Industria y Comercio
SUPERINTENDENCIA

Radicación	14 46974 0000 0000	2014-03-05 15:04:14	
Trámite	11	PCT-PATENTE DE INVENCION CAPITULOS I Y II	
Evento	379	FASE NACIONAL INCLUIDO CAPITULO II	
Actuación	411	PRESENTACION	Folios: 111
Dependencia	2020	DIRECCION DE NUEVAS CREACIONES	
Pago #1	25018	$ 440800	

DIRECCIÓN DE NUEVAS CREACIONES SOLICITUD FASE NACIONAL - PCT

1	**Titulo de la Invención**(200 caractares o espacios máximos)
	METODO PARA INCREMENTAR LA ASIMILACION DE CO2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS

2	**Datos del Solicitante / Titular**			
	Nombre:	AGRESEARCH LIMITED	**Dirección Electrónica:**	notificacionespatentes@lloredacamacho.com
	Dirección:	5th Floor Tower Block Ruakura Research, Centre East Street, Hamilton, 3240, Nueva Zelanda.	**Domicilio/País de constitución:**	NUEVA ZELANDA - WAIKATO - HAMILTON
	Identificación: ☐ CEDULA DE CIUDADANIA ☐ CEDULA DE EXTRANJERIA ☒ EMPRESA EXTRANJERA ☐ NIT ☐ PASAPORTE **Número:** 90004801099-			

3	**Solicitantes**		
	Apellidos - Nombres o Razón Social	**Tipo**	**Identificación**
1.	AGRESEARCH LIMITED	NN	90004801099

4	Datos del Inventor			
	Nombre:	Nicholas John ROBERTS	**Dirección Electrónica:**	notificacionespatentes@lloredacamacho.com
	Dirección:	97 McKay Line, RD 7 Feilding, Wellington, 4777, Nueva Zelanda	**Domicilio/País de constitución:**	NUEVA ZELANDA - WELLINGTON - WELLINGTON
	Identificación: ☐ CEDULA DE CIUDADANIA ☐ CEDULA DE EXTRANJERIA ☐ EMPRESA EXTRANJERA ☐ NIT ☐ PASAPORTE ☒ No Aplica **Número:** -			

5	Inventor(es)	
	Apellidos - Nombres	**Domicilio**
1.	ROBERTS Nicholas John	NUEVA ZELANDA
2.	SCOTT Richard William	NUEVA ZELANDA
3.	WINICHAYAKUL Somrutai	NUEVA ZELANDA
4.	ROLDAN Marissa	NUEVA ZELANDA

6	Datos Inventor(es)			
	País de Residencia	**Departamento/Estado**	**Ciudad**	**Dirección**
1.	NUEVA ZELANDA	WELLINGTON	WELLINGTON	97 McKay Line, RD 7 Feilding, Wellington, 4777, Nueva Zelanda

2.	NUEVA ZELANDA	WELLINGTON	WELLINGTON	25 Colombo Street, Palmerston North, Wellington, 4410, Nueva Zelanda
3.	NUEVA ZELANDA	WELLINGTON	WELLINGTON	26 Geneva Terrace, Kelvin Grove, Palmerston North, Wellington, 4414, Nueva Zelanda
4.	NUEVA ZELANDA	WELLINGTON	WELLINGTON	12 Brooklyn Heights Drive, Kelvin Grove, Palmerston North, Wellington, 4414, Nueva Zelanda

7	**Datos del Representante Legal / Apoderado**		
Nombre:	ALICIA LLOREDA RICAURTE	**Dirección Electrónica:**	notificacionespatentes@lloredacamacho.com
Dirección:	calle 72 no 5-83	**Domicilio/País de constitución:**	COLOMBIA - BOGOTA D.C. - BOGOTA D.C.

Identificación:

☒ CEDULA DE CIUDADANIA ☐ CEDULA DE EXTRANJERIA

☐ EMPRESA EXTRANJERA ☐ NIT

☐ PASAPORTE

Número: 53215-

Presentación de Poder

Año de Radicación	
Número de Radicación	

8	**Datos Solicitud: PCT / WO**		
Número Solicitud:	PCT/NZ2012/000138	**Fecha Solicitud:**	03/08/2012
Número Publicacion:	WO 2013/022353 A1	**Fecha Publicacion:**	14/02/2013

9	**Declaraciones de prioridad**	☒ SI ☐ NO

	(33) País de origen	Codigo del país	(31) No. Solicitud	(32) Fecha
1.	ESTADOS UNIDOS DE AMÉRICA	US	61/515,610	05/08/2011

10	Reivindicaciones			
Número reivindicaciones:	69	**Pago Reivindicaciones:**	No	

11 Reducción de tasas.

Declaro que carezco de medios económicos para presentar la solicitud de patente.

☐ SI ☒ NO

Nota: En caso de ser persona natural y carecer de medios económicos, y por lo tanto, aplique la reducción de tasas a que se refiere la resolucón vigente en tarifas, debe firmar la presente solicitud bajo la gravedad de juramento.

☐ Micro, pequeñas y medianas empresas
☐ Universidades públicas o privadas
☐ Entidades sin ánimo de lucro

Debe aportar los documentos que se indican en el numeral 17 de anexos

12 Documentos Anexos

☒ Reivindicaciones
☒ Descripción
☒ Dibujos y/o Figuras
☒ Resumen
☐ Certificado Depósito Material Biológico
☐ Uso de Conocimiento tradicional
☒ Listado de secuencias
☐ Artes finales 12 x 12 cm
☐ Poderes, si fuere el caso
☐ Copia de la primera solicitud si se reivindica prioridad
☐ Traducción simple de la primera solicitud, si se reivindica prioridad
☒ Otros Anexos

MODIFICACIONES A LAS REIVINDICACIONES PRESENTADAS EN FASE INTERNACIONAL

(INDICADAS EN EL REPORTE PRELIMINAR SOBRE PATENTABILIDAD)

Atentamente me permito informar a ese Despacho que el Capítulo Reivindicatorio tal como fue presentado originalmente en Fase Internacional fue modificado en dicha instancia bajo el Reporte Preliminar Internacional sobre Patentabilidad.

Las modificaciones realizadas se encuentran plenamente sustentadas en la descripción de la presente invención y en este sentido, no configuran una ampliación del invento o de la divulgación contenida en la solicitud presentada.

METODO PARA INCREMENTAR LA ASIMILACIÓN DE CO_2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS

MEMORIA DESCRIPTIVA

Campo Técnico

El invento se refiere a métodos para la producción de células y plantas fotosintéticas con tasas de asimilación de CO_2 acrecentadas. El invento se refiere también a métodos para la producción de aceite de las plantas.

Antecedentes

La creciente población global presenta demandas de cultivos que tengan un mayor rendimiento con una producción mejorada (asimilación fotosintética del carbono).

La ribulosa bifosfato carboxilasa (Rubisco) es la enzima clave responsable de la asimilación fotosintética del carbono. En presencia de O_2, Rubisco realiza también una reacción de oxigenasa que inicia el ciclo fotorrespiratorio que tiene como resultado una pérdida indirecta de la célula de nitrógeno y CO_2 fijados que necesitan ser recuperados. Una modificación genética para aumentar la especificidad de Rubisco para el CO_2 en relación con el O_2 y para acrecentar la tasa catalítica de Rubisco en plantas de cultivo tendría un gran valor agronómico. Parry *y otros*, (2003) revisaron el avance hasta la fecha, concluyendo que hay todavía muchas barreras técnicas por superar y hasta la fecha todas las tentativas de la ingeniería genética no han logrado hasta aquí producir una mejor Rubisco (Peterhansel y otros, 2008).

En la naturaleza, una cantidad de plantas superiores (plantas C4) han desarrollado mecanismos que requieren energía para aumentar la concentración de CO_2 intracelular en estrecha proximidad a Rubisco aumentando con ello la proporción de reacciones de carboxilasa. Por ejemplo, el maíz ha logrado esto por una manipulación de la arquitectura de la planta que permite un proceso inicial de fijación del CO_2, conocido como metabolismo C4. La desventaja agronómica de esta modificación evolucionada es un aumento en la fibra de la hoja que tiene como consecuencia que las hojas de plantas C4 son comparativamente poco digeribles. Se piensa que la fotosíntesis C4 es un producto de evolución convergente que se ha desarrollado en ocasiones separadas en taxones muy diferentes. Sin embargo, esta adaptación es posible sólo para organismos multicelulares (y no para organismos unicelulares fotosintéticos tales como algas). Las algas tienen una variedad de diferentes mecanismos para concentrar el CO_2; sin embargo, parece que hay un continuo en

el grado en el cual el mecanismo de concentración de CO_2 (CCM) se expresa en respuesta a la concentración de carbono inorgánico disuelto (DIC) externo, llevando mayores concentraciones a un mayor grado de supresión de la actividad del CCM. Dos reseñas han cubierto los CCMs en algas así como también su modulación y sus mecanismos y son incorporadas a la presente por referencia (Giordano, Beardall y otros 2005; Moroney e Ynalvez 2007). Las plantas vasculares que constituían actualmente el porcentaje mayor de los alimentos básicos de la dieta humana son plantas C3 (arroz y tubérculos) y no plantas C4. De una manera similar, muchos de los cultivos de semillas oleosas (canola, girasol, cártamo) y muchos cultivos de alimentos para animales de carne y leche (legumbres, cereales, soya, pastos forrajeros) son plantas C3.

El aumento en la eficiencia de asimilación del CO_2, debería aumentar por lo tanto concurrentemente la tolerancia al estrés abiótico y la eficiencia en el uso del nitrógeno y aportaría un beneficio agronómico significativo a las plantas C3 y los microorganismos fotosintéticos.

Por lo tanto, los mecanismos para elevar la concentración del CO_2 en el cloroplasto, reducir la fotorrespiración y posteriormente aumentar la tolerancia al estrés abiótico y la productividad sería de un beneficio agronómico significativo para plantas C3 y microorganismos fotosintéticos.

Un objetivo del invento es proveer métodos para incrementar la tasa de asimilación del CO_2 en células y plantas fotosintéticas, y métodos para producir células y plantas fotosintéticas con una tasa de asimilación del CO_2 acrecentada.

En la naturaleza las plantas que florecen almacenan eficientemente la energía en sus semillas a través de la acumulación de aceite, a saber, triacilglicerol (TAG) y lo almacenan en cuerpos oleosos discretos incrustando una monocapa de proteína de fosfolípidos alrededor del cuerpo oleoso. Estos cultivos de semillas se han usado en una variedad de aplicaciones agrícolas como forraje y más recientemente también como una fuente de materia prima para los biocombustibles. En una base en peso, los lípidos tienen aproximadamente el doble del contenido de energía de las proteínas o los carbohidratos y por lo tanto se ha puesto un enfoque substancial en elevar el contenido de aceite de varias especies, más particularmente, de plantas.

Lamentablemente, las semillas de plantas representan un porcentaje muy pequeño de la biomasa total de la planta y con la demanda de una mayor productividad agrícola y de energías alternativas, se reconoce que la producción actual de aceite a partir de una cantidad de cultivos de semillas dedicados a esto es insuficiente. Los esfuerzos de investigación se han enfocado no sólo en aumentar la productividad de la producción de aceite dentro de las

semillas de plantas sino también de la producción de aceite en otros tipos y especies de células.

La reproducción tradicional y la mutagénesis han ofrecido éxitos crecientes en esta área; sin embargo, la ingeniería genética ha hecho los mayores avances en la modificación de organismos para producir niveles de aceite elevados. Aunque ciertos grupos han trabajado a lo largo de varias partes de la vía de síntesis de aceite para regular positivamente la producción de aceite dentro de la semilla, otros grupos se han enfocado en aumentar el aceite en tipos de células que representan una porción mayor de la biomasa.

Avances significativos han sido realizados a través de la expresión de oleosinas modificadas que incluyen cisteínas introducidas en plantas en forma artificial. En WO 2011/053169 los solicitantes han demostrado un incremento significativo en el nivel del aceite producido en las hojas. Sin embargo serían más beneficiosos métodos para aumentar más el nivel en tejidos de plantas, y así aumentar la cantidad de aceite extraíble de las plantas.

Por lo tanto, otro objetivo del invento es proveer métodos para incrementar el nivel de producción de aceite en tejidos/órganos de plantas y/o métodos para incrementar la producción de aceite de las plantas.

Reseña del Invento

El invento provee métodos para incrementar la tasa de asimilación de CO_2 en células y plantas fotosintéticas. El invento comprende reducir el reciclado de lípidos y/o expresar oleosinas modificadas con residuos de cisteína introducidos artificialmente en células y plantas fotosintéticas.

El invento provee también métodos para incrementar la producción de aceite en plantas, por medio de la expresión de oleosinas modificadas con residuos de cisteína introducidos artificialmente en los tejidos/órganos no fotosintéticos de las plantas. Los solicitantes han mostrado de manera sorprendente que los tejidos/órganos no fotosintéticos de plantas que expresan tales oleosinas modificadas acumulan aceite hasta un nivel más alto que otros tejidos de la planta. El método incluye también opcionalmente el paso de extraer el aceite de tejidos/órganos no fotosintéticos de la planta, o procesar los tejidos/órganos no fotosintéticos ricos en aceite en materias primas para alimentos de animales o para biocombustibles.

En el primer aspecto el invento provee un método para producir una célula fotosintética con una tasa acrecentada de asimilación de CO_2, comprendiendo el método por lo menos uno de los pasos:

a) modificar genéticamente la célula fotosintética para reducir o evitar el reciclado de lípidos, y

b) transformar la célula fotosintética con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

En una realización el método comprende el paso de modificar genéticamente la célula fotosintética para reducir o evitar el reciclado de lípidos.

En otra realización, el método comprende el paso de transformar la célula fotosintética con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

En una realización la célula es modificada genéticamente para evitar el reciclado de lípidos, transformando la célula fotosintética con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

En una realización preferida la oleosina modificada es expresada en la célula fotosintética. En una realización la expresión de la oleosina modificada produce la tasa acrecentada de asimilación de CO_2. En una realización, la expresión de la oleosina modificada reduce o evita el reciclado de lípidos en la célula fotosintética. En una realización preferida el reciclado de lípidos reducido o evitado produce el incremento en la asimilación de CO_2.

En una realización más el reciclado de lípidos es iniciado por la acción de lipasas que liberan ácidos grasos libres de un esqueleto de glicerol. En una realización más el reciclado de lípidos es impulsado por la reincorporación de ácidos grasos en esqueletos de glicerol dentro del retículo endoplásmico de la célula.

En una realización la tasa de asimilación de CO_2 es acrecentada por lo menos en 1%, más preferentemente por lo menos en 2%, más preferentemente por lo menos en 3%, más preferentemente por lo menos en 4%, más preferentemente por lo menos en 5%, más preferentemente por lo menos en 10%, más preferentemente por lo menos en 15%, más preferentemente por lo menos en 20%, más preferentemente por lo menos en 25%, más preferentemente por lo menos en 30%, más preferentemente por lo menos 35%, más preferentemente por lo menos en 40%, más preferentemente por lo menos en 45%, más preferentemente por lo menos en 50%, en relación con una planta de control.

En una realización el incremento en la tasa de asimilación de CO_2 está en el intervalo de 1% a 50%, más preferentemente de 2% a 40%, más preferentemente de 3% a 30%, más preferentemente de 4% a 25%, más preferentemente de 5% a 20%, en relación con una planta de control.

En una realización el incremento en la asimilación de CO_2 es el resultado de una concentración elevada de CO_2 en el cloroplasto.

Oleosina modificada

El término oleosina incluye también esteroleosina y caolesina. La oleosina modificada puede seleccionarse por lo tanto de una oleosina modificada, una caolesina modificada o una esteroleosina modificada. En una realización la oleosina modificada es una oleosina modificada. En otra realización la oleosina modificada es una caolesina modificada. En otra realización la oleosina modificada es una esteroleosina modificada. En la presente se describe ejemplos de cada tipo de oleosina (oleosina, caolesina y esteroleosina).

En una realización, la oleosina modificada incluye por lo menos dos cisteínas, de las cuales por lo menos una es introducida artificialmente. En una realización más, la oleosina modificada incluye por lo menos desde dos hasta por lo menos trece (es decir 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o más) cisteínas introducidas artificialmente. En una realización las cisteínas son introducidas artificialmente en la región hidrófila N-terminal de la oleosina, o en la región hidrófila C-terminal de la oleosina. En una realización más la oleosina modificada incluye por lo menos una cisteína en la región hidrófila N-terminal, y por lo menos una cisteína en la región hidrófila C-terminal. En una realización más las cisteínas están distribuidas de una manera substancialmente uniforme a lo largo de las regiones hidrófilas N-terminal y C-terminal de la oleosina. En una realización más las cisteínas están distribuidas uniformemente a lo largo de las regiones hidrófilas N-terminal y C-terminal de la oleosina.

Otros fenotipos asociados de la célula fotosintética

En una realización más, además de la tasa acrecentada de asimilación de CO_2 el método produce una célula fotosintética que tiene por lo menos una de:

a) una tasa acrecentada de fotosíntesis, y

b) eficiencia acrecentada en el uso del agua, y

c) una tasa de crecimiento acrecentada

d) concentración acrecentada de CO_2 en el cloroplasto

e) una reducción en la tasa de fotorrespiración,

f) tolerancia acrecentada a altas temperaturas,

g) tolerancia acrecentada a una alta concentración de oxígeno,

h) eficiencia acrecentada en el uso de nitrógeno, y

i) reducción en la pérdida de carbono fijado.

Preferentemente la célula fotosintética producida tiene todos los puntos a) a i).

Promotores

En una realización el polinucleótido está ligado operativamente a un polinucleótido promotor.

En una realización el promotor es capaz de impulsar la expresión del polinucleótido en una célula fotosintética. En una realización el promotor impulsa la expresión del polinucleótido preferentemente en células fotosintéticas. En una realización el promotor es un promotor preferido de una célula fotosintética. En una realización más el promotor es un promotor específico de una célula fotosintética. En una realización más el promotor es un promotor regulado por la luz.

El polinucleótido es parte de una construcción genética

En una realización el polinucleótido es transformado como parte de una construcción genética. Preferentemente la construcción genética es una construcción de expresión. Preferentemente la construcción de expresión incluye el polinucleótido ligado operativamente al promotor. En una realización más el polinucleótido está ligado operativamente a una secuencia de terminación.

La célula fotosintética es transformada también con una enzima sintetizadora de TAG

En una realización más la célula fotosintética es modificada también genéticamente para expresar una enzima sintetizadora de triacilglicerol (TAG). En una realización más la célula fotosintética es modificada genéticamente para que comprenda una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG). En una realización más la célula fotosintética comprende una construcción de expresión que incluye una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG).

En una realización más el ácido nucleico está ligado operativamente a un polinucleótido promotor.

En una realización el promotor es capaz de impulsar la expresión del polinucleótido en la célula fotosintética. En una realización el promotor es un promotor preferido de una célula fotosintética. En una realización más el promotor es un promotor específico de una célula fotosintética. En una realización más el promotor es un promotor regulado por la luz.

Los expertos en la especialidad comprenderán que el polinucleótido que codifica la oleosina modificada y la secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG) pueden ser colocados en la misma construcción o en construcciones separadas para ser transformados en la célula hospedadora. La expresión de cada uno puede ser impulsada por promotores iguales o diferentes, los que pueden estar incluidos en la construcción que se va a transformar. Los expertos en la especialidad comprenderán también

que alternativamente el polinucleótido y el ácido nucleico pueden ser transformados en la célula sin un promotor, pero la expresión de cualquiera del polinucleótido o del ácido nucleico o de ambos podría ser impulsada por un promotor o por promotores endógenos a la célula transformada.

Tipos de células fotosintéticas

La célula fotosintética puede ser de cualquier tipo. En una realización la célula fotosintética es una célula procariota. En una realización más la célula fotosintética es una célula eucariota. En una realización la célula fotosintética es seleccionada de una célula bacteriana fotosintética, una célula de levadura fotosintética, una célula fotosintética de hongo, una célula de alga fotosintética, y una célula vegetal. En una realización la célula fotosintética es una célula bacteriana. En una realización más la célula fotosintética es una célula de levadura. En una realización más la célula fotosintética es una célula de hongo. En una realización más la célula fotosintética es una célula de alga.

La célula fotosintética es una célula de alga

En una realización preferida la célula fotosintética es una célula de alga. En una realización la célula fotosintética de alga es una célula de alga seleccionada de una de las siguientes divisiones: Chlorophyta (algas verdes), Rhodophyta (algas rojas), Phaeophyceae (algas pardas), Bacillariophycaeae (diatomeas), y Dinoflagellata (dinoflageladas).

En una realización la célula de alga muestra un incremento de la tasa de crecimiento, en relación con una célula de alga de control, a una concentración elevada de oxígeno (O_2).

En una realización más, la concentración de O_2 es elevada por lo menos hasta 1,1 veces la saturación de aire, más preferentemente 1,5 veces la saturación de aire, más preferentemente por lo menos 2 veces la saturación de aire, más preferentemente por lo menos 4 veces la saturación de aire, más preferentemente por lo menos 8 veces la saturación de aire, más preferentemente por lo menos 16 veces la saturación de aire.

En una realización más, el incremento de la tasa de crecimiento de la célula de alga es por lo menos de 10%, más preferentemente por lo menos de 20%, más preferentemente por lo menos de 30%, más preferentemente por lo menos de 40%, más preferentemente por lo menos de 50%, más preferentemente por lo menos de 60% más preferentemente por lo menos de 70% más preferentemente por lo menos de 80%, más preferentemente por lo menos de 90%, más preferentemente por lo menos de 100% más que la tasa de crecimiento de una célula de alga de control a la misma concentración de O_2.

En una realización más el incremento en la tasa de crecimiento de la célula de alga está en el intervalo de 10% a aproximadamente 130%, más preferentemente 20% a 120%, más preferentemente 30% a 110%, más preferentemente 40% a 100%, más preferentemente 50% a 90%, más que la tasa de crecimiento de una célula de alga de control a la misma concentración de O_2.

La célula fotosintética es parte de una planta.

En una realización preferida más la célula fotosintética es una célula vegetal. En una realización preferida la célula vegetal es parte de una planta.

De este modo, el invento provee un método para producir una célula o planta fotosintética, con una tasa acrecentada de asimilación de CO_2.

Promotores para plantas

En una realización el promotor ligado operativamente al polinucleótido es capaz de impulsar la expresión del polinucleótido en un tejido fotosintético de una planta. En una realización el promotor es un promotor preferido de una célula fotosintética. En otra realización el promotor es un promotor específico de una célula fotosintética. En una realización más el promotor es capaz de impulsar la expresión del polinucleótido en un tejido fotosintético vegetativo de una planta. En una realización más el promotor es capaz de impulsar la expresión del polinucleótido en una hoja de una planta.

Otros fenotipos asociados para plantas

En una realización más, además de la tasa acrecentada de asimilación de CO_2 la planta tiene también por lo menos uno de:

a) una tasa acrecentada de fotosíntesis, y

b) eficiencia acrecentada del uso de agua, y

c) una tasa acrecentada de crecimiento.

Preferentemente la planta tiene todos los puntos a) a c).

En una realización más, además de la tasa acrecentada de asimilación de CO_2 la planta tiene también por lo menos uno de:

d) incremento de biomasa,

e) retardo de la floración,

f) incremento de la concentración de CO_2 en el cloroplasto,

g) una reducción en la tasa de fotorrespiración,

h) incremento del rendimiento de semillas, frutos u órganos de almacenamiento,

i) incremento en la tolerancia a la sequía,

j) incremento en la tolerancia a alta temperatura,

k) incremento en la tolerancia a una elevada concentración de oxígeno,

l) incremento en la eficiencia en el uso de nitrógeno, y

m) reducción en la pérdida de carbono fijado.

Preferentemente la planta tiene todos los puntos a) a m).

En una realización la biomasa es acrecentada por lo menos en 5%, preferentemente por lo menos en 10%, preferentemente por lo menos en 20%, preferentemente por lo menos en 30%, preferentemente por lo menos en 40%, preferentemente por lo menos en 50%, preferentemente por lo menos en 60% en relación con una planta de control.

En una realización el incremento en biomasa está en el intervalo de 2% a 100%, preferentemente 4% a 90%, preferentemente 6% a 80%, preferentemente 8% a 70%, preferentemente 10% a 60% en relación con una planta de control.

Tipos de plantas

En una realización la planta es una planta C3.

En una realización la planta es seleccionada de: arroz, soya, trigo, cebada, avena, mijo, cebada, papa, canola, girasol y cártamo.

Las plantas preferidas incluyen aquellas de los siguientes géneros*: Oryza, Glycine, Hordeum, Secale, Avena, Pennisetum, Setaria, Panicum, Eleusine, Solanum, Brassica, Helianthus* y*Carthamus.*

Las especies preferidas de *Oryza* incluyen *Oryza sativa* y *Oryza minuta.*

Las especies preferidas de*Glycine* incluyen *Glycine max* y *Glycine wightii (conocida también como Neonotonia wightii).* Una especie de *Glycine* preferidaparticularmente es *Glycine max,* conocida comúnmente como semilla de soya. Una especie de *Glycine* preferidaparticularmente es *Glycine wightii,* conocida comúnmente como semilla de soya perenne.

Una especie preferida de *Hordeum* es *Hordeum vulgare.*

Las especies preferidas de *Triticum* incluyen *Triticum aestivum, Triticum durum* y *Triticum monococcum.*

Una especie preferida de *Secale* es *Secale cereal.*

Una especie preferida de *Avena* es *Avena sativa.*

Las especies preferidas de mijo incluyen *Pennisetum glaucum, Setaria italica, Panicum miliaceum* y *Eleusine coracana.*

Las especies preferidas de *Sorghum* incluyen *Sorghum bicolor* y *Sorghum propinquum.*

Las especies preferidas de *Solanum* incluyen *Solanum habrochaites, Solanum lycopersicum, Solanum nigrum,* y *Solanum tuberosum.*

Las especies preferidas de *Brassica* incluyen*Brassica napus, Brassica campestris* y*Brassica Rapa.*

Las especies preferidas de*Helianthus* incluyen *Helianthus annuus* y *Helianthus argophyllus.*

Una especie preferida de*Carthamus*es*Carthamus tinctorius.*

En un aspecto más el invento provee un método para producir aceite, comprendiendo el método los pasos de:

a) proveer una planta que comprende un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente,

b) cultivar la planta para que exprese la oleosina modificada y produzca aceite en sus tejidos/órganos no fotosintéticos.

Nivel de producción total de lípidos en tejidos/órganos no fotosintéticos.

En una realización la planta acumula más lípido total en sus tejidos/órganos no fotosintéticos que una planta de control.

En una realización más la planta acumula por lo menos 10%, más preferentemente por lo menos 20%, más preferentemente por lo menos 50%, más preferentemente por lo menos 100%, más preferentemente 150%, más preferentemente 200%, más preferentemente 250%, más preferentemente 300%, más preferentemente 350%, más preferentemente 400%, más preferentemente 450%, más preferentemente 500%, más lípido total en sus tejidos/órganos no fotosintéticos que una planta de control.

En una realización la planta produce lípido total en sus tejidos/órganos no fotosintéticos en el intervalo de 100% a 900%, más preferentemente 200% a 800%, más preferentemente 300% a 700%, más preferentemente 400% a 600%, más que una planta de control.

Nivel de producción de aceite en tejidos/órganos no fotosintéticos.

En una realización la planta acumula más aceite en sus tejidos/órganos no fotosintéticos que una planta de control.

En una realización la planta acumula por lo menos 1,2x, por lo menos 1,5x, por lo menos 2x, más preferentemente por lo menos 3x, más preferentemente por lo menos 4x, más preferentemente por lo menos 5x, más preferentemente por lo menos 6x, más preferentemente por lo menos 7x, más preferentemente por lo menos 8x, más preferentemente por lo menos 9x, más preferentemente por lo menos 10x, más preferentemente por lo menos 11x, más

preferentemente por lo menos 12x, más preferentemente por lo menos 13x, más preferentemente por lo menos 14x, más preferentemente por lo menos 15x, más aceite en sus tejidos/órganos no fotosintéticos que una planta de control.

En una realización la planta produce aceite en sus tejidos/órganos no fotosintéticos en el intervalo de 3x a 15x, más preferentemente 4x a 14x, más preferentemente 5x a 13x, más preferentemente 6x a 12x, más preferentemente 7x a 11x, más preferentemente 8x a 10x más que una planta de control.

Las plantas de control adecuadas incluyen versiones no transformadas o de tipo salvaje de una planta de la misma variedad y/o especie que la planta transformada usada en el método del invento. Las plantas de control adecuadas incluyen también plantas de la misma variedad y/o especie que la planta transformada, las cuales son transformadas con una construcción de control. Las plantas de control adecuadas incluyen también plantas que no han sido transformadas con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente. Las plantas de control adecuadas incluyen también plantas que no expresan una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

En una realización preferida se produce aceite en un nivel más alto en los tejidos/órganos no fotosintéticos que en otros o tejidos/órganos de la planta.

Preferentemente el incremento en el nivel de producción de aceite en los tejidos/órganos no fotosintéticos es producido por la expresión de la oleosina modificada en los tejidos/órganos no fotosintéticos.

Tejidos/órganos no fotosintéticos

En una realización el tejido/órgano no fotosintético se selecciona de tejidos/órganos de la planta que están bajo tierra. En una realización más, el tejido/órgano que está bajo tierra es seleccionado de raíz, tubérculo, bulbo, tallo bulboso y rizoma. En una realización más el tejido/órgano no fotosintético es seleccionado de raíz, tubérculo, bulbo, tallo bulboso, rizoma y endosperma. En una realización más el tejido/órgano no fotosintético es una raíz.

Modificación genética

En una realización el método incluye el paso de transformar la planta con el polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

Promotores

En una realización el polinucleótido es ligado operativamente a un promotor de polinucleótido.

En una realización el promotor es capaz de impulsar la expresión del polinucleótido en los tejidos no fotosintéticos de la planta. En una realización el

promotor es un promotor constitutivo. En una realización el promotor es un promotor preferido de tejido no fotosintético. En una realización más el promotor es un promotor preferido de raíz. En una realización más el promotor es un promotor específico de raíz. En una realización más el promotor es un promotor preferido de tubérculo. En una realización más el promotor es un promotor específico de tubérculo. En una realización más el promotor es un promotor preferido de bulbo. En una realización más el promotor es un promotor específico de bulbo. En una realización más el promotor es un promotor preferido de tallo bulboso. En una realización más el promotor es un promotor específico de tallo bulboso. En una realización más el promotor es un promotor preferido de rizoma. En una realización más el promotor es un promotor específico de rizoma. En una realización más el promotor es un promotor preferido de endosperma. En una realización más el promotor es un promotor específico de endosperma.

El polinucleótido es parte de una construcción genética

En una realización el polinucleótido es transformado como parte de una construcción genética. Preferentemente la construcción genética es una construcción de expresión. Preferentemente la construcción de expresión incluye el polinucleótido ligado operativamente al promotor. En una realización más el polinucleótido está ligado operativamente a una secuencia de terminación.

La planta es transformada también con una enzima sintetizadora de TAG

En una realización más la planta es modificada también genéticamente para expresar una enzima sintetizadora de triacilglicerol (TAG). En una realización más la planta es modificada genéticamente para que comprenda una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG). En una realización más la planta comprende una construcción de expresión que incluye una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG).

En una realización más el ácido nucleico está ligado operativamente a un polinucleótido promotor.

En una realización el promotor es capaz de impulsar la expresión del polinucleótido en los tejidos no fotosintéticos de la planta. En una realización el promotor es un promotor constitutivo. En una realización el promotor es un promotor preferido de tejido no fotosintético. En una realización más el promotor es un promotor preferido de raíz. En una realización más el promotor es un promotor específico de raíz. En una realización más el promotor es un promotor preferido de tubérculo. En una realización más el promotor es promotor específico de tubérculo. En una realización más el promotor es un promotor preferido de tallo

bulboso. En una realización más el promotor es un promotor específico de tallo bulboso. En una realización más el promotor es un promotor preferido de rizoma. En una realización más el promotor es un promotor específico de rizoma. En una realización más el promotor es un promotor preferido de endosperma. En una realización más el promotor es un promotor específico de endosperma.

Los expertos en la especialidad comprenderán que el polinucleótido que codifica la oleosina modificada y la secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG) pueden ser colocados en la misma construcción o en construcciones separadas para ser transformados en la célula huésped. La expresión de cada uno puede ser impulsada por promotores iguales o diferentes que pueden estar incluidos en la construcción que va a ser transformada. Los expertos en la especialidad comprenderán también que alternativamente el polinucleótido y el ácido nucleico pueden ser transformados dentro de la célula sin un promotor, pero la expresión del polinucleótido o del ácido nucleico o de ambos podría ser impulsada por un promotor, o promotores, endógenos a la planta transformada.

Métodos que incluyen otros pasos de procesamiento

En una realización el método comprende el paso adicional de procesar el tejido/órgano no fotosintético de la planta para formar una materia prima para alimento de animales.

En una realización más el método comprende el paso adicional de procesar el tejido/órgano no fotosintético de la planta para formar una materia prima para biocombustibles.

En una realización más el método comprende el paso adicional de extraer aceite del tejido/órgano no fotosintético de la planta.

En una realización más el método comprende el paso adicional de procesar el tejido/órgano no fotosintético para formar una fracción de aceite.

En una realización más el aceite del tejido/órgano no fotosintético es procesado para formar un combustible, aceite oleoquímico o nutritivo o cosmético, un ácido graso poliinsaturado (PUFA) o una combinación de estos.

Tejido/órgano no fotosintético de la planta producido por el método del invento

En una realización más el invento provee un tejido/órgano no fotosintético de una planta producido por un método del invento. En una realización el tejido/órgano no fotosintético es seleccionado de tejidos/órganos de la planta que están bajo la tierra. En una realización más el tejido/órgano bajo la tierra se selecciona de raíz, tubérculo, bulbo, tallo bulboso y rizoma. En una realización más el tejido/órgano no fotosintético se selecciona de raíz, tubérculo,

bulbo, tallo bulboso, rizoma y endosperma. En una realización más el tejido/órgano no fotosintético es una raíz.

En una realización el tejido/órgano no fotosintético contiene por lo menos 100%, más preferentemente 150%, más preferentemente 200%, más preferentemente 250%, más preferentemente 300%, más preferentemente 350%, más preferentemente 400%, más preferentemente 450%, más preferentemente 500%, más lípido total que el tejido/órgano no fotosintético de una planta de control.

En una realización el tejido/órgano no fotosintético contiene 100% a 900%, más preferentemente 200% a 800%, más preferentemente 300% a 700%, más preferentemente 400% a 600%, más lípido total que el correspondiente tejido/órgano no fotosintético de una planta de control.

Nivel de producción de aceite en tejidos/órganos no fotosintéticos.

En una realización el tejido/órgano no fotosintético contiene por lo menos 2x, más preferentemente 3x, más preferentemente 4x, más preferentemente 5x, más preferentemente 6x, más preferentemente 7x, más preferentemente 8x, más preferentemente 9x, más preferentemente 10x, más preferentemente 11x, más preferentemente 12x, más preferentemente 13x, más preferentemente 14x, más preferentemente 15x, más aceite que el correspondiente tejido/órgano no fotosintético de una planta de control.

En una realización el tejido/órgano no fotosintético contiene 3x a 15x, más preferentemente 4x a 14x, más preferentemente 5x a 13x , más preferentemente 6x a 12x , más preferentemente 7x a 11x , más preferentemente 8x a 10x más aceite que el correspondiente tejido/órgano no fotosintético de una planta de control.

Las plantas de control adecuadas incluyen versiones no transformadas o de tipo salvaje de una planta de la misma variedad y/o especie que la planta transformada usada en el método del invento. Las plantas de control adecuadas incluyen también plantas de la misma variedad y/o especie de la planta transformada que son transformadas con una construcción de control. Las plantas de control adecuadas incluyen también plantas que no han sido transformadas con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente. Las plantas de control adecuadas incluyen también plantas que no expresan una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

Preferentemente el incremento en nivel de producción de aceite es causado por la expresión de la oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

Alimento para animales que comprende tejido/órgano no fotosintético del invento

En un aspecto más el invento provee un alimento para animales que comprende el tejido/órgano no fotosintético del invento.

Materia prima para biocombustible que comprende tejido/órgano no fotosintético del invento

En un aspecto más el invento provee una materia prima para biocombustible que comprende el tejido/órgano no fotosintético del invento.

Fuente de oleosinas y plantas

Las oleosinas modificadas pueden ser oleosinas naturales modificadas. Las plantas de las cuales se derivan las secuencias de oleosinas no modificadas pueden ser derivadas de cualquier especie de planta que contenga oleosinas y secuencias de polinucleótidos que codifican oleosinas.

Las células vegetales, en las que se expresan las oleosinas modificadas, pueden ser de cualquier especie de planta. Las plantas, en las que se expresan las oleosinas modificadas, pueden ser de cualquier especie de planta.

En una realización la célula vegetal o planta, es derivada de una especie de planta gimnosperma. En una realización más la célula vegetal o planta, es derivada de una especie de planta angiosperma. En una realización más, la célula vegetal o planta es derivada de una especie de planta dicotiledónea. En una realización más la célula vegetal o planta es derivada de una especie de planta monocotiledónea.

Las especies de plantas preferidas son aquellas que producen tubérculos (tallos modificados) tales como la especie *Solanum* pero sin estar limitadas a ésta. Otras especies de plantas preferidas son aquellas que producen bulbos (hojas de almacenamiento subterráneas) tales como las especies *Lilaceae, Amaryllis, Hippeastrum, Narcissus, Iridaceae,* y *Oxalis* sin estar limitadas a éstas. Otras especies de plantas preferidas son aquellas que producen tallos bulbosos (tallos hinchados subterráneos) tales como las especies *Musa, Elocharis, Gladiolus* y *Colocasia* sin estar limitadas a éstas. Otras especies de plantas preferidas son aquellas que producen rizomas (tallo de almacenamiento subterráneos) tales como las especies *Asparagus, Zingiber* y *Bambuseae* sin estar limitadas a ellas. Otras especies preferidas son aquellas que producen un endosperma substancial en sus semillas, tales como maíz y sorgo, pero sin estar limitadas a estos.

Las plantas preferidas incluyen aquellas de los siguientes géneros: *Brassica, Solanum, Raphanus, Allium, Foeniculum, Lilaceae, Amaryllis,*

Hippeastrum, Narcissus, Iridaceae, Oxalis, Musa, Eleocharis, Gladiolus, Colocasia, Asparagus, Zingiber y *Bambuseae.*

Una especie preferida de *Brassica* es *Brassica rapa var. rapa* (nabo).

Las especies preferidas de *Solanum* son aquellas que producen tubérculos. Una especie preferida de *Solanum* es *Solanum tuberosum* (papa).

Las especies preferidas de *Raphanus* incluyen *Raphanus raphanistrum*, *Raphanus caudatu y Raphanus sativus*. Una especie preferida de *Raphanus* es *Raphanus sativus*(rábano).

Las especies preferidas de *Allium* incluyen: *Allium cepa* (cebolla, chalota), *Alliumfistulosum* (escalonia)*, Alliumschoenoprasum* (cebollinos)*, Allium tuberosum* (cebollino chino chives)*, Alliumampeloprasum* (puerro, kurrat, ajo de cabeza grande, cebolla perla)*, Allium*sativum (ajo) *y Alliumchinense* (rakkyo). Una especie preferida de *Allium* es *Allium cepa* (cebolla).

Las especies preferidas de *Musa*incluyen: *Musa acuminata* y *Musa balbisiana*. Una especie preferida de *Musa* es *Musaacuminata* (banana, plátano).

Una especie preferida de *Zingiber* es *Zingiber officinale* (gengibre).

Una especie preferida de *Oxalis* es *Oxalis tuberosa*(ñame).

Una especie preferida de *Colocasia* es *Colocasia esculenta*(taro).

Otro género preferido es *Zea*. Una especie preferida de *Zea* es *Zea mays*.

Otro género preferido es *Sorghum*. Una especie preferida de *Sorghum* es *Sorghum bicolor*.

Otras plantas preferidas son especies de plantas para forraje de un grupo que comprende pero no está limitado a los siguientes géneros: Zea, *Lolium, Hordium, Miscanthus, Saccharum,Festuca*, *Dactylis*, *Bromus*, *Thinopyrum*, *Trifolium*, *Medicago*, *Pheleum*, *Phalaris*, *Holcus*, *Glycine*, *Lotus*, *Plantago* y *Cichorium*.

Otras plantas preferidas son las plantas leguminosas. La planta leguminosa o parte de ella puede comprender cualquier planta en la familia de las plantas Leguminosae o Fabaceae. Por ejemplo, las plantas pueden ser seleccionadas de legumbres forrajeras que incluyen, alfalfa, trébol; leucaena; legumbres de granos que incluyen, porotos, lentejas, lupinos, arvejas, cacahuetes, soya; legumbres de flor que incluyen lupino, legumbres para uso farmacéutico o industrial; y especies de legumbres para abono verde o para barbecho.

Un género preferido particularmente es *Trifolium*. Las especies de *Trifolium* preferidas incluyen *Trifolium repens*; *Trifolium arvense*; *Trifolium affine*; y *Trifolium occidentale*. Una especie de *Trifolium* preferida particularmente es *Trifolium repens*.

Otro género preferido es *Medicago*. Las especies de *Medicago* preferidas incluyen *Medicago sativa* y *Medicago truncatula*. Una especie de *Medicago* preferida particularmente es *Medicago sativa*, conocida comúnmente como alfalfa.

Otro género preferido es *Glycine*. Las especies de*Glycine* preferidas incluyen *Glycine max* y *Glycine wightii* (conocida también como *Neonotonia wightii*). Una especie de*Glycine* preferida particularmente es *Glycine max*, conocida comúnmente como semilla de soja. Una especie de *Glycine* particularmente preferida es *Glycine wightii*, conocida comúnmente como soya perenne.

Otro género preferido es *Vigna*. Una especie de *Vigna* preferida particularmente es *Vigna unguiculata* conocida comúnmente como caupí.

Otro género preferido es *Mucana*. Las especies *Mucana* preferidas incluyen *Mucana pruniens*. Una especie de *Mucana* preferida particularmente es *Mucana pruniens* conocida comúnmente como frejol terciopelo.

Otro género preferido es *Arachis*. Una especie de *Arachis* preferida particularmente es *Arachis glabrata* conocida comúnmente como maní perenne.

Otro género preferido es *Pisum*. Una especie de *Pisum* preferida particularmente es *Pisum sativum* conocida comúnmente como arveja.

Otro género preferido es *Lotus*. Las especies preferidas de Lotus incluyen *Lotus corniculatus*, *Lotus pedunculatus*, *Lotus glabar*, *Lotus tenues yLotus uliginosus*. Una especie preferida de *Lotus* es *Lotus corniculatus* conocida comúnmente es Trébol. Otra especie preferida es *Lotus glabar* conocida comúnmente como trébol de hoja angosta. Otra especie de *Lotus* es *Lotus pedunculatus* conocida comúnmente como trébol grande. Otra especie de *Lotus* preferida es *Lotus tenuis* conocida comúnmente como trébol delgado.

Otro género preferido es *Brassica*. Una especie de *Brassica* preferida es *Brassica oleracea*, conocida comúnmente como berza forrajera y repollo.

Otras especies preferidas son los cultivos de semillas oleaginosas que incluyen, pero no están limitadas a los siguientes géneros: *Brassica*, *Carthumus*, *Helianthus*, *Zea* y *Sesamum*.

Un género preferido de semilla oleaginosa es *Brassica*. Una especie preferida de semilla oleaginosa es *Brassica napus*.

Un género preferido de semilla oleaginosa es *Brassica.* Una especie preferida de semilla oleaginosa es *Brassica oleraceae.*

Un género preferido de semilla oleaginosa es *Carthamus*. Una especie preferida de semilla oleaginosa es *Carthamus tinctorius*.

Un género preferido de semilla oleaginosa es *Helianthus*. Una especie preferida de semilla oleaginosa es *Helianthus annuus*.

Un género preferido de semilla oleaginosa es *Zea*. Una especie preferida de semilla oleaginosa es *Zea mays*.

Un género preferido de semilla oleaginosa es *Sesamum*. Una especie preferida de semilla oleaginosa es *Sesamum indicum*.

Un género preferido para ensilaje es *Zea*. Una especie preferida para ensilaje es *Zea mays*.

Un género preferido para producir grano es *Hordeum*. Una especie preferida para producir grano es *Hordeum vulgare*.

Un género preferido para pastoreo es *Lolium*. Una especie para pastoreo preferida es *Lolium perenne*.

Un género preferido para pastoreo es *Lolium*. Una especie preferida para pastoreo es *Lolium arundinaceum*.

Un género preferido para pastoreo es *Trifolium*. Una especie preferida para pastoreo es *Trifolium repens*.

Un género preferido para pastoreo es *Hordeum*. Una especie preferida para pastoreo es *Hordeum vulgare*.

Las plantas preferidas incluyen también plantas para forraje, o de alimento para animales. Tales plantas incluyen pero no están limitadas a los siguientes géneros: *Miscanthus, Saccharum, Panicum.*

Un género preferido para biocombustible es *Miscanthus*. Una especie preferida para biocombustible es *Miscanthus giganteus*.

Un género preferido para biocombustible es *Saccharum*. Una especie preferida para biocombustible es *Saccharum officinarum*.

Un género preferido para biocombustible es *Panicum*. Una especie preferida para biocombustible es *Panicum virgatum*.

Descripción Detallada del Invención

Cuando en esta memoria descriptiva se ha citado memorias de patente, otros documentos externos, u otras fuentes de información, esto es generalmente con el propósito de proveer un contexto para tratar sobre las características del invento. A menos que se exprese específicamente otra cosa, la referencia a tales documentos externos no debe ser interpretada como una admisión de que tales documentos, o tales fuentes de información, en cualquier

jurisdicción, son parte de la especialidad anterior, o forman parte del conocimiento general común en la especialidad.

El término “que comprende”, como se usa en esta memoria, significa “que consta por lo menos en parte de”. Cuando se interpreta cada afirmación en esta memoria que incluye el término “que comprende”, pueden estar también presentes otras características diferentes a aquella o aquellas a las que precede el término. Términos relacionados tales como “comprenden” y “comprende” deben ser interpretados de la misma manera.

Los lípidos tienen en una base de peso por peso aproximadamente el doble del contenido de energía que las proteínas o los carbohidratos. La mayor parte de los lípidos del mundo son producidos por plantas y la forma de lípido más densa es como un triacilglicerol (TAG). Las plantas dicotiledóneas pueden acumular hasta aproximadamente 60% del peso de sus semillas como TAG que es usado posteriormente como una fuente de energía para la germinación. Por lo tanto, una cantidad de esfuerzos han sido dirigidos al uso de semillas ricas en aceites para producir en forma sustentable suficientes lípidos para alimento de animales y como materia prima de biocombustibles.

Dado que hay sólo una cantidad limitada de TAG que puede ser producido por semillas, se están realizando propuestas alternativas para producir lípido adicional (preferentemente TAGs) en tejidos vegetativos. La mayoría de estas propuestas han perseguido la regulación positiva o la sobreexpresión de una o varias enzimas en la vía Kennedy en las hojas de plantas para sintetizar TAG. Sin embargo, típicamente la mayoría de los lípidos adicionales producidos por esta propuesta son movilizados nuevamente dentro de la planta por una combinación de lipasas y β–oxidación, lo que tiene por resultado un aumento limitado en el contenido de lípidos (generalmente 2–4% del DM).

El TAG producido en semillas en desarrollo está contenido típicamente dentro de estructuras discretas denominadas cuerpos oleosos (COs) que son altamente estables y se mantienen como orgánulos discretos empaquetados fuertemente, sin conglutinarse aún cuando las células se deshidratan o experimentan condiciones de congelamiento (Siloto y otros, 2006; Shimada y otros, 2008). Los COs constan de un núcleo de TAG rodeado por una monocapa de fosfolípidos incrustada con emulsionantes proteínicos. Estos últimos constituyen 0,5–3,5% del CO; de esto, 80–90% es oleosina y el resto consiste predominantemente en proteínas de fijación a calcio (caolesina) y proteínas de fijación a esterol (esteroleosina) (Lin y Tzen, 2004). Las propiedades emulsionantes de las oleosinas se derivan de sus tres dominios funcionales que constan de un brazo N–terminal anfipático, un núcleo hidrófobo central altamente

conservado (~72 residuos) y un brazo C–terminal anfipático. Similarmente, la caolesina y la esteroleosina poseen brazos N y C–terminales hidrófilos y su propio núcleo hidrófobo conservado.

Se especuló previamente que la expresión constitutiva de oleosina o polioleosina (fusiones de oleosinas de cabeza a cola en tándem) con enzimas sintetizadoras de TAGs en las hojas daría por resultado la formación de cuerpos oleosos estables que llevarían a la acumulación de TAG. Sin embargo, hemos encontrado posteriormente, que la oleosina y las polioleosinas son ineficaces y promueven la acumulación de TAG cuando son coexpresadas con DGAT1 en hojas de plantas (Roberts *y otros,* datos no publicados).

Se ha mostrado (WO 2011/053169) que la expresión de oleosinas modificadas con cisteínas introducidas artificialmente puede producir un incremento en el nivel de aceite de las hojas de las plantas. Sin embargo, los presentes solicitantes han mostrado ahora en forma sorprendente que es posible acumular niveles significativamente más altos en los tejidos no fotosintéticos de las plantas que en los otros tejidos de las plantas.

Cuerpos oleosos

Los COs fluctúan generalmente entre 0,5–2,5 µm de diámetro y constan de un núcleo de TAG rodeado por una monocapa de fosfolípidos incrustada de emulsionantes proteínicos – predominantemente oleosinas (Tzen y otros, 1993; Tzen, y otros 1997). Los COs constan sólo de 0,5–3,5% de proteína; de ésta 80–90% es oleosina y el resto consiste predominantemente en proteínas de fijación a calcio (caleosina) y proteínas de fijación esterol (esteroleosina) (Lin y Tzen, 2004). La razón de oleosina a TAG dentro de la célula vegetal influye en el tamaño y el número de cuerpos oleosos dentro de la célula (Sarmiento y otros, 1997; Siloto y otros, 2006).

Si bien los COs son producidos naturalmente predominantemente en las semillas y el polen de muchas plantas, también se encuentran en algunos otros órganos (por ejemplo, tubérculos específicos).

Las oleosinas son proteínas comparativamente pequeñas (15 24 kDa) que están incrustadas en la superficie de los COs.

Biohidrogenación

Se ha demostrado que el perfil lipídico de los alimentos para animales rumiantes influye a su vez en el perfil lipídico de los productos cárneos y lácteos (Demeyer y Doreau, 1999). Diferentes plantas tienen diferentes perfiles lipídicos; alimentando selectivamente a los animales sólo con plantas que tienen el perfil lipídico deseado, es posible influir positivamente en el perfil lipídico de los productos cárneos y lácteos corriente abajo. En los rumiantes la composición final

de los lípidos de la carne y la leche no sólo es influenciada por los lípidos de la dieta sino que también es influenciada fuertemente por la biohidrogenación (Jenkins y McGuire 2006; Firkins y otros, 2006; Lock y Bauman, 2004). La biohidrogenación es la hidrogenación de compuestos no reducidos (tales como grasas insaturadas) por la biota presente en el rumen. La biohidrogenación puede ser evitada/retardada encapsulando los lípidos en una proteína o en proteínas que proporcionan resistencia a la degradación microbiana (Jenkins y Bridges 2007). La prevención de la biohidrogenación encapsulando triacilglicéridos en una polioleosina o en oleosinas *in planta* fue informada por Scott y otros, (2007), Cookson y otros, (2009) y Roberts y otros, (2008).

Oleosinas

Las oleosinas son proteínas comparativamente pequeñas (15 a 24 kDa) que permiten que los COs se conviertan en orgánulos discretos empaquetados fuertemente sin conglutinarse cuando las células se desecan o experimentan condiciones de congelamiento (Leprince *y otros,* 1998; Siloto *y otros,* 2006; Slack *y otros,* 1980; Shimada *y otros,* 2008).

Las oleosinas tienen tres dominios funcionales que consisten en un brazo N–terminal anfipático, un núcleo hidrófobo central altamente conservado (~72 residuos) y un brazo anfipático C–terminal. El modelo topológico aceptado es uno en el cual los brazos anfipáticos N– y C–terminales están ubicados en el exterior de los COs y el núcleo hidrófobo central está ubicado dentro del CO (Huang, 1992; Loer y Herman, 1993; Murphy 1993). Los residuos cargados negativamente de los brazos anfipáticos N– y C–terminales están expuestos al exterior acuoso, en tanto que los residuos cargados positivamente están expuestos al interior del CO y enfrentan a los lípidos cargados negativamente. Así, los brazos anfipáticos con su carga negativa orientada hacia afuera son responsables de mantener a los COs como entidades individuales a través del impedimento estérico y la repulsa electrostática tanto *in vivo* como en preparados aislados (Tzen y otros, 1992). El brazo anfipático N–terminal es altamente variable y por lo tanto ninguna estructura secundaria específica puede describir todos los ejemplos. En comparación, el brazo C–terminal contiene un dominio α–helicoidal de 30–40 residuos (Tzen y otros, 2003). El núcleo central está muy conservado y se cree que es la región hidrófoba más larga que según se sabe ocurre en la naturaleza; en el centro se encuentra un motivo de nudo conservado de prolina de 12 residuos que incluye tres residuos de prolina espaciados (ver reseñas en Frandsen y otros, 2001; Tzen y otros, 2003). Aún no se conoce bien la estructura secundaria, terciaria y cuaternaria del dominio central. Existe evidencia de modelado, transformación de Fourier-infrarroja (FT–IR) y

dicromismo circular (CD) para una cantidad de configuraciones diferentes (Ver una reseña en Roberts y otros, 2008).

Las propiedades de las principales oleosinas están relativamente conservadas entre las plantas y están caracterizadas por lo siguiente:

- proteína de 15–25kDa que corresponde a aproximadamente 140–230 residuos de aminoácidos.
- La secuencia de proteínas puede ser dividida casi igualmente a lo largo de su longitud en 4 partes que corresponden a una región hidrófila N–terminal, dos regiones hidrófobas centrales (unidas por un nudo o botón de prolina) y una región hidrófila C–terminal.
- La topología de la oleosina es atribuida a sus propiedades físicas que incluyen un núcleo hidrófobo plegado flanqueado por dominios hidrófilos. Esta configuración confiere una naturaleza anfipática a la oleosina que tiene como resultado que el dominio hidrófobo esté incrustado en la monocapa de fosfolípidos (Tzen y otros, 1992) mientras que los dominios hidrófilos que lo flanquean están expuestos al ambiente acuoso del citoplasma.
- Típicamente las oleosinas no contienen cisteínas.

Las oleosinas preferidas para usar en el invento son aquellas que contienen un dominio central de aproximadamente 70 residuos de aminoácidos no polares (incluyendo un nudo de prolina) no interrumpido por residuos cargados, flanqueado por dos brazos hidrófilos.

El término “oleosina”, como se usa en la presente, también incluye esteroleosina y caolesina.

Esteroleosinas

Las esteroleosinas comprenden un segmento de anclaje N–terminal que contiene dos α-hélices anfipáticas (912 residuos en cada hélice) conectadas por una región hidrófoba de anclaje de 14 residuos. El dominio de deshidrogenasa soluble contiene un subdominio de ligadura NADP+– y un subdominio de ligadura a esterol. La distinción aparente entre las esteroleosinas A y B ocurre en sus diversos subdominios de ligadura a esterol (Lin y Tzen, 2004). Las esteroleosinas tienen un botón de prolina en su dominio hidrófobo y contienen una deshidrogenasa que se fija a esterol en uno de sus brazos hidrófilos.

Caolesinas

Las caolesinas (Frandsen *y otros,* 2001)tienen un nudo de prolina ligeramente diferente al de las oleosinas básicas, y contienen un motivo

que se fija a calcio y varios sitios de fosforilación potenciales en los brazos hidrófilos. De una manera similar a la oleosina, se propone que la caolesina tiene tres dominios estructurales, en donde los brazos N– y C–terminales son hidrófilos mientras que el dominio central es hidrófobo y actúa como el ancla del cuerpo oleoso. El dominio hidrófilo N–terminal consta de un motivo de mano EF que se fija a calcio hélice–bucle–hélice de 28 residuos que incluye un residuo de glicina invariable como un punto de cambio estructural y cinco residuos conservados que contienen oxígeno como ligandos que se fijan a calcio (Chen *y otros,* 1999; Frandsen *y otros,* 2001). El dominio hidrófilo C–terminal contiene varios sitios de fosforilación y cerca del extremo C se encuentra una cisteína invariable que no está involucrada en ninguna unión intra– o inter–disulfuro (Peng, 2004). Los terminales hidrófilos N y C de caoleosina son aproximadamente 3 veces más grandes que aquellos de la oleosina (Lin y Tzen, 2004). Se cree que el dominio hidrófobo consta de una α-hélice anfipática y una región de anclaje (que incluye un nudo de prolina).

En la siguiente Tabla 1 se muestra ejemplos de secuencias de oleosinas (oleosina, esteroleosina y caleosina) adecuadas para que sean modificadas para usar en el invento por la adición de por lo menos una cisteína introducida artificialmente. Las secuencias (tanto de polinucleótidos como de polipéptidos) son proporcionadas en el Listado de Secuencias.

Tabla 1

Oleosina	**Especie**	**Nº de acceso de cADN**	**SEQ ID NO:**	**Nº de acceso de proteína**	**SEQ ID NO:**
Oleosina	*S. indicum*	AF302907	34	AAG23840	35
Oleosina	*S. indicum*	U97700	36	AAB58402	37
Oleosina	*A. thaliana*	X62353	38	CAA44225	39
Oleosina	*A. thaliana*	BT023738	40	AAZ23930	41
Oleosina	*H. annuus*	X62352.1	42	CAA44224.1	43
Oleosina	*B. napus*	X82020.1	44	CAA57545.1	45
Oleosina	*Z. mays*	NM_001153560.1	46	NP_001147032.1	47
Oleosina	*O.sativa*	AAL40177.1	48	AAL40177.1	49
Oleosina	*B.oleracea*	AF117126.1	50	AAD24547.1	51
Oleosina	*C. arabica*	AY928084.1	52	AAY14574.1	53

Esteroleosina	*S. indicum*	AAL13315	54	AAL13315	55
Esteroleosina	*A. napus*	EU678274	56	ACG69522	57
Esteroleosina	*Z. mays*	NM_001159142.1	58	NP_001152614.1	59
Esteroleosina	*B. napus*	EF143915.1	60	ABM30178.1	61
Caolesina	*S. indicum*	AF109921	62	AAF13743	63
Caleosina	*G. max*	AF004809	64	AAB71227	65
Caleosina	*Z. mays*	NM_001158434.1	66	NP_001151906	67
Caleosina	*B. napus*	AY966447.1	68	AAY40837	69
Caleosina	*C. revoluta*	FJ455154.1	70	ACJ70083	71
Caleosina	*C. sativus*	EU232173.1	72	ABY56103.1	73

Las oleosinas, esteroleosinas y caleosinas son bien conocidas por los expertos en la especialidad. Otras secuencias de muchas especies diferentes pueden ser identificadas sin dificultad por métodos bien conocidos por los expertos en la especialidad. Por ejemplo, se puede identificar fácilmente otras secuencias mediante una búsqueda cruzada en las bases de datos del sistema Entrez del Centro Nacional de Información Biotecnológica (NCBI) disponible en http:// www.ncbi.nlm.nih.gov/sites/gquery) usando cualquiera de los términos oleosina, esteroleosina y caolesina.

Biosíntesis de lípidos de plantas

Todas las células vegetales producen ácidos grasos a partir de acetil–CoA por una vía común localizada en plástidos. Aunque una parte de las cadenas acilo recién sintetizadas es usada después para la biosíntesis de lípidos dentro del plástido (la vía procariótica), una parte principal es exportada al citosol para el ensamblaje de glicerolípidos en el retículo endoplasmático (ER) u otros sitios (la vía eucariótica). Además, algunos de los glicerolípidos extraplastidiales retornan al plástido, lo que tiene como resultado un entremezclado considerable entre el plástido y los depósitos de lípidos en el RE (Ohlrogge y Jaworski 1997).

La descripción más simple de la vía plastidial de la biosíntesis de ácidos grasos consiste en dos sistemas de enzimas: la acetil–CoA carboxilasa (ACCasa) y la ácido graso sintasa (FAS). La ACCasa cataliza la formación de la malonil–CoA a partir de la acetil–CoA, y la FAS transfiere la fracción malonilo a la proteína portadora de acilo (ACP) y cataliza la extensión de la cadena acilo creciente con malonilo–ACP.

La reacción inicial en la síntesis de ácidos grasos es catalizada por 3–cetoacil–ACP III (KAS III) lo que tiene como resultado la condensación de acetil–CoA y malonil–ACP. Las condensaciones subsiguientes son catalizadas por KAS I y KAS II. Antes que comience un ciclo subsiguiente de síntesis de ácidos grasos, el intermediario 3–cetoacil–ACP es reducido a la acil–ACP saturada en las reacciones de FAS restantes, catalizadas secuencialmente por la 3–cetoacil–ACP reductasa, la 3–hidroxiacil–ACP dehidrasa y la enoil–ACP reductasa.

Los productos finales de FAS son usualmente las ACPs 16:0 y 18:0, y la composición final de los ácidos grasos de una célula vegetal es determinada en gran parte por las actividades de varias enzimas que usan estas acil–ACPs en la fase de terminación de la síntesis de ácidos grasos. La estearoil–ACP desaturasa modifica el producto final de FAS por inserción de un doble enlace cis en la posición 9 de la ACP C18:0. Las reacciones de la síntesis de ácidos grasos son terminadas por hidrólisis o transferencia de la cadena acilo de la ACP. La hidrólisis es catalizada por las acil–ACP tioesterasas, de las cuales hay dos tipos principales: una tioesterasa relativamente específica para ACP 18:1 y una segunda más específica para las acil–ACPs saturadas. Los ácidos grasos que han sido liberados de las ACPs por las tioesterasas dejan el plástido y entran en la vía de los lípidos eucarióticos, en donde son esterificados primero a glicerolípidos en el RE. Las acil transferasas en el plástido, en contraste con las tioesterasas, terminan la síntesis de ácidos grasos transesterificando las fracciones acilo de ACP a glicerol, y son una parte esencial de la vía de lípidos procarióticos que lleva al ensamblaje de glicerolípidos del plástido.

Biosíntesis de triacilglicerol

El único paso comprometido en la biosíntesis de TAG es el último, es decir, la adición de un tercer ácido graso a un diacilglicerol existente, generando así TAG. En las plantas este paso es realizado principalmente (pero no en forma exclusiva) por una de cinco enzimas sintetizadoras de TAG (ubicadas principalmente en el RE) que incluyen: acil CoA: diacilglicerol aciltransferasa (DGAT1); una acil CoA no relacionada: diacilglicerol acil transferasa (DGAT2); una DGAT (DGAT3) soluble que tiene menos de 10% de identidad con la DGAT1 o la DGAT2 (Saha y otros, 2006); la fosfatidilcolina–esterol O–aciltransferasa (PDAT); y una cera sintasa (WSD1, Li y otros, 2008). Las proteínas DGAT1 y DGAT2 son codificadas por dos familias de genes distintas, DGAT1 que contiene aproximadamente 500 aminoácidos y 10 dominios transmembrana predichos y DGAT2 que tiene sólo 320 aminoácidos y dos dominios transmembrana (Shockey y otros, 2006).

El término “enzima sintetizadora de triacilglicerol” o “enzima sintetizadora de TAG”, como se usa en la presente, significa una enzima capaz de catalizar la adición de un tercer ácido graso a un diacilglicerol existente, generando así TAG. Las enzimas sintetizadoras de TAG preferidas incluyen, pero no están limitadas a: acil CoA: diacilglicerol aciltransferasa1 (DGAT1); diacilglicerol acil transferasa2 (DGAT2); fosfatidilcolina–esterol O–aciltransferasa (PDAT) y la forma soluble citosólica de DGAT (DGAT soluble o DGAT3).

Dado que las DGAT1 y DGAT2 endógenas parecen desempeñar roles en las hojas maduras y senescentes (Kaup y otros 2002; Shockey y otros 2006), es probable que las plantas posean una cantidad de mecanismos de retroalimentación para controlar su actividad. Por cierto, Zou y otros, (2008) identificaron recientemente una secuencia de consenso (X–Leu–X–Lys–X–X–Ser–X–X–X–Val) dentro de las secuencias de DGAT1 (TmDGAT1) de *Tropaeolum majus* (capuchina de jardín) como un motivo de objetivo típico de los miembros de la proteína quinasa–1 relacionada con SNF1 (SnRK1) siendo Ser el residuo para la fosforilación. Las proteínas SnRK1 son una clase de Ser/Thr proteína quinasas que han estado implicadas en forma creciente en la regulación global del metabolismo de carbono en plantas, por ejemplo, la inactivación de sacarosa fosfato sintasa por fosforilación (Halford & Hardie 1998). Zou *y otros* (2008) se propusieron demostrar que la obliteración de un sitio de fosforilación de SnRK1 en DGAT1 por mutación de un solo punto (Ser197Ala de TmDGAT1) llevaba a la acumulación de niveles significativamente mayores de TAG en la semilla. Esta mutación aumentó la actividad en 38–80%, lo que llevó a un aumento de 20–50% en el contenido de aceite en una base por semilla en Arabidopsis.

La fosfolípido:DGA aciltransferasa (PDAT) forma TAG a partir de una molécula de fosfolípido y una molécula de diacilglicerol. La PDAT es muy activa cuando se expresa en levaduras pero no aumenta apreciablemente los rendimientos de TAG cuando se expresa en semillas de plantas. La PDAT y una DAG:DAG transacilasa propuesta son enzimas sintetizadoras de lípidos neutros que producen TAG, pero no son consideradas parte de la vía Kennedy.

Se ha encontrado una combinación de cera éster sintasa y enzima DGAT (WS/DGAT) en todas las procariotas que producen lípidos neutros estudiados hasta ahora. La WS/DAGAT tiene una actividad extraordinariamente amplia en una variedad de ácidos grasos, alcoholes e incluso tioles inusuales. Esta enzima tiene una región putativa que abarca la membrana, pero no muestra homología de secuencia con las familias DGAT1 y DGAT2 de eucariotas o la WE sintasa de jojoba (Jojoba es la única eucariota que se ha encontrado que acumula éster de cera).

Debería señalarse que la Lecitina–Colesterol AcilTransferasa (LCAT) y la Acil–coenzima:Colesterol AcilTransferasa (ACAT) son enzimas que producen ésteres de esterol (una forma de lípido neutro), no TAGs.

En aplicaciones que requieren el aumento de los lípidos neutros la evidencia sugiere que la mayor actividad y la especificidad más amplia de la DGAT1 en relación con la DGAT2 es preferencial. Donde se prefiere un ácido graso específico, tal como un PUFA de cadena larga, se puede aplicar aún la DGAT1, siempre que acepte el ácido graso elegido. Las plantas incorporan generalmente los PUFAs de cadena larga en la posición sn–2. No se sabe si esto se debe a la alta actividad de LPAT o la baja actividad de DGAT1 en este substrato. Para una especificidad mejorada por los PUFAs, puede ser preferible una DGAT2 que prefiera estos ácidos grasos, o las propiedades de DGAT1 podrían ser alteradas usando la evolución dirigida o un procedimiento equivalente.

En la siguiente Tabla 2 se provee ejemplos de estas enzimas que sintetizan TAG, adecuadas para usar en los métodos y composiciones del invento, de miembros de varias especies de plantas. Las secuencias (tanto de polinucleótidos como de polipéptidos) se proveen en el Listado de Secuencias.

Tabla 2

Enzima sintetizadora de TAG	**Especie**	**Nº de acceso de cADN**	**SEQ ID NO:**	**Nº de acceso de la proteína**	**SEQ ID NO:**
DGAT1	*A. thaliana*	NM_127503	74	P_179535	75
DGAT1	*T. majus*	AY084052	76	AM03340	77
DGAT1	*Z. mays*	EU039830	78	BV91586	79
DGAT2	*A. thaliana*	NM_115011	80	P_566952	81
DGAT2	*B. napus*	FJ858270	82	C090187	83
DGAT3 (DGAT soluble)	*A. hypogaea*	AY875644	84	AX62735	85
PDAT	*A. thaliana*	NM_121367	86	P_196868	87
PDAT	*R. communis*	XM_002521304	88	XP_002521350	89

El invento contempla también el uso de enzimas sintetizadoras de TAG modificadas, que son modificadas (por ejemplo, en su secuencia por substituciones, inserciones o adiciones y similares) para alterar su especificidad y/o actividad.

Acumulación de TAG en hojas

Una reciente investigación en el terreno de 302 especies de angiospermas en la región central norte de los EE.UU. encontró que 24% tiene gotitas conspicuas de aceite citosólico en las hojas, usualmente con una gran gota de aceite por célula mesófila (Lersten *y otros,* 2006 [de Slocombe *y otros,* 2009]). Se cree que el rol citosólico del TAG en la hoja está involucrado en el almacenamiento de carbono y/o en la remodelación de los lípidos de la membrana (ver una reseña, en Slocombe *y otros,* 2009). Por cierto, en hojas senescentes, los ácidos grasos plastidiales son divididos antes en TAG para una movilización ulterior, y se cree que DGAT1 es instrumental en este proceso (Kaup *y otros,* 2002).

Ha habido varios intentos de manipulación genética de las plantas para que acumulen niveles elevados de TAG en sus hojas. El éxito de estos ha sido algo limitado por el nivel relativamente bajo de TAG acumulado y en algunos casos la mayor parte del TAG se acumuló solamente en hojas senescentes, limitando así la flexibilidad de la cosecha y la proporción de TAG que se acumuló en el cultivo en cualquier momento dado (Bouvier–Nave *y otros*, 2001; Xu *y otros,* 2005; Winichayakul *y otros,* 2008; Andrianov y otros, 2010; Slocombe *y otros,* 2009 y referencias en los mismos).

Hasta la fecha los intentos para acumular TAG en las hojas se han enfocado predominantemente en tres genes candidatos particulares que incluyen la sobreexpresión de DGAT (biosíntesis de TAG), la mutación de TGD1 o CTS (que tiene como resultado la prevención de la removilización de los lípidos), y la sobreexpresión de LEC1, LEC2 y WRI1 (factores transcripcionales involucrados en el almacenamiento de aceite y la acumulación de proteínas en las semillas en desarrollo). La sobreexpresión de TAG y otras enzimas sintetizadoras de lípidos neutros se basa en la presencia de suficiente substrato; se supone que en la hoja en expansión y/o madura éste es provisto por el plástido (cloroplasto en el caso de la hoja) que sintetiza lípidos para las membranas. En hojas fotosintéticas de Arabidopsis se ha estimado que el recambio de los lípidos de la membrana es un 4% del total de ácidos grasos por día (Bao *y otros*, 2000). En hojas senescentes, las membranas plastidiales existentes proveen la mayor parte de los ácidos grasos para ser divididos en TAG antes de una movilización ulterior.

La sobreexpresión del gen de DGAT1 de Arabidopsis en hojas de tabaco tiene como resultado un aumento en la acumulación de TAG (Bouvier–Nave *y otros,* 2001); esto fue repetido y cuantificado más tarde por Andrianov y otros, (2010). Ellos calcularon que el nivel de TAG aumentó 20 veces y llevó a duplicar el contenido de lípidos de ~3% a ~6% de materia seca en hojas maduras. Se logró un aumento adicional a 6,8% por la sobreexpresión de LEC2 (un regulador maestro de la maduración de las semillas y el almacenamiento de aceite en las semillas) en hojas maduras usando el promotor Alc inducible (Andrianov y otros, 2010). No se dio ninguna estimación del TAG extraíble, ni se realizó ningún cálculo de la acumulación de TAG en hojas en expansión.

Las mutaciones en una proteína del tipo de permeasa TRIGALACTOSILDIACILGLICEROL (TGD1), en Arabidopsis thaliana causaron la acumulación de TAGs, oligogalactolípidos y fosfatidato; esto fue acompañado por una alta incidencia de abortos de embriones y un crecimiento general de las plantas comparativamente pobre (Xu *y otros*, 2005).

Winichayakul *y otros*, (2008) sobreexpresaron DGAT1 de Arabidopsis thaliana en las hojas de ballico (*Lolium perenne*) y encontraron que esto llevó a una elevación de 50% del total de lípidos extraíbles de las hojas (de ~4% a 6% de materia seca). Además, el elevado nivel de lípidos estaba presente en hojas nuevas generadas por cosechas repetidas espaciadas de 2 a 3 semanas, indicando que las nuevas hojas emergentes también eran capaces de acumular lípido adicional. Sin embargo, típicamente el elevado nivel de lípidos en estas hojas comenzó a declinar a los niveles del tipo salvaje cuando las hojas tenían más de 2 semanas, indicando que los lípidos estaban siendo removilizados a través de catabolismo (liberación del esqueleto de glicerol por lipasa seguido por β–oxidación).

Slocombe *y otros*, (2009) demostraron que las mutaciones en el transportador ABC peroxisomal CTS (cts–2) llevó a una acumulación hasta de 1,4% de TAG en las hojas, particularmente al comienzo de la senescencia. Ellas expresaron también LEC2 ectópicamente durante la senescencia en el antecedente de cts–2; si bien esto no elevó la acumulación general de TAG con respecto a la mutante cts–2, ciertamente aumentó la acumulación de TAG en la especie del tipo de aceite de semilla en tejido senescente. Si bien cts–2 bloquea la desintegración de los ácidos grasos, llevó también a un fenotipo riguroso. Slocombe *y otros*, (2009) concluyeron que los ácidos grasos de la membrana reciclados pueden ser redirigidos a TAG expresando el programa de semillas en el tejido senescente o por un bloqueo en la descomposición de los ácidos grasos.

Scott *y otros,* (2007) reivindicaron que la coexpresión de una enzima sintetizadora de triacilglicérido y polioleosina (dos o más unidades de oleosina fusionadas en una configuración en tándem de cabeza a cola) permitiría el almacenamiento de lípidos en una célula vegetal. De una manera similar, Cookson *y otros,* (2009) reivindicaron que la producción de una sola oleosina y una enzima sintetizadora de TAG dentro de las partes vegetativas de una planta llevaría a un incremento en el número de cuerpos oleosos y TAG en el tejido vegetativo. El uso de cualquiera de estas técnicas lleva a un aumento máximo del contenido de lípidos (no necesariamente en la forma de TAG) hasta aproximadamente 50%. Además este nivel empieza a declinar a medida que las hojas maduran; típicamente en hojas de más de 2 semanas (datos no publicados).

Por lo tanto, el grado en el que se puede acumular TAG en tejidos vegetativos parece estar limitado en cierta medida por el hecho de que la maquinaria endógena de recuperación del carbono fijado cataboliza el TAG.

Senescencia foliar – reciclado de lípidos por medio de los intermediarios de TAG

La senescencia foliar es una secuencia de eventos altamente controlada que llevan últimamente a la muerte de las células, tejidos y finalmente del órgano completo. Esto incluye el reclutamiento regulado de nutrientes junto con su translocación desde el tejido senescente a otros tejidos que aún están creciendo y desarrollándose. El cloroplasto es el primer orgánulo de las células del mesófilo en mostrar síntomas de senescencia y aunque la descomposición de las membranas de los tilacoides se inicia tempranamente en la cascada de la senescencia foliar, la envoltura del cloroplasto permanece relativamente intacta hasta las últimas etapas de la senescencia. DGAT1 es regulada positivamente durante la senescencia de las hojas de Arabidopsis y esto se correlaciona temporalmente con el incremento en los niveles de los ácidos grasos que contienen TAG encontrados comúnmente en los galactolípidos del cloroplasto. El reclutamiento del carbono de la membrana de las hojas senescentes, en particular los cloroplastos senescentes, hacia partes de la planta en crecimiento es una característica clave de la senescencia foliar, e involucra la desesterificación de los lípidos de los tilacoides y la conversión de los ácidos grasos libres resultantes en sacarosa móvil en el floema. La desesterificación de los lípidos en las membranas de los tilacoides parece estar mediada por una o más galactolipasas inducidas por la senescencia. La formación del TAG parece ser una etapa intermedia en la movilización de carbono desde los lípidos de membrana a la sacarosa móvil en el floema durante la senescencia (Kaup *y otros*, 2002).

