtgcc caacaccacc accctcatcc tccagaagcg caaggggttc gtgcgcacag	660
ccctccagca tggggctcat ctggctccca ccttactttt tggggaaact gaggtgatg	720
atcaggtgct gttccataag gatagcagga tgtacaagtt ccagagctgc ttccgcogta	780
tctttggttt ctactgttgt gtcttctatg gacaaaagctt ctgtcaaggc tccactgggc	840
tcoctgccata ctccagcct attgtcactg tggttgggga gcctctgccca ctgccccaaa	900
ttgaaaagcc aagccaggag atgggtggaca aataccatgc actttatatg gatgctctgc	960
acaaaactggt cgaccagcat aagacccact atggctgctc agagacccaa aagctgtttt	1020
tcctgtgaat gaaggtactg catgcccagg agcacaggag tgccctgcctt gaagaagaga	1080
ctcatctgcc actaaccaaa gacaggcagg agatgagga ggttatatgt ggtaggggag	1140
ggcatgagga attccttctt tgccttcttg ccacagggtc cttacaggaa ttctttctga	1200
agagctgcac acagtcattc ctcaaaggag ggcaccttag tgccctcat gctggggcct	1260
gatgcctggt catcatttga gtcctctggg acacattagc agtcaactgca tattccccca	1320
gcccctgggc aactatctac tttctgtctc tgtggatttg ccaattctga acatttata	1380
taaatggaat aatacaataa aaaaaaaaaa a	1411

<210> 210  
 <211> 2488  
 <212> ADN  
 <213> *Bos taurus*

5

<400> 210

ES 2 640 100 T3

gccgtggccg tggccaagge tccgctttcc gccgggtgctc cggcccctct ggcggcccgt 60  
 cctcogcac cctcggcogt agcgggtccc gcgctccct gctcggcgcg aagccctggc 120  
 cccgggggco ggggcatggg ccagggggcg ggggtctggc ggcttccgc ggggccccga 180  
 cctgtactgg cttcgtcatg aagaccctca tagccgcta ctccggggtc ctgcgaggca 240  
 ctggctccag catcctctct gccctccagg acctgttttc tgtcacttgg ctcaataggt 300  
 ccaaggtaga gaagcagctc caagtcactc cgggtgctaca atgggtcctg tctttcctcg 360  
 tgctgggagt ggctgcagc gtcactctca tgtacacatt ctgcaccgat tgctggctca 420  
 ttgocgtgct ctacttcacc tggctggtgt ttgactggaa cacaccaag aaaggtggca 480  
 ggaggtcaca gtgggtccga aactgggtg tgtggcgcta ctttcgagac tactttocca 540  
 ttcagctggt gaagacacac aacttactga ccagcaggaa ctacatctt gggtaccatc 600  
 cccatggcat catgggcctg ggtgccttct gcaactcag cacagaggcc acagaagtaa 660  
 gcaagaagtt ccttggcata aggccctacc tggccacgct ggccggcaac ttccggatgc 720  
 cagtgtgcg ggagtacctg atgtctggag gcatctgcc agtgaaccgg gacaccatag 780  
 actacttgct ttcaaagaat gggagtggca atgccatcat catcgtggtg gggggcgcgg 840  
 ctgaatcct gagctccatg cccggcaaga atgcagtcac cctgcgcaat cgcaagggct 900  
 ttgtgaaact ggocctgcgc catggagccg acctggttcc cacctactcc tttggggaga 960  
 atgaggtgta caagcaggtg atctttgagg agggctcctg gggccggtgg gtgcagaaga 1020  
 agttccagaa gtacattggc tttgccccat gcatcttcca tggctcaggc ctcttctcct 1080  
 ctgacacctg gggctggtg ccctactcca agccatcac cactgtcgtg ggcgagccca 1140  
 ttaccatccc caggctggag cggccgagc agcaggacat cgacctgtac caogccatgt 1200  
 acgtgcaagc cctggtgaag ctcttogacc agcataagac caagttcggc ctcccggaga 1260  
 ccgaggtcct ggaggtgaac tgagcctgtc tgcaggggoc agcttctggg aggaaccatc 1320  
 tgcaaatcgt tttctaccaa gttctcaagt gctttttgtt ctgtaaattt ggaagcatca 1380  
 tgggtgtctg tgggttattt aaaagaaatt ataataattt tgtaaaccg ttacaatgtt 1440

ES 2 640 100 T3

aggtcttttt taagaaggag aaaggaata tttcaagctc tctcagttcc agtttgctct 1500  
 gtccaggtg ggcgctgcc catctgggcc tttatggttt cacaaccctt cttctctcct 1560  
 cccccaaatt acagagaaaa ttcagtcctg gttaactggg gaagaaggac agtcattagt 1620  
 gactcgggcc agttagatta tttgcctttt ggccctgagg gatgaggggc agaagcctcc 1680  
 tctagtacaa acatctctat gccagctac ccaaggcttg actgcaggac taaaccctct 1740  
 tgccaaggaa agacacagag gaaacaggtg ttoccatgct gaagagtgca gggataagca 1800  
 cctggaatat tcaagcttga ttctcctttc tgccaccccc atggccccag tctctgaaat 1860  
 ctgagcctgg actggcctcc aaagagagga tcaggggtgc agtggtcctg tgctggagaa 1920  
 tggagctgct cagatgaacc atctcttggg ttatcaagcc ccatcatggt ctggttggag 1980  
 tgactggttt tcctcagggg gctgatgaca cagaccaga cccagcacag acctgcctt 2040  
 ggccctgccag tggcctcgtt ttgcttaacc cacatcctca gcagtggatg cgaggaaggg 2100  
 gtcccctttc ctgtttacag aggggcctga gcagttagct ggtaccagaa cagaagttcc 2160  
 ccgaacccaa acctcatgta tttgtgcctt ttctgagagg ggggccaggg gggacagcta 2220  
 gtactctgtg tactctgctt tctcttgatg agatcattat accatgtcag actttcgtat 2280  
 attcctaaat ggataaatga aaacaaggtg cctctatgag ttattgtag ggtcacatta 2340  
 gcatcttgct gctgatgatt ggggagacca ctgcccagc tgggctgcct gctaagccct 2400  
 ctccacttt ctocctcagc ccaaatggcc ctttgtatat tagttatcaa ttgctgtgta 2460  
 ataaattact ccaaaaactg gtggctgc 2488

- 5
- <210> 211
  - <211> 1330
  - <212> ADN
  - <213> *Mus musculus*
  - <400> 211



ES 2 640 100 T3

```

ggtcaggggc gcgpcgtgag gcpctttct gcacggccgt gacgtgcatt ggcttcagca      60
tgaagaccct catcgcogcc tactcogggg tcctgcgggg tgagcgtcgg gcggaagctg      120
cccgcagcga aaacaagaat aaaggatctg ccctgtcacg cgaggggtct gggcgatggg      180
gcactggctc cagcatcctc tcagccctcc aagacatctt ctctgtcacc tggctcaaca      240
gatctaaggt ggaaaaacag ctgcaggtca tctcagtact acaatgggtc ctatccttcc      300
tggtgctagg agtggcctgc agtgtcatcc tcatgtacac cttctgcaca gactgctggc      360
tgatagctgt gctctacttc acctggctgg catttgactg gaacacgccc aagaaaggtg      420
gcaggagatc gcagtgggtg cgaaactggg ccgtgtggcg ctacttccga gactactttc      480
ccatccagct ggtgaagaca cacaacctgc tgaccaccag gaactatatac tttggatacc      540
acccccatgg catcatgggc ctgggtgcct tctgtaactt cagcacagag gctactgaag      600
tcagcaagaa gtttcctggc ataaggccct atttggctac gttggctggg aacttccgga      660
tgectgtgct tocgaggtac ctgatgtctg gaggcctctg ccctgtcaac cgagacacca      720
tagactactt gctctccaag aatgggagtg gcaatgctat catcatcgtg gtgggaggtg      780
cagctgagtc cctgagctoc atgcctggca agaacgcagt cacctgaag aaccgcaaag      840
gctttgtgaa gctggccctg cgccatggag ctgatctggt toccacttat tcctttggag      900
agaatgaggt atacaagcag gtgatctttg aggagggttc ctggggccga tgggtccaga      960
agaagttcca gaagtatatt ggtttcgccc cctgcatctt ccatggccga ggcctcttct      1020
cctctgacac ctgggggctg gtgcctact ccaagcccat caccacgctc gtgggggagc      1080
ccatcaactgt cccaagctg gagcaccgca ccagaaaga catcgacctg taccatgccca      1140
tgtacatgga ggccctggtg aagctctttg acaatcaca gaccaaattt ggccttcag      1200
agactgaggt gctggaggtg aactgaccca gccgcgcgt gccagctcct gggagggagc      1260
actgcagatc cttttctacc gagttctoga gtgcattttg ttctgtaaat ttggaagcgt      1320
catgggtgct                                     1330