Oleosinas modificadas manipuladas genéticamente para que incluyan cisteínas introducidas artificialmente

Las oleosinas modificadas para usar en los métodos del invento, se modifican para que contengan por lo menos un residuo de cisteína introducido artificialmente. Preferentemente las oleosinas manipuladas genéticamente contienen por lo menos dos cisteínas.

Se puede usar varios métodos bien conocidos por los expertos en la especialidad en la producción de las oleosinas modificadas con cisteínas introducidas artificialmente.

Tales métodos incluyen mutagénesis dirigida al sitio (US 6.448.048) en la cual el polinucleótido que codifica una oleosina es modificado para introducir una cisteína en la proteína de oleosina codificada.

Alternativamente, el polinucleótido que codifica las oleosinas modificadas puede ser sintetizado en su totalidad.

Metodología adicional para producir oleosinas modificadas y para usar en los métodos del invento, se proporciona en la sección de los Ejemplos.

La cisteína introducida puede ser un aminoácido adicional (es decir, una inserción) o puede reemplazar a un aminoácido existente (es decir, un reemplazo). Preferentemente la cisteína introducida reemplaza un aminoácido existente. En una realización preferida el aminoácido reemplazado es un residuo cargado. Preferentemente se predice que el residuo cargado está en los dominios hidrófilos y en consecuencia es probable que se ubiquen en la superficie del cuerpo oleoso.

Las regiones/brazos hidrófilas e hidrófobas de la oleosina pueden ser identificadas fácilmente por los expertos en la técnica usando metodología estándar (por ejemplo: Kyte y Doolitle (1982).

Las oleosinas modificadas para usar en los métodos del invento están preferentemente en un intervalo de peso molecular de 5 a 50 kDa, más preferentemente de 10 a 40 kDa, más preferentemente de 15 a 25 kDa.

Las oleosinas modificadas para usar en los métodos del invento están preferentemente en el intervalo de tamaño de 100 a 300 aminoácidos, más preferentemente de 110 a 260 aminoácidos, más preferentemente de 120 a 250 aminoácidos, más preferentemente de 130 a 240 aminoácidos, más preferentemente de 140 a 230 aminoácidos.

Preferentemente las oleosinas modificadas comprenden una región hidrófila N–terminal, dos regiones hidrófobas centrales (unidas por un nudo o glóbulo de prolina) y una región hidrófila C–terminal.

Preferentemente la longitud de las oleosinas modificadas puede ser dividida en cuatro partes casi iguales que corresponden a la región (o brazo) hidrófila N–terminal, las dos regiones centrales hidrófobas (unidas por un nudo o glóbulo de prolina) y una región (o rama) hidrófila C–terminal.

Preferentemente la topología de la oleosina modificada es atribuida a sus propiedades físicas que incluyen un centro hidrófobo plegado flanqueado por dominios hidrófilos.

Preferentemente las oleosinas modificadas pueden ser formadas en cuerpos oleosos cuando se combinan con triacilglicerol (TAG) y fosfolípido.

Preferentemente la topología confiere a la oleosina modificada una naturaleza anfipática que tiene como resultado que el dominio hidrófobo queda encajado en la monocapa de fosfolípido del cuerpo oleoso mientras que los dominios hidrófilos flanqueantes son expuestos al ambiente acuoso fuera del cuerpo oleoso, tal como en el citoplasma.

En una realización la oleosina modificada para usar en el método del invento comprende una secuencia que tiene por lo menos 70% de

identidad con el dominio hidrófobo de cualquiera de las secuencias de proteína de la oleosina a las que se hizo referencia en la Tabla 1 anterior.

En una realización la oleosina modificada para usar en el método del invento comprende una secuencia que tiene por lo menos 70% de identidad con cualquiera de las secuencias de proteína a las que se hizo referencia en la Tabla 1 anterior.

En otra realización la oleosina modificada es esencialmente igual a cualquiera de las oleosinas a las que se hizo referencia en la Tabla 1 anterior, con excepción de la cisteína o cisteínas adicionales introducidas artificialmente.

En otra realización la oleosina modificada del invento o usada en el método del invento, comprende una secuencia que tiene por lo menos 70% de identidad con la secuencia de oleosina de la SEQ ID NO: 16.

En otra realización la oleosina modificada tiene la misma secuencia de aminoácidos que aquella de la SEQ ID NO: 16, con excepción de la cisteína o cisteínas adicionales introducidas artificialmente.

En otra realización la oleosina modificada tiene la secuencia de aminoácidos de cualquiera de las SEQ ID NOS: 16 a 20.

Resumen de la fotosíntesis

El proceso general con el cual las algas y plantas utilizan la luz para sintetizar compuestos orgánicos se denomina fotosíntesis (Figura 19). La fotosíntesis abarca una serie compleja de reacciones que incluyen absorción de luz, producción de energía almacenada y reducción de potencia (las Reacciones Luminosas). Incluye también una vía enzimática en múltiples pasos que las utiliza para convertir CO_2 y agua en carbohidratos (el ciclo de Calvin, Figura 20). En las plantas, las reacciones biofísicas y bioquímicas de la fotosíntesis ocurren dentro de un solo cloroplasto (fotosíntesis C3) pero pueden ser separadas también en cloroplastos de tipos de células diferentes (fotosíntesis C4).

La fijación de carbono es una reacción de reducción-oxidación donde la fotosíntesis provee la energía para impulsar este proceso, así como también los electrones requeridos para convertir CO_2 en carbohidrato (Figura 19). Estos dos procesos se realizan a través de una secuencia diferente de reacciones químicas y en diferentes compartimientos celulares. En la primera etapa, se utiliza la luz para generar las moléculas para almacenamiento de energía ATP y NADPH. Las membranas tilacoides contienen los complejos fotosintéticos de multiproteínas, Fotosistemas I y II (PSI y PSII), los cuales incluyen los centros de reacción responsables de convertir la energía luminosa en energía de enlaces químicos (a través de una cadena de transferencia de electrones). La cadena fotosintética de

transferencia de electrones mueve los electrones desde el agua hacia adentro del lumen de los tilacoides a compuestos solubles con actividad reductora-oxidante en el estroma. Un subproducto de este proceso (Reacción de Hill) es el oxígeno.

La segunda parte del ciclo fotosintético es la fijación de CO_2 en azúcares (Ciclo de Calvin, Figura 20); esto ocurre en el estroma y utiliza las moléculas ATP y NADPH generadas en la reacción luminosa.

Rubisco

La ribulosa bifosfato carboxilasa (Rubisco) es la enzima clave responsable de la asimilación fotosintética de carbono al catalizar la reacción de CO_2 con ribulosa 1,5 biofosfato (RuBP) para formar dos moléculas de ácido D-fosfoglicérico (PGA) (Parry y otros, 2003). Como Rubisco opera muy lentamente, catalizando solamente la reacción de unas pocas moléculas por segundo, se requiere grandes cantidades de la enzima; en consecuencia, Rubisco integra 30-50% de la proteína soluble en las hojas (Bock y Khan, 2004). La modificación genética para acrecentar la tasa catalítica de Rubisco tendría gran importancia. Parry y otros, (2003) revisaron el avance hasta la fecha, concluyendo que hay todavía muchas barreras técnicas por superar y hasta la fecha todas las tentativas de manipulación genética para producir una mejor Rubisco han fracasado.

En presencia de O_2, Rubisco efectúa también una reacción de oxigenasa, la cual inicia el ciclo fotorrespiratorio o C2 (Figura 21) por la formación de fosfoglicolato y 3-fosfoglicerato (3-PGA). El reciclado de fosfoglicolato tiene como resultado una pérdida indirecta desde la célula de nitrógeno y CO_2 fijados, los que necesitan ser recuperados. La modificación genética para incrementar la especificidad de Rubisco para CO_2 con respecto a O_2 y para incrementar la tasa catalítica de Rubisco en plantas de cultivo tendría gran importancia agronómica. Parry *y otros*, (2003) revisaron el avance hasta la fecha, concluyendo que hay aún muchas barreras técnicas por superar y hasta hoy todas las manipulaciones genéticas no han logrado producir una mejor Rubisco (Peterhansel y otros, 2008). Se ha demostrado, además, que se requiere fotorrespiración en las plantas C3 para proteger las plantas de fotooxidación bajo una alta intensidad luminosa (Kozaki y Takeba 1996).

Ciclos C3 y C2

En las plantas C3 en condiciones atmosféricas, aproximadamente tres de cuatro reacciones enzimáticas de Rubisco en las plantas C3 fijan CO_2 (reacción de carboxilasa, ciclo C3, Figura 20). Sin embargo, la cuarta reacción cataliza una reacción de oxigenasa (Figura 3) la que tiene indirectamente como resultado una pérdida neta de CO_2 y NH_4^+ fijados y la producción de una cantidad de metabolitos intermedios a través del ciclo C2 (fotorrespiración) (Figura 22). Por

último, esto provoca un costo metabólico substancial a través de la nueva fijación de CO_2 y NH_4^+, así como también el reciclado de los compuestos intermedios. Por otra parte, cuando las plantas C3 experimentan estrés hídrico y/o temperaturas elevadas, la porción de las reacciones de oxigenasa a carboxilasa se eleva debido al elevado O_2 dentro de la hoja. Sin embargo, se ha demostrado que se requiere fotorrespiración en las plantas C3 para proteger las plantas de fotooxidación bajo una alta intensidad luminosa (Kozaki y Takeba, 1996) y al parecer proporciona gran parte de la potencia reductora requerida para la asimilación de NO_3^- en la hoja (Rachmilevitch *y otros*, 2004).

Los organismos capaces de fotosíntesis oxigénica iniciaron su evolución en una atmósfera extremadamente diferente (Giordano y otros, 2005). Uno de los cambios más impresionantes ha sido la elevación en la razón de O_2:CO_2 donde la competencia entre estos dos gases por el sitio activo de Rubisco se ha hecho progresivamente restrictiva para la tasa de fijación de carbono. Sin embargo, algunos han sugerido que el cambio gradual parece haber proporcionado una falta de presión evolutiva para Rubisco con una alta afinidad hacia CO_2 o una Rubisco sin actividad de oxigenasa. Ciertamente, se considera que las Rubisco de plantas son evolutivamente más recientes que las Rubisco de algas y así ellas son mucho más selectivas hacia CO_2 en relación con O_2. Las modificaciones genéticas para incrementar la especificidad de Rubisco hacia CO_2 en relación con O_2 han fracasado (Parry, Andralojc y otros, 2003).

Un rol significativo del ciclo de carbono fotosintético oxidativo C_2 o vía fotorrespiratoria es el reciclado del 2-fosfoglicolato (2PG) producido por la actividad de oxigenasa de Rubisco (Tolbert 1997). 2PG es tóxico para la célula; en consecuencia es desfosforilada rápidamente (a través de fosfoglicolato fosfatasa, PGP) a glicolato (Tolbert y otros, 1983). Además, se ha demostrado que la fotorrespiración es requerida en las plantas C3 para proteger las plantas de fotooxidación bajo una alta intensidad luminosa (Kozaki y Takeba 1996).

Las enzimas que oxidan glicolato a glicoxilato en la vía fotorrespiratoria son caracterizadas en dos grupos estructuralmente diferentes. En las plantas superiores, la glicolato oxigenasa que contiene FMN, GOX (EC 1.1.3.15), localizada en el peroxisoma) cataliza la oxidación de glicolato usando oxígeno molecular como el aceptor de electrones terminal y tiene una estereoespecificidad hacia L-lactato como un substrato alternativo. En contraste, la glicolato deshidrogenasa, GDH (EC 1.1.99.14) ha sido caracterizada sólo por su reacción enzimática que no requiere oxígeno y su estereoespecificidad hacia D-lactato como un substrato alternativo. En la mayoría de las algas, el glicolato es oxidado en las mitocondrias usando una GDH monomérica que depende de

cofactores orgánicos. La capacidad de la reacción parece ser limitada por los cofactores orgánicos y por lo tanto, muchas algas excretan glicolato, en el medio en condiciones de crecimiento fotorrespiratorio (Bari y otros, 2009; Colman y otros, 1974). *GDH* en *C. reinhardtii* es un gen ubicado en las mitocondrias, de baja respuesta al CO_2 (Nakamura y otros, 2005). Otros homólogos de GDH incluyen la denominada glicolato oxidasa (GOX) de *E. coli* y otras bacterias. En *E. coli*, el complejo GOX se compone de tres subunidades funcionales, GlcD, GlcE, y GlcF de la cuales GlcD y GlcE comparten una secuencia de aminoácidos muy conservada que incluye una región putativa de unión a flavina. En la proteína GlcF se ha reconocido dos motivos CxxCxxCxxxCP muy conservados, los cuales representan los típicos grupos de hierro-azufre 2x[4Fe-4S], encontrados también en la subunidad de deshidrogenasa anaeróbica G3P, y homólogos de ubiquinona oxidoreductasa de procariotas y eucariotas (Nakamura y otros, 2005).

Ciclo C4

No todas las plantas usan Rubisco para generar 3-PGA como el primer intermediario fotosintético estable. El maíz, la caña de azúcar, numerosos pastos tropicales y algunas plantas dicotiledóneas (por ejemplo, Amaranto) usan inicialmente fosfoenolpiruvato para fijar el carbono, formando ácidos orgánicos con 4 carbonos (plantas C_4). Las plantas C4 evitan el ciclo C2 a través de modificaciones en su arquitectura que comprende dos tipos diferentes de células que contienen cloroplastos, células mesófilas y células de la vaina del haz, lo cual aísla a Rubisco en un entorno relativamente rico en CO_2 incrementando con ello la proporción de las reacciones de carboxilasa. Esto permite a estas plantas usar inicialmente fosfoenolpiruvato para fijar carbono, formando ácidos orgánicos con 4 carbonos (de aquí plantas C_4). Así, el metabolismo C4 comprende fijar carbono inorgánico en un tipo de célula (mesófila), transportarlo a un tipo de célula blindada parcialmente del oxígeno atmosférico (vaina del haz), y liberar el carbono inorgánico cerca de Rubisco en este ambiente privado de oxígeno.

Las hojas de las plantas C_4 muestran una anatomía inusual que comprende dos tipos diferentes de células que contienen cloroplastos, células mesófilas y células de la vaina del haz donde las células mesófilas rodean las células de la vaina del haz las cuales rodean a su vez el tejido vascular; los cloroplastos de las células mesófilas contienen todos los complejos transmembrana requeridos para las reacciones luminosas de la fotosíntesis pero poco o nada de Rubisco, en tanto que los cloroplastos de las células de la vaina del haz carecen de tilacoides apilados y contienen poco PSII. Las plantas C_4 concentran CO_2 en las células de la vaina del haz reprimiendo eficazmente la actividad de oxigenasa de las Rubisco y eliminando la fotorrespiración.

El oxaloacetato es generado de HCO_3^- y fosfoenolpiruvato (PEP) por la fosfoenolpiruvato carboxilasa (PEPC) en el citosol de las células mesófilas. Se utiliza el Ion HCO_3^- ya que su equilibrio acuoso es preferido al CO_2 gaseoso. Además, la PEP carboxilasa no puede fijar oxígeno, que tiene una estructura 3D similar a aquella del CO_2 pero no de HCO_3^-. Dependiendo de la planta C_4, el oxaloacetato es oxidado a malato o condensado con glutamato para formar aspartato y α-cetoglutarato. El malato y el aspartato son transportados hacia adentro de las células de la vaina del haz y descarboxilados liberando CO_2, el cual queda disponible después para Rubisco y para ser incorporado en el ciclo de Calvin.

El inconveniente agronómico de esta modificación desarrollada es un incremento en la fibra de las hojas que tiene como resultado una digestibilidad comparativamente mala de las hojas de las plantas C4 (por ejemplo maíz, caña de azúcar, numerosos pastos tropicales y algunas plantas dicotiledóneas tal como el amaranto). Hasta la fecha, la modificación de una planta C3 para emular el proceso C4 total está más allá de la biotecnología actual. Además, las tentativas de manipular genéticamente Rubisco para obliterar la actividad de la oxigenasa o para reducir la afinidad hacia O_2 han fracasado (ver reseña en Peterhansel *y otros*, 2008).

Interacción con la asimilación de nitrato

La reducción de la fotorrespiración mediante manipulación de CO_2 atmosférico a través de largos períodos ha llevado a la reducción inesperada de la asimilación de nitrato en las plantas C3 (Rachmilevitch y otros, 2004). Hay una cantidad de posibles explicaciones que incluyen la disminución de la energía reductora disponible, la reducción de ferredoxina y NADH, de las cuales la primera es requerida para la nitrato reductasa y la glitamato sintetasa en tanto que la segunda es requerida para la reducción de NO_3^- (donde la NADH es producida durante el paso fotorrespiratorio de la glicina descarboxilasa en las mitocondrias). Además, el transporte de NO_2^- desde el citosol hacia el interior del cloroplasto involucra la difusión neta de HNO_2 o el cotransporte de protones y NO_2^- a través de la membrana del cloroplasto. Esto requiere que el estroma sea más alcalino que el citosol, pero la gradiente del pH es disipada un tanto por niveles elevados de CO_2. Rachmilevitch y otros (2004) concluyeron que la actividad de la nitrato reductasa por sí sola no era limitante para la asimilación de nitrato con fotorrespiración reducida. Ellos concluyeron también que la forma del nitrógeno disponible para la planta era lo que determinaba el grado en el cual niveles elevados de CO_2 tendrían como consecuencia un incremento en la producción primaria neta, es decir, donde NH_4^+ es la forma de nitrógeno dominante. Esto

sugeriría que en ausencia de prácticas que cambien la fertilización agronómica, las legumbres se beneficiarán más con la reducción de la fotorrespiración ya que la simbiosis de rhizobium y leguminosas tiene como resultado la fijación del nitrógeno atmosférico en la forma de NH_4^+ en vez de NO_3^-.

Esfuerzos anteriores para manipular genéticamente niveles más altos de CO_2 en cloroplastos y reducción de fotorrespiración en plantas C3

Numerosas investigaciones han sido realizadas en plantas superiores para abordar las limitaciones de la fotorrespiración. Esencialmente sólo una de ellas parece tener aplicaciones potenciales en la adaptación a las plantas superiores.Un reciente bypass fotorrespiratorio que incrementó la eficiencia de reciclado del glicolato fue manipulada genéticamente con éxito en Arabidopsis y tuvo como resultado un incremento de 30% en biomasa de las hojas (Kebeish y otros, 2007). Kebeish y otros (2007) transformaron Arabidopsis para que expresara tres genes de *E. coli*: glicolato deshidrogenasa (GDH), glioxilato carboxiligasa (GCL), semialdehído tartrónico reductasa (TSR) en sus cloroplastos (Figura 23). Combinados, estos genes reciclaron glicolato a glicerato en el cloroplasto, en otras palabras sin que se involucrara el peroxisoma o la mitocondria. La GDH de E. coli es un heterotrímero que consta de *glc*D, *glc*E y *glc*F que tiene como resultado plantas con un incremento de 30% en biomasa de las hojas al fin del período de crecimiento (Figura 24). Esta vía incluyó un paso de liberación de CO_2 de los cloroplastos que redujo adicionalmente la actividad de oxigenasa de Rubisco *in vivo*. Además, se creyó que con el bypass se ahorraba energía y equivalentes reductores ya que no tiene como resultado la liberación de amonio y se ahorra la energía de oxidación de glicolato en los equivalentes reductores y no es consumida durante la formación de H_2O_2 (Maurino y Peterhansel 2010). Peterhansel (2011) concluyó que para transformar realmente una planta C3 en una planta C4 se requerirá la transferencia eficiente de múltiples genes.

Biosíntesis de lípidos de plantas

Todas las células vegetales producen ácidos grasos de acetil-CoA por una vía común localizada en plástidos (Figura 25). Una porción de las cadenas acilo recién sintetizadas es utilizada después para la biosíntesis de lípidos dentro del plástido (la vía procariota); sin embargo, una porción importante es exportada hacia el interior del citosol para el ensamblaje de glicerolípidos en el retículo endoplásmico (RE) u otros sitios (la vía eucariota). Además, algunos de los glicerolípidos extraplastidiales retornan al plástido, lo que tiene como resultado una mezcla considerable entre el plástido y los depósitos de lípidos en el RE (Ohlrogge y Jaworski 1997).

La descripción más simple de la vía plastidial de la biosíntesis de ácidos grasos consta de dos sistemas de enzimas: acetil-CoA carboxilasa (ACCasa) y ácido graso sintasa (FAS). La ACCasa cataliza la formación de malonil-CoA de acetil-CoA, y la FAS transfiere la fracción malonilo a la proteína portadora de acilo (ACP) y cataliza la extensión de la cadena creciente de acilo con malonil-ACP.

La reacción inicial de la síntesis de ácidos grasos es catalizada por 3-cetoacil-ACP III (KAS III) lo que tiene como resultado la condensación de acetil-CoA y malonil-ACP. Condensaciones posteriores son catalizadas por KAS I y KAS II. Antes que comience un ciclo siguiente de síntesis de ácidos grasos, la 3-cetoacil-ACP intermedio es reducida a acil-ACP saturada en el resto de las reacciones de la FAS, es catalizada secuencialmente por la 3-cetoacil-ACP reductasa, 3 hidroxiacil-ACP deshidrasa, y la enoil-ACP reductasa.

Los productos finales de FAS son usualmente ACP 16:0 y 18:0, y la composición final de ácidos grasos de una célula vegetal es determinada en gran parte por las actividades de varias enzimas que utilizan estas acil-ACPs en la fase de terminación de la síntesis de ácidos grasos. La estearoil-ACP desaturasa modifica el producto final de la FAS por inserción de un doble enlace cis en la posición 9 de la ACP C18:0. Las reacciones de síntesis de ácidos grasos son terminadas por hidrólisis o transferencia de la cadena acilo de la ACP. La hidrólisis es catalizada por acil-ACP tioesterasas, de las que hay dos tipos principales: una tioesterasa relativamente específica para ACP 18:1 y una segunda más específica para las acil-ACPs saturadas. Los ácidos grasos que han sido liberados de ACPs por tioesterasas dejan el plástido y entran en la vía eucariótica del metabolismo de los lípidos, donde ellos son esterificados principalmente a glicerolípidos en el RE. En contraste con las tioesterasas, las acil transferasas en el plástido terminan la síntesis de ácidos grasos transesterificando las fracciones acilo de ACP a glicerol, y ellas son una parte esencial de la vía de lípidos de procariotas que lleva al ensamblaje de glicerolípidos del plástido.

Enlace predicho entre elevación de la biosíntesis de lípidos, niveles más altos de CO_2 en el cloroplasto y reducción de fotorrespiración en el cloroplasto

En las semillas verdes se descubrió recientemente que Rubisco sin el ciclo de Calvin pasa por alto la parte superior de la glicólisis en los plástidos y proporciona una mayor eficiencia en el uso del carbono que permite volver a fijar el CO_2 formado por el complejo plastídico piruvato deshidrogenasa (Schwender y otros, 2004). La acetil-CoA producida en los plástidos a partir del piruvato es activada a malonil-CoA; el grupo malonilo es transferido posteriormente a la ACP

dando origen a malonil-ACP, el substrato primario del complejo ácido graso sintasa. La formación de malonil-CoA es el paso obligado en la síntesis de ácidos grasos y es catalizada por el complejo plastídico acetil-CoA carboxilasa altamente regulado (Nikolau y otros, 2003).

Se ha especulado que cuando las hojas sintetizan triacilglicérido (TAG) la nueva fijación del CO_2 liberado por la activación de piruvato a malonil-CoA será fijada de nuevo por fotosíntesis (Durret y otros 2008). Los ácidos grasos sintetizados en el plástido son transportados al RE y acilados secuencialmente en un esqueleto de glicerol a través de la via Kennedy. Esto culmina en la producción de TAG a través de la sobreexpresión de la enzima DGAT. En este caso el ácido 3-fosfoglicérico es sintetizado por Rubisco (sin el ciclo de Calvin) en vez de una transformación de azúcares. La transformación siguiente de éste a acetil-CoA (a través del intermediario piruvato) tiene como resultado la liberación de CO_2 en el cloroplasto (Figura 26). Esto incrementa la presión parcial de CO_2 en relación con O_2 en el cloroplasto, reduciendo así la proporción de los ciclos C2 a C3 iniciados por Rubisco. Sin embargo, se ha descubierto que el catabolismo siguiente de este TAG anula esta ventaja (Winichayakul y otros, 2008). La sobreexpresión de DGAT lleva a la acumulación de TAG que es degradado posteriormente por lipasas, teniendo como resultado la liberación de ácidos grasos libres. Algunos de estos ácidos grasos libres son catabolizados por β-oxidación en el peroxisoma, en tanto que otros originan un ciclo inútil volviendo a entrar en el RE donde ellos son reincorporados en el TAG (Figura 27). Este ciclo inútil resultante reduce la demanda de síntesis *de-novo* de nuevos lípidos; posteriormente el nivel de CO_2 reciclado dentro del cloroplasto se reduce a los niveles del tipo salvaje (o se acerca a estos), lo que lleva a la reanudación de la razón de tipo salvaje de los ciclos C2 a C3 que está siendo efectuada por Rubisco dentro de la célula fotosintética C3.

Sin estar limitados por la teoría, los solicitantes proponen el siguiente modelo para el incremento observado en la asimilación de CO_2. La coexpresión de DGAT y una oleosina modificada que contiene residuos de cisteína manipulados genéticamente lleva a la acumulación de TAG que es encapsulado por la oleosina modificada que contiene los residuos de cisteína manipulados genéticamente (Figura 28). Esto impide la degradación de TAG por lipasas y así impide también el reciclado inútil de lípidos. En consecuencia, esto asegura una demanda continua de síntesis *de-novo* de lípidos y la siguiente elevación de la presión parcial de CO_2 en la célula fotosintética, lo cual tiene a su vez como resultado una supresión continuada de ciclos C2 en relación con los ciclos C3 y una elevación de la tasa de asimilación CO_2.

Posteriormente, esto debería tener como resultado una cantidad de beneficios para todos los organismos multicelulares y unicelulares fijando inicialmente el carbono usando la vía fotosintética de C3, los que incluyen:

- Incremento en la concentración de CO_2 en el cloroplasto
- Reducción de la fotorrespiración
- Elevación de la biomasa
- Elevación del rendimiento de semillas/frutos/órganos de almacenamiento
- Elevación de la eficiencia en el uso del agua
- Elevación de la tolerancia a la sequía
- Elevación de la tolerancia al oxígeno
- Elevación de la eficiencia en el uso del nitrógeno
- Reducción de la pérdida de carbono fijado
- Retardo de la floración

Tejidos/órganos no fotosintéticos

El término tejidos/órganos no fotosintéticos significa tejidos u órganos de la planta que no experimentan una fotosíntesis substantiva durante el ciclo de vida normal de la planta.

Los expertos en la especialidad comprenden que se puede hacer que aún los tejidos/órganos no fotosintéticos experimenten fotosíntesis por exposición a la luz, pero cuando lo hacen, el nivel de la fotosíntesis no es “substantivo” y es insignificante en relación con aquel realizado por tejidos fotosintéticos normales.

En una realización el tejido/órgano no fotosintético se selecciona de tejidos/órganos de la planta bajo tierra. En otra realización el tejido/órgano bajo tierra se selecciona de raíz, tubérculo, bulbo, tallo bulboso y rizoma. En una realización más el tejido/órgano no fotosintético se selecciona de raíz, tubérculo, bulbo, tallo bulboso rizoma y endosperma. En una realización más el tejido/órgano no fotosintético es una raíz.

Promotores de tejidos/órganos específicos y preferidos

Un promotor preferido de tejidos/órganos es un promotor que impulsa la expresión de un polinucleótido ligado operativamente en un tejido/órgano particular a un nivel más alto que en otros tejidos/órganos. Un promotor específico de tejidos es un promotor que impulsa la expresión de un polinucleótido ligado operativamente en forma específica en un tejido/órgano particular. Aún con los promotores específicos de tejidos/órganos, hay generalmente un pequeño grado de expresión por lo menos en otro tejido. Por definición, un promotor específico de tejidos es también un promotor preferido de tejidos.

Tejidos vegetativos

Los tejidos *v*egetativos incluyen brotes, hojas, raíces, tallos. Un tejido vegetativo preferido es una hoja.

Promotores específicos de tejido vegetativo

Un ejemplo de un promotor específico de tejido vegetativo se encuentra en US 6.229.067; y US 7.629.454; y US 7.153.953; y US 6.228.643.

Promotores específicos de polen

Un ejemplo de un promotor específico de polen se encuentra en US 7.141.424; y US 5.545.546; y US 5.412.085; y US 5.086.169; y US 7.667.097.

Promotores específicos de semillas

Un ejemplo de un promotor específico de semillas se encuentra en US 6.342.657; y US 7.081.565; y US 7.405.345; y US 7.642.346; y US 7.371.928.

Promotores específicos de fruta

Un ejemplo de un promotor específico de fruta se encuentra en US 5.536.653; y US 6.127.179; y US 5.608.150; y US 4.943.674.

Promotores preferidos de tejidos no fotosintéticos

Los promotores preferidos de tejidos no fotosintéticos incluyen aquellos expresados preferentemente en órganos/tejidos no fotosintéticos de la planta.

Los promotores preferidos de tejidos no fotosintéticos pueden incluir también promotores reprimidos por la luz.

Promotores reprimidos por la luz

Un ejemplo de un promotor reprimido por la luz se encuentra en US 5.639.952 y en US 5.656.496.

Promotores específicos de raíces

Un ejemplo de un promotor específico de raíces se encuentra en US 5.837.848; y US 2004/0067506 y US 2001/0047525.

Promotores específicos de tubérculos

Un ejemplo de un promotor específico de tubérculos se encuentra en US 6.184.443.

Promotores específicos de bulbos

Un ejemplo de un promotor específico de bulbos se encuentra en Smeets *y otros*, (1997) Plant Physiol. 113:765-771.

Promotores preferidos de rizomas

Un ejemplo de un promotor preferido de rizomas se encuentra en Seong Jang *y otros*, (2006) Plant Physiol. 142:1148-1159.

Promotores específicos de endosperma

Un ejemplo de un promotor específico de endosperma se encuentra en US 7.745.697.

Promotores de tallos bulbosos

Un ejemplo de un promotor capaz de impulsar la expresión en un tallo bulboso se encuentra en Schenk *y otros*, (2001) Plant Molecular Biology, 47:399-412.

Promotores preferidos de tejidos fotosintéticos

Los promotores preferidos de tejidos fotosintéticosincluyen aquellos que se expresan preferentemente en tejidos fotosintéticosde las plantas. Los tejidos fotosintéticos de la planta incluyen hojas, tallos, brotes y partes de la planta que están encima de la tierra. Los promotores preferidos de tejidos fotosintéticosincluyen promotores regulados por la luz.

Promotores regulados por la luz

Los expertos en la especialidad conocen numerosos promotores regulados por la luz que incluyen, por ejemplo, promotores que se unen a clorofila a/b (Cab) y promotores de Pequeña Subunidad de Rubisco (SSU). Un ejemplo de un promotor regulado por la luz se encuentra en US 5.750.385. Regulado por la luz en este contexto significa inducible por la luz o inducido por la luz.

Términos relativos

Los términos relativos tales como acrecentado y reducido como se usan en la presente con respecto a las plantas, son en relación con una planta de control. Las plantas de control adecuadas incluyen versiones no transformadas o de tipo salvaje de una planta de la misma variedad y/o especie que la planta transformada usada en el método del invento. Las plantas de control adecuadas incluyen también plantas de la misma variedad y/o especie que la planta transformada que son transformadas con una construcción de control. Las construcciones de control adecuadas incluyen construcciones de vectores vacíos, conocidas por los expertos en la especialidad. Las plantas de control adecuadas incluyen también plantas que no han sido transformadas con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente. Las plantas de control adecuadas incluyen también plantas que no expresan una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

El término “lípido total” como se usa en la presente incluye grasas, aceites, ceras, esteroles, lípidos de glicerol, monoglicéridos, diglicéridos, fosfolípidos, monogalactolípidos, digalactolípidos, fosfatidilcolinas,

fosfatidiletanolamina, fosfatidilglicerol, fosfatidil-inositol, sulfoguinovosil-diacilglicerol, y triglicéridos.

El término “aceite” como se usa en la presente se refiere preferentemente a triacilglicerol (TAG)

El término “biomasa” se refiere al tamaño y/o masa y/o número de órganos vegetativos de la planta en una edad o etapa de desarrollo particular. Así una planta con incremento de biomasa tiene un mayor tamaño y/o masa y/o número de órganos vegetativos que una planta de control adecuada de la misma edad o en una etapa de desarrollo equivalente. El incremento de biomasa puede comprender también un incremento en la tasa de crecimiento y/o tasa de formación de órganos vegetativos durante algunos o todos los períodos del ciclo de vida de una planta en relación con un control adecuado. Así, el incremento de biomasa puede tener como resultado un adelanto en el tiempo que demora tal planta en alcanzar una cierta etapa de desarrollo.

Los términos “rendimiento de semillas”, “rendimiento de frutos” y “rendimiento de órganos” se refiere al tamaño y/o masa y/o número de semillas, frutos u órganos producidos por una planta. Así una planta con un incremento en el rendimiento de las semillas, frutos u órganos tiene un incremento de tamaño y/o masa y/o número de semillas, frutos u órganos respectivamente, en relación con una planta de control a la misma edad o en una etapa de desarrollo equivalente.

Con los términos “tolerancia acrecentada a la sequía” y “eficiencia acrecentada en el uso del agua” o equivalentes gramaticales de estos, se tiene la intención de describir una planta que se comporta más favorablemente en cualquier aspecto del crecimiento o desarrollo en o después de condiciones subóptimas de hidratación, que plantas de control en las misma condiciones.

Con el término “tolerancia acrecentada a las altas temperaturas” o equivalentes gramaticales de éste, se tiene la intención de describir una planta que se comporta más favorablemente en cualquier aspecto del crecimiento o desarrollo en o después de condiciones de temperatura elevada subóptimas, que plantas de control en las misma condiciones.

El término “tolerancia acrecentada a una alta concentración de oxígeno” o equivalentes gramaticales de éste, está destinado a describir una planta que se comporta más favorablemente en cualquier aspecto del crecimiento y desarrollo en o después de condiciones de una concentración elevada de oxígeno subóptima, que plantas de control en las misma condiciones.

El término “eficiencia acrecentada en el uso de nitrógeno” o equivalentes gramaticales de éste, está destinado a describir una planta que se comporta más favorablemente en cualquier aspecto del crecimiento y desarrollo

en o después de condiciones de una reducción del nitrógeno subóptima, que una planta de control en las mismas condiciones.

El término “tasa acrecentada de asimilación de CO_2” o equivalentes gramaticales de éste, está destinado a describir una planta que asimila más CO_2 en cualesquiera condiciones dadas que una planta de control en las mismas condiciones.

El término “tasa acrecentada de fotosíntesis” o equivalentes gramaticales de éste, está destinado a describir una planta que acumula más fotosintasa en cualesquiera condiciones dadas que una planta de control en las misma condiciones.

El término “tasa acrecentada decrecimiento” o equivalentes gramaticales de éste, está destinado a describir una planta que crece más rápidamente en cualesquiera condiciones dadas que una planta de control en las mismas condiciones.

El término “floración retardada” o equivalentes gramaticales de éste, está destinado a describir una planta que florece más tarde en cualesquiera condiciones dadas que una planta de control en las misma condiciones.

El término “concentración acrecentada de CO_2 en el cloroplasto” o equivalentes gramaticales de éste, está destinado a describir una planta que tiene una concentración de CO_2 más alta en el cloroplasto en cualesquiera condiciones dadas que una planta de control en las mismas condiciones.

El término “tasa reducida de fotorrespiración” o equivalentes gramaticales de éste, está destinado a describir una planta que muestra menos fotorrespiración en cualesquiera condiciones dadas que una planta de control en las mismas condiciones.

El término “pérdida reducida del carbono fijado” o equivalentes gramaticales de éste, está destinado a describir una planta que pierde menos carbono fijado en cualesquiera condiciones dadas que una planta de control en las mismas condiciones.

Polinucleótidos y fragmentos

El término “polinucleótido(s)”, tal como se usa en la presente, significa un polímero de desoxirribonucleótido o ribonucleótido monocatenario o bicatenario de cualquier longitud, pero preferentemente de por lo menos 15 nucleótidos, e incluye como ejemplos no limitativos, secuencias codificadoras y no codificadoras de un gen, complementos de secuencias con sentido y antisentido, exones, intrones, ADN genómico, cADN, pre–mARN, mARN, rARN, siARN, miARN, tARN, ribozimas, polipéptidos recombinantes, secuencias de ADN o ARN

naturales aisladas y purificadas, secuencias de ARN y ADN sintéticas, sondas de ácido nucleico, cebadores y fragmentos.

Un “fragmento” de una secuencia de polinucleótidos proporcionada en la presente es una subsecuencia de nucleótidos contiguos que es capaz de hibridación específica a un objetivo de interés, por ejemplo, una secuencia que tiene por lo menos 15 nucleótidos de longitud. Los fragmentos del invento comprenden 15 nucleótidos, preferentemente, por lo menos 16 nucleótidos, más preferentemente por lo menos 17 nucleótidos, más preferentemente por lo menos 18 nucleótidos, más preferentemente por lo menos 19 nucleótidos, más preferentemente por lo menos 20 nucleótidos, más preferentemente por lo menos 21 nucleótidos, más preferentemente por lo menos 22 nucleótidos, más preferentemente por lo menos 23 nucleótidos, más preferentemente por lo menos 24 nucleótidos, más preferentemente por lo menos 25 nucleótidos, más preferentemente por lo menos 26 nucleótidos, más preferentemente por lo menos 27 nucleótidos, más preferentemente por lo menos 28 nucleótidos, más preferentemente por lo menos 29 nucleótidos, más preferentemente por lo menos 30 nucleótidos, más preferentemente por lo menos 31 nucleótidos, más preferentemente por lo menos 32 nucleótidos, más preferentemente por lo menos 33 nucleótidos, más preferentemente por lo menos 34 nucleótidos, más preferentemente por lo menos 35 nucleótidos, más preferentemente por lo menos 36 nucleótidos, más preferentemente por lo menos 37 nucleótidos, más preferentemente por lo menos 38 nucleótidos, más preferentemente por lo menos 39 nucleótidos, más preferentemente por lo menos 40 nucleótidos, más preferentemente por lo menos 41 nucleótidos, más preferentemente por lo menos 42 nucleótidos, más preferentemente por lo menos 43 nucleótidos, más preferentemente por lo menos 44 nucleótidos, más preferentemente por lo menos 45 nucleótidos, más preferentemente por lo menos 46 nucleótidos, más preferentemente por lo menos 47 nucleótidos, más preferentemente por lo menos 48 nucleótidos, más preferentemente por lo menos 49 nucleótidos, más preferentemente por lo menos 50 nucleótidos, más preferentemente por lo menos 51 nucleótidos, más preferentemente por lo menos 52 nucleótidos, más preferentemente por lo menos 53 nucleótidos, más preferentemente por lo menos 54 nucleótidos, más preferentemente por lo menos 55 nucleótidos, más preferentemente por lo menos 56 nucleótidos, más preferentemente por lo menos 57 nucleótidos, más preferentemente por lo menos 58 nucleótidos, más preferentemente por lo menos 59 nucleótidos, más preferentemente por lo menos 60 nucleótidos, más preferentemente por lo menos 61 nucleótidos, más

preferentemente por lo menos 62 nucleótidos, más preferentemente por lo menos 63 nucleótidos, más preferentemente por lo menos 64 nucleótidos, más preferentemente por lo menos 65 nucleótidos, más preferentemente por lo menos 66 nucleótidos, más preferentemente por lo menos 67 nucleótidos, más preferentemente por lo menos 68 nucleótidos, más preferentemente por lo menos 69 nucleótidos, más preferentemente por lo menos 70 nucleótidos, más preferentemente por lo menos 71 nucleótidos, más preferentemente por lo menos 72 nucleótidos, más preferentemente por lo menos 73 nucleótidos, más preferentemente por lo menos 74 nucleótidos, más preferentemente por lo menos 75 nucleótidos, más preferentemente por lo menos 76 nucleótidos, más preferentemente por lo menos 77 nucleótidos, más preferentemente por lo menos 78 nucleótidos, más preferentemente por lo menos 79 nucleótidos, más preferentemente por lo menos 80 nucleótidos, más preferentemente por lo menos 81 nucleótidos, más preferentemente por lo menos 82 nucleótidos, más preferentemente por lo menos 83 nucleótidos, más preferentemente por lo menos 84 nucleótidos, más preferentemente por lo menos 85 nucleótidos, más preferentemente por lo menos 86 nucleótidos, más preferentemente por lo menos 87 nucleótidos, más preferentemente por lo menos 88 nucleótidos, más preferentemente por lo menos 89 nucleótidos, más preferentemente por lo menos 90 nucleótidos, más preferentemente por lo menos 91 nucleótidos, más preferentemente por lo menos 92 nucleótidos, más preferentemente por lo menos 93 nucleótidos, más preferentemente por lo menos 94 nucleótidos, más preferentemente por lo menos 95 nucleótidos, más preferentemente por lo menos 96 nucleótidos, más preferentemente por lo menos 97 nucleótidos, más preferentemente por lo menos 98 nucleótidos, más preferentemente por lo menos 99 nucleótidos, más preferentemente por lo menos 100 nucleótidos, más preferentemente por lo menos 150 nucleótidos, más preferentemente por lo menos 200 nucleótidos, más preferentemente por lo menos 250 nucleótidos, más preferentemente por lo menos 300 nucleótidos, más preferentemente por lo menos 350 nucleótidos, más preferentemente por lo menos 400 nucleótidos, más preferentemente por lo menos 450 nucleótidos y en la forma más preferida por lo menos 500 nucleótidos de nucleótidos contiguos de un polinucleótido divulgado. Se puede usar un fragmento de una secuencia de polinucleótidos en antisentido, interferencia de ARN (ARNi), silenciamiento de genes, tecnología de triple hélice o de ribozimas, o como un cebador, una sonda, incluido en un microarray, o usado en métodos de selección a base de polinucleótidos del invento.

El término “cebador” se refiere a un polinucleótido corto, que usualmente tiene un grupo 3’OH libre, que se hibrida a un molde y se usa para activar la polimerización de un polinucleótido complementario al objetivo.

El término “sonda” se refiere a un polinucleótido corto que se usa para detectar una secuencia de polinucleótidos que es complementaria con la sonda, en un ensayo basado en hibridación. La sonda puede consistir en un “fragmento” de un polinucleótido como se define en la presente.

Polipéptidos y fragmentos

El término “polipéptido”, como se usa en la presente, abarca cadenas de aminoácidos de cualquier longitud pero preferentemente por lo menos de 5 aminoácidos, que incluyen proteínas de longitud completa, en las cuales los residuos de aminoácido son ligados por enlaces peptídicos covalentes. Los polipéptidos del presente invento, o usados en los métodos del invento, pueden ser productos naturales purificados o pueden ser producidos en forma parcial o total usando técnicas recombinantes o de síntesis. El término puede referirse a un polipéptido, un agregado de un polipéptido tal como un dímero u otro multímero, un polipéptido de fusión, un fragmento de polipéptido, una variante de polipéptido, o derivado de este.

Un “fragmento” de un polipéptido es una subsecuencia del polipéptido que realiza una función que es necesaria para la actividad biológica y/o proporciona la estructura tridimensional del polipéptido. El término puede referirse a un polipéptido, un agregado de polipéptidos tal como un dímero u otro multímero, un polipéptido de fusión, un fragmento de polipéptido, una variante de polipéptido o derivados de éste capaz de realizar la actividad enzimática anterior.

El término “aislado” aplicado a las secuencias de polinucleótidos o polipéptidos divulgadas en la presente se usa para referirse a secuencias que son extraídas de su entorno celular natural. Una molécula aislada puede obtenerse por cualquier método o combinación de métodos que incluyen técnicas bioquímicas, recombinantes y sintéticas.

El término “recombinante” se refiere a una secuencia de polinucleótidos que es extraída de las secuencias que lo rodean en su contexto natural y/o es recombinada con secuencias que no están presentes en su contexto natural.

Una secuencia de polipéptidos “recombinante” es producida por la traducción de una secuencia de polinucleótidos “recombinante”.

El término “derivado de” con respecto a los polinucleótidos o polipéptidos del invento que son derivados de un género o especie particular, significa que el polinucleótido o polipéptido tiene la misma secuencia que un

polinucleótido o polipéptido natural en ese género o especie. Por lo tanto, el polinucleótido o polipéptido, derivado de un género o especie particular, puede ser producido en forma sintética o recombinante.

Variantes

Como se usa en la presente, el término “variante” se refiere a secuencias de polinucleótidos o polipéptidos diferentes de las secuencias identificadas específicamente, en donde uno o más residuos de nucleótidos o aminoácidos es suprimido, substituido o añadido. Las variantes pueden ser variantes alélicas naturales, o variantes no naturales. Las variantes pueden ser de la misma especie o de especies diferentes y pueden abarcar homólogos, parálogos y ortólogos. En ciertas realizaciones, las variantes de los polipéptidos y los polipéptidos del invento poseen actividades biológicas que son iguales o similares a aquellas de los polipéptidos o polipéptidos del invento. El término “variante” con referencia a los polipéptidos y polipéptidos abarca todas las formas de polipéptidos y polipéptidos que se define en la presente.

Variantes de polinucleótidos

Las secuencias de variantes de polinucleótidos exhiben preferentemente por lo menos 50%, más preferentemente por lo menos 51%, más preferentemente por lo menos 52%, más preferentemente por lo menos 53%, más preferentemente por lo menos 54%, más preferentemente por lo menos 55%, más preferentemente por lo menos 56%, más preferentemente por lo menos 57%, más preferentemente por lo menos 58%, más preferentemente por lo menos 59%, más preferentemente por lo menos 60%, más preferentemente por lo menos 61%, más preferentemente por lo menos 62%, más preferentemente por lo menos 63%, más preferentemente por lo menos 64%, más preferentemente por lo menos 65%, más preferentemente por lo menos 66%, más preferentemente por lo menos 67%, más preferentemente por lo menos 68%, más preferentemente por lo menos 69%, más preferentemente por lo menos 70%, más preferentemente por lo menos 71%, más preferentemente por lo menos 72%, más preferentemente por lo menos 73%, más preferentemente por lo menos 74%, más preferentemente por lo menos 75%, más preferentemente por lo menos 76%, más preferentemente por lo menos 77%, más preferentemente por lo menos 78%, más preferentemente por lo menos 79%, más preferentemente por lo menos 80%, más preferentemente por lo menos 81%, más preferentemente por lo menos 82%, más preferentemente por lo menos 83%, más preferentemente por lo menos 84%, más preferentemente por lo menos 85%, más preferentemente por lo menos 86%, más preferentemente por lo menos 87%, más preferentemente por lo menos 88%, más preferentemente por lo menos 89%, más preferentemente por lo menos 90%, más preferentemente por lo menos 91%, más

preferentemente por lo menos 92%, más preferentemente por lo menos 93%, más preferentemente por lo menos 94%, más preferentemente por lo menos 95%, más preferentemente por lo menos 96%, más preferentemente por lo menos 97%, más preferentemente por lo menos 98%, y en la forma más preferida, por lo menos 99% de identidad con una secuencia del presente invento. La identidad es encontrada a lo largo de una ventana comparativa de por lo menos 20 posiciones de nucleótidos, preferentemente por lo menos 50 posiciones de nucleótidos, más preferentemente por lo menos 100 posiciones de nucleótidos, y en la forma más preferida, a lo largo de toda la longitud de un polinucleótido del invento.

La identidad de la secuencia de polinucleótidos puede determinarse de la siguiente manera. La secuencia de polinucleótidos del presente es comparada con una secuencia de polinucleótidos candidata usando BLASTN (del conjunto de programas BLAST, versión 2.2.5 [Nov 2002]) en bl2seq (Tatiana A. Tatusova, Thomas L. Madden (1999), "Blast 2 sequences – a new tool for comparing protein and nucleotide sequences", FEMS Microbiol Lett. 174:247–250), que está disponible públicamente en NCBI (ftp://ftp.ncbi.nih.gov/blast/). Se utiliza los parámetros predeterminados de bl2seq excepto porque se debe desactivar la filtración de las partes de baja complejidad.

La identidad de las secuencias de polinucleótidos puede ser examinada usando los siguientes parámetros de línea de comando unix:

bl2seq –i nucleotideseq1 –j nucleotideseq2 –F F –p blastn

El parámetro –F F desactiva la filtración de las secciones de baja complejidad. El parámetro –p selecciona el algoritmo apropiado para el par de secuencias. El programa bl2wseq informa la identidad de secuencia como el número y el porcentaje de nucleótidos idénticos en una línea “Identidades = “.

La identidad de la secuencia de polinucleótidos se puede calcular también a lo largo de toda la longitud de la superposición entre una secuencia de polinucleótidos candidata y una sujeto usando los programas de alineamiento global de secuencias (por ejemplo, Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443–453). Una implementación completa del algoritmo de alineamiento global de Needleman–Wunsch se encuentra en el programa needle del paquete EMBOSS (Rice,P. Longden,I. y Bleasby,A. EMBOSS: The European Molecular Biology Open Software Suite, Trends in Genetics junio de 2000, vol 16, No 6. pp.276–277) que se puede obtener en http://www.hgmp.mrc.ac.uk/Software/EMBOSS/. El servidor European Bioinformatics Institute también proporciona el dispositivo para realizar alineamientos globales EMBOSS–needle entre las dos secuencias en línea en http:/www.ebi.ac.uk/emboss/align/.

Alternativamente se puede usar el programa GAP que calcula un alineamiento global óptimo de dos secuencias sin penalizar las brechas terminales. GAP se describe en el siguiente documento: Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227–235.

Un método preferido para calcular el % de identidad de secuencia del polinucleótido se basa en el alineamiento de las secuencias que se va a comparar usando Clustal X (Jeanmougin y otros, 1998, Trends Biochem. Sci. 23, 403–5).

Las variantes de polinucleótidos del presente invento también abarcan aquellas que exhiben una semejanza a una o más de las secuencias identificadas específicamente que probablemente conserva(n) la equivalencia funcional de aquellas secuencias y que no se podría esperar en forma razonable que haya(n) ocurrido en forma casual. Tal semejanza de secuencia con respecto a los polipéptidos puede determinarse usando el programa bl2seq disponible al público en el conjunto de programas BLAST (versión 2.2.5 [Nov 2002]) de NCBI (ftp://ftp.ncbi.nih.gov/blast/).

La semejanza de las secuencias de polinucleótidos puede ser examinada usando los siguientes parámetros de línea del comando unix:

bl2seq –i nucleotideseq1 –j nucleotideseq2 –F F –p tblastx

El parámetro –F F desactiva la filtración de las secciones de baja complejidad. El parámetro –p selecciona el algoritmo apropiado para el par de secuencias. Este programa encuentra regiones de semejanza entre las secuencias y para cada una de tales regiones informa un “valor E” que es el número esperado de veces que se podría esperar ver tal concordancia casual en una base de datos de un tamaño de referencia fijo que contiene secuencias aleatorias. El tamaño de esta base de datos es ajustado por defecto en el programa bl2seq. Para valores de E pequeños, mucho menores de uno, el valor E es aproximadamente la probabilidad de tal concordancia aleatoria.

Las secuencias de los polinucleótidos variantes exhiben preferentemente un valor E menor de 1 x 10 –6, más preferentemente menor de 1 x 10 –9, más preferentemente menor de 1 x 10 –12, más preferentemente menor de 1 x 10 –15, más preferentemente menor de 1 x 10 –18, más preferentemente menor de 1 x 10 –21, más preferentemente menor de 1 x 10 –30, más preferentemente menor de 1 x 10 –40, más preferentemente menor de 1 x 10 –50, más preferentemente menor de 1 x 10 –60, más preferentemente menor de 1 x 10 –70, más preferentemente menor de 1 x 10 –80, más preferentemente menor de 1

x 10 –90 y en la forma más preferida menor de 1 x 10–100 cuando son comparadas con cualquiera de las secuencias identificadas específicamente.

Alternativamente, las variantes de polinucleótidos del presente invento, o usadas en los métodos del invento, se hibridan a las secuencias de polinucleótidos especificadas, o sus complementos en condiciones rigurosas.

El término "se hibrida en condiciones rigurosas", y sus equivalentes gramaticales, se refiere a la capacidad de una molécula de polinucleótido de hibridarse a una molécula de polinucleótido de objetivo (tal como una molécula de polinucleótido de objetivo inmovilizada en una transferencia de ADN o ARN, tal como una transferencia Southern o transferencia Northern) en condiciones definidas de temperatura y concentración de sal. La capacidad de hibridarse en condiciones de hibridación rigurosas se puede determinar efectuando la hibridación inicial en condiciones menos rigurosas y aumentando después la rigurosidad hasta la rigurosidad deseada.

Con respecto a las moléculas de polinucleótido mayores de aproximadamente 100 bases de longitud, las condiciones típicas de hibridación rigurosa son a no más de 25 a 30°C (por ejemplo, 10°C) bajo la temperatura de fusión (Tm) del duplex nativo (ver generalmente, Sambrook *y otros,* Eds, 1987, Molecular Cloning, A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press; Ausubel *y otros*, 1987, Current Protocols in Molecular Biology, Greene Publishing). La Tm para las moléculas de polinucleótido mayores de aproximadamente 100 bases puede ser calculada por la fórmula Tm = 81,5 + 0,41% (G + C–log (Na+). (Sambrook *y otros,* Eds, 1987, Molecular Cloning, A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press; Bolton y McCarthy, 1962, PNAS 84:1390). Las condiciones rigurosas típicas para un polinucleótido de más de 100 bases de longitud serían condiciones de hibridación tales como prelavado en una solución de 6X SSC, 0,2% de SDS; hibridación a 65°C, 6X SSC, 0,2% de SDS durante toda la noche; seguido por dos lavados de 30 minutos cada uno en 1X SSC, 0,1% de SDS a 65°C y dos lavados de 30 minutos cada uno en 0,2X SSC, 0,1% de SDS a 65°C.

Con respecto a las moléculas de polinucleótido que tienen una longitud menor de 100 bases, el ejemplo de condiciones de hibridación rigurosas son 5 a 10°C bajo la Tm. En promedio, la Tm de una molécula de polinucleótido de menos de 100 pb se reduce aproximadamente en (500/ longitud de oligonucleótidos)^{o}C.

Con respecto a los mímicos de ADN conocidos como ácidos nucleicos peptídicos (PNAs) (Nielsen *y otros*, Science. 1991 Dic 6;254(5037):1497–500) los valores de Tm son más altos que aquellos de los

híbridos de ADN–ADN o ADN–ARN, y pueden calcularse usando la fórmula descrita en Giesen *y otros*, Nucleic Acids Res. 1998 Nov 1;26(21):5004–6. Como ejemplo, las condiciones rigurosas de hibridación para un híbrido de ADN–PNA que tiene una longitud menor de 100 bases son 5 a 10°C bajo la Tm.

Los polinucleótidos variantes del presente invento, o usados en los métodos del invento, abarcan también los polinucleótidos que difieren de las secuencias del invento pero que como una consecuencia de la degeneración del código genético, codifican un polipéptido que tiene actividad similar a un polipéptido codificado por un polinucleótido del presente invento. Una alteración de la secuencia que no cambia la secuencia de aminoácidos del polipéptido es una “variación silente”. Excepto para ATG (metionina) y TGG (triptófano), se puede cambiar otros codones para el mismo aminoácido por técnicas reconocidas en la especialidad, por ejemplo, para optimizar la expresión del codón en un organismo huésped particular.

Se incluye también en el invento las alteraciones en la secuencia de polinucleótidos que tienen como resultado substituciones conservadoras de uno o varios aminoácidos en la secuencia de polipéptidos codificada sin alterar significativamente su actividad biológica. Un profesional experto estará al tanto de los métodos para obtener substituciones de aminoácidos fenotípicamente silentes (ver, por ejemplo, Bowie *y otros*, 1990, Science 247, 1306).

Las variantes de polinucleótidos debidas a variaciones silentes y substituciones conservadoras en la secuencia de polipéptidos codificada pueden determinarse usando el programa bl2seq disponible al público del conjunto de programas BLAST (versión 2.2.5 [Nov 2002]) de NCBI (ftp://ftp.ncbi.nih.gov/blast/) por medio del algoritmo tblastx descrito previamente.

Variantes de polipéptidos

El término “variante” con referencia a polipéptidos comprende polipéptidos naturales, producidos en forma recombinante y sintética. Las secuencias de polipéptidos variantes exhiben preferentemente por lo menos 50%, más preferentemente por lo menos 51%, más preferentemente por lo menos 52%, más preferentemente por lo menos 53%, más preferentemente por lo menos el 54%, más preferentemente por lo menos 55%, más preferentemente por lo menos 56%, más preferentemente por lo menos 57%, más preferentemente por lo menos 58%, más preferentemente por lo menos 59%, más preferentemente por lo menos 60%, más preferentemente por lo menos 61%, más preferentemente por lo menos 62%, más preferentemente por lo menos 63%, más preferentemente por lo menos 64%, más preferentemente por lo menos 65%, más preferentemente por lo menos

66%, más preferentemente por lo menos 67%, más preferentemente por lo menos 68%, más preferentemente por lo menos 69%, más preferentemente por lo menos 70%, más preferentemente por lo menos 71%, más preferentemente por lo menos 72%, más preferentemente por lo menos 73%, más preferentemente por lo menos 74%, más preferentemente por lo menos 75%, más preferentemente por lo menos 76%, más preferentemente por lo menos 77%, más preferentemente por lo menos 78%, más preferentemente por lo menos 79%, más preferentemente por lo menos 80%, más preferentemente por lo menos 81%, más preferentemente por lo menos 82%, más preferentemente por lo menos 83%, más preferentemente por lo menos 84%, más preferentemente por lo menos 85%, más preferentemente por lo menos 86%, más preferentemente por lo menos 87%, más preferentemente por lo menos 88%, más preferentemente por lo menos 89%, más preferentemente por lo menos 90%, más preferentemente por lo menos 91%, más preferentemente por lo menos 92%, más preferentemente por lo menos 93%, más preferentemente por lo menos 94%, más preferentemente por lo menos 95%, más preferentemente por lo menos 96%, más preferentemente por lo menos 97%, más preferentemente por lo menos 98%, y en la forma más preferida por lo menos 99% de identidad con una secuencia del presente invento. La identidad se encuentra a lo largo de una ventana comparativa de por lo menos 20 posiciones de aminoácidos, preferentemente por lo menos 50 posiciones de aminoácidos, más preferentemente por lo menos 100 posiciones de aminoácidos, y en la forma más preferida, a lo largo de toda la longitud de un polipéptido del invento.

La identidad de secuencia de los polipéptidos puede determinarse de la siguiente manera. La secuencia de los presentes polipéptidos se compara con una secuencia de polipéptidos candidata usando BLASTP (del conjunto de programas BLAST, versión 2.2.5 [Nov 2002]) en bl2seq, que está disponible al público en NCBI (ftp://ftp.ncbi.nih.gov/blast/). Se utiliza los parámetros predeterminados de bl2seq salvo porque se debe desactivar la filtración de las regiones de baja complejidad.

La identidad de secuencia de los polipéptidos puede calcularse también a lo largo de la longitud completa de la superposición entre una secuencia de polinucleótidos candidata y la presente usando los programas de alineamiento global. EMBOSS–needle (disponible en http:/www.ebi.ac.uk/emboss/align/) y GAP (Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227–235) como se describió antes son también programas de alineamiento global de secuencias adecuados para calcular la identidad de secuencia de los polipéptidos.

Un método preferido para calcular el % de identidad de secuencia de los polipéptidos se basa en alinear las secuencias que deben ser comparadas usando Clustal X (Jeanmougin y otros, 1998, Trends Biochem. Sci. 23, 403–5.)

Las variantes de los polipéptidos del presente invento, o usadas en los métodos del invento, abarcan también aquellas que exhiben una semejanza a una o más de las secuencias identificadas específicamente que es probable que conserven la equivalencia funcional de esas secuencias y que no se podría esperar en forma razonable que hayan ocurrido en forma casual. Tal semejanza de las secuencias con respecto a los polipéptidos puede determinarse usando el programa bl2seq disponible al público en el conjunto de programas BLAST (versión 2.2.5 [Nov 2002]) de NCBI (ftp://ftp.ncbi. nih.gov/blast/). La semejanza de las secuencias de los polipéptidos puede examinarse usando los siguientes parámetros de línea de comando unix:

bl2seq –i peptideseq1 –j peptideseq2 –F F –p blastp

Las secuencias de polipéptidos variantes exhiben preferentemente un valor de E menor de 1 x 10–6, más preferentemente menor de 1 x 10–9, más preferentemente menor de 1 x 10–12, más preferentemente menor de 1 x 10 –15, más preferentemente menor de 1 x 10–18, más preferentemente menor de 1 x 10–21, más preferentemente menor de 1 x 10 –30, más preferentemente menor de 1 x 10 –40, más preferentemente menor de 1 x 10–50, más preferentemente menor de 1 x 10–60, más preferentemente menor de 1 x 10–70, más preferentemente menor de 1 x 10–80, más preferentemente menor de 1 x 10–90 y en la forma más preferida de 1x10–100 cuando se comparan con cualquiera de las secuencias identificadas específicamente.

El parámetro –F F desactiva la filtración de las secciones de baja complejidad. El parámetro –p selecciona el algoritmo apropiado para el par de secuencias. Este programa encuentra regiones de semejanza entre las secuencias y para cada una de tales regiones informa un “valor E” que es el número esperado de veces que se podría esperar ver tal concordancia por casualidad en una base de datos de un tamaño de referencia fijo que contiene secuencias aleatorias. Para valores de E pequeños, mucho menores de 1, esta es aproximadamente la probabilidad de tal concordancia aleatoria.

También están incluidas en el invento las substituciones conservadoras de uno o varios aminoácidos de una secuencia de polipéptidos descrita sin alterar significativamente su actividad biológica. Un profesional experto conocerá los métodos para preparar substituciones de aminoácidos

fenotípicamente silentes (ver, por ejemplo, Bowie y otros, 1990, Science 247, 1306).

Construcciones, vectores y componentes de estos

La expresión "construcción genética" se refiere a una molécula de polinucleótido, habitualmente ADN bicatenario, que puede haber sido insertada en otra molécula de polinucleótido (la molécula de polinucleótido insertada) tal como una molécula de ADNc pero sin limitarse a esto. Una construcción genética puede contener los elementos necesarios que permiten transcribir la molécula de polinucleótido insertada, y, opcionalmente, traducir el transcrito en un polipéptido. La molécula de polinucleótido insertada puede ser derivada de la célula huésped, o puede ser derivada de una célula u organismo diferente y/o puede ser un polinucleótido recombinante. Una vez dentro de la célula huésped la construcción genética puede integrarse en el ADN cromosómico del huésped. La construcción genética puede ser ligada a un vector.

El término “vector” se refiere a una molécula de polinucleótido, usualmente ADN bicatenario, que se usa para transportar la construcción genética hacia adentro de una célula huésped. El vector puede ser capaz de replicación por lo menos en un sistema huésped adicional, tal como *E. coli*.

El término "construcción de expresión" se refiere a una construcción genética que incluye los elementos necesarios que permiten transcribir la molécula de polinucleótido insertada, y, opcionalmente, traducir el transcrito en un polipéptido. Una construcción de expresión comprende normalmente en una dirección 5’ a 3’:

a) un promotor funcional en la célula huésped dentro de la cual se transformará la construcción,

b) el polinucleótido que va a ser expresado, y

c) un terminador funcional en la célula huésped dentro de la cual se transformará la construcción.

El término “región codificadora” o “marco de lectura abierto” (ORF) se refiere a la cadena de sentido de una secuencia de ADN genómico o una secuencia de ADNc que es capaz de producir un producto de transcripción y/o un polipéptido bajo el control de secuencias reguladoras apropiadas. En algunos casos, la secuencia codificadora puede ser identificada por la presencia de un codón de comienzo de traducción 5’ y un codón de terminación de traducción 3’. Cuando es insertada en una construcción genética, una “secuencia codificadora” es capaz de ser expresada cuando está ligada operativamente a secuencias promotoras y terminadoras.

“Ligada operativamente” significa que la secuencia que va a ser expresada es colocada bajo el control de elementos reguladores que incluyen promotores, elementos reguladores específicos del tejido, elementos reguladores temporales, potenciadores, represores y terminadores.

El término “región no codificadora” se refiere a secuencias no traducidas que están corriente arriba del sitio de inicio de la traducción y corriente abajo del sitio de terminación de la traducción. Estas secuencias son designadas también, respectivamente, la 5’ UTR y la 3’ UTR. Estas regiones incluyen elementos requeridos para iniciar y terminar la transcripción, para la estabilidad del ARNm, y para regular la eficiencia de la traducción.

Los terminadores son secuencias que terminan la transcripción, y se encuentran en los extremos no traducidos 3’ de los genes corriente abajo de la secuencia traducida. Los terminadores son determinantes importantes de la estabilidad del ARNm y se ha encontrado en algunos casos que tienen funciones reguladoras espaciales.

El término “promotor” se refiere a elementos reguladores cis no transcritos corriente arriba de la región codificadora que regula la transcripción del gen. Los promotores comprenden elementos iniciadores cis que especifican el sitio de inicio de la transcripción, y cajas conservadas tales como la caja TATA, y motivos que son ligados por factores de transcripción. Los intrones dentro de las secuencias codificadoras pueden regular también la transcripción e influir en el procesamiento postranscripcional (lo que incluye empalme, terminación y poliadenilación).

Un promotor puede ser homólogo con respecto al polinucleótido que va a ser expresado. Esto significa que el promotor y el polinucleótido se encuentran ligados operativamente en la naturaleza.

Alternativamente el promotor puede ser heterólogo con respecto al polinucleótido que va a ser expresado. Esto significa que el promotor y el polinucleótido no se encuentran ligados operativamente en la naturaleza.

Un “transgén” es un polinucleótido que es tomado de un organismo e introducido en un organismo diferente por transformación. El transgén puede ser derivado de la misma especie o de una especie diferente a la especie del organismo en el cual el transgén es introducido.

Una “repetición invertida” es una secuencia que se repite, donde la segunda mitad de la repetición está en la cadena complementaria, por ejemplo,

(5’)GATCTA.....TAGATC(3’)

(3’)CTAGAT.....ATCTAG(5’)

La transcripción ininterrumpida producirá un transcrito que experimenta el apareamiento de bases complementario para formar una estructura de horquilla con la condición que exista un espaciador de 3–5 pb entre las regiones repetidas.

Células huésped

Las células huésped pueden ser derivadas, por ejemplo, de organismos bacterianos, hongos, levaduras, insectos, mamíferos, algas o vegetales. Las células huésped también pueden ser células sintéticas. Las células huésped preferidas son células eucariotas. Una célula huésped particularmente preferida es una célula vegetal.

Una “planta transgénica” se refiere a una planta que contiene material genético nuevo como resultado de la manipulación o transformación genética. El material genético nuevo puede ser derivado de una planta de la misma especie que la planta transgénica resultante o de una especie diferente.

Métodos para aislar o producir polinucleótidos

Las moléculas de polinucleótidos del invento pueden ser aisladas usando una variedad de técnicas conocidas por personas que tienen una destreza normal en la especialidad. A modo de ejemplo, tales polipéptidos pueden ser aislados mediante el uso de la reacción en cadena de la polimerasa (PCR) descrita en Mullis y otros, Eds. 1994 The Polymerase Chain Reaction, Birkhauser, incorporada en la presente por referencia. Los polipéptidos del invento pueden ser amplificados usando cebadores, como se define en la presente, derivados de las secuencias de polinucleótidos del invento.

Otros métodos para aislar los polinucleótidos del invento incluyen el uso de todos los polipéptidos, o porciones de estos, que tienen la secuencia expuesta en la presente como sondas de hibridación. Se puede usar la técnica para hibridar sondas de polinucleótidos marcados a polinucleótidos inmovilizados en soportes sólidos tales como filtros de nitrocelulosa o membranas de nylon, para escrutar las bibliotecas genómicas o de ADNc. Los ejemplos de condiciones de hibridación y lavado son: hibridación durante 20 horas a 65°C en 5,0 X SSC, 0,5% de dodecilsulfato de sodio, 1 X de solución de Denhardt; lavado (tres lavados de veinte minutos cada uno a 55°C) en 1,0 X SSC, 1% (p/v) de dodecilsulfato de sodio, y opcionalmente un lavado (durante veinte minutos) en 0,5 X SSC, 1% (p/v) de dodecilsulfato de sodio, a 60°C. Se puede realizar otro lavado opcional (durante veinte minutos) en condiciones de 0,1 X SSC, 1% (p/v) de dodecilsulfato de sodio, a 60°C.

Los fragmentos de polinucleótidos del invento pueden ser producidos por técnicas bien conocidas en la especialidad tales como digestión

por endonucleasa de restricción, síntesis de oligonucleótidos y amplificación por PCR.

Se puede usar una secuencia parcial de polinucleótidos, en métodos bien conocidos en la especialidad para identificar la secuencia de polinucleótidos de longitud completa correspondiente. Tales métodos incluyen métodos basados en PCR, 5'RACE (Frohman MA, 1993, Methods Enzymol. 218: 340–56) y el método basado en hibridación, métodos basados en computador/base de datos. Además, a modo de ejemplo, la PCR inversa permite la adquisición de secuencias desconocidas, que flanquean las secuencias de polinucleótidos divulgadas en la presente, comenzando con los cebadores sobre la base de una región conocida (Triglia y otros, 1998, Nucleic Acids Res 16, 8186, incorporada en la presente por referencia). El método usa varias enzimas de restricción para generar un fragmento adecuado en la región conocida de un gen. El fragmento es circularizado después por ligamiento intramolecular y usado como un molde de PCR. Los cebadores divergentes son diseñados a partir de la región conocida. Para ensamblar físicamente clones de longitud completa, se puede utilizar propuestas de biología molecular estándar (Sambrook y otros, Molecular Cloning: A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press, 1987).

Cuando se produce una planta transgénica a partir de una especie particular puede ser útil transformar tal planta con una secuencia o secuencias derivadas de esa especie. El beneficio puede ser aliviar preocupaciones públicas relacionadas con la transformación de especies cruzadas para generar organismos transgénicos. Además, cuando el resultado deseado es la regulación negativa de un gen, puede ser necesario utilizar una secuencia idéntica (o al menos muy similar) a aquella de la planta cuya expresión se desea reducir. Por estas razones entre otras, es deseable poder identificar y aislar ortólogos de un gen particular en varias especies de plantas diferentes.

Las variantes (que incluyen ortólogos) pueden ser identificadas por los métodos descritos.

Métodos para identificar variantes

Métodos físicos

Los polipéptidos variantes pueden ser identificados usando métodos basados en PCR (Mullis y otros, Eds. 1994 The

Polymerase Chain Reaction, Birkhauser). Típicamente, la secuencia de polinucleótidos de un cebador, útil para amplificar variantes de moléculas de polinucleótidos del invento por PCR, puede basarse en una secuencia que codifica una región conservada de la correspondiente secuencia de aminoácidos.

Alternativamente, se puede emplear métodos para escrutar bibliotecas, bien conocidos por los expertos en la especialidad (Sambrook y otros, Molecular Cloning: A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press, 1987). Cuando se identifica las variantes de la secuencia de la sonda, típicamente se reducirá normalmente la rigurosidad de hibridación y/o lavado con respecto a cuando se busca concordancias de secuencia exactas.

Las variantes de polipéptidos pueden ser identificadas también por métodos físicos, por ejemplo escrutando bibliotecas de expresión con el uso de anticuerpos cultivados contra los polipéptidos del invento (Sambrook y otros, Molecular Cloning: A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press, 1987) o identificando polipéptidos de fuentes naturales con la ayuda de tales anticuerpos

Métodos basados en computador

Las secuencias variantes del invento, que incluyen variantes de polinucleótidos y de polipéptidos, pueden ser identificadas también por métodos basados en computador bien conocidos por los expertos en la especialidad, usando algoritmos de alineamiento de secuencia de dominio público y herramientas de búsqueda de semejanza de secuencias para buscar en bases de datos de secuencias (las bases de datos de dominio público incluyen Genbank, EMBL, Swiss–Prot, PIR y otras). Ver, por ejemplos, en Nucleic Acids Res. 29: 1–10 y 11–16, 2001, ejemplos de recursos en línea. Las búsquedas de semejanza recuperan y alinean secuencias de objetivo para compararlas con una secuencia que va a ser analizada (es decir, una secuencia de consulta). Los algoritmos de comparación de secuencias usan matrices de puntuación para asignar una puntuación general a cada uno de los alineamientos.

Un ejemplo de familia de programas útiles para identificar variantes en las base de datos de secuencia es el conjunto de programas BLAST (versión 2.2.5 [Nov 2002]) que incluye BLASTN, BLASTP, BLASTX, tBLASTN y tBLASTX, los que están disponibles para el público en (ftp://ftp.ncbi.nih.gov/blast/) o en el National Center for Biotechnology Information (NCBI), National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894, EE.UU. El servidor NCBI proporciona también el dispositivo para usar los programas para identificar numerosas bases de datos de secuencia disponibles al público. BLASTN compara una secuencia de nucleótidos de consulta contra una base de datos de secuencias de nucleótidos. BLASTP compara una secuencia de aminoácidos de consulta

contra una base de datos de secuencias de proteína. BLASTX compara una secuencia de nucleótidos de consulta traducida en todos los marcos de lectura contra una base de datos de secuencias de proteína. tBLASTN compara una secuencia de proteína de consulta contra una base de datos de secuencias de nucleótidos traducida dinámicamente en todos los marcos de lectura. tBLASTX compara las traducciones de seis marcos de una secuencia de nucleótidos de consulta contra las traducciones de seis marcos de una secuencia de nucleótidos de la base de datos. Los programas BLAST pueden ser usados con parámetros predeterminados o los parámetros pueden ser alterados según se requiera para refinar el escrutinio.

El uso de la familia BLAST de algoritmos, que incluye BLASTN, BLASTP, y BLASTX, se describe en la publicación de Altschul y otros, Nucleic Acids Res. 25: 3389–3402, 1997.

Los “aciertos” en una o más secuencias de la base de datos por una secuencia de consulta producidos por BLASTN, BLASTP, BLASTX, tBLASTN, tBLASTX, o un algoritmo similar, alinean e identifican porciones similares de las secuencias. Los aciertos se disponen en el orden del grado de semejanza y la longitud de la superposición de las secuencias. Los aciertos para una secuencia de la base de datos representan generalmente una superposición sólo sobre una fracción de la longitud de secuencia de la secuencia de consulta.

Los algoritmos BLASTN, BLASTP, BLASTX, tBLASTN y tBLASTX también producen valores “Esperados” para los alineamientos. El valor Esperado (E) indica el número de aciertos que es "esperable" observar por causalidad cuando se busca en una base de datos del mismo tamaño que contiene secuencias contiguas aleatorias. El valor Esperado se usa como un umbral de significancia para determinar si el acierto a una base de datos indica una semejanza verdadera. Por ejemplo, un valor E de 0,1 asignado a un acierto de un polinucleótido se interpreta con el significado que en una base de datos del tamaño de la base de datos escrutada, se podría esperar ver 0,1 concordancias con respecto a la porción alineada de la secuencia con una puntuación similar simplemente por azar. Para las secuencias que tienen un valor E de 0,01 o menos con respecto a las porciones alineadas y apareadas, la probabilidad de encontrar una concordancia por azar en esa base de datos es de 1% o menos usando el algoritmo BLASTN, BLASTP, BLASTX, tBLASTN o tBLASTX.

Los alineamientos de múltiples secuencias de un grupo de secuencias relacionadas pueden llevarse a cabo con CLUSTALW (Thompson, J.D., Higgins, D.G. y Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions–

specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673–4680, http://www–igbmc.u–strasbg.fr/BioInfo/ ClustalW/ Top.html) o T–COFFEE (Cedric Notredame, Desmond G. Higgins, Jaap Heringa, T–Coffee: A novel method for fast and accurate multiple sequence alignment, J. Mol. Biol. (2000) 302: 205–217)) o PILEUP, que usa alineamientos progresivos, por pares (Feng y Doolittle, 1987, J. Mol. Evol. 25, 351).

Se dispone de las aplicaciones del programa de reconocimiento de patrón para encontrar motivos o secuencias distintivas. Por ejemplo, MEME (Multiple Em for Motif Elicitation) encuentra motivos y secuencias distintivas en un conjunto de secuencias, y MAST (Motif Alignment and Search Tool) usa estos motivos para identificar motivos similares o iguales en las secuencias de consulta. Los resultados de MAST se proporcionan como una serie de alineamientos con los datos estadísticos apropiados y un panorama visual de los motivos encontrados. MEME y MAST fueron desarrollados en la Universidad de California, San Diego.

PROSITE (Bairoch y Bucher, 1994, Nucleic Acids Res. 22, 3583; Hofmann *y otros*, 1999, Nucleic Acids Res. 27, 215) es un método para identificar las funciones de proteínas no caracterizadas traducidas de secuencias genómicas o de ADNc. La base de datos PROSITE (www.expasy.org/prosite) contiene patrones y perfiles biológicamente significativos y está diseñada para que pueda ser utilizada con herramientas de computación apropiadas para asignar una nueva secuencia a una familia de proteínas conocida o para determinar cual(es) dominio(s) conocido(s) está(n) presente(s) en la secuencia (Falquet *y otros*, 2002, Nucleic Acids Res. 30, 235). Prosearch es una herramienta que puede buscar en las bases de datos SWISS–PROT y EMBL con un patrón de secuencia o firma dado.

Métodos para aislar polipéptidos

Los polipéptidos del invento, o usados en los métodos del invento, que incluyen los polipéptidos variantes, pueden prepararse usando métodos de síntesis de péptidos bien conocidos en la especialidad tales como síntesis directa de péptidos usando técnicas de fase sólida (por ejemplo Stewart y otros, 1969, en Solid–Phase Peptide Synthesis, WH Freeman Co, San Francisco California, o síntesis automatizada, por ejemplo usando un sintetizador de péptidos de Applied Biosystems 431A (Foster City, California). Durante tales síntesis se puede producir también formas mutadas de los polipéptidos.

Los polipéptidos y los polipéptidos variantes del invento, o usados en los métodos del invento, pueden ser purificados también de fuentes naturales usando una variedad de técnicas que son bien conocidas en la materia (por

ejemplo Deutscher, 1990, Ed, Methods in Enzymology, Vol. 182, Guide to Protein Purification).

Alternativamente, los polipéptidos y polipéptidos variantes del invento, o usados en los métodos del invento, pueden expresarse en forma recombinante en células huésped adecuadas y separados de las células como se describe a continuación.

Métodos para producir construcciones y vectores

Las construcciones genéticas del presente invento comprenden una o más secuencias de los polinucleótidos del invento y/o polinucleótidos que codifican polipéptidos del invento, y pueden ser útiles para transformar, por ejemplo, organismos de bacterias, hongos, insectos, mamíferos o plantas. Se tiene la intención que las construcciones genéticas del invento incluyan las construcciones de expresión definidas en la presente.

Los métodos para producir y usar las construcciones genéticas y vectores son bien conocidos en la especialidad y se describen generalmente en Sambrook *y otros*, Molecular Cloning: A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press, 1987; Ausubel *y otros*, Current Protocols in Molecular Biology, Greene Publishing, 1987.

Métodos para producir células huésped que comprenden polinucleótidos, construcciones o vectores

El invento proporciona una célula huésped que comprende una construcción genética o vector del invento.

Las células huésped que comprenden construcciones genéticas del invento, tales como construcciones de expresión son útiles en métodos bien conocidos en la especialidad (por ejemplo en Sambrook *y otros*, Molecular Cloning: A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press, 1987; Ausubel *y otros*, Current Protocols in Molecular Biology, Greene Publishing, 1987) para la producción recombinante de los polipéptidos del invento. Tales métodos pueden comprender el cultivo de las células huésped en un medio apropiado en condiciones adecuadas o que conduzcan a la expresión de un polipéptido del invento. El polipéptido recombinante expresado, que opcionalmente puede ser secretado en el cultivo, puede ser separado después del medio, células huésped o medio de cultivo por métodos bien conocidos en la especialidad (por ejemplo Deutscher, Ed, 1990, Methods in Enzymology, Vol 182, Guide to Protein Purification).

Métodos para producir células vegetales y plantas que comprenden construcciones y vectores

El invento proporciona además células vegetales que comprenden una construcción genética del invento, y células vegetales modificadas para alterar la expresión de un polinucleótido o polipéptido del invento, o usado en los métodos del invento. Las plantas que comprenden tales células forman también un aspecto del invento.

Los métodos para transformar células vegetales, plantas y porciones de éstas con polipéptidos han sido descritas en Draper *y otros*, 1988, Plant Genetic Transformation and Gene Expression. A Laboratory Manual. Blackwell Sci. Pub. Oxford, p. 365; Potrykus y Spangenburg, 1995, Gene Transfer to Plants. Springer–Verlag, Berlín; y Gelvin *y otros*, 1993, Plant Molecular Biol. Manual. Kluwer Acad. Pub. Dordrecht. Una revisión de las plantas transgénicas que incluyen las técnicas de transformación, es proporcionada en Galun and Breiman, 1997, Transgenic Plants. Imperial College Press, Londres.

Métodos para la manipulación genética de las plantas

Se dispone de numerosas estrategias de transformación de plantas (por ejemplo en Birch, 1997, Ann Rev Plant Phys Plant Mol Biol, 48, 297, Hellens RP, y otros, (2000) Plant Mol Biol 42: 819–32, Hellens R y otros, Plant Meth 1: 13). Por ejemplo, se puede diseñar estrategias para incrementar la expresión de un polinucleótido/polipéptido en una célula vegetal, órgano y/o en una etapa de desarrollo particular de la planta que/cuando no es expresado normalmente o es expresado ectópicamente en una célula, tejido, órgano y/o en un estado de desarrollo particular de la planta en el que/cuando no se expresa normalmente. El polinucleótido/polipéptido expresado puede ser derivado de la especie de planta que va a ser transformada o puede ser derivado de una especie de planta diferente.

Se puede diseñar las estrategias de transformación para reducir la expresión de un polinucleótido/polipéptido en una célula vegetal, tejido, órgano y/o en un estado de desarrollo particular que/cuando se expresa normalmente. Tales estrategias se conocen como estrategias de silenciamiento génico.

Las construcciones genéticas para la expresión de los genes en las plantas transgénicas incluyen normalmente promotores para impulsar la expresión de uno o más polinucleótidos clonados, terminadores y secuencias de marcador seleccionables para detectar la presencia de la construcción genética en la planta transformada.

Los promotores adecuados para usar en las construcciones de este invento son funcionales en una célula, tejido u órgano de una planta monocotiledónea o dicotiledónea e incluyen promotores específicos de células,

tejidos y órganos, promotores específicos del ciclo celular, promotores temporales, promotores inducibles, promotores constitutivos que son activos en la mayor parte de los tejidos vegetales, y promotores recombinantes. La elección del promotor dependerá de la expresión temporal y espacial del polinucleótido clonado, tal como se desea. Los promotores pueden ser aquellos asociados normalmente con un transgén de interés, o promotores que son derivados de genes de otras plantas, virus, y bacterias y hongos patógenos de las plantas. Los expertos en la especialidad podrán seleccionar, sin experimentación excesiva, los promotores que sean adecuados para usar en la modificación y modulación de rasgos de la planta usando construcciones genéticas que comprenden las secuencias de polinucleótidos del invento. Los ejemplos de promotores constitutivos de plantas incluyen el promotor CaMV 35S, el promotor de nopalina sintasa, el promotor de octopina sintasa y el promotor Ubi 1 del maíz. Los promotores de plantas que son activos en tejidos específicos, que responden a señales de desarrollo interno o apremios abióticos o bióticos externos han sido descritos en la bibliografía científica. Se describe ejemplos de promotores, tal como en WO 02/00894, que es incorporada en la presente por referencia.

Los ejemplos de terminadores que son usados comúnmente en la construcción genética para transformación de la planta incluyen, por ejemplo, el terminador del virus del mosaico de coliflor (CaMV) 35S, los terminadores de nopalina sintasa u octopina sintasa de *Agrobacterium tumefaciens*, el terminador del gen de zeína de *Zea mays*, el terminador de ADP–glucosa pirofosforilasa de *Oryza sativa* y el terminador de PI–II de *Solanum tuberosum*.

Los marcadores seleccionables usados comúnmente en la transformación de plantas incluyen el gen de neomicina fosfotransferasa II (NPT II) que confiere resistencia a kanamicina, el gen aadA que confiere resistencia a espectinomicina y estreptomicina, la fosfinotricina acetil transferasa(gen *bar*) para la resistencia a Ignite (AgrEvo) y Basta (Hoechst), y el gen de higromicina fosfotransferasa (hpt) para la resistencia a higromicina.

Está contemplado también el uso de construcciones genéticas que comprenden genes informadores (secuencias codificadoras que expresan una actividad que es extraña para el huésped, usualmente una actividad enzimática y/o una señal visible, (por ejemplo, luciferasa, GUS, GFP) que pueden usarse en análisis de la expresión del promotor en plantas y en tejidos de plantas. La bibliografía sobre genes informadores es examinada por Herrera–Estrella *y otros*, 1993, Nature 303, 209, y Schrott, 1995, en: Gene Transfer to Plants (Potrykus, T., Spangenberg. Eds) Springer Verlag. Berline, pp. 325–336.

Las siguientes son publicaciones representativas que divulgan protocolos de transformación genética que pueden ser usados para transformar genéticamente las siguientes especies de plantas: arroz (Alam y otros, 1999, Plant Cell Rep. 18, 572); manzana (Yao y otros, 1995, Plant Cell Reports 14, 407–412); maíz (patentes de EE.UU. Nos. de Serie 5.177.010 y 5.981.840); trigo (Ortiz y otros, 1996, Plant Cell Rep. 15, 1996, 877); tomate (Patente de EE.UU. No. de Serie 5.159.135); papa (Kumar y otros, 1996 Plant J. 9:821); yuca (Li y otros, 1996 Nat. Biotechnology 14, 736); lechuga (Michelmore y otros, 1987, Plant Cell Rep. 6, 439); tabaco (Horsch y otros, 1985, Science 227, 1229); algodón (patentes de EE.UU. Nos. de Serie 5.846.797 y 5.004.863); pastos (patentes de EE.UU. Nos 5.187.073 y 6.020.539); menta (Niu y otros, 1998, Plant Cell Rep. 17, 165); plantas cítricas (Pena y otros, 1995, Plant Sci.104, 183); alcaravea (Krens y otros, 1997, Plant Cell Rep, 17, 39); banana (patente de EE.UU. No. de Serie 5.792.935); soya (patentes de EE.UU. Nos. 5.416.011; 5.569.834; 5.824.877; 5.563.455 y 5.968.830); piña (patente de EE.UU. No. de Serie 5.952.543); álamo (patente de EE.UU. No. 4.795.855); monocotiledóneas en general (patentes de EE.UU. Nos. 5.591.616 y 6.037.522); brassica (patentes de EE.UU. Nos. 5.188.958; 5.463.174 y 5.750.871); cereales (patente de EE.UU. No. 6.074.877); pera (Matsuda y otros, 2005, Plant Cell Rep. 24(1):45–51); Prunus (Ramesh y otros, 2006 Plant Cell Rep. 25(8):821–8; Song y Sink 2005 Plant Cell Rep. 2006;25(2):117–23; Gonzalez Padilla y otros, 2003 Plant Cell Rep.22(1):38–45); frutilla (Oosumi y otros, 2006 Planta. 223(6):1219–30; Folta y otros, 2006 Planta Apr 14; PMID: 16614818), rosa (Li y otros, 2003), Rubus (Graham y otros, 1995 Methods Mol Biol. 1995;44:129–33), tomate (Dan y otros, 2006, Plant Cell Reports V25:432–441), manzana (Yao y otros, 1995, *Plant Cell Rep.*14, 407–412), Canola (Brassica napus L.) (Cardoza y Stewart, 2006 Methods Mol Biol. 343:257–66), cártamo (Orlikowska y otros, 1995, Plant Cell Tissue and Organ Culture 40:85–91), ballico (Altpeter y otros, 2004 Developments in Plant Breeding 11(7):255–250), arroz (Christou y otros, 1991 Nature Biotech. 9:957–962), maíz (Wang y otros 2009 In: Handbook of Maize pp. 609–639) y *Actinidia eriantha* (Wang y otros, 2006, Plant Cell Rep. 25,5: 425–31). El invento contempla también la transformación de otras especies. Los métodos y protocolos adecuados están disponibles en la literatura científica.

Plantas

El término “planta” está destinado a incluir una planta completa, cualquier parte de una planta, una semilla, un fruto, propágulos y la progenie de una planta.

El término ‘propágulo’ significa cualquier parte de una planta que puede ser usada en la reproducción o propagación, ya sea sexual o asexual, e incluye semillas y estacas.

Las plantas del invento pueden ser cultivadas y autofecundadas o cruzadas con una cepa de planta diferente y se puede identificar los híbridos resultantes con las características fenotípicas deseadas. Se puede cultivar dos o más generaciones para asegurar que las características fenotípicas del sujeto se mantengan y se hereden en forma estable. Las plantas resultantes de tales propuestas de cultivo estándar forman también un aspecto del presente invento.

Abreviaturas

Oleosina (u Ole)_0–0 significa una oleosina sin cisteínas manipuladas genéticamente.

Oleosina (u Ole)_1–1 significa una oleosina con una cisteína manipulada genéticamente en cada brazo hidrófilo.

Oleosina (u Ole)_1–3 significa una oleosina con una cisteína manipulada genéticamente en el brazo hidrófilo N–terminal y tres cisteínas manipuladas genéticamente en el brazo hidrófilo C–terminal.

Oleosina (u Ole)_3–1 significa una oleosina con tres cisteínas manipuladas genéticamente en el brazo hidrófilo N–terminal y una cisteína manipulada genéticamente en el brazo hidrófilo C–terminal.

Oleosina (u Ole)_3–3 significa una oleosina con tres cisteínas manipuladas genéticamente en el brazo hidrófilo N–terminal y tres cisteínas manipuladas genéticamente en el brazo hidrófilo C–terminal.

Oleosina (u Ole)_5–6 significa una oleosina con cinco cisteínas manipuladas genéticamente en el brazo hidrófilo N–terminal y seis cisteínas manipuladas genéticamente en el brazo hidrófilo C–terminal.

Oleosina (u Ole)_6–7 significa una oleosina con seis cisteínas manipuladas genéticamente en el brazo hidrófilo N–terminal y siete cisteínas manipuladas genéticamente en el brazo hidrófilo C–terminal.

BREVE DESCRIPCIÓN DE LAS FIGURAS

La Figura 1 muestra la secuencia de la construcción de Oleosin_0–0 y DGAT1 (S205A). CaMV35 es el promotor 35S del Virus Mosaico de Coliflor. attB1 es el sitio de recombinación GATEWAY™. UBQ10 es el intrón del

gen UBQ10de A. thaliana. El terminador OCS es el terminador de octopina sintasa.

La Figura 2 muestra la disposición de la construcción de Oleosin_1–1 y DGAT1 (S205A), transformada en Arabidopsis thaliana.

La Figura 3 muestra la secuencia de la construcción de Oleosin_1–3 y DGAT1 (S205A). CaMV35 es el promotor 35S del Virus Mosaico de Coliflor. attB1 es el sitio de recombinación GATEWAY™. UBQ10 es el intrón del gen UBQ10 de A. thaliana. El terminador OCS es el terminador de octopina sintasa.

La Figura 4 muestra la construcción de Oleosin_3–1 y DGAT1 (S205A). CaMV35 es el promotor 35S del Virus Mosaico de Coliflor. attB1 es el sitio de recombinación GATEWAY™. UBQ10 es el intrón del gen UBQ10 de A. thaliana. El terminador OCS es el terminador de octopina sintasa.

La Figura 5 muestra la construcción de Oleosin_3–3 y DGAT1 (S205A). CaMV35 es el promotor 35S del Virus Mosaico de Coliflor. attB1 es el sitio de recombinación GATEWAY™. UBQ10 es el intrón del gen UBQ10 de A. thaliana. El terminador OCS es el terminador de octopina sintasa.

La Figura 6 muestra un mapa de la construcción pRSh1 utilizada para transformar plantas. El mapa muestra la disposición de las oleosinas, con cisteínas introducidas artificialmente (en este caso Oleo_3–3) bajo el control del promotor CaMV35s, así como también DGAT1 (S205A) de Arabidopsis thaliana también bajo el control del promotor CaMV35s. Por supuesto, es posible substituir Oleo_3–3 y DGAT1 respectivamente por otras secuencias de oleosinas y secuencias de enzimas sintetizadoras de TAG.

La Figura 7 muestra una comparación de dot blot de anticuerpos anti oleosina de semillas de sésamo ligados a oleosina recombinante de semillas de sésamo purificada, con y sin residuos de cisteína manipulados genéticamente.

La Figura 8 muestra un análisis de inmunoblot para detectar las proteínas oleosina cisteína expresadas en *E. Coli,*en cuerpos oleosos artificiales (COAs). Para cada carril se cargó un volumen igual de COA (7,5 µL que incluía 2x SDS colorante de carga sin agente reductor) La concentración mM de GSSG se indica arriba de cada carril.

La Figura 9 muestra análisis de SDS y de SDS–UREA PAGE/inmunoblot de Ole–0–0, Ole–1–1 y Ole–3–3 expresado en *E. coli*. Las muestras fueron preparadas de cuerpos de inclusión (IB) y de cuerpos oleosos artificiales (COAs) en presencia y ausencia de agentes de reducción (DTT y β–ME) o de agente de oxidación (GSSG), donde se cargó iguales cantidades de

proteína en los carriles adyacentes.

La Figura 10 muestra análisis de inmunoblot de acumulación de oleosina (Oleo_0–0, Oleo_1–3, Oleo_3–1, y Oleo_3–3, SEQ ID NOs 11–20) en las semillas de *Arabidopsis thaliana* transgénica que expresaban tanto DGAT1 (S205A) como una oleosina de sésamo bajo el control de promotores CaMV35S.

La Figura 11 muestra análisis de inmunoblot de acumulación de oleosinas (Oleo_0–0, Oleo_1–3, Oleo_3–1, y Oleo_3–3, SEQ ID NOs 11–20) en los cuerpos oleosos de *Arabidopsis thaliana* transgénica que expresaban DGAT1 (S205A) y una oleosina de sésamo bajo el control de promotores CaMV35S. La aparición de las bandas de oleosina oligomérica (dimérica y trimérica) en presencia de agente de oxidación (+) indica que los enlaces disulfuro tienen la capacidad de formarse en la parte externa de los cuerpos oleosos nativos.

La Figura 12 muestra análisis de inmunoblot de acumulación de oleosinas (Oleo_0–0, Oleo_1–3, Oleo_3–1, y Oleo_3–3, SEQ ID NOs 11–20) en las hojas de *Arabidopsis thaliana* transgénica que expresan DGAT1 (S205A) y una oleosina de sésamo bajo el control de promotores CaMV35S.

La Figura 13 muestra inmunoblot de acumulación de oleosina recombinante (flecha negra) en hojas de Arabidopsis transgénica.

La Figura 14 muestra resultados de FAMES GC/MS que muestran la acumulación de lípidos adicionales (flechas negras) en hojas de Arabidopsis que sobreexpresan DGAT1 (S205A) y Ole_3,3.

La Figura 15 muestra resultados de GC/MS para el perfil lipídico total de hojas de líneas de tipo salvaje e independientes de Arabidopsis transgénica que contienen DGAT1 (S205A) y Ole_3,3. La flecha gris indica el estándar interno. Las flechas negras indican lípidos neutros adicionales (ésteres de cera, ésteres de esterol y TAGs). Las flechas abiertas muestran tres líneas (41S, 18A y 47C) que acumulan cantidades sustanciales de lípidos neutros en sus hojas en comparación con el tipo salvaje (y la línea 50A).

La Figura 16 presenta resultados de GC/MS que muestran el perfil total de TAG de Arabidopsis de tipo salvaje y transgénico (que contienen DGAT1 (S205A) y Ole_3,3) a las 2, 3, 4 y 5 semanas después de la germinación. Las flechas negras indican TAGs adicionales encontrados en hojas transgénicas pero no en las de tipo salvaje.

La Figura 17 presenta resultados de FAMES GC/MS que muestran los perfiles lipídicos totales en hojas de *Trifolium repens* de tipo salvaje y transgénico (que contienen DGAT1 (S205A) y Ole_3,3).

La Figura 18 presenta resultados de FAMES GC/MS que muestran los perfiles lipídicos de C18:1 y C18:2 en hojas de *Trifolium repens* de tipo salvaje y transgénico (que contienen DGAT1 (S205A) y Ole_3,3).

La Figura 19 muestra una presentación esquemática del orden de eventos en la fotosíntesis, que incluye la Reacción de Hill (Reacción Luminosa) y la fijación de carbono (Ciclo de Calvin).

La Figura 20 muestra una presentación esquemática del Ciclo de Calvin (C3). Los segmentos de color gris claro, de gris algo más oscuro y del gris más oscuro muestran las reacciones de carboxilación, reducción y regeneración, respectivamente. Para 3 moléculas de CO_2 fijado se dispone de una molécula de gliceraldehido 3-fosfato (GAP) para biosíntesis y energía.

La ecuación general para la fotosíntesis por algas y plantas (donde el donante de electrones es agua) es:

$$2n\ CO_2 + 2n\ H_2O + \text{fotones} \rightarrow 2(CH_2O)n + 2n\ O_2$$

La Figura 21 muestra una presentación esquemática de la reacción de oxigenasa de Rubisco.

La Figura 22 muestra la vía fotorrespiratoria en la planta superior. El ciclo de Calvin se muestra de color gris sombreado y muestra el punto de retorno del glicolato reciclado (ahora en la forma de 3-fosfoglicerato).

La Figura 23 muestra un bypass fotorrespiratorio según Kebeish y otros, (2007). El área sombreada muestra el efecto de obviar los pasos que involucran normalmente al peroxisoma y también a las mitocondrias, llevando a una elevación de la concentración de CO_2 en el cloroplasto, así como también a un reciclado más eficiente de glicolato.

La Figura 24 muestra una comparación de patrones de crecimiento de Arabidopsis transgénica. A) tipo salvaje; B) planta transformada sólo con GDH; C) planta transformada con GDH, GCL y TSR. (Kebeish *y otros,* 2007).

La Figura 25muestra una presentación esquemática de la biosíntesis de triacilglicéridos en organismos fotosintéticos. Los ácidos grasos son sintetizados en el plástido transportado al retículo endoplásmico y acilados secuencialmente sobre un esqueleto de glicerol por medio de la vía Kennedy; esto culmina en la producción de triacilglicérido por medio de sobreexpresión de la enzima DGAT.

La Figura 26muestra una presentación esquemática de la influencia de la biosíntesis lipídica continua en la hoja transgénica. Los ácidos grasos son sintetizados en el plástido transportado al retículo endoplásmico, acilados secuencialmente sobre un esqueleto de glicerol por medio de la vía Kennedy; esto

culmina en la producción de triacilglicérido por medio de sobreexpresión de la enzima DGAT. En este caso el ácido 3-fosfoglicérico es sintetizado por Rubisco (sin el ciclo de Calvin) en vez de la transformación de azúcares. La transformación siguiente de éste a acetil-CoA (a través del piruvato intermediario) tiene resultado la liberación de CO_2 en el cloroplasto. Esto incrementa la presión parcial de CO_2 en relación con O_2 en el cloroplasto, reduciendo así la proporción de los ciclos C2 a C3 iniciados por Rubisco e incrementando la tasa de asimilación de CO_2.

La Figura 27muestra una presentación esquemática del catabolismo de TAG no protegido producido en la hoja transgénica. La sobreexpresión de DGAT lleva a la acumulación de TAG que es degradado posteriormente por las lipasas, teniendo como resultado la liberación de ácidos grasos libres. Algunos de estos ácidos grasos libres son catabolizados por β-oxidación en el peroxisoma, en tanto que otros originan un ciclo inútil reingresando en el retículo endoplásmico donde ellos se vuelven a incorporar en el TAG. Este ciclo inútil redujo la demanda de síntesis *de-novo* de nuevos lípidos; posteriormente el nivel de reciclado de CO_2 dentro del cloroplasto se reduce a los niveles del tipo salvaje o se acerca a estos, lo que lleva a la reanudación de la razón de tipo salvaje de los ciclos C2 a C3 que está siendo efectuada por Rubisco dentro de la célula fotosintética C3.

La Figura 28muestra una presentación esquemática de la influencia de evitar el catabolismo de TAG en la fotorrespiración en la hoja transgénica. La sobreexpresión de DGAT lleva a la acumulación de TAG, el cual es encapsulado posteriormente por oleosina coexpresada que contiene residuos de cisteína manipulados genéticamente. Esto impide la degradación de TAG por las lipasas y así evita también el reciclaje inútil de lípidos (ver cruces). Por consiguiente, hay una demanda continua de síntesis lipídica *de novo* y una presión parcial elevada de CO_2 en la célula fotosintética, lo que tiene como resultado una supresión continua de los ciclos C2 en relación con los ciclos C3.

La Figura 29muestra una comparación de los patrones de crecimiento de Arabidopsis transgénica. A, C, E) tipo salvaje; B, D, F) planta transformada con DGAT1 (S205A) y Ole_3,3. A y B) 20 días después de la germinación; C y D) 30 días después de la germinación; E y F)72 días después de la germinación.

El panel izquierdo de la Figura 30muestra la tasa de fijación de CO_2 en el aire para el tipo salvaje (WT) y plantas transformadas con DGAT1-Ole_3,3 (T) y con bajo O_2 para el tipo salvaje (WTO2) y plantas transformadas con DGAT1-Ole_3,3 (TO2). El panel derecho muestra el % de cambio de la tasa de fijación

deCO_2 para el tipo salvaje (WT) y las plantas transformadas con DGAT1-Ole_3,3 (T) cuando son colocadas en un ambiente con bajo O_2.

El panel izquierdo de la Figura 31 muestra la Eficiencia Intrínseca en el Uso del Agua (iWUE) en el aire para el tipo salvaje (WT) y plantas transformadas con DGAT1-Ole_3,3 (T) y con bajo O_2 para el tipo salvaje (WTO2) y para plantas transformadas con DGAT1-Ole_3,3 (TO2). El panel derecho muestra el % de cambio en iWUE para el tipo salvaje (WT) y plantas transformadas con DGAT1-Ole_3,3 (T) cuando son colocadas en un ambiente con bajo O_2.

El panel izquierdo de la Figura 32 muestra la Conductancia Estomática en aire para el tipo salvaje (WT) y para plantas transformadas con DGAT1-Ole_3,3 (T) y en bajo O_2 para el tipo salvaje (WTO2) y para plantas transformadas con DGAT1-Ole_3,3 (TO2). El panel derecho muestra el % de cambio en Conductancia Estomática para el tipo salvaje (WT) y para plantas transformadas con DGAT1-Ole_3,3 (T) cuando son colocadas en un ambiente con bajo O_2.

La Figura 33 muestra la Densidad Estomática para la superficie adaxial del tipo salvaje (WT AD), la superficie adaxial de plantas transformadas con DGAT1-Ole_3,3 (T AD), la superficie abaxial del tipo salvaje (WT AB), la superficie abaxial de plantas transformadas con DGAT1-Ole_3,3 (T AB).

La Figura 34 muestra las diferencias en el tamaño de las plantas entre plantas de control del tipo salvaje y plantas transformadas con DGAT1-Ole_3,3 (DAG = Días Después de la Germinación).

El panel izquierdo de la Figura 35 muestra la cantidad total (como % de peso seco) para cada especie principal de lípido en raíces de plantas del tipo salvaje (barras negras) y en raíces de plantas transformadas con DGAT1-Ole_3,3 (barras grises). El panel derecho muestra cada especie principal de lípido como un % del total de lípidos en las raíces de plantas del tipo salvaje (barras negras) y en las raíces de plantas transformadas con DGAT1-Ole_3,3 (barras grises).

La Figura 36muestra cuatro trazos desplazados y superpuestos.

El trazo A muestra TAG extraído de 100 mg de raíces de plantas transformadas con DGAT1-Ole_3,3.

El trazo B muestra TAG extraído de 100 mg de hojas de plantas transformadas con DGAT1-Ole_3,3.

El trazo C muestra TAG extraído de 100 mg de raíces de plantas del tipo salvaje.

El trazo D muestra TAG extraído de 100 mg de hojas de plantas del tipo salvaje.

La Figura 37 muestra A) immunoblot de hojas de tipo salvaje, D1o0-0 y tres líneas independientes D1o-3-3 sondeadas con anticuerpos anti-oleosina 28 días después de sembrar (DAS); B) tasas de asimilación de CO_2 en hojas de las mismas plantas en A. Las barras representan valores medios ± d.s. de experimentos separados (*n* = 8), importancia estadística entre WT y líneas D1o3-3 calculadas por prueba t de Student de 2 colas, para D1o3-3#18,#41 y #47 *P* = 9,15E-09, 0,01698 y 8,7E-08, respectivamente; C) fotos representativas de las mismas plantas analizadas en A y B.

EJEMPLOS

Este invento será ilustrado ahora refiriéndose a los siguientes ejemplos no limitativos.

EJEMPLO 1:Creación de anticuerpos de conejo anti–oleosina de semillas de sésamo

Generación de anticuerpos de conejo anti–oleosina de semillas de sésamo

Una oleosina de longitud completa de semillas de sésamo que contenía un His tag C–terminal (la secuencia de nucleótidos se muestra en la SEQ ID NO: 1) fue expresada en *E. coli*, y se preparó cuerpos de inclusión por técnicas estándar. Los cuerpos de inclusión fueron solubilizados en Tampón Ligante (tampón fosfato 100 mM pH 8,0, NaCl 500 mM, urea 8 M e imidazol 10 mM) y cargados en una columna de cromatografía por afinidad de iones metálicos (IMAC) (Invitrogen) que contenía Ni agarosa en equilibrio. Las proteínas no ligadas fueron retiradas de la columna mediante lavado con seis volúmenes de Tampón de Lavado (tampón fosfato 100 mM pH 8,0, NaCl 500 mM, urea 6 M e imidazol 50 mM). La proteína fue eluída en alícuotas de 1 vol. de Tampón de Elución (tampón fosfato 100 mM pH 8,0, NaCl 500 mM, urea 6M e imidazol 250 mM). Las fracciones eluídas fueron analizadas mediante SDS–PAGE/teñido de Coomassie, y se midió la concentración de proteína usando el Ensayo de Bradford. Se mezcló 265 µg de la proteína oleosina recombinante purificada por IMAC con una cantidad igual de Adyuvante de Freund Completo hasta un volumen final de 0,5 mL. Después de recolectar el presangrado, se administró la primera inyección en múltiples sitios a través de la parte posterior del cuello y el área del hombro de un conejo. Se entregó inyecciones de refuerzo que contenían 77 µg de la oleosina purificada a las tres y a las siete semanas después de la inyección primaria, y se extrajo una sangría de testeo de ~3 mL para el análisis preliminar a las nueve semanas. Se conservó el suero mediante la adición de fenol 0,25% v/v y de merthiolate 0,01% v/v, y se almacenó en alícuotas de 200 µL a –20 °C.

La sensibilidad de los anticuerpos de conejo anti–oleosina de semillas de sésamo fue evaluada mediante immunodotting, que indicó que se podía detectar regularmente 0,25 ng de oleosina de semillas de sésamo con una dilución 1/ 2000 del anticuerpo (Figura 7).

EJEMPLO 2: Diseño y expresión en *E. coli* de oleosinas modificadas que contienen uno o más residuos de cisteína introducidos artificialmente

Diseño de construcción para expresión en E. coli

Se diseñó una cantidad de construcciones de oleosinas modificadas para ser expresadas en *E. Coli.* Estascontenían uno o tres residuos de cisteína en los brazos hidrófilos N–terminal y C–terminal. Las construcciones se basaron en la secuencia de nucleótidos y en una secuencia de polipéptidos traducida de una oleosina de semillas de sésamo, clon AF091840 de GenBank que no contiene ningún residuo de cisteína (SEQ ID NO: 16).

Todos los clones fueron subclonados en pET29b usando los sitios NdeI/XhoI modificados genéticamente. Además, se añadió una secuencia codificadora ProTrp a la región codificadora del extremo 3' del brazo hidrófilo C–terminal para imitar los residuos de aminoácidos codificados por el sitio NcoI modificado antes genéticamente por Peng y otros (2006) Stability enhancement of native and artificial oil bodies by genipin crosslink. Patente de Taiwan I 250466.

Las proteínas oleosina–cisteína mutadas para que incluyeran residuos de cisteína en las dos regiones hidrófilas N– y C–terminales aquí descritas, son designadas Ole–1–1, Ole–1–3, Ole–3–1, y Ole–3–3 (SEQ ID NOS 2, 3, 4, y 5 respectivamente), correspondiendo los dígitos numerales primero y segundo a la cantidad de enlaces disulfuro en el terminal N– y C–, respectivamente. Como control se utilizó la oleosina estándar sin los residuos de cisteína, y fue designada como Ole–0–0 (SEQ ID NO 1).

Las cisteínas fueron substituidas por residuos cargados de los cuales se predijo que estarían sobre la superficie de los cuerpos oleosos y se indican a continuación.

Cisteína simple N–terminal	(Ole–1–x)	Glu3Cys	
Cisteína triple N–terminal	(Ole–3–x)	Glu3Cys	Arg12Cys Gln23Cys
Cisteína simple C–terminal	(Ole–x–1)	Gln137Cys	
Cisteína triple C–terminal	(Ole–x–3)	Gln112Cys	Lys123Cys Gln137Cys

Las construcciones fueron diseñadas para que pudieran ser subclonadas de una manera relativamente sencilla a partir del esqueleto (pCR4Blunt-TOPO) proporcionado por GENEART en pET29b (Novogen) por medio de digestión de NcoI/XhoI y ligación. Esto colocó la secuencia codificadora de la oleosina corriente abajo de la fusión pET29 S•tag N–terminal y corriente arriba del His tag C–terminal (Figuras 1–5 y SEQ ID NOS 1–10). Las secuencias de oleosina y de oleosina modificada utilizadas han sido resumidas en la Tabla de Resumen de Secuencias.

Expresión en E. coli y purificación de oleosinas modificadas que contienen por lo menos una cisteína introducida artificialmente

La expresión de las oleosinas recombinantes de semillas de sésamo (con y sin cisteínas modificadas genéticamente) en el sistema de expresión de *E. coli* fue evaluada mediante SDS–PAGE/teñido azul brillante de Coomassie y SDS–PAGE/análisis de inmunoblot, utilizando anticuerpos cultivados contra la oleosina de semillas de sésamo (descrita en el Ejemplo 1).

La expresión de oleosina recombinante modificada fue inducida en un cultivo de *E. coli* recién inoculado de 10 mL (BL21 Rosetta–Gami) que contenía una secuencia codificadora de oleosina (con o sin residuos de cisteína modificada genéticamente) en el vector de expresión pET29. El cultivo fue desarrollado a 37°C, 220 rpm, hasta la fase logarítmica media (OD_{600}0,5 – 0,7); se indujo la expresión mediante la adición de IPTG hasta una concentración final de 1 mM. El cultivo inducido fue incubado a 37°C, 220 rpm, durante otras 2–3 horas. Dadas las propiedades de la oleosina modificada, los solicitantes no intentaron expresarla en una forma soluble sino que en vez de esto eligieron extraerla de los cuerpos de inclusión. Se transfirió alícuotas (1 mL) del cultivo a tubos de microfuga de 1,5 mL, y se formó una pelotilla de células mediante centrifugación (2655 × *g* durante 5 minutos a 4 °C).

Las células en la pelotilla fueron resuspendidas en reactivo BugBuster® (Merck) a 5 mL/g de células de la pelotilla húmedas, con la adición de DNasa a 40 µg/mL y fueron mezcladas suavemente sobre un rotador durante 30 minutos seguido por centrifugación a 8.000 g durante 10 minutos a 4°C. La pelotilla de células resultante fue tratada nuevamente con BugBuster® y DNasa como antes. La proteína soluble restante y los restos de células suspendidos fueron separados de los cuerpos de inclusión insolubles por centrifugación a 8.000 g durante 10 minutosa 4°C.

Las oleosinas recombinantes siguieron siendo purificadas de los cuerpos de inclusión utilizando un procedimiento adaptado de D'Andréa *y otros* (2007). Brevemente, se lavó el preparado de cuerpos de inclusión mediante

resuspensión en tampón de carbonato de sodio 200 mM pH 11 (5 mL por gramo de pelotilla de células original) y se volvió a formar la pelotilla por centrifugación a 8000 × g durante diez minutos a 4°C. La pelotilla de cuerpos de inclusión lavada fue suspendida nuevamente en 5 mL de tampón de carbonato de sodio 200 mM por gramo de pelotilla, y se añadió a 9 volúmenes de una mezcla de cloroformo:metanol recién preparada (5:4 v/v) dando una razón final de 5:4:1 (cloroformo:metanol:tampón). La suspensión fue mezclada suavemente durante cinco minutos, con lo cual se formó una mezcla monofásica de aspecto lechoso; ésta fue centrifugada a 10,000 ×g durante diez minutos a 4°C, y el sobrenadante que contenía la oleosina modificada fue separado cuidadosamente de la pelotilla y transferido a un tubo nuevo. Se secó alícuotas del material sobrenadante bajo una corriente de nitrógeno, y la proteína fue solubilizada de nuevo en urea 8 M y cuantificada mediante Qubit™ (Invitrogen).

EJEMPLO 3: Uso de anticuerpos anti–oleosina de semillas de sésamo para ligar oleosina de semillas de sésamo con cisteínas introducidas artificialmente

Se utilizó un dot–blot para comparar la capacidad de los anticuerpos (Abs) anti–oleosina de semillas de sésamo descritos en el Ejemplo 1 para ligarse a oleosina sin cisteínas contra las oleosinas que contienen cisteínas (descritas en el Ejemplo 2). Se aplicó una serie de diluciones de 12 a 0,25 ng de Ole–0–0, Ole–1–3 y Ole–3–1 purificadas sobre una membrana de transferencia Hybond–P PVDF previamente equilibrada. Esta fue incubada con los anticuerpos anti–oleosina de semillas de sésamo a 1:2000 como el Ab primario. Después la transferencia fue incubada con el Ab secundario adecuado y desarrollada por quimioluminiscencia (Figura 7). Los resultados indican que sobre un inmunoblot, los anticuerpos anti–oleosina de semillas de sésamo son más sensibles hasta en un orden de magnitud a la oleosina sin residuos de cisteína que a las oleosinas con residuos de cisteína. Como consecuencia de las diferentes sensibilidades fue necesario cargar diferentes cantidades de proteína recombinante sobre los geles para el análisis por inmunoblot. A pesar de la carga no uniforme de los carriles es aún posible comparar las diferentes oleosinas entre carriles en términos de su distribución relativa entre formas monoméricas y oligoméricas.

EJEMPLO 4: Creación de cuerpos oleosos artificiales con oleosinas modificadas expresadas en *E. coli* que contienen por lo menos una cisteína introducida artificialmente y alteración el grado de reticulación

Preparación de cuerpos oleosos artificiales

Después se preparó cuerpos oleosos artificiales (COAs) secando alícuotas del sobrenadante descrito en el Ejemplo 3, calculadas para contener 150 µg o 1 mg de oleosina recombinante.

El proceso de generación de los COAs implicó combinar PL, TAG, y la oleosina recombinante/oleosina modificada. En la ausencia de fuertes agentes caotrópicos, la fuerza disruptiva requerida para disociar las oleosinas recombinantes individuales de la fracción purificada implicó varios ciclos alternados de sonicación y enfriamiento. Esto se logró solubilizando las muestras de 150 µg y de 1 mg de oleosina/oleosina modificada en 20 µL de cloroformo que contenía 150 µg de PL (Sigma, Cat#P3644) y mezclado con 60 µL de aceite de semillas de sésamo purificado (Tzen y Huang 1992) y 940 µL de tampón AOB (tampón fosfato de sodio 50 mM pH 8,0, NaCl 100 mM). La mezcla completa fue sonicada después tres veces durante 30 segundos (Sonics & Materials Vibra~Cell VC600, 600 W, 20 kHz; sonda de micropunta ahusada de 1/8”, ajuste de potencia #3).

Los solicitantes descubrieron también que el procedimiento de purificación podía ser aumentado en forma progresiva con éxito, y cuando se utilizó una pelotilla de células de 50 gramos como el material de partida fue necesario substituir la corriente de nitrógeno por un evaporador giratorio de vacío para extraer el cloroformo y la mayor parte del metanol. En este punto la mayor parte de la oleosina/oleosina modificada se precipitó separándose del solvente azeotrópico, y fue separada por centrifugación a 12.000 g durante 10 min.

Los cuerpos de inclusión fueron suspendidos en 1 mL de tampón AOB II (fosfato de sodio 50 mM pH 8,0, NaCl 100 mM, β–mercaptoetanol 20 mM, DTT 10 mM y aceite de sésamo al 5% [v/v], y sonicados después 4x. Los COAs fueron concentrados por centrifugación a 12,000 rpm durante 10 minutos, y esto tuvo como resultado la formación de una suspensión de COAs que recubrió la fracción acuosa. La fracción acuosa subyacente fue extraída mediante pipeta, y los COAs restantes fueron lavados (para extraer las proteínas solubles y los agentes reductores) agitando suavemente en 1 mL de tampón AOB III (fosfato de sodio 50 mM, pH 8,0, NaCl 100 mM). Después de lavar se volvió a concentrar los COAs por centrifugación, y se extrajo la fracción acuosa subyacente, después fueron suspendidos nuevamente por la acción de un vórtice en tampón AOB IV

(tampón fosfato de sodio 50 mM, pH 8,0, NaCl 100 mM, GSSG 1 mM) y los COAs fueron almacenados a 4°C para análisis posteriores.

La Ole–0–0 recombinante, y todas las variaciones de las oleosina–cisteínas fueron expresadas con éxito y se situaron en cuerpos de inclusión de *E. coli* (Figura 9). Ole–0–0 estaba predominantemente presente como un monómero (tanto en los cuerpos de inclusión como en los COAs); ésta emigró fraccionadamente más rápido que el marcador de 20kDa de peso molecular (en SDS y SDS–UREA PAGE tanto reductores como no reductores). También había presentes dos bandas inmunorreactivas que emigraron más lentamente, de alrededor de 35 y 36 kDa, que probablemente corresponden a dos formas de oleosina dimérica. Si bien no se predice que Ole–0–0 contenga algunos residuos de cisteína, la intensidad total y la razón de los dos dímeros aparentes fueron influenciadas por la presencia de los agentes de reducción (β–ME @ 5% del tampón de carga de la muestra y DTT 10 mM).

En los cuerpos de inclusión, la forma predominante de Ole–1–1 es monomérica. Una sola forma dimérica pareció estar presente, y ésta no fue influenciada por los agentes reductores ni por la urea. Ole–1–1 de los COAs (generada en la presencia de agente de reducción y después en la presencia de agente de oxidación) mostró un gran incremento en la razón de dímero a monómero, así como también en la formación de oligómeros triméricos, tetraméricos y probablemente pentaméricos (el enfoque electroforético de estos oligómeros mejoró considerablemente en el gel de SDS–UREA). La remoción del GSSG y la reintroducción de agentes de reducción en los COAs tuvo como resultado que sólo se vio la presencia de monómero y dímero en proporciones similares a las vistas en los cuerpos de inclusión. Los COAs generados con Ole–1–1 (en la ausencia tanto de agentes de reducción como de GSSG) mostraron la presencia de porciones casi iguales de monómero y dímero y una pequeña cantidad de trímero, lo que indica que las condiciones en las cuales se forman los COAs tienen algún potencial reductor. La siguiente adición de GSSG tuvo como resultado un incremento en las porciones oligoméricas así como también la aparición de una forma tetramérica.

Aunque el monómero fue la forma predominante de Ole–3–3 en los cuerpos de inclusión, un porcentaje comparativamente elevado estuvo también presente en múltiples formas oligoméricas. La proporción de oligómeros declinó hasta un grado pequeño con la adición de agente reductor y levemente más por la adición de agentes reductores y caotrópicos. Las formas oligoméricas de Ole–3–3 que eran más grandes que un trímero se resolvieron insuficientemente

cuando se extrajo la proteína recombinante de los COA. La creación de grandes formas oligoméricas fue promovida por la adición de GSSG y en la ausencia de agentes de reducción y caotrópicos, una porción de estas formas oligoméricas no logró ingresar en el gel de apilamiento. Combinados, estos resultados indican que en los COAs, Ole–3–3 estaba altamente reticulada y que la posición de las reticulaciones era más variable en comparación con la Ole–3–3 recuperada de los cuerpos de inclusión. Esto sugiere que a pesar de una reticulación preexistente considerable (dentro de los cuerpos de inclusión), en el COA Ole–3–3 tiene acceso a un alto número de acompañantes potenciales para la reticulación. De igual manera para Ole–1–3 y Ole–3–1, el número de especies reticuladas aumentó cuando había más de una cisteína en una o en ambas regiones hidrófilas (Figuras 8 y 9).

Podría anticiparse que en el SDS–PAGE no reductor, los oligómeros que contienen el mismo número de olefinas pero con los enlaces disulfuro en diferentes posiciones, emigrarían en forma diferente unos respecto de otros. De hecho, esto puede observarse en la Figura 8 en la cual los datos indican que la posición de los brazos de la oleosina, uno respecto del otro, está en posiciones diferentes sobre el cuerpo oleoso. Por ejemplo, Ole–1–1 sólo puede formar un enlace disulfuro por brazo y éste tiene que formarse en la misma posición, en tanto que la presencia de tres cisteínas permite la formación de más de un enlace disulfuro pero también permite que los enlaces disulfuro se formen con diferentes grados de superposición de los brazos hidrófilos, además de tener múltiples olefinas ligadas al mismo brazo (Figuras 8 y 9).

La adición de SDS y agentes de reducción (DTT y β–ME) redujo el número de complejos poliméricos (Figura 9). La adición de SDS y de urea tuvo como resultado un patrón similar al de SDS sólo, con la salvedad de que las múltiples formas diméricas resueltas previamente emigraron como una sola forma, y las formas triméricas y tetraméricas parecen ser más abundantes, posiblemente porque también están emigrando como bandas individuales, lo que incrementa la intensidad de manera correspondiente (Figura 9). En contraste, la presencia de SDS, de agente de reducción y urea tuvo como resultado menos formas oligoméricas de Ole–1–1 y Ole–1–3 pero no de Ole–3–1 ni de Ole–3–3 (Figura 9). En el caso de Ole–3–1 y Ole–3–3 parece que la urea no desnaturaliza por completo las oleosinas de disulfuro, y que de hecho puede impedir la reducción completa de los enlaces disulfuro. Podría ser que estos enlaces se formen durante la generación de los cuerpos de inclusión (sería necesario ver preparados de cuerpos de inclusión reducidos y no reducidos). Por otra parte, la presencia de la oleosina dimérica formada en ausencia de residuos de cisteína

manipulados genéticamente (Figuras 8 y 9) indica que alguna oligomerización se debe a otros tipos de atracción, por ejemplo un fuerte enlace hidrófobo que no es destruido por completo por SDS pero que puede ser destruido casi por completo por la combinación de SDS y urea (Figuras 8 y 9).

El efecto de acrecentar la cantidad de sitios de reticulación potenciales en un péptido de oleosina en la integridad de los COAs y en la estabilidad de la emulsión puede ser evaluada como sigue.

Determinación cuantitativa de la integridad de los COAs

La evaluación de la estabilidad e integridad de los COAs utilizando absorbencia (OD_{600}), conteo directo de COAs mediante un hemocitómetro, o evaluación visual de la coalescencia por microscopía, demostró ser sumamente variable e influenciada, entre otras cosas por el grado de agitación del muestreo preliminar, la cantidad de muestra extraída y el tiempo dejado bajo el microscopio. Para evitar esto, los solicitantes idearon un método sencillo para cuantificar la cantidad de TAG liberado desde los COAs hacia el medio circundante durante diversos de tratamientos como un medio para comparar la integridad. Se preparó cantidades esencialmente iguales (basadas en la estimación de TAG por FAMES–GC/MS y en la determinación de proteína según Bradford) de preparados de COAs hasta un volumen total de 200 µL usando tampón AOB (que contenía Proteinasa K [PNK] cuando fuera adecuado a una razón de 1:1 de PNK:proteínas totales en muestras de CO o COA en tubos con inserto de vidrio para GC de 250 µL, y se los tapó con una tapa de material plástico. Después del tratamiento (temperatura elevada o exposición a PNK) se añadió a la muestra 15 µL de aceite de pescado (Vitamax®, Australia), y se mezcló en un dispositivo de vórtice seguido por centrifugación a 5.200 g por un minuto. La adición de aceite de pescado seguida por introducción en un dispositivo de vórtice permitió que cualquier TAG que se hubiera fugado de los COAs se mezclara con el aceite de pescado añadido y flotara mediante una breve centrifugación. Se tomó una muestra de 4 µL de la fase de aceite y se sometió a una esterificación metílica de ácido graso (FAME) y después se analizó mediante GC–MS (números de modelo de Shimadzu, equipado con una columna capilar 50mQC2/BPX70–0,25 GC (SGE) como describen Browse *y otros* (1986). En la ausencia de adición de aceite de pescado la cantidad de TAG que se había fugado desde los COAs fue demasiado pequeña para formar una capa visible de la que se pudiera tomar muestras aún después de la centrifugación; en tal caso el volumen máximo habría sido de 6 µL. Los perfiles lipídicos muy diferentes del

aceite de pescado y el aceite de sésamo nos permitieron distinguir fácilmente el TAG que se había fugado del TAG añadido.

Usando los estándares internos C15:0 y C17:0 los solicitantes pueden calcular las cantidades absolutas de C18:2 (el principal lípido en el aceite de semillas de sésamo) recuperado después del tratamiento.

Determinación de la integridad de los COAs y de la estabilidad de la emulsión a temperatura elevada

Las emulsiones de aceite en agua son menos estables a temperaturas elevadas; por ello es interesante investigar si las oleosinas modificadas con cantidades variadas de cisteínas introducidas influyen en la integridad de los COAs a temperatura elevada. Para lograr esto los solicitantes determinaron la integridad (usando el método descrito antes) de los COs y COAs (que contenían diferentes oleosinas) en un tampón fosfato (tampón fosfato de sodio 50 mM pH 8, NaCl 100 mM) a 95˚C. Los COAs son calentados durante dos horas. La integridad se determina como se determinó anteriormente.

El efecto de razones más altas de oleosina reticulada:TAG en la estabilidad de los COAs en el fluido ruminal puede evaluarse como sigue.

Determinación de la estabilidad de los COAs en fluido ruminal

Uno de los objetivos del disulfuro era proveer algún grado de protección de la biohidrogenación por la microflora ruminal. La evaluación de la estabilidad de los COAs con fluido ruminal puede efectuarse como sigue. Se añade los COAs a un volumen igual (25 µL) de fluido ruminal. Se incuba las muestras a 39°C durante 0, 15, 30, 60, 120 y 240 min; al final de la incubación se añade un volumen igual de tampón de carga (Invitrogen), se mezcla y se calienta a 70°C durante 10 minutos. Se compara 15 µL de cada mezcla de muestra/tampón de carga mediante SDS–PAGE/inmunoblot. La integridad se determina como se describió anteriormente.

Análisis de integridad de los COAs en Proteinasa K

Para investigar la influencia de la oleosina modificada en un entorno altamente degradante controlable y repetible, (usando el método descrito anteriormente) se determina la integridad de los COAs (que contienen diferentes oleosinas modificadas) después de incubar en un tampón fosfato (tampón fosfato-Na 50 mM pH 8, NaCl 100 mM) que contiene 1:1 (g/g proteína) Proteinasa K (Invitrogen) a 37°C durante 4 horas. Aunque la máxima actividad de la Proteinasa K se alcanza a una temperatura inferior a 65 °C, se utiliza la temperatura más baja

a fin de reducir la influencia de la temperatura en la inestabilidad de los COAs. La integridad se determina como se describió anteriormente.

EJEMPLO 5: Diseño y expresión *in planta* de oleosina modificada que contiene una o más cisteínas introducidas artificialmente

Diseño de construcción para la expresión in planta

Los solicitantes sintetizaron secuencias codificadoras individuales para la oleosina de semillas de sésamo (basadas en el clon AF091840 de GenBank) con diferentes cantidades de cisteínas en los brazos N– y C–terminales. La secuencia codificadora estaba flanqueada por un sitio 5' NotI y un sitio 3' NdeI. Se sintetizó un casete aceptor separado que contenía un sitio attL1, un sitio NotI y un sitio NdeI seguido por una secuencia nos de terminación, un promotor CaMV35s orientado hacia adelante, el DGAT1 (S205A) de*Arabidopsis thaliana* (SEQ ID NOs 11–20 y Figuras 1–5**)** y más su propio intrón UBQ10, un sitio attL2. Las oleosinas de semillas de sésamo con diferentes cantidades de cisteínas fueron transferidas individualmente al casete aceptor por intermedio de los sitios NotI y NdeI. Cada uno de estos casetes completados fue transferido después a un vector binario de planta pRSh1, Figura 6 (Winichayakul y otros, 2008) mediante la reacción de recombinación de LR. Esto colocó a la oleosina corriente abajo de un promotor CaMV35S (que ya estaba contenido en pRSh1) y colocó un terminador nos (que ya estaba contenido en pRSh1) corriente abajo del DGAT1 (S205A) de Arabidopsis (Figuras 1–5). Las secuencias de nucleótidos que codifican las oleosinas de semillas de sésamo (con cisteínas) y DGAT1 fueron optimizadas para ser expresadas en *Arabidopsis thaliana*; esto incluyó optimización de la frecuencia de los codones, el contenido de GC, la remoción de sitios de corte y empalme crípticos, la remoción de secuencias de inestabilidad de mRNA, la remoción de sitios potenciales de reconocimiento de poliadenilación, y la adición del codón de parada de tetranucleótido (Brown y otros, 1990; Beelman y Parker, 1995; Rose, 2004; Rose y Beliakoff, 2000; Norris y otros, 1993).

Debe observarse que la secuencia de oleosina utilizada se da sólo a modo de ejemplo. Se podría manipular genéticamente cualesquiera secuencias de oleosina o esteroleosina o caoleosina para que contengan regiones reticulantes. Las secuencias codificadoras de los ORF completos (después del corte y empalme) fueron verificadas comparándolas con la repetición de la secuencia traducida de la oleosina original, y se encontró que era idéntica con respecto a las regiones codificadoras de oleosina.

Transformación de Arabidopsis thaliana con oleosinas de semillas de sésamo que contienen cisteínas

La transformación de *Arabidopsis thaliana* var Columbia (con construcciones descritas anteriormente), los análisis de semillas T2 para determinar la oleosina modificada, el análisis de inmunoblot de cuerpos oleosos de *Arabidopsis thaliana* que contenían oleosina de semillas de sésamo con diferentes cantidades de cisteínas fueron realizadas como se describió anteriormente (Scott y otros, 2007).

Se utilizó los métodos de inmersión floral (Clough, 1998) y de caída floral (Martínez–Trujillo, 2004) en la transformación de Arabidopsis por *Agrobacterium tumefaciens* GV3101 que contenía las construcciones binarias. Se recolectó semillas de T1 de las plantas tratadas, fueron hechas germinar y seleccionadas rociando a los 14 días y 21 días después de la germinación con Basta®. Las plantas T1 resistentes a Basta® (71, 62 y 23 transformantes que contenían la oleosina individual de las semillas de sésamo, y construcciones de oleosina modificada, respectivamente) fueron trasplantadas, dejadas autofertilizarse, formar semillas y se recolectó las semillas T2. Cantidades iguales de extracto de semillas de plantas de Arabidopsis resistente a Basta® fueron analizadas mediante SDS–PAGE/inmunoblot con los anticuerpos anti–oleosina de semillas de sésamo; en la mayoría de las muestras (Figura 10) se observó oleosina recombinante de semillas de sésamo y oleosina modificada del tamaño adecuado. Se llevó a cabo un análisis de transferencia Southern en líneas T2 seleccionadas para determinar la cantidad de sitios de inserción.

EJEMPLO 6: Extracción y purificación de cuerpos oleosos con oleosinas modificadas que contienen por lo menos una cisteína introducida artificialmente de las semillas de *Arabidopsis thaliana*

Preparados de cuerpos oleosos crudos de semillas de Arabidopsis thaliana

Se elaboró preparados de COs crudos de semillas de plantas producidas como se describe en el Ejemplo 5, moliendo 200 mg de semillas con un mortero y pisón que contenía una punta de espátula de arena y 750 µL de Tampón de Extracción (tampón fosfato 10 mM, pH 7,5 que contenía sucrosa 600 mM) o mediante homogeneización de 25 mg de semillas en 300 µL de tampón de extracción, utilizando un homogeneizador Wiggenhauser D–130. Se añadió otros 750 µL de Tampón de Extracción, y la suspensión presente en el mortero fue transferida a un tubo de microfuga de 2 mL mientras que la punta del homogeneizador se lavó con 1 mL de Tampón de Extracción, y se añadió este volumen a la semilla homogeneizada. Después se centrifugó las muestras a 20.000 × *g* durante cinco minutos; esto dejó una pelotilla y un sobrenadante

acuoso que estaba recubierto por una capa aceitosa inmiscible que contenía cuerpos oleosos intactos y fracturados, así como también TAG libre. La capa de aceite superior fue empujada suavemente hacia el costado del tubo, y se descartó la capa acuosa y la pelotilla de material. Después se volvió a suspender la capa de aceite desde el costado del tubo colocándola en un dispositivo de vórtice en tampón de extracción y fue colocada en un tubo de microfuga fresco de 2 mL. Se completó el volumen final hasta 0,5 mL con tampón de extracción.

Preparados de Cuerpos Oleosos purificados de semillas de Arabidopsis thaliana y residuos de cisteína reticulantes entre las oleosinas manipuladas genéticamente

Se trituró 25 mg de semillas de *Arabidopsis* (de plantas transformadas, como se describe en el Ejemplo 5) en 300 µl de tampón de extracción (tampón fosfato 10 mM, pH 7,5 que contenía sucrosa 600 mM) usando un Homogenizador Wiggenhauser D–130. Se trituró la semilla hasta que quedó completamente molida, y la muestra tuvo un aspecto “cremoso” y espumoso a medida que se liberaba almidón de las semillas. La punta del homogeneizador fue lavada en 1 ml de tampón, y se añadió este volumen a la semilla triturada. Se preparó muestras hasta este momento en lotes de 4, las que fueron centrifugadas después a 14.000 rpm durante 5 minutos. Se utilizó una punta fina para cargar gel para empujar suavemente la capa de aceite hacia el costado del tubo, y se retiró la capa acuosa introduciéndola en un tubo fresco. La capa de aceite fue resuspendida desde el costado del tubo utilizando tampón de extracción, y fue colocada en un tubo fresco de 2 ml. El volumen final fue completado hasta 0,5 ml (leído en el costado del tubo) con tampón de extracción, las muestras fueron divididas en dos, se añadió agente de oxidación (GSSG 3mM) a un tubo y se incubó a temperatura ambiente durante 10 minutos. Después se añadió preparados de cuerpos oleosos a un volumen igual de 2 x tampón de carga de gel, y se hirvió por 5 minutos antes de cargar sobre un gel.

Las muestras fueron procesadas en NuPAGE® Novex 4–12% Bis–Tris Midi Gels (Invitrogen) prevaciado, en un sistema con un equipo para gel Criterion (Bio–Rad), o en una gradiente de gel Bis–Tris NuPAGE® Novex 12% 1,0 mm, de 15 pocillos, cat# NP0343BOX, con tampón de elución NuPAGE® MES SDS (sólo para Bis–Tris Geles) (20X), cat# NP0002–02, o sobre geles Tris–HCl vaciados a mano. Los geles fueron teñidos con SafeStain (Invitrogen) para mostrar la proteína total cargada o transferida sobre la membrana de nitrocelulosa, utilizando el sistema iBlot (Invitrogen). En cada caso, el control negativo fue una muestra extraída de semilla de Columbia de tipo salvaje, y el control positivo fue el mismo método de extracción (aunque la molienda se realizó mediante mortero y

pisón) llevado a cabo en semillas de sésamo de tipo salvaje. Se cargó sobre el gel 10 µl de cada muestra y de control negativo, y se utilizó 5 µl para el control positivo.

Después de la transferencia la membrana fue bloqueada en una solución de leche descremada en polvo al 12,5 % en TBST (Tris 50 mM pH 7,4, NaCl 100 mM, Tween al 0,2%) por lo menos durante 1,5 horas. Después se lavó la membrana en TBST 3 x 5 minutos antes de incubar con anticuerpo primario (anti–sésamo) a 1/1.000 en TBST durante una hora a temperatura ambiente. Después de tres lavados más con TBST, se llevó a cabo la incubación con un anticuerpo secundario (anti–conejo) a 1/5000 durante una hora a temperatura ambiente. La membrana experimentó tres lavados adicionales y después se desarrolló la señal usando el protocolo de quimioluminiscencia estándar.

La Figura 11 muestra la acumulación de unidades de oleosina de semillas de sésamo sobre los cuerpos oleosos bajo el control del promotor CaMV35S. Puede verse que se descubrió que la oleosina y la polioleosina recombinantes se acumularon en las semillas de *Arabidopsis thaliana* y se dirigieron correctamente a los cuerpos oleosos (Figura 11). Además, se puede observar que en la presencia de agente de oxidación durante 10 minutos las oleosinas recombinantes que contenían cisteínas formaron reticulaciones como lo evidenció la aparición de oligómeros y la correspondiente desaparición de las formas monoméricas en estas muestras y no en los cuerpos oleosos de tipo salvaje o transgénicos no oxidados.

El efecto de incrementar el número de sitios potenciales de reticulación en un péptido de oleosina en la integridad del CO *in planta* y en la estabilidad de la emulsión, puede evaluarse como sigue.

Determinación cuantitativa de la integridad de los COs

La evaluación de la estabilidad e integridad de los COs utilizando absorbencia (OD_{600}), conteo directo de los COAs mediante un hemocitómetro, o evaluación visual de la coalescencia por microscopía, demostró ser sumamente variable y entre otras cosas fue influenciada por el grado de agitación del muestreado preliminar, la cantidad de muestra extraída y el tiempo dejado bajo el microscopio. Para evitar esto los solicitantes idearon un método simple para cuantificar la cantidad de TAG liberado de los COs hacia el medio circundante durante diversos tratamientos como un medio para comparar la integridad. Cantidades esencialmente iguales (basándose en la estimación de TAG por FAMES–GC/MS y en la determinación de proteína según Bradford) de preparados de CO fueron completadas hasta un volumen total de 200 µL utilizando tampón AOB (que contenía Proteinasa K [PNK] cuando fue apropiado

con una razón 1:1 de PNK:total de proteínas en las muestras de CO en tubos de vidrio con inserciones para GC de 250 µL, que+ fueron tapados con una tapa de material plástico. Después del tratamiento (temperatura elevada o exposición a PNK) se añadió a la muestra 15 µL de aceite de pescado (Vitamax®, Australia), y se mezcló en un dispositivo de vórtice seguido por centrifugación a 5.200 g por un minuto. La adición de aceite de pescado seguida por tratamiento en el dispositivo de vórtice permitió que cualquier TAG que se hubiera fugado de los COs se mezclara con el aceite de pescado añadido y fuera hecho flotar con una breve centrifugación. Se tomó muestras de 4 µL de la fase de aceite, y se las sometió a esterificación metílica de ácido graso (FAME), después de lo cual fueron analizadas mediante GC–MS (número de modelo de Shimadzu, equipado con una columna capilar 50mQC2/BPX70–0,25 GC (SGE) como se describe en Browse *y otros* (1986). En la ausencia de aceite de pescado añadido, la cantidad de TAG que se había fugado de los COs fue demasiado pequeña para formar una capa visible de la que se pudiera tomar muestras, aún después de la centrifugación; en tal caso el volumen máximo hubiera sido de 6 µL. Los perfiles lipídicos muy diferentes del aceite de pescado y el aceite de sésamo nos permitieron distinguir fácilmente entre el TAG que se había fugado y el TAG añadido.

Utilizando los estándares internos C15:0 y C17:0 los solicitantes pueden calcular las cantidades absolutas de C18:2 (el lípido principal en el aceite de semillas de sésamo) recuperado después del tratamiento.

Determinación de integridad de los COs y de la estabilidad de la emulsión a temperatura elevada

Las emulsiones de aceite en agua son menos estables a temperaturas elevadas; por ello es interesante investigar si las oleosinas modificadas con diversas cantidades de cisteínas introducidas influyen en la integridad de los COs y los COAs a una temperatura elevada. Para lograr esto los solicitantes determinaron la integridad (usando el método descrito anteriormente) de COs (que contienen diferentes oleosinas) en un tampón fosfato (tampón fosfato de sodio 50 mM pH 8, NaCl 100 mM) a 95˚C. Los COAs fueron calentados durante dos horas. La integridad fue determinada como antes.

El efecto de razones más elevadas de oleosina reticulada:TAG para incrementar la estabilidad de los OBs en el fluido ruminal puede evaluarse como sigue.

Determinación de integridad de los COs en fluido ruminal

Uno de los objetivos del disulfuro era proveer algún grado de protección de la biohidrogenación por la microflora ruminal. La evaluación de la estabilidad de los COs con fluido ruminal puede efectuarse como sigue. Se añade

COs a un volumen igual (25 µL) de fluido ruminal. Las muestras son incubadas a 39°C durante 0, 15, 30, 60, 120 y 240 minutos; al final de la incubación se añade un volumen igual de tampón de carga (Invitrogen), se mezcla y se calienta a 70°C durante 10 minutos. Se compara 15 µL de cada mezcla de muestra/tampón de carga mediante SDS–PAGE/inmunoblot. La integridad es determinada como antes.

Análisis de integridad de los COs en Proteinasa K

Para investigar la influencia de la oleosina modificada en un entorno altamente degradante controlable y repetible, se determina la integridad (usando el método descrito anteriormente) de COAs (que contienen diferentes oleosinas modificadas) después de incubar en un tampón fosfato (tampón fosfato de sodio 50 mM pH 8, NaCl 100 mM) que contiene 1:1 (g/g proteína) Proteinasa K (Invitrogen) a 37 °C durante 4 horas. Aunque se alcanza la máxima actividad de la Proteinasa K a menos de 65 °C, se utiliza la temperatura más baja a fin de reducir la influencia de la temperatura en la inestabilidad de los COs. La integridad es determinada como antes.

EJEMPLO 7: Producción de cuerpos oleosos en las hojas de *Arabidopsis thaliana*

Para producir cuerpos oleosos en tejido vegetativo, es necesario producir triacilglicerol en tal tejido (por ejemplo, hojas).

Producción de triacilglicerol en las porciones vegetativas de la planta

En la mayoría de las plantas (incluso en *Lolium perenne*) la mayor parte de los lípidos de las hojas están fijados a un esqueleto de glicerol y existen como diacilgliceroles. Estos están incorporados en bicapas de lípidos donde ellos funcionan como membranas de múltiples orgánulos subcelulares o como la membrana de la célula misma. La mayor parte de las bicapas de lípidos en la hoja está constituida por la membrana tilacoidal del cloroplasto. Una cantidad menor de lípidos de las hojas existe como ceras epicuticulares y un porcentaje aún más pequeño está presente en forma de triacilglicerol (TAG).

La mayoría de las plantas sintetizan y almacenan TAG en embriones en desarrollo y células de polen donde es utilizado después para proveer energía catabolizable durante la germinación y el crecimiento de los tubos de polen. Las plantas dicotiledóneas pueden acumular hasta aproximadamente el 60% del peso de sus semillas como TAG. Habitualmente, este nivel es considerablemente más bajo en las semillas de monocotiledóneas donde la principal forma de almacenamiento de energía consiste en carbohidratos (por ejemplo, almidón). El único paso obligado en la biosíntesis de TAG es el último, es decir, la adición de un tercer ácido graso a un diacilglicerol existente, con lo cual

se genera TAG. En las plantas este paso es llevado a cabo por una de tres enzimas que incluyen: acil CoA:diacilglicerol aciltransferasa (DGAT1), una acil CoA: diacilglicerol acil transferasa no relacionada (DGAT2), y fosfolípido:diacilglicerol aciltransferasa (Zou *y otros*, 1999; Bouvier–Navé *y otros*, 2000; Dahlqvist y otros, 2000; Lardizabal y otros, 2001). La sobreexpresión de la región transcrita de cualquiera de estos genes en las porciones vegetativas de las plantas conduce a la formación de gotitas de TAG en el citoplasma de las células de las hojas, como se ha demostrado mediante la sobreexpresión de: DGAT1 de Arabidopsis en el tabaco, por Bouvier–Navé y otros, (2000); DGAT2 del árbol de tung en levadura y tabaco por Shockey y otros, (2006); PDAT de Arabidopsis en Arabidopsis por Stahl y otros, (2004). En algunos casos se mostró que la sobreexpresión de DGAT1 de Arabidopsis acrecentaba el nivel total de lípidos pero no necesariamente por la acumulación de TAG, por ejemplo en raíces vellosas de *Lotus japonicus* (Bryan y otros, 2004) y en hojas de *Lolium perenne* (Cookson y otros, 2009).

Para demostrar la acumulación de TAG en las hojas de estas plantas se puede comparar la cantidad total de extracción de lípido de las hojas de estas plantas con aquella de plantas no transformadas o de plantas transformadas con el vector binario vacío, asegurándose que las plantas sean cultivadas en las mismas condiciones ambientales y que las hojas muestreadas sean fisiológicamente equivalentes. Con los estándares internos adecuados se puede lograr la cuantificación de la extracción total de lípido usando el análisis de FAMES GC–MS (descrito por Winichayakul y otros, 2008, Delivery of grasses with high levels of unsaturated, protected fatty acids. Proceedings of the New Zealand Grassland Association, 70:211–216). Alternativamente, se puede extraer los lípidos totales mediante el método de Folsch (Folsch y otros, 1957 J. Folsch, M. Lees y G.A. Slone–Stanley, A simple method for the determination of total lipid extraction and purification, Journal of Biological Chemistry 226 (1957), pp. 497–507) y cuantificarlo usando estándares internos apropiados con un GC–MS equipado con un Restek (Restek Corp., Bellefonte, PA) columna RTX65TG.

Se tomó muestras de hojas de plantas que sobreexpresaban el DGAT1 de *A. thaliana* (S205A) y la construcción de oleosina de semillas de sésamo (ya sea Oleo_0–0, Oleo_1–1, Oleo_1–3, Oleo_3–1 u Oleo_3–3, SEQ ID NOs 11–20, Figuras 1–5) y fueron analizadas mediante SDS–PAGE/inmunoblot usando antisuero policlonal anti–oleosina de semillas de sésamo. Puede observarse que se descubrió que la oleosina recombinante se acumula en las hojas de *Arabidopsis thaliana* (Figura 12).

La expresión y acumulación simultáneas de la proteína oleosina/oleosina modificada en la misma célula (por ejemplo, células de hoja) tendrá como resultado la producción de cuerpos oleosos de triglicéridos encapsulados por una monocapa de fosfolípido incrustada con oleosina; esto ha sido demostrado con una oleosina no modificada en levadura (Ting y otros, 1997) y semillas (Abell y otros, 2004).

Preparaciones de cuerpos oleosos a partir de las hojas de Arabidopsis thaliana transgénica

Es posible extraer cuerpos oleosos de las hojas de Arabidopsis thaliana transgénica que expresa DGAT1 (S205A) y la construcción de oleosina de semillas de sésamo (ya sea Oleo_0–0, Oleo_1–1, Oleo_1–3, Oleo_3–1 u Oleo_3–3, SEQ ID NOs 11–20, Figuras 1–5).

Se puede evaluar el efecto de incrementar la cantidad de sitios potenciales de reticulación en un péptido de olesina en los COs de tales plantas midiendo la integridad de los COs y la estabilidad de la emulsión, como se describe en el Ejemplo 6.

Diseño y construcción de oleosinas que contienen más de tres residuos de cisteína en cada brazo hidrófilo

Las líneas Ole–3,3 tenían grados substanciales de altos niveles lipídicos en la forma de TAGs cuando fueron coexpresadas con DGAT1 (S205A) mientras que las líneas que contenían Ole–0,0 no tenían altos niveles lipídicos superiores al control que sobreexpresa DGAT1. Las Ole–1,1, Ole–1,3 y Ole–3,1 mostraron que había una correlación entre el nivel de acumulación de lípidos en la hoja y el incremento en el número de cisteínas manipuladas genéticamente en cada brazo (Tabla 3).

Tabla 3. Composición de ácidos grasos (como % de peso seco) de hojas de Arabidopsis que expresan vector de control, DGAT1 (S205A) solo, o DGAT1 (S205A) y diferentes formas de oleosina (que no contienen ninguna cisteína adicional o hasta 3 cisteínas adicionales en cada brazo hidrófilo).

Perfil de ácido graso	**Vector de control**	**DGAT1 SOLO DGAT1SA #2**	**DGAT1 + OLE 0–0 (#11)**	**DGAT1 +OLE 1–1 (#9)**	**DGAT1 +OLE 1–3 (#5)**	**DGAT1 +OLE 3–1 (#18)**	**DGAT1 +OLE 3–3 (#47)**

C16:0	0,55 ± 0,035	0,55 ± 0,001	0,54 ± 0,014	0,57 ± 0,001	0,68 ± 0,042*	0,62 ± 0,084	0,95 ± 0,049*
C16:1	0,085 ± 0,007	0,105 ± 0,007	0,11 ± 0,001	0,13 ± 0,014	0,1 ± 0,001	0,135 ± 0,021	0,11 ± 0,001
C16:3	0,34 ± 0,021	0,41 ± 0,028	0,42 ± 0,007	0,48 ± 0,028	0,51 ± 0,035	0,55 ± 0,071*	0,62 ± 0,049*
C18:1	0,095 ± 0,007	0,075 ± 0,007	0,1 ± 0,001	0,185± 0,007*	0,34 ± 0,007*	0,2 ± 0,014*	0,61 ± 0,014*
C18:2	0,55 ± 0,014	0,46 ± 0,035	0,56 ± 0,014	0,77 ± 0,049*	0,97 ± 0,007*	0,79 ± 0,113*	1,82 ± 0,113*
C18:3	1,67 ± 0,056	1,91 ± 0,028	1,78 ± 0,014	1,68 ± 0,028	1,74 ± 0,014	1,9 ± 0,28	2,29 ± 0,056*
C20:0	No detectado	No detectado	No detectado	No detectado	No detectado	No detectado	0,054 ± 0,003

La correlación entre el incremento en lípido total (que demostró ser TAG) y la cantidad de cisteínas manipuladas genéticamente en los dominios hidrófilos indicó que la cantidad de cisteínas puede ser una manera de influir en el nivel de TAG deseado. Por lo tanto, se diseñó nuevas construcciones que contenían más de tres cisteínas por brazo hidrófilo. Aunque no es posible colocar un número infinito de cisteínas/brazo hidrófilo, las limitaciones incluyen:

- Longitud de los brazos – si se añadiera residuos adicionales para dejar espacio para las cisteínas, entonces se reduciría eventualmente el grado de interacción de los dominios hidrófobos, ya que su capacidad para ponerse en contacto sería limitada por su libertad de moverse en el CO.
- Mantención de la proporción de los residuos anfipáticos +, – y – ; si el equilibrio de estos residuos y la distribución de estos residuos se altera drásticamente, es probable que los brazos hidrófilos no interactúen efectivamente con la superficie del CO, y así no proporcionarían ninguna protección contra las lipasas o contra la coalescencia.

- Disponibilidad de azufre – el incremento de la cantidad de cisteínas por molécula de oleosina puede colocar a la planta bajo un estrés nutricional si el azufre es limitante.

La cisteína-oleosina original fue manipulada genéticamente para que portara 3 cisteínas no apareadas espaciadas con relativa uniformidad en cada brazo, reemplazando aminoácidos y predominantemente aquellos de los que se podía predecir que eran neutros o cargados pero no hidrófobos.