```

<210> 212  
 <211> 314  
 <212> PRT  
 <213> *Arabidopsis thaliana*  
 <400> 212

5

ES 2 640 100 T3

Met Gly Gly Ser Arg Glu Phe Arg Ala Glu Glu His Ser Asn Gln Phe  
 1 5 10 15

His Ser Ile Ile Ala Met Ala Ile Trp Leu Gly Ala Ile His Phe Asn  
 20 25 30

Val Ala Leu Val Leu Cys Ser Leu Ile Phe Leu Pro Pro Ser Leu Ser  
 35 40 45

Leu Met Val Leu Gly Leu Leu Ser Leu Phe Ile Phe Ile Pro Ile Asp  
 50 55 60

His Arg Ser Lys Tyr Gly Arg Lys Leu Ala Arg Tyr Ile Cys Lys His  
 65 70 75 80

Ala Cys Asn Tyr Phe Pro Val Ser Leu Tyr Val Glu Asp Tyr Glu Ala  
 85 90 95

Phe Gln Pro Asn Arg Ala Tyr Val Phe Gly Tyr Glu Pro His Ser Val  
 100 105 110

ES 2 640 100 T3

Leu Pro Ile Gly Val Val Ala Leu Cys Asp Leu Thr Gly Phe Met Pro  
 115 120 125

Ile Pro Asn Ile Lys Val Leu Ala Ser Ser Ala Ile Phe Tyr Thr Pro  
 130 135 140

Phe Leu Arg His Ile Trp Thr Trp Leu Gly Leu Thr Ala Ala Ser Arg  
 145 150 155 160

Lys Asn Phe Thr Ser Leu Leu Asp Ser Gly Tyr Ser Cys Val Leu Val  
 165 170 175

Pro Gly Gly Val Gln Glu Thr Phe His Met Gln His Asp Ala Glu Asn  
 180 185 190

Val Phe Leu Ser Arg Arg Arg Gly Phe Val Arg Ile Ala Met Glu Gln  
 195 200 205

Gly Ser Pro Leu Val Pro Val Phe Cys Phe Gly Gln Ala Arg Val Tyr  
 210 215 220

Lys Trp Trp Lys Pro Asp Cys Asp Leu Tyr Leu Lys Leu Ser Arg Ala  
 225 230 235 240

Ile Arg Phe Thr Pro Ile Cys Phe Trp Gly Val Phe Gly Ser Pro Leu  
 245 250 255

Pro Cys Arg Gln Pro Met His Val Val Val Gly Lys Pro Ile Glu Val  
 260 265 270

Thr Lys Thr Leu Lys Pro Thr Asp Glu Glu Ile Ala Lys Phe His Gly  
 275 280 285

Gln Tyr Val Glu Ala Leu Arg Asp Leu Phe Glu Arg His Lys Ser Arg  
 290 295 300

Val Gly Tyr Asp Leu Glu Leu Lys Ile Leu  
 305 310

<210> 213  
 <211> 340  
 <212> PRT  
 <213> *Ricinus communis*

<400> 213

Met Gly Glu Glu Ala Asn His Asn Asn Asn Asn Asn Ile Asn Ser  
 1 5 10 15

5

10

ES 2 640 100 T3

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

ES 2 640 100 T3

Phe Trp Gly Val Leu Gly Ser His Leu Pro Leu Gln Arg Pro Met His  
 275 280 285

Val Val Val Gly Lys Pro Ile Glu Val Lys Gln Asn Pro Gln Pro Thr  
 290 295 300

Val Glu Glu Val Ser Glu Val Gln Gly Gln Phe Val Ala Ala Leu Lys  
 305 310 315 320

Asp Leu Phe Glu Arg His Lys Ala Arg Val Gly Tyr Ala Asp Leu Thr  
 325 330 335

Leu Glu Ile Leu  
 340

<210> 214  
 <211> 322  
 <212> PRT  
 <213> *Vernicia fordii*

5

<400> 214

Met Gly Met Val Glu Val Lys Asn Glu Glu Glu Val Thr Ile Phe Lys  
 1 5 10 15

Ser Gly Glu Ile Tyr Pro Thr Asn Ile Phe Gln Ser Val Leu Ala Leu  
 20 25 30

Ala Ile Trp Leu Gly Ser Phe His Phe Ile Leu Phe Leu Val Ser Ser  
 35 40 45

Ser Ile Phe Leu Pro Phe Ser Lys Phe Leu Leu Val Ile Gly Leu Leu  
 50 55 60

Leu Phe Phe Met Val Ile Pro Ile Asn Asp Arg Ser Lys Leu Gly Gln  
 65 70 75 80

Cys Leu Phe Ser Tyr Ile Ser Arg His Val Cys Ser Tyr Phe Pro Ile  
 85 90 95

Thr Leu His Val Glu Asp Ile Asn Ala Phe Arg Ser Asp Arg Ala Tyr  
 100 105 110

Val Phe Gly Tyr Glu Pro His Ser Val Phe Pro Ile Gly Val Met Ile  
 115 120 125

Leu Ser Leu Gly Leu Ile Pro Leu Pro Asn Ile Lys Phe Leu Ala Ser  
 130 135 140

10

ES 2 640 100 T3

Ser Ala Val Phe Tyr Thr Pro Phe Leu Arg His Ile Trp Ser Trp Cys  
 145 150 155 160

Gly Leu Thr Pro Ala Thr Arg Lys Asn Phe Val Ser Leu Leu Ser Ser  
 165 170 175

Gly Tyr Ser Cys Ile Leu Val Pro Gly Gly Val Gln Glu Thr Phe Tyr  
 180 185 190

Met Lys Gln Asp Ser Glu Ile Ala Phe Leu Lys Ala Arg Arg Gly Phe  
 195 200 205

Ile Arg Ile Ala Met Gln Thr Gly Thr Pro Leu Val Pro Val Phe Cys  
 210 215 220

Phe Gly Gln Met His Thr Phe Lys Trp Trp Lys Pro Asp Gly Glu Leu  
 225 230 235 240

Phe Met Lys Ile Ala Arg Ala Ile Lys Phe Thr Pro Thr Ile Phe Trp  
 245 250 255

Gly Val Leu Gly Thr Pro Leu Pro Phe Lys Asn Pro Met His Val Val  
 260 265 270

Val Gly Arg Pro Ile Glu Val Lys Gln Asn Pro Gln Pro Thr Ala Glu  
 275 280 285

Glu Val Ala Glu Val Gln Arg Glu Phe Ile Ala Ser Leu Lys Asn Leu  
 290 295 300

Phe Glu Arg His Lys Ala Arg Val Gly Tyr Ser Asp Leu Lys Leu Glu  
 305 310 315 320

Ile Phe

- <210> 215
- <211> 355
- <212> PRT
- <213> *Mortierella ramanniana*

5

<400> 215

Met Ala Ser Lys Asp Gln His Leu Gln Gln Lys Val Lys His Thr Leu  
 1 5 10 15

Glu Ala Ile Pro Ser Pro Arg Tyr Ala Pro Leu Arg Val Pro Leu Arg  
 20 25 30

10

ES 2 640 100 T3

Arg Arg Leu Gln Thr Leu Ala Val Leu Leu Trp Cys Ser Met Met Ser  
 35 40 45

Ile Cys Met Phe Ile Phe Phe Phe Leu Cys Ser Ile Pro Val Leu Leu  
 50 55 60

Trp Phe Pro Ile Ile Leu Tyr Leu Thr Trp Ile Leu Val Trp Asp Lys  
 65 70 75 80

Ala Pro Glu Asn Gly Gly Arg Pro Ile Arg Trp Leu Arg Asn Ala Ala  
 85 90 95

Trp Trp Lys Leu Phe Ala Gly Tyr Phe Pro Ala His Val Ile Lys Glu  
 100 105 110

Ala Asp Leu Asp Pro Ser Lys Asn Tyr Ile Phe Gly Tyr His Pro His  
 115 120 125

Gly Ile Ile Ser Met Gly Ser Phe Cys Thr Phe Ser Thr Asn Ala Thr  
 130 135 140

Gly Phe Asp Asp Leu Phe Pro Gly Ile Arg Pro Ser Leu Leu Thr Leu  
 145 150 155 160