Presuntamente, la oleosina necesita tener un cierto nivel de carga negativa y en el C-terminal esto parece haber sido logrado por K(lys); por consiguiente continuar la estrategia de hacer un trueque de residuos cargados o neutros por cisteínas adicionales puede tener como resultado una mala estabilidad en términos de evitar la coalescencia. Por otra parte, en la región hidrófila N-terminal parece que quedan demasiado pocos residuos entre las cisteínas manipuladas genéticamente para permitir seguir substituyendo residuos mientras se mantiene el espaciamiento y la oscilación entre aminoácidos con carga positiva y negativa. Por esto, para los terminales N- y C- se agregó residuos adicionales (cisteínas) en vez de substituir los residuos existentes con cisteínas. Alternativamente, se podría haber usado una oleosina con brazos hidrófilos más largos.

Se diseñó también otras dos construcciones (Ole-5,6 y Ole-6,7). Estas no estaban desequilibradas a propósito en términos de los residuos de cisteína por brazo, sino que estaban organizadas para intentar dar típicamente 4–5 residuos entre cada cisteína. De hecho, para incrementar las cisteínas a 6 en el brazo N–terminal, fue necesario generar residuos adicionales (contrariamente a la substitución de residuos existentes); esto se logra replicando los primeros 6 residuos de la Ole–3,3.

En lugar de diseñar secuencias de nucleótidos completamente nuevas, se añadió el triplete TGT para codificar cisteína (donde fuera apropiado) para generar Ole_5,6. Para los residuos de glutamina adicionales se utilizó el codón triplete GGA. Para los 6 residuos N–terminales adicionales en Ole_6–7 se replicó el N-terminal de Ole_3,3, y se fusionó en el marco.

Se diseñó una estrategia de subclonación para que fuese idéntica a las cisteínas de oleosina iniciales, es decir, fue subclonada en el oleoaceptor por NotI/NdeI. En seguida ésta fue recombinada por reacción de LR para obtener pRSH1 (Winichayakul y otros, 2008). Esencialmente, se coloca DGAT1 (S205A) de Arabidopsis y oleosina bajo sus propios promotores CaMV35s y terminadores OCS. Los clones DGA1 y de oleosina contienen un intrón UBQ10.

Se utilizó NetGene2 para predecir el patrón de corte y empalme de Ole_5,6 y Ole_6,7. Se predijo que ambas tenían un solo sitio dador y aceptor en la cadena directa (se predijo que ambas tendrían una probabilidad de reconocimiento muy alta), y ningún sitio en la cadena complementaria.

Los datos indican que las oleosinas que contienen cisteínas 1,3 o 3,1 acumulan de hecho niveles detectables de TAG pero éste es ciertamente inferior al de las oleosinas con cisteínas 3,3 (la 1,1 acumuló vestigios en tanto que la 0,0 no lo hizo). Esto sugiere aún más fuertemente que es probable que las oleosinas 5,6 y 6,7 acumulen TAG en un grado aún mayor que la construcción 3,3. Los primeros datos de las construcciones 5,6 y 6,7 estarán disponibles pronto.

Transformación de oleosinas que contienen cisteínas manipuladas genéticamente y DGAT1 en Arabidopsis thaliana de tipo salvaje

Cinco construcciones híbridas de disulfuro–oleosina/DGAT1 (S205A) y un control (construcción que contiene DGAT1 (S205A) pero no oleosina) fueron transferidas después al vector binario de planta pRSh1 (Winichayakul y otros, 2008) y transformadas en Arabidopsis thaliana de tipo salvaje utilizando transformación mediada por Agrobacterium.

Se siguió una modificación del método tradicional de inmersión floral ya que se ha informado que la inmersión floral tiende a dañar las silicuas en desarrollo debido a la presencia de detergente en los inóculos (Martínez–Trujillo y otros, 2004). Por lo tanto, se llevó a cabo una inoculación gota a gota en cada flor, utilizando una micropipeta. Se repitió la inoculación después de una semana para introducir el inóculo en las flores recién desarrolladas. Las semillas fueron recolectadas cuando las silicuas se habían secado, después fueron limpiadas y plantadas para hacer un escrutinio de los transformantes.

El escrutinio de los transformantes se llevó a cabo mediante selección BASTA, y los transformantes homocigotos fueron seleccionados utilizando análisis de razón de segregación para la resistencia a BASTA.

Transformación de oleosinas que contienen cisteínas manipuladas genéticamente y DGAT1 en Trifolium repens de tipo salvaje

La transformación en *Trifolium repens* (trébol blanco) se llevó a cabo de acuerdo con el procedimiento de Voisey *y otros*, (1994).

Se pesó las semillas para proveer aproximadamente 400 – 500 cotiledones (es decir 200 – 250 semillas) para que fueran disecadas (0,06 g = 100 semillas). En un tubo de centrífuga, las semillas fueron enjuagadas con etanol al 70% por un minuto. Fueron esterilizadas superficialmente en blanqueador (cloro disponible 5%) sacudiendo sobre una mezcladora circular durante 15 minutos,

seguido por cuatro lavados en agua estéril. Se dejó las semillas saturándose durante toda la noche a 4˚C.

Las mismas construcciones utilizadas para transformar Arabidopsis (arriba) fueron mantenidas en la cepa Agrobacterium GV3101 e inoculadas en 25 mL de caldo MGL (Tabla 4) que contenía espectinomicina a una concentración de 100 mg/L. Los cultivos fueron desarrollados durante toda la noche (16 horas) sobre un agitador giratorio (200 rpm) a 28°C. Los cultivos bacterianos fueron cosechados por centrifugación (3000 x g, 10 minutos). Se extrajo el sobrenadante y las células fueron resuspendidas en una solución de 5 mL de $MgSO_4$ 10 mM.

Los cotiledones fueron disecados de las semillas usando un microscopio de disección. Primero se retiró el tegumento y el endosperma. Los cotiledones fueron separados del radical con el escalpelo, colocando la cuchilla entre los cotiledones y cortando a través del pecíolo restante. Los explantes cotiledonarios fueron cosechados sobre un disco de filtro estéril en medio CR7.

Para la transformación, se dispensó una alícuota de 3 µl de suspensión de Agrobacterium sobre cada cotiledón disecado. Las placas fueron selladas y cultivadas a 25˚C durante un fotoperíodo de 16 horas. Después de un período de de cocultivo de 72 horas, los cotiledones transformados fueron transferidos a placas que contenían medio CR7 complementado con glufosinato de amonio (2,5 mg/L) y timentina (300 mg/L) y se los devolvió a la sala de cultivo.

Después de la regeneración de los brotes, los explantes fueron transferidos a medio CR5 complementado con glufosinato de amonio (2,5 mg/L) y timentina (300 mg/L). Los brotes que se estaban regenerando fueron subcultivados tres veces por semana con una solución que contenía medio CR5 fresco.

A medida que ocurría la formación de las raíces, se transfirió las plántulas a tinas que contenían medio CR0 con una selección de glufosinato de amonio. En esta etapa, se dividió grandes racimos de regenerantes en plántulas individuales. Las plantas enteras, que habían echado raíces, y que habían crecido bajo selección fueron colocadas luego en cilindros de turba estéril. Una vez establecidas en los cilindros de turba, las plantas fueron transferidas después al invernadero.

Tabla 4. Composiciones de los medios utilizados para la transformación de *Trifolium repens*.

A.	CR#0 Sales MS Vitaminas B5 Sucrosa 30 g/L pH 5,8 (KOH) agar 8,0 g/L
	CR#5 Sales MS Vitaminas B5 sucrosa 30 g/L BA 0,1 mg/L NAA 0,05 mg/L pH 5,8 (KOH) agar 8,0 g/L
B.	CR#7 Sales MS Vitaminas B5 sucrosa 30 g/L BA I,0 mg/L NAA 0,05mg/L pH 5,8 (KOH) agar 8,0 g/L

C.	MGL Manitol 5,0 g/L Acido L glutámico 1,0 g/L KH_2PO_4 250 mg/L $MgSO_4$ 100 mg/L NaCl 100 mg/L Biotina 100 mg/L Bactotriptona 5,0 g/L Extracto de levadura 2,5 g/L pH 7,0 (NaOH)

Los resultados de FAMES GC/MS mostraron que el *Trifolium repens* transgénico (que contenía DGAT1 (S205A) y Ole_3,3 u Ole_5,6 u Ole 6,7) tenía perfiles lipídicos totales elevados de las hojas en comparación con el tipo salvaje (Figura 17). Había una correlación general entre el nivel más elevado de lípidos de las hoja y el número más alto de cisteínas manipuladas genéticamente en la oleosina.

Los resultados de FAMES GC/MS mostraron que el *Trifolium repens* transgénico (que contenía DGAT1 (S205A) y Ole_3,3 u Ole_5,6 u Ole 6,7) tenía perfiles lipídicos elevados C18:1 y C18:2 de las hojas en comparación con el tipo salvaje, como también se ve en Arabidopsis (Figura 18). El nivel más alto C18:1 y C18:2 de las hojas se encontró en plantas transformadas con la oleosina que contenía la cantidad más elevada de cisteínas manipuladas genéticamente.

Determinación del ensamblaje de cuerpos oleosos en hojas (y semillas)

Se realizó un escrutinio adicional mediante análisis de inmunoblot (con un anticuerpo anti-oleosina de semillas de sésamo, (Scott y otros, 2007) para determinar las líneas que sobreexpresaban la proteína oleosina. Usando este método, se extrajo cuerpos oleosos (COs) de semillas T2 de transformantes putativos usando gradiente de densidades de sucrosa, o se extrajo la proteína total de hojas en tampón desnaturalizante/reductor y se separó las proteínas en SDS–PAGE, se las transfirió a una membrana de nitrocelulosa, y se las estimuló con un anticuerpo cultivado contra la oleosina de sésamo (Scott y otros, 2007).

Se extrajo cuerpo oleoso crudo (CO) de 25 mg de semillas en 500 µL de tampón OB (fosfato de sodio 10 mM, pH 7,5 que contenía sucrosa 600 mM). Después de centrifugar a 13.000 x g, se succionó cuidadosamente la capa acuosa y se resuspendió la masa de grasa en 200 µL de tampón OB sin agitar la pelotilla en el fondo. Se agregó a cada 20 µL de extracto de CO 4 x colorante de carga y 10 x de agente reductor, fueron calentados a 70°C durante cinco minutos, y cargados sobre gel de poliacrilamida al 4–12% para análisis de inmunoblot. La transferencia fue incubada en un anticuerpo de α–sésamo oleosina (1°Ab) a una dilución de 1:750 durante una hora, y durante otra hora en anticuerpo secundario (1:10,000).

La oleosina se expresa naturalmente en semillas y no en las hojas. Sin embargo, y dado que hemos expresado DGAT1 conjuntamente con oleosina bajo el control de promotores CaMV35S, podía preverse que esto permitiría la acumulación en las hojas de niveles detectables de oleosina. Las hojas de líneas transformadas con alta expresión de oleosina recombinante en las semillas (identificadas mediante análisis de inmunoblot) fueron analizadas por inmunoblot utilizando anticuerpos cultivados contra la oleosina de sésamo.

En la siguiente Tabla 5 se consigna la cantidad de transformantes putativos generados y la cantidad de plantas que expresan oleosina recombinante en las semillas y hojas.

Tabla 5

ID de Construcción del gen	Cantidad de transformantes putativos (en base a la resistencia a BASTA)	Número de semillas de líneas que fueron analizadas por inmunoblot (anticuerpo anti oleosina de semillas de sésamo)	Número de líneas con una banda inmunorreactiva positiva del tamaño adecuado en el extracto de semillas	Número de líneas con una banda inmunorreactiva positiva del tamaño adecuado en el extracto de hojas
pRSh1–DGAT1(S205A) control	8	N/A	N/A	
pRSh1–DGAT1(S205A)–Ole_0–0	14	8	7	3
pRSh1–DGAT1(S205A)–Ole_1–1	22	2	1	1
pRSh1–DGAT1(S205A)–Ole_1–3	20	0	0	1
pRSh1–DGAT1(S205A)–Ole_3–1	23	8	4	2
pRSh1–DGAT1(S205A)–Ole_3–3	54	22	16	5

Debe observarse que el nivel de oleosina recombinante que se acumuló en las hojas fue considerablemente más bajo que en las semillas. Sin

embargo, la proporción de líneas individuales que acumularon niveles detectables en las hojas fue mucho mayor que cuando la oleosina fue expresada sola (Roberts Lab, datos no publicados) lo que indica que la coexpresión de DGAT1 y oleosina en la hoja ha llevado a la acumulación de niveles más altos de oleosina.

Análisis de hojas de plantas transgénicas que coexpresan DGAT1 (S205A) y oleosinas de disulfuro

Se hizo germinar semillas de líneas homocigotas que sobreexpresaban la proteína oleosina en las semillas para permitir que se desarrollaran durante 2, 3, 4 o 5 semanas. Se cosechó suficiente material de hoja para análisis de FAMES- GC–MS, así como también por GC–MS utilizando una columna Restek RTX 65–TG que permite la separación e identificación de ácidos grasos libres, diacilglicéridos, ésteres de cera, ésteres de esterol, y triacilglicéridos sin derivación.

Preparación de material para análisis de FAMES GC/MS

Se colocó 10 mg de polvo de hojas secadas por congelación en un tubo de 13 x 100 mm con tapa roscada, se añadió 10 µL de estándar interno no metilado (C15:0 FA, 4 mg/mL disuelto en heptano). A esta mezcla se añadió 1 mL del reactivo de HCl metanólico (1 mL de solución 3 M diluido a 1 M usando metanol seco que había sido tratado con 2,2–dimetoxipropano al 5% como un eliminador de agua. Luego se sometió el tubo a un barrido con N_2 gaseoso y después se selló inmediatamente con un tapón revestido con Teflón, y se calentó a 80°C durante una hora. Después que los tubos se hubieron enfriado a la temperatura ambiente se añadió 10 µL de estándar premetilado (4 mg/mL de C17:0 disuelto en heptano). A esta mezcla se añadió 0,6 mL de heptano y 1,0 mL de NaCl al 0,9% (peso/v), y se mezcló intensamente en un dispositivo de vórtice. Después de centrifugar a 500 rpm durante un minuto a temperatura ambiente, se recolectó 100 µL de la capa superior (que contenía heptanos) y se los transfirió a un inserto de vidrio de fondo plano dentro de un frasco marrón para análisis de GC/MS.

Análisis de FAMES GC/MS

Se realizó el análisis de FAMES GC/MS usando la columna capilar SGE BPX70 (50 m x 0,22 mm x 0,25 µm). La condición de GC–MS fue la siguiente: se programó la temperatura de 80°C a 150°C a razón de 15°C/minuto y después de 250°C a 8°C/minuto y se mantuvo isotérmica durante 10 minutos. Las muestras fueron inyectadas en un modo dividido; caudal total de 28,4 mL/minuto; caudal de columna de 0,82 mL/minuto; y un caudal de purga de 3,0 mL/minuto. Se mantuvo la presión a 150 kPa; la temperatura de la fuente de iones era de 200°C, y se mantuvo una temperatura de interfaz de 260°C. Los compuestos de objetivo

se obtuvieron mediante espectrometría de masas en un modo de barrido, partiendo a 50 m/z y terminando a 350 m/z.

Extracción de TAG y lípido polar

El TAG fue extraído usando un método modificado de Ruiz–López y otros, (2003). Brevemente, para cada análisis de TAG, se colocó entre 34 y 80 mg de polvo de hojas secadas por congelación en un tubo calibrado de 13 mm con tapa roscada y se pesó, se añadió 2,4 mL de NaCl 0,17 M en MeOH, y se mezcló en un dispositivo de vórtice. Después de agregar 4,8 mL de heptano y 10 µL de estándar interno (C14:0, 10 µg.µL–1), la suspensión fue mezclada suavemente e incubada sin agitación en un baño María a 80°C durante dos horas. Después de enfriar a la temperatura ambiente, la fase superior (que contenía lípidos) fue transferida a un tubo fresco con tapa roscada y evaporada hasta sequedad bajo una corriente de N gaseoso. Finalmente, se resuspendió el polvo en 100 µL de heptanos, se mezcló intensamente y se transfirió después a un inserto de vidrio de fondo plano colocado dentro de un frasco de vidrio marrón para análisis de TAG.

Análisis de TAG con GC–MS

El análisis de TAG se llevó a cabo en un GC de Hewlett Packard (HP) y un MS de Shimadzu Scientific Instruments Inc. (QP2010). Todos los análisis fueron realizados con una columna capilar RESTEK de MXT–65TG (65% de difenilo – 35% de dimetil polisiloxano, de 30,0 m x 0,10 µm de espesor x 0,25 mm de diámetro) en modo de ionización Electron Impact (EI). Como gas portador se utilizó helio. Todas las muestras fueron inyectadas en un modo sin división de flujo, en alícuotas de 1,0 µl, y con un caudal de columna de 1,2 mL.min–1. Se programó la cromatografía de gases desde 200 a 370°C a 15°C.min–1 y se mantuvo isotérmica a 370°C durante 15 minutos. La temperatura del orificio del inyector de muestras se mantuvo a 350°C, la temperatura del horno de columna a 200°C, con una presión de 131,1 kPa y un caudal de purga de 3,0 mL.min–1. Las condiciones espectrométricas de masas fueron las siguientes: la temperatura de la fuente de iones se mantuvo a 260°C durante los tiempos de ejecución de GC–MS, los espectros de masas se obtuvieron con un voltaje de ionización de 70 eV con una corriente de emisión de 60 µA y una temperatura de interfaz de 350°C. El modo de adquisición fue mediante barrido con una velocidad de 5000, 0,25 segundo por barrido. Se recolectó picos de cromatografía con una razón de masa a carga desde 45 m/z hasta 1.090 m/z partiendo a los 9 minutos y terminando a los 25 minutos.

EJEMPLO 8: Otras oleosinas, caoleosinas y esteroleosinas modificadas genéticamente para contener residuos de cisteína adicionales en los brazos hidrófilos N– y C– terminales

Los solicitantes han usado la misma estrategia que para la oleosina de semillas de sésamo, número de acceso AAD42942, (es decir, substituyendo residuos cargados de los que se predijo que estarían sobre la superficie de los COs con cisteínas) para manipular genéticamente cisteínas en los brazos hidrófilos N– y C– terminales de oleosinas, caoleosinas y esteroleosinas. En algunos casos ha sido posible substituir solamente los aminoácidos cargados negativamente (ácido glutámico y ácido aspártico) que están separados entre sí con relativa uniformidad. En el caso de la oleosina de sésamo AAD42942 fue necesario a veces transigir sobre la substitución de la carga. En los ejemplos que siguen debe observarse que dos caoleosinas (AAB71227 y AAF13743) contienen dos cisteínas endógenas en su brazo C–terminal. Estas fueron dejadas inalteradas en la manipulación genética.

Para determinar la posición de la substitución de aminoácidos cada proteína fue alineada con la oleosina de sésamo (AAD42942) en su forma original, así como también las formas que contenían 1 o 3 cisteínas por brazo hidrófilo (es decir ole_0,0, ole_1,1, ole_3,1, ole_1,3, ole_3,3). Los ácidos glutámicos y ácidos aspárticos potenciales en el N– o C–terminal de cada uno de los brazos hidrófilos (determinado mediante gráficos de hidrofobicidad) fueron resaltados después con recuadros grises, así como también los residuos lisina, arginina y glutamina relevantes (todos los cuales fueron alterados con éxito en la oleosina de sésamo (AAD42942). Después se consideró la mutación de estos residuos a cisteína junto con la separación entre ellos. Las sustituciones finales se muestran después con la secuencia de péptidos original y solamente la secuencia modificada genéticamente. En este caso sólo 3 cisteínas fueron manipuladas genéticamente en cada brazo; sin embargo, el número podría haber sido mayor o menor. Una propuesta alternativa habría sido trabajar con cada proteína aislada y simplemente empezar identificando las regiones hidrófilas por un gráfico de hidrofobicidad y luego empezar el proceso de substitución con el aminoácido cargado más apropiado.

La siguiente Tabla 6 muestra otras oleosinas y caoleosinas que fueron modificadas por los solicitantes para introducir cisteínas en las porciones hidrófilas.

Tabla 6

Tipo de proteína	**Planta de origen**	**Número de acceso**	**SEQ ID NO**
Oleosina	*Brassica oleraceae* (oleosina de polen)	CAA65272.1	90
Oleosina	Maíz	NP_001147032.1	91
Oleosina	Arroz	AAL40177.1	92
Caoleosina	Sésamo	AAF13743	93
Caoleosina	Soya	AAB71227	94
Caoleosina	Maíz	NP_001151906	95
Esteroleosina	Sésamo	AAL13315	96
Esteroleosina	Brassica napus	ACG69522	97
Esteroleosina	Maíz	NP_001152614.1	98

Tabla 7: Referencia a los SEQ ID NO en las oleosinas modificadas

Tipo de proteína	**Planta de origen**	**Número de acceso**	**SEQ ID NO**
Oleosina	*Brassica oleraceae* (oleosina de polen)	X96409	99
Oleosina	Maíz	NP_001147032.1	100
Oleosina	Arroz	AAL40177.1	101
Caoleosina	Sésamo	AAF13743	102
Caoleosina	Soya	AAB71227	103
Caoleosina	Maíz	NP_001151906	104
Esteroleosina	Sésamo	AAL13315	105
Esteroleosina	Brassica napus	ACG69522	106
Esteroleosina	Maíz	NP_001152614.1	107

La secuencia modificada puede expresarse como se describe en los ejemplos precedentes para producir cuerpos oleosos, emulsiones, células hospedantes transgénicas, plantas, etc., y para someter a prueba las propiedades de cada uno de ellos.

No se tiene la intención de limitar el alcance del invento sólo a los ejemplos mencionados anteriormente. Como podría apreciar un experto en la especialidad, muchas variaciones son posibles sin apartarse del alcance del invento.

EJEMPLO 9: Incremento de la producción de biomasa mediante la elevación en la concentración de CO_2 en el cloroplasto, la elevación de la tasa de asimilación de CO_2 y la elevación de la eficiencia intrínseca de uso del agua en las hojas

Los solicitantes han usado la misma estrategia del Ejemplo 7 (evitando el catabolismo de TAG en la hoja, lo que a su vez asegura que haya un reciclado continuo de CO_2 del piruvato mientras es usado por el complejo piruvato deshidrogenasa para generar Acetil-CoA para la biosíntesis de lípidos) no sólo para incrementar la tasa de asimilación de CO_2 sino también para elevar la eficiencia intrínseca en el uso del agua. El efecto neto de esto es elevar la presión parcial de CO_2 comparado con O_2 en el cloroplasto. Se midió las tasas de asimilación de CO_2 en aire (que contiene aproximadamente 79% de N_2, 21% de O_2 y 400ppm de CO_2) y en una mezcla de gases calibrada con O_2 reducido (que contiene 98% de N_2, 2% de O_2 y 400ppm de CO_2) para demostrar que las plantas que tenían TAG protegido (mediante la coexpresión de DGAT y cisteína oleosina) tenían niveles reducidos de fotorrespiración.

Calibraciones de IRGA

Se midió tasas de fotosíntesis para 6 plantas de tipo salvaje (WT) y 6 plantas transgénica (T) (DGAT+Ole3,3), a 200 µmol m^{-2} s^{-1} PAR (condición de crecimiento), entre las 11:00 y las 16:00 el 24 de Mayo de 2011 usando un sistema de fotosíntesis portátil (Li6400, LiCor Inc, Lincoln, NE, EE.UU.) equipado con una cámara para hoja estándar de 2x3 cm, termopar para las hojas y una fuente de luz LED azul-roja. La eficiencia intrínseca en el uso del agua fue calculada de la razón de fotosíntesis/conductancia (Osmond y otros, 1980). La temperatura del conjunto fue mantenida a 20°C, la razón de estomas fue establecida en 1,6 y el déficit en la presión de vapor fue entre 0,6 y 0,9 kPa. Para mediciones en condiciones no fotorrespiratorias, se conectó un tanque de oxígeno al 2% (certificado) en nitrógeno a la entrada del Li-6400.

Análisis estadístico

Se usó en este estudio una estadística de prueba t estándar (R 2.12) para comparar entre tratamientos o entre genotipos. Se obtuvo los puntos medios de cada genotipo, junto con el error estándar promedio de la diferencia entre dos puntos medios (SED).

Fotosíntesis neta y eficiencia intrínseca en el uso del agua

Las tasas de fotosíntesis fueron significativamente mayores (en condiciones fotorrespiratorias) en las plantas transgénicas comparadas con plantas de tipo salvaje; de una manera similar, las tasas de fotosíntesis fueron significativamente mayores (en condiciones no fotorrespiratorias) en las plantas

transgénicas comparadas con las plantas de tipo salvaje. (Figura 30, panel izquierdo). Las plantas con DGAT1-Ole_3,3 tuvieron mayores incrementos en fotosíntesis cuando se eliminó completamente la fotorrespiración usando un entorno con bajo O_2 comparadas con las de tipo salvaje (Figura 30, panel derecho), mostrando así que las plantas con DGAT1-Ole_3,3 tienen tasas elevadas de asimilación de CO_2 comparadas con plantas del tipo salvaje.

Las mediciones de la eficiencia intrínseca en el uso del agua fueron significativamente mayores en condiciones no fotorrespiratorias que con concentración de oxígeno ambiental para los genotipos WT y T (Figura 31, panel izquierdo). Con niveles ambientales de O_2 la iWUE fue consistentemente mayor en las plantas transformadas con DGAT1-Ole_3,3 que en las plantas del tipo salvaje; esto fue demostrado además por el hecho que las plantas con DGAT1-Ole_3,3 tuvieron menores incrementos en iWUE cuando se eliminó completamente la fotorrespiración usando un entorno con bajo O_2 (Figura 31, panel derecho), mostrando así que las plantas con DGAT1-Ole_3,3 tienen una iWUE más alta comparadas con plantas del tipo salvaje.

La conductancia estomática fue significativamente mayor en plantas del tipo salvaje que en plantas transformadas con DGAT1-Ole_3,3 en condiciones de oxígeno ambiental (Figura 32, panel izquierdo). En las plantas del tipo salvaje, la conductancia estomática se redujo levemente en condiciones no fotorrespiratorias en tanto que en las plantas transformadas con DGAT1-Ole_3,3 la conductancia estomática se incrementó en más de 25% comparando con las condiciones ambientales (Figura 32, panel derecho). La conductancia estomática para las plantas del tipo salvaje y aquellas transformadas con DGAT1-Ole_3,3 fue aproximadamente igual cuando la fotorrespiración fue eliminada por las condiciones de bajo O_2 (Figura 32, panel izquierdo).

La densidad estomática de las plantas del tipo salvaje y de las plantas transformadas con DGAT1-Ole_3,3 fue similar (Figura 33).

La consecuencia de transformar plantas con DGAT1-Ole_3,3 fue una elevación en las tasas de asimilación fotosintética, eficiencia acrecentada en el uso del agua y reducción de la conductancia estomática, lo que llevó a un incremento en la tasa de crecimiento que se vio en una mayor biomasa, mostrada en la siguiente Tabla 8 y en la Figura 34.

	Peso seco promedio de hoja/planta (mg) 20 días después de	Peso seco promedio de hoja/planta (mg) 30 días después de

	germinación	germinación
Tipo salvaje	30,9 ± 3,5	61,1 ± 3,3
OLE1_3,3	47,6 ± 4,4	90,0 ± 6,1

Las plantas que sobreexpresaron DGAT1(S205A) y OLE1_3,3 tuvieron aproximadamente 50% más biomasa que el tipo salvaje; esto incluyó el inicio de la emergencia de los tallos florales (~20 días después de la germinación) como también el desarrollo mediano a tardío de los tallos florales (~35 días después de la germinación).

EJEMPLO 10: Producción de TAG en las raíces de *Arabidopsis thaliana*

Se extrajo raíces de plantas con Ole_3,3 y del tipo salvaje usando los mismos procedimientos descritos en el Ejemplo 7. El análisis cuantitativo de FAMES (Figura 35 panel izquierdo) mostró que el contenido total de lípidos de las raíces con Ole_3,3 era 8,2% del DM mientras que el contenido total de lípidos de las raíces del tipo salvaje era 1,7% del DM. FAMES mostró también que el perfil lipídico de las raíces de las Ole_3,3 no era demasiado diferente del tipo salvaje (Figura 34, panel derecho). El cambio más notorio fue la proporción de C18:1 que fue de 4.0% en las raíces del tipo salvaje y se elevó más de cuatro veces a 18.1% del total de ácidos grasos en las raíces de plantas con DGAT1-Ole_3,3. A pesar del contenido total similar de lípidos (~8%) de las hojas y raíces de las plantas DGAT1-Ole_3,3, el análisis de TAG demostró que una mayor porción del total de lípidos en las raíces era TAG cuando se comparó con el material de hojas (Figura 36). Esto se debió probablemente a que una porción mucho mayor del total de lípidos en las hojas consistía en lípidos de membrana (predominantemente membrana tilacoidal).

EJEMPLO 11:

Los incrementos en la tasa de asimilación de CO_2 y la biomasa se correlacionan con la acumulación de Ole_3-3 en las hojas de *Arabidopsis thaliana*

El total de proteínas solubles extraídas de hojas 35 días después de sembrar (DAS) de plantas del tipo salvaje (WT), D1o0-0 (transformada para expresar DGAT1 (S205A) de *A. thaliana* y Ole_0,0 como se mencionó en los Ejemplos anteriores) y tres líneas independientes de plantas D1o3-3 (mencionada como DGAT1-Ole_3,3 en los Ejemplos anteriores) fueron sometidas a SDS-PAGE e immunoblot según los Ejemplos 1, 3, 6 y 7. La transferencia fue sondeada con anticuerpos policlonales anti-o3-3 (Figura 37A). Se midió las tasas de asimilación

de CO_2 en hojas del mismo grupo de plantas por IRGA, según el Ejemplo 9 (Figura 37B) y se tomó fotos representativas de la plantas a los 28 DAS (Figura 37C). El análisis de Immunoblot de las líneas independientes D1o3-3 indica que el nivel de acumulación de o3-3 cisteína-oleosina recombinante en la hoja se correlaciona positivamente con la tasa de asimilación de CO_2 y el tamaño de los rosetones (Figura 37A-C).

REFERENCIAS

Abell *y otros,* (2004). Plant J, 37: 461-70.

Altschul *y otros,* (1997) Nucleic Acids Res. 25: 3389-3402,

Andrianov *y otros,* (2010). Plant Biotechnol J. 8(3):277-87.

Ausubel *y otros,* (1987) Current Protocols in Molecular Biology, Greene Publishing Bairoch y Bucher (1994), Nucleic Acids Res. 22, 3583

Bao *y otros*, (2000) Plant J. 22(1):39-50.

Birch (1997) Ann Rev Plant Phys Plant Mol Biol, 48, 297

Bolton y McCarthy (1962) PNAS 84:1390

Bowie *y otros,* (1990). Science 247, 1306

Bouvier-Nave *y otros,* (2000) *Eur. J. Biochem.* 267, 85-96.

Bryan *y otros,* (2007) Modification of fatty acid biosintesis. United States Patent 20070118927.

Capuano *y otros,* (2007). Biotechnol Adv. 25:203-206.

Notredame *y otros,* (2000). J. Mol. Biol. 302: 205-217

Chen *y otros,* (1999). Plant Cell Physiol, 40:1079-1086.Chiang *y otros,* (2005). J Agric Food Chem 53:4799-804.

Chiang *y otros,* (2007). Protein Expr Purif. 52:14-8.

Cookson *y otros,* (2009). Improvements in and relating to oil production. PCT/NZ2008/000085 WO/2008/130248

Dahlqvist *y otros,* (2000). Proc Natl Acad Sci U S A. 97, 6487-6492.

Deckers *y otros,* (2003). United States patent US 6582710

Demeyer y Doreau, (1999). Proc Nutr Soc. 58(3):593-607.

Deutscher (1990) Ed, Methods in Enzymology, Vol. 182, Guide to Protein Purification

Draper *y otros,* 1988, Plant Genetic Transformación and Gene Expression. A Laboratory Manual. Blackwell Sci. Pub. Oxford, p. 365

Falquet *y otros,* 2002, Nucleic Acids Res. 30, 235

Feng y Doolittle, 1987, J. Mol. Evol. 25, 351

Firkins *y otros,* (2006). J Dairy Sci. 89 Suppl 1:E31-51. Review. Greenspan.

Frandsen *y otros,* (2001). Physiologia Plantarum, 112:301-307.

Frohman (1993). Methods Enzymol. 218: 340-56

Galun y Breiman (1997). Transgenic Plants. Imperial College Press, London

Gelvin *y otros,* 1993, Plant Molecular Biol. Manual. Kluwer Acad. Pub. Dordrecht

Giesen *y otros,* Nucleic Acids Res. 1998 Nov 1;26(21):5004-6

Halford & Hardie (1998). Plant Mol Biol. 37:735-48. Review

Harada *y otros,* (2002). OLEOSIN/PHOSPHOLIPID COMPLEX AND PROCESS FOR PRODUCING THE SAME. World Patent WO/2002/026788.

Hellens *y otros,* (2000). Plant Mol Biol 42: 819-32

Hellens *y otros,* (2005). Plant Meth 1: 13

Herrera-Estrella *y otros,* (1993). Nature 303, 209

Hofmann *y otros,* (1999). Nucleic Acids Res. 27, 215

Hou *y otros,* (2003). J Dairy Sci; 86: 424-8.

Huang (1992). Ann. Rev. Plant Physiol. Plant Mol. Biol. 43:177-200.

Huang, X. (1994) Computer Applications in the Biosciences 10, 227-235

Jeanmougin *y otros,* (1998) Trends Biochem. Sci. 23, 403-5.

Jenkins y Bridges (2007). Eur. J. Lipid Sci. Technol. 109:778-789.

Jenkins y McGuire (2006). J Dairy Sci. 89(4):1302-10. Review.

Kaup *y otros,* (2002) Plant Physiol. 129(4):1616-26.

Kyte y Doolitle (1982)J. Mol. Biol. 157:105-132

Lanfranco L. (2003). Riv Biol. 96(1):31-54.

Lardizabal *y otros,* (2001). J.B.C. 276, 38862-38869.

Leprince *y otros,* (1998). Plant 204 109–119.

Lin y Tzen. (2004). Plant Physiology and Biochemistry. 42:601-608.

Lock y Bauman (2004). Lipids. 39(12):1197-206. Review.

Loer y Herman (1993). Plant Physiol. 101(3):993-998.

Mayer y Fowler (1985). J Cell Biol. 100(3):965-73.

Mekhedov *y otros,* (2000). Plant Physiol. 122(2):389-402).

Mullis *y otros,* Eds. 1994 The Polymerase Chain Reaction, Birkhauser

Murphy (1993). Prog. Lipid Res. 32:247-280.

Needleman y Wunsch, (1970) J. Mol. Biol. 48, 443-453

Nielsen *y otros,* Science. (1991) 254(5037):1497-500

Ohlrogge y Jaworski (1997). Annu Rev Plant Physiol Plant Mol Biol. 48:109-136.

Papapostolou y Howorka (2009). Mol Biosyst. 5(7):723-32. Review.

Peng (2004). Development and application of artificial sesame oil body. Ph.D. Dissertation. National ChunHsing University Graduate Institute of Biotechnology. Taichung, Taiwan.

Peng *y otros,* (2006). Stability enhancement of native and artificial oil bodies by genipin crosslink. Taiwan patent I250466.

Peng *y otros,* (2004). J Biotechnol 2004; 111: 51-7.

Potrykus y Spangenburg (1995). Gene Transfer to Plants. Springer-Verlag, Berlin

Roberts *y otros,* (2008). The Open Biotechnology Journal 2:13-21.

Roux *y otros,* (2004). J Agric Food Chem. 52(16):5245-9.

Scott *y otros,* (2007). Polyoleosins. WO2007045019.

Saha *y otros,* (2006). Plant Physiol. 141(4):1533-43.

Sambrook *y otros,* Eds, 1987, Molecular Cloning, A Laboratory Manual, 2a. Ed. Cold Spring Harbor Press

Sarmiento *y otros,* (1997). Plant J. 11(4):783-96.

Schrott (1995) In: Gene Transfer to Plants (Potrykus, T, Spangenberg. Eds) Springer Verlag. Berlín, pp. 325-336

Shimada *y otros,* (2008). Plant J. 55(5):798-809.

Shockey *y otros,* (2006). Plant Cell., 18,2294-2313.

Siloto *y otros,* (2006). Plant Cell., 18(8):1961-74.

Slack *y otros,* (1980). Biochem J. 190(3):551–561.

Slocombe *y otros,* (2009). Plant Biotechnol J. 7(7):694-703.

Stahl *y otros,* (2004). Plant Physiology, 135: 1324-1335.

Stewart *y otros,* (1969), In: Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco California

Thompson *y otros,* (1994) Nucleic Acids Research, 22:4673-4680

Ting *y otros,* (1997). J Biol Chem, 272: 3699-3706.

Tadege *y otros,* (2005). Trends Plant Sci. 10(5):229-35.

Triglia *y otros,* 1998, Nucleic Acids Res 16, 8186

Tzen *y otros,* (1992). J. Biol. Chem. 267: 15626-34

Tzen *y otros,* (2003). Adv Plant Physiol, 6: 93-104.

Tzen *y otros,* (1997). J Biochem. 121(4):762-8.

Voisey *y otros,* (1994). Plant Cell Reports 13: 309 314.

Winichayakul *y otros,* (2008). Proc. NZGA, 70:211-216

Xu *y otros,* (2005). Plant Cell. 17(11):3094-110.

Zou *y otros,* (1999). Plant J. 19, 645-653.

Zou *y otros,* (2008). Plant Biotech. J. 6(8):799-818.

Bari y otros, (2009). J. Exp. Bot. 55:623-630.

Bock & Khan (2004). Trends in Biotech. 22:311-318.

Chisti (2007). Biotech. Adv. 25:294-306.

Colman y otros, (1974). Plant Phys, 53: 395-397.

Giordano y otros, (2005). Ann. Rev. Pl. Biol. 56:99-131.

Kebeish y otros, (2007). Nature Biotech, 25:593-599.

Kozaki & Takeba (1996). Nature, 384:557-560.

Nakamura y otros, (2005). Can. J. Bot. 83:820-833.
Parry y otros, (2003). J. Exp. Bot, 54:1321-1333.
Peterhansel y otros, (2008). Fotochem. Fotobiol. 84: 1317-1323.
Tolbert (1997). Ann. Rev. Pl. Phys. Pl. Molec. Biol. 48:1-25.
Tolbert y otros, (1983). Pl. Physiol. 72:1075-1083.

RESUMEN DEL LISTADO DE SECUENCIAS

SEQ ID NO:	**TIPO**	**ESPECIE**	**COMENTARIOS**
1	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 0,0, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His tag C-terminal)
2	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 1,1, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His tag C-terminal)
3	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 1.3, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI agrega sitio de clivaje de trombina S tag N-terminal y His tag C-terminal)
4	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 3,1, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His tag C-terminal)
5	Polinucleóti	Artificial	Secuencia de nucleótidos de

	do		disulfuro de oleosina 3,3, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His tag C-terminal)
6	Polipéptido	Artificial	Secuencia de peptidos de disulfuro de oleosina 0,0, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His Tag C-terminal)
7	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 1,1, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His Tag C-terminal)
8	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 1,3, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His Tag C-terminal)
9	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 3,1, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI (agrega sitio de clivaje de trombina S tag N-terminal y His Tag C-terminal)
10	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 3,3, que fue clonada en pET29b usando sitios de restricción NdeI y XhoI ((agrega sitio de clivaje de trombina S tag N-terminal y His

			Tag C-terminal)
11	Polinucleótido	Artificial	(Secuencia de nucleótidos de disulfuro de oleosina 0,0 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.)
12	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 1,1 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.
13	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 1,3 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.
14	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 3,1 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.
15	Polinucleótido	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 3,3 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.)
16	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 0,0, que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s..

17	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 1,1 que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s
18	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 1,3, que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s
19	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 3,1, que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s
20	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 3,3 que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s
21	Polinucleóti do	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 5,6 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.
22	Polinucleóti do	Artificial	Secuencia de nucleótidos de disulfuro de oleosina 6,7 que incluye secuencia de Kozac e intrón UBQ10, que fue transformada en *Arabidopsis thaliana* bajo el control del promotor CaMV35s.
23	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 5,6, que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s

24	Polipéptido	Artificial	Secuencia de péptidos de disulfuro de oleosina 6,7, que fue transformada en *Arabidopsis thaliana*bajo el control del promotor CaMV35s
25	Polinucleótido	Artificial	Oleoaceptor (contiene terminador OCS, promotor CAMV35S, DGAT1 (S205A) de Arabidopsis e intrón UBQ10)
26	Polinucleótido	Artificial	Oleosina_0,0 y DGAT1 (S205A) en pRSH1
27	Polinucleótido	Artificial	Oleosina _1,1 y DGAT1 (S205A) en pRSH1
28	Polinucleótido	Artificial	Oleosina _1,3 y DGAT1 (S205A) en pRSH1
29	Polinucleótido	Artificial	Oleosina _3,1 y DGAT1 (S205A) en pRSH1
30	Polinucleótido	Artificial	Oleosina _3,3 y DGAT1 (S205A) en pRSH1
31	Polinucleótido	Artificial	Oleosina _5,6 y DGAT1(S205A) en pRSH1
32	Polinucleótido	Artificial	Oleosina _6,7 y DGAT1 (S205A) en pRSH1
33	Polipéptido	Artificial	DGAT1 (S205A)
34	Polinucleótido	*S. indicum*	Oleosina - AF302807
35	Polipéptido	*S. indicum*	Oleosina - AAG23840
36	Polinucleótido	*S. indicum*	Oleosina - U97700
37	Polipéptido	*S. indicum*	Oleosina - AAB58402
38	Polinucleótido	*A. thaliana*	Oleosina - X62353
39	Polipéptido	*A. thaliana*	Oleosina - CAA44225
40	Polinucleótido	*A. thaliana*	Oleosina - BT023738

41	Polipéptido	*A. thaliana*	Oleosina - AAZ23930
42	Polinucleótido	*H. annuus*	Oleosina - X62352.1
43	Polipéptido	*H. annuus*	Oleosina - CAA44224.1
44	Polinucleótido	*B. napus*	Oleosina - X82020.1
45	Polipéptido	*B. napus*	Oleosina - CAA57545.1
46	Polinucleótido	*Z. mays*	Oleosina - NM_001153560.1
47	Polipéptido	*Z. mays*	Oleosina - NP_001147032.1
48	Polinucleótido	*O. sativa*	Oleosina - L76464
49	Polipéptido	*O. sativa*	Oleosina - AAL40177.1
50	Polinucleótido	*B.oleracea*	Oleosina - AF117126.1
51	Polipéptido	*B.oleracea*	Oleosina - AAD24547.1
52	Polinucleótido	*C. arabica*	Oleosina - AY928084.1
53	Polipéptido	*C. arabica*	Oleosina - AAY14574.1
54	Polinucleótido	*S. indicum*	Esteroleosina - AF421889
55	Polipéptido	*S. indicum*	Esteroleosina - AAL13315
56	Polinucleótido	*B. napus*	Esteroleosina - EU678274
57	Polipéptido	*B. napus*	Esteroleosina - ACG69522
58	Polinucleótido	*Z. mays*	Esteroleosina - NM_001159142.1
59	Polipéptido	*Z. mays*	Esteroleosina - NP_001152614.1
60	Polinucleótido	*B. napus*	Esteroleosina - EF143915.1
61	Polipéptido	*B. napus*	Esteroleosina - ABM30178.1
62	Polinucleótido	*S. indicum*	Caleosina - AF109921
63	Polipéptido	*S. indicum*	Caleosina- AAF13743
64	Polinucleótido	*G. max*	Caleosina - AF004809
65	Polipéptido	*G. max*	Caleosina- AAB71227

66	Polinucleótido	*Z. mays*	Caleosina - NM_001158434.1
67	Polipéptido	*Z. mays*	Caleosina- NP_001151906
68	Polinucleótido	*B. napus*	Caleosina- AY966447.1
69	Polipéptido	*B. napus*	Caleosina - AAY40837.1
70	Polinucleótido	*C. revoluta*	Caleosina - FJ455154.1
71	Polipéptido	*C. revoluta*	Caleosina - ACJ70083.1
72	Polinucleótido	*C. sativus*	Caleosina- EU232173.1
73	Polipéptido	*C. sativus*	Caleosina - ABY56103.1
74	Polinucleótido	*A. thaliana*	DGAT1 - NM_127503
75	Polipéptido	*A. thaliana*	DGAT1 - NP_179535
76	Polinucleótido	*T. majus*	DGAT1 - AY084052
77	Polinucleotido	*T. majus*	DGAT1 - AAM03340
78	Polinucleótido	*Z. mays*	DGAT1 – EU039830.1
79	Polipéptido	*Z. mays*	DGAT1 - ABV91586.1
80	Polinucleótido	*A. thaliana*	DGAT2 - NM_115011
81	Polipéptido	*A. thaliana*	DGAT2 – NP_566952.1
82	Polinucleótido	*B. napus*	DGAT2 – FJ858270
83	Polipéptido	*B. napus*	DGAT2 - AC090187.1
84	Polinucleótido	*A. hypogaea*	DGAT3 (DGAT soluble) – AY875644
85	Polipéptido	*A. hypogaea*	DGAT3 (DGAT soluble) - AAX62735.1
86	Polinucleótido	*A. thaliana*	PDAT – NM_121367
87	Polipéptido	*A. thaliana*	PDAT - NP_196868.1
88	Polinucleótido	*R. communis*	PDAT – XM_002521304

89	Polipéptido	*R. communis*	PDAT - XP_002521350
90	Polipéptido	*B.oleraceae*	Oleosina - CAA65272.1
91	Polipéptido	*Z. mays*	Oleosina– NP_001147032.1
92	Polipéptido	*O. sativa*	Oleosina - AAL40177.1
93	Polipéptido	*S. indicum*	Caleosina- AAF13743
94	Polipéptido	*G. Max*	Caleosina - AAB71227
95	Polipéptido	*Z. mays*	Caleosina - NP_001151906
96	Polipéptido	*S. indicum*	Esteroleosina- AAL13315
97	Polipéptido	Brassica napus	EsteroleosinaACG69522
98	Polipéptido	*Z. mays*	EsteroleosinaNP_001152614.1
99	Polipéptido	*Brassica oleraceae*	Oleosina de polen modificada – CAA65272.1
100	Polipéptido	*Zea mays*	Oleosina modificada - NP_001147032.1
101	Polipéptido	*Oryza sativa*	Oleosina modificada - AAL40177.1
102	Polipéptido	*S. indicum*	Caoleosina modificada - AAF13743
103	Polipéptido	*G. soja*	Caoleosina modificada - AAB71227
104	Polipéptido	*Z. mays*	Caoleosina modificada – NP_001151906
105	Polipéptido	*S. indicum*	Esteroleosina modificada - AAL13315
106	Polipéptido	*Brassica napus*	Esteroleosina modificada ACG69522
107	Polipéptido	*Z. mays*	Esteroleosina modificada - NP_001152614.1

CAPÍTULO REIVINDICATORIO QUE CONTIENE LAS MODIFICACIONES HECHAS EN FASE INTERNACIONAL INDICADAS EN EL REPORTE PRELIMINAR INTERNACIONAL SOBRE PATENTABILIDAD

REIVINDICACIONES

1. Un método para producir una célula fotosintética con una tasa acrecentada de asimilación de CO_2, el método comprende por lo menos uno de los pasos:

a) modificar genéticamente la célula fotosintética para reducir o evitar el reciclado de lípidos, y

b) transformar la célula fotosintética con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

2. El método de la reivindicación 1, en donde la célula es modificada genéticamente para evitar el reciclado de lípidos, transformando la célula fotosintética con un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

3. El método de la reivindicación 1 o 2, en donde la oleosina modificada es expresada en la célula fotosintética.

4. El método de la reivindicación 3, en donde la expresión de la oleosina modificada produce la tasa acrecentada de asimilación de CO_2.

5. El método de la reivindicación 3, en donde la expresión de la oleosina modificada reduce o evita el reciclado de lípidos en la célula fotosintética.

6. El método de cualquiera de las reivindicaciones precedentes, en donde el reciclado de lípidos reducido o evitado produce el incremento de la asimilación de CO_2.

7. El método de cualquiera de las reivindicaciones precedentes, en donde la tasa de asimilación de CO_2 es acrecentada por lo menos en un 5% en relación con una planta de control.

8. El método de cualquiera de las reivindicaciones precedentes, en donde el incremento en la tasa de asimilación de CO_2 está en el intervalo de 1% a 50% en relación con una planta de control.

9. El método de cualquiera de las reivindicaciones precedentes, en donde la oleosina modificada incluye por lo menos dos cisteínas, de las cuales por lo menos una es introducida artificialmente.

10. El método de cualquiera de las reivindicaciones precedentes, en donde cada una de las oleosinas modificadas incluye:

i) por lo menos dos cisteínas introducidas artificialmente,

ii) por lo menos tres cisteínas introducidas artificialmente,

iii) por lo menos cuatro cisteínas introducidas artificialmente,

iv) por lo menos cinco cisteínas introducidas artificialmente,

v) por lo menos seis cisteínas introducidas artificialmente,

vi) por lo menos siete cisteínas introducidas artificialmente,

vii) por lo menos ocho cisteínas introducidas artificialmente,

viii) por lo menos nueve cisteínas introducidas artificialmente,

ix) por lo menos diez cisteínas introducidas artificialmente,

x) por lo menos once cisteínas introducidas artificialmente,

xi) por lo menos doce cisteínas introducidas artificialmente,

xii) por lo menos trece cisteínas introducidas artificialmente, o

xiii) por lo menos catorce cisteínas introducidas artificialmente.

11. El método de cualquiera de las reivindicaciones precedentes, en donde la oleosina modificada incluye por lo menos una cisteína en la región hidrófila N-terminal, y por lo menos una cisteína en la región hidrófila C-terminal.

12. El método de cualquiera de las reivindicaciones precedentes, en donde las cisteínas están distribuidas de una manera substancialmente uniforme entre las regiones hidrófilas N-terminal y C-terminal de la oleosina.

13. El método de cualquiera de las reivindicaciones precedentes, en donde además de la tasa de asimilación de CO_2 acrecentada, el método produce una célula fotosintética que tiene por lo menos una de las siguientes:

a) una tasa acrecentada de fotosíntesis,

b) eficiencia acrecentada en el uso del agua,

c) una tasa de crecimiento acrecentada,

d) concentración acrecentada de CO_2 en el cloroplasto,

e) una reducción en la tasa de fotorrespiración,

f) tolerancia acrecentada a altas temperaturas,

g) tolerancia acrecentada a una alta concentración de oxígeno,

h) eficiencia acrecentada en el uso del nitrógeno, e

i) reducción en la pérdida de carbono fijado.

14. El método de cualquiera de las reivindicaciones precedentes, en donde el polinucleótido está ligado operativamente a un polinucleótido promotor.

15. El método de la reivindicación 14, en donde el promotor es un promotor preferido de una célula fotosintética.

16. El método de la reivindicación 14, en donde el promotor es un promotor regulado por la luz.

17. El método de cualquiera de las reivindicaciones precedentes, en donde el polinucleótido es parte de una construcción genética.

18. El método de cualquiera de las reivindicaciones precedentes en donde la célula fotosintética es modificada también genéticamente para que exprese una enzima sintetizadora de triacilglicerol (TAG).

19. El método de la reivindicación 18, donde la célula fotosintética es modificada genéticamente para que comprenda una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG).

20. El método de la reivindicación 18, donde la célula fotosintética comprende una construcción genética que incluye una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG) y en la cual el ácido nucleico está ligado operativamente a un polinucleótido promotor.

21. El método de la reivindicación 20, donde el promotor ligado operativamente al ácido nucleico es un promotor preferido de una célula fotosintética.

22. El método de la reivindicación 20, donde el promotor ligado operativamente al ácido nucleico es un promotor regulado por la luz.

23. El método de cualquiera de las reivindicaciones precedentes, donde la célula fotosintética es seleccionada de una célula fotosintética bacteriana, una célula fotosintética de levadura, una célula fotosintética de hongo, una célula fotosintética de alga, y una célula vegetal.

24. El método de cualquiera de las reivindicaciones precedentes, donde la célula fotosintética es una célula de alga seleccionada de una de las siguientes divisiones: Chlorophyta (algas verdes), Rhodophyta (algas rojas), Phaeophyceae (algas marrón), Bacillariophycaeae (diatomeas), y Dinoflagellata (dinoflageladas).

25. El método de la reivindicación 24, donde la célula de alga muestra una tasa de crecimiento acrecentada, en relación con una célula de alga de control, con una concentración elevada de oxígeno (O_2).

26. El método de la reivindicación 25, donde la concentración de O_2 es elevada por lo menos a 1,1 veces la saturación del aire.

27. El método de la reivindicación 25 o 26, donde la tasa de crecimiento acrecentada de la célula de alga es por lo menos 10% mayor que la tasa de crecimiento de una célula de alga de control a la misma concentración de O_2.

28. El método de cualquiera de las reivindicaciones 1 a 21, donde la célula fotosintética es una célula vegetal.

29. El método de la reivindicación 28, donde la célula vegetal es parte de una planta.

30. El método de la reivindicación 29 o 30, en donde el promotor ligado operativamente al polinucleótido o al ácido nucleico es un promotor preferido de tejido fotosintético.

31. El método de la reivindicación 29 o 30, en donde el promotor ligado operativamente al polinucleótido o al ácido nucleico es un promotor regulado por la luz.

32. El método de cualquiera de las reivindicaciones 29 a 31, en donde, además de la tasa acrecentada de asimilación de CO_2 la planta tiene también por lo menos uno de:

a) incremento en la tasa de fotosíntesis,

b) incremento en la eficiencia de uso del agua,

c) un incremento en la tasa de crecimiento

d) incremento de biomasa,

e) retardo de la floración,

f) incremento de la concentración de CO_2 en el cloroplasto,

g) una reducción en la tasa de fotorrespiración,

h) incremento del rendimiento de semillas, frutos u órganos de almacenamiento,

i) incremento en la tolerancia a la sequía,

j) incremento en la tolerancia a alta temperatura,

k) incremento en la tolerancia a una elevada concentración de oxígeno,

l) incremento en la eficiencia en el uso del nitrógeno, y

m) reducción en la pérdida de carbono fijado.

33. El método de la reivindicación 32, en donde la planta tiene todos los puntos de a) a m).

34. El método de cualquiera de las reivindicaciones 29 a 33, en donde la planta es una planta C3.

35. Un método para producir aceite, el método comprende los pasos de:

a) proveer una planta que comprende un polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente,

b) cultivar la planta para que exprese la oleosina modificada y produzca aceite en sus tejidos/órganos no fotosintéticos,

en donde la planta acumula por lo menos 150% más lípido total en sus tejidos/órganos no fotosintéticos que una planta de control.

36. El método de la reivindicación 35, donde la planta acumula más aceite en sus tejidos/órganos no fotosintéticos que una planta de control.

37. El método de la reivindicación 35, donde la planta acumula por lo menos 1,2x más aceite en sus tejidos/órganos no fotosintéticos que una planta de control.

38. El método de cualquiera de las reivindicaciones 35 a 37, donde el aceite es producido a un nivel más alto en sus tejidos/órganos no fotosintéticos que en una planta de control.

39. El método de cualquiera de las reivindicaciones 35 a 38, donde el incremento en el nivel de producción de aceite en los tejidos/órganos no fotosintéticos es producido por la expresión de la oleosina modificada en los tejidos/órganos.

40. El método de cualquiera de las reivindicaciones 35 a 39, donde el tejido/órgano no fotosintético es seleccionado de tejidos/órganos de la planta que están bajo tierra.

41. El método de la reivindicación 40, donde el tejido/órgano bajo tierra es seleccionado de raíz, tubérculo, bulbo, tallo bulboso y rizoma.

42. El método de cualquiera de las reivindicaciones 35 a 39, donde el tejido/órgano no fotosintético es seleccionado de raíz, tubérculo, bulbo, tallo bulboso, rizoma y endosperma.

43. El método de cualquiera de las reivindicaciones 35 a 39, donde el tejido/órgano fotosintético es raíz.

44. El método de cualquiera de las reivindicaciones 35 a 43, que incluye el paso de transformar la planta con el polinucleótido que codifica una oleosina modificada que incluye por lo menos una cisteína introducida artificialmente.

45. El método de cualquiera de las reivindicaciones 35 a 44, donde el polinucleótido está ligado operativamente a un polinucleótido promotor.

46. El método de la reivindicación 45, donde el promotor es un promotor preferido de tejido no fotosintético.

47. El método la reivindicación 46, donde el promotor es seleccionado de un promotor preferido de raíz, un promotor preferido de tubérculo, un promotor preferido de bulbo, un promotor preferido de tallo bulboso, un promotor preferido de rizoma y un promotor preferido de endosperma.

48. El método de cualquiera de las reivindicaciones 35 a 47, donde la parte del polinucleótido de una construcción genética o de una construcción de expresión que incluye el polinucleótido está ligada operativamente al promotor.

49. El método de cualquiera de las reivindicaciones 35 a 48, donde la planta es modificada también genéticamente para que exprese una enzima sintetizadora de triacilglicerol (TAG).

50. El método de cualquiera de las reivindicaciones 35 a 49, donde la planta es modificada genéticamente para que comprenda una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG).

51. El método de cualquiera de las reivindicaciones 35 a 50, donde la planta comprende una construcción de expresión que incluye una secuencia de ácido nucleico que codifica una enzima sintetizadora de triacilglicerol (TAG), y en el cual el ácido nucleico está ligado operativamente a un polinucleótido promotor.

52. El método de la reivindicación 51, donde el promotor ligado operativamente a un polinucleótido promotor es un promotor preferido de tejido no fotosintético.

53. El método de la reivindicación 52, donde el promotor ligado operativamente a un polinucleótido promotor es seleccionado de un promotor preferido de raíz, un promotor preferido de tubérculo, un promotor preferido de bulbo, un promotor preferido de tallo bulboso, un promotor preferido de rizoma y un promotor preferido de endosperma.

54. El método de cualquiera de las reivindicaciones 35 a 53, que comprende el paso adicional de procesar el tejido/órgano no fotosintético de la planta para formar un material de alimentación para animales.

55. El método de cualquiera de las reivindicaciones 35 a 53, que comprende el paso adicional de procesar el tejido/órgano no fotosintético de la planta para formar una materia prima para biocombustible.

56. El método de cualquiera de las reivindicaciones 35 a 53, **CARACTERIZADO** porque comprende el paso adicional de extraer aceite del tejido/órgano no fotosintético de la planta.

57. El método de cualquiera de las reivindicaciones 5 a 53, que comprende el paso adicional de procesar el tejido/órgano no fotosintético para formar una fracción de aceite.

58. El método de cualquiera de las reivindicaciones 5 a 53, donde el aceite del tejido/órgano no fotosintético es procesado para formar un combustible, un aceite oleoquímico, nutricional o cosmético, un ácido graso poliinsaturado (PUFA) o una combinación de estos.

59. Un tejido/órgano no fotosintético de una planta, que es producido por un método de cualquiera de las reivindicaciones 35 a 58.

60. El tejido/órgano no fotosintético de la reivindicación 56, que es seleccionado de un tejido/órgano de la planta que está bajo tierra.

61. El tejido/órgano no fotosintético de la reivindicación 61, que es seleccionado de una raíz, tubérculo, bulbo, tallo bulboso y rizoma.

62. El tejido/órgano no fotosintético de la reivindicación 59, que es seleccionado de una raíz, tubérculo, bulbo, tallo bulboso, rizoma y endosperma.

63. El tejido/órgano no fotosintético de la reivindicación 59, que es una raíz.

64. El tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 63, que tiene más lípido total que el tejido/órgano no fotosintético correspondiente de una planta de control.

65. El tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 63, que tiene por lo menos un 10% más de lípido total en sus tejidos/u órganos no fotosintético que una planta de control.

66. El tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 65, que tiene más aceite que el correspondiente tejido/órgano no fotosintético de una planta de control.

67. El tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 65, que tiene por lo menos 1,2x más aceite que el correspondiente tejido/órgano no fotosintético de una planta de control.

68. Un alimento para animales que comprende el tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 67.

69. Una materia prima para biocombustible que comprende el tejido/órgano no fotosintético de cualquiera de las reivindicaciones 61 a 67.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISIÓN DE NUEVAS CREACIONES
EXTRACTO PARA PUBLICACIÓN
SOLICITUD DE PATENTE DE INVENCIÓN PCT

(21) No. de solicitud: (22) Fecha de solicitud: 05-03-14 (30) Prioridad (31) ***No. Prioridad*** 61/515,610 (32) ***Fecha*** 05-08-11 (33) ***País*** US	(51) Int Cl: **C12N 15/29** (71) Solicitante(s): **AGRESEARCH LIMITED** (72) Inventor(es): ROBERTS, Nicholas John; SCOTT, Richard William; WINICHAYAKUL, Somrutai y ROLDAN, Marissa. (74) Apoderado: **ALICIA LLOREDA RICAURTE**
(85) Fecha límite inicio fase nacional: 05-03-14 (86) Datos relativos a la presentación de la solicitud PCT Fecha de presentación de la solicitud: *03-08-12* No. de solicitud: ***PCT/NZ2012/000138***	(87) Publicación internacional Fecha: ***14-02-13*** No. publicación: ***WO 2013/022353 A1***

(54) Titulo:

METODO PARA INCREMENTAR LA ASIMILACIÓN DE CO_2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS

Resumen - Reivindicación(es):

Resumen.-

El invento provee métodos para acrecentar la tasa de asimilación de CO_2 en células y plantas fotosintéticas. El invento comprende reducir o evitar el reciclado de lípidos y/o expresar oleosinas modificadas con residuos de cisteína introducidos artificialmente en las células y plantas fotosintéticas. El invento provee también métodos para acrecentar producción de aceite en plantas, mediante la expresión de oleosinas modificadas con residuos de cisteína introducidos artificialmente en los tejidos/órganos no fotosintéticos de las plantas. El método incluye también opcionalmente el paso de extraer el aceite de los tejidos/órganos no fotosintéticos de la planta, o procesar los tejidos/órganos no fotosintéticos ricos en aceite para formar alimentos para animales o materia prima de biocombustibles.

Sector: Ingeniería química y de petróleos.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISIÓN DE NUEVAS CREACIONES
TARJETA ARCHIVO TEMÁTICO

PATENTE DE INVENCION ____ PCT X MODELO DE UTILIDAD ____ DISEÑO INDUSTRIAL ___

(21) No. de solicitud: (22) Fecha de solicitud: 05-03-14
(51) Clasificación Internacional: C12N 15/29

(71) Solicitante(s): **AGRESEARCH LIMITED**

(72) Inventor(es): ROBERTS, Nicholas John; SCOTT, Richard William; WINICHAYAKUL, Somrutai y ROLDAN, Marissa.

(74) Apoderado: **ALICIA LLOREDA RICAURTE**

(30) Prioridad	(31) ***No. Prioridad***	(32) ***Fecha***	(33) ***País***
	61/515,610	05-08-11	US

(85) Fecha límite inicio fase nacional: 05-03-14
(86) Datos relativos a la presentación de la solic. PCT
Fecha de presentación de la solicitud: *03-08-12*
No. de solicitud: ***PCT/NZ2012/000138***

(87) Publicación internacional:
Fecha: ***14-02-13***
No. Publicación: ***WO 2013/022353 A1***

(54) Titulo: METODO PARA INCREMENTAR LA ASIMILACIÓN DE CO_2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS.

(57) Resumen - Reivindicación(es):

Resumen.-

El invento provee métodos para acrecentar la tasa de asimilación de CO_2 en células y plantas fotosintéticas. El invento comprende reducir o evitar el reciclado de lípidos y/o expresar oleosinas modificadas con residuos de cisteína introducidos artificialmente en las células y plantas fotosintéticas. El invento provee también métodos para acrecentar producción de aceite en plantas, mediante la expresión de oleosinas modificadas con residuos de cisteína introducidos artificialmente en los tejidos/órganos no fotosintéticos de las plantas. El método incluye también opcionalmente el paso de extraer el aceite de los tejidos/órganos no fotosintéticos de la planta, o procesar los tejidos/órganos no fotosintéticos ricos en aceite para formar alimentos para animales o materia prima de biocombustibles.

Sector: Ingeniería química y de petróleos.

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
DIVISIÓN DE NUEVAS CREACIONES

TARJETA ARCHIVO PROPIETARIOS

PATENTE DE INVENCION ____ PCT __X__ MODELO DE UTILIDAD ____ DISEÑO INDUSTRIAL __

(21) No. de solicitud: (22) Fecha de solicitud: 05-03-14
(51) Clasificación Internacional: C12N 15/29

(71) Solicitante(s): **AGRESEARCH LIMITED**

(72) Inventor (es): ROBERTS, Nicholas John; SCOTT, Richard William; WINICHAYAKUL, Somrutai y ROLDAN, Marissa.

(74) Apoderado: **ALICIA LLOREDA RICAURTE**

(30) Prioridad	(31) ***No. Prioridad***	(32) ***Fecha***	(33) ***País***
	61/515,610	05-08-11	US

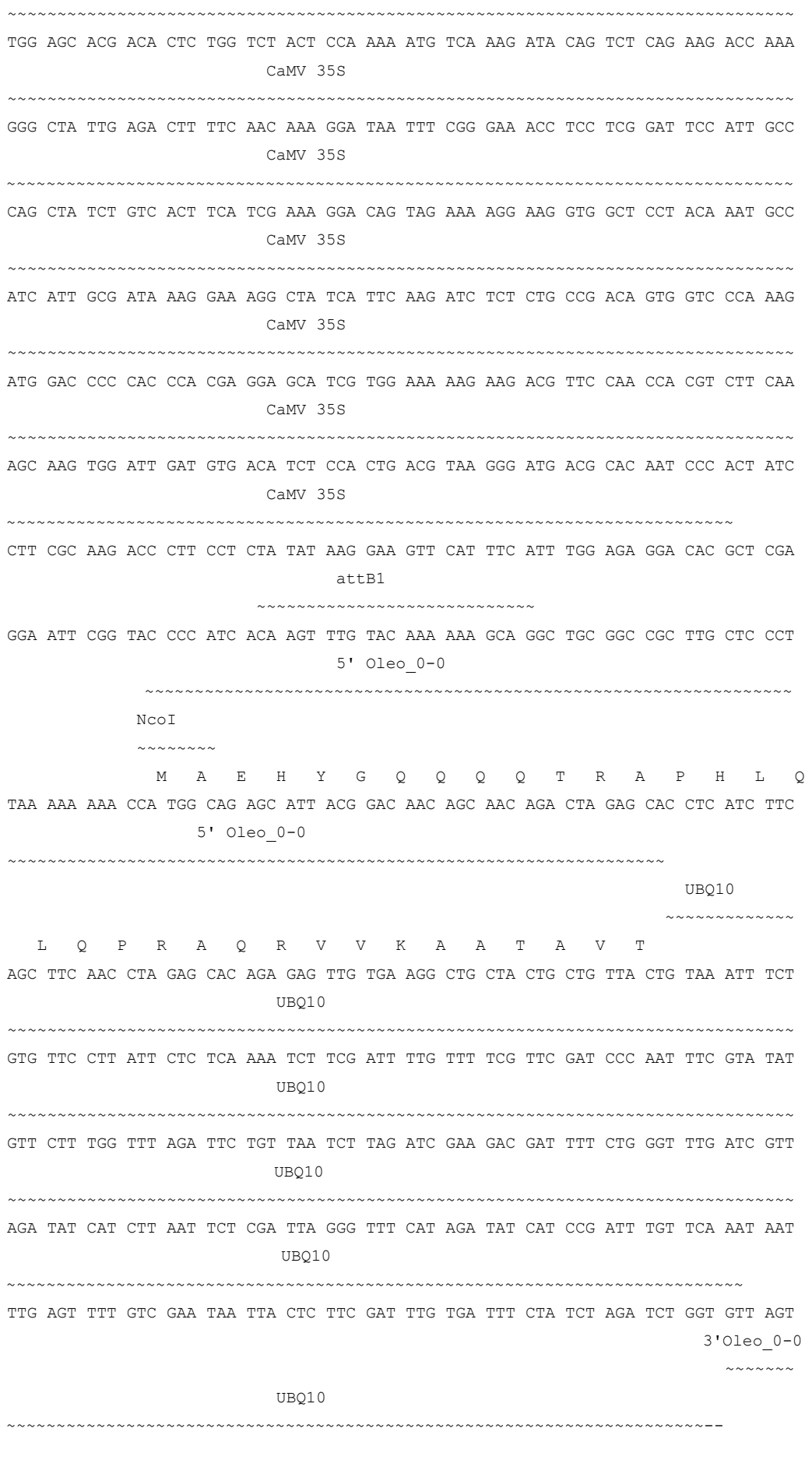
(85) Fecha límite inicio fase nacional: 05-03-14
(86) Datos relativos a la presentación de la solicitud PCT
Fecha de presentación de la solicitud: *03-08-12*
No. de solicitud: ***PCT/NZ2012/000138***

(87) Publicación internacional
Fecha: ***14-02-13***
No. publicación: ***WO 2013/022353 A1***

(54) Titulo: **METODO PARA INCREMENTAR LA ASIMILACIÓN DE CO_2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS.**

Figura 1

```
                                                      CaMV 35S
                                               ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                               T CGA CGA ATT AAT TCC AAT CCC ACA AAA ATC
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC TCA TAT
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG TTG TAC
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~--
AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT TAC AGA
                                 CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA CTT CAT
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG CCC ACC
                                 CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC AGC CCC
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA CGA TCT
                                CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA GAA GGT
                                  CaMV 35S
-~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA CGC AGC
                                  CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT TCC CAA
                                  CaMV 35
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA GAA AGA
                                  CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT TCA TAA
                                  CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC TAA GGC
                                  CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA CTG GCG
                                  CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA ACA TGG
                                  CaMV 35S
```


TGG AGC ACG ACA CTC TGG TCT ACT CCA AAA ATG TCA AAG ATA CAG TCT CAG AAG ACC AAA
CaMV 35S
GGG CTA TTG AGA CTT TTC AAC AAA GGA TAA TTT CGG GAA ACC TCC TCG GAT TCC ATT GCC
CaMV 35S
CAG CTA TCT GTC ACT TCA TCG AAA GGA CAG TAG AAA AGG AAG GTG GCT CCT ACA AAT GCC
CaMV 35S
ATC ATT GCG ATA AAG GAA AGG CTA TCA TTC AAG ATC TCT CTG CCG ACA GTG GTC CCA AAG
CaMV 35S
ATG GAC CCC CAC CCA CGA GGA GCA TCG TGG AAA AAG AAG ACG TTC CAA CCA CGT CTT CAA
CaMV 35S
AGC AAG TGG ATT GAT GTG ACA TCT CCA CTG ACG TAA GGG ATG ACG CAC AAT CCC ACT ATC
CaMV 35S
CTT CGC AAG ACC CTT CCT CTA TAT AAG GAA GTT CAT TTC ATT TGG AGA GGA CAC GCT CGA
attB1
GGA ATT CGG TAC CCC ATC ACA AGT TTG TAC AAA AAA GCA GGC TGC GGC CGC TTG CTC CCT
5' Oleo_0-0
NcoI
M A E H Y G Q Q Q Q T R A P H L Q
TAA AAA AAA CCA TGG CAG AGC ATT ACG GAC AAC AGC AAC AGA CTA GAG CAC CTC ATC TTC
5' Oleo_0-0
UBQ10
L Q P R A Q R V V K A A T A V T
AGC TTC AAC CTA GAG CAC AGA GAG TTG TGA AGG CTG CTA CTG CTG TTA CTG TAA ATT TCT
UBQ10
GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT CCC AAT TTC GTA TAT
UBQ10
GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT CTG GGT TTG ATC GTT
UBQ10
AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG ATT TGT TCA AAT AAT
UBQ10
TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT AGA TCT GGT GTT AGT
3'Oleo_0-0
UBQ10

```
                                                                                A   G
TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT GAT TAA CAG GCT GGA
                          3' Oleo_0-0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 G   S   L   L   V   L   S   G   L   T   L   A   G   T   V   I   A   L   T   I
GGA TCT CTT CTT GTT CTC TCT GGA CTT ACT CTC GCT GGA ACT GTT ATC GCT CTC ACT ATC
                          3' Oleo_0-0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   T   P   L   L   V   I   F   S   P   V   L   V   P   A   V   I   T   I   F
GCT ACA CCT CTT CTC GTT ATC TTC TCT CCT GTT CTC GTT CCT GCT GTG ATC ACT ATC TTC
                          3' Oleo_0-0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 L   L   G   A   G   F   L   A   S   G   G   F   G   V   A   A   L   S   V   L
CTT CTC GGA GCT GGA TTT CTT GCT TCT GGT GGA TTT GGA GTT GCT GCT CTC TCT GTT CTC
                          3' Oleo_0-0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 S   W   I   Y   R   Y   L   T   G   K   H   P   P   G   A   D   Q   L   E   S
TCT TGG ATC TAC AGA TAC CTC ACT GGA AAA CAT CCT CCA GGT GCT GAT CAA CTT GAG TCT
                          3' Oleo_0-0
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   K   T   K   L   A   S   K   A   R   E   M   K   D   R   A   E   Q   F   S
GCT AAG ACT AAG CTC GCT TCT AAG GCT AGA GAG ATG AAG GAT AGA GCA GAG CAA TTC TCT
                3' Oleo_0-0                                   terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 Q   Q   P   V   A   G   S   Q   T   S   *   *
CAA CAG CCT GTT GCT GGA TCT CAG ACT TCT TAA TGA ACA TAT GGT CCT GCT TTA ATG AGA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT GCG AGA CGC CTA TGA TCG CAT GAT ATT TGC TTT CAA TTC TGT TGT GCA CGT TGT AAA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA CCT GAG CAT GTG TAG CTC AGA TCC TTA CCG CCG GTT TCG GTT CAT TCT AAT GAA TAT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATC ACC CGT TAC TAT CGT ATT TTT ATG AAT AAT ATT CTC CGT TCA ATT TAC TGA TTG TAC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCT ACT ACT TAT ATG TAC AAT ATT AAA ATG AAA ACA ATA TAT TGT GCT GAA TAG GTT TAT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGC GAC ATC TAT GAT AGA GCG CCA CAA TAA CAA ACA ATT GCG TTT TAT TAT TAC AAA TCC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAT TTT AAA AAA AGC GGC AGA ACC GGT CAA ACC TAA AAG ACT GAT TAC ATA AAT CTT ATT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA ATT TCA AAA GGC CCC AGG GGC TAG TAT CTA CGA CAC ACC GAG CGG CGA ACT AAT AAC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GTT CAC TGA AGG GAA CTC CGG TTC CCC GCC GGC GCG CAT GGG TGA GAT TCC TTG AAG TTG
                          terminador OCS
```

~~~~~~

AGT ATT GGC CGT CCG CTC TAC CGA AAG TTA CGG GCA CCA TTC AAC CCG GTC CAG CAC GGC

terminador OCS

~~~~~~

GGC CGG GTA ACC GAC TTG CTG CCC CGA GAA TTA TGC AGC ATT TTT TTG GTG TAT GTG GGC

terminador OCS

~~~~~~

CCC AAA TGA AGT GCA GGT CAA ACC TTG ACA GTG ACG ACA AAT CGT TGG GCG GGT CCA GGG

terminador OCS

~~~~~~

CGA ATT TTG CGA CAA CAT GTC GAG GCT CAG CAG GAC CTG CAG GCA TGC AAG CTA GCT TAC

terminador OCS

~~~~~~

CaMV35S

-~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

TAG TGA TGC ATA TTC TAT AGT GTC ACC TAA ATC TTC GAC GAA TTA ATT CCA ATC CCA CAA

CaMV35S

~~~~~~

AAA TCT GAG CTT AAC AGC ACA GTT GCT CCT CTC AGA GCA GAA TCG GGT ATT CAA CAC CCT

CaMV35S

~~~~~~

CAT ATC AAC TAC TAC GTT GTG TAT AAC GGT CCA CAT GCC GGT ATA TAC GAT GAC TGG GGT

CaMV35S

~~~~~~

TGT ACA AAG GCG GCA ACA AAC GGC GTT CCC GGA GTT GCA CAC AAG AAA TTT GCC ACT ATT

CaMV35S

~~~~~~

ACA GAG GCA AGA GCA GCA GCT GAC GCG TAC ACA ACA AGT CAG CAA ACA GAC AGG TTG AAC

CaMV35S

~~~~~~

TTC ATC CCC AAA GGA GAA GCT CAA CTC AAG CCC AAG AGC TTT GCT AAG GCC CTA ACA AGC

CaMV35S

~~~~~~

CCA CCA AAG CAA AAA GCC CAC TGG CTC ACG CTA GGA ACC AAA AGG CCC AGC AGT GAT CCA

CaMV35S

~~~~~~

GCC CCA AAA GAG ATC TCC TTT GCC CCG GAG ATT ACA ATG GAC GAT TTC CTC TAT CTT TAC

CaMV35S

~~~~~~

GAT CTA GGA AGG AAG TTC GAA GGT GAA GGT GAC GAC ACT ATG TTC ACC ACT GAT AAT GAG

CaMV35S

~~~~~~

AAG GTT AGC CTC TTC AAT TTC AGA AAG AAT GCT GAC CCA CAG ATG GTT AGA GAG GCC TAC

CaMV35S

~~~~~~

GCA GCA GGT CTC ATC AAG ACG ATC TAC CCG AGT AAC AAT CTC CAG GAG ATC AAA TAC CTT

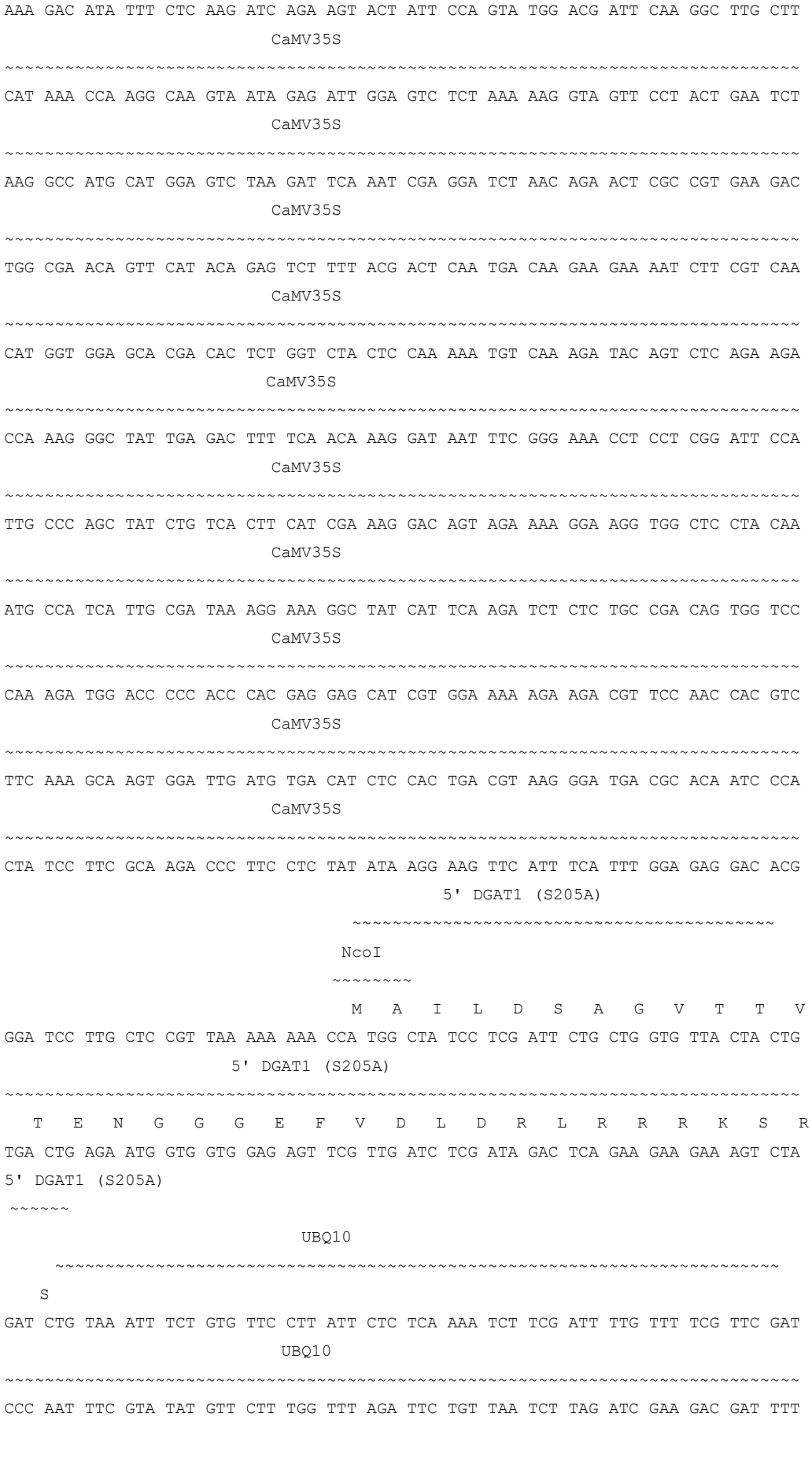
CaMV35S

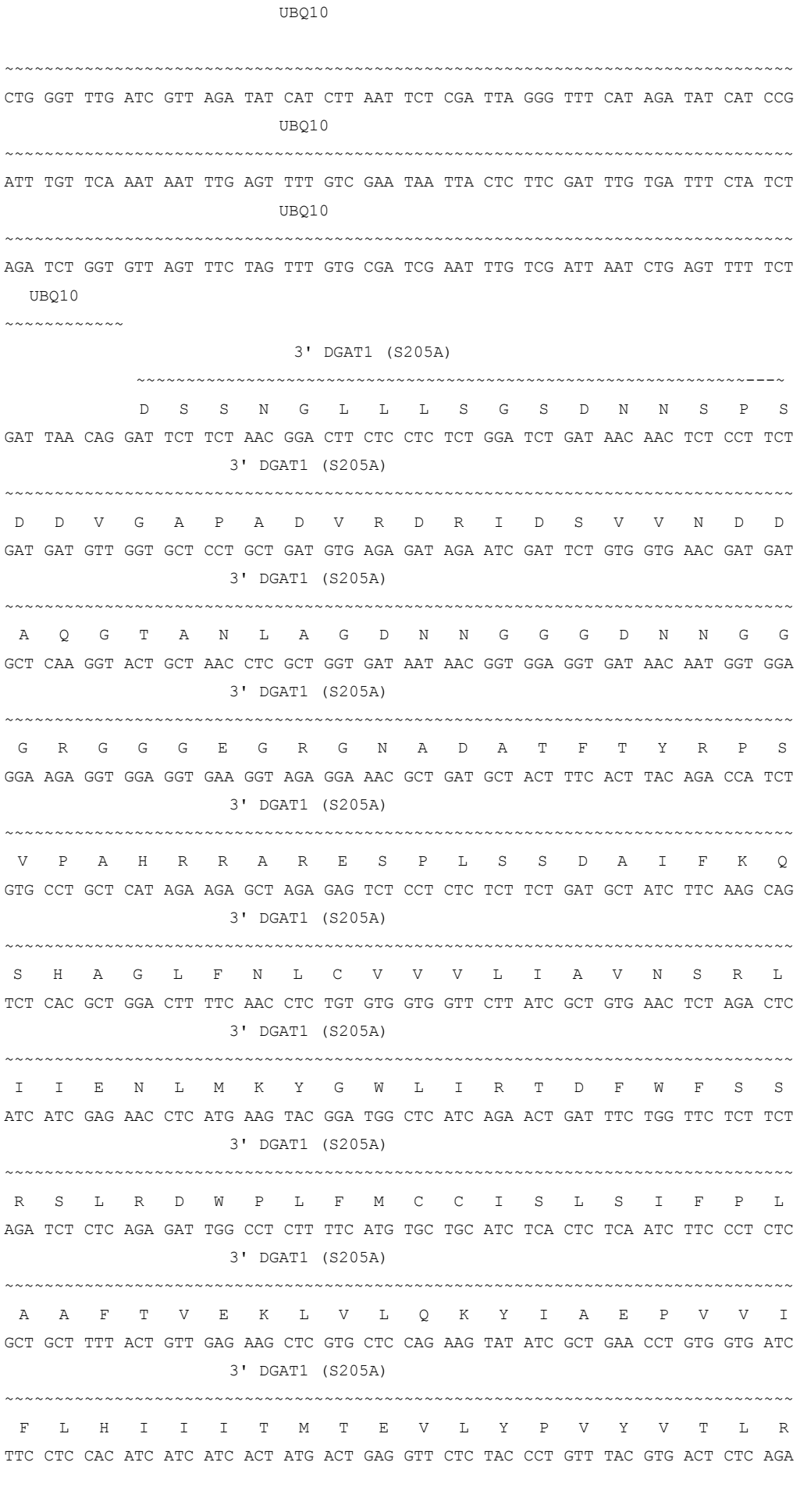
~~~~~~

CCC AAG AAG GTT AAA GAT GCA GTC AAA AGA TTC AGG ACT AAT TGC ATC AAG AAC ACA GAG

CaMV35S

~~~~~~
~~~~~~


AAA GAC ATA TTT CTC AAG ATC AGA AGT ACT ATT CCA GTA TGG ACG ATT CAA GGC TTG CTT
CaMV35S
CAT AAA CCA AGG CAA GTA ATA GAG ATT GGA GTC TCT AAA AAG GTA GTT CCT ACT GAA TCT
CaMV35S
AAG GCC ATG CAT GGA GTC TAA GAT TCA AAT CGA GGA TCT AAC AGA ACT CGC CGT GAA GAC
CaMV35S
TGG CGA ACA GTT CAT ACA GAG TCT TTT ACG ACT CAA TGA CAA GAA GAA AAT CTT CGT CAA
CaMV35S
CAT GGT GGA GCA CGA CAC TCT GGT CTA CTC CAA AAA TGT CAA AGA TAC AGT CTC AGA AGA
CaMV35S
CCA AAG GGC TAT TGA GAC TTT TCA ACA AAG GAT AAT TTC GGG AAA CCT CCT CGG ATT CCA
CaMV35S
TTG CCC AGC TAT CTG TCA CTT CAT CGA AAG GAC AGT AGA AAA GGA AGG TGG CTC CTA CAA
CaMV35S
ATG CCA TCA TTG CGA TAA AGG AAA GGC TAT CAT TCA AGA TCT CTC TGC CGA CAG TGG TCC
CaMV35S
CAA AGA TGG ACC CCC ACC CAC GAG GAG CAT CGT GGA AAA AGA AGA CGT TCC AAC CAC GTC
CaMV35S
TTC AAA GCA AGT GGA TTG ATG TGA CAT CTC CAC TGA CGT AAG GGA TGA CGC ACA ATC CCA
CaMV35S
CTA TCC TTC GCA AGA CCC TTC CTC TAT ATA AGG AAG TTC ATT TCA TTT GGA GAG GAC ACG
5' DGAT1 (S205A)
NcoI
M A I L D S A G V T T V
GGA TCC TTG CTC CGT TAA AAA AAA CCA TGG CTA TCC TCG ATT CTG CTG GTG TTA CTA CTG
5' DGAT1 (S205A)
T E N G G G E F V D L D R L R R R K S R
TGA CTG AGA ATG GTG GTG GAG AGT TCG TTG ATC TCG ATA GAC TCA GAA GAA GAA AGT CTA
5' DGAT1 (S205A)
UBQ10
S
GAT CTG TAA ATT TCT GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT
UBQ10
CCC AAT TTC GTA TAT GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT


UBQ10
CTG GGT TTG ATC GTT AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG
UBQ10
ATT TGT TCA AAT AAT TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT
UBQ10
AGA TCT GGT GTT AGT TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT
UBQ10
3' DGAT1 (S205A)
D S S N G L L L S G S D N N S P S
GAT TAA CAG GAT TCT TCT AAC GGA CTT CTC CTC TCT GGA TCT GAT AAC AAC TCT CCT TCT
3' DGAT1 (S205A)
D D V G A P A D V R D R I D S V V N D D
GAT GAT GTT GGT GCT CCT GCT GAT GTG AGA GAT AGA ATC GAT TCT GTG GTG AAC GAT GAT
3' DGAT1 (S205A)
A Q G T A N L A G D N N G G G D N N G G
GCT CAA GGT ACT GCT AAC CTC GCT GGT GAT AAT AAC GGT GGA GGT GAT AAC AAT GGT GGA
3' DGAT1 (S205A)
G R G G G E G R G N A D A T F T Y R P S
GGA AGA GGT GGA GGT GAA GGT AGA GGA AAC GCT GAT GCT ACT TTC ACT TAC AGA CCA TCT
3' DGAT1 (S205A)
V P A H R R A R E S P L S S D A I F K Q
GTG CCT GCT CAT AGA AGA GCT AGA GAG TCT CCT CTC TCT TCT GAT GCT ATC TTC AAG CAG
3' DGAT1 (S205A)
S H A G L F N L C V V V L I A V N S R L
TCT CAC GCT GGA CTT TTC AAC CTC TGT GTG GTG GTT CTT ATC GCT GTG AAC TCT AGA CTC
3' DGAT1 (S205A)
I I E N L M K Y G W L I R T D F W F S S
ATC ATC GAG AAC CTC ATG AAG TAC GGA TGG CTC ATC AGA ACT GAT TTC TGG TTC TCT TCT
3' DGAT1 (S205A)
R S L R D W P L F M C C I S L S I F P L
AGA TCT CTC AGA GAT TGG CCT CTT TTC ATG TGC TGC ATC TCA CTC TCA ATC TTC CCT CTC
3' DGAT1 (S205A)
A A F T V E K L V L Q K Y I A E P V V I
GCT GCT TTT ACT GTT GAG AAG CTC GTG CTC CAG AAG TAT ATC GCT GAA CCT GTG GTG ATC
3' DGAT1 (S205A)
F L H I I I T M T E V L Y P V Y V T L R
TTC CTC CAC ATC ATC ATC ACT ATG ACT GAG GTT CTC TAC CCT GTT TAC GTG ACT CTC AGA

3' DGAT1 (S205A)

~~~~~~

C D S A F L S G V T L M L L T C I V W L

TGC GAT TCT GCT TTC CTC TCT GGT GTT ACT CTT ATG CTC CTC ACT TGC ATT GTG TGG CTT

3' DGAT1 (S205A)

~~~~~~

K L V S Y A H T S Y D I R S L A N A A D

AAG CTC GTG TCT TAC GCT CAC ACT TCT TAC GAT ATC AGA TCT CTC GCT AAC GCT GCT GAT

3' DGAT1 (S205A)

~~~~~~

K A N P E V S Y Y V S L K S L A Y F M V

AAG GCT AAC CCT GAA GTG TCT TAC TAC GTG TCT CTC AAG TCT CTC GCT TAC TTC ATG GTT

3' DGAT1 (S205A)

~~~~~~

A P T L C Y Q P S Y P R S A C I R K G W

GCT CCT ACA CTT TGT TAC CAG CCA TCT TAC CCT AGA TCT GCT TGC ATT AGA AAG GGA TGG

3' DGAT1 (S205A)

~~~~~~

V A R Q F A K L V I F T G F M G F I I E

GTG GCA AGA CAA TTC GCT AAG TTG GTG ATC TTC ACT GGA TTC ATG GGA TTC ATC ATC GAG

3' DGAT1 (S205A)

~~~~~~

Q Y I N P I V R N S K H P L K G D L L Y

CAG TAC ATC AAC CCT ATT GTG AGA AAC TCT AAG CAC CCT CTC AAG GGT GAT CTT CTC TAC

3' DGAT1 (S205A)

~~~~~~

A I E R V L K L S V P N L Y V W L C M F

GCT ATC GAG AGA GTT CTT AAG CTC TCT GTG CCT AAC CTT TAT GTG TGG CTC TGC ATG TTC

3' DGAT1 (S205A)

~~~~~~

Y C F F H L W L N I L A E L L C F G D R

TAC TGT TTC TTC CAC CTC TGG CTT AAC ATC CTT GCT GAG TTG CTT TGC TTC GGA GAT AGA

3' DGAT1 (S205A)

~~~~~~

E F Y K D W W N A K S V G D Y W R M W N

GAG TTC TAC AAG GAT TGG TGG AAC GCT AAG TCT GTT GGA GAT TAT TGG AGA ATG TGG AAC

3' DGAT1 (S205A)

~~~~~~

M P V H K W M V R H I Y F P C L R S K I

ATG CCT GTG CAT AAG TGG ATG GTG CGT CAC ATC TAC TTC CCT TGC CTC AGA TCT AAG ATC

3' DGAT1 (S205A)

~~~~~~

P K T L A I I I A F L V S A V F H E L C

CCT AAG ACT CTC GCT ATC ATT ATC GCT TTC CTC GTG TCT GCT GTT TTC CAT GAG TTG TGT

3' DGAT1 (S205A)

~~~~~~

I A V P C R L F K L W A F L G I M F Q V

ATC GCT GTT CCT TGC AGA CTT TTC AAG CTT TGG GCT TTC CTC GGA ATC ATG TTC CAG GTT

3' DGAT1 (S205A)

~~~~~~

P L V F I T N Y L Q E R F G S T V G N M
~~~~~~

```
CCA CTC GTG TTC ATC ACT AAC TAC CTC CAA GAG AGA TTC GGA TCT ACT GTT GGA AAC ATG
                      3' DGAT1 (S205A)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
  I   F   W   F   I   F   C   I   F   G   Q   P   M   C   V   L   L   Y   Y   H
ATT TTC TGG TTC ATT TTC TGC ATC TTC GGA CAG CCT ATG TGC GTT CTC CTC TAC TAC CAC
          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 D   L   M   N   R   K   G   S   M   S   *   *
GAT CTC ATG AAC AGA AAG GGA TCT ATG TCT TAA TGA AGG ATC CAC CCA GCT TTC TTG TAC
                                                                  attB2
                                                                 ocs
                                                               ~~~~~~~~~~~~~~~~~~~~~~
AAA GTG GTG ATG GGT TCG AAA TCG ATA AGC TTG GAT CCT CTA GAG TCC TGC TTT AAT GAG
~~~~~~~~~
 attB2
                              ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATA TGC GAG ACG CCT ATG ATC GCA TGA TAT TTG CTT TCA ATT CTG TTG TGC ACG TTG TAA
                              ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA ACC TGA GCA TGT GTA GCT CAG ATC CTT ACC GCC GGT TTC GGT TCA TTC TAA TGA ATA
                              ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT CAC CCG TTA CTA TCG TAT TTT TAT GAA TAA TAT TCT CCG TTC AAT TTA CTG ATT GTA
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC TAC TAC TTA TAT GTA CAA TAT TAA AAT GAA AAC AAT ATA TTG TGC TGA ATA GGT TTA
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CGA CAT CTA TGA TAG AGC GCC ACA ATA ACA AAC AAT TGC GTT TTA TTA TTA CAA ATC
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA TTT TAA AAA AAG CGG CAG AAC CGG TCA AAC CTA AAA GAC TGA TTA CAT AAA TCT TAT
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA AAT TTC AAA AGG CCC CAG GGG CTA GTA TCT ACG ACA CAC CGA GCG GCG AAC TAA TAA
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGT TCA CTG AAG GGA ACT CCG GTT CCC CGC CGG CGC GCA TGG GTG AGA TTC CTT GAA GTT
                                ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAG TAT TGG CCG TCC GCT CTA CCG AAA GTT ACG GGC ACC ATT CAA CCC GGT CCA GCA CGG
                                ocs
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 CGG CCG GGT AAC CGA CTT GCT GCC CCG AGA ATT ATG CAG CAT TTT TTT GGT GTA TGT GGG
                                ocs
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 CCC CAA ATG AAG TGC AGG TCA AAC CTT GAC AGT GAC GAC AAA TCG TTG GGC GGG TCC AGG
                                ocs
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GCG AAT TTT GCG ACA ACA TGT CGA GGC TCA GCA GGA CCT GCA GGC ATG CAA GCT AGC TTA
```

ocs

~~~~~~

CTA GTG ATG CAT ATT CTA TAG TGT CAC CTA AAT CT
~~~~~~

Figura 2

CaMV 35S

T CGA CGA ATT AAT TCC AAT CCC ACA AAA ATC

CaMV 35S

TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC TCA TAT

CaMV 35S

CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG TTG TAC

CaMV 35S

AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT TAC AGA

CaMV 35S

GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA CTT CAT

CaMV 35S

CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG CCC ACC

CaMV 35S

AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC AGC CCC

CaMV 35S

AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA CGA TCT

CaMV 35S

AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA GAA GGT

CaMV 35S

TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA CGC AGC

CaMV 35S

AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT TCC CAA

CaMV 35S

GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA GAA AGA

CaMV 35S

CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT TCA TAA

CaMV 35S

ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC TAA GGC

CaMV 35S

CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA CTG GCG

CaMV 35S

AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA ACA TGG

CaMV 35S

~~~~~~
TGG AGC ACG ACA CTC TGG TCT ACT CCA AAA ATG TCA AAG ATA CAG TCT CAG AAG ACC AAA
CaMV 35S
~~~~~~
GGG CTA TTG AGA CTT TTC AAC AAA GGA TAA TTT CGG GAA ACC TCC TCG GAT TCC ATT GCC
CaMV 35S
~~~~~~
CAG CTA TCT GTC ACT TCA TCG AAA GGA CAG TAG AAA AGG AAG GTG GCT CCT ACA AAT GCC
CaMV 35S
~~~~~~
ATC ATT GCG ATA AAG GAA AGG CTA TCA TTC AAG ATC TCT CTG CCG ACA GTG GTC CCA AAG
CaMV 35S
~~~~~~
ATG GAC CCC CAC CCA CGA GGA GCA TCG TGG AAA AAG AAG ACG TTC CAA CCA CGT CTT CAA
CaMV 35S
~~~~~~
AGC AAG TGG ATT GAT GTG ACA TCT CCA CTG ACG TAA GGG ATG ACG CAC AAT CCC ACT ATC
CaMV 35S
~~~~~~
CTT CGC AAG ACC CTT CCT CTA TAT AAG GAA GTT CAT TTC ATT TGG AGA GGA CAC GCT CGA
attB1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GGA ATT CGG TAC CCC ATC ACA AGT TTG TAC AAA AAA GCA GGC TGC GGC CGC TTG CTC CCT
5' Oleo_1-1
~~~~~~
NcoI
~~~~~~~~
M A C H Y G Q Q Q Q T R A P H L Q
TAA AAA AAA CCA TGG CAT GTC ATT ACG GAC AAC AGC AAC AGA CTA GAG CAC CTC ATC TTC
5' Oleo_1-1
~~~~~~
UBQ10
~~~~~~~~~~~~~
· L Q P R A Q R V V K A A T A V T
AGC TTC AAC CTA GAG CAC AGA GAG TTG TGA AGG CTG CTA CTG CTG TTA CTG TAA ATT TCT
UBQ10
~~~~~~
GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT CCC AAT TTC GTA TAT
UBQ10
~~~~~~
GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT CTG GGT TTG
UBQ10
~~~~~~
AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG ATT TGT TCA AAT AAT
UBQ10
~~~~~~
TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT AGA TCT GGT GTT AGT
3'Oleo_1-1
~~~~~~~

UBQ10
~~~~~~~

~~~~~~

A G

TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT GAT TAA CAG GCT GGA

3' Oleo_1-1

~~~~~~

G S L L V L S G L T L A G T V I A L T I

GGA TCT CTT CTT GTT CTC TCT GGA CTT ACT CTC GCT GGA ACT GTT ATC GCT CTC ACT ATC

3' Oleo_1-1

~~~~~~

A T P L L V I F S P V L V P A V I T I F

GCT ACA CCT CTT CTC GTT ATC TTC TCT CCT GTT CTC GTT CCT GCT GTG ATC ACT ATC TTC

3' Oleo_1-1

~~~~~~

L L G A G F L A S G G F G V A A L S V L

CTT CTC GGA GCT GGA TTT CTT GCT TCT GGT GGA TTT GGA GTT GCT GCT CTC TCT GTT CTC

3' Oleo_1-1

~~~~~~

S W I Y R Y L T G K H P P G A D Q L E S

TCT TGG ATC TAC AGA TAC CTC ACT GGA AAA CAT CCT CCA GGT GCT GAT CAA CTT GAG TCT

3' Oleo_1-1

~~~~~~

A K T K L A S K A R E M K D R A E Q F S

GCT AAG ACT AAG CTC GCT TCT AAG GCT AGA GAG ATG AAG GAT AGA GCA GAG CAA TTC TCT

3' Oleo_1-1 terminador OCS

~~~~~~

C Q P V A G S Q T S * *

TGT CAG CCT GTT GCT GGA TCT CAG ACT TCT TAA TGA ACA TAT GGT CCT GCT TTA ATG AGA

terminador OCS

~~~~~~

TAT GCG AGA CGC CTA TGA TCG CAT GAT ATT TGC TTT CAA TTC TGT TGT GCA CGT TGT AAA

terminador OCS

~~~~~~

AAA CCT GAG CAT GTG TAG CTC AGA TCC TTA CCG CCG GTT TCG GTT CAT TCT AAT GAA TAT

terminador OCS

~~~~~~

ATC ACC CGT TAC TAT CGT ATT TTT ATG AAT AAT ATT CTC CGT TCA ATT TAC TGA TTG TAC

terminador OCS

~~~~~~

CCT ACT ACT TAT ATG TAC AAT ATT AAA ATG AAA ACA ATA TAT TGT GCT GAA TAG GTT TAT

terminador OCS

~~~~~~

AGC GAC ATC TAT GAT AGA GCG CCA CAA TAA CAA ACA ATT GCG TTT TAT TAT TAC AAA TCC

terminador OCS

~~~~~~

AAT TTT AAA AAA AGC GGC AGA ACC GGT CAA ACC TAA AAG ACT GAT TAC ATA AAT CTT ATT

terminador OCS

~~~~~~

CAA ATT TCA AAA GGC CCC AGG GGC TAG TAT CTA CGA CAC ACC GAG CGG CGA ACT AAT AAC

terminador OCS

~~~~~~

GTT CAC TGA AGG GAA CTC CGG TTC CCC GCC GGC GCG CAT GGG TGA GAT TCC TTG AAG TTG

terminador OCS
~~~~~~

~~~~~~

AGT ATT GGC CGT CCG CTC TAC CGA AAG TTA CGG GCA CCA TTC AAC CCG GTC CAG CAC GGC

terminador OCS

~~~~~~

GGC CGG GTA ACC GAC TTG CTG CCC CGA GAA TTA TGC AGC ATT TTT TTG GTG TAT GTG GGC

terminador OCS

~~~~~~

CCC AAA TGA AGT GCA GGT CAA ACC TTG ACA GTG ACG ACA AAT CGT TGG GCG GGT CCA GGG

terminador OCS

~~~~~~

CGA ATT TTG CGA CAA CAT GTC GAG GCT CAG CAG GAC CTG CAG GCA TGC AAG CTA GCT TAC

terminador OCS

~~~~~~

CaMV35S

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

TAG TGA TGC ATA TTC TAT AGT GTC ACC TAA ATC TTC GAC GAA TTA ATT CCA ATC CCA CAA

CaMV35S

~~~~~~

AAA TCT GAG CTT AAC AGC ACA GTT GCT CCT CTC AGA GCA GAA TCG GGT ATT CAA CAC CCT

CaMV35S

~~~~~~

CAT ATC AAC TAC TAC GTT GTG TAT AAC GGT CCA CAT GCC GGT ATA TAC GAT GAC TGG GGT

CaMV35S

~~~~~~

TGT ACA AAG GCG GCA ACA AAC GGC GTT CCC GGA GTT GCA CAC AAG AAA TTT GCC ACT ATT

CaMV35S

~~~~~~

ACA GAG GCA AGA GCA GCA GCT GAC GCG TAC ACA ACA AGT CAG CAA ACA GAC AGG TTG AAC

CaMV35S

~~~~~~

TTC ATC CCC AAA GGA GAA GCT CAA CTC AAG CCC AAG AGC TTT GCT AAG GCC CTA ACA AGC

CaMV35S

~~~~~~

CCA CCA AAG CAA AAA GCC CAC TGG CTC ACG CTA GGA ACC AAA AGG CCC AGC AGT GAT CCA

CaMV35S

~~~~~~

GCC CCA AAA GAG ATC TCC TTT GCC CCG GAG ATT ACA ATG GAC GAT TTC CTC TAT CTT TAC

CaMV35S

~~~~~~

GAT CTA GGA AGG AAG TTC GAA GGT GAA GGT GAC GAC ACT ATG TTC ACC ACT GAT AAT GAG

CaMV35S

~~~~~~

AAG GTT AGC CTC TTC AAT TTC AGA AAG AAT GCT GAC CCA CAG ATG GTT AGA GAG GCC TAC

CaMV35S

~~~~~~

GCA GCA GGT CTC ATC AAG ACG ATC TAC CCG AGT AAC AAT CTC CAG GAG ATC AAA TAC CTT

CaMV35S

~~~~~~

CCC AAG AAG GTT AAA GAT GCA GTC AAA AGA TTC AGG ACT AAT TGC ATC AAG AAC ACA GAG

CaMV35S
~~~~~~

~~~~~~
AAA GAC ATA TTT CTC AAG ATC AGA AGT ACT ATT CCA GTA TGG ACG ATT CAA GGC TTG CTT
CaMV35S
~~~~~~
CAT AAA CCA AGG CAA GTA ATA GAG ATT GGA GTC TCT AAA AAG GTA GTT CCT ACT GAA TCT
CaMV35S
~~~~~~
AAG GCC ATG CAT GGA GTC TAA GAT TCA AAT CGA GGA TCT AAC AGA ACT CGC CGT GAA GAC
CaMV35S
~~~~~~
TGG CGA ACA GTT CAT ACA GAG TCT TTT ACG ACT CAA TGA CAA GAA GAA AAT CTT CGT CAA
CaMV35S
~~~~~~
CAT GGT GGA GCA CGA CAC TCT GGT CTA CTC CAA AAA TGT CAA AGA TAC AGT CTC AGA AGA
CaMV35S
~~~~~~
CCA AAG GGC TAT TGA GAC TTT TCA ACA AAG GAT AAT TTC GGG AAA CCT CCT CGG ATT CCA
CaMV35S
~~~~~~
TTG CCC AGC TAT CTG TCA CTT CAT CGA AAG GAC AGT AGA AAA GGA AGG TGG CTC CTA CAA
CaMV35S
~~~~~~
ATG CCA TCA TTG CGA TAA AGG AAA GGC TAT CAT TCA AGA TCT CTC TGC CGA CAG TGG TCC
CaMV35S
~~~~~~
CAA AGA TGG ACC CCC ACC CAC GAG GAG CAT CGT GGA AAA AGA AGA CGT TCC AAC CAC GTC
CaMV35S
~~~~~~
TTC AAA GCA AGT GGA TTG ATG TGA CAT CTC CAC TGA CGT AAG GGA TGA CGC ACA ATC CCA
CaMV35S
~~~~~~
CTA TCC TTC GCA AGA CCC TTC CTC TAT ATA AGG AAG TTC ATT TCA TTT GGA GAG GAC ACG
5' DGAT1 (S205A)
~~~~~~
NcoI
~~~~~~~~
M A I L D S A G V T T V
GGA TCC TTG CTC CGT TAA AAA AAA CCA TGG CTA TCC TCG ATT CTG CTG GTG TTA CTA CTG
5' DGAT1 (S205A)
~~~~~~
· T E N G G G E F V D L D R L R R R K S R
TGA CTG AGA ATG GTG GTG GAG AGT TCG TTG ATC TCG ATA GAC TCA GAA GAA GAA AGT CTA
5' DGAT1 (S205A)
~~~~~~
UBQ10
~~~~~~
· S
GAT CTG TAA ATT TCT GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT
UBQ10
~~~~~~
CCC AAT TTC GTA TAT GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT
~~~~~~

```
                                UBQ10

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CTG GGT TTG ATC GTT AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG
                                UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATT TGT TCA AAT AAT TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT
                                UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGA TCT GGT GTT AGT TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT
  UBQ10
~~~~~~~~~~~~
                                3' DGAT1 (S205A)
                ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                 D   S   S   N   G   L   L   L   S   G   S   D   N   N   S   P   S
GAT TAA CAG GAT TCT TCT AAC GGA CTT CTC CTC TCT GGA TCT GAT AAC AAC TCT CCT TCT
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 D   D   V   G   A   P   A   D   V   R   D   R   I   D   S   V   V   N   D   D
GAT GAT GTT GGT GCT CCT GCT GAT GTG AGA GAT AGA ATC GAT TCT GTG GTG AAC GAT GAT
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   Q   G   T   A   N   L   A   G   D   N   N   G   G   G   D   N   N   G   G
GCT CAA GGT ACT GCT AAC CTC GCT GGT GAT AAT AAC GGT GGA GGT GAT AAC AAT GGT GGA
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 G   R   G   G   G   E   G   R   G   N   A   D   A   T   F   T   Y   R   P   S
GGA AGA GGT GGA GGT GAA GGT AGA GGA AAC GCT GAT GCT ACT TTC ACT TAC AGA CCA TCT
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 V   P   A   H   R   R   A   R   E   S   P   L   S   S   D   A   I   F   K   Q
GTG CCT GCT CAT AGA AGA GCT AGA GAG TCT CCT CTC TCT TCT GAT GCT ATC TTC AAG CAG
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 S   H   A   G   L   F   N   L   C   V   V   V   L   I   A   V   N   S   R   L
TCT CAC GCT GGA CTT TTC AAC CTC TGT GTG GTG GTT CTT ATC GCT GTG AAC TCT AGA CTC
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 I   I   E   N   L   M   K   Y   G   W   L   I   R   T   D   F   W   F   S   S
ATC ATC GAG AAC CTC ATG AAG TAC GGA TGG CTC ATC AGA ACT GAT TTC TGG TTC TCT TCT
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 R   S   L   R   D   W   P   L   F   M   C   C   I   S   L   S   I   F   P   L
AGA TCT CTC AGA GAT TGG CCT CTT TTC ATG TGC TGC ATC TCA CTC TCA ATC TTC CCT CTC
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   A   F   T   V   E   K   L   V   L   Q   K   Y   I   A   E   P   V   V   I
GCT GCT TTT ACT GTT GAG AAG CTC GTG CTC CAG AAG TAT ATC GCT GAA CCT GTG GTG ATC
                          3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 F   L   H   I   I   I   T   M   T   E   V   L   Y   P   V   Y   V   T   L   R
TTC CTC CAC ATC ATC ATC ACT ATG ACT GAG GTT CTC TAC CCT GTT TAC GTG ACT CTC AGA
```

3' DGAT1 (S205A)

~~~~~~

C D S A F L S G V T L M L L T C I V W L

TGC GAT TCT GCT TTC CTC TCT GGT GTT ACT CTT ATG CTC CTC ACT TGC ATT GTG TGG CTT

3' DGAT1 (S205A)

~~~~~~

K L V S Y A H T S Y D I R S L A N A A D

AAG CTC GTG TCT TAC GCT CAC ACT TCT TAC GAT ATC AGA TCT CTC GCT AAC GCT GCT GAT

3' DGAT1 (S205A)

~~~~~~

K A N P E V S Y Y V S L K S L A Y F M V

AAG GCT AAC CCT GAA GTG TCT TAC TAC GTG TCT CTC AAG TCT CTC GCT TAC TTC ATG GTT

3' DGAT1 (S205A)

~~~~~~

A P T L C Y Q P S Y P R S A C I R K G W

GCT CCT ACA CTT TGT TAC CAG CCA TCT TAC CCT AGA TCT GCT TGC ATT AGA AAG GGA TGG

3' DGAT1 (S205A)

~~~~~~

V A R Q F A K L V I F T G F M G F I I E

GTG GCA AGA CAA TTC GCT AAG TTG GTG ATC TTC ACT GGA TTC ATG GGA TTC ATC ATC GAG

3' DGAT1 (S205A)

~~~~~~

Q Y I N P I V R N S K H P L K G D L L Y

CAG TAC ATC AAC CCT ATT GTG AGA AAC TCT AAG CAC CCT CTC AAG GGT GAT CTT CTC TAC

3' DGAT1 (S205A)

~~~~~~

A I E R V L K L S V P N L Y V W L C M F

GCT ATC GAG AGA GTT CTT AAG CTC TCT GTG CCT AAC CTT TAT GTG TGG CTC TGC ATG TTC

3' DGAT1 (S205A)

~~~~~~

Y C F F H L W L N I L A E L L C F G D R

TAC TGT TTC TTC CAC CTC TGG CTT AAC ATC CTT GCT GAG TTG CTT TGC TTC GGA GAT AGA

3' DGAT1 (S205A)

~~~~~~

E F Y K D W W N A K S V G D Y W R M W N

GAG TTC TAC AAG GAT TGG TGG AAC GCT AAG TCT GTT GGA GAT TAT TGG AGA ATG TGG AAC

3' DGAT1 (S205A)

~~~~~~

M P V H K W M V R H I Y F P C L R S K I

ATG CCT GTG CAT AAG TGG ATG GTG CGT CAC ATC TAC TTC CCT TGC CTC AGA TCT AAG ATC

3' DGAT1 (S205A)

~~~~~~

P K T L A I I I A F L V S A V F H E L C

CCT AAG ACT CTC GCT ATC ATT ATC GCT TTC CTC GTG TCT GCT GTT TTC CAT GAG TTG TGT

3' DGAT1 (S205A)

~~~~~~

I A V P C R L F K L W A F L G I M F Q V

ATC GCT GTT CCT TGC AGA CTT TTC AAG CTT TGG GCT TTC CTC GGA ATC ATG TTC CAG GTT

3' DGAT1 (S205A)

~~~~~~

P L V F I T N Y L Q E R F G S T V G N M
~~~~~~

```
CCA CTC GTG TTC ATC ACT AAC TAC CTC CAA GAG AGA TTC GGA TCT ACT GTT GGA AAC ATG
                     3' DGAT1 (S205A)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 I   F   W   F   I   F   C   I   F   G   Q   P   M   C   V   L   L   Y   Y   H
ATT TTC TGG TTC ATT TTC TGC ATC TTC GGA CAG CCT ATG TGC GTT CTC CTC TAC TAC CAC
         3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 D   L   M   N   R   K   G   S   M   S   *   *
GAT CTC ATG AAC AGA AAG GGA TCT ATG TCT TAA TGA AGG ATC CAC CCA GCT TTC TTG TAC
                                                                  attB2
                                                                 ocs
                                                               ~~~~~~~~~~~~~~~~~~~~~~
AAA GTG GTG ATG GGT TCG AAA TCG ATA AGC TTG GAT CCT CTA GAG TCC TGC TTT AAT GAG
~~~~~~~~~
 attB2
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATA TGC GAG ACG CCT ATG ATC GCA TGA TAT TTG CTT TCA ATT CTG TTG TGC ACG TTG TAA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA ACC TGA GCA TGT GTA GCT CAG ATC CTT ACC GCC GGT TTC GGT TCA TTC TAA TGA ATA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT CAC CCG TTA CTA TCG TAT TTT TAT GAA TAA TAT TCT CCG TTC AAT TTA CTG ATT GTA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC TAC TAC TTA TAT GTA CAA TAT TAA AAT GAA AAC AAT ATA TTG TGC TGA ATA GGT TTA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CGA CAT CTA TGA TAG AGC GCC ACA ATA ACA AAC AAT TGC GTT TTA TTA TTA CAA ATC
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA TTT TAA AAA AAG CGG CAG AAC CGG TCA AAC CTA AAA GAC TGA TTA CAT AAA TCT TAT
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA AAT TTC AAA AGG CCC CAG GGG CTA GTA TCT ACG ACA CAC CGA GCG GCG AAC TAA TAA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGT TCA CTG AAG GGA ACT CCG GTT CCC CGC CGG CGC GCA TGG GTG AGA TTC CTT GAA GTT
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAG TAT TGG CCG TCC GCT CTA CCG AAA GTT ACG GGC ACC ATT CAA CCC GGT CCA GCA CGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGG CCG GGT AAC CGA CTT GCT GCC CCG AGA ATT ATG CAG CAT TTT TTT GGT GTA TGT GGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC CAA ATG AAG TGC AGG TCA AAC CTT GAC AGT GAC GAC AAA TCG TTG GGC GGG TCC AGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GCG AAT TTT GCG ACA ACA TGT CGA GGC TCA GCA GGA CCT GCA GGC ATG CAA GCT AGC TTA
```

ocs

~~~~~~

CTA GTG ATG CAT ATT CTA TAG TGT CAC CTA AAT CT
~~~~~~

Figura 3

CaMV 35S

T CGA CGA ATT AAT TCC AAT CCC ACA AAA ATC

CaMV 35S

TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC TCA TAT

CaMV 35S

CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG TTG TAC

CaMV 35S

AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT TAC AGA

CaMV 35S

GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA CTT CAT

CaMV 35S

CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG CCC ACC

CaMV 35S

AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC AGC CCC

CaMV 35S

AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA CGA TCT

CaMV 35S

AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA GAA GGT

CaMV 35S

TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA CGC AGC

CaMV 35S

AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT TCC CAA

CaMV 35S

GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA GAA AGA

CaMV 35S

CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT TCA TAA

CaMV 35S

ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC TAA GGC

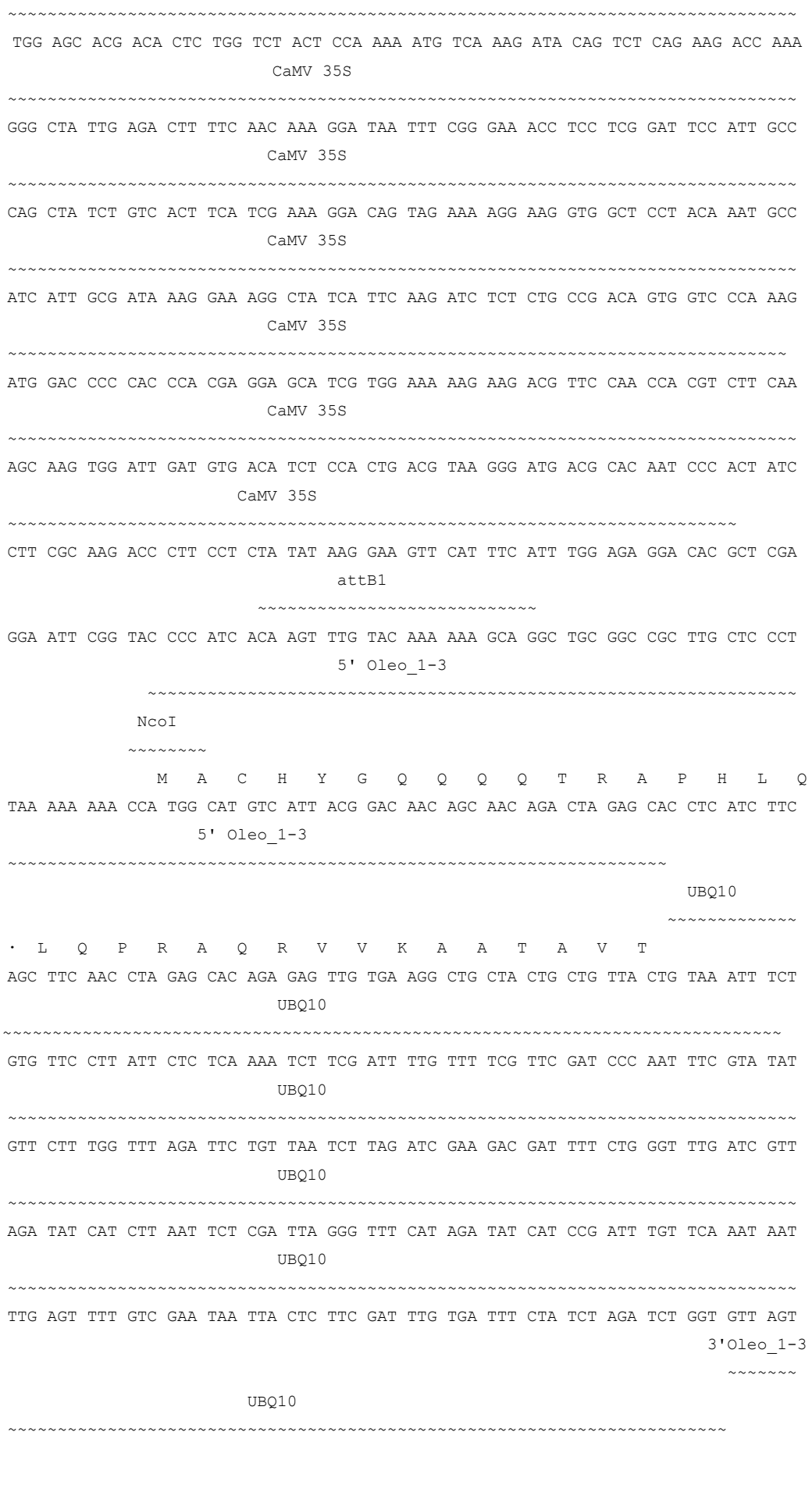
CaMV 35S

CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA CTG GCG

CaMV 35S

AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA ACA TGG

CaMV 35S


TGG AGC ACG ACA CTC TGG TCT ACT CCA AAA ATG TCA AAG ATA CAG TCT CAG AAG ACC AAA
CaMV 35S
GGG CTA TTG AGA CTT TTC AAC AAA GGA TAA TTT CGG GAA ACC TCC TCG GAT TCC ATT GCC
CaMV 35S
CAG CTA TCT GTC ACT TCA TCG AAA GGA CAG TAG AAA AGG AAG GTG GCT CCT ACA AAT GCC
CaMV 35S
ATC ATT GCG ATA AAG GAA AGG CTA TCA TTC AAG ATC TCT CTG CCG ACA GTG GTC CCA AAG
CaMV 35S
ATG GAC CCC CAC CCA CGA GGA GCA TCG TGG AAA AAG AAG ACG TTC CAA CCA CGT CTT CAA
CaMV 35S
AGC AAG TGG ATT GAT GTG ACA TCT CCA CTG ACG TAA GGG ATG ACG CAC AAT CCC ACT ATC
CaMV 35S
CTT CGC AAG ACC CTT CCT CTA TAT AAG GAA GTT CAT TTC ATT TGG AGA GGA CAC GCT CGA
attB1
GGA ATT CGG TAC CCC ATC ACA AGT TTG TAC AAA AAA GCA GGC TGC GGC CGC TTG CTC CCT
5' Oleo_1-3
NcoI
M A C H Y G Q Q Q Q T R A P H L Q
TAA AAA AAA CCA TGG CAT GTC ATT ACG GAC AAC AGC AAC AGA CTA GAG CAC CTC ATC TTC
5' Oleo_1-3
UBQ10
· L Q P R A Q R V V K A A T A V T
AGC TTC AAC CTA GAG CAC AGA GAG TTG TGA AGG CTG CTA CTG CTG TTA CTG TAA ATT TCT
UBQ10
GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT CCC AAT TTC GTA TAT
UBQ10
GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT CTG GGT TTG ATC GTT
UBQ10
AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG ATT TGT TCA AAT AAT
UBQ10
TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT AGA TCT GGT GTT AGT
3'Oleo_1-3
UBQ10

                                                                                A   G
TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT GAT TAA CAG GCT GGA
                        3' Oleo_1-3
~~~~~~
 G   S   L   L   V   L   S   G   L   T   L   A   G   T   V   I   A   L   T   I
GGA TCT CTT CTT GTT CTC TCT GGA CTT ACT CTC GCT GGA ACT GTT ATC GCT CTC ACT ATC
                        3' Oleo_1-3
~~~~~~
 A   T   P   L   L   V   I   F   S   P   V   L   V   P   A   V   I   T   I   F
GCT ACA CCT CTT CTC GTT ATC TTC TCT CCT GTT CTC GTT CCT GCT GTG ATC ACT ATC TTC
                        3' Oleo_1-3
~~~~~~
 L   L   G   A   G   F   L   A   S   G   G   F   G   V   A   A   L   S   V   L
CTT CTC GGA GCT GGA TTT CTT GCT TCT GGT GGA TTT GGA GTT GCT GCT CTC TCT GTT CTC
                        3' Oleo_1-3
~~~~~~
 S   W   I   Y   R   Y   L   T   G   K   H   P   P   G   A   D   C   L   E   S
TCT TGG ATC TAC AGA TAC CTC ACT GGA AAA CAT CCT CCA GGT GCT GAT TGT CTT GAG TCT
                        3' Oleo_1-3
~~~~~~
 A   K   T   K   L   A   S   C   A   R   E   M   K   D   R   A   E   Q   F   S
GCT AAG ACT AAG CTC GCT TCT TGT GCT AGA GAG ATG AAG GAT AGA GCA GAG CAA TTC TCT
                        3' Oleo_1-3   terminador OCS
~~~~~~
 C   Q   P   V   A   G   S   Q   T   S   *   *T
TGT CAG CCT GTT GCT GGA TCT CAG ACT TCT TAA TGA ACA TAT GGT CCT GCT TTA ATG AGA
                        terminador OCS
~~~~~~
TAT GCG AGA CGC CTA TGA TCG CAT GAT ATT TGC TTT CAA TTC TGT TGT GCA CGT TGT AAA
                        terminador OCS
~~~~~~
AAA CCT GAG CAT GTG TAG CTC AGA TCC TTA CCG CCG GTT TCG GTT CAT TCT AAT GAA TAT
                        terminador OCS
~~~~~~
ATC ACC CGT TAC TAT CGT ATT TTT ATG AAT AAT ATT CTC CGT TCA ATT TAC TGA TTG TAC
                        terminador OCS
~~~~~~
CCT ACT ACT TAT ATG TAC AAT ATT AAA ATG AAA ACA ATA TAT TGT GCT GAA TAG GTT TAT
                        terminador OCS
~~~~~~
AGC GAC ATC TAT GAT AGA GCG CCA CAA TAA CAA ACA ATT GCG TTT TAT TAT TAC AAA TCC
                        terminador OCS
~~~~~~
AAT TTT AAA AAA AGC GGC AGA ACC GGT CAA ACC TAA AAG ACT GAT TAC ATA AAT CTT ATT
                        terminador OCS
~~~~~~
CAA ATT TCA AAA GGC CCC AGG GGC TAG TAT CTA CGA CAC ACC GAG CGG CGA ACT AAT AAC
                        terminador OCS
~~~~~~
GTT CAC TGA AGG GAA CTC CGG TTC CCC GCC GGC GCG CAT GGG TGA GAT TCC TTG AAG TTG
                        terminador OCS

~~~~~~

AGT ATT GGC CGT CCG CTC TAC CGA AAG TTA CGG GCA CCA TTC AAC CCG GTC CAG CAC GGC

terminador OCS

~~~~~~

GGC CGG GTA ACC GAC TTG CTG CCC CGA GAA TTA TGC AGC ATT TTT TTG GTG TAT GTG GGC

terminador OCS

~~~~~~

CCC AAA TGA AGT GCA GGT CAA ACC TTG ACA GTG ACG ACA AAT CGT TGG GCG GGT CCA GGG

terminador OCS

~~~~~~

CGA ATT TTG CGA CAA CAT GTC GAG GCT CAG CAG GAC CTG CAG GCA TGC AAG CTA GCT TAC

terminador OCS

~~~~~~

CaMV35S

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

TAG TGA TGC ATA TTC TAT AGT GTC ACC TAA ATC TTC GAC GAA TTA ATT CCA ATC CCA CAA

CaMV35S

~~~~~~

AAA TCT GAG CTT AAC AGC ACA GTT GCT CCT CTC AGA GCA GAA TCG GGT ATT CAA CAC CCT

CaMV35S

~~~~~~

CAT ATC AAC TAC TAC GTT GTG TAT AAC GGT CCA CAT GCC GGT ATA TAC GAT GAC TGG GGT

CaMV35S

~~~~~~

TGT ACA AAG GCG GCA ACA AAC GGC GTT CCC GGA GTT GCA CAC AAG AAA TTT GCC ACT ATT

CaMV35S

~~~~~~

ACA GAG GCA AGA GCA GCA GCT GAC GCG TAC ACA ACA AGT CAG CAA ACA GAC AGG TTG AAC

CaMV35S

~~~~~~

TTC ATC CCC AAA GGA GAA GCT CAA CTC AAG CCC AAG AGC TTT GCT AAG GCC CTA ACA AGC

CaMV35S

~~~~~~

CCA CCA AAG CAA AAA GCC CAC TGG CTC ACG CTA GGA ACC AAA AGG CCC AGC AGT GAT CCA

CaMV35S

~~~~~~

GCC CCA AAA GAG ATC TCC TTT GCC CCG GAG ATT ACA ATG GAC GAT TTC CTC TAT CTT TAC

CaMV35S

~~~~~~

GAT CTA GGA AGG AAG TTC GAA GGT GAA GGT GAC GAC ACT ATG TTC ACC ACT GAT AAT GAG

CaMV35S

~~~~~~

AAG GTT AGC CTC TTC AAT TTC AGA AAG AAT GCT GAC CCA CAG ATG GTT AGA GAG GCC TAC

CaMV35S

~~~~~~

GCA GCA GGT CTC ATC AAG ACG ATC TAC CCG AGT AAC AAT CTC CAG GAG ATC AAA TAC CTT

CaMV35S

~~~~~~

CCC AAG AAG GTT AAA GAT GCA GTC AAA AGA TTC AGG ACT AAT TGC ATC AAG AAC ACA GAG

CaMV35S

~~~~~~

AAA GAC ATA TTT CTC AAG ATC AGA AGT ACT ATT CCA GTA TGG ACG ATT CAA GGC TTG CTT
CaMV35S
~~~~~~
CAT AAA CCA AGG CAA GTA ATA GAG ATT GGA GTC TCT AAA AAG GTA GTT CCT ACT GAA TCT
CaMV35S
~~~~~~
AAG GCC ATG CAT GGA GTC TAA GAT TCA AAT CGA GGA TCT AAC AGA ACT CGC CGT GAA GAC
CaMV35S
~~~~~~
TGG CGA ACA GTT CAT ACA GAG TCT TTT ACG ACT CAA TGA CAA GAA GAA AAT CTT CGT CAA
CaMV35S
~~~~~~
CAT GGT GGA GCA CGA CAC TCT GGT CTA CTC CAA AAA TGT CAA AGA TAC AGT CTC AGA AGA
CaMV35S
~~~~~~
CCA AAG GGC TAT TGA GAC TTT TCA ACA AAG GAT AAT TTC GGG AAA CCT CCT CGG ATT CCA
CaMV35S
~~~~~~
TTG CCC AGC TAT CTG TCA CTT CAT CGA AAG GAC AGT AGA AAA GGA AGG TGG CTC CTA CAA
CaMV35S
~~~~~~
ATG CCA TCA TTG CGA TAA AGG AAA GGC TAT CAT TCA AGA TCT CTC TGC CGA CAG TGG TCC
CaMV35S
~~~~~~
CAA AGA TGG ACC CCC ACC CAC GAG GAG CAT CGT GGA AAA AGA AGA CGT TCC AAC CAC GTC
CaMV35S
~~~~~~
TTC AAA GCA AGT GGA TTG ATG TGA CAT CTC CAC TGA CGT AAG GGA TGA CGC ACA ATC CCA
CaMV35S
~~~~~~
CTA TCC TTC GCA AGA CCC TTC CTC TAT ATA AGG AAG TTC ATT TCA TTT GGA GAG GAC ACG
5' DGAT1 (S205A)
~~~~~~
NcoI
~~~~~~~~
M A I L D S A G V T T V
GGA TCC TTG CTC CGT TAA AAA AAA CCA TGG CTA TCC TCG ATT CTG CTG GTG TTA CTA CTG
5' DGAT1 (S205A)
~~~~~~
· T E N G G G E F V D L D R L R R R K S R
TGA CTG AGA ATG GTG GTG GAG AGT TCG TTG ATC TCG ATA GAC TCA GAA GAA GAA AGT CTA
5' DGAT1 (S205A)
~~~~~~
UBQ10
~~~~~~
· S
GAT CTG TAA ATT TCT GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT
UBQ10
~~~~~~
CCC AAT TTC GTA TAT GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT

UBQ10

~~~~~~

CTG GGT TTG ATC GTT AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG

UBQ10

~~~~~~

ATT TGT TCA AAT AAT TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT

UBQ10

~~~~~~

AGA TCT GGT GTT AGT TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT

UBQ10

~~~~~~~~~~~~

3' DGAT1 (S205A)

~~~~~~

D S S N G L L L S G S D N N S P S

GAT TAA CAG GAT TCT TCT AAC GGA CTT CTC CTC TCT GGA TCT GAT AAC AAC TCT CCT TCT

3' DGAT1 (S205A)

~~~~~~

D D V G A P A D V R D R I D S V V N D D

GAT GAT GTT GGT GCT CCT GCT GAT GTG AGA GAT AGA ATC GAT TCT GTG GTG AAC GAT GAT

3' DGAT1 (S205A)

~~~~~~

A Q G T A N L A G D N N G G G D N N G G

GCT CAA GGT ACT GCT AAC CTC GCT GGT GAT AAT AAC GGT GGA GGT GAT AAC AAT GGT GGA

3' DGAT1 (S205A)

~~~~~~

G R G G G E G R G N A D A T F T Y R P S

GGA AGA GGT GGA GGT GAA GGT AGA GGA AAC GCT GAT GCT ACT TTC ACT TAC AGA CCA TCT

3' DGAT1 (S205A)

~~~~~~

V P A H R R A R E S P L S S D A I F K Q

GTG CCT GCT CAT AGA AGA GCT AGA GAG TCT CCT CTC TCT TCT GAT GCT ATC TTC AAG CAG

3' DGAT1 (S205A)

~~~~~~

S H A G L F N L C V V V L I A V N S R L

TCT CAC GCT GGA CTT TTC AAC CTC TGT GTG GTG GTT CTT ATC GCT GTG AAC TCT AGA CTC

3' DGAT1 (S205A)

~~~~~~

I I E N L M K Y G W L I R T D F W F S S

ATC ATC GAG AAC CTC ATG AAG TAC GGA TGG CTC ATC AGA ACT GAT TTC TGG TTC TCT TCT

3' DGAT1 (S205A)

~~~~~~

R S L R D W P L F M C C I S L S I F P L

AGA TCT CTC AGA GAT TGG CCT CTT TTC ATG TGC TGC ATC TCA CTC TCA ATC TTC CCT CTC

3' DGAT1 (S205A)

~~~~~~

A A F T V E K L V L Q K Y I A E P V V I

GCT GCT TTT ACT GTT GAG AAG CTC GTG CTC CAG AAG TAT ATC GCT GAA CCT GTG GTG ATC

3' DGAT1 (S205A)

~~~~~~

F L H I I I T M T E V L Y P V Y V T L R

TTC CTC CAC ATC ATC ATC ACT ATG ACT GAG GTT CTC TAC CCT GTT TAC GTG ACT CTC AGA

3' DGAT1 (S205A)

~~~~~~

C D S A F L S G V T L M L L T C I V W L

TGC GAT TCT GCT TTC CTC TCT GGT GTT ACT CTT ATG CTC CTC ACT TGC ATT GTG TGG CTT

3' DGAT1 (S205A)

~~~~~~

K L V S Y A H T S Y D I R S L A N A A D

AAG CTC GTG TCT TAC GCT CAC ACT TCT TAC GAT ATC AGA TCT CTC GCT AAC GCT GCT GAT

3' DGAT1 (S205A)

~~~~~~

K A N P E V S Y Y V S L K S L A Y F M V

AAG GCT AAC CCT GAA GTG TCT TAC TAC GTG TCT CTC AAG TCT CTC GCT TAC TTC ATG GTT

3' DGAT1 (S205A)

~~~~~~

A P T L C Y Q P S Y P R S A C I R K G W

GCT CCT ACA CTT TGT TAC CAG CCA TCT TAC CCT AGA TCT GCT TGC ATT AGA AAG GGA TGG

3' DGAT1 (S205A)

~~~~~~

V A R Q F A K L V I F T G F M G F I I E

GTG GCA AGA CAA TTC GCT AAG TTG GTG ATC TTC ACT GGA TTC ATG GGA TTC ATC ATC GAG

3' DGAT1 (S205A)

~~~~~~

Q Y I N P I V R N S K H P L K G D L L Y

CAG TAC ATC AAC CCT ATT GTG AGA AAC TCT AAG CAC CCT CTC AAG GGT GAT CTT CTC TAC

3' DGAT1 (S205A)

~~~~~~

A I E R V L K L S V P N L Y V W L C M F

GCT ATC GAG AGA GTT CTT AAG CTC TCT GTG CCT AAC CTT TAT GTG TGG CTC TGC ATG TTC

3' DGAT1 (S205A)

~~~~~~

Y C F F H L W L N I L A E L L C F G D R

TAC TGT TTC TTC CAC CTC TGG CTT AAC ATC CTT GCT GAG TTG CTT TGC TTC GGA GAT AGA

3' DGAT1 (S205A)

~~~~~~

E F Y K D W W N A K S V G D Y W R M W N

GAG TTC TAC AAG GAT TGG TGG AAC GCT AAG TCT GTT GGA GAT TAT TGG AGA ATG TGG AAC

3' DGAT1 (S205A)

~~~~~~

M P V H K W M V R H I Y F P C L R S K I

ATG CCT GTG CAT AAG TGG ATG GTG CGT CAC ATC TAC TTC CCT TGC CTC AGA TCT AAG ATC

3' DGAT1 (S205A)

~~~~~~

P K T L A I I I A F L V S A V F H E L C

CCT AAG ACT CTC GCT ATC ATT ATC GCT TTC CTC GTG TCT GCT GTT TTC CAT GAG TTG TGT

3' DGAT1 (S205A)

~~~~~~

I A V P C R L F K L W A F L G I M F Q V

ATC GCT GTT CCT TGC AGA CTT TTC AAG CTT TGG GCT TTC CTC GGA ATC ATG TTC CAG GTT

3' DGAT1 (S205A)

~~~~~~

P L V F I T N Y L Q E R F G S T V G N M
~~~~~~

```
CCA CTC GTG TTC ATC ACT AAC TAC CTC CAA GAG AGA TTC GGA TCT ACT GTT GGA AAC ATG
                     3' DGAT1 (S205A)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 I   F   W   F   I   F   C   I   F   G   Q   P   M   C   V   L   L   Y   Y   H
ATT TTC TGG TTC ATT TTC TGC ATC TTC GGA CAG CCT ATG TGC GTT CTC CTC TAC TAC CA
         3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 D   L   M   N   R   K   G   S   M   S   *   *
GAT CTC ATG AAC AGA AAG GGA TCT ATG TCT TAA TGA AGG ATC CAC CCA GCT TTC TTG TAC
                                                                attB2
                                                              ocs
                                                           ~~~~~~~~~~~~~~~~~~~~~
AAA GTG GTG ATG GGT TCG AAA TCG ATA AGC TTG GAT CCT CTA GAG TCC TGC TTT AAT GAG
~~~~~~~~~
attB2
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATA TGC GAG ACG CCT ATG ATC GCA TGA TAT TTG CTT TCA ATT CTG TTG TGC ACG TTG TAA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA ACC TGA GCA TGT GTA GCT CAG ATC CTT ACC GCC GGT TTC GGT TCA TTC TAA TGA ATA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT CAC CCG TTA CTA TCG TAT TTT TAT GAA TAA TAT TCT CCG TTC AAT TTA CTG ATT GTA
                            ocs

CCC TAC TAC TTA TAT GTA CAA TAT TAA AAT GAA AAC AAT ATA TTG TGC TGA ATA GGT TTA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CGA CAT CTA TGA TAG AGC GCC ACA ATA ACA AAC AAT TGC GTT TTA TTA TTA CAA ATC
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA TTT TAA AAA AAG CGG CAG AAC CGG TCA AAC CTA AAA GAC TGA TTA CAT AAA TCT TAT
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA AAT TTC AAA AGG CCC CAG GGG CTA GTA TCT ACG ACA CAC CGA GCG GCG AAC TAA TAA
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGT TCA CTG AAG GGA ACT CCG GTT CCC CGC CGG CGC GCA TGG GTG AGA TTC CTT GAA GTT
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAG TAT TGG CCG TCC GCT CTA CCG AAA GTT ACG GGC ACC ATT CAA CCC GGT CCA GCA CGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGG CCG GGT AAC CGA CTT GCT GCC CCG AGA ATT ATG CAG CAT TTT TTT GGT GTA TGT GGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC CAA ATG AAG TGC AGG TCA AAC CTT GAC AGT GAC GAC AAA TCG TTG GGC GGG TCC AGG
                            ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GCG AAT TTT GCG ACA ACA TGT CGA GGC TCA GCA GGA CCT GCA GGC ATG CAA GCT AGC TTA
```

ocs

~~~~~~

CTA GTG ATG CAT ATT CTA TAG TGT CAC CTA AAT CT
~~~~~~

Figura 4

CaMV 35S

TCG ACG AAT TAA TTC CAA TCC CAC AAA AAT

CaMV 35S

CTG AGC TTA ACA GCA CAG TTG CTC CTC TCA GAG CAG AAT CGG GTA TTC AAC ACC CTC ATA

CaMV 35S

TCA ACT ACT ACG TTG TGT ATA ACG GTC CAC ATG CCG GTA TAT ACG ATG ACT GGG GTT GTA

CaMV 35S

CAA AGG CGG CAA CAA ACG GCG TTC CCG GAG TTG CAC ACA AGA AAT TTG CCA CTA TTA CAG

CaMV 35S

AGG CAA GAG CAG CAG CTG ACG CGT ACA CAA CAA GTC AGC AAA CAG ACA GGT TGA ACT TCA

CaMV 35S

TCC CCA AAG GAG AAG CTC AAC TCA AGC CCA AGA GCT TTG CTA AGG CCC TAA CAA GCC CAC

CaMV 35S

CAA AGC AAA AAG CCC ACT GGC TCA CGC TAG GAA CCA AAA GGC CCA GCA GTG ATC CAG CCC

CaMV 35S

CAA AAG AGA TCT CCT TTG CCC CGG AGA TTA CAA TGG ACG ATT TCC TCT ATC TTT ACG ATC

CaMV 35S

TAG GAA GGA AGT TCG AAG GTG AAG GTG ACG ACA CTA TGT TCA CCA CTG ATA ATG AGA AGG

CaMV 35S

TTA GCC TCT TCA ATT TCA GAA AGA ATG CTG ACC CAC AGA TGG TTA GAG AGG CCT ACG CAG

CaMV 35S

CAG GTC TCA TCA AGA CGA TCT ACC CGA GTA ACA ATC TCC AGG AGA TCA AAT ACC TTC CCA

CaMV 35S

AGA AGG TTA AAG ATG CAG TCA AAA GAT TCA GGA CTA ATT GCA TCA AGA ACA CAG AGA AAG

CaMV 35S

ACA TAT TTC TCA AGA TCA GAA GTA CTA TTC CAG TAT GGA CGA TTC AAG GCT TGC TTC ATA

CaMV 35S

AAC CAA GGC AAG TAA TAG AGA TTG GAG TCT CTA AAA AGG TAG TTC CTA CTG AAT CTA AGG

CaMV 35S

CCA TGC ATG GAG TCT AAG ATT CAA ATC GAG GAT CTA ACA GAA CTC GCC GTG AAG ACT GGC

CaMV35S

GAA CAG TTC ATA CAG AGT CTT TTA CGA CTC AAT GAC AAG AAG AAA ATC TTC GTC AAC ATG

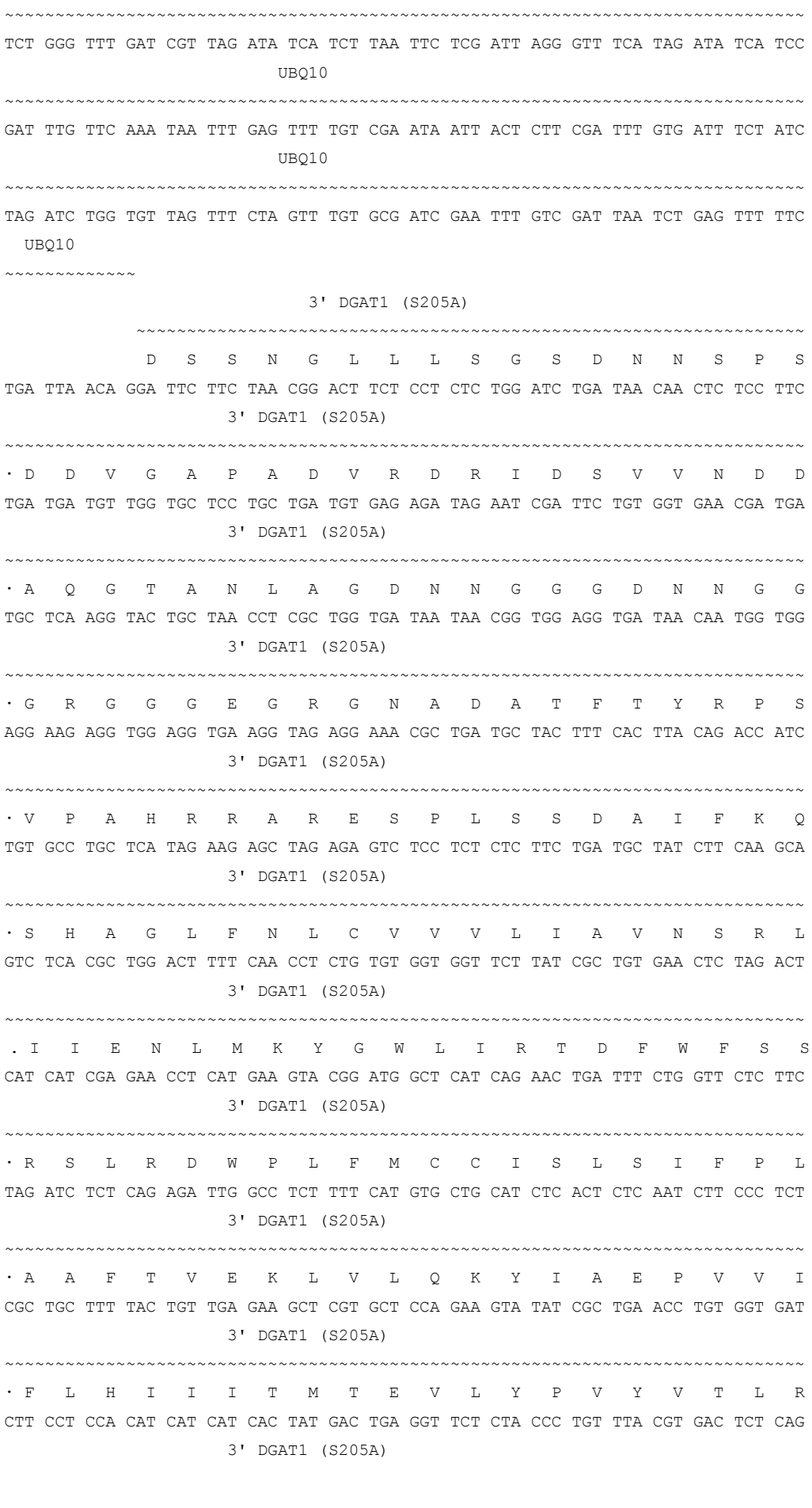
CaMV 35S

```
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GTG GAG CAC GAC ACT CTG GTC TAC TCC AAA AAT GTC AAA GAT ACA GTC TCA GAA GAC CAA
                         CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGG GCT ATT GAG ACT TTT CAA CAA AGG ATA ATT TCG GGA AAC CTC CTC GGA TTC CAT TGC
                          CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCA GCT ATC TGT CAC TTC ATC GAA AGG ACA GTA GAA AAG GAA GGT GGC TCC TAC AAA TGC
                          CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAT CAT TGC GAT AAA GGA AAG GCT ATC ATT CAA GAT CTC TCT GCC GAC AGT GGT CCC AAA
                          CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAT GGA CCC CCA CCC ACG AGG AGC ATC GTG GAA AAA GAA GAC GTT CCA ACC ACG TCT TCA
                          CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAG CAA GTG GAT TGA TGT GAC ATC TCC ACT GAC GTA AGG GAT GAC GCA CAA TCC CAC TAT
                        CaMV 35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCT TCG CAA GAC CCT TCC TCT ATA TAA GGA AGT TCA TTT CAT TTG GAG AGG ACA CGC TCG
                                  attB1
                          ~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGG AAT TCG GTA CCC CAT CAC AAG TTT GTA CAA AAA AGC AGG CTG CGG CCG CTT GCT CCC
                                  5' Oleo_3-1
                ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
              NcoI
             ~~~~~~~~
                 M   A   C   H   Y   G   Q   Q   Q   Q   T   C   A   P   H   L
TTA AAA AAA ACC ATG GCA TGT CAT TAC GGA CAA CAG CAA CAG ACT TGT GCA CCT CAT CTT
                    5' Oleo_3-1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                                                      UBQ10
                                                                    ~~~~~~~~~~~
 Q   L   Q   P   R   A   C   R   V   V   K   A   A   T   A   V   T
CAG CTT CAA CCT AGA GCA TGT AGA GTT GTG AAG GCT GCT ACT GCT GTT ACT GTA AAT TTC
                           UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TGT GTT CCT TAT TCT CTC AAA ATC TTC GAT TTT GTT TTC GTT CGA TCC CAA TTT CGT ATA
                           UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TGT TCT TTG GTT TAG ATT CTG TTA ATC TTA GAT CGA AGA CGA TTT TCT GGG TTT GAT CGT
                           UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG ATA TCA TCT TAA TTC TCG ATT AGG GTT TCA TAG ATA TCA TCC GAT TTG TTC AAA TAA
                           UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TTT GAG TTT TGT CGA ATA ATT ACT CTT CGA TTT GTG ATT TCT ATC TAG ATC TGG TGT TAG
                                                                      3'Oleo_3-1
                                                                         ~~~~~~
                         UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
                                                                                A   G
TTT CTA GTT TGT GCG ATC GAA TTT GTC GAT TAA TCT GAG TTT TTC TGA TTA ACA GGC TGG
                          3' Oleo_3-1
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· G   S   L   L   V   L   S   G   L   T   L   A   G   T   V   I   A   L   T   I
AGG ATC TCT TCT TGT TCT CTC TGG ACT TAC TCT CGC TGG AAC TGT TAT CGC TCT CAC TAT
                          3' Oleo_3-1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· A   T   P   L   L   V   I   F   S   P   V   L   V   P   A   V   I   T   I   F
CGC TAC ACC TCT TCT CGT TAT CTT CTC TCC TGT TCT CGT TCC TGC TGT GAT CAC TAT CTT
                          3' Oleo_3-1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· L   L   G   A   G   F   L   A   S   G   G   F   G   V   A   A   L   S   V   L
CCT TCT CGG AGC TGG ATT TCT TGC TTC TGG TGG ATT TGG AGT TGC TGC TCT CTC TGT TCT
                          3' Oleo_3-1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· S   W   I   Y   R   Y   L   T   G   K   H   P   P   G   A   D   Q   L   E   S
CTC TTG GAT CTA CAG ATA CCT CAC TGG AAA ACA TCC TCC AGG TGC TGA TCA ACT TGA GTC
                          3' Oleo_3-1
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· A   K   T   K   L   A   S   K   A   R   E   M   K   D   R   A   E   Q   F   S
TGC TAA GAC TAA GCT CGC TTC TAA GGC TAG AGA GAT GAA GGA TAG AGC AGA GCA ATT CTC
            3' Oleo_3-1                                       terminador OCS
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
· C   Q   P   V   A   G   S   Q   T   S   *   *
TTG TCA GCC TGT TGC TGG ATC TCA GAC TTC TTA ATG AAC ATA TGG TCC TGC TTT AAT GAG
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATA TGC GAG ACG CCT ATG ATC GCA TGA TAT TTG CTT TCA ATT CTG TTG TGC ACG TTG TAA
                          terminador OCS
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA ACC TGA GCA TGT GTA GCT CAG ATC CTT ACC GCC GGT TTC GGT TCA TTC TAA TGA ATA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT CAC CCG TTA CTA TCG TAT TTT TAT GAA TAA TAT TCT CCG TTC AAT TTA CTG ATT GTA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC TAC TAC TTA TAT GTA CAA TAT TAA AAT GAA AAC AAT ATA TTG TGC TGA ATA GGT TTA
                           terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CGA CAT CTA TGA TAG AGC GCC ACA ATA ACA AAC AAT TGC GTT TTA TTA TTA CAA ATC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA TTT TAA AAA AAG CGG CAG AAC CGG TCA AAC CTA AAA GAC TGA TTA CAT AAA TCT TAT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA AAT TTC AAA AGG CCC CAG GGG CTA GTA TCT ACG ACA CAC CGA GCG GCG AAC TAA TAA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGT TCA CTG AAG GGA ACT CCG GTT CCC CGC CGG CGC GCA TGG GTG AGA TTC CTT GAA GTT
                          terminador OCS
```