Thr Ser Asn Phe Asn Ile Pro Leu Tyr Arg Asp Tyr Leu Met Ala Cys  
 165 170 175

Gly Leu Cys Ser Val Ser Lys Thr Ser Cys Gln Asn Ile Leu Thr Lys  
 180 185 190

Gly Gly Pro Gly Arg Ser Ile Ala Ile Val Val Gly Gly Ala Ser Glu  
 195 200 205

Ser Leu Asn Ala Arg Pro Gly Val Met Asp Leu Val Leu Lys Arg Arg  
 210 215 220

Phe Gly Phe Ile Lys Ile Ala Val Gln Thr Gly Ala Ser Leu Val Pro  
 225 230 235 240

Thr Ile Ser Phe Gly Glu Asn Glu Leu Tyr Glu Gln Ile Glu Ser Asn  
 245 250 255

Glu Asn Ser Lys Leu His Arg Trp Gln Lys Lys Ile Gln His Ala Leu  
 260 265 270

Gly Phe Thr Met Pro Leu Phe His Gly Arg Gly Val Phe Asn Tyr Asp  
 275 280 285

ES 2 640 100 T3

Phe Gly Leu Leu Pro His Arg His Pro Ile Tyr Thr Ile Val Gly Lys  
 290 295 300

Pro Ile Pro Val Pro Ser Ile Lys Tyr Gly Gln Thr Lys Asp Glu Ile  
 305 310 315 320

Ile Arg Glu Leu His Asp Ser Tyr Met His Ala Val Gln Asp Leu Tyr  
 325 330 335

Asp Arg Tyr Lys Asp Ile Tyr Ala Lys Asp Arg Val Lys Glu Leu Glu  
 340 345 350

Phe Val Glu  
 355

- <210> 216
- <211> 388
- <212> PRT
- <213> *Homo sapiens*
- <400> 216

Met Lys Thr Leu Ile Ala Ala Tyr Ser Gly Val Leu Arg Gly Glu Arg  
 1 5 10 15

Gln Ala Glu Ala Asp Arg Ser Gln Arg Ser His Gly Gly Pro Ala Leu  
 20 25 30

Ser Arg Glu Gly Ser Gly Arg Trp Gly Thr Gly Ser Ser Ile Leu Ser  
 35 40 45

Ala Leu Gln Asp Leu Phe Ser Val Thr Trp Leu Asn Arg Ser Lys Val  
 50 55 60

Glu Lys Gln Leu Gln Val Ile Ser Val Leu Gln Trp Val Leu Ser Phe  
 65 70 75 80

Leu Val Leu Gly Val Ala Cys Ser Ala Ile Leu Met Tyr Ile Phe Cys  
 85 90 95

Thr Asp Cys Trp Leu Ile Ala Val Leu Tyr Phe Thr Trp Leu Val Phe  
 100 105 110

Asp Trp Asn Thr Pro Lys Lys Gly Gly Arg Arg Ser Gln Trp Val Arg  
 115 120 125

Asn Trp Ala Val Trp Arg Tyr Phe Arg Asp Tyr Phe Pro Ile Gln Leu  
 130 135 140



ES 2 640 100 T3

Val Lys Thr His Asn Leu Leu Thr Thr Arg Asn Tyr Ile Phe Gly Tyr  
145 150 155 160

His Pro His Gly Ile Met Gly Leu Gly Ala Phe Cys Asn Phe Ser Thr  
165 170 175

Glu Ala Thr Glu Val Ser Lys Lys Phe Pro Gly Ile Arg Pro Tyr Leu  
180 185 190

Ala Thr Leu Ala Gly Asn Phe Arg Met Pro Val Leu Arg Glu Tyr Leu  
195 200 205

Met Ser Gly Gly Ile Cys Pro Val Ser Arg Asp Thr Ile Asp Tyr Leu  
210 215 220

Leu Ser Lys Asn Gly Ser Gly Asn Ala Ile Ile Ile Val Val Gly Gly  
225 230 235 240

Ala Ala Glu Ser Leu Ser Ser Met Pro Gly Lys Asn Ala Val Thr Leu  
245 250 255

Arg Asn Arg Lys Gly Phe Val Lys Leu Ala Leu Arg His Gly Ala Asp  
260 265 270

Leu Val Pro Ile Tyr Ser Phe Gly Glu Asn Glu Val Tyr Lys Gln Val  
275 280 285

Ile Phe Glu Glu Gly Ser Trp Gly Arg Trp Val Gln Lys Lys Phe Gln  
290 295 300

Lys Tyr Ile Gly Phe Ala Pro Cys Ile Phe His Gly Arg Gly Leu Phe  
305 310 315 320

Ser Ser Asp Thr Trp Gly Leu Val Pro Tyr Ser Lys Pro Ile Thr Thr  
325 330 335

Val Val Gly Glu Pro Ile Thr Ile Pro Lys Leu Glu His Pro Thr Gln  
340 345 350

Gln Asp Ile Asp Leu Tyr His Thr Met Tyr Met Glu Ala Leu Val Lys  
355 360 365

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

Leu Glu Val Asn

385

5 <210> 217  
<211> 328  
<212> PRT  
<213> *Homo sapiens*  
  
<400> 217

ES 2 640 100 T3

Met Ala His Ser Lys Gln Pro Ser His Phe Gln Ser Leu Met Leu Leu  
1 5 10 15

Gln Trp Pro Leu Ser Tyr Leu Ala Ile Phe Trp Ile Leu Gln Pro Leu  
20 25 30

Phe Val Tyr Leu Leu Phe Thr Ser Leu Trp Pro Leu Pro Val Leu Tyr  
35 40 45

Phe Ala Trp Leu Phe Leu Asp Trp Lys Thr Pro Glu Arg Gly Gly Arg  
50 55 60

Arg Ser Ala Trp Val Arg Asn Trp Cys Val Trp Thr His Ile Arg Asp  
65 70 75 80

Tyr Phe Pro Ile Thr Ile Leu Lys Thr Lys Asp Leu Ser Pro Glu His  
85 90 95

Asn Tyr Leu Met Gly Val His Pro His Gly Leu Leu Thr Phe Gly Ala  
100 105 110

Phe Cys Asn Phe Cys Thr Glu Ala Thr Gly Phe Ser Lys Thr Phe Pro  
115 120 125

Gly Ile Thr Pro His Leu Ala Thr Leu Ser Trp Phe Phe Lys Ile Pro  
130 135 140

Phe Val Arg Glu Tyr Leu Met Ala Lys Gly Val Cys Ser Val Ser Gln  
145 150 155 160

Pro Ala Ile Asn Tyr Leu Leu Ser His Gly Thr Gly Asn Leu Val Gly  
165 170 175

Ile Val Val Gly Gly Val Gly Glu Ala Leu Gln Ser Val Pro Asn Thr  
180 185 190

Thr Thr Leu Ile Leu Gln Lys Arg Lys Gly Phe Val Arg Thr Ala Leu  
195 200 205

Gln His Gly Ala His Leu Val Pro Thr Phe Thr Phe Gly Glu Thr Glu

ES 2 640 100 T3

```

          210                215                220

Val Tyr Asp Gln Val Leu Phe His Lys Asp Ser Arg Met Tyr Lys Phe
225                230                235                240

Gln Ser Cys Phe Arg Arg Ile Phe Gly Phe Tyr Cys Cys Val Phe Tyr
                245                250                255

Gly Gln Ser Phe Cys Gln Gly Ser Thr Gly Leu Leu Pro Tyr Ser Arg
                260                265                270

Pro Ile Val Thr Val Val Gly Glu Pro Leu Pro Leu Pro Gln Ile Glu
                275                280                285

Lys Pro Ser Gln Glu Met Val Asp Lys Tyr His Ala Leu Tyr Met Asp
290                295                300

Ala Leu His Lys Leu Phe Asp Gln His Lys Thr His Tyr Gly Cys Ser
305                310                315                320

Glu Thr Gln Lys Leu Phe Phe Leu
                325
    
```