~~~~~~
GAG TAT TGG CCG TCC GCT CTA CCG AAA GTT ACG GGC ACC ATT CAA CCC GGT CCA GCA CGG
terminador OCS
~~~~~~
CGG CCG GGT AAC CGA CTT GCT GCC CCG AGA ATT ATG CAG CAT TTT TTT GGT GTA TGT GGG
terminador OCS
~~~~~~
CCC CAA ATG AAG TGC AGG TCA AAC CTT GAC AGT GAC GAC AAA TCG TTG GGC GGG TCC AGG
terminador OCS
~~~~~~
GCG AAT TTT GCG ACA ACA TGT CGA GGC TCA GCA GGA CCT GCA GGC ATG CAA GCT AGC TTA
terminador OCS
~~~~~~
CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CTA GTG ATG CAT ATT CTA TAG TGT CAC CTA AAT CTT CGA CGA ATT AAT TCC AAT CCC ACA
CaMV35S
~~~~~~
AAA ATC TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC
CaMV35S
~~~~~~
TCA TAT CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG
CaMV35S
~~~~~~
TTG TAC AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT
CaMV35S
~~~~~~
TAC AGA GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA
CaMV35S
~~~~~~
CTT CAT CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG
CaMV35S
~~~~~~
CCC ACC AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC
CaMV35S
~~~~~~
AGC CCC AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA
CaMV35S
~~~~~~
CGA TCT AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA
CaMV35S
~~~~~~
GAA GGT TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA
CaMV35S
~~~~~~
CGC AGC AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT
CaMV35S
~~~~~~
TCC CAA GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA
CaMV35S
~~~~~~

```
GAA AGA CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA TAA ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAA GGC CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CTG GCG AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ACA TGG TGG AGC ACG ACA CTC TGG TCT ACT CCA AAA ATG TCA AAG ATA CAG TCT CAG AAG
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ACC AAA GGG CTA TTG AGA CTT TTC AAC AAA GGA TAA TTT CGG GAA ACC TCC TCG GAT TCC
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATT GCC CAG CTA TCT GTC ACT TCA TCG AAA GGA CAG TAG AAA AGG AAG GTG GCT CCT ACA
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAT GCC ATC ATT GCG ATA AAG GAA AGG CTA TCA TTC AAG ATC TCT CTG CCG ACA GTG GTC
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCA AAG ATG GAC CCC CAC CCA CGA GGA GCA TCG TGG AAA AAG AAG ACG TTC CAA CCA CGT
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CTT CAA AGC AAG TGG ATT GAT GTG ACA TCT CCA CTG ACG TAA GGG ATG ACG CAC AAT CCC
                            CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ACT ATC CTT CGC AAG ACC CTT CCT CTA TAT AAG GAA GTT CAT TTC ATT TGG AGA GGA CAC
CaMV35S                                    5' DGAT1 (S205A)
~                                 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                NcoI
                               ~~~~~~~~
                                   M   A   I   L   D   S   A   G   V   T   T
GGG ATC CTT GCT CCG TTA AAA AAA ACC ATG GCT ATC CTC GAT TCT GCT GGT GTT ACT ACT
                      5' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 V   T   E   N   G   G   G   E   F   V   D   L   D   R   L   R   R   R   K   S
GTG ACT GAG AAT GGT GGT GGA GAG TTC GTT GAT CTC GAT AGA CTC AGA AGA AGA AAG TCT
5' DGAT1 (S205A)
~~~~~~~~
                                UBQ10
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 R   S
AGA TCT GTA AAT TTC TGT GTT CCT TAT TCT CTC AAA ATC TTC GAT TTT GTT TTC GTT CGA
                            UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCC CAA TTT CGT ATA TGT TCT TTG GTT TAG ATT CTG TTA ATC TTA GAT CGA AGA CGA TTT
                            UBQ10
```


UBQ10
UBQ10
UBQ10
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)
3' DGAT1 (S205A)

~~~~~~

·C D S A F L S G V T L M L L T C I V W L

ATG CGA TTC TGC TTT CCT CTC TGG TGT TAC TCT TAT GCT CCT CAC TTG CAT TGT GTG GCT

3' DGAT1 (S205A)

~~~~~~

·K L V S Y A H T S Y D I R S L A N A A D

TAA GCT CGT GTC TTA CGC TCA CAC TTC TTA CGA TAT CAG ATC TCT CGC TAA CGC TGC TGA

3' DGAT1 (S205A)

~~~~~~

·K A N P E V S Y Y V S L K S L A Y F M V

TAA GGC TAA CCC TGA AGT GTC TTA CTA CGT GTC TCT CAA GTC TCT CGC TTA CTT CAT GGT

3' DGAT1 (S205A)

~~~~~~

·A P T L C Y Q P S Y P R S A C I R K G W

TGC TCC TAC ACT TTG TTA CCA GCC ATC TTA CCC TAG ATC TGC TTG CAT TAG AAA GGG ATG

3' DGAT1 (S205A)

~~~~~~

·V A R Q F A K L V I F T G F M G F I I E

GGT GGC AAG ACA ATT CGC TAA GTT GGT GAT CTT CAC TGG ATT CAT GGG ATT CAT CAT CGA

3' DGAT1 (S205A)

~~~~~~

·Q Y I N P I V R N S K H P L K G D L L Y

GCA GTA CAT CAA CCC TAT TGT GAG AAA CTC TAA GCA CCC TCT CAA GGG TGA TCT TCT CTA

3' DGAT1 (S205A)

~~~~~~

·A I E R V L K L S V P N L Y V W L C M F

CGC TAT CGA GAG AGT TCT TAA GCT CTC TGT GCC TAA CCT TTA TGT GTG GCT CTG CAT GTT

3' DGAT1 (S205A)

~~~~~~

·Y C F F H L W L N I L A E L L C F G D R

CTA CTG TTT CTT CCA CCT CTG GCT TAA CAT CCT TGC TGA GTT GCT TTG CTT CGG AGA TAG

3' DGAT1 (S205A)

~~~~~~

·E F Y K D W W N A K S V G D Y W R M W N

AGA GTT CTA CAA GGA TTG GTG GAA CGC TAA GTC TGT TGG AGA TTA TTG GAG AAT GTG GAA

3' DGAT1 (S205A)

~~~~~~

·M P V H K W M V R H I Y F P C L R S K I

CAT GCC TGT GCA TAA GTG GAT GGT GCG TCA CAT CTA CTT CCC TTG CCT CAG ATC TAA GAT

3' DGAT1 (S205A)

~~~~~~

·P K T L A I I I A F L V S A V F H E L C

CCC TAA GAC TCT CGC TAT CAT TAT CGC TTT CCT CGT GTC TGC TGT TTT CCA TGA GTT GTG

3' DGAT1 (S205A)

--~~~~~~

.I A V P C R L F K L W A F L G I M F Q V

TAT CGC TGT TCC TTG CAG ACT TTT CAA GCT TTG GGC TTT CCT CGG AAT CAT GTT CCA GGT

3' DGAT1 (S205A)

~~~~~~

·P L V F I T N Y L Q E R F G S T V G N M

TCC ACT CGT GTT CAT CAC TAA CTA CCT CCA AGA GAG ATT CGG ATC TAC TGT TGG AAA CAT

3' DGAT1 (S205A)

~~~~~~

· I F W F I F C I F G Q P M C V L L Y Y H

GAT TTT CTG GTT CAT TTT CTG CAT CTT CGG ACA GCC TAT GTG CGT TCT CCT CTA CTA CCA

3' DGAT1 (S205A)

~~~~~~

· D L M N R K G S M S * *

CGA TCT CAT GAA CAG AAA GGG ATC TAT GTC TTA ATG AAG GAT CCA CCC AGC TTT CTT GTA

attB2

ocs

~~~~~~~~~~~~~~~~~~~

CAA AGT GGT GAT GGG TTC GAA ATC GAT AAG CTT GGA TCC TCT AGA GTC CTG CTT TAA TGA

~~~~~~~~~~

attB2

ocs

~~~~~~

GAT ATG CGA GAC GCC TAT GAT CGC ATG ATA TTT GCT TTC AAT TCT GTT GTG CAC GTT GTA

ocs

~~~~~~

AAA AAC CTG AGC ATG TGT AGC TCA GAT CCT TAC CGC CGG TTT CGG TTC ATT CTA ATG AAT

ocs

~~~~~~

ATA TCA CCC GTT ACT ATC GTA TTT TTA TGA ATA ATA TTC TCC GTT CAA TTT ACT GAT TGT

ocs

~~~~~~

ACC CTA CTA CTT ATA TGT ACA ATA TTA AAA TGA AAA CAA TAT ATT GTG CTG AAT AGG TTT

ocs

~~~~~~

ATA GCG ACA TCT ATG ATA GAG CGC CAC AAT AAC AAA CAA TTG CGT TTT ATT ATT ACA AAT

ocs

~~~~~~

CCA ATT TTA AAA AAA GCG GCA GAA CCG GTC AAA CCT AAA AGA CTG ATT ACA TAA ATC TTA

ocs

~~~~~~

TTC AAA TTT CAA AAG GCC CCA GGG GCT AGT ATC TAC GAC ACA CCG AGC GGC GAA CTA ATA

ocs

~~~~~~

ACG TTC ACT GAA GGG AAC TCC GGT TCC CCG CCG GCG CGC ATG GGT GAG ATT CCT TGA AGT

ocs

~~~~~~

TGA GTA TTG GCC GTC CGC TCT ACC GAA AGT TAC GGG CAC CAT TCA ACC CGG TCC AGC ACG

ocs

~~~~~~

GCG GCC GGG TAA CCG ACT TGC TGC CCC GAG AAT TAT GCA GCA TTT TTT TGG TGT ATG TGG

ocs

~~~~~~

GCC CCA AAT GAA GTG CAG GTC AAA CCT TGA CAG TGA CGA CAA ATC GTT GGG CGG GTC CAG

ocs

~~~~~~

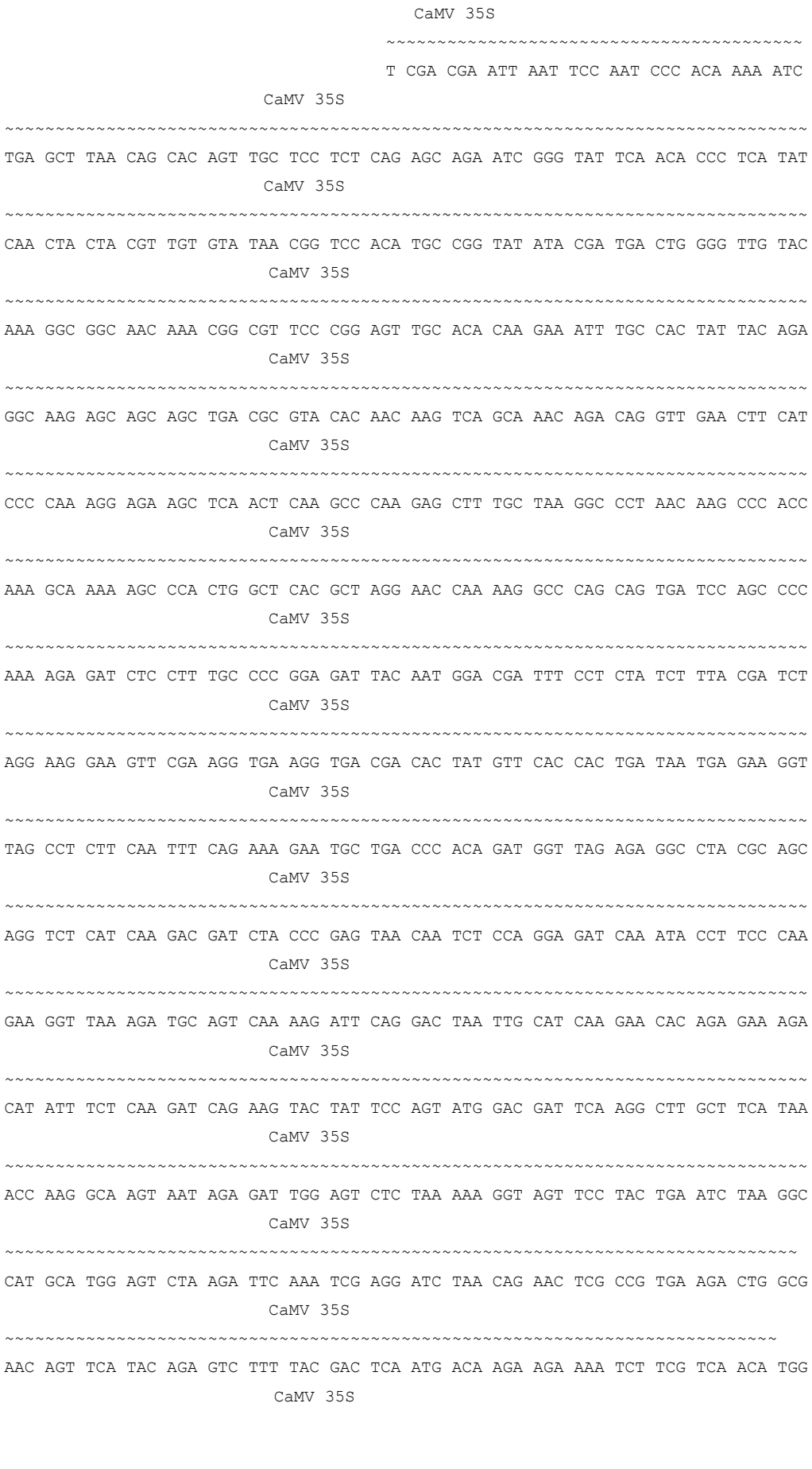
GGC GAA TTT TGC GAC AAC ATG TCG AGG CTC AGC AGG ACC TGC AGG CAT GCA AGC TAG CTT

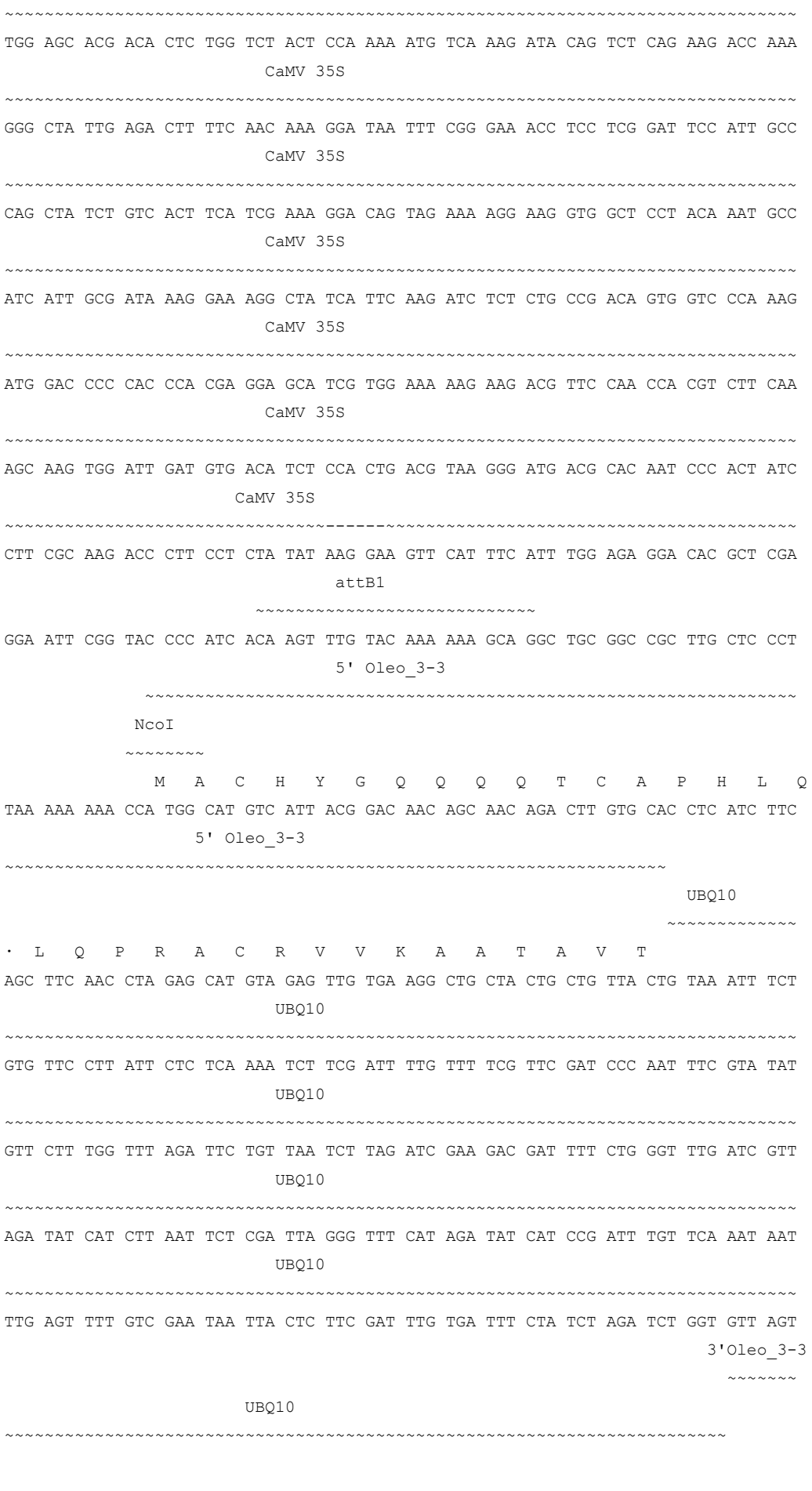
ocs

~~~~~~

ACT AGT GAT GCA TAT TCT ATA GTG TCA CCT AAA TCT
~~~~~~

Figura 5


CaMV 35S
T CGA CGA ATT AAT TCC AAT CCC ACA AAA ATC
CaMV 35S
TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC TCA TAT
CaMV 35S
CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG TTG TAC
CaMV 35S
AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT TAC AGA
CaMV 35S
GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA CTT CAT
CaMV 35S
CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG CCC ACC
CaMV 35S
AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC AGC CCC
CaMV 35S
AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA CGA TCT
CaMV 35S
AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA GAA GGT
CaMV 35S
TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA CGC AGC
CaMV 35S
AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT TCC CAA
CaMV 35S
GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA GAA AGA
CaMV 35S
CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT TCA TAA
CaMV 35S
ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC TAA GGC
CaMV 35S
CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA CTG GCG
CaMV 35S
AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA ACA TGG
CaMV 35S


TGG AGC ACG ACA CTC TGG TCT ACT CCA AAA ATG TCA AAG ATA CAG TCT CAG AAG ACC AAA
CaMV 35S
GGG CTA TTG AGA CTT TTC AAC AAA GGA TAA TTT CGG GAA ACC TCC TCG GAT TCC ATT GCC
CaMV 35S
CAG CTA TCT GTC ACT TCA TCG AAA GGA CAG TAG AAA AGG AAG GTG GCT CCT ACA AAT GCC
CaMV 35S
ATC ATT GCG ATA AAG GAA AGG CTA TCA TTC AAG ATC TCT CTG CCG ACA GTG GTC CCA AAG
CaMV 35S
ATG GAC CCC CAC CCA CGA GGA GCA TCG TGG AAA AAG AAG ACG TTC CAA CCA CGT CTT CAA
CaMV 35S
AGC AAG TGG ATT GAT GTG ACA TCT CCA CTG ACG TAA GGG ATG ACG CAC AAT CCC ACT ATC
CaMV 35S
CTT CGC AAG ACC CTT CCT CTA TAT AAG GAA GTT CAT TTC ATT TGG AGA GGA CAC GCT CGA
attB1
GGA ATT CGG TAC CCC ATC ACA AGT TTG TAC AAA AAA GCA GGC TGC GGC CGC TTG CTC CCT
5' Oleo_3-3
NcoI
M A C H Y G Q Q Q Q T C A P H L Q
TAA AAA AAA CCA TGG CAT GTC ATT ACG GAC AAC AGC AAC AGA CTT GTG CAC CTC ATC TTC
5' Oleo_3-3
UBQ10
· L Q P R A C R V V K A A T A V T
AGC TTC AAC CTA GAG CAT GTA GAG TTG TGA AGG CTG CTA CTG CTG TTA CTG TAA ATT TCT
UBQ10
GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT CCC AAT TTC GTA TAT
UBQ10
GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT CTG GGT TTG ATC GTT
UBQ10
AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG ATT TGT TCA AAT AAT
UBQ10
TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT AGA TCT GGT GTT AGT
3'Oleo_3-3
UBQ10

```
                                                                                A   G
TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT GAT TAA CAG GCT GGA
                          3' Oleo_3-3
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 G   S   L   L   V   L   S   G   L   T   L   A   G   T   V   I   A   L   T   I
GGA TCT CTT CTT GTT CTC TCT GGA CTT ACT CTC GCT GGA ACT GTT ATC GCT CTC ACT ATC
                          3' Oleo_3-3
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   T   P   L   L   V   I   F   S   P   V   L   V   P   A   V   I   T   I   F
GCT ACA CCT CTT CTC GTT ATC TTC TCT CCT GTT CTC GTT CCT GCT GTG ATC ACT ATC TTC
                          3' Oleo_3-3
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 L   L   G   A   G   F   L   A   S   G   G   F   G   V   A   A   L   S   V   L
CTT CTC GGA GCT GGA TTT CTT GCT TCT GGT GGA TTT GGA GTT GCT GCT CTC TCT GTT CTC
                          3' Oleo_3-3
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 S   W   I   Y   R   Y   L   T   G   K   H   P   P   G   A   D   C   L   E   S
TCT TGG ATC TAC AGA TAC CTC ACT GGA AAA CAT CCT CCA GGT GCT GAT TGT CTT GAG TCT
                          3' Oleo_3-3
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   K   T   K   L   A   S   C   A   R   E   M   K   D   R   A   E   Q   F   S
GCT AAG ACT AAG CTC GCT TCT TGT GCT AGA GAG ATG AAG GAT AGA GCA GAG CAA TTC TCT
                          3' Oleo_3-3                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 C   Q   P   V   A   G   S   Q   T   S   *   *
TGT CAG CCT GTT GCT GGA TCT CAG ACT TCT TAA TGA ACA TAT GGT CCT GCT TTA ATG AGA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT GCG AGA CGC CTA TGA TCG CAT GAT ATT TGC TTT CAA TTC TGT TGT GCA CGT TGT AAA
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA CCT GAG CAT GTG TAG CTC AGA TCC TTA CCG CCG GTT TCG GTT CAT TCT AAT GAA TAT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATC ACC CGT TAC TAT CGT ATT TTT ATG AAT AAT ATT CTC CGT TCA ATT TAC TGA TTG TAC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCT ACT ACT TAT ATG TAC AAT ATT AAA ATG AAA ACA ATA TAT TGT GCT GAA TAG GTT TAT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AGC GAC ATC TAT GAT AGA GCG CCA CAA TAA CAA ACA ATT GCG TTT TAT TAT TAC AAA TCC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAT TTT AAA AAA AGC GGC AGA ACC GGT CAA ACC TAA AAG ACT GAT TAC ATA AAT CTT ATT
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA ATT TCA AAA GGC CCC AGG GGC TAG TAT CTA CGA CAC ACC GAG CGG CGA ACT AAT AAC
                          terminador OCS
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GTT CAC TGA AGG GAA CTC CGG TTC CCC GCC GGC GCG CAT GGG TGA GAT TCC TTG AAG TTG
                          terminador OCS
```

~~~~~~

AGT ATT GGC CGT CCG CTC TAC CGA AAG TTA CGG GCA CCA TTC AAC CCG GTC CAG CAC GGC

terminador OCS

GGC CGG GTA ACC GAC TTG CTG CCC CGA GAA TTA TGC AGC ATT TTT TTG GTG TAT GTG GGC

terminador OCS

CCC AAA TGA AGT GCA GGT CAA ACC TTG ACA GTG ACG ACA AAT CGT TGG GCG GGT CCA GGG

terminador OCS

CGA ATT TTG CGA CAA CAT GTC GAG GCT CAG CAG GAC CTG CAG GCA TGC AAG CTA GCT TAC

terminador OCS

CaMV35S

TAG TGA TGC ATA TTC TAT AGT GTC ACC TAA ATC TTC GAC GAA TTA ATT CCA ATC CCA CAA

CaMV35S

AAA TCT GAG CTT AAC AGC ACA GTT GCT CCT CTC AGA GCA GAA TCG GGT ATT CAA CAC CCT

CaMV35S

CAT ATC AAC TAC TAC GTT GTG TAT AAC GGT CCA CAT GCC GGT ATA TAC GAT GAC TGG GGT

CaMV35S

TGT ACA AAG GCG GCA ACA AAC GGC GTT CCC GGA GTT GCA CAC AAG AAA TTT GCC ACT ATT

CaMV35S

ACA GAG GCA AGA GCA GCA GCT GAC GCG TAC ACA ACA AGT CAG CAA ACA GAC AGG TTG AAC

CaMV35S

TTC ATC CCC AAA GGA GAA GCT CAA CTC AAG CCC AAG AGC TTT GCT AAG GCC CTA ACA AGC

CaMV35S

CCA CCA AAG CAA AAA GCC CAC TGG CTC ACG CTA GGA ACC AAA AGG CCC AGC AGT GAT CCA

CaMV35S

GCC CCA AAA GAG ATC TCC TTT GCC CCG GAG ATT ACA ATG GAC GAT TTC CTC TAT CTT TAC

CaMV35S

GAT CTA GGA AGG AAG TTC GAA GGT GAA GGT GAC GAC ACT ATG TTC ACC ACT GAT AAT GAG

CaMV35S

AAG GTT AGC CTC TTC AAT TTC AGA AAG AAT GCT GAC CCA CAG ATG GTT AGA GAG GCC TAC

CaMV35S

GCA GCA GGT CTC ATC AAG ACG ATC TAC CCG AGT AAC AAT CTC CAG GAG ATC AAA TAC CTT

CaMV35S

CCC AAG AAG GTT AAA GAT GCA GTC AAA AGA TTC AGG ACT AAT TGC ATC AAG AAC ACA GAG

CaMV35S
~~~~~~

```
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA GAC ATA TTT CTC AAG ATC AGA AGT ACT ATT CCA GTA TGG ACG ATT CAA GGC TTG CTT
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAT AAA CCA AGG CAA GTA ATA GAG ATT GGA GTC TCT AAA AAG GTA GTT CCT ACT GAA TCT
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAG GCC ATG CAT GGA GTC TAA GAT TCA AAT CGA GGA TCT AAC AGA ACT CGC CGT GAA GAC
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TGG CGA ACA GTT CAT ACA GAG TCT TTT ACG ACT CAA TGA CAA GAA GAA AAT CTT CGT CAA
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAT GGT GGA GCA CGA CAC TCT GGT CTA CTC CAA AAA TGT CAA AGA TAC AGT CTC AGA AGA
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCA AAG GGC TAT TGA GAC TTT TCA ACA AAG GAT AAT TTC GGG AAA CCT CCT CGG ATT CCA
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TTG CCC AGC TAT CTG TCA CTT CAT CGA AAG GAC AGT AGA AAA GGA AGG TGG CTC CTA CAA
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATG CCA TCA TTG CGA TAA AGG AAA GGC TAT CAT TCA AGA TCT CTC TGC CGA CAG TGG TCC
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA AGA TGG ACC CCC ACC CAC GAG GAG CAT CGT GGA AAA AGA AGA CGT TCC AAC CAC GTC
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TTC AAA GCA AGT GGA TTG ATG TGA CAT CTC CAC TGA CGT AAG GGA TGA CGC ACA ATC CCA
                          CaMV35S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CTA TCC TTC GCA AGA CCC TTC CTC TAT ATA AGG AAG TTC ATT TCA TTT GGA GAG GAC ACG
                                          5' DGAT1 (S205A)
                                  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
                                 NcoI
                                ~~~~~~~~
                                   M   A   I   L   D   S   A   G   V   T   T   V
GGA TCC TTG CTC CGT TAA AAA AAA CCA TGG CTA TCC TCG ATT CTG CTG GTG TTA CTA CTG
                      5' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
·  T   E   N   G   G   G   E   F   V   D   L   D   R   L   R   R   R   K   S   R
TGA CTG AGA ATG GTG GTG GAG AGT TCG TTG ATC TCG ATA GAC TCA GAA GAA GAA AGT CTA
5' DGAT1 (S205A)
~~~~~~
                               UBQ10
      ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
·  S
GAT CTG TAA ATT TCT GTG TTC CTT ATT CTC TCA AAA TCT TCG ATT TTG TTT TCG TTC GAT
                             UBQ10
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC AAT TTC GTA TAT GTT CTT TGG TTT AGA TTC TGT TAA TCT TAG ATC GAA GAC GAT TTT
```

UBQ10

~~~~~~
CTG GGT TTG ATC GTT AGA TAT CAT CTT AAT TCT CGA TTA GGG TTT CAT AGA TAT CAT CCG
UBQ10
~~~~~~
ATT TGT TCA AAT AAT TTG AGT TTT GTC GAA TAA TTA CTC TTC GAT TTG TGA TTT CTA TCT
UBQ10
~~~~~~
AGA TCT GGT GTT AGT TTC TAG TTT GTG CGA TCG AAT TTG TCG ATT AAT CTG AGT TTT TCT
UBQ10
~~~~~~~~~~~~
3' DGAT1 (S205A)
~~~~~~
D S S N G L L L S G S D N N S P S
GAT TAA CAG GAT TCT TCT AAC GGA CTT CTC CTC TCT GGA TCT GAT AAC AAC TCT CCT TCT
3' DGAT1 (S205A)
~~~~~~
D D V G A P A D V R D R I D S V V N D D
GAT GAT GTT GGT GCT CCT GCT GAT GTG AGA GAT AGA ATC GAT TCT GTG GTG AAC GAT GAT
3' DGAT1 (S205A)
~~~~~~
A Q G T A N L A G D N N G G G D N N G G
GCT CAA GGT ACT GCT AAC CTC GCT GGT GAT AAT AAC GGT GGA GGT GAT AAC AAT GGT GGA
3' DGAT1 (S205A)
~~~~~~
G R G G G E G R G N A D A T F T Y R P S
GGA AGA GGT GGA GGT GAA GGT AGA GGA AAC GCT GAT GCT ACT TTC ACT TAC AGA CCA TCT
3' DGAT1 (S205A)
~~~~~~
V P A H R R A R E S P L S S D A I F K Q
GTG CCT GCT CAT AGA AGA GCT AGA GAG TCT CCT CTC TCT TCT GAT GCT ATC TTC AAG CAG
3' DGAT1 (S205A)
~~~~~~
S H A G L F N L C V V V L I A V N S R L
TCT CAC GCT GGA CTT TTC AAC CTC TGT GTG GTG GTT CTT ATC GCT GTG AAC TCT AGA CTC
3' DGAT1 (S205A)
~~~~~~
I I E N L M K Y G W L I R T D F W F S S
ATC ATC GAG AAC CTC ATG AAG TAC GGA TGG CTC ATC AGA ACT GAT TTC TGG TTC TCT TCT
3' DGAT1 (S205A)
~~~~~~
R S L R D W P L F M C C I S L S I F P L
AGA TCT CTC AGA GAT TGG CCT CTT TTC ATG TGC TGC ATC TCA CTC TCA ATC TTC CCT CTC
3' DGAT1 (S205A)
~~~~~~
A A F T V E K L V L Q K Y I A E P V V I
GCT GCT TTT ACT GTT GAG AAG CTC GTG CTC CAG AAG TAT ATC GCT GAA CCT GTG GTG ATC
3' DGAT1 (S205A)
~~~~~~
F L H I I I T M T E V L Y P V Y V T L R

```
TTC CTC CAC ATC ATC ATC ACT ATG ACT GAG GTT CTC TAC CCT GTT TAC GTG ACT CTC AGA
                      3' DGAT1 (S205A)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 C   D   S   A   F   L   S   G   V   T   L   M   L   L   T   C   I   V   W   L
TGC GAT TCT GCT TTC CTC TCT GGT GTT ACT CTT ATG CTC CTC ACT TGC ATT GTG TGG CTT
                      3' DGAT1 (S205A)
 ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 K   L   V   S   Y   A   H   T   S   Y   D   I   R   S   L   A   N   A   A   D
AAG CTC GTG TCT TAC GCT CAC ACT TCT TAC GAT ATC AGA TCT CTC GCT AAC GCT GCT GAT
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 K   A   N   P   E   V   S   Y   Y   V   S   L   K   S   L   A   Y   F   M   V
AAG GCT AAC CCT GAA GTG TCT TAC TAC GTG TCT CTC AAG TCT CTC GCT TAC TTC ATG GTT
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   P   T   L   C   Y   Q   P   S   Y   P   R   S   A   C   I   R   K   G   W
GCT CCT ACA CTT TGT TAC CAG CCA TCT TAC CCT AGA TCT GCT TGC ATT AGA AAG GGA TGG
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 V   A   R   Q   F   A   K   L   V   I   F   T   G   F   M   G   F   I   I   E
GTG GCA AGA CAA TTC GCT AAG TTG GTG ATC TTC ACT GGA TTC ATG GGA TTC ATC ATC GAG
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 Q   Y   I   N   P   I   V   R   N   S   K   H   P   L   K   G   D   L   L   Y
CAG TAC ATC AAC CCT ATT GTG AGA AAC TCT AAG CAC CCT CTC AAG GGT GAT CTT CTC TAC
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 A   I   E   R   V   L   K   L   S   V   P   N   L   Y   V   W   L   C   M   F
GCT ATC GAG AGA GTT CTT AAG CTC TCT GTG CCT AAC CTT TAT GTG TGG CTC TGC ATG TTC
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 Y   C   F   F   H   L   W   L   N   I   L   A   E   L   L   C   F   G   D   R
TAC TGT TTC TTC CAC CTC TGG CTT AAC ATC CTT GCT GAG TTG CTT TGC TTC GGA GAT AGA
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 E   F   Y   K   D   W   W   N   A   K   S   V   G   D   Y   W   R   M   W   N
GAG TTC TAC AAG GAT TGG TGG AAC GCT AAG TCT GTT GGA GAT TAT TGG AGA ATG TGG AAC
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 M   P   V   H   K   W   M   V   R   H   I   Y   F   P   C   L   R   S   K   I
ATG CCT GTG CAT AAG TGG ATG GTG CGT CAC ATC TAC TTC CCT TGC CTC AGA TCT AAG ATC
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 P   K   T   L   A   I   I   I   A   F   L   V   S   A   V   F   H   E   L   C
CCT AAG ACT CTC GCT ATC ATT ATC GCT TTC CTC GTG TCT GCT GTT TTC CAT GAG TTG TGT
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 I   A   V   P   C   R   L   F   K   L   W   A   F   L   G   I   M   F   Q   V
ATC GCT GTT CCT TGC AGA CTT TTC AAG CTT TGG GCT TTC CTC GGA ATC ATG TTC CAG GTT
                      3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

```
 P   L   V   F   I   T   N   Y   L   Q   E   R   F   G   S   T   V   G   N   M
CCA CTC GTG TTC ATC ACT AAC TAC CTC CAA GAG AGA TTC GGA TCT ACT GTT GGA AAC ATG
                    3' DGAT1 (S205A)

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 I   F   W   F   I   F   C   I   F   G   Q   P   M   C   V   L   L   Y   Y   H
ATT TTC TGG TTC ATT TTC TGC ATC TTC GGA CAG CCT ATG TGC GTT CTC CTC TAC TAC CAC
         3' DGAT1 (S205A)
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
 D   L   M   N   R   K   G   S   M   S   *   *
GAT CTC ATG AAC AGA AAG GGA TCT ATG TCT TAA TGA AGG ATC CAC CCA GCT TTC TTG TAC
                                                                  attB2
                                                                 ocs
                                                              ~~~~~~~~~~~~~~~~~~~~
AAA GTG GTG ATG GGT TCG AAA TCG ATA AGC TTG GAT CCT CTA GAG TCC TGC TTT AAT GAG
~~~~~~~~~
 attB2
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
ATA TGC GAG ACG CCT ATG ATC GCA TGA TAT TTG CTT TCA ATT CTG TTG TGC ACG TTG TAA
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
AAA ACC TGA GCA TGT GTA GCT CAG ATC CTT ACC GCC GGT TTC GGT TCA TTC TAA TGA ATA
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAT CAC CCG TTA CTA TCG TAT TTT TAT GAA TAA TAT TCT CCG TTC AAT TTA CTG ATT GTA
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC TAC TAC TTA TAT GTA CAA TAT TAA AAT GAA AAC AAT ATA TTG TGC TGA ATA GGT TTA
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TAG CGA CAT CTA TGA TAG AGC GCC ACA ATA ACA AAC AAT TGC GTT TTA TTA TTA CAA ATC
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CAA TTT TAA AAA AAG CGG CAG AAC CGG TCA AAC CTA AAA GAC TGA TTA CAT AAA TCT TAT
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
TCA AAT TTC AAA AGG CCC CAG GGG CTA GTA TCT ACG ACA CAC CGA GCG GCG AAC TAA TAA
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGT TCA CTG AAG GGA ACT CCG GTT CCC CGC CGG CGC GCA TGG GTG AGA TTC CTT GAA GTT
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
GAG TAT TGG CCG TCC GCT CTA CCG AAA GTT ACG GGC ACC ATT CAA CCC GGT CCA GCA CGG
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CGG CCG GGT AAC CGA CTT GCT GCC CCG AGA ATT ATG CAG CAT TTT TTT GGT GTA TGT GGG
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
CCC CAA ATG AAG TGC AGG TCA AAC CTT GAC AGT GAC GAC AAA TCG TTG GGC GGG TCC AGG
                           ocs
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
```

GCG AAT TTT GCG ACA ACA TGT CGA GGC TCA GCA GGA CCT GCA GGC ATG CAA GCT AGC TTA

ocs

~~~~~~

CTA GTG ATG CAT ATT CTA TAG TGT CAC CTA AAT CT
~~~~~~

Figura 6

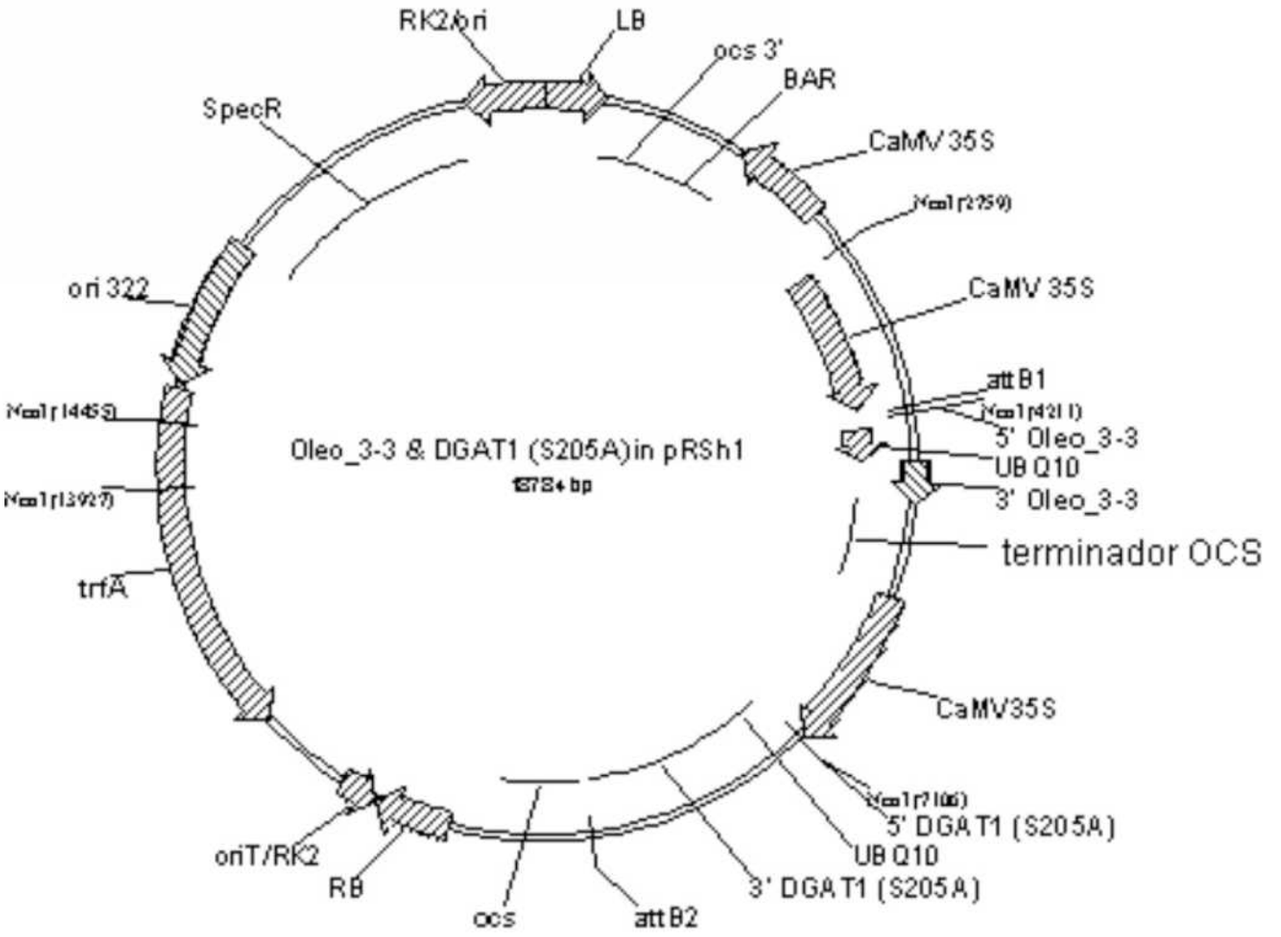

RK2/ori
LB
ocs 3'
BAR
SpecR
CaMV 35S
CaMV 35S
ori 322
attB1
5' Oleo_3-3
UB Q10
3' Oleo_3-3
Oleo_3-3 & DGAT1 (S205A)in pRSh1
terminador OCS
trfA
CaMV35S
5' DGAT1 (S205A)
UB Q10
3' DGAT1 (S205A)
oriT/RK2
RB
ocs
attB2

Figura 7

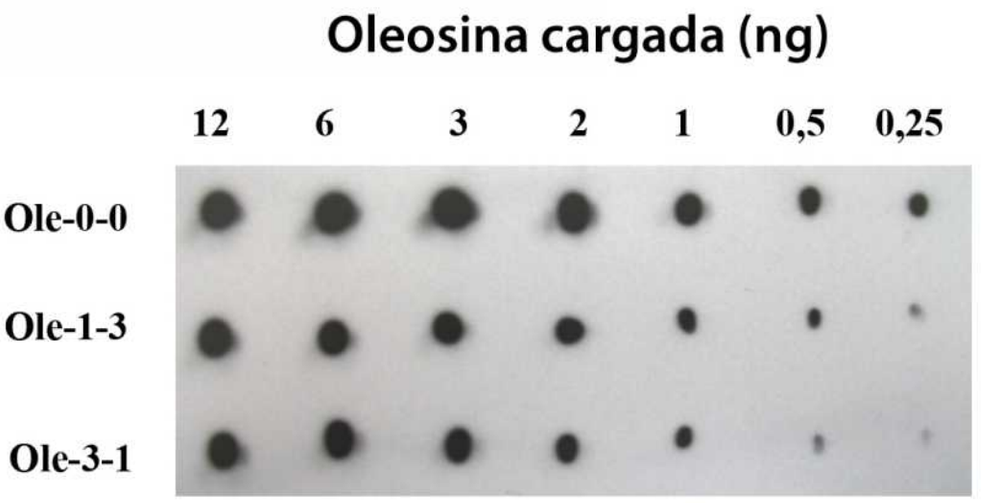


Figura 8

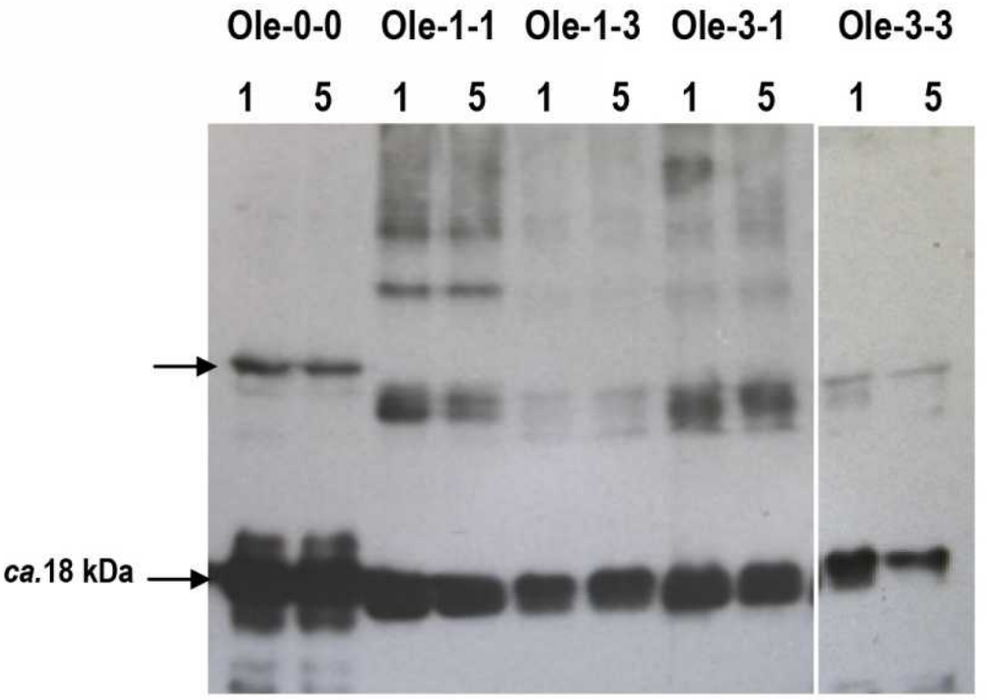


Figura 9

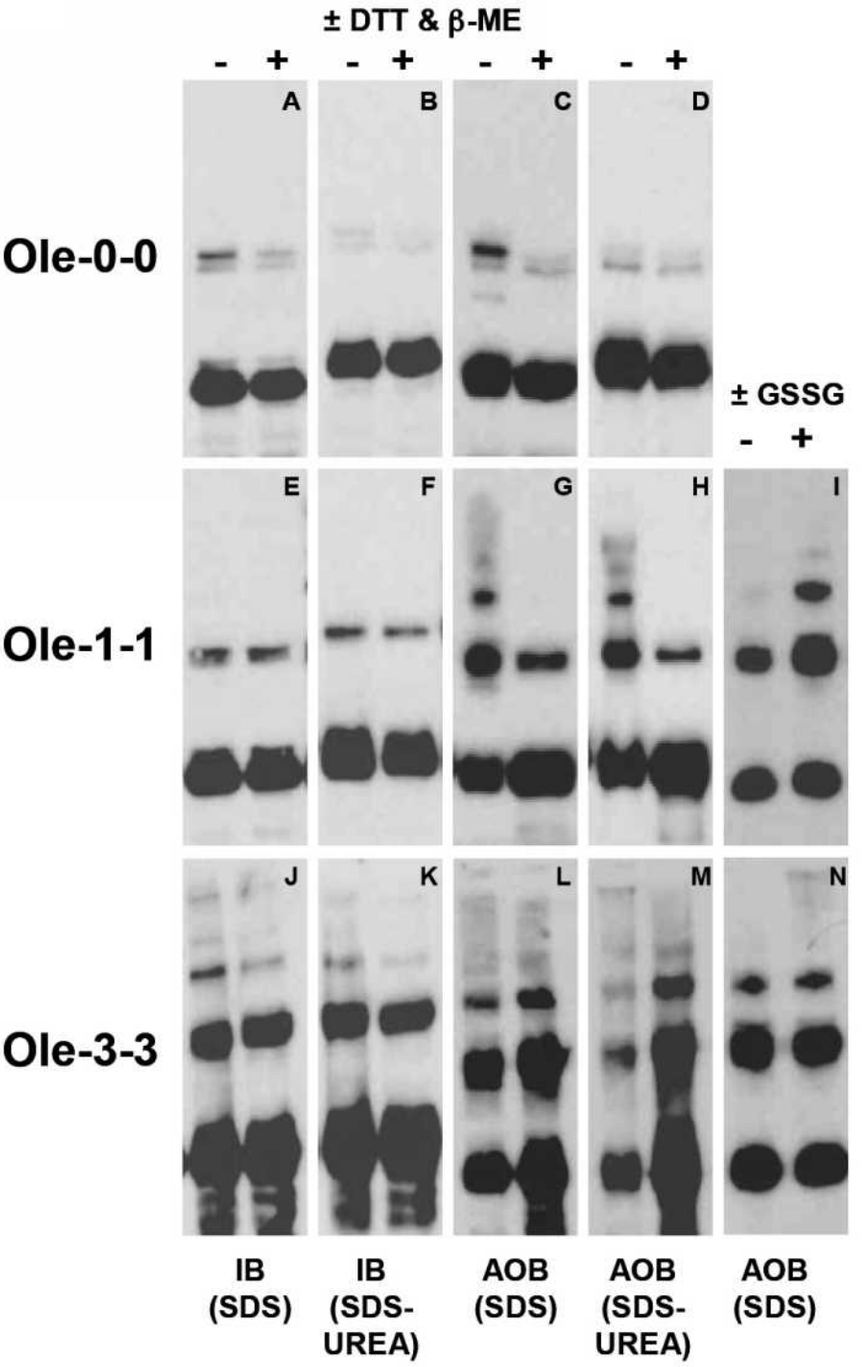

± DTT & β-ME
- + - + - + - +
A
B
C
D
Ole-0-0
± GSSG
- +
E
F
G
H
I
Ole-1-1
J
K
L
M
N
Ole-3-3
IB (SDS)
IB (SDS-UREA)
AOB (SDS)
AOB (SDS-UREA)
AOB (SDS)

Figura 10

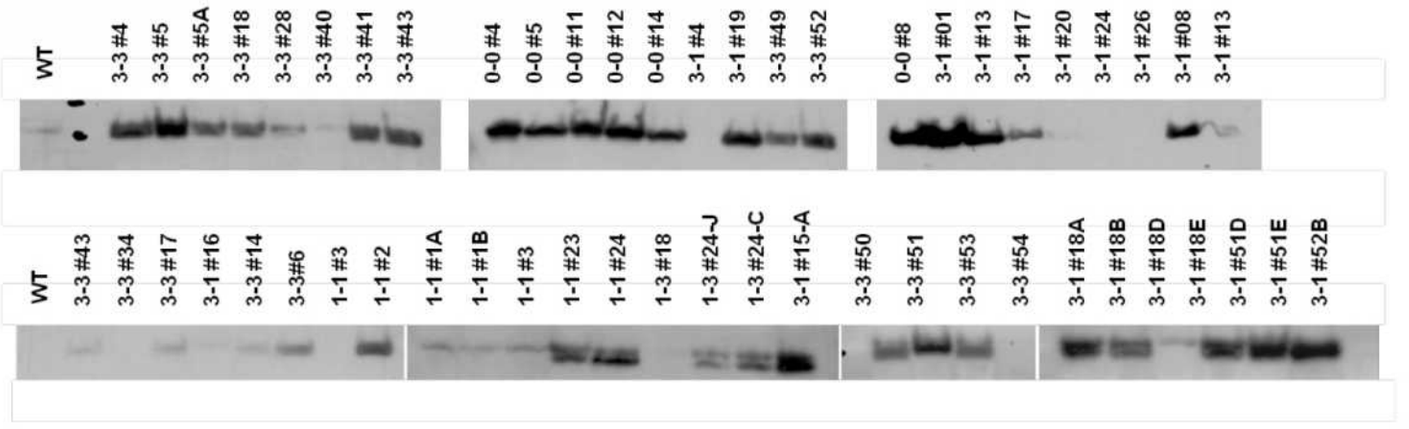

WT
3-3 #4
3-3 #5
3-3 #5A
3-3 #18
3-3 #28
3-3 #40
3-3 #41
3-3 #43
0-0 #4
0-0 #5
0-0 #11
0-0 #12
0-0 #14
3-1 #4
3-1 #19
3-3 #49
3-3 #52
0-0 #8
3-1 #01
3-1 #13
3-1 #17
3-1 #20
3-1 #24
3-1 #26
3-1 #08
3-1 #13
WT
3-3 #43
3-3 #34
3-3 #17
3-1 #16
3-3 #14
3-3#6
1-1 #3
1-1 #2
1-1 #1A
1-1 #1B
1-1 #3
1-1 #23
1-1 #24
1-3 #18
1-3 #24-J
1-3 #24-C
3-1 #15-A
3-3 #50
3-3 #51
3-3 #53
3-3 #54
3-1 #18A
3-1 #18B
3-1 #18D
3-1 #18E
3-1 #51D
3-1 #51E
3-1 #52B

Figura 11

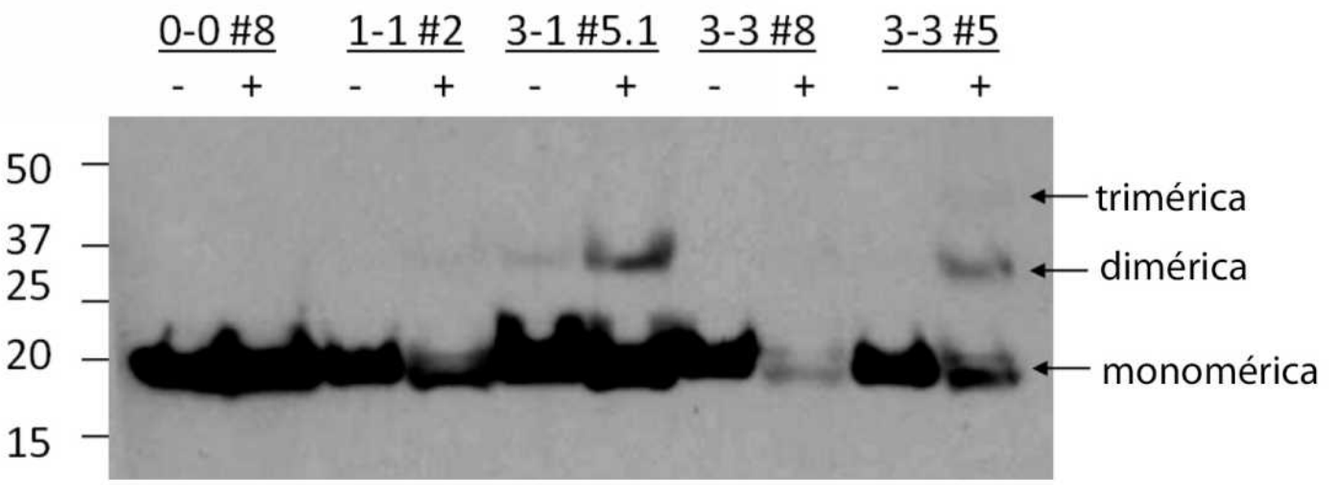

0-0 #8
1-1 #2
3-1 #5.1
3-3 #8
3-3 #5
- + - + - + - + - +
50
37
25
20
15
trimérica
dimérica
monomérica

Figura 12

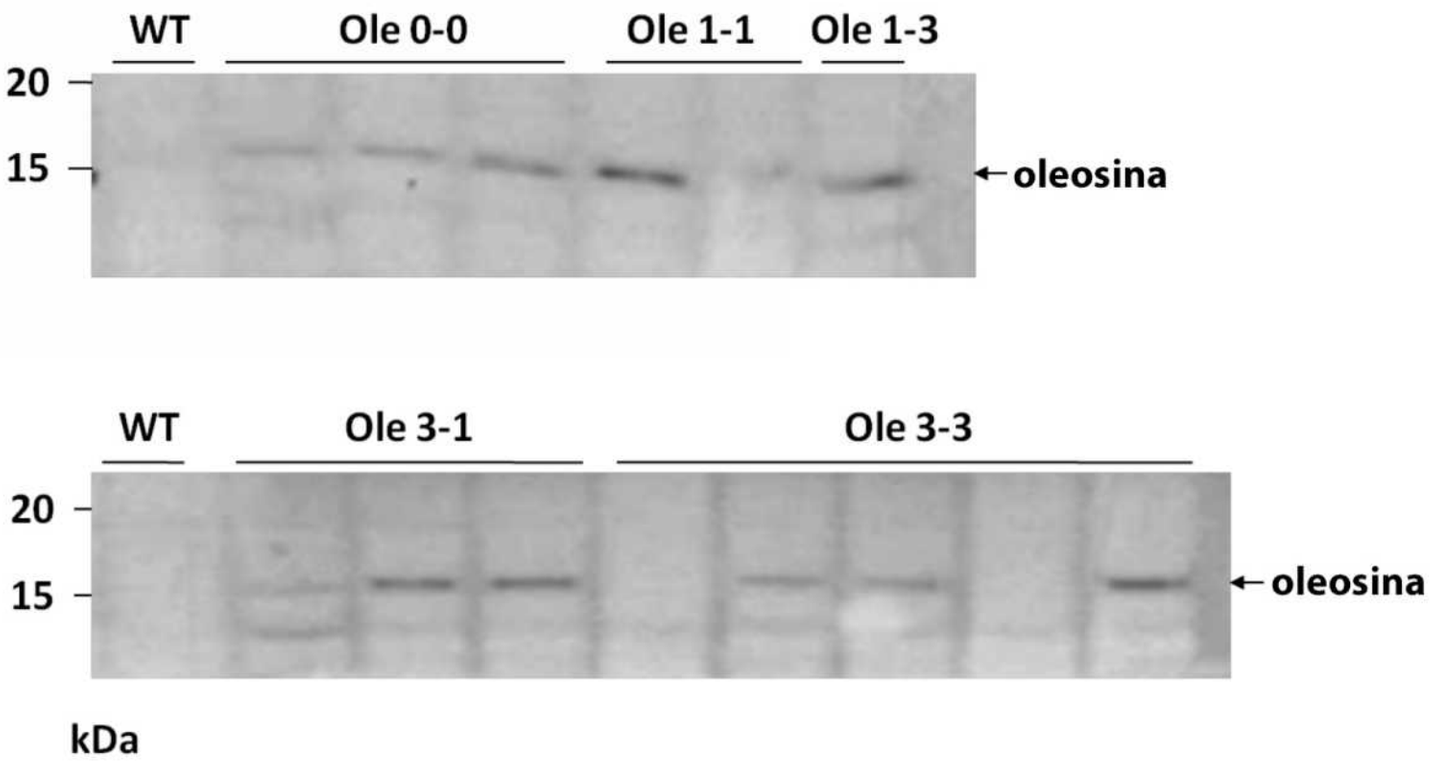

WT
Ole 0-0
Ole 1-1
Ole 1-3
20
15
oleosina
WT
Ole 3-1
Ole 3-3
20
15
oleosina
kDa

Figura 13

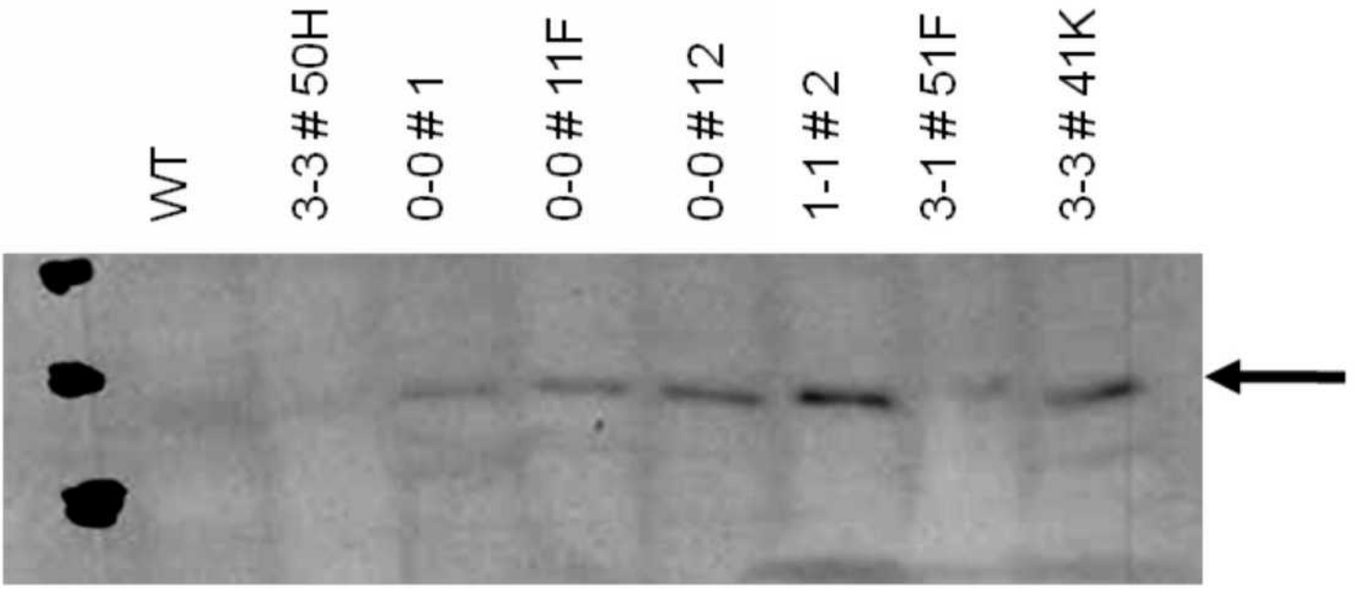


Figura 14

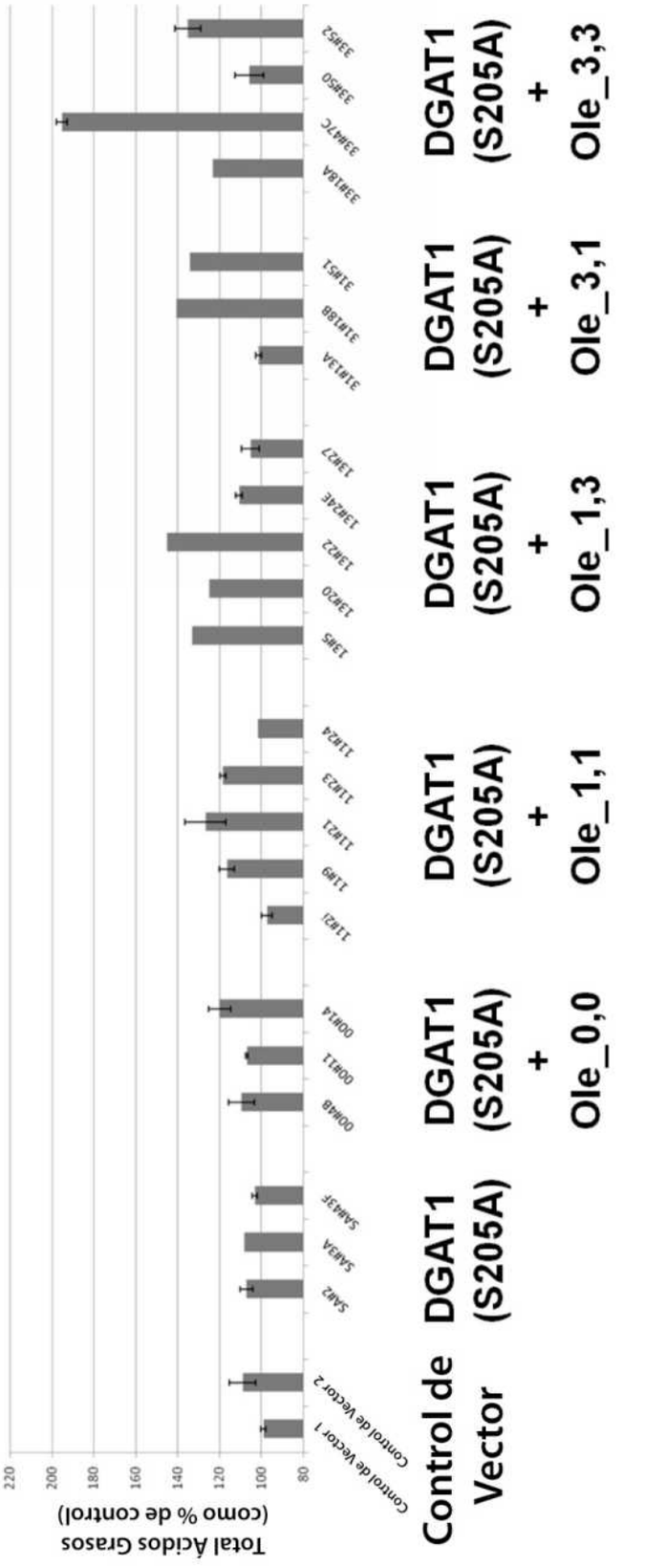

Total Ácidos Grasos
(como % de control)
220
200
180
160
140
120
100
80
Control de Vector
DGAT1 (S205A)
DGAT1 (S205A) + Ole_0,0
DGAT1 (S205A) + Ole_1,1
DGAT1 (S205A) + Ole_1,3
DGAT1 (S205A) + Ole_3,1
DGAT1 (S205A) + Ole_3,3

Figura 15

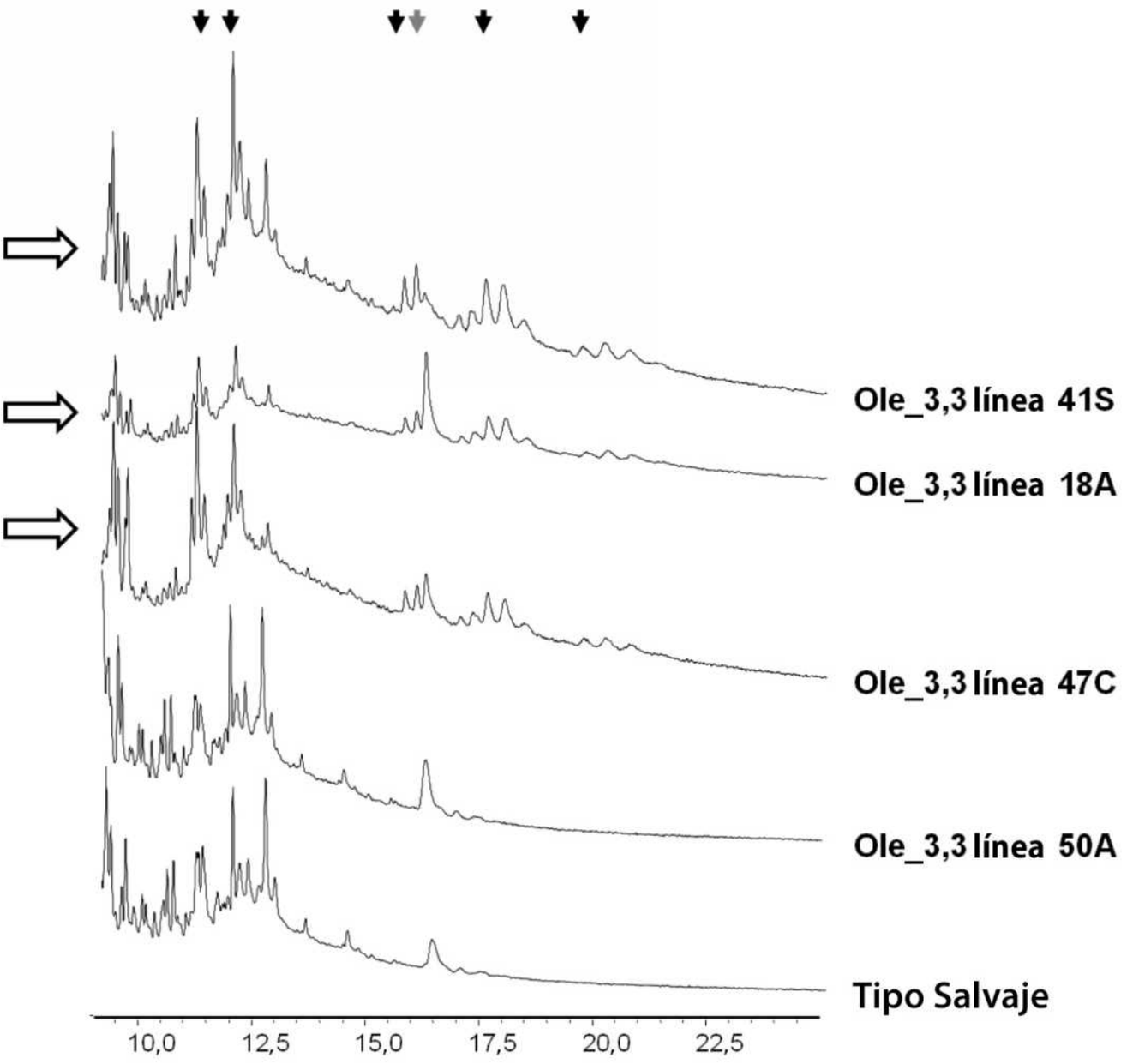
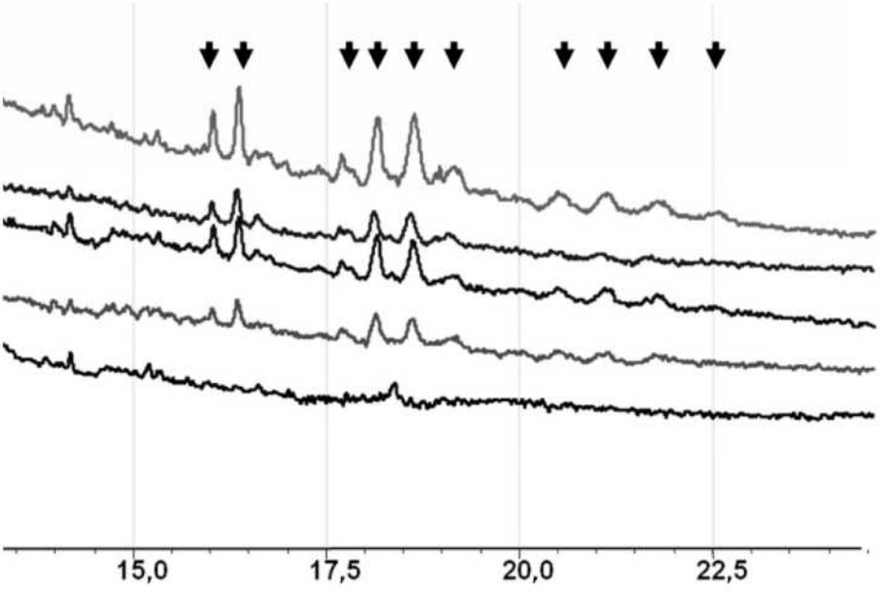

Ole_3,3 línea 41S
Ole_3,3 línea 18A
Ole_3,3 línea 47C
Ole_3,3 línea 50A
Tipo Salvaje
10,0
12,5
15,0
17,5
20,0
22,5

Figura 16



Ole _3,3. 5 semanas después de la germinación, (97 mg DW)

Ole _3,3. 4 semanas después de la germinación, (101 mg DW)

Ole _3,3. 3 semanas después de la germinación, (81 mg DW)

Ole _3,3. 2 semanas después de la germinación, (75 mg DW)

Tipo Salvaje. 5 semanas después de la germinación, (100 mg DW)

Figura 17

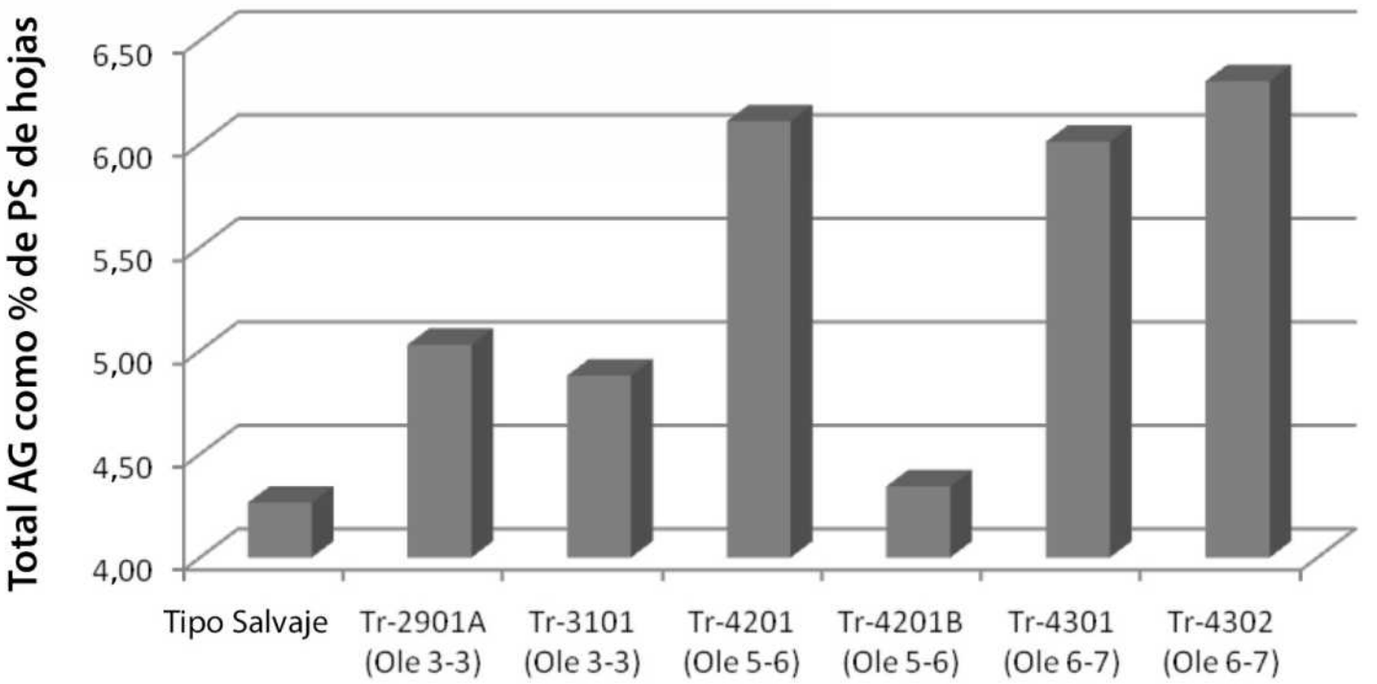

Total AG como % de PS de hojas
6,50
6,00
5,50
5,00
4,50
4,00
Tipo Salvaje
Tr-2901A (Ole 3-3)
Tr-3101 (Ole 3-3)
Tr-4201 (Ole 5-6)
Tr-4201B (Ole 5-6)
Tr-4301 (Ole 6-7)
Tr-4302 (Ole 6-7)

Figura 18

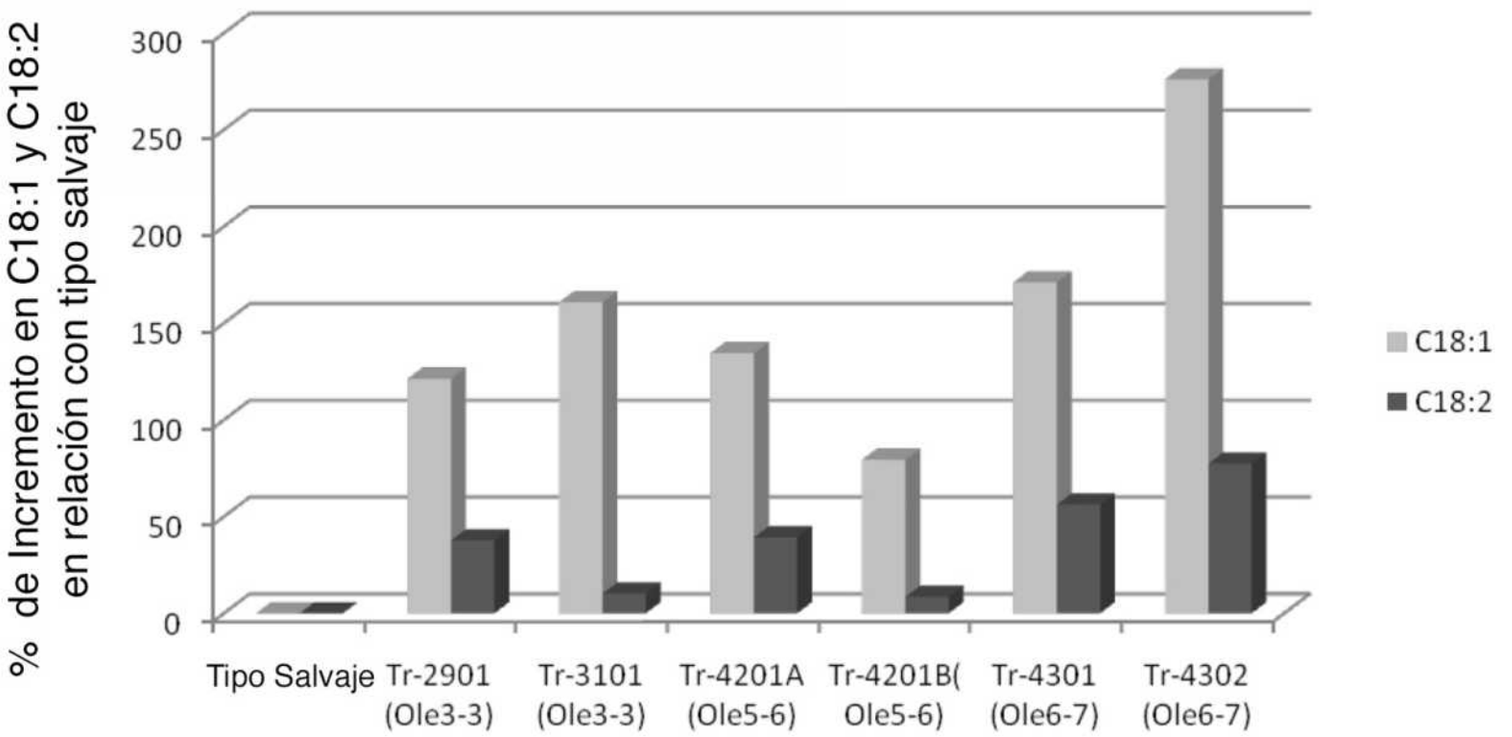

% de Incremento en C18:1 y C18:2 en relación con tipo salvaje
300
250
200
150
100
50
0
Tipo Salvaje
Tr-2901 (Ole3-3)
Tr-3101 (Ole3-3)
Tr-4201A (Ole5-6)
Tr-4201B(Ole5-6)
Tr-4301 (Ole6-7)
Tr-4302 (Ole6-7)
C18:1
C18:2

Figura 19

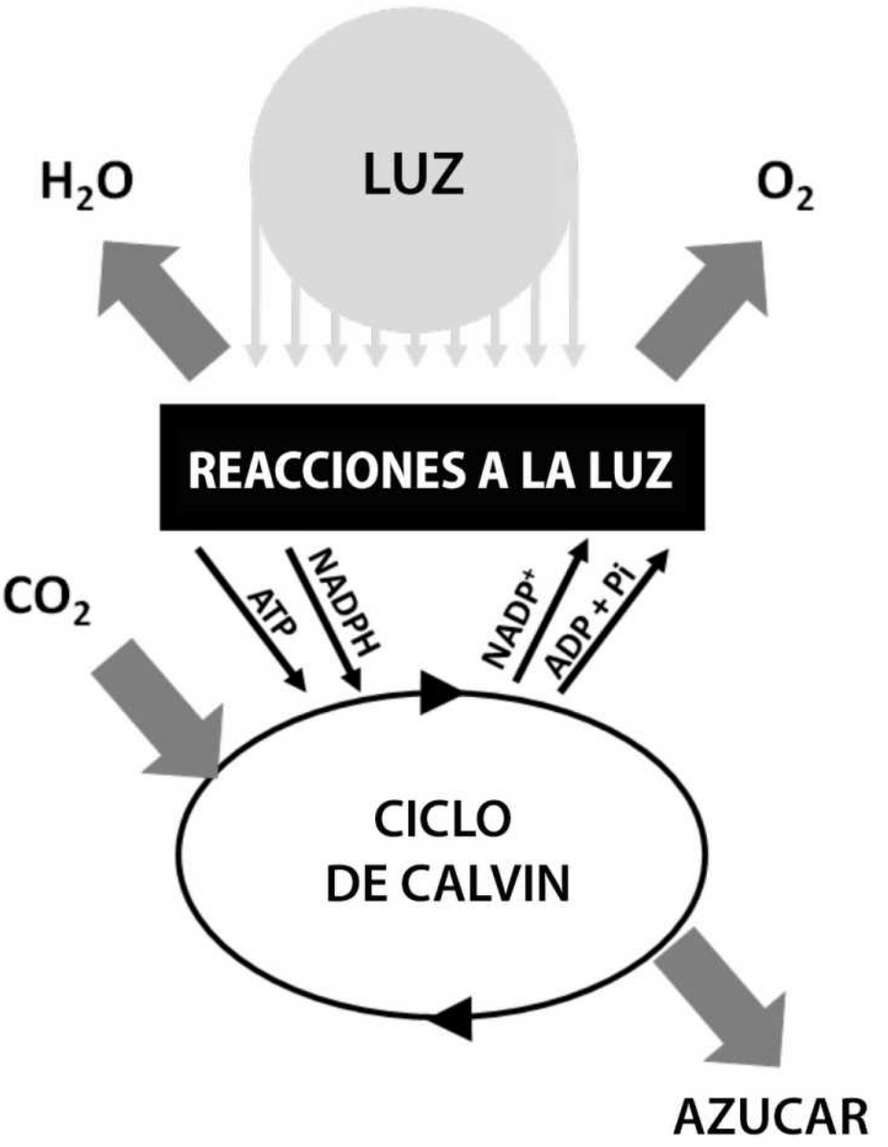

H_2O
LUZ
O_2
REACCIONES A LA LUZ
CO_2
ATP
NADPH
$NADP^+$
ADP + Pi
CICLO
DE CALVIN
AZUCAR

Figura 20

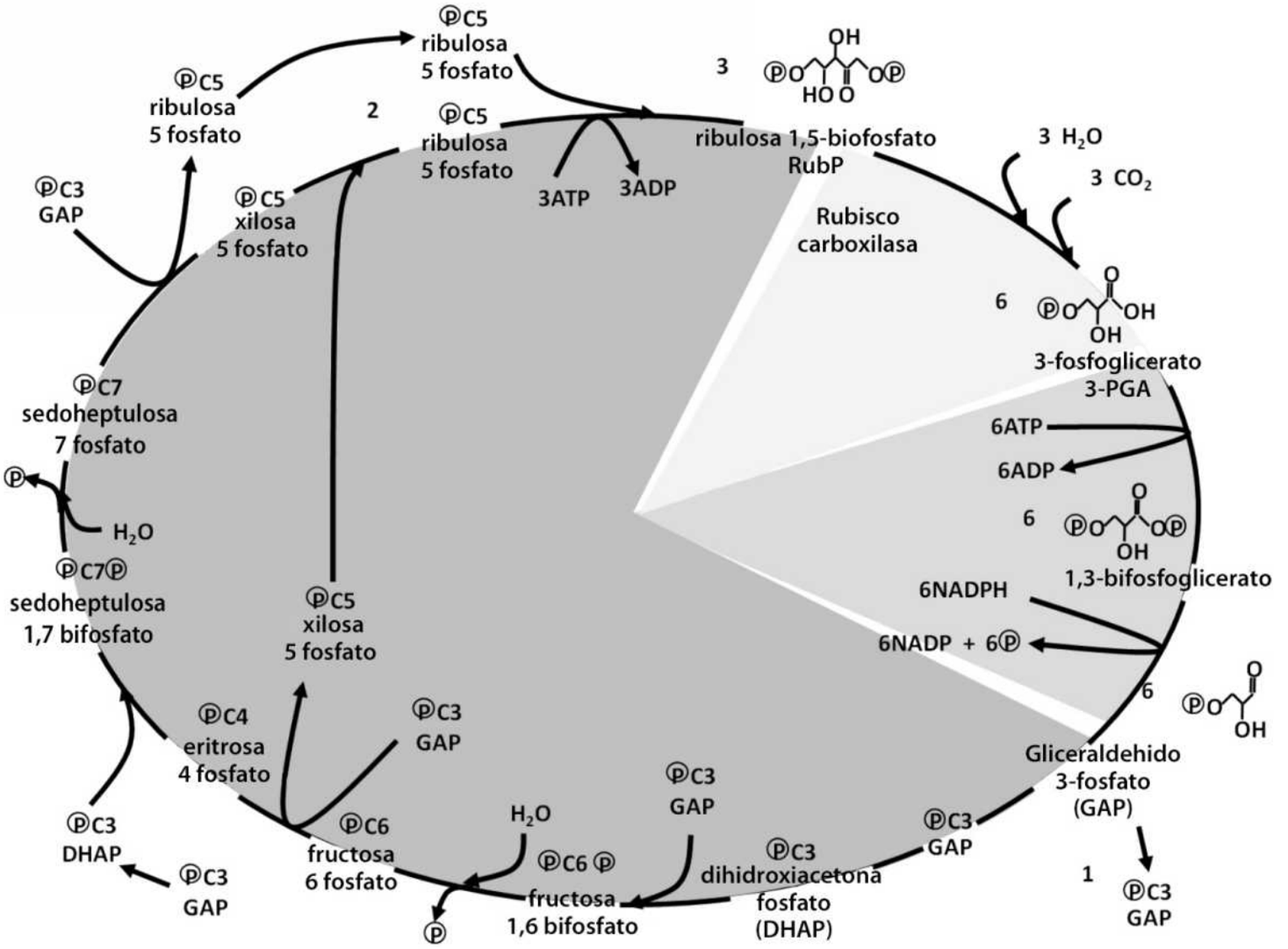

Ⓟ C5
ribulosa
5 fosfato
3
OH
Ⓟ O
HO O
O Ⓟ
2
Ⓟ C5
ribulosa
5 fosfato
Ⓟ C5
ribulosa
5 fosfato
ribulosa 1,5-biofosfato
RubP
3 H2O
3 CO2
Rubisco
carboxilasa
3ATP
3ADP
Ⓟ C3
GAP
Ⓟ C5
xilosa
5 fosfato
6
Ⓟ O
OH
OH
3-fosfoglicerato
3-PGA
6ATP
6ADP
Ⓟ C7
sedoheptulosa
7 fosfato
Ⓟ
H2O
6
Ⓟ O
O Ⓟ
OH
1,3-bifosfoglicerato
Ⓟ C7 Ⓟ
sedoheptulosa
1,7 bifosfato
6NADPH
6NADP + 6Ⓟ
Ⓟ C5
xilosa
5 fosfato
6
Ⓟ O
OH
Ⓟ C4
eritrosa
4 fosfato
Ⓟ C3
GAP
Gliceraldehido
3-fosfato
(GAP)
Ⓟ C3
GAP
H2O
Ⓟ C3
GAP
Ⓟ C3
DHAP
Ⓟ C6
fructosa
6 fosfato
Ⓟ C6 Ⓟ
fructosa
1,6 bifosfato
Ⓟ C3
dihidroxiacetona
fosfato
(DHAP)
Ⓟ C3
GAP
1
Ⓟ C3
GAP
Ⓟ
Ⓟ C3
GAP

Figura 21

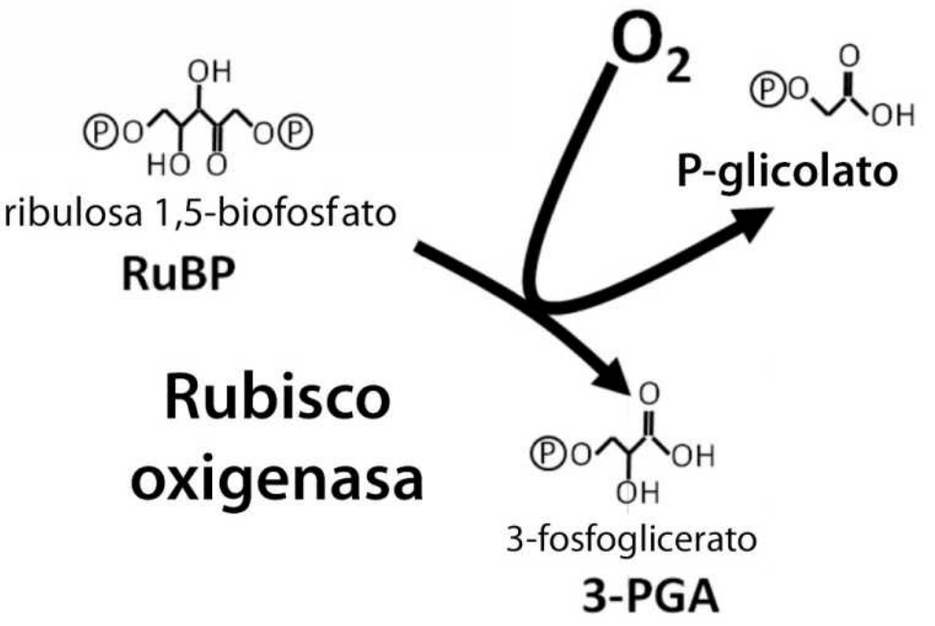

OH
PO
OP
HO O
ribulosa 1,5-biofosfato
RuBP
O2
O
PO
OH
P-glicolato
Rubisco
oxigenasa
O
PO
OH
OH
3-fosfoglicerato
3-PGA

Figura 22

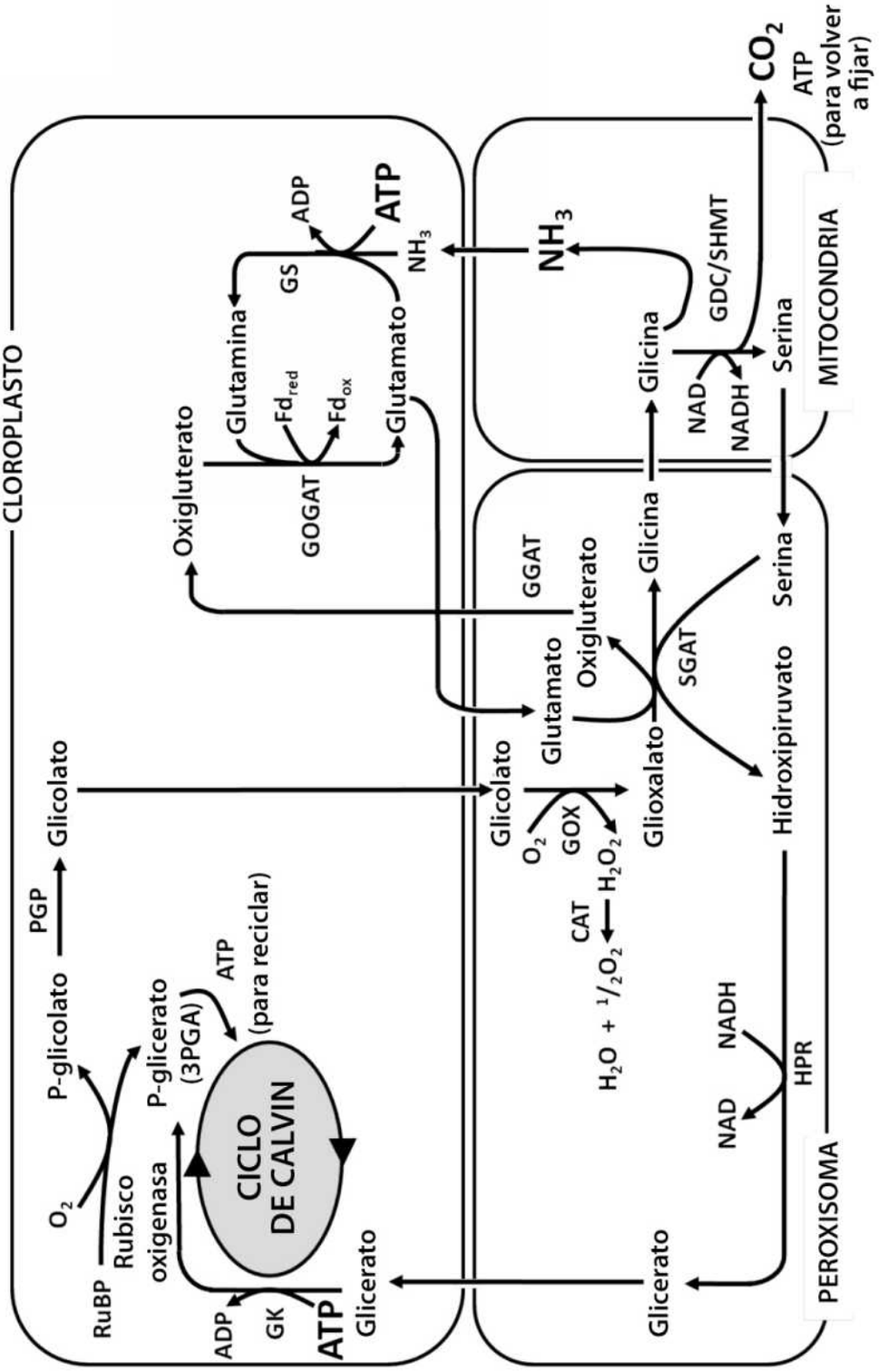

CLOROPLASTO
PEROXISOMA
MITOCONDRIA
CICLO DE CALVIN
RuBP
Rubisco oxigenasa
O_2
P-glicolato
PGP
Glicolato
P-glicerato (3PGA)
ATP (para reciclar)
ADP
GK
ATP
Glicerato
Oxigluterato
Glutamina
Fd_{red}
Fd_{ox}
GOGAT
Glutamato
GS
ADP
ATP
NH_3
GGAT
Glicolato
O_2
GOX
H_2O_2
CAT
$H_2O + {}^1/_2O_2$
Glioxalato
Glutamato
Oxigluterato
SGAT
Glicina
Serina
Hidroxipiruvato
HPR
NADH
NAD
Glicerato
NH_3
Glicina
GDC/SHMT
NAD
NADH
Serina
CO_2
ATP (para volver a fijar)

Figura 23

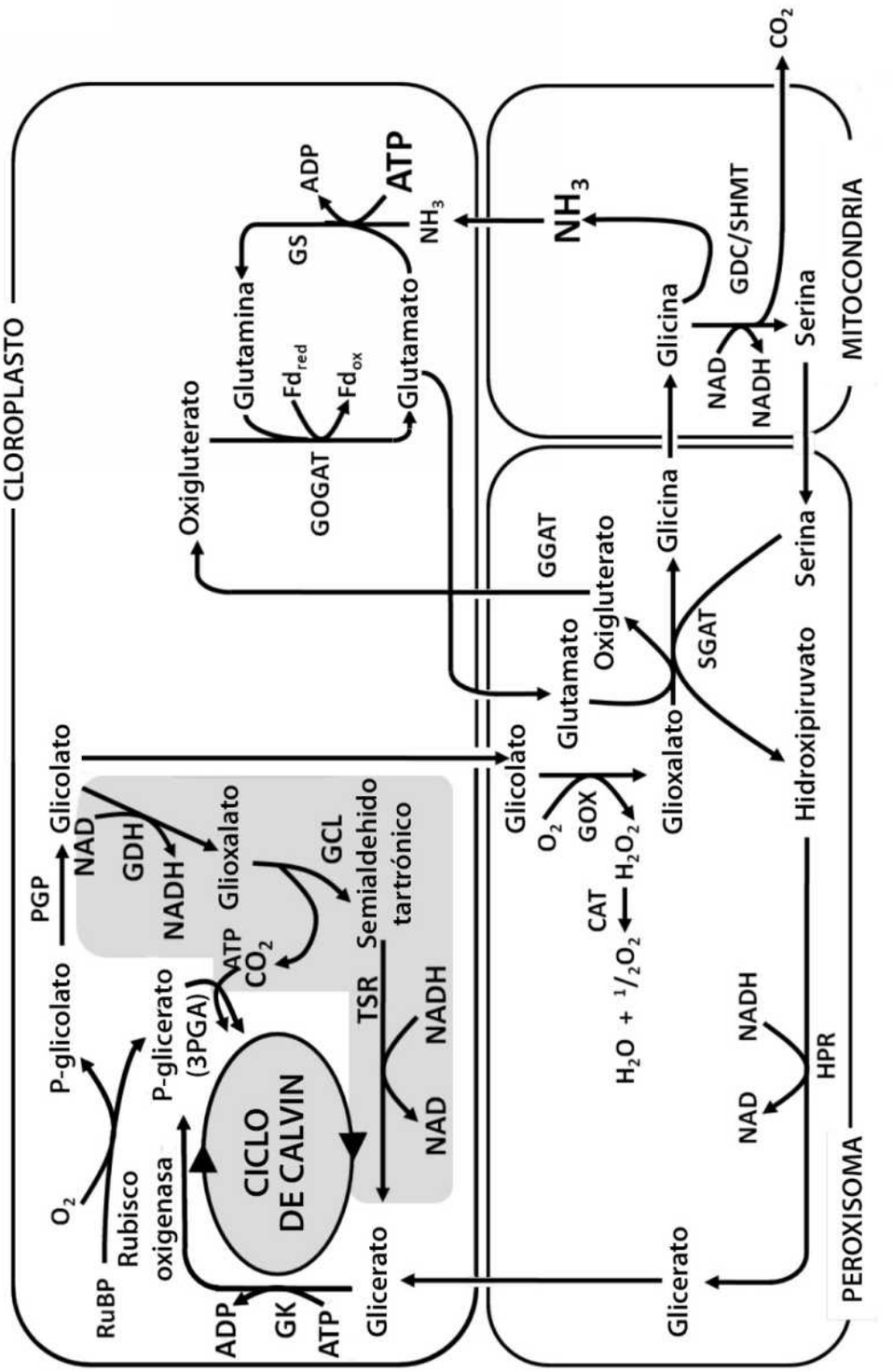
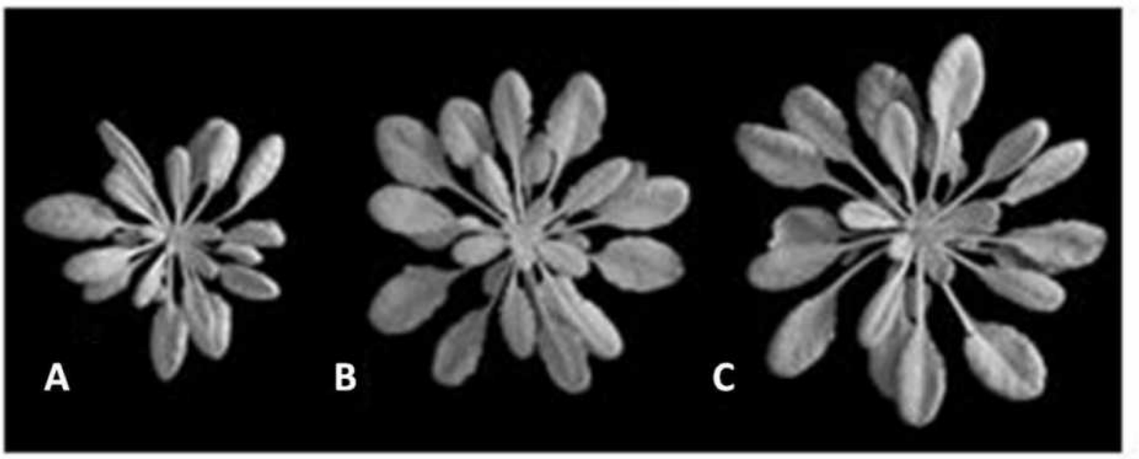

CLOROPLASTO
CICLO DE CALVIN
RuBP
Rubisco oxigenasa
O_2
P-glicolato
PGP
Glicolato
NAD
GDH
NADH
Glioxalato
GCL
Semialdehido tartrónico
ATP
CO_2
P-gliceratо (3PGA)
TSR
NADH
NAD
Glicerato
ADP
GK
ATP
Oxigluterato
Glutamina
Fd_{red}
Fd_{ox}
GOGAT
GS
ADP
ATP
NH_3
Glutamato
GGAT
Glicolato
O_2
GOX
H_2O_2
CAT
$H_2O + {}^1/_2O_2$
Glioxalato
Glutamato
Oxigluterato
SGAT
Glicina
Serina
Hidroxipiruvato
HPR
NADH
NAD
Glicerato
PEROXISOMA
NH_3
Glicina
GDC/SHMT
CO_2
NAD
NADH
Serina
MITOCONDRIA

Figura 24



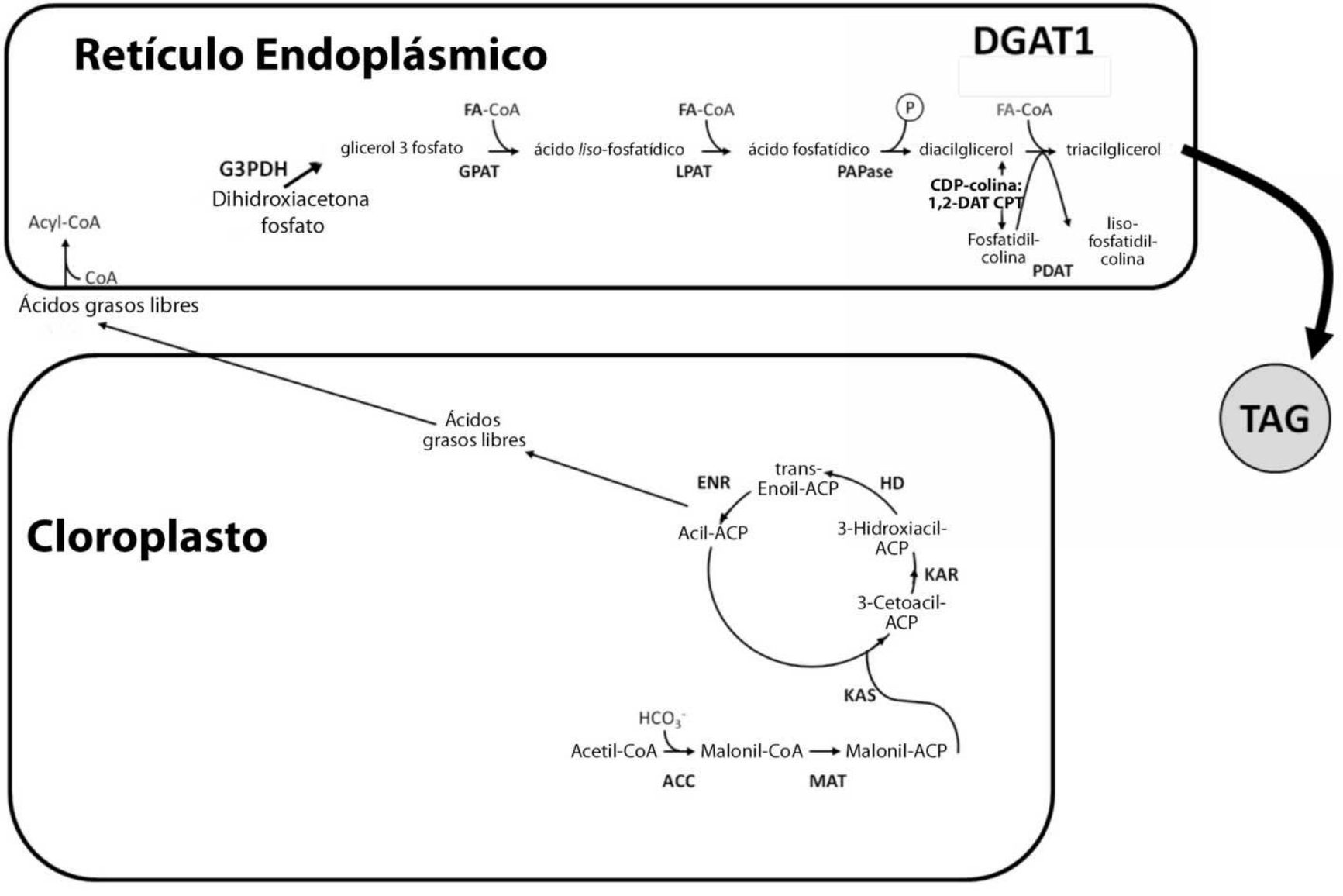

Retículo Endoplásmico
DGAT1
FA-CoA
FA-CoA
P
FA-CoA
glicerol 3 fosfato
ácido liso-fosfatídico
ácido fosfatídico
diacilglicerol
triacilglicerol
G3PDH
Dihidroxiacetona fosfato
GPAT
LPAT
PAPase
CDP-colina: 1,2-DAT CPT
Acyl-CoA
CoA
Fosfatidil-colina
liso-fosfatidil-colina
PDAT
Ácidos grasos libres
TAG
Ácidos grasos libres
Cloroplasto
ENR
trans-Enoil-ACP
HD
Acil-ACP
3-Hidroxiacil-ACP
KAR
3-Cetoacil-ACP
KAS
HCO3-
Acetil-CoA
Malonil-CoA
Malonil-ACP
ACC
MAT

Figura 25

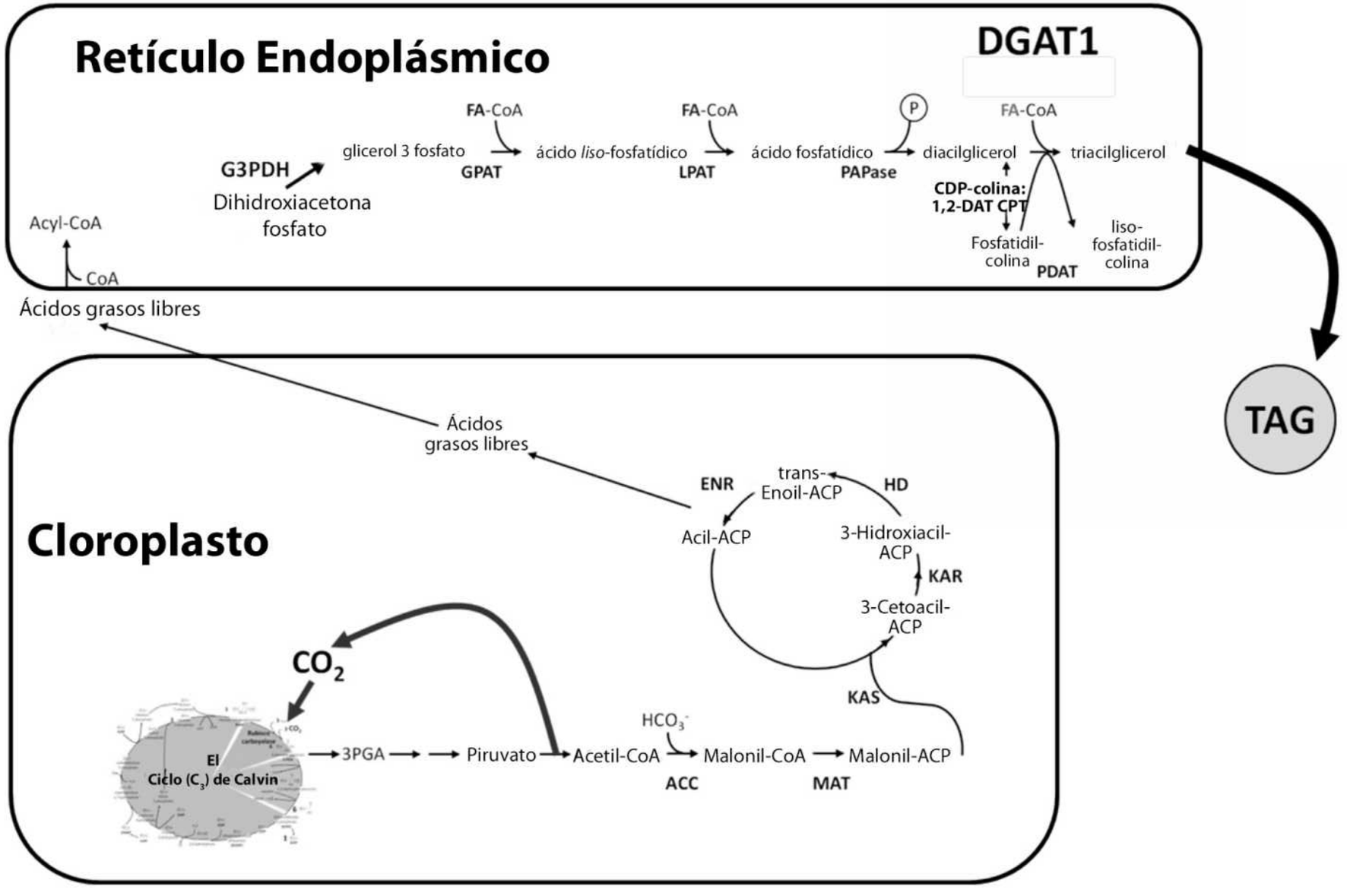

Retículo Endoplásmico
DGAT1
FA-CoA
FA-CoA
P
FA-CoA
glicerol 3 fosfato
ácido liso-fosfatídico
ácido fosfatídico
diacilglicerol
triacilglicerol
G3PDH
GPAT
LPAT
PAPase
Dihidroxiacetona fosfato
CDP-colina: 1,2-DAT CPT
Acyl-CoA
CoA
Fosfatidil-colina
liso-fosfatidil-colina
PDAT
Ácidos grasos libres
TAG
Ácidos grasos libres
ENR
trans-Enoil-ACP
HD
Cloroplasto
Acil-ACP
3-Hidroxiacil-ACP
KAR
3-Cetoacil-ACP
KAS
CO2
HCO3-
El Ciclo (C3) de Calvin
3PGA
Piruvato
Acetil-CoA
Malonil-CoA
Malonil-ACP
ACC
MAT

Figura 26

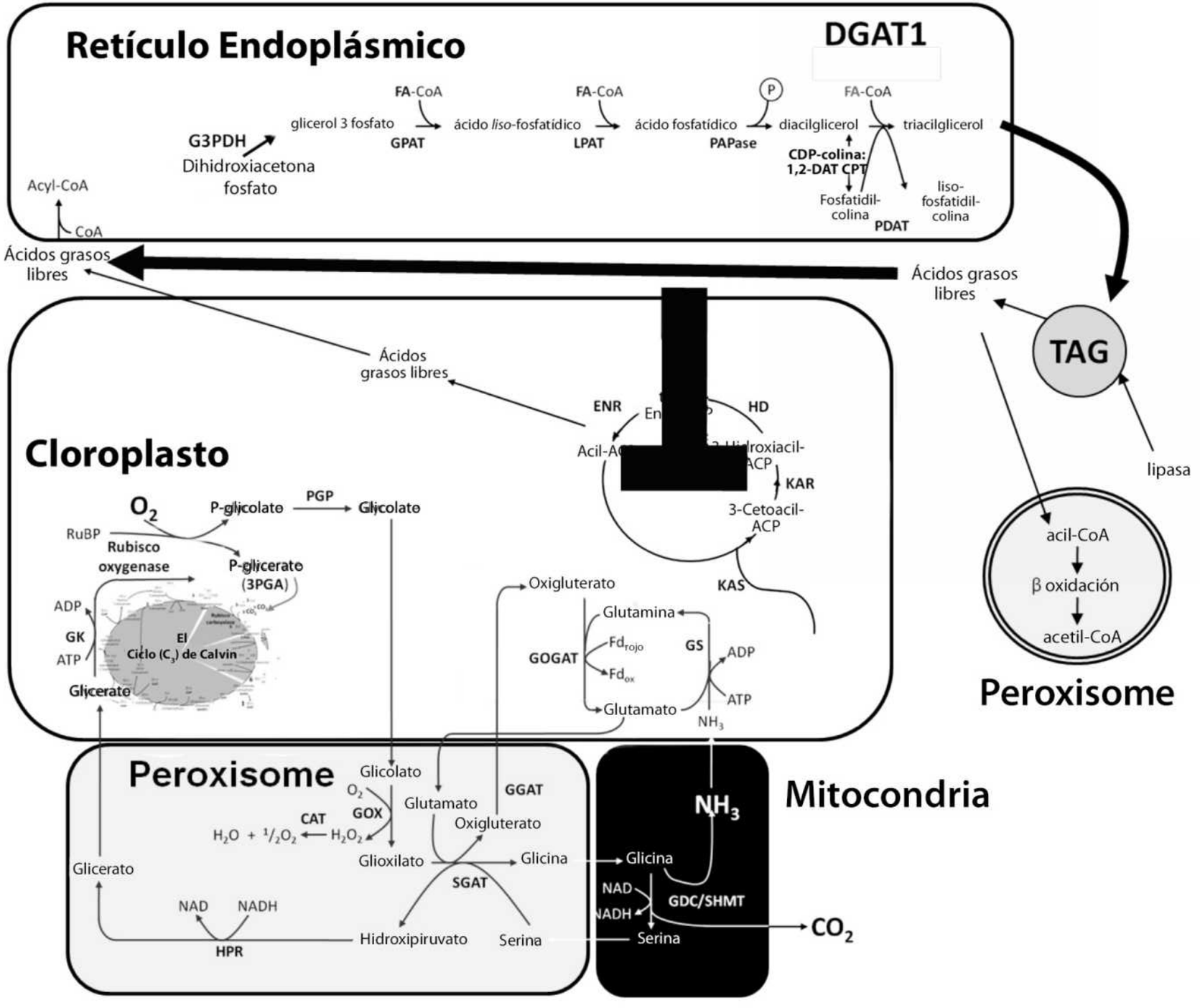

Retículo Endoplásmico
DGAT1
FA-CoA
FA-CoA
P
FA-CoA
glicerol 3 fosfato
ácido liso-fosfatídico
ácido fosfatídico
diacilglicerol
triacilglicerol
G3PDH
Dihidroxiacetona fosfato
GPAT
LPAT
PAPase
CDP-colina: 1,2-DAT CPT
Acyl-CoA
CoA
Fosfatidil-colina
liso-fosfatidil-colina
PDAT
Ácidos grasos libres
Ácidos grasos libres
TAG
Ácidos grasos libres
ENR
HD
KAR
3-Cetoacil-ACP
KAS
Cloroplasto
O2
P-glicolato
PGP
Glicolato
RuBP
Rubisco oxygenase
P-glicerato (3PGA)
ADP
GK
ATP
El Ciclo (C3) de Calvin
Glicerato
Oxigluterato
Glutamina
GOGAT
Fdrojo
Fdox
GS
ADP
ATP
Glutamato
NH3
lipasa
acil-CoA
β oxidación
acetil-CoA
Peroxisome
Peroxisome
Glicolato
O2
GOX
CAT
H2O + 1/2O2
H2O2
Glutamato
GGAT
Oxigluterato
Glioxilato
Glicina
SGAT
Glicerato
NAD
NADH
HPR
Hidroxipiruvato
Serina
Mitocondria
NH3
Glicina
NAD
NADH
GDC/SHMT
CO2
Serina

Figura 27

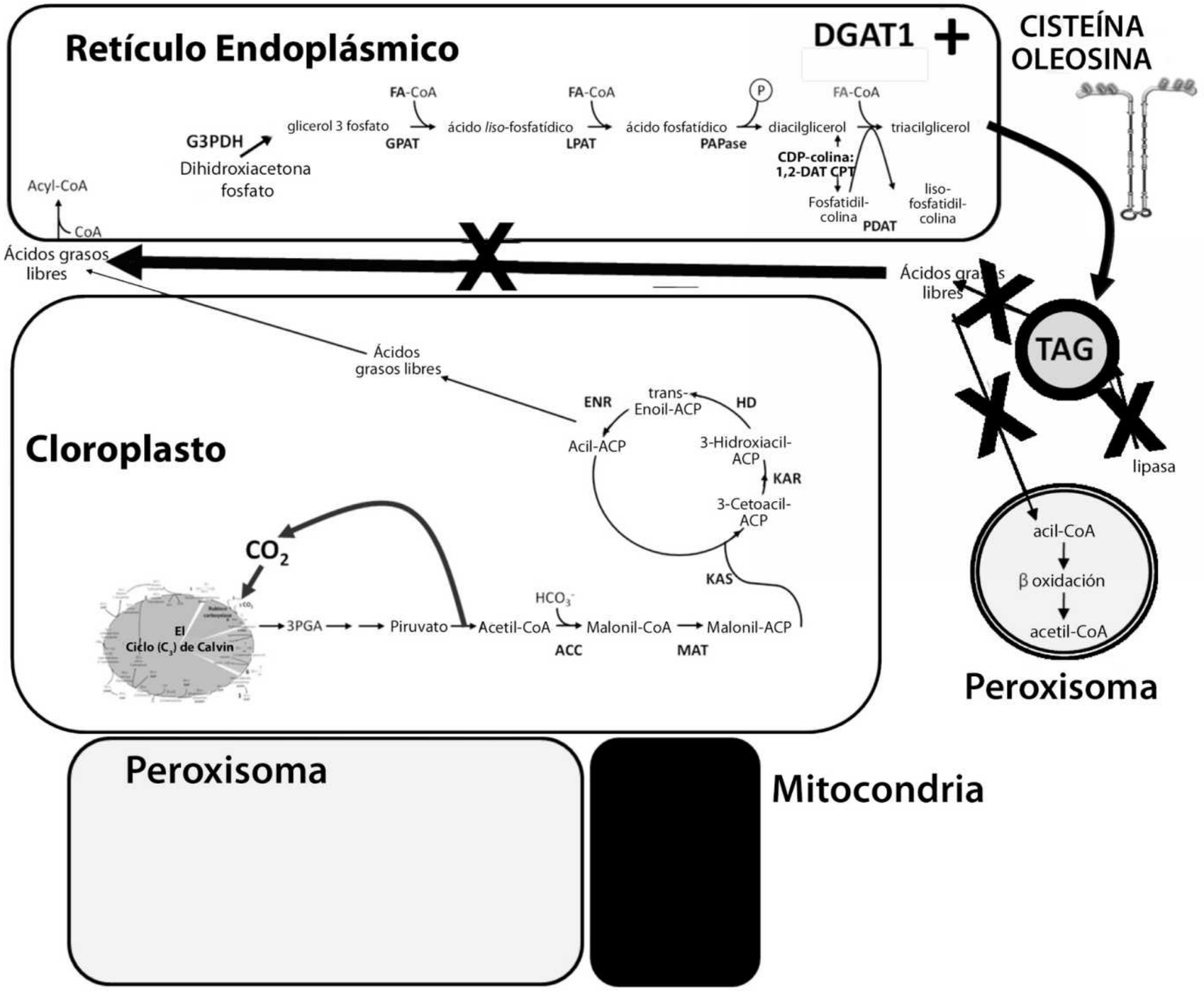

Retículo Endoplásmico
DGAT1 +
CISTEÍNA
OLEOSINA
FA-CoA
FA-CoA
FA-CoA
glicerol 3 fosfato
GPAT
ácido liso-fosfatídico
LPAT
ácido fosfatídico
PAPase
diacilglicerol
triacilglicerol
G3PDH
Dihidroxiacetona fosfato
CDP-colina:
1,2-DAT CPT
Fosfatidil-colina
PDAT
liso-fosfatidil-colina
Acyl-CoA
CoA
Ácidos grasos libres
Ácidos grasos libres
TAG
lipasa
Ácidos grasos libres
Cloroplasto
ENR
trans-Enoil-ACP
HD
Acil-ACP
3-Hidroxiacil-ACP
KAR
3-Cetoacil-ACP
KAS
CO_2
El Ciclo (C_3) de Calvin
3PGA
Piruvato
Acetil-CoA
HCO_3^-
ACC
Malonil-CoA
MAT
Malonil-ACP
acil-CoA
β oxidación
acetil-CoA
Peroxisoma
Peroxisoma
Mitocondria

Figura 28

Figura 29



WT | Control de Vector | DGAT1 | DGAT1 + Ole_0,0 | DGAT1 + Ole_1,1 | DGAT1 + Ole_1,3 | DGAT1 + Ole_3,1 | DGAT1 + Ole_3,3

Figura 30

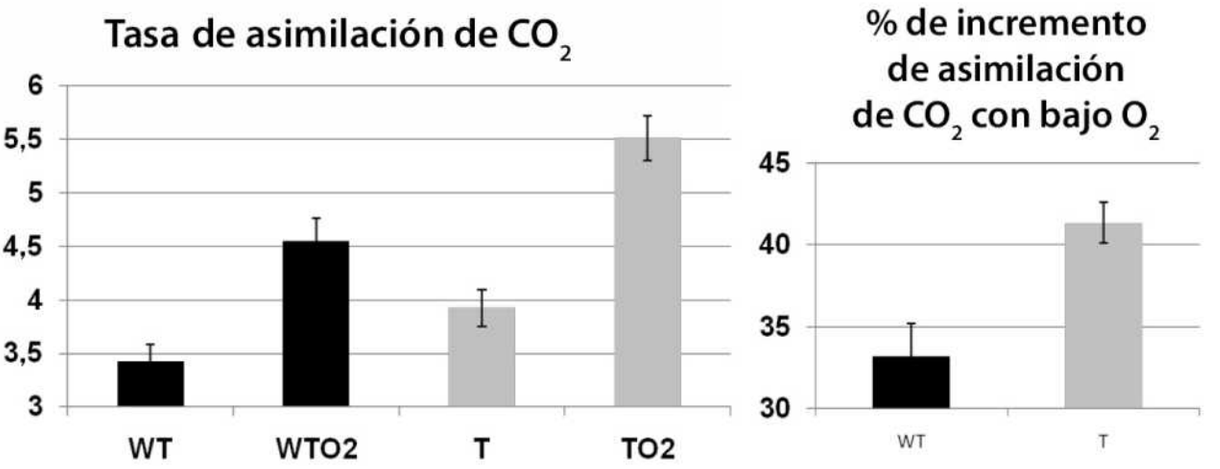

Tasa de asimilación de CO_2
6
5,5
5
4,5
4
3,5
3
WT
WTO2
T
TO2
% de incremento de asimilación de CO_2 con bajo O_2
45
40
35
30
WT
T

Figura 31

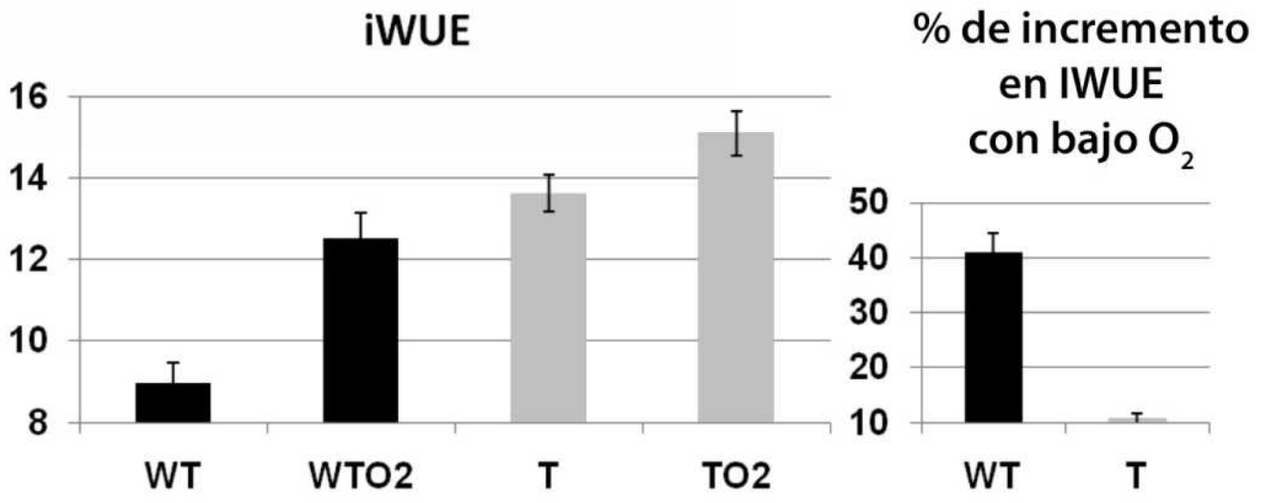

iWUE
16
14
12
10
8
WT
WTO2
T
TO2
% de incremento en IWUE con bajo O_2
50
40
30
20
10
WT
T

Figura 32

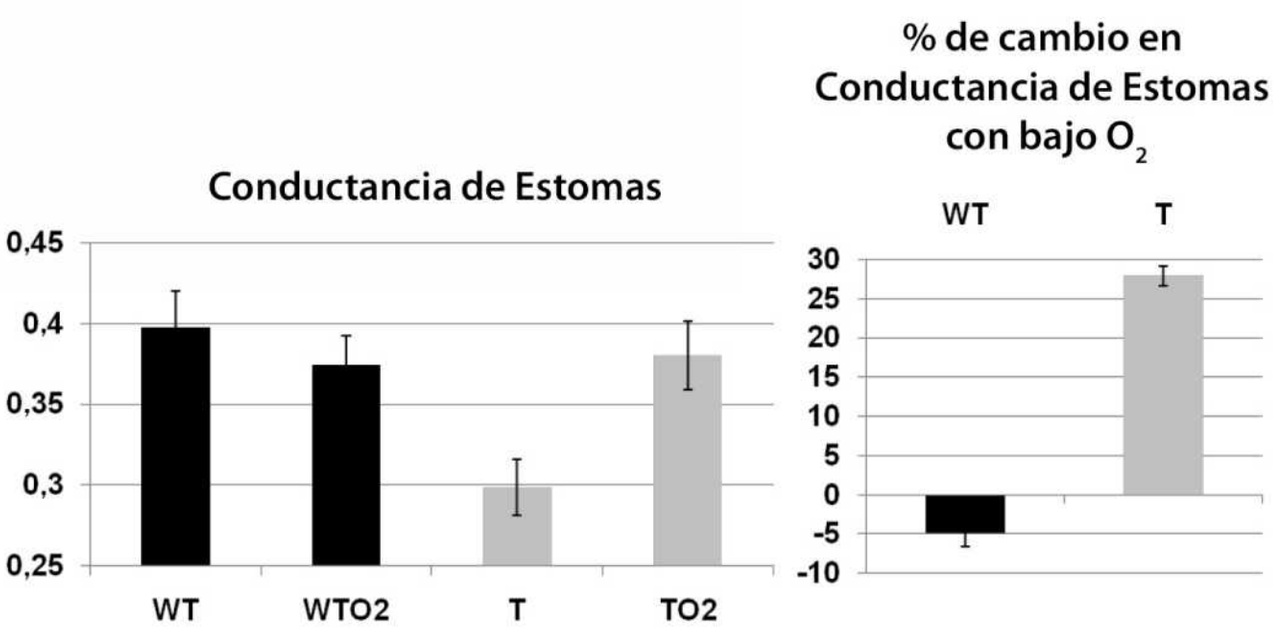

Conductancia de Estomas
0,45
0,4
0,35
0,3
0,25
WT
WTO2
T
TO2
% de cambio en Conductancia de Estomas con bajo O_2
WT
T
30
25
20
15
10
5
0
-5
-10

Figura 33

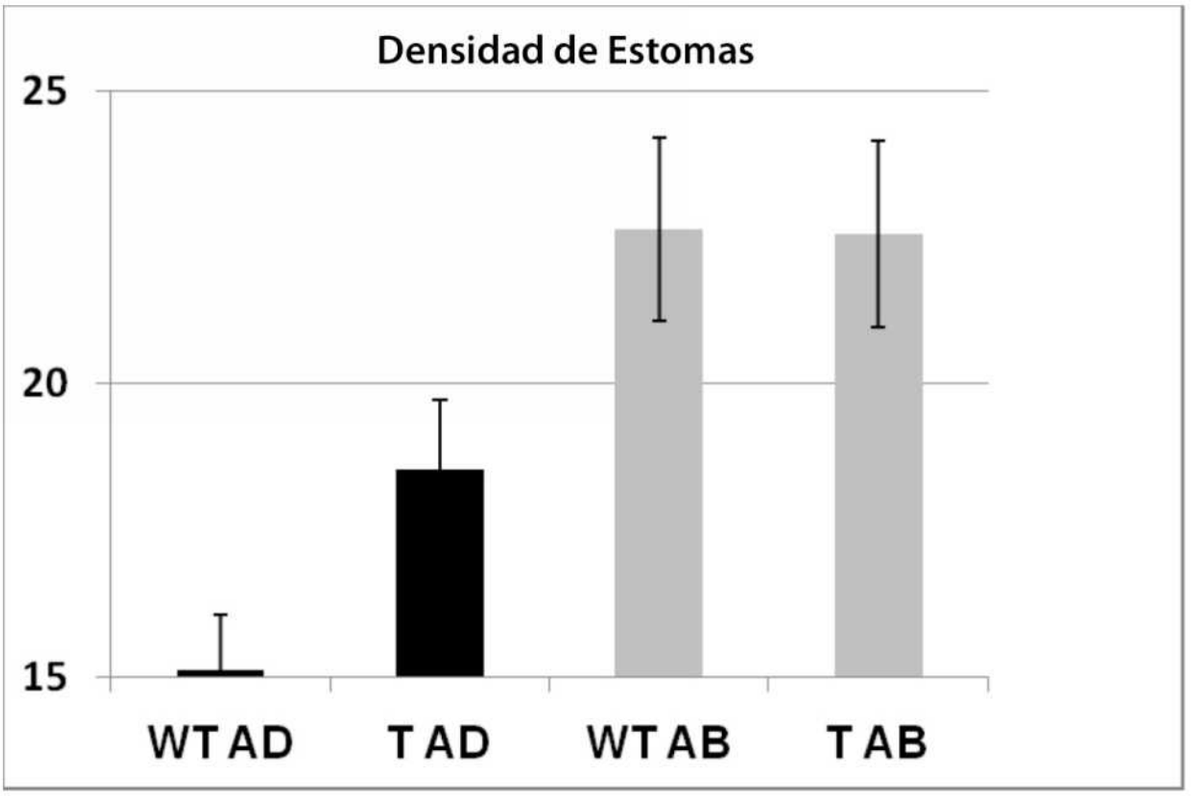


Figura 34

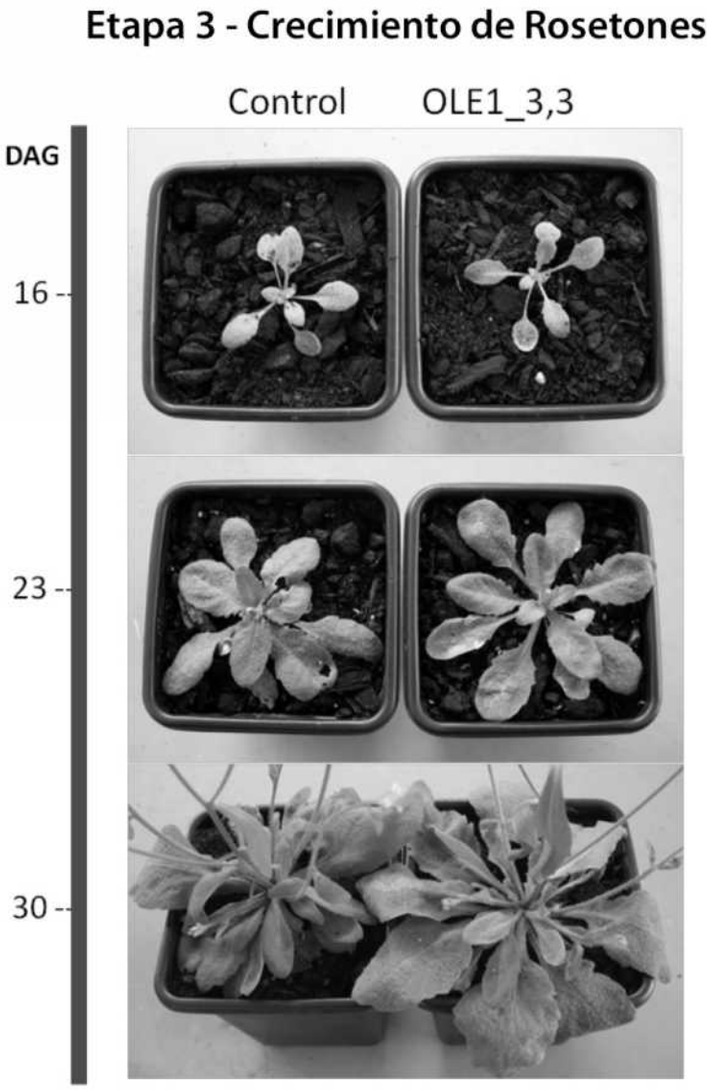

Etapa 3 - Crecimiento de Rosetones
Control
OLE1_3,3
DAG
16
23
30

Figura 35

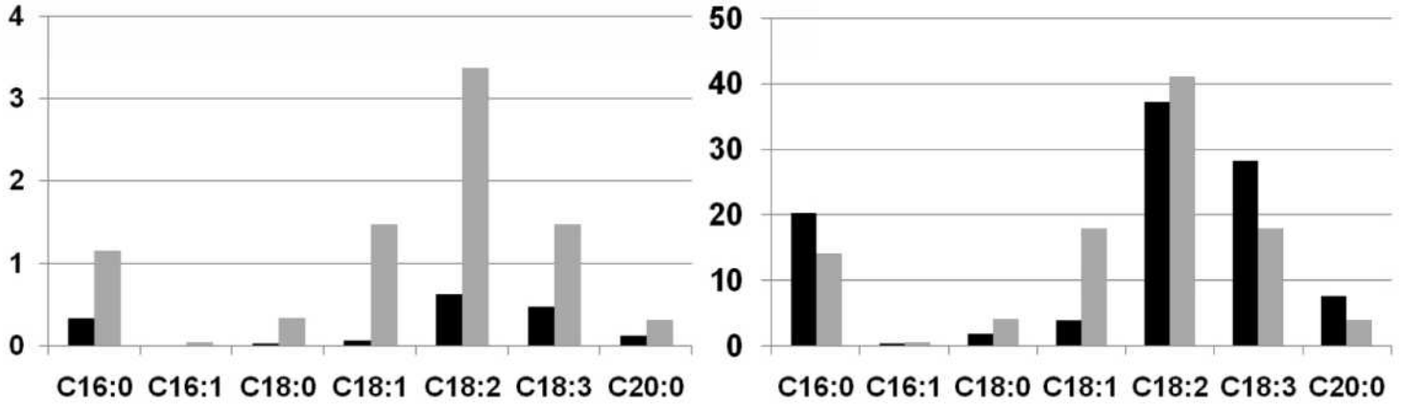

4
3
2
1
0
C16:0 C16:1 C18:0 C18:1 C18:2 C18:3 C20:0
50
40
30
20
10
0
C16:0 C16:1 C18:0 C18:1 C18:2 C18:3 C20:0

Figura 36

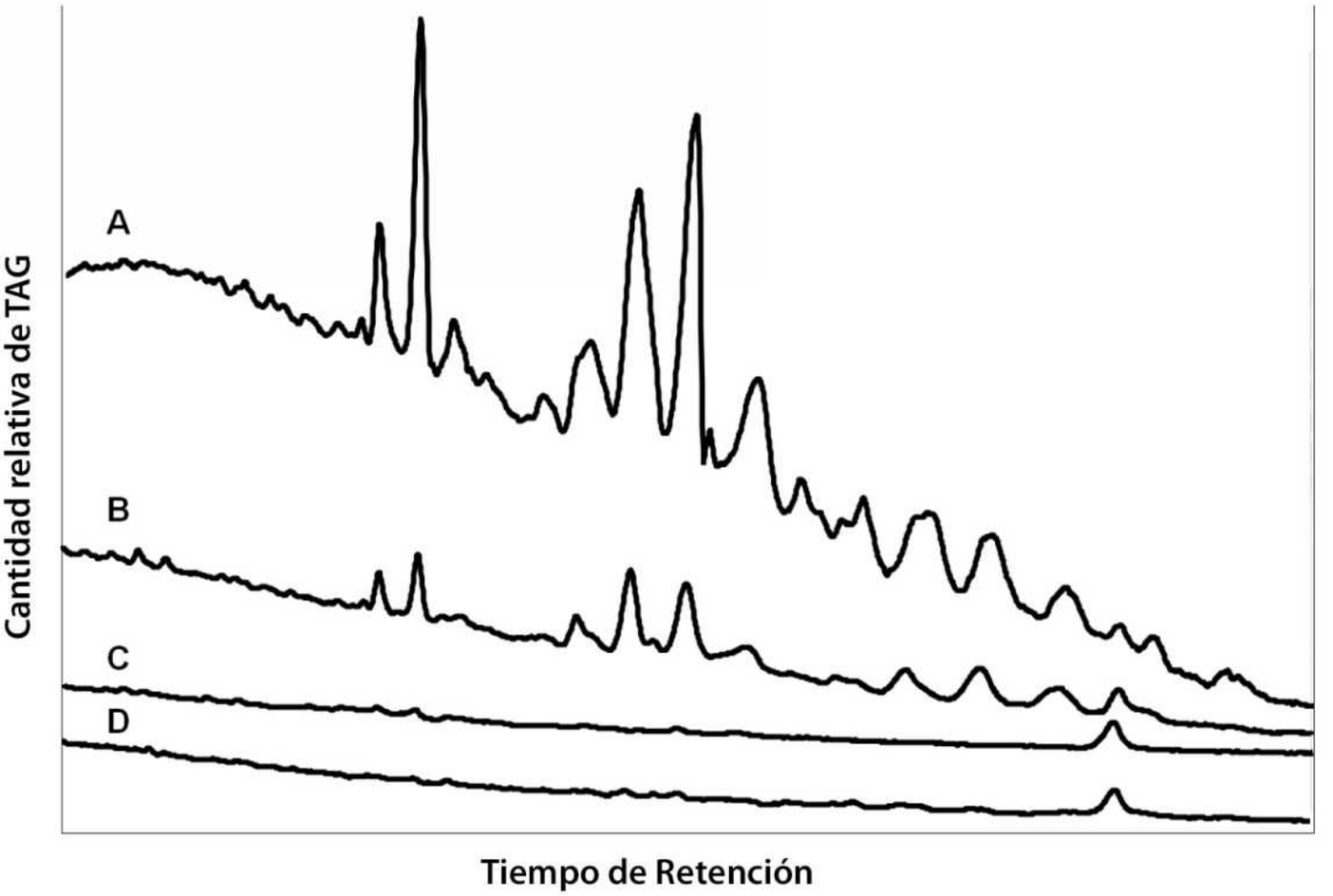
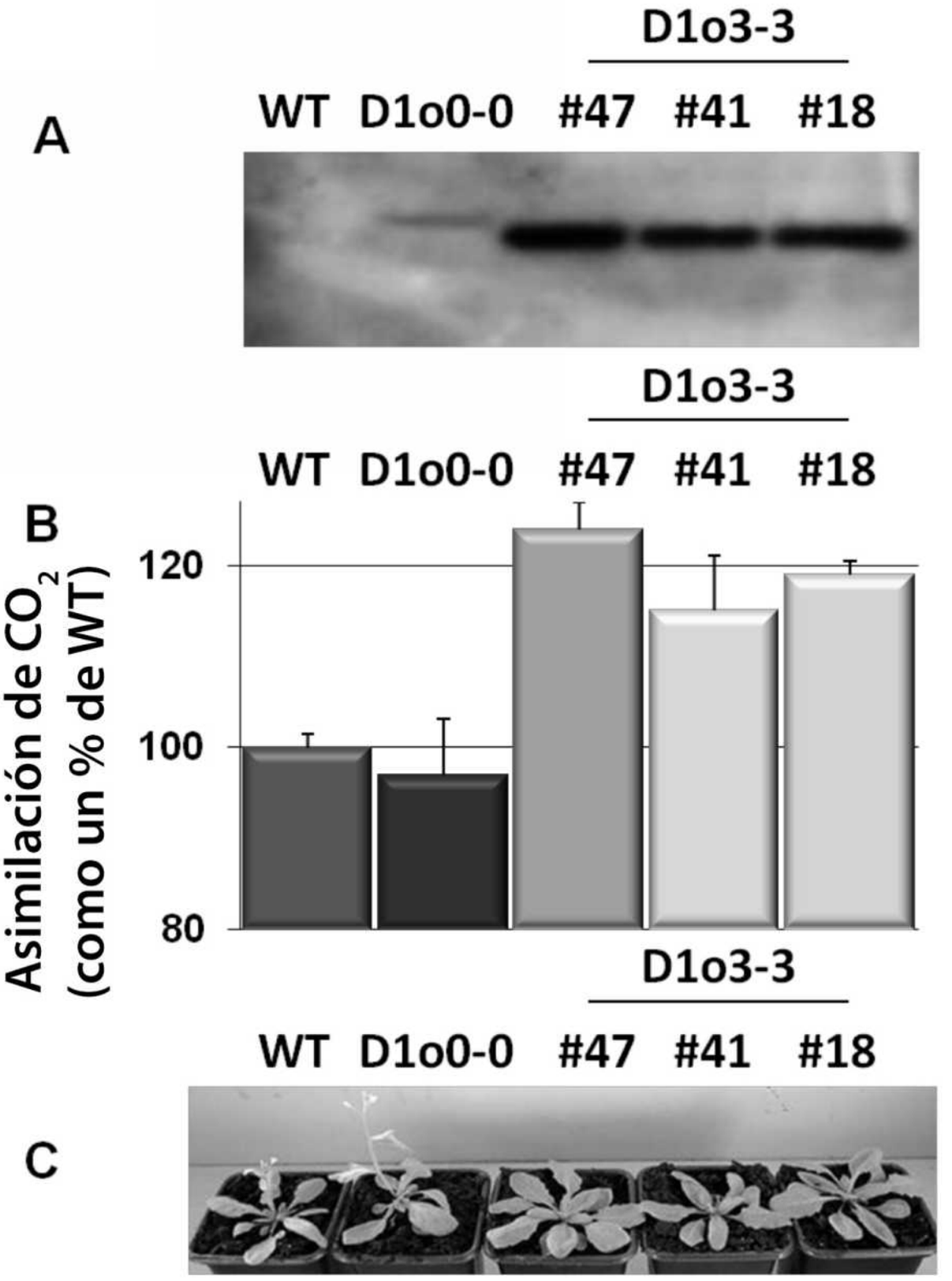


Figura 37


A
D1o3-3
WT
D1o0-0
#47
#41
#18
B
D1o3-3
WT
D1o0-0
#47
#41
#18
Asimilación de CO_2 (como un % de WT)
120
100
80
C
D1o3-3
WT
D1o0-0
#47
#41
#18

TRADUCCION OFICIAL No. 2014-02-18-107
DE UN DOCUMENTO ESCRITO EN INGLES

NICHOLAS JOHN ROBERTS
RICHARD WILLIAM SCOTT
SOMRUTAI WINICHAYAKUL
MARISSA ROLDAN
AGRESEARCH LIMITED

ESCRITURA DE TRASPASO DE DERECHOS DE PROPIEDAD INTELECTUAL

PARTES

NICHOLAS JOHN ROBERTS, ciudadano de Australia de 97 McKay Line, RD 7, Feilding 4777, Nueva Zelanda ("Primer Cedente")

RICHARD WILLIAM SCOTT, ciudadano de Nueva Zelanda de 25 Colombo Street, Palmerston North 4410, Nueva Zelanda ("Segundo Cedente")

SOMRUTAI WINICHAYAKUL, ciudadano de Tailandia de 26 Geneva Terrace, Kelvin Grove, Palmerston North 4414, Nueva Zelanda ("Tercer Cedente")

MARISSA ROLDAN, ciudadana de Nueva Zelanda de 12 Brooklyn Heights Drive, Kelvin Grove, Palmerston North 4414, Nueva Zelanda ("Cuarto Cedente")

AGRESEARCH LIMITED, una sociedad de Nueva Zelanda cuyo domicilio registrado está ubicado en 5th Floor, Tower Block, Ruakura Research Centre, East Street, Hamilton ("Cesionario")

INTRODUCCION

A. Los Cedentes han ideado o contribuido a la Invención, ya sea durante el curso del empleo con el Cesionario o bajo nombramiento del Cesionario.

B. Los Cedentes reconocen que el Cesionario es o debe ser el propietario legítimo y beneficioso de la Invención y los Derechos de Propiedad Intelectual.

C. Cada uno de los Cedentes ha acordado traspasar, y el Cesionario ha acordado aceptar, la Invención y dichos Derechos de Propiedad Intelectual con sujeción a los términos y condiciones de esta escritura.

ACUERDOS

1. DEFINICIONES

1.1 En esta escritura, incluyendo la Introducción, las siguientes palabras tendrán los siguientes significados:

"Cedentes" significa el Primer, Segundo, Tercero y Cuarto Cedente;

"Obra Protegida" significa:

(a) todas las obras protegidas y todos los derechos en la naturaleza de obra protegida, en un trabajo artístico, literario y otro original; y

(b) un derecho de bases de datos,

que comprenda o se relacione con la Invención según exista en cualquier parte del mundo;

"Derechos de Diseño" significa todos los derechos en y de los diseños que se apliquen a artículos de o relacionados con la Invención según existan en cualquier parte del mundo incluyendo, pero sin limitarse a:

(a) el derecho a solicitar y obtener protección para dichos diseños en relación con dichos artículos y los derechos conferidos por dicha protección cuando sea otorgado;

(b) el derecho a reivindicar prioridad bajo una convención o acuerdo internacional incluyendo la Convención de París (según sea modificada) de dicha solicitud o solicitudes referidas en el párrafo (a) anterior;

(c) un derecho de diseño no registrado; y

(d) un derecho de diseño de topografía de semi-conductores o circuitos integrados;

"Derechos de Propiedad Intelectual" significan todos los derechos de propiedad intelectual e industrial (ya sea que sea protegidos por estatuto, de hecho o en equidad) en y de la Invención según pudiesen existir en cualquier parte del mundo y ya sea que estén registrados o no o registrables incluyendo, pero sin limitarse a, Derechos de Patente, Obras Protegidas y Derechos de Diseño;

"Invención" significa cualquiera y todas las invenciones sujeto de la Solicitud de Patente;

"Know-How" / "Conocimiento Experto" significa información, conocimiento, experiencia, datos y diseños en posesión de uno de los Cedentes de naturaleza confidencial y no de dominio del público relacionada con la Invención o el proceso para hacerla o usarla;

"Solicitud de Patente" significa la solicitud de patente identificada en el Anexo; y

"Derechos de Patente" significa todos los derechos de patente en y de la Invención según pudiesen existir o lleguen a existir en cualquier parte del mundo incluyendo, pero sin limitarse a:

(a) el derecho a solicitar y obtener patentes u otras formas similares de protección respecto de la invención en cualquier país;

(b) el derecho a reivindicar prioridad bajo una convención o acuerdo internacional incluyendo la Convención de París (según sea modificada) de dicha solicitud o solicitudes referidas en el párrafo (a) anterior;

(c) todos los derechos conferidos por dicha patente(s) o formas similares de protección cuando sean otorgados; y

(d) la Solicitud de Patente junto con:

  (i) una patente que sea otorgada de conformidad con la Solicitud de Patente; y

  (ii) una patente(s) futura y solicitud(es) de patente que se base en o derive prioridad de o tenga reivindicaciones equivalentes a la Solicitud de Patente en cualquier país del

mundo (incluyendo divisionales, continuaciones, continuaciones en parte, certificados de protección complementarios, reexpediciones, extensiones, patentes de innovación y patentes pequeñas).

2. TRASPASO

2.1 Traspaso: En consideración de la suma de NZ$1.00 pagada por el Cesionario a cada uno de los Cedentes, el recibo y suficiencia de lo cual se reconoce por el presente por cada uno de los Cedentes, cada uno de los Cedentes por el presente traspasa al Cesionario absolutamente todos sus derechos, título e interés en y de:

(a) la Invención;

(b) los Derechos de Propiedad Intelectual; y

(c) el Know-How/Conocimiento Experto y el derecho completo, sin restricciones y exclusivo en todo el mundo para usar el Know-How/Conocimiento Experto para un propósito (incluyendo la presentación de solicitudes de patente respecto de todo o parte de ella).

2.2 Derechos de acción: Los traspasos efectuados de conformidad con la cláusula 2.1 incluyen, sin limitación, el traspaso y transferencia de:

(a) todos los derechos de acción, poderes y beneficios que surjan de la propiedad de los Derechos de Propiedad Intelectual incluyendo, sin limitación, el derecho a demandar por perjuicios y otros remedios jurídicos y equitativos respecto de todas las causas de acción que surjan antes, en o después de la fecha de esta escritura; y

(b) todos los derechos de propiedad de cualquier material que forme parte del Know-How/Conocimiento Experto.

2.3 Derechos morales: Los Cedentes renuncian a todos sus derechos morales que surjan de la Invención en todo el mundo, en la medida en que puedan hacerlo jurídicamente.

3. KNOW-HOW/CONOCIMIENTO EXPERTO

3.1 Disposición de Know-How/Conocimiento Experto: Cada uno de los Cedentes, a solicitud del Cesionario y en la medida en que sea excepcional:

(a) revelarán al Cesionario por escrito todo el Know-How/Conocimiento Experto conocido por ellos;

(b) proporcionarán toda otra asistencia e información razonable según sea razonablemente necesaria con el fin de ayudarle al Cesionario, o a su nominatario, para desarrollar y hacer o usar la Invención; y

(c) si aplica, proporcionarán todas las versiones originales de la Invención y cualquier material que forme parte del Know-How/Conocimiento Experto.

3.2 Confidencialidad:

(a) Los Cedentes acuerdan tratar toda la información relacionada con la Invención, Know-How/Conocimiento Experto y los Derechos de Propiedad Intelectual como secreto y confidencial.

(b) Después de la suscripción de esta escritura, los Cedentes no usarán, revelarán o publicarán la información referida en el párrafo (a) anterior sin el previo consentimiento por escrito del Cesionario.

(c) Estas obligaciones de confidencialidad no se extenderán a ninguna información que es o llegue a estar generalmente disponible al público por medio de un acto o defecto de los Cedentes. Si alguno de los Cedentes tiene conocimiento de alguna pérdida de reserva o confidencialidad actual o potencial respecto a dicha información, dichos Cedentes le notificarán oportunamente al Cesionario sobre dicho aspecto.

4. MEJORAS

4.1 Revelación: Después de la fecha de esta escritura, cada uno de los Cedentes, mientras sea empleado del Cesionario o bajo nombramiento del Cesionario, revelará de inmediato al Cesionario, todas las mejoras en, modificaciones de o adiciones a:

(a) la Invención;

(b) trabajo artístico, literario u otro relacionado con la Invención; y

(c) diseño que será aplicado a los artículos de o relacionados con la Invención,

ideados, creados, diseñados, contribuidos a o adquiridos por ellos ("Mejoras").

4.2 Propiedad de las Mejoras:

(a) El Cesionario poseerá exclusivamente todas las Mejoras y todos los derechos de propiedad intelectual en dichas Mejoras. En la medida en que una Mejora y dichos derechos de propiedad intelectual no le otorguen en su creación al Cesionario pero

le otorguen a uno o más de los Cedentes, dichos Cedentes mantendrán dichas Mejoras y derechos de propiedad intelectual en fiducia para el Cesionario.

(b) Los Cedentes a quienes el párrafo (a) anterior aplique suscribirán en un momento, después de la solicitud razonable del Cesionario y a expensas del Cesionario, de manera apropiada todos los documentos necesarios para:

(i) confirmar la propiedad del Cesionario de las Mejoras y todos los derechos de propiedad intelectual en dichas Mejoras; o

(ii) presentarán una solicitud protectora para dichas Mejoras y derechos de propiedad intelectual y/o defenderán dicha solicitud protectora.

5. SUSCRIPCION DE DOCUMENTOS Y ACCIONES ADICIONALES

5.1 Acciones adicionales: Si el Cesionario les solicita, cada uno de los Cedentes, a expensas del Cesionario, suscribirá todos los documentos, dará dicha asistencia y hará todos los otros actos y cosas según sean necesarios o deseables para:

(a) solicitar y obtener o (si el Cesionario lo considera apropiado) unirse al Cesionario para solicitar y obtener protección de los Derechos de Propiedad Intelectual, con el entendimiento y la intención de que todos los derechos, título e interés en y de dichas solicitudes y protección otorgada es otorgada al Cesionario;

(b) otorgar dicha protección referida en el párrafo (a) en el Cesionario;

(c) modificar, mantener o renovar dicha protección referida en el párrafo (a);

(d) asistir al Cesionario para;

(i) permitir la suscripción de los Derechos de Propiedad Intelectual (incluyendo la obtención de todos los remedios según estén disponibles por infracción de los Derechos de Propiedad Intelectual);

(ii) derrotar el desafío a la validez de los Derechos de Propiedad Intelectual;

(iii) defender cualquier procedimiento de oposición entablada por un tercero respecto de los Derechos de Propiedad Intelectual; o

(iv) conducir procedimientos de oposición respecto a una solicitud para protección de propiedad intelectual por un tercero cuando dicha solicitud pudiese afectar adversamente la capacidad del Cesionario para explotar los Derechos de Propiedad Intelectual; y

(e) de lo contrario implementar y llevar a cabo sus obligaciones bajo esta escritura.