5

<210> 218  
 <211> 361  
 <212> PRT  
 <213> *Bos taurus*

<400> 218

ES 2 640 100 T3

Met Lys Thr Leu Ile Ala Ala Tyr Ser Gly Val Leu Arg Gly Thr Gly  
 1 5 10 15

Ser Ser Ile Leu Ser Ala Leu Gln Asp Leu Phe Ser Val Thr Trp Leu  
 20 25 30

Asn Arg Ala Lys Val Glu Lys Gln Leu Gln Val Ile Ser Val Leu Gln  
 35 40 45

Trp Val Leu Ser Phe Leu Val Leu Gly Val Ala Cys Ser Val Ile Leu  
 50 55 60

Met Tyr Thr Phe Cys Thr Asp Cys Trp Leu Ile Ala Val Leu Tyr Phe  
 65 70 75 80

Thr Trp Leu Val Phe Asp Trp Asn Thr Pro Lys Lys Gly Gly Arg Arg  
 85 90 95

Ser Gln Trp Val Arg Asn Trp Ala Val Trp Arg Tyr Phe Arg Asp Tyr

ES 2 640 100 T3

				100						105					110		
Phe	Pro	Ile	Gln	Leu	Val	Lys	Thr	His	Asn	Leu	Leu	Thr	Ser	Arg	Asn		
			115						120						125		
Tyr	Ile	Phe	Gly	Tyr	His	Pro	His	Gly	Ile	Met	Gly	Leu	Gly	Ala	Phe		
			130						135						140		
Cys	Asn	Phe	Ser	Thr	Glu	Ala	Thr	Glu	Val	Ser	Lys	Lys	Phe	Pro	Gly		
						150									155		160
Ile	Arg	Pro	Tyr	Leu	Ala	Thr	Leu	Ala	Gly	Asn	Phe	Arg	Met	Pro	Val		
									170						175		
Leu	Arg	Glu	Tyr	Leu	Met	Ser	Gly	Gly	Ile	Cys	Pro	Val	Asn	Arg	Asp		
									185						190		
Thr	Ile	Asp	Tyr	Leu	Leu	Ser	Lys	Asn	Gly	Ser	Gly	Asn	Ala	Ile	Ile		
									200						205		
Ile	Val	Val	Gly	Gly	Ala	Ala	Glu	Ser	Leu	Ser	Ser	Met	Pro	Gly	Lys		
															210		
Asn	Ala	Val	Thr	Leu	Arg	Asn	Arg	Lys	Gly	Phe	Val	Lys	Leu	Ala	Leu		
															225		240
Arg	His	Gly	Ala	Asp	Leu	Val	Pro	Thr	Tyr	Ser	Phe	Gly	Glu	Asn	Glu		
															245		255
Val	Tyr	Lys	Gln	Val	Ile	Phe	Glu	Glu	Gly	Ser	Trp	Gly	Arg	Trp	Val		
															260		270
Gln	Lys	Lys	Phe	Gln	Lys	Tyr	Ile	Gly	Phe	Ala	Pro	Cys	Ile	Phe	His		
															275		280
Gly	Arg	Gly	Leu	Phe	Ser	Ser	Asp	Thr	Trp	Gly	Leu	Val	Pro	Tyr	Ser		
															290		300
Lys	Pro	Ile	Thr	Thr	Val	Val	Gly	Glu	Pro	Ile	Thr	Ile	Pro	Arg	Leu		
															305		310
Glu	Arg	Pro	Thr	Gln	Gln	Asp	Ile	Asp	Leu	Tyr	His	Ala	Met	Tyr	Val		
															325		330
Gln	Ala	Leu	Val	Lys	Leu	Phe	Asp	Gln	His	Lys	Thr	Lys	Phe	Gly	Leu		
															340		345
																	350

ES 2 640 100 T3

Pro Glu Thr Glu Val Leu Glu Val Asn  
 355 360

5  
 <210> 219  
 <211> 388  
 <212> PRT  
 <213> *Mus musculus*  
 <400> 219

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

10

ES 2 640 100 T3

Met Ser Gly Gly Ile Cys Pro Val Asn Arg Asp Thr Ile Asp Tyr Leu  
 210 215 220

Leu Ser Lys Asn Gly Ser Gly Asn Ala Ile Ile Ile Val Val Gly Gly  
 225 230 235 240

Ala Ala Glu Ser Leu Ser Ser Met Pro Gly Lys Asn Ala Val Thr Leu  
 245 250 255

Lys Asn Arg Lys Gly Phe Val Lys Leu Ala Leu Arg His Gly Ala Asp  
 260 265 270

Leu Val Pro Thr Tyr Ser Phe Gly Glu Asn Glu Val Tyr Lys Gln Val  
 275 280 285

Ile Phe Glu Glu Gly Ser Trp Gly Arg Trp Val Gln Lys Lys Phe Gln  
 290 295 300

Lys Tyr Ile Gly Phe Ala Pro Cys Ile Phe His Gly Arg Gly Leu Phe  
 305 310 315 320

Ser Ser Asp Thr Trp Gly Leu Val Pro Tyr Ser Lys Pro Ile Thr Thr  
 325 330 335

Val Val Gly Glu Pro Ile Thr Val Pro Lys Leu Glu His Pro Thr Gln  
 340 345 350

Lys Asp Ile Asp Leu Tyr His Ala Met Tyr Met Glu Ala Leu Val Lys  
 355 360 365

Leu Phe Asp Asn His Lys Thr Lys Phe Gly Leu Pro Glu Thr Glu Val  
 370 375 380

Leu Glu Val Asn  
 385

<210> 220  
 <211> 3  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia conservada

<400> 220

Tyr Phe Pro  
 1

<210> 221



<211> 4  
 <212> PRT  
 <213> Secuencia artificial  
 5 <220>  
 <223> Secuencia conservada  
 <400> 221

**His Pro His Gly**  
**1**

10 <210> 222  
 <211> 4  
 <212> PRT  
 15 <213> Secuencia artificial  
 <220>  
 <223> Secuencia conservada  
 20 <400> 222

**Glu Pro His Ser**  
**1**

25 <210> 223  
 <211> 24  
 <212> PRT  
 <213> Secuencia artificial  
 30 <220>  
 <223> Secuencia conservada  
 35 <220>  
 <221> X  
 <222> (2)..(2)  
 <223> Cualquier aminoácido  
 40 <220>  
 <221> X  
 <222> (5)..(5)  
 <223> Cualquier aminoácido  
 45 <220>  
 <221> X  
 <222> (6)..(6)  
 <223> Lisina (K) o Arginina (R)  
 50 <220>  
 <221> X  
 <222> (7)..(7)  
 <223> Cualquier aminoácido  
 55 <220>  
 <221> X  
 <222> (9)..(11)  
 <223> Cualquier aminoácido  
 60 <220>  
 <221> X  
 <222> (13)..(15)  
 <223> Cualquier aminoácido  
 <220>

ES 2 640 100 T3

<221> X  
 <222> (16)..(16)  
 <223> Leucina (L) o Valina (V)

5

<220>  
 <221> X  
 <222> (19)..(21)  
 <223> Cualquier aminoácido

10

<220>  
 <221> X  
 <222> (24)..(24)  
 <223> Ácido glutámico (E) o Glutamina (Q)

15

<400> 223

**Arg Xaa Gly Phe Xaa Xaa Xaa Ala Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa**  
**1 5 10 15**

**Val Pro Xaa Xaa Xaa Phe Gly Xaa**  
**20**

<210> 224  
 <211> 8  
 <212> PRT  
 <213> Secuencia artificial

20

<220>  
 <223> Secuencia conservada

25

<220>  
 <221> X  
 <222> (3)..(3)  
 <223> Cualquier aminoácido

30

<220>  
 <221> X  
 <222> (5)..(7)  
 <223> Cualquier aminoácido

35

<400> 224

**Phe Leu Xaa Leu Xaa Xaa Xaa Asn**  
**1 5**

40

<210> 225  
 <211> 118  
 <212> PRT  
 <213> Secuencia artificial

45

<220>  
 <223> Secuencia conservada

50

<400> 225

ES 2 640 100 T3

Ala Leu Val Val Ala Asn His Gln Ser Phe Leu Asp Pro Leu Val Leu  
1 5 10 15

Ser Ala Leu Leu Pro Arg Lys Gly Gly Arg Val Arg Phe Val Ala Lys  
20 25 30

Lys Glu Leu Phe Tyr Val Pro Leu Leu Gly Trp Leu Leu Arg Leu Leu  
35 40 45

Gly Ala Ile Phe Ile Asp Arg Glu Asn Gly Arg Leu Ala Arg Ala Ala  
50 55 60

Leu Arg Glu Ala Val Arg Leu Leu Arg Asp Gly Gly Trp Leu Leu Ile  
65 70 75 80

Phe Pro Glu Gly Thr Arg Ser Arg Pro Gly Lys Leu Leu Pro Phe Lys  
85 90 95

Lys Gly Ala Ala Arg Leu Ala Leu Glu Ala Gly Val Pro Ile Val Pro  
100 105 110

Val Ala Ile Arg Gly Thr  
115

- 5 <210> 226
- <211> 187
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Secuencia conservada
- <220>
- <221> X
- <222> (15)..(15)
- <223> Cualquier aminoácido
- 15 <220>
- <221> X
- <222> (18)..(18)
- <223> Cualquier aminoácido
- 20 <220>
- <221> X
- <222> (23)..(23)
- <223> Cualquier aminoácido
- 25 <220>
- <221> X
- <222> (25)..(26)
- <223> Cualquier aminoácido
- 30 <220>
- <221> X
- <222> (28)..(30)
- <223> Cualquier aminoácido

5	<220> <221> X <222> (32)..(33) <223> Cualquier aminoácido
10	<220> <221> X <222> (35)..(38) <223> Cualquier aminoácido
15	<220> <221> X <222> (41)..(41) <223> Cualquier aminoácido
20	<220> <221> X <222> (46)..(48) <223> Cualquier aminoácido
25	<220> <221> X <222> (53)..(53) <223> Cualquier aminoácido
30	<220> <221> X <222> (55)..(57) <223> Cualquier aminoácido
35	<220> <221> X <222> (61)..(61) <223> Cualquier aminoácido
40	<220> <221> X <222> (67)..(67) <223> Cualquier aminoácido
45	<220> <221> X <222> (72)..(72) <223> Cualquier aminoácido
50	<220> <221> X <222> (74)..(77) <223> Cualquier aminoácido
55	<220> <221> X <222> (79)..(79) <223> Cualquier aminoácido
60	<220> <221> X <222> (114)..(114) <223> Cualquier aminoácido
65	<220> <221> X <222> (127)..(128) <223> Cualquier aminoácido

5  
<220>  
<221> X  
<222> (136)..(136)  
<223> Cualquier aminoácido

10  
<220>  
<221> X  
<222> (139)..(142)  
<223> Cualquier aminoácido

15  
<220>  
<221> X  
<222> (144)..(144)  
<223> Cualquier aminoácido

20  
<220>  
<221> X  
<222> (150)..(150)  
<223> Cualquier aminoácido

25  
<220>  
<221> X  
<222> (164)..(165)  
<223> Cualquier aminoácido

30  
<220>  
<221> X  
<222> (167)..(172)  
<223> Cualquier aminoácido

<400> 226

ES 2 640 100 T3

Ala Val Phe Asp Lys Asp Gly Thr Leu Thr Glu Asp Asp Thr Xaa Phe  
 1 5 10 15

Leu Xaa Tyr Leu Leu Lys Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Leu Xaa  
 20 25 30

Xaa Asp Xaa Xaa Xaa Xaa Gly Ser Xaa Leu Thr Leu Ser Xaa Xaa Xaa  
 35 40 45

Asp Leu Leu Glu Xaa Leu Xaa Xaa Xaa Gly Gly Ile Xaa Val Ile Gly  
 50 55 60

Leu Ala Xaa Arg Tyr Leu Glu Xaa Leu Xaa Xaa Xaa Xaa Glu Xaa Ala  
 65 70 75 80

Lys Leu Phe Glu Gly Phe Ile Lys Pro Asp Ala Ala Glu Leu Leu Lys  
 85 90 95

Glu Leu His Glu Ala Gly Leu Arg Val Val Val Leu Thr Gly Asp Pro  
 100 105 110

Arg Xaa Ile Ala Lys Pro Val Ala Lys Glu Leu Gly Ile Asp Xaa Xaa  
 115 120 125

Asn Val Leu Ala Thr Glu Leu Xaa Asp Glu Xaa Xaa Xaa Xaa Val Xaa  
 130 135 140

Gly Arg Ile Thr Gly Xaa Leu Asp Lys Ala Arg Ala Val Glu Arg Leu  
 145 150 155 160

Val Val Leu Xaa Xaa Lys Xaa Xaa Xaa Xaa Xaa Xaa Val Val Ala Ile  
 165 170 175

Gly Asp Ser Ala Asn Asp Leu Pro Ala Leu Lys  
 180 185

<210> 227  
 <211> 190  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> Secuencia conservada

10

<400> 227

ES 2 640 100 T3

Ile Lys Ala Val Val Phe Asp Lys Asp Gly Thr Leu Thr Asp Gly Glu  
 1 5 10 15

Pro Pro Ile Ala Glu Ala Ile Val Glu Ala Ala Ala Glu Leu Gly Leu  
 20 25 30

Pro Leu Leu Leu Pro Leu Glu Glu Val Glu Lys Leu Leu Gly Arg Gly  
 35 40 45

Val Glu Gly Ile Glu Arg Ile Leu Leu Glu Gly Gly Leu Thr Ala Glu  
 50 55 60

Leu Leu Leu Glu Leu Glu Gly Glu Leu Ala Ala Gly Lys Thr Ala Val  
 65 70 75 80

Leu Val Ala Leu Asp Gly Glu Val Leu Gly Leu Ile Ala Leu Ala Asp  
 85 90 95

Lys Leu Tyr Pro Gly Ala Arg Glu Ala Leu Lys Ala Leu Lys Glu Arg  
 100 105 110

Gly Ile Lys Val Ala Ile Leu Thr Asn Gly Asp Arg Ala Asn Ala Glu  
 115 120 125

Ala Val Leu Glu Ala Leu Gly Leu Ala Asp Leu Phe Asp Val Ile Val  
 130 135 140

Asp Ser Asp Asp Val Gly Pro Val Lys Pro Lys Pro Glu Ile Phe Leu  
 145 150 155 160

Lys Ala Leu Glu Arg Leu Gly Val Lys Pro Glu Glu Val Leu Met Val  
 165 170 175

Gly Asp Gly Val Asn Asp Ala Pro Ala Leu Ala Ala Ala Gly  
 180 185 190

<210> 228  
 <211> 15  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> Secuencia conservada

<400> 228

Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 1 5 10 15

ES 2 640 100 T3

5  
<210> 229  
<211> 6  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Secuencia conservada

10  
<220>  
<221> X  
<222> (2)..(2)  
<223> Cualquier aminoácido

15  
<220>  
<221> X  
<222> (4)..(4)  
<223> Cualquier aminoácido

20  
<220>  
<221> X  
<222> (5)..(5)  
<223> Treonina (T) o Valina (V)

25  
<220>  
<221> X  
<222> (6)..(6)  
<223> Leucina (L) o Valina (V)

30  
<400> 229

**Asp Xaa Asp Xaa Xaa Xaa**

**1**

**5**

35  
<210> 230  
<211> 8  
<212> PRT  
<213> Secuencia artificial

40  
<220>  
<223> Secuencia conservada

45  
<220>  
<221> X  
<222> (2)..(2)  
<223> 17 a 20 aminoácidos; donde los aminoácidos pueden ser cualquier aminoácido

50  
<220>  
<221> X  
<222> (3)..(3)  
<223> Glicina (G) o Serina (S)

55  
<220>  
<221> X  
<222> (4)..(4)  
<223> Ácido aspártico (D) o Serina (S)