5.2 Poder: Cada uno de los Cedentes por el presente irrevocablemente nombra al Cesionario como su apoderado con pleno poder para actuar en su nombre para llevar a cabo cualquiera de los asuntos descritos en la cláusula 5.1:

(a) en la medida en que dichos Cedentes fallen en llevar a cabo dichos asuntos después de haber sido llamados para ello por el Cesionario; o

(b) si el Cesionario no puede, después de hacer averiguaciones razonables y apropiadas, localizar a dichos Cedentes para solicitarles la ejecución de dicho asunto.

6. GENERAL

6.1 Renuncia: Ni la falla ni la demora de alguna de las partes en ejercer un derecho, poder o privilegio bajo esta escritura operará como una renuncia de dicho derecho, poder o privilegio, ni ningún ejercicio individual o parcial excluirá ningún otro o ejercicio adicional de un poder, derecho o privilegio bajo esta escritura.

6.2 Copias facsímiles:

(a) Esta escritura podrá ser suscrita en cualquier número de copias (incluyendo copias facsímiles) todas las cuales, al ser tomadas en su conjunto, constituirán uno y el mismo instrumento.

(b) Una de las partes podrá celebrar esta escritura al suscribir una de las copias. Las partes reconocen que esta escritura podrá ser suscrita con base en un intercambio de copias facsímiles y confirmar que su respectiva suscripción de esta escritura por dichos medios será una suscripción válida y suficiente.

6.3 Ley que rige: La formación, validez, construcción y suscripción de esta escritura se regirá por y se interpretará de conformidad con las leyes de Nueva Zelanda. Las partes irrevocablemente acuerdan que los Tribunales de Nueva Zelanda tendrán jurisdicción no exclusiva para oír y determinar todos los conflictos bajo o en conexión con esta escritura. Las partes irrevocablemente renuncian a cualquier objeción en cuanto a Nueva Zelanda como el foro para procedimientos sobre la base de foro inconveniente o bases similares.

FIRMADO COMO ESCRITURA

FIRMADO por NICHOLAS JOHN ROBERTS

en presencia de: (ilegible)

Firma

20/9/11
Fecha

TESTIGO
Firma: (ilegible)
Nombre: (ilegible)
Dirección: Palmerston North
Ocupación: Administrador Científico

FIRMADO por RICHARD WILLIAM SCOTT
en presencia de:

(ilegible)
Firma
20/9/11
Fecha

TESTIGO
Firma: (ilegible)
Nombre: (ilegible)
Dirección: Palmerston North
Ocupación: Administrador Científico

FIRMADO por SOMRUTAI WINICHAYAKUL
en presencia de:

(ilegible)
Firma
20/9/11
Fecha

TESTIGO
Firma: (ilegible)
Nombre: (ilegible)
Dirección: Palmerston North
Ocupación: Administrador Científico

FIRMADO por MARISSA ROLDAN
en presencia de:

(ilegible)
Firma
20/9/11

Fecha

TESTIGO

Firma: (ilegible)

Nombre: (ilegible)

Dirección: Palmerston North

Ocupación: Administrador Científico

FIRMADO por AGRESEARCH LIMITED por: (ilegible)

Firma del Director / Signatario Autorizado

Peter Scott

Nombre del Director

23/9/11

Fecha

_____

Firma del Director

_____

Nombre del Director

_____

Fecha

ANEXO

SOLICITUD DE PATENTE

País	Solicitud No.	Título	Fecha de Presentación
Estados Unidos de América	61/515,610	METODO PARA AUMENTAR LA ASIMILACION DE CO2 Y LA PRODUCCION DE ACEITE EN ORGANISMOS NO FOTOSINTETICOS	5 de agosto de 2011

Traducido según mi mejor saber y entender, este día 18 de febrero de 2014, en la ciudad de Bogotá, D.C., por Anita Escobar de Tamayo, Traductora Oficial, Res. 2177 de 1989, Ministerio de Justicia, República de Colombia.

TRADUCTORA OFICIAL
RES. 2177/89 MINJUSTICIA

NICHOLAS JOHN ROBERTS

RICHARD WILLIAM SCOTT

SOMRUTAI WINICHAYAKUL

MARISSA ROLDAN

AGRESEARCH LIMITED

DEED OF ASSIGNMENT OF INTELLECTUAL PROPERTY RIGHTS

PARTIES

NICHOLAS JOHN ROBERTS, an Australian citizen of 97 McKay Line, RD 7, Feilding 4777, New Zealand ("First Assignor")

RICHARD WILLIAM SCOTT, a New Zealand citizen of 25 Colombo Street, Palmerston North 4410, New Zealand ("Second Assignor")

SOMRUTAI WINICHAYAKUL, a Thai citizen of 26 Geneva Terrace, Kelvin Grove, Palmerston North 4414, New Zealand ("Third Assignor")

MARISSA ROLDAN, a New Zealand citizen of 12 Brooklyn Heights Drive, Kelvin Grove, Palmerston North 4414, New Zealand ("Fourth Assignor")

AGRESEARCH LIMITED, a New Zealand company whose registered office is located at 5th Floor, Tower Block, Ruakura Research Centre, East Street, Hamilton ("Assignee")

INTRODUCTION

A. The Assignors have devised or contributed to the Invention, either during the course of employment with the Assignee or under a commission from the Assignee.

B. The Assignors acknowledge that the Assignee is or should be the legal and beneficial owner of the Invention and the Intellectual Property Rights.

C. Each of the Assignors has agreed to assign, and the Assignee has agreed to accept, the Invention and such Intellectual Property Rights subject to the terms and conditions of this deed.

COVENANTS

1. DEFINITIONS

1.1 In this deed, including the Introduction, the following words will have the following meanings:

"Assignors" means the First, Second, Third and Fourth Assignor;

"Copyright" means:

(a) all copyrights and all rights in the nature of copyright, in any original artistic, literary and other works; and

(b) any database rights,

comprising or relating to the Invention as may exist anywhere in the world;

"Design Rights" means all rights in and to the designs to be applied to articles of or relating to the Invention as may exist anywhere in the world including, but not limited to:

(a) the right to apply for and obtain protection for such designs in relation to such articles and the rights conferred by such protection when granted;

(b) the right to claim priority under any international convention or agreement including the Paris Convention (as amended) from any such application or applications referred to in paragraph (a) above;

(c) any unregistered design rights; and

(d) any semi-conductor topography or integrated circuit layout rights;

"Intellectual Property Rights" means all industrial and intellectual property rights (whether protectable by statute, at common law or in equity) in and to the Invention as may exist anywhere in the world and whether or not registered or registrable including, but not limited to, the Patent Rights, Copyright and Design Rights;

"Invention" means any and all inventions the subject of the Patent Application;

"Know-How" means any information, knowledge, experience, data and designs in the possession of either of the Assignors of a confidential nature and not in the public domain relating to the Invention or the process for making or using it;

"Patent Application" means the patent application identified in the Schedule; and

"Patent Rights" means all patent rights in and to the Invention as may exist or come into existence anywhere in the world including, but not limited to:

(a) the right to apply for and obtain patents or other similar forms of protection in respect of the Invention in any country;

(b) the right to claim priority under any international convention or agreement including the Paris Convention (as amended) from any such application or applications referred to in paragraph (a) above;

(c) all rights conferred by any such patent(s) or similar forms of protection when granted; and

(d) the Patent Application together with:

(i) any patent that may be granted pursuant to the Patent Application; and

(ii) any future patent(s) and patent application(s) that are based on or derive priority from or have equivalent claims to the Patent Application in any country in the world (including divisionals, continuations, continuations in part, supplementary protection certificates, reissues, extensions, innovation patents and petty patents).

2. ASSIGNMENT

2.1 **Assignment:** In consideration of the sum of NZ$1.00 paid by the Assignee to each of the Assignors, the receipt and sufficiency of which is hereby acknowledged by each of the Assignors, each of the Assignors hereby assigns to the Assignee absolutely all of their rights, title and interest in and to:

(a) the Invention;

(b) the Intellectual Property Rights; and

(c) the Know-How and the full, unfettered and exclusive worldwide right to use the Know-How for any purpose whatsoever (including the filing of patent applications in respect of all or any part of it).

2.2 **Rights of action:** The assignments effected by clause 2.1 include, without limitation, the assignment and transfer of:

(a) all rights of action, powers and benefits arising from the ownership of the Intellectual Property Rights including, without limitation, the right to sue for damages and other legal and equitable remedies in respect of all causes of action arising before, on or after the date of this deed; and

(b) all rights of ownership of any materials that form part of the Know-How.

2.3 **Moral rights:** The Assignors waive all of their moral rights arising from the Invention throughout the world, to the extent that they may lawfully do so.

3. KNOW-HOW

3.1 **Provision of Know-How:** Each of the Assignors will, at the request of the Assignee and to the extent outstanding:

(a) disclose to the Assignee in writing all Know-How known to them;

(b) provide all other reasonable assistance and information as may be reasonably necessary in order to assist the Assignee, or its nominee, to develop and make or use the Invention; and

(c) where possible, provide all original versions of the Invention and any materials that form part of the Know-How.

3.2 **Confidentiality:**

(a) The Assignors agree to treat all information relating to the Invention, the Know-How and the Intellectual Property Rights as secret and confidential.

(b) Following the execution of this deed, the Assignors will not use, disclose or publish the information referred to in paragraph (a) above without the Assignee's prior written consent.

(c) These obligations of confidentiality will not extend to any information that is or becomes generally available to the public through no act or default of the Assignors. If either of the Assignors becomes aware of any actual or potential loss of secrecy or confidentiality in respect of such information, such of the Assignors will promptly advise the Assignee of such matter.

4. IMPROVEMENTS

4.1 **Disclosure:** Following the date of this deed, each of the Assignors, while employed by the Assignee or under any commission from the Assignee, will immediately disclose to the Assignee all improvements in, modifications of or additions to:

(a) the Invention;

(b) any original artistic, literary or other works relating to the Invention; and

(c) any designs to be applied to the articles of or relating to the Invention,

devised, created, designed, contributed to or acquired by them ("Improvements").

4.2 **Ownership of Improvements:**

(a) The Assignee will exclusively own all Improvements and all intellectual property rights in such Improvements. To the extent that any Improvements and such intellectual property rights do not on their creation vest in the Assignee but vest in one or more of the Assignors, such Assignors will hold such Improvements and intellectual property rights on trust for the Assignee.

(b) The Assignors to whom paragraph (a) above applies will at any time, upon the Assignee's reasonable request and at the Assignee's expense, appropriately execute all documents necessary to:

(i) confirm the Assignee's ownership of the Improvements and all intellectual property rights in such Improvements; or

(ii) file a protective application for such Improvements and intellectual property rights and/or defend such protective application.

5. EXECUTION OF DOCUMENTS AND FURTHER ACTIONS

5.1 **Further actions:** If requested by the Assignee, each of the Assignors will, at the Assignee's expense, execute all documents, give such assistance and do all other acts and things as may be necessary or desirable to:

(a) apply for and obtain or (if the Assignee thinks fit) join with the Assignee in applying for and obtaining protection of the Intellectual Property Rights, with the understanding and the intention that all rights, title and interest in and to such applications and any granted protection is to vest in the Assignee;

(b) vest any such protection referred to in paragraph (a) in the Assignee;

(c) amend, maintain or renew any such protection referred to in paragraph (a);

(d) assist the Assignee to:

(i) enable enforcement of any of the Intellectual Property Rights (including obtaining all remedies as may be available for infringement of the Intellectual Property Rights);

(ii) defeat any challenge to the validity of any of the Intellectual Property Rights;

(iii) defend any opposition proceedings brought by a third party in respect of the Intellectual Property Rights; or

(iv) conduct opposition proceedings in respect of any application for intellectual property protection by a third party where such application may adversely affect the Assignee's ability to exploit the Intellectual Property Rights; and

(e) otherwise implement and carry out their obligations under this deed.

5.2 **Power of attorney:** Each of the Assignors hereby irrevocably appoints the Assignee as their attorney with full power to act in their name and on their behalf in fulfilling any of the matters set out in clause 5.1:

(a) to the extent that such of the Assignors fails to do any of such matters after being called upon to do so by the Assignee; or

(b) if the Assignee is unable, after making reasonable and proper inquiries, to locate such of the Assignors to request the fulfilment of such matters.

6. GENERAL

6.1 **Waiver:** No failure or delay by any party in exercising any right, power or privilege under this deed will operate as a waiver of such right, power or privilege, nor will any single or partial exercise preclude any other or further exercise of any right, power or privilege under this deed.

6.2 **Facsimile counterparts:**

(a) This deed may be executed in any number of counterparts (including facsimile copies) all of which, when taken together, will constitute one and the same instrument.

(b) A party may enter into this deed by executing any counterpart. The parties acknowledge that this deed may be executed on the basis of an exchange of facsimile copies and confirm that their respective execution of this deed by such means will be a valid and sufficient execution.

6.3 **Governing law:** The formation, validity, construction and performance of this deed will be governed by and construed in accordance with the laws of New Zealand. The parties irrevocably agree that the Courts of New Zealand will have non-exclusive jurisdiction to hear and determine all disputes under or in connection with this deed. The parties irrevocably waive any objections to New Zealand as the forum for proceedings on the grounds of forum non-conveniens or any similar grounds.

SIGNED AS A DEED

SIGNED by **NICHOLAS JOHN ROBERTS** in the presence of:

Signature

20-9-11
Date

WITNESS
Signature:
Name:
Address:
Occupation: Scene Administrator

SIGNED by **RICHARD WILLIAM SCOTT** in the presence of:

Signature

20-9-11
Date

WITNESS
Signature:
Name:
Address:
Occupation: Scene Administrator

SIGNED by **SOMRUTAI WINICHAYAKUL** in the presence of:

Signature

20-9-11
Date

WITNESS
Signature:
Name:
Address:
Occupation: Scene Administrator

SIGNED by **MARISSA ROLDAN** in the presence of:

Signature

20-9-11

Date

WITNESS

Signature:

Name:

Address: Palmerston North

Occupation: Dairnz Administrator

SIGNED by **AGRESEARCH LIMITED** by:

Signature of Director Authorised Signatory

Peter Scott

Name of Director

23/9/2011

Date

Signature of Director

Name of Director

Date

SCHEDULE

PATENT APPLICATION

Country	Application No.	Title	Date Filed
United States of America	61/515,610	METHODS FOR INCREASING CO2 ASSIMILATION AND OIL PRODUCTION IN PHOTOSYNTHETIC ORGANISMS	5 August 2011

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference 660921HCF	**FOR FURTHER ACTION**	see Form PCT/ISA/220 as well as, where applicable, item 5 below.
International application No. **PCT/NZ2012/000138**	International filing date *(day/month/year)* 03 August 2012	(Earliest) Priority Date *(day/month/year)* 05 August 2011

Applicant

AGRESEARCH LIMITED ET AL

This international search report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This international search report consists of a total of **7** sheets.

☐ It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**

a. With regard to the **language,** the international search was carried out on the basis of:

☒ The international application in the language in which it was filed.

☐ A translation of the international application into , which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

b. ☐ This international search report has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43.6*bis*(a)).

c. ☒ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, see Box No. I.

2. ☐ **Certain claims were found unsearchable** (See Box No. II).

3. ☒ **Unity of invention is lacking** (See Box No. III).

4. With regard to the **title,**

☒ the text is approved as submitted by the applicant.

☐ the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**

☒ the text is approved as submitted by the applicant.

☐ the text has been established, according to Rule 38.2, by this Authority as it appears in Box No. IV. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. With regard to the **drawings,**

a. the figure of the **drawings** to be published with the abstract is Figure No. **30**

☐ as suggested by the applicant.

☒ as selected by this Authority, because the applicant failed to suggest a figure.

☐ as selected by this Authority, because this figure better characterizes the invention.

b. ☐ none of the figures is to be published with the abstract.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ2012/000138

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:

   a. (means)

   ☐ on paper

   ☐ in electronic form

   b. (time)

   ☐ in the international application as filed

   ☐ together with the international application in electronic form

   ☐ subsequently to this Authority for the purposes of search

2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

A sequence listing in both paper and electronic form was filed, but it was not used for the purposes of search and opinion.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ2012/000138

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

See Supplemental Box for Details

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NZ2012/000138

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/29 (2006.01) C12N 15/82 (2006.01) C12N 1/13 (2006.01) C07K 14/415 (2006.01) A01H 5/00 (2006.01) A01H 13/00 (2006.01) C12P 5/00 (2006.01) A23K 1/00 (2006.01)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Databases searched: MEDLINE, AGRICOLA, BIOSIS, CAPLUS, EMBASE, EPODOC, WPI. Keywords: oleosin, polyoleosin, CO2, carbon, assimilate, fix, recycle, lipid, fatty acid, release, oil, seedoil, triacylglycerol (and like terms).

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 December 2012	17 December 2012
Name and mailing address of the ISA/AU	**Authorised officer**
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustralia.gov.au Facsimile No.: +61 2 6283 7999	Alison Knight AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 6283 2847

INTERNATIONAL SEARCH REPORT C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		International application No. PCT/NZ2012/000138
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/053169 A1 (AGRESEARCH LIMITED) 05 May 2011 Example 5; Example 6; lines 23-27, p4; lines 16-17, p11	35-71
A	EKMAN, A. et al. Elevated atmospheric CO(2) concentration and diurnal cycle induce changes in lipid composition in Arabidopsis thaliana. The New Phytologist. 2007, 174(3): 591-599. Abstract	1-34
A	SUEN, Y. et al. Total lipid production of the green alga Nannochloropsis sp. QII under different nitrogen regimes. Journal of Phycology. 1987, 23: 289-296. Abstract	1-34
X	WO 2007/045019 A2 (AGRICULTURE VICTORIA SERVICES PTY LTD et al.) 26 April 2007 line 25, p11- line 10, p12; lines 25-27, p12; lines 25-28, p13;	35; 44; 46-47; 50-53; 58-65;
X	VAN ROOIJEN, G.J.H. et al. Structural Requirements of the Oleosin domains for subcellular targeting to the oil body. Plant Physiology. 1995, 109: 1353-1361. Abstract; Materials and Methods p1353-1354; Materials and Methods, p1356-1357;	35; 44; 46-47; 50; 58-59; 61-65;
E	US 2012/0278951 A1 (ROBERTS, N et. al) 01 November 2012 Abstract; lines 25-28, page 5; lines 1-8, p6;	1-71

Supplemental Box

Continuation of: **Box III**

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are two different inventions based on the following features that separate the claims into distinct groups:

Invention 1: Methods for producing a photosynthetic cell with an increased rate of CO_2 assimilation comprising genetically modifying the photosynthetic cell to reduce or prevent lipid recycling and/or transforming a photosynthetic cell with a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine. Invention 1 is defined by claims 1-34 (fully).

Invention 2: Methods for producing oil comprising providing a plant comprising a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine and cultivating the plant to express said oleosin in its non-photosynthetic tissues; a non-photosynthetic tissue/organ of a plant produced by the method; an animal feed or biofuel comprising said non-photosynthetic tissue/organ. Invention 2 is defined by claims 35-71 (fully).

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention.

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The only feature common to all of the claimed inventions and which provides a technical relationship among them is a modified oleosin including at least one artificially introduced cysteine.

However this feature does not make a contribution over the prior art because it is disclosed in:

D1: WO 2011/053169 A1 (AGRESEARCH LIMITED) 5 May 2011

Therefore in light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

NOTE: Despite the lack of unity, in the interest of expediting examination, all inventions have been searched and examined so far as they relate to a modified oleosin (see Article 5 objection).

INTERNATIONAL SEARCH REPORT Information on patent family members	International application No. **PCT/NZ2012/000138**

This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document/s Cited in Search Report		**Patent Family Member/s**	
Publication Number	**Publication Date**	**Publication Number**	**Publication Date**
WO 2011/053169 A1	05 May 2011	AR 078858 A1	07 Dec 2011
		AU 2010313865 A1	10 May 2012
		CA 2778150 A1	05 May 2011
		CN 102741411 A	17 Oct 2012
		EP 2494051 A1	05 Sep 2012
		MX 2012004677 A	25 Jul 2012
		TW 201127957 A	16 Aug 2011
		US 2012278951 A1	01 Nov 2012
		UY 32994 A	31 May 2011
		WO 2011053169 A1	05 May 2011
WO 2007/045019 A2	26 Apr 2007	AU 2006303870 A1	26 Apr 2007
		CA 2626390 A1	26 Apr 2007
		EP 1934350 A2	25 Jun 2008
		NZ 567603 A	30 Mar 2012
		US 2009133160 A1	21 May 2009
		US 8309794 B2	13 Nov 2012
		WO 2007045019 A2	26 Apr 2007
US 2012/0278951 A1	01 Nov 2012	AR 078858 A1	07 Dec 2011
		AU 2010313865 A1	10 May 2012
		CA 2778150 A1	05 May 2011
		CN 102741411 A	17 Oct 2012
		EP 2494051 A1	05 Sep 2012
		MX 2012004677 A	25 Jul 2012
		TW 201127957 A	16 Aug 2011
		US 2012278951 A1	01 Nov 2012
		UY 32994 A	31 May 2011
		WO 2011053169 A1	05 May 2011

End of Annex

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 February 2013 (14.02.2013)

WIPO | PCT

(10) International Publication Number
WO 2013/022353 A1

(51) International Patent Classification:
C12N 15/29 (2006.01) *A01H 5/00* (2006.01)
C12N 15/82 (2006.01) *A01H 13/00* (2006.01)
C12N 1/13 (2006.01) *C12P 5/00* (2006.01)
C07K 14/415 (2006.01) *A23K 1/00* (2006.01)

(21) International Application Number:
PCT/NZ2012/000138

(22) International Filing Date:
3 August 2012 (03.08.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/515,610 5 August 2011 (05.08.2011) US

(71) Applicant *(for all designated States except US)*: **AGRESEARCH LIMITED** [NZ/NZ]; 5th Floor Tower Block Ruakura Research, Centre East Street, Hamilton, 3240 (NZ).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **ROBERTS, Nicholas John** [AU/NZ]; 97 McKay Line, RD 7, Feilding, Wellington, 4777 (NZ). **SCOTT, Richard William** [NZ/NZ]; 25 Colombo Street, Palmerston North, Wellington, 4410 (NZ). **WINICHAYAKUL, Somrutai** [TH/NZ]; 26 Geneva Terrace, Kelvin Grove, Palmerston North, Wellington, 4414 (NZ). **ROLDAN, Marissa** [NZ/NZ]; 12 Brooklyn Heights Drive, Kelvin Grove, Palmerston North, Wellington, 4414 (NZ).

(74) Agents: **FOOTE, Humphrey, C** et al.; A J Park, Level 22, State Insurance Tower, 1 Willis Street, P O Box 949, Wellington, 6015 (NZ).

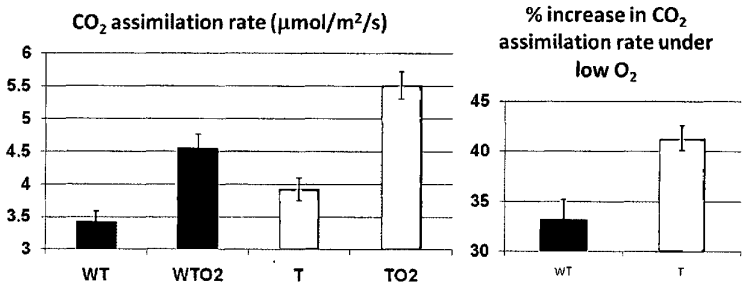
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) **Title**: METHODS FOR INCREASING CO_2 ASSIMILATION AND OIL PRODUCTION IN PHOTOSYNTHETIC ORGANISMS

Figure 30



(57) **Abstract**: The invention provides methods for increasing the rate of CO_2 assimilation in photosynthetic cells and plants. The invention involves reducing or preventing lipid recycling, and/or expressing modified oleosins with artificially introduced cysteine residues in the photosynthetic cells and plants. The invention also provides methods for increasing oil production in plants, via expression of modified oleosins with artificially introduced cysteine residues in the non-photosynthetic tissues/organs of plants. The method also optionally includes the step of extracting the oil from the non-photosynthetic tissues/organs of the plant, or processing the oil rich non-photosynthetic tissues/organs into animal or biofuel feedstocks.

WO 2013/022353 A1

PATENT COOPERATION TREATY

From the:
INTERNATIONAL SEARCHING AUTHORITY

To:

A J PARK
Level 22, State Insurance Tower
1 Willis Street
P O Box 949
6015 Wellington New Zealand

PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing *(day/month/year)*
17 December 2012

Applicant's or agent's file reference
660921HCF

FOR FURTHER ACTION
See paragraph 2 below

International application No.	International filing date *(day/month/year)*	Priority date *(day/month/year)*
PCT/NZ2012/000138	03 August 2012	05 August 2011

International Patent Classification (IPC) or both national classification and IPC
C12N 15/29 (2006.01) C12N 15/82 (2006.01) C12N 1/13 (2006.01) C07K 14/415 (2006.01) A01H 5/00 (2006.01) A01H 13/00 (2006.01) C12P 5/00 (2006.01) A23K 1/00 (2006.01)

Applicant
AGRESEARCH LIMITED ET AL

1. This opinion contains indications relating to the following items:

- [X] Box No. I — Basis of the opinion
- [X] Box No. II — Priority
- [ ] Box No. III — Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [X] Box No. IV — Lack of unity of invention
- [X] Box No. V — Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- [X] Box No. VI — Certain documents cited
- [ ] Box No. VII — Certain defects in the international application
- [X] Box No. VIII — Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

Name and mailing address of the ISA	Date of completion of this opinion	**Authorised Officer**
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA Email address: pct@ipaustralia.gov.au Facsimile No. +61 2 6283 7999	17 December 2012	Alison Knight AUSTRALIAN PATENT OFFICE (ISO 9001 Quality Certified Service) Telephone No. +61 2 6283 2847

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/NZ2012/000138

Box No. I Basis of this opinion

1. With regard to the **language,** this opinion has been established on the basis of:

   [X] The international application in the language in which it was filed

   [ ] A translation of the international application into, , which is the language of a translation furnished for the purposes of international search (under Rules 12.3(a) and 23.1(b)).

2. [ ] This opinion has been established taking into account **the rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(a))

3. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing filed or furnished:

   a. (means)

      [ ] on paper

      [ ] in electronic form

   b. (time)

      [ ] in the international application as filed

      [ ] together with the international application in electronic form

      [ ] subsequently to this Authority for the purposes of search

4. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

A sequence listing in both paper and electronic form was filed, but it was not used for the purposes of search and opinion.

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/NZ2012/000138

Box No. II Priority

1. ☐ The validity of the priority claim has not been considered because the International Searching Authority does not have in its possession a copy of the earlier application whose priority has been claimed or, where required, a translation of that earlier application. This opinion has nevertheless been established on the assumption that the relevant date (Rules 43*bis*.1 and 64.1) is the claimed priority date.

2. ☐ This opinion has been established as if no priority had been claimed due to the fact that the priority claim has been found invalid (Rules 43*bis*.1 and 64.1). Thus for the purposes of this opinion, the international filing date indicated above is considered to be the relevant date.

3. Additional observations, if necessary:

The claimed priority date is found to be valid.

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY

International application No.

PCT/NZ2012/000138

Box No. IV Lack of unity of invention

1. ☐ In response to the invitation (Form PCT/ISA/206) to pay additional fees the applicant has, within the applicable time limit:

   ☐ paid additional fees

   ☐ paid additional fees under protest and, where applicable, the protest fee

   ☐ paid additional fees under protest but the applicable protest fee was not paid

   ☐ not paid additional fees

2. ☒ This Authority found that the requirement of unity of invention is not complied with and chose not to invite the applicant to pay additional fees.

3. This Authority considers that the requirement of unity of invention in accordance with Rule 13.1, 13.2 and 13.3 is

   ☐ complied with

   ☒ not complied with for the following reasons:

See Supplemental Box for Details

4. Consequently, this opinion has been established in respect of the following parts of the international application:

   ☒ all parts

   ☐ the parts relating to claims Nos.

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY	International Application No. PCT/NZ2012/000138

Box No. V **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement**

1. Statement

Novelty (N)	Claims **1-34; 42-43; 45; 48-49; 54-55;**	**YES**
	Claims **35-41; 44; 46-47; 50-53; 56-71;**	**NO**
Inventive step (IS)	Claims **1-34**	**YES**
	Claims **35-71**	**NO**
Industrial applicability (IA)	Claims **1-71**	**YES**
	Claims **NONE**	**NO**

The inventions searched and examined by this Authority relate to methods for producing a photosynthetic cell with an increased rate of CO_2 assimilation (Invention 1) and methods for producing oil (Invention 2). Both the method of Invention 1 and the method of Invention 2 comprise the provision of a photosynthetic cell (Invention 1) or plant (Invention 2) comprising a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine. Invention 2 also includes a non-photosynthetic tissue/organ of a plant produced by the method of Invention 2 and an animal feed or biofuel comprising said non-photosynthetic tissue/organ.

2. CITATIONS AND EXPLANATIONS:

CITATIONS

D1: WO 2011/053169 A1 (AGRESEARCH LIMITED) 05 May 2011
D2: EKMAN, A. et al. Elevated atmospheric CO(2) concentration and diurnal cycle induce changes in lipid composition in Arabidopsis thaliana. The New Phytologist. 2007, 174(3): 591-599.
D3: SUEN, Y. et al. Total lipid production of the green alga Nannochloropsis sp. QII under different nitrogen regimes. Journal of Phycology. 1987, 23: 289-296.
D4: WO 2007/045019 A2 (AGRICULTURE VICTORIA SERVICES PTY LTD et al.) 26 April 2007
D5: VAN ROOIJEN, G.J.H. et al. Structural Requirements of the Oleosin domains for subcellular targeting to the oil body. Plant Physiology. 1995, 109: 1353-1361.

NOVELTY (N)

Invention 1 relates to methods for producing a photosynthetic cell with an increased rate of CO_2 assimilation via transforming a photosynthetic cell with a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine. While photosynthetic cells transformed with said modified oleosins exist in the art, as disclosed for example in D1, the art does not disclose that such modifications can increase the rate of CO_2 assimilation. Similarly, while the art discloses that elevated CO_2 levels induce changes in the composition of leaves in *Arabidopsis thaliana* (D2, Abstract), and that CO_2 fixation causes enhanced lipid biosynthesis (D3, Abstract), a link between oleosin and CO_2 fixation specifically is not disclosed. Therefore claims 1-34 meet the requirements of Novelty under Article 33(2) of the PCT.

Invention 2 is not novel in light of any of D1, D4 or D5 alone.

D1 discloses methods of producing, and extracting oil bodies from, the seeds of the plant *Arabidopsis thaliana*, where the plant is genetically modified to produce modified oleosins with artificially introduced cysteine residues and is cultivated to produce said modified oleosin in the seeds, among other tissues (Examples 5 and 6 of D1). Said method disclosed in D1 anticipates the method recited in present claim 35 because the present application defines "non-photosynthetic tissue" to include endosperm (lines 21-26, p10 of present description), which is a part of the seed. Moreover, the scope of present claim 35 is not limited to methods in which the modified oleosin is only expressed in the non-photosynthetic tissues of the plant; provision of a plant that produces oil in said tissues but expresses the modified oleosin in a different tissue would still fall within the scope of present claim 35. D1 also discloses a method in which the host cell is additionally genetically modified to express a triacylglycerol (TAG) synthesising enzyme (lines 23-27, p4) and a method in which the plant producing the modified oleosin is processed into an animal feed or biofuel (lines 16-17, p11). Further, it is considered that the additional features recited in present claims 37; 39; 67; 69; in which the

plant of present claim 35 accumulates 1.2x more oil or 10% more lipid in non-photosynthetic tissues than a control plant, would be an inevitable result of the skilled artisan carrying out the teaching of D1. Therefore claims 35-41; 44; 46-47; 50-53; 56-71; are not novel in light of D1.

D4 discloses a method of producing repeat oleosins in the oil bodies of a eukaryotic or prokaryotic cell and purifying these repeat oleosins from the oil bodies (line 25, p11- line 10, p12). D4 also discloses that in a preferred embodiment, the oleosin repeat unit may be from white clover or sesame seed (lines 25-27, p12). Since the oleosin sequences from white clover and sesame each contain at least one cysteine residue (see accession numbers 289719012 and AAG23840, respectively), oleosin repeat units derived from said sequences would include an artificially introduced cysteine residue in some embodiments. Thus, D4 discloses a method with all the features of the method defined in independent claim 35 of the present application. D4 further discloses that the repeat oleosins may be co-expressed with TAG synthesising enzymes such as DGAT1, DGAT2, or PDAT, in plant vegetative organs (lines 25-28, p13).Therefore claims 35; 44; 46-47; 50-53; 58-65; are not novel in light of the disclosure in D4, for similar reasons as for D1.

D5 discloses a method in which oleosin from *A. thaliana* is fused to the sequence of *Escherichia coli* beta-glucuronidase (GUS) and expressed in the oil bodies of seeds of *Brassica napus* (Abstract; Materials and Methods p1353-1354). The oil bodies containing the transgene are then extracted from the seeds of the transgenic plant (Materials and Methods, p1356-1357). Since the GUS sequence encodes several cysteine residues, and this sequence is fused to the oleosin sequence and expressed in the oil bodies of the seeds of a plant, D5 similarly discloses a method for producing oil in the non-photosynthetic tissues of a plant as recited in claim 35 of the present application. Therefore claims 35; 44; 46-47; 50; 58-59; 61-65; are not novel in light of D5.

In summary, claims 35-41; 44; 46-47; 50-53; 56-71; do not meet the requirements of Novelty under Article 33(2) of the PCT because the prior art discloses all the features of these claims.

In contrast, claims 1-34; 42-43; 45; 48-49; 54-55; meet the requirements of Novelty under Article 33(2) of the PCT. The prior art published before the priority date does not disclose the features defined in these claims.

INVENTIVE STEP (IS)

Claims 35-41; 44; 46-47; 50-53; 56-71; are already considered to be not novel, and cannot be considered inventive in light of the prior art disclosures as discussed above.

D1 discloses all the features of invention 2 defined in claims 35-41; 44; 46-47; 50-53; 56-71; as discussed above. The additional features of invention 2 as recited in dependent claims 42-43; 45; 48-49; 54-55; such as specific tissues in which the oil is produced and specific methods of transforming/generating a plant comprising and expressing a modified oleosin, are mere design choices and as such, do not add an inventive step over and above the teachings of D1.

In summary, claims 35-71; do not meet the requirements of Inventive Step under Article 33(2) of the PCT because the prior art discloses all the features of these claims or the additional features not disclosed in the prior art are mere design choices and do not add an inventive step.

In contrast, none of the above prior art documents either individually or in combination obviously suggests to a person skilled in the art a method of increasing the rate of CO_2 assimilation by transforming a photosynthetic cell with a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine. Therefore, claims 1-34 meet the requirements of Inventive step under Article 33(2) of the PCT.

INDUSTRIAL APPLICABILITY (IA)

The invention defined in the claims is considered to meet the requirements of industrial applicability under Article 33(4) of the PCT because it can be used in industry.

E Category Document/s listed in Box VI:

There is a document(s) listed as an E category document in Box VI because it:

- discloses subject matter that is of particular relevance to this application and
- is an earlier application or patent but published on or after the international filing date of this application

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY	International Application No. PCT/NZ2012/000138

Under the PCT, only documents published before the priority date of the instant application can deprive the claims of that application of novelty.

However, the relevance of a document published after the priority date of the application under consideration is dependent on national law in individual countries.

E category document/s may become significant in the National Phase.

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY	International Application No.
Box No. VI Certain documents cited	**PCT/NZ2012/000138**

1. Certain published documents (Rules 43*bis*.1 and 70.10)

Application No. Patent No.	Publication date *(day/month/year)*	Filing date *(day/month/year)*	Priority date (valid claim) *(day/month/year)*
E : US 2012/0278951 A1	01 November 2012	30 April 2012	30 October 2009

See Supplemental Box for Details

2. Non-written disclosures (Rules 43*bis*.1 and 70.9)

Kind of non-written disclosure	Date of non-written disclosure *(day/month/year)*	Date of written disclosure referring to non-written disclosure *(day/month/year)*

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY	International Application No.
Box No. VIII Certain observations in the international application	**PCT/NZ2012/000138**

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

Present claim 1 relates to a method for producing a photosynthetic cell with an increased rate of CO_2 assimilation comprising (in one alternative) genetically modifying the cell to reduce or prevent lipid recycling. However, the present description only provides evidence that said outcome may be achieved via transforming the cell to include a modified oleosin with at least one artificial cysteine residue (Example 7; Example 9; Example 11).

PCT Article 5 requires that the description shall disclose the invention in a manner sufficiently clear and complete for the invention to be carried out by a person skilled in the art. The PCT Guidelines state that the disclosure of the claimed invention is considered sufficiently clear and complete if it provides information which is sufficient to allow the invention to be carried out by a person skilled in the art as of the international filing date, without undue experimentation.

Although a reasonable amount of trial and error is permissible, a person skilled in the art must, on the basis of the disclosure of the claimed invention and the general knowledge, be able to carry out the invention across the entire scope of the claim without undue experimentation.

As explained above, the present description only has sufficient information for the skilled artisan to carry out said method by transforming a photosynthetic cell with a modified oleosin including at least one artificial cysteine. There is no evidence provided in the present description that modification of lipid recycling via any other method, as encompassed by the scope of claim 1, would lead to an increase in CO_2 assimilation, nor any suggestion of how to achieve said outcome by another mechanism. Indeed the present description specifically hypothesises that it is the co-expression of a modified oleosin containing engineered cysteine residue/s and enzymes such as DGAT that causes an increase in the assimilation of CO_2 (line 32 p34-line 4 p35). Thus, the skilled artisan could not carry out the invention across the entire scope of claim 1 without undue experimentation.

Therefore the description does not meet the requirements of Article 5 of the PCT with regards to it being sufficiently clear and complete for the invention to be carried out by a person skilled in the art. This authority has thus limited the search and examination to the inventions so far as they relate to a modified oleosin.

Claim 48 is unclear and appears to be missing an antecedent. For the purposes of examination this Authority has construed claim 48 to read "The method of claim 47 in which the promoter is non-photosynthetic tissue preferred".

WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY	International Application No.
Supplemental Box	PCT/NZ2012/000138

Continuation of: **Box IV**

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are two different inventions based on the following features that separate the claims into distinct groups:

Invention 1: Methods for producing a photosynthetic cell with an increased rate of CO_2 assimilation comprising genetically modifying the photosynthetic cell to reduce or prevent lipid recycling and/or transforming a photosynthetic cell with a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine. Invention 1 is defined by claims 1-34 (fully).

Invention 2: Methods for producing oil comprising providing a plant comprising a polynucleotide encoding a modified oleosin including at least one artificially introduced cysteine and cultivating the plant to express said oleosin in its non-photosynthetic tissues; a non-photosynthetic tissue/organ of a plant produced by the method; an animal feed or biofuel comprising said non-photosynthetic tissue/organ. Invention 2 is defined by claims 35-71 (fully).

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention.

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The only feature common to all of the claimed inventions and which provides a technical relationship among them is a modified oleosin including at least one artificially introduced cysteine.

However this feature does not make a contribution over the prior art because it is disclosed in:

D1: WO 2011/053169 A1 (AGRESEARCH LIMITED) 5 May 2011

Therefore in light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

NOTE: Despite the lack of unity, in the interest of expediting examination, all inventions have been searched and examined so far as they relate to a modified oleosin (see Article 5 objection).

Continuation of: **Box VI**

US 2012/0278951 discloses methods for increasing the rate of CO_2 assimilation in photosynthetic cells and plants, by reducing or preventing lipid recycling, and/or expressing modified oleosins with artificially introduced cysteine residues in photosynthetic cells and plants (Abstract; lines 25-28, page 5). US 2012/0278951 also discloses methods for increasing oil production in plants, via expression of modified oleosins with artificially introduced cysteine residues in the non-photosynthetic tissues/organs of plants and processing said organs into animal or biofuel feedstocks (lines 1-8, p6). Therefore, this document is relevant to claims 1-71 of the present application.

4197737_1
SEQUENCE LISTING

<110> AGRESEARCH LIMITED

<120> MODIFIED OIL ENCAPSULATING PROTEINS AND USES THEREOF

<130> 660921

<150> 61/515,610
<151> 2011-08-05

<160> 107

<170> PatentIn version 3.5

<210> 1
<211> 579
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 1
atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60

accctggtgc cacgcggttc catggctgag cattatggtc aacaacagca gaccagggcg 120

cctcacctgc agctgcagcc gcgcgcccag cgggtagtga aggcggccac cgccgtgaca 180

gccggcggct cgcttctcgt cctctctggc ctcactttag ccggaactgt tattgcgctc 240

accatcgcca ctccgctgct tgtgatcttt agccccgttc tggtgccggc ggtcataacc 300

attttcttgc tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt 360

gtgctgtcgt ggatttacag atatctgaca gggaaacacc cgccgggggc ggatcagctg 420

gaatcggcaa agacgaagct ggcgagcaag gcgcgagaga tgaaggatag ggcagagcag 480

ttctcgcagc agcctgttcc atggctgata tcggatccga attcgagctc cgtcgacaag 540

cttgcggccg cactcgagca ccaccaccac caccactga 579

<210> 2
<211> 531
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 2
atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60

accctggtgc cacgcggttc catggcttgt cattatggtc aacaacagca gaccagggcg 120

cctcacctgc agctgcagcc gcgcgcccag cgggtagtga aggcggccac cgccgtgaca 180

gccggcggat cccttctcgt cctctctggc ctcactttag ccggaactgt tattgcgctc 240

accatcgcca ctccgctgct tgtgatcttt agccccgttc tggtgccggc ggtcataacc 300

attttcttgc tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt 360

gtgctgtcgt ggatttacag atatctgaca gggaaacacc cgccgggggc ggatcagctg 420

gaatcggcaa agacgaagct ggcgagcaag gcgcgagaga tgaaggatag ggcagagcag 480

ttctcgtgcc agcctgttcc gtggctcgag caccaccacc accaccactg a 531

<210> 3
<211> 531
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 3
atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60

accctggtgc cacgcggttc catggcttgt cattatggtc aacaacagca gaccagggcg 120

cctcacctgc agctgcagcc gcgcgcccag cgggtagtga aggcggccac cgccgtgaca 180

gccggcggat cccttctcgt cctctctggc ctcactttag ccggaactgt tattgcgctc 240

accatcgcca ctccgctgct tgtgatcttt agccccgttc tggtgccggc ggtcataacc 300

attttcttgc tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt 360

gtgctgtcgt ggatttacag atatctgaca gggaaacacc cgccgggggc ggattgcctg 420

gaatcggcaa agacgaagct ggcgagctgt gcgcgagaga tgaaggatag ggcagagcag 480

ttctcgtgcc agcctgttcc gtggctcgag caccaccacc accaccactg a 531

<210> 4
<211> 531
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 4
atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60

accctggtgc cacgcggttc catggcttgt cattatggtc aacaacagca gacctgcgcg 120

cctcacctgc agctgcagcc gcgcgcctgt cgggtagtga aggcggccac cgccgtgaca 180

gcgggcggat cccttctcgt cctctctggc ctcactttag ccggtaccgt tattgcgctc 240

accatcgcca ctccgctgct tgtgatcttt agccccgttc tggttccggc ggtcataacc 300

attttcttgc tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt 360

gtgctgtcgt ggatttacag atatctgaca gggaaacacc cgccgggggc ggatcagctg 420

gaatcggcaa agacgaagct ggcgagcaag gcgcgagaga tgaaggatag ggcagagcag 480

ttctcgtgcc agcctgttcc gtggctcgag caccaccacc accaccactg a 531

<210> 5
<211> 531
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 5

4197737_1

atgaaagaaa ccgctgctgc taaattcgaa cgccagcaca tggacagccc agatctgggt 60

accctggtgc cacgcggttc catggcttgt cattatggtc aacaacagca gacctgcgcg 120

cctcacctgc agctgcagcc gcgcgcctgt cgggtagtga aggcggccac cgccgtgaca 180

gcgggcggat cccttctcgt cctctctggc ctcactttag ccggtaccgt tattgcgctc 240

accatcgcca ctccgctgct tgtgatcttt agccccgttc tggttccggc ggtcataacc 300

attttcttgc tgggtgcggg ttttctggca tccggaggct tcggcgtggc ggcgctgagt 360

gtgctgtcgt ggatttacag atatctgaca gggaaacacc cgccgggggc ggattgcctg 420

gaatcggcaa agacgaagct ggcgagctgt gcgcgagaga tgaaggatag ggcagagcag 480

ttctcgtgcc agcctgttcc gtggctcgag caccaccacc accaccactg a 531

<210> 6
<211> 193
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 6

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Glu His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
50 55 60

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

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
85 90 95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

4197737_1

Phe Ser Gln Gln Pro Val Pro Trp Leu Ile Ser Asp Pro Asn Ser Ser
165 170 175

Ser Val Asp Lys Leu Ala Ala Ala Leu Glu His His His His His His
180 185 190

His

<210> 7
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 7

Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1 5 10 15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Cys His Tyr
20 25 30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
35 40 45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
50 55 60

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

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
85 90 95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

Phe Ser Cys Gln Pro Val Pro Trp Leu Glu His His His His His His
165 170 175

<210> 8
<211> 176
<212> PRT

4197737_1

<213> Artificial Sequence

<220>
<223> Artificial

<400> 8

```
Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1               5                  10                  15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Cys His Tyr
            20                  25                  30

Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu Gln Leu Gln Pro Arg
        35                  40                  45

Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
    50                  55                  60

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

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
                85                  90                  95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
            100                 105                 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
        115                 120                 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Cys Leu Glu Ser Ala Lys
    130                 135                 140

Thr Lys Leu Ala Ser Cys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145                 150                 155                 160

Phe Ser Cys Gln Pro Val Pro Trp Leu Glu His His His His His His
                165                 170                 175
```

<210> 9
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 9

```
Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1               5                  10                  15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Cys His Tyr
            20                  25                  30
```

4197737_1

```
Gly Gln Gln Gln Gln Thr Cys Ala Pro His Leu Gln Leu Gln Pro Arg
        35                  40                  45

Ala Cys Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
    50                  55                  60

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

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
                85                  90                  95

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
            100                 105                 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
        115                 120                 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln Leu Glu Ser Ala Lys
    130                 135                 140

Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145                 150                 155                 160

Phe Ser Cys Gln Pro Val Pro Trp Leu Glu His His His His His His
                165                 170                 175
```

<210> 10
<211> 176
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 10

```
Met Lys Glu Thr Ala Ala Ala Lys Phe Glu Arg Gln His Met Asp Ser
1               5                   10                  15

Pro Asp Leu Gly Thr Leu Val Pro Arg Gly Ser Met Ala Cys His Tyr
            20                  25                  30

Gly Gln Gln Gln Gln Thr Cys Ala Pro His Leu Gln Leu Gln Pro Arg
        35                  40                  45

Ala Cys Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser
    50                  55                  60

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

Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro
                85                  90                  95
```

4197737_1

Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly
100 105 110

Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr
115 120 125

Leu Thr Gly Lys His Pro Pro Gly Ala Asp Cys Leu Glu Ser Ala Lys
130 135 140

Thr Lys Leu Ala Ser Cys Ala Arg Glu Met Lys Asp Arg Ala Glu Gln
145 150 155 160

Phe Ser Cys Gln Pro Val Pro Trp Leu Glu His His His His His His
165 170 175

<210> 11
<211> 766
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 11
ttgctccctt aaaaaaaacc atggctgagc attatggtca acaacagcag accagggcgc 60

ctcacctgca gctgcagccg cgcgcccagc gggtagtgaa ggcggccacc gccgtgacag 120

taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 180

ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 240

ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 300

tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 360

ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 420

caggccggcg gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg 480

ctcaccatcg ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata 540

accattttct tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg 600

agtgtgctgt cgtggattta cagatatctg acagggaaac acccgccggg ggcggatcag 660

ctggaatcgg caaagacgaa gctggcgagc aaggcgcgag agatgaagga tagggcagag 720

cagttctcgc agcagcctgt tgcggggtct caaacttctt aatgaa 766

<210> 12
<211> 766
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 12
ttgctccctt aaaaaaaacc atggcttgtc attatggtca acaacagcag accagggcgc 60

ctcacctgca gctgcagccg cgcgcccagc gggtagtgaa ggcggccacc gccgtgacag 120

taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 180

ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 240

ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 300

tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 360

ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 420

caggccggcg gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg 480

ctcaccatcg ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata 540

accattttct tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg 600

agtgtgctgt cgtggattta cagatatctg acagggaaac acccgccggg ggcggatcag 660

ctggaatcgg caaagacgaa gctggcgagc aaggcgcgag agatgaagga tagggcagag 720

cagttctcgt gtcagcctgt tgcggggtct caaacttctt aatgaa 766

<210> 13
<211> 766
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 13
ttgctccctt aaaaaaaacc atggcttgtc attatggtca acaacagcag accagggcgc 60

ctcacctgca gctgcagccg cgcgcccagc gggtagtgaa ggcggccacc gccgtgacag 120

taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 180

ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 240

ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 300

tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 360

ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 420

caggccggcg gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg 480

ctcaccatcg ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata 540

accattttct tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg 600

agtgtgctgt cgtggattta cagatatctg acagggaaac acccgccggg ggcggattgt 660

ctggaatcgg caaagacgaa gctggcgagc tgtgcgcgag agatgaagga tagggcagag 720

cagttctcgt gtcagcctgt tgcggggtct caaacttctt aatgaa 766

<210> 14
<211> 766
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 14

ttgctccctt aaaaaaaacc atggcttgtc attatggtca acaacagcag acctgtgcgc 60

ctcacctgca gctgcagccg cgcgcctgtc gggtagtgaa ggcggccacc gccgtgacag 120

taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 180

ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 240

ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 300

tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 360

ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 420

caggccggcg gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg 480

ctcaccatcg ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata 540

accattttct tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg 600

agtgtgctgt cgtggattta cagatatctg acagggaaac acccgccggg ggcggatcag 660

ctggaatcgg caaagacgaa gctggcgagc aaggcgcgag agatgaagga tagggcagag 720

cagttctcgt gtcagcctgt tgcggggtct caaacttctt aatgaa 766

<210> 15
<211> 766
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 15
ttgctccctt aaaaaaaacc atggcttgtc attatggtca acaacagcag acctgtgcgc 60

ctcacctgca gctgcagccg cgcgcctgtc gggtagtgaa ggcggccacc gccgtgacag 120

taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 180

ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 240

ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 300

tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 360

ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 420

caggccggcg gctcgcttct cgtcctctct ggcctcactt tagccggaac tgttattgcg 480

ctcaccatcg ccactccgct gcttgtgatc tttagccccg ttctggtgcc ggcggtcata 540

accattttct tgctgggtgc gggttttctg gcatccggag gcttcggcgt ggcggcgctg 600

agtgtgctgt cgtggattta cagatatctg acagggaaac acccgccggg ggcggattgt 660

ctggaatcgg caaagacgaa gctggcgagc tgtgcgcgag agatgaagga tagggcagag 720

cagttctcgt gtcagcctgt tgcggggtct caaacttctt aatgaa 766

<210> 16
<211> 145
<212> PRT
<213> Artificial Sequence

<220>

<223> Artificial

<400> 16

```
Met Ala Glu His Tyr Gly Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                  10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65                  70                  75                  80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
                85                  90                  95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
            100                 105                 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
        115                 120                 125

Asp Arg Ala Glu Gln Phe Ser Gln Gln Pro Val Ala Gly Ser Gln Thr
    130                 135                 140

Ser
145
```

<210> 17
<211> 145
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 17

```
Met Ala Cys His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1               5                  10                  15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
            20                  25                  30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
        35                  40                  45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
    50                  55                  60
```

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Cys Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser
145

<210> 18
<211> 145
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 18

Met Ala Cys His Tyr Gly Gln Gln Gln Gln Thr Arg Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Gln Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Cys
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Cys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Cys Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser
145

<210> 19
<211> 145
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 19

Met Ala Cys His Tyr Gly Gln Gln Gln Gln Thr Cys Ala Pro His Leu
1 5 10 15

Gln Leu Gln Pro Arg Ala Cys Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Gln
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Lys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Cys Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser
145

<210> 20
<211> 145
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 20

Met Ala Cys His Tyr Gly Gln Gln Gln Gln Thr Cys Ala Pro His Leu
1 5 10 15

4197737_1

Gln Leu Gln Pro Arg Ala Cys Arg Val Val Lys Ala Ala Thr Ala Val
20 25 30

Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr Leu Ala Gly
35 40 45

Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val Ile Phe Ser
50 55 60

Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu Gly Ala Gly
65 70 75 80

Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser Val Leu Ser
85 90 95

Trp Ile Tyr Arg Tyr Leu Thr Gly Lys His Pro Pro Gly Ala Asp Cys
100 105 110

Leu Glu Ser Ala Lys Thr Lys Leu Ala Ser Cys Ala Arg Glu Met Lys
115 120 125

Asp Arg Ala Glu Gln Phe Ser Cys Gln Pro Val Ala Gly Ser Gln Thr
130 135 140

Ser
145

<210> 21
<211> 810
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 21
ggtaccgcgg ccgcttgctc ccttaaaaaa aaccatggca tgtcattacg gatgtggaca 60

acagcaacag acttgtgcac ctcatctttg tcagcttcaa cctagagcat gtagagttgt 120

gaaggctgct actgctgtta ctgtaaattt ctgtgttcct tattctctca aaatcttcga 180

ttttgttttc gttcgatccc aatttcgtat atgttctttg gtttagattc tgttaatctt 240

agatcgaaga cgattttctg ggtttgatcg ttagatatca tcttaattct cgattagggt 300

ttcatagata tcatccgatt tgttcaaata atttgagttt tgtcgaataa ttactcttcg 360

atttgtgatt tctatctaga tctggtgtta gtttctagtt tgtgcgatcg aatttgtcga 420

ttaatctgag tttttctgat taacaggctg gaggatctct tcttgttctc tctggactta 480

ctctcgctgg aactgttatc gctctcacta tcgctacacc tcttctcgtt atcttctctc 540

ctgttctcgt tcctgctgtg atcactatct tccttctcgg agctggattt cttgcttctg 600

gtggatttgg agttgctgct ctctctgttc tctcttggat ctacagatac ctcactggat 660

gtaaacatcc tccaggtgct gattgtcttg agtctgcttg taagactaag ctcgcttctt 720

gtgctagaga gatgaaggat tgtagagcag agcaattctc ttgtcagcct gttgctggat 780

ctcagacttc ttaatgaaca tatggagctc 810

<210> 22
<211> 831
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 22
ggtaccgcgg ccgcttgctc ccttaaaaaa aaccatggca tgtcattacg gaatggcatg 60

tcattacgga tgtggacaac agcaacagac ttgtgcacct catctttgtc agcttcaacc 120

tagagcatgt agagttgtga aggctgctac tgctgttact gtaaatttct gtgttcctta 180

ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa tttcgtatat gttctttggt 240

ttagattctg ttaatcttag atcgaagacg attttctggg tttgatcgtt agatatcatc 300

ttaattctcg attagggttt catagatatc atccgatttg ttcaaataat ttgagttttg 360

tcgaataatt actcttcgat ttgtgatttc tatctagatc tggtgttagt ttctagtttg 420

tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta acaggctgga ggatctcttc 480

ttgttctctc tggacttact ctcgctggaa ctgttatcgc tctcactatc gctacacctc 540

ttctcgttat cttctctcct gttctcgttc ctgctgtgat cactatcttc cttctcggag 600

ctggatttct tgcttctggt ggatttggag ttgctgctct ctctgttctc tcttggatct 660

acagatacct cactggatgt aaacatcctc caggtgctga ttgtcttgag tctgcttgta 720

agactaagct cgcttcttgt gctagagaga tgaaggattg tagagcagag caattctctt 780

gtcagcctgt tgctggatgt tctcagactt cttaatgaac atatggagct c 831

<210> 23
<211> 151
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 23

Met Ala Cys His Tyr Gly Cys Gly Gln Gln Gln Thr Cys Ala Pro
1 5 10 15

His Leu Cys Gln Leu Gln Pro Arg Ala Cys Arg Val Val Lys Ala Ala
20 25 30

Thr Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Gly Leu Thr
35 40 45

Leu Ala Gly Thr Val Ile Ala Leu Thr Ile Ala Thr Pro Leu Leu Val
50 55 60

```
Ile Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Phe Leu Leu
65                  70                  75                  80

Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe Gly Val Ala Ala Leu Ser
                85                  90                  95

Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr Gly Cys Lys His Pro Pro
            100                 105                 110

Gly Ala Asp Cys Leu Glu Ser Ala Cys Lys Thr Lys Leu Ala Ser Cys
        115                 120                 125

Ala Arg Glu Met Lys Asp Cys Arg Ala Glu Gln Phe Ser Cys Gln Pro
    130                 135                 140

Val Ala Gly Ser Gln Thr Ser
145                 150
```

<210> 24
<211> 158
<212> PRT
<213> Artificial Sequence

<220>
<223> Artificial

<400> 24

```
Met Ala Cys His Tyr Gly Met Ala Cys His Tyr Gly Cys Gly Gln Gln
1               5                   10                  15

Gln Gln Thr Cys Ala Pro His Leu Cys Gln Leu Gln Pro Arg Ala Cys
            20                  25                  30

Arg Val Val Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Leu Leu
        35                  40                  45

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

Ala Thr Pro Leu Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Val
65                  70                  75                  80

Ile Thr Ile Phe Leu Leu Gly Ala Gly Phe Leu Ala Ser Gly Gly Phe
                85                  90                  95

Gly Val Ala Ala Leu Ser Val Leu Ser Trp Ile Tyr Arg Tyr Leu Thr
            100                 105                 110

Gly Cys Lys His Pro Pro Gly Ala Asp Cys Leu Glu Ser Ala Cys Lys
        115                 120                 125

Thr Lys Leu Ala Ser Cys Ala Arg Glu Met Lys Asp Cys Arg Ala Glu
    130                 135                 140
```

4197737_1

Gln Phe Ser Cys Gln Pro Val Ala Gly Cys Ser Gln Thr Ser
145 150 155

<210> 25
<211> 4104
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 25
ggtaccacaa gtttgtacaa aaaagcaggc tgcggccgca aaattttcat atggtcctgc 60

tttaatgaga tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc 120

acgttgtaaa aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc 180

taatgaatat atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta 240

ctgattgtac cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga 300

ataggtttat agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta 360

ttacaaatcc aattttaaaa aaagcggcag aaccggtcaa acctaaaaga ctgattacat 420

aaatcttatt caaatttcaa aaggccccag gggctagtat ctacgacaca ccgagcggcg 480

aactaataac gttcactgaa gggaactccg gttccccgcc ggcgcgcatg ggtgagattc 540

cttgaagttg agtattggcc gtccgctcta ccgaaagtta cgggcaccat tcaacccggt 600

ccagcacggc ggccgggtaa ccgacttgct gccccgagaa ttatgcagca ttttttggt 660

gtatgtgggc cccaaatgaa gtgcaggtca aaccttgaca gtgacgacaa atcgttgggc 720

gggtccaggg cgaattttgc gacaacatgt cgaggctcag caggacctgc aggcatgcaa 780

gctagcttac tagtgatgca tattctatag tgtcacctaa atcttcgacg aattaattcc 840

aatcccacaa aaatctgagc ttaacagcac agttgctcct ctcagagcag aatcgggtat 900

tcaacaccct catatcaact actacgttgt gtataacggt ccacatgccg gtatatacga 960

tgactggggt tgtacaaagg cggcaacaaa cggcgttccc ggagttgcac acaagaaatt 1020

tgccactatt acagaggcaa gagcagcagc tgacgcgtac acaacaagtc agcaaacaga 1080

caggttgaac ttcatcccca aaggagaagc tcaactcaag cccaagagct ttgctaaggc 1140

cctaacaagc ccaccaaagc aaaaagccca ctggctcacg ctaggaacca aaaggcccag 1200

cagtgatcca gccccaaaag agatctcctt tgccccggag attacaatgg acgatttcct 1260

ctatctttac gatctaggaa ggaagttcga aggtgaaggt gacgacacta tgttcaccac 1320

tgataatgag aaggttagcc tcttcaattt cagaaagaat gctgacccac agatggttag 1380

agaggcctac gcagcaggtc tcatcaagac gatctacccg agtaacaatc tccaggagat 1440

caaatacctt cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta attgcatcaa 1500

gaacacagag aaagacatat ttctcaagat cagaagtact attccagtat ggacgattca 1560

aggcttgctt cataaaccaa ggcaagtaat agagattgga gtctctaaaa aggtagttcc 1620

tactgaatct aaggccatgc atggagtcta agattcaaat cgaggatcta acagaactcg 1680

ccgtgaagac tggcgaacag ttcatacaga gtcttttacg actcaatgac aagaagaaaa 1740

tcttcgtcaa catggtggag cacgacactc tggtctactc caaaaatgtc aaagatacag 1800

tctcagaaga ccaaagggct attgagactt ttcaacaaag gataatttcg ggaaacctcc 1860

tcggattcca ttgcccagct atctgtcact tcatcgaaag gacagtagaa aaggaaggtg 1920

gctcctacaa atgccatcat tgcgataaag gaaaggctat cattcaagat ctctctgccg 1980

acagtggtcc caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc 2040

caaccacgtc ttcaaagcaa gtggattgat gtgacatctc cactgacgta agggatgacg 2100

cacaatccca ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg 2160

agaggacacg ggatccttgc tccgttaaaa aaaaccatgg ctatcctcga ttctgctggt 2220

gttactactg tgactgagaa tggtggtgga gagttcgttg atctcgatag actcagaaga 2280

agaaagtcta gatctgtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt 2340

ttcgttcgat cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga 2400

agacgatttt ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag 2460

atatcatccg atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg 2520

atttctatct agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct 2580

gagtttttct gattaacagg attcttctaa cggacttctc ctctctggat ctgataacaa 2640

ctctccttct gatgatgttg gtgctcctgc tgatgtgaga gatagaatcg attctgtggt 2700

gaacgatgat gctcaaggta ctgctaacct cgctggtgat aataacggtg gaggtgataa 2760

caatggtgga ggaagaggtg gaggtgaagg tagaggaaac gctgatgcta ctttcactta 2820

cagaccatct gtgcctgctc atagaagagc tagagagtct cctctctctt ctgatgctat 2880

cttcaagcag tctcacgctg gacttttcaa cctctgtgtg gtggttctta tcgctgtgaa 2940

ctctagactc atcatcgaga acctcatgaa gtacggatgg ctcatcagaa ctgatttctg 3000

gttctcttct agatctctca gagattggcc tcttttcatg tgctgcatct cactctcaat 3060

cttccctctc gctgctttta ctgttgagaa gctcgtgctc cagaagtata tcgctgaacc 3120

tgtggtgatc ttcctccaca tcatcatcac tatgactgag gttctctacc ctgtttacgt 3180

gactctcaga tgcgattctg ctttcctctc tggtgttact cttatgctcc tcacttgcat 3240

tgtgtggctt aagctcgtgt cttacgctca cacttcttac gatatcagat ctctcgctaa 3300

cgctgctgat aaggctaacc ctgaagtgtc ttactacgtg tctctcaagt ctctcgctta 3360

cttcatggtt gctcctacac tttgttacca gccatcttac cctagatctg cttgcattag 3420

aaagggatgg gtggcaagac aattcgctaa gttggtgatc ttcactggat tcatgggatt 3480

catcatcgag cagtacatca accctattgt gagaaactct aagcaccctc tcaagggtga 3540

tcttctctac gctatcgaga gagttcttaa gctctctgtg cctaaccttt atgtgtggct 3600

ctgcatgttc tactgtttct tccacctctg gcttaacatc cttgctgagt tgctttgctt 3660

cggagataga gagttctaca aggattggtg gaacgctaag tctgttggag attattggag 3720

aatgtggaac atgcctgtgc ataagtggat ggtgcgtcac atctacttcc cttgcctcag 3780

atctaagatc cctaagactc tcgctatcat tatcgctttc ctcgtgtctg ctgttttcca 3840

tgagttgtgt atcgctgttc cttgcagact tttcaagctt tgggctttcc tcggaatcat 3900

gttccaggtt ccactcgtgt tcatcactaa ctacctccaa gagagattcg gatctactgt 3960

tggaaacatg attttctggt tcattttctg catcttcgga cagcctatgt gcgttctcct 4020

ctactaccac gatctcatga acagaaaggg atctatgtct taatgaagga tccacccagc 4080

tttcttgtac aaagtggtga gctc 4104

<210> 26
<211> 18784
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 26
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60

atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120

gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420

atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480

ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540

aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600

aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660

gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720

agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780

ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840

aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900

cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960

cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020

cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080

cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140

gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200

catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260

cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320

gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380

ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440

ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500

cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560

gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620

cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680

ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740

tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800

acttgctttg aagacgtggt tggaacgtct tcttttttcca cgatgctcct cgtgggtggg 1860

ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920

caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980

atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040

atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100

tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160

ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220

tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280

tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340

gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400

atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460

gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520

acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580

atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640

attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700

attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760

tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820

tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880

caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940

aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000

ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060

ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120

aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180

aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240

aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300

tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360

aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420

gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480
catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540
accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600
catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660
aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720
tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780
gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840
cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900
atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960
atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020
agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080
cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140
ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200
taaaaaaaac catggcagag cattacggac aacagcaaca gactagagca cctcatcttc 4260
agcttcaacc tagagcacag agagttgtga aggctgctac tgctgttact gtaaatttct 4320
gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa tttcgtatat 4380
gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg tttgatcgtt 4440
agatatcatc ttaattctcg attagggttt catagatatc atccgatttg ttcaaataat 4500
ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc tggtgttagt 4560
ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta acaggctgga 4620
ggatctcttc ttgttctctc tggacttact ctcgctggaa ctgttatcgc tctcactatc 4680
gctacacctc ttctcgttat cttctctcct gttctcgttc ctgctgtgat cactatcttc 4740
cttctcggag ctggatttct tgcttctggt ggatttggag ttgctgctct ctctgttctc 4800
tcttggatct acagatacct cactggaaaa catcctccag gtgctgatca acttgagtct 4860
gctaagacta agctcgcttc taaggctaga gagatgaagg atagagcaga gcaattctct 4920
caacagcctg ttgctggatc tcagacttct taatgaacat atggtcctgc tttaatgaga 4980
tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 5040
aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 5100
atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtac 5160
cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga ataggtttat 5220
agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta ttacaaatcc 5280
aattttaaaa aaagcggcag aaccggtcaa acctaaaaga ctgattacat aaatcttatt 5340
caaatttcaa aaggccccag gggctagtat ctacgacaca ccgagcggcg aactaataac 5400
gttcactgaa gggaactccg gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg 5460

agtattggcc gtccgctcta ccgaaagtta cgggcaccat tcaacccggt ccagcacggc 5520

ggccgggtaa ccgacttgct gccccgagaa ttatgcagca ttttttggt gtatgtgggc 5580

cccaaatgaa gtgcaggtca aaccttgaca gtgacgacaa atcgttgggc gggtccaggg 5640

cgaattttgc gacaacatgt cgaggctcag caggacctgc aggcatgcaa gctagcttac 5700

tagtgatgca tattctatag tgtcacctaa atcttcgacg aattaattcc aatcccacaa 5760

aaatctgagc ttaacagcac agttgctcct ctcagagcag aatcgggtat tcaacaccct 5820

catatcaact actacgttgt gtataacggt ccacatgccg gtatatacga tgactggggt 5880

tgtacaaagg cggcaacaaa cggcgttccc ggagttgcac acaagaaatt tgccactatt 5940

acagaggcaa gagcagcagc tgacgcgtac acaacaagtc agcaaacaga caggttgaac 6000

ttcatcccca aaggagaagc tcaactcaag cccaagagct ttgctaaggc cctaacaagc 6060

ccaccaaagc aaaaagccca ctggctcacg ctaggaacca aaaggcccag cagtgatcca 6120

gccccaaaag agatctcctt tgccccggag attacaatgg acgatttcct ctatctttac 6180

gatctaggaa ggaagttcga aggtgaaggt gacgacacta tgttcaccac tgataatgag 6240

aaggttagcc tcttcaattt cagaaagaat gctgacccac agatggttag agaggcctac 6300

gcagcaggtc tcatcaagac gatctacccg agtaacaatc tccaggagat caaatacctt 6360

cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta attgcatcaa gaacacagag 6420

aaagacatat ttctcaagat cagaagtact attccagtat ggacgattca aggcttgctt 6480

cataaaccaa ggcaagtaat agagattgga gtctctaaaa aggtagttcc tactgaatct 6540

aaggccatgc atggagtcta agattcaaat cgaggatcta acagaactcg ccgtgaagac 6600

tggcgaacag ttcatacaga gtcttttacg actcaatgac aagaagaaaa tcttcgtcaa 6660

catggtggag cacgacactc tggtctactc caaaaatgtc aaagatacag tctcagaaga 6720

ccaaagggct attgagactt ttcaacaaag gataatttcg ggaaacctcc tcggattcca 6780

ttgcccagct atctgtcact tcatcgaaag gacagtagaa aaggaaggtg gctcctacaa 6840

atgccatcat tgcgataaag gaaaggctat cattcaagat ctctctgccg acagtggtcc 6900

caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 6960

ttcaaagcaa gtggattgat gtgacatctc cactgacgta agggatgacg cacaatccca 7020

ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggacacg 7080

ggatccttgc tccgttaaaa aaaaccatgg ctatcctcga ttctgctggt gttactactg 7140

tgactgagaa tggtggtgga gagttcgttg atctcgatag actcagaaga agaaagtcta 7200

gatctgtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt ttcgttcgat 7260

cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga agacgatttt 7320

ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag atatcatccg 7380

atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg atttctatct 7440

agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct gagtttttct 7500

gattaacagg attcttctaa cggacttctc ctctctggat ctgataacaa ctctccttct 7560

gatgatgttg gtgctcctgc tgatgtgaga gatagaatcg attctgtggt gaacgatgat 7620
gctcaaggta ctgctaacct cgctggtgat aataacggtg gaggtgataa caatggtgga 7680
ggaagaggtg gaggtgaagg tagaggaaac gctgatgcta ctttcactta cagaccatct 7740
gtgcctgctc atagaagagc tagagagtct cctctctctt ctgatgctat cttcaagcag 7800
tctcacgctg gacttttcaa cctctgtgtg gtggttctta tcgctgtgaa ctctagactc 7860
atcatcgaga acctcatgaa gtacggatgg ctcatcagaa ctgatttctg gttctcttct 7920
agatctctca gagattggcc tcttttcatg tgctgcatct cactctcaat cttccctctc 7980
gctgctttta ctgttgagaa gctcgtgctc cagaagtata tcgctgaacc tgtggtgatc 8040
ttcctccaca tcatcatcac tatgactgag gttctctacc ctgtttacgt gactctcaga 8100
tgcgattctg ctttcctctc tggtgttact cttatgctcc tcacttgcat tgtgtggctt 8160
aagctcgtgt cttacgctca cacttcttac gatatcagat ctctcgctaa cgctgctgat 8220
aaggctaacc ctgaagtgtc ttactacgtg tctctcaagt ctctcgctta cttcatggtt 8280
gctcctacac tttgttacca gccatcttac cctagatctg cttgcattag aaagggatgg 8340
gtggcaagac aattcgctaa gttggtgatc ttcactggat tcatgggatt catcatcgag 8400
cagtacatca accctattgt gagaaactct aagcaccctc tcaagggtga tcttctctac 8460
gctatcgaga gagttcttaa gctctctgtg cctaaccttt atgtgtggct ctgcatgttc 8520
tactgtttct tccacctctg gcttaacatc cttgctgagt tgctttgctt cggagataga 8580
gagttctaca aggattggtg gaacgctaag tctgttggag attattggag aatgtggaac 8640
atgcctgtgc ataagtggat ggtgcgtcac atctacttcc cttgcctcag atctaagatc 8700
cctaagactc tcgctatcat tatcgctttc ctcgtgtctg ctgttttcca tgagttgtgt 8760
atcgctgttc cttgcagact tttcaagctt tgggctttcc tcggaatcat gttccaggtt 8820
ccactcgtgt tcatcactaa ctacctccaa gagagattcg gatctactgt tggaaacatg 8880
attttctggt tcattttctg catcttcgga cagcctatgt gcgttctcct ctactaccac 8940
gatctcatga acagaaaggg atctatgtct taatgaagga tccacccagc tttcttgtac 9000
aaagtggtga tgggttcgaa atcgataagc ttggatcctc tagagtcctg ctttaatgag 9060
atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa 9120
aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata 9180
tatcacccgt tactatcgta tttttatgaa taatattctc cgttcaattt actgattgta 9240
ccctactact tatatgtaca atattaaaat gaaaacaata tattgtgctg aataggttta 9300
tagcgacatc tatgatagag cgccacaata acaaacaatt gcgttttatt attacaaatc 9360
caattttaaa aaaagcggca gaaccggtca aacctaaaag actgattaca taaatcttat 9420
tcaaatttca aaaggcccca ggggctagta tctacgacac accgagcggc gaactaataa 9480
cgttcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt ccttgaagtt 9540
gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaacccgg tccagcacgg 9600

cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg tgtatgtggg 9660

ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg cgggtccagg 9720

gcgaattttg cgacaacatg tcgaggctca gcaggacctg caggcatgca agctagctta 9780

ctagtgatgc atattctata gtgtcaccta aatctgcggc cgcctgcagg tcgatatggg 9840

agagctccca acgcgttgga tgcatagctt gagtattcta tagtgtcacc taaatagctt 9900

ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 9960

caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 10020

cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 10080

gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggg ctgagtggct 10140

ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg gcttgtcccg cgtcatcggc 10200

gggggtcata acgtgactcc cttaattctc cgctcatgat cagattgtcg tttcccgcct 10260

tcagtttaaa ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag 10320

cgtttattag aataatcgga tatttaaaag ggcgtgaaaa ggtttatccg ttcgtccatt 10380

tgtatgtgca tgccaaccac agggttcccc tcgggagtgc ttggcattcc gtgcgataat 10440

gacttctgtt caaccaccca aacgtcggaa agcctgacga cggagcagca ttccaaaaag 10500

atcccttggc tcgtctgggt cggctagaag gtcgagtggg ctgctgtggc ttgatccctc 10560

aacgcggtcg cggacgtagc gcagcgccga aaaatcctcg atcgcaaatc cgacgctgtc 10620

gaaaatcgtg atctgcttgt cgctctttcg gccgacgtcc tggccagtca tcacgcgcca 10680

aagttccgtc acaggatgat ctggcgcgag ttgctggatc tcgccttcaa tccgggtctg 10740

tggcgggaac tccacgaaaa tatccgaacg cagcaagatg tcgacggatc ttttccgctg 10800

cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc tttttcgcac 10860

gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc caacggcgtc 10920

agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc ttcactgtcc 10980

cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg ccggctaccg 11040

ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc aggggtgatg 11100

ctgccaactt actgatttag tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt 11160

tctatcagct gtccctcctg ttcagctact gacggggtgg tgcgtaacgg caaaagcacc 11220

gccggacatc agcgctatct ctgctctcac tgccgtaaaa catggcaact gcagttcact 11280

tacaccgctt ctcaacccgg tacgcaccag aaaatcattg atatggccat gaatggcgtt 11340

ggatgccggg caacagcccg cattatgggc gttggcctca acacgatttt acgtcactta 11400

aaaaactcag gccgcagtcg gtaacctcgc gcatacagcc gggcagtgac gtcatcgtct 11460

gcgcggaaat ggacgaacag tggggctatg tcggggctaa atcgcgccag cgctggctgt 11520

tttacgcgta tgacagtctc cggaagacgg ttgttgcgca cgtattcggt gaacgcacta 11580

tggcgacgct ggggcgtctt atgagcctgc tgtcaccctt tgacgtggtg atatggatga 11640

cggatggctg gccgctgtat gaatcccgcc tgaagggaaa gctgcacgta atcagcaagc 11700

gatatacgca gcgaattgag cggcataacc tgaatctgag gcagcacctg gcacggctgg 11760
gacggaagtc gctgtcgttc tcaaaatcgg tggagctgca tgacaaagtc atcgggcatt 11820
atctgaacat aaaacactat caataagttg gagtcattac ccaaccagga agggcagccc 11880
acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa aggcggcggc 11940
ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca aaatcacggg 12000
cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc tgggccgcct 12060
gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt tcggtgatgc 12120
cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg gcaaggtcat 12180
gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa aacggccggg 12240
gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc gacttcgcgg 12300
agctggtatt cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg 12360
tctacgggac cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg 12420
ggtcaaatca ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag 12480
ggtgaatgaa tcggacgttt