60  
<220>  
<221> X  
<222> (5)..(7)  
<223> Cualquier aminoácido

<220>



ES 2 640 100 T3

<221> X  
<222> (8)..(8)  
<223> Ácido aspártico (D) o Asparagina (N)

5 <400> 230

**Lys Xaa Xaa Xaa Xaa Xaa Xaa Xaa**  
**1 5**

## REIVINDICACIONES

1. Un organismo transgénico no humano o una parte del mismo o una semilla transgénica de una planta que comprenden uno o más polinucleótidos exógenos que codifican una monoacilglicerol aciltransferasa (MGAT), en donde el organismo transgénico no humano o una parte del mismo o una semilla tienen un nivel aumentado de uno o más lípidos polares en comparación con un organismo o una parte del mismo o una semilla correspondientes que carecen de los uno o más polinucleótidos exógenos, en donde el organismo transgénico no humano es una planta, un alga, una levadura o un hongo y los uno o más lípidos no polares incluyen triacilglicerol (TAG).
2. El organismo transgénico no humano o una parte del mismo o una semilla de la reivindicación 1, en donde los uno o más polinucleótidos exógenos codifican una monoacilglicerol aciltransferasa 2 (MGAT2).
3. El organismo transgénico no humano o una parte del mismo o una semilla de las reivindicaciones 1 o 2, en donde el nivel de los uno o más lípidos no polares del organismo transgénico no humano o una parte del mismo o una semilla es al menos un 0,5 % (p/p) mayor basado en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en el organismo correspondiente o una parte del mismo o una semilla.
4. El organismo transgénico no humano o una parte del mismo o una semilla de una cualquiera de las reivindicaciones 1 a 4, que se caracterizan por uno o más de:
- (i) en donde los uno o más polinucleótidos exógenos codifican una MGAT y una glicerol-3-fosfato aciltransferasa (GPAT), una MGAT y una diacilglicerol aciltransferasa (DGAT) o las tres de una MGAT, una GPAT y una DGAT, preferentemente en donde la GPAT también tiene actividad fosfatasa para producir monoacilglicerol (MAG) y preferentemente en donde la DGAT es una diacilglicerol aciltransferasa 2 (DGAT2);
  - (ii) en donde el contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo o una semilla está aumentado en comparación con el organismo o una parte del mismo correspondientes;
  - (iii) en donde el contenido total de lípidos no polares del organismo transgénico no humano o una parte del mismo o una semilla es al menos un 0,5 % (p/p) mayor basado en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en el organismo o una parte del mismo o una semilla correspondientes;
  - (iv) en donde el contenido en ácidos grasos poliinsaturados (PUFA) del lípido del organismo transgénico no humano o una parte del mismo o una semilla está aumentado en comparación con el lípido del organismo correspondiente o una parte del mismo o una semilla;
  - (v) en donde el contenido de un ácido graso poliinsaturado (PUFA) en el lípido del organismo transgénico no humano o una parte del mismo o una semilla está aumentado en comparación con el lípido del organismo o una parte del mismo o una semilla correspondientes, en donde el PUFA comprende ácido eicosadienoico (EDA), ácido araquidónico (ARA), ácido alfa linoléico (ALA), ácido estearidónico (SDA), ácido eicosatrienoico (ETE), ácido eicosatetraenoico (ETA), ácido eicosapentaenoico (EPA), ácido docosapentaenoico (DPA), ácido docosahexaenoico (DHA) o una combinación de dos o más de los mismos;
  - (vi) en donde la parte es una semilla o un fruto;
  - (vii) en donde el organismo transgénico no humano o una parte del mismo o una semilla es de *Brassica sp.*, *Gossypium hirsutum*, *Linum usitatissimum*, *Helianthus sp.*, *Carthamus tinctorius*, *Glycine max*, *Zea mays*, *Arabidopsis thaliana*, *Sorghum bicolor*, *Sorghum vulgare*, *Avena sativa*, *Triforium sp.*, *Camelina sativa*, *Miscanthus x giganteus* o *Miscanthus sinensis* o de lupino, cacahuete o girasol;
  - (viii) en donde el contenido total de aceite o el contenido total de ácidos grasos de la semilla es de entre un 0,5 % (p/p) y un 25 % (p/p) mayor basado en peso que en una semilla correspondiente que carece de los uno o más polinucleótidos exógenos;
  - (ix) en donde el organismo transgénico no humano o una parte del mismo o una semilla son una parte vegetativa de una planta, una parte aérea de una planta, una parte verde de planta, una hoja de planta o un tallo de planta;
  - (x) en donde la parte es una parte vegetativa de una planta y el contenido de triacilglicerol (TAG), diacilglicerol (DAG), TAG y DAG o de monoacilglicerol (MAG) de la parte vegetativa de planta es al menos un 10 % (p/p) sobre una base relativa que el contenido de TAG, DAG, TAG y DAG o MAG de una parte vegetativa de planta correspondiente que carece de los uno o más polinucleótidos exógenos;
  - (xi) en donde al menos un 60 % (% en moles) del contenido de ácidos grasos del lípido del organismo transgénico no humano o una parte del mismo o una semilla es de ácido oleico; y
  - (xii) en donde el nivel de los uno o más lípidos no polares y/o el contenido total de lípidos no polares del organismo transgénico o una parte del mismo o una semilla es al menos un 0,5 % (p/p) mayor basado en peso y/o al menos un 1 % (p/p) mayor sobre una base relativa que en un organismo o una parte del mismo o una semilla correspondientes que carecen de los uno o más polinucleótidos exógenos pero que comprenden un polinucleótido exógeno que codifica una diacilglicerol aciltransferasa 1 (DGAT1) de *Arabidopsis thaliana*.
5. El organismo transgénico no humano o una parte del mismo o una semilla de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en donde el lípido no polar es TAG, DAG, TAG y DAG, MAG, PUFA total o un PUFA específico que es ácido eicosadienoico (EDA), ácido araquidónico (ARA), ácido alfa linoléico (ALA), ácido estearidónico (SDA), ácido eicosatrienoico (ETE), ácido eicosatetraenoico (ETA), ácido eicosapentaenoico (EPA), ácido docosapentaenoico (DPA), ácido docosahexaenoico (DHA) o una combinación de dos o más de los mismos.

6. Un método para producir lípido extraído, comprendiendo el método las etapas de:

i) obtener un organismo transgénico no humano o una parte del mismo o una semilla de acuerdo con una cualquiera de las reivindicaciones 1 a 5 y

ii) extraer el lípido del organismo transgénico no humano o una parte del mismo, produciendo de este modo el lípido extraído.

7. El método de la reivindicación 6, en donde el organismo transgénico no humano o una parte del mismo o una semilla es semilla de canola que comprende uno o más de lo siguiente:

i) al menos un 45 % de aceite de semilla basado en peso,

ii) un contenido relativo de aceite de semilla al menos un 5 % (p/p) mayor que en una semilla correspondiente,

iii) un contenido relativo de DAG al menos un 10 % (p/p) mayor que en una semilla correspondiente, y

iv) un contenido relativo de TAG al menos un 5 % (p/p) mayor que en una semilla correspondiente.

8. El método de la reivindicación 6, en donde el organismo transgénico no humano o una parte del mismo o una semilla es semilla de maíz que comprende uno o más de lo siguiente:

i) al menos un 5 % de aceite de semilla basado en peso,

ii) un contenido relativo de aceite de semilla al menos un 5 % (p/p) mayor que en una semilla correspondiente,

iii) un contenido relativo de DAG al menos un 10 % (p/p) mayor que en una semilla correspondiente, y

iv) un contenido relativo de TAG al menos un 5 % (p/p) mayor que en una semilla correspondiente.

9. El método de la reivindicación 6, en donde el organismo transgénico no humano o una parte del mismo o una semilla es semilla de soja que comprende uno o más de lo siguiente:

i) al menos un 20 % de aceite de semilla basado en peso,

ii) un contenido relativo de aceite de semilla al menos un 5 % (p/p) mayor que en una semilla correspondiente,

iii) un contenido relativo de DAG al menos un 10 % (p/p) mayor que en una semilla correspondiente, y

iv) un contenido relativo de TAG al menos un 5 % (p/p) mayor que en una semilla correspondiente.

10. El método de una cualquiera de las reivindicaciones 6 a 9, en donde el método comprende además uno o más de cosechar la semilla de una planta transgénica, obtener el aceite de semilla prensando la semilla, purificar el aceite de semilla y producir ésteres de alquilo, tales como ésteres de metilo de los ácidos grasos del lípido extraído.

11. Un método para obtener una célula con capacidad mejorada para producir uno o más lípidos no polares, comprendiendo el método:

i) introducir en una célula uno o más polinucleótidos exógenos que codifican

a) una MGAT,

b) una MGAT y una GPAT

c) una MGAT y una DGAT o

d) una MGAT, una GPAT y una DGAT,

en donde dichos uno o más polinucleótidos exógenos están ligados operativamente a uno o más promotores que tienen la capacidad de dirigir la expresión de dichos uno o más polinucleótidos exógenos en la célula, en donde la célula es una célula vegetal, de alga, de levadura o fúngica y en donde los uno o más lípidos no polares incluyen triacilglicerol (TAG),

ii) expresar el uno o más polinucleótidos exógenos en la célula,

iii) analizar el contenido de lípidos de la célula y

iv) seleccionar una célula que tenga un nivel aumentado de dichos uno o más lípidos no polares en comparación con una célula correspondiente que no contiene los polinucleótidos exógenos.

12. El método de la reivindicación 11, en donde la célula seleccionada tiene una o más de las características del organismo o la parte del mismo o la semilla definidos en una cualquiera de las reivindicaciones 1 a 5.

13. Uso de uno o más polinucleótidos exógenos que codifican:

i) una MGAT,

ii) una MGAT y una GPAT

iii) una MGAT y una DGAT, o

iv) una MGAT, una GPAT y una DGAT,

para producir un organismo no humano o una parte del mismo o una semilla transgénicos con capacidad mejorada para producir uno o más lípidos no polares en comparación con un organismo o una parte del mismo o una semilla

correspondientes que carecen de los uno o más polinucleótidos exógenos, en donde el organismo transgénico no humano es una planta, un alga, una levadura o un hongo y los uno o más lípidos no polares incluyen triacilglicerol (TAG).

- 5 14. Uso del organismo no humano o una parte del mismo o una semilla transgénicos de una cualquiera de las reivindicaciones 1 a 5 para la fabricación de ésteres de alquilo, tales como metil ésteres de ácidos grasos.
- 10 15. Uso de un polinucleótido exógeno que codifica una monoacilglicerol aciltransferasa (MGAT) en un organismo transgénico no humano o una parte del mismo o una semilla transgénica de una planta para la producción de triacilglicerol (TAG), en donde el organismo transgénico no humano es una planta, un alga, una levadura o un hongo.

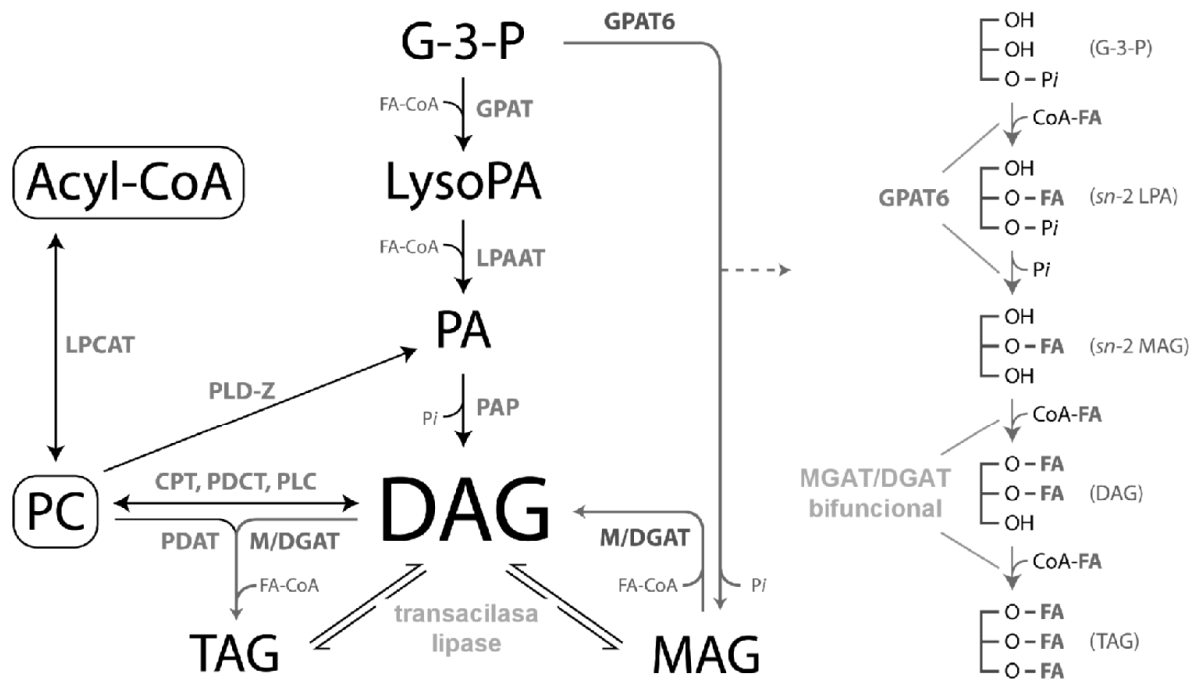
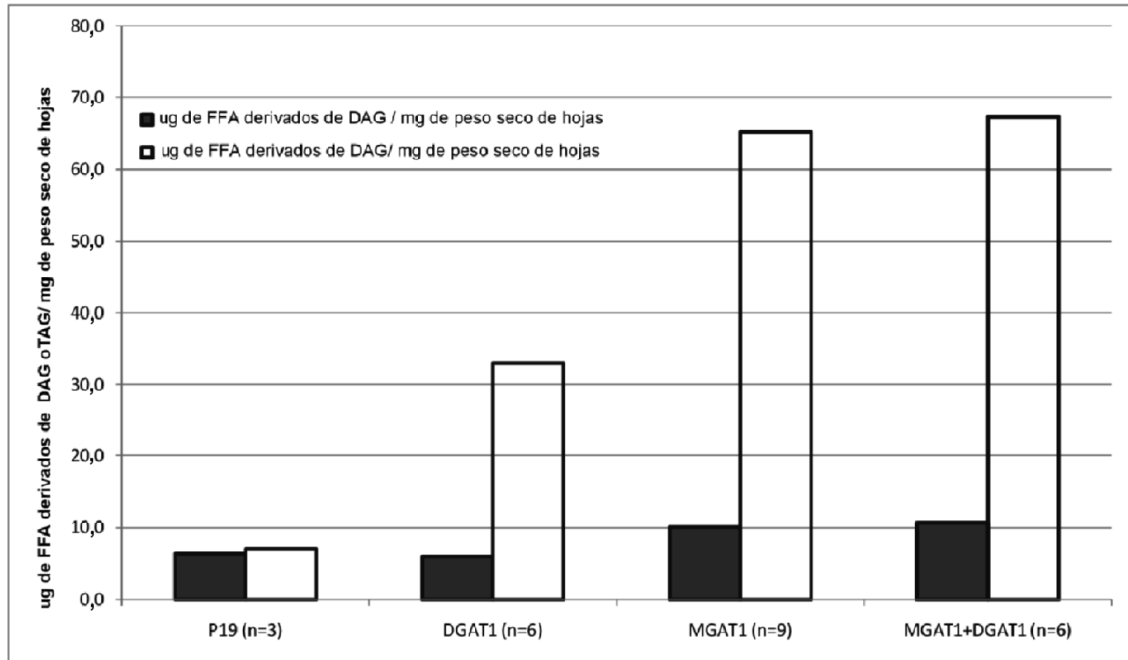
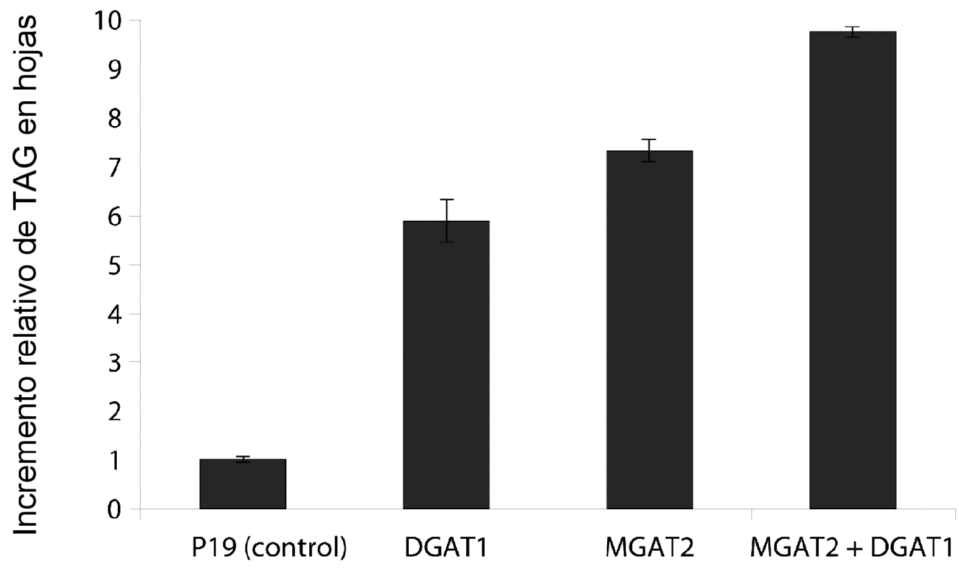


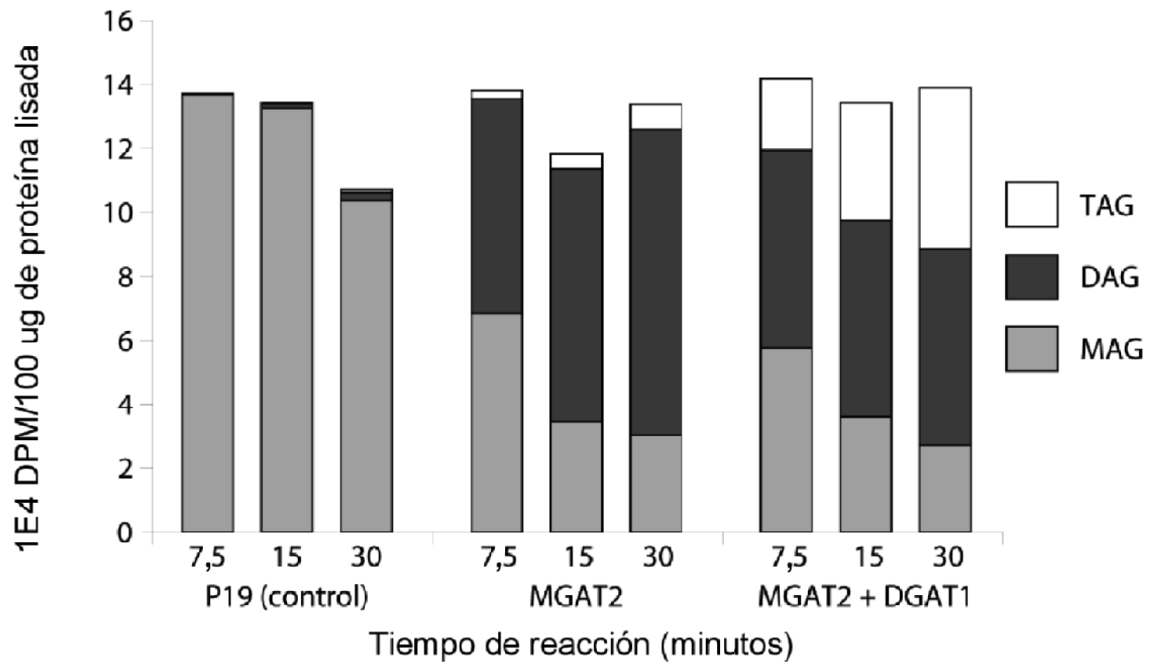
Figura 1



**Figura 2**

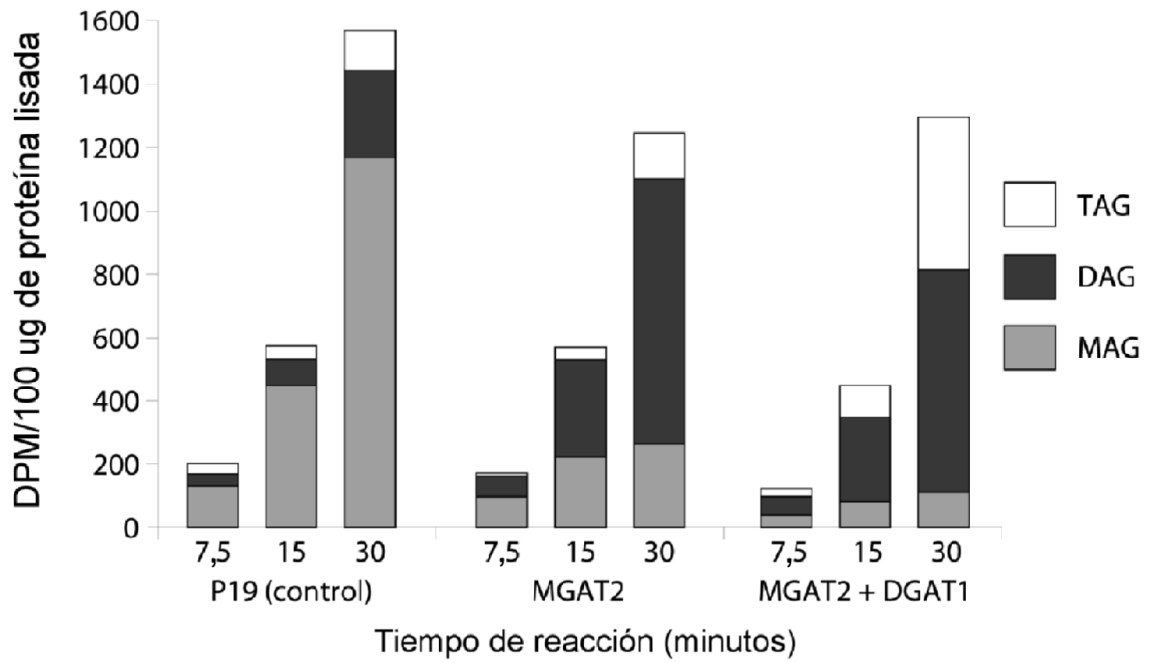


**Figura 3**

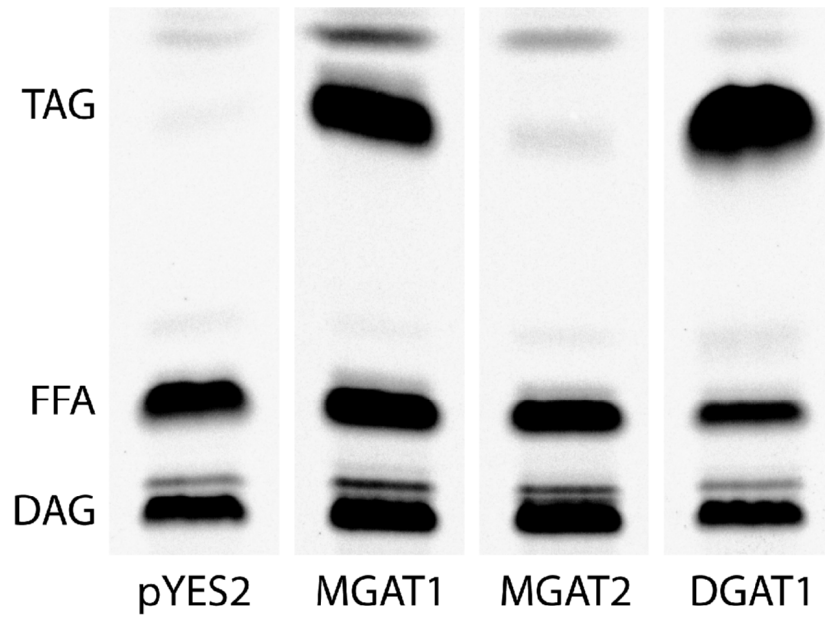


**Figura 4**





**Figura 5**



**Figura 6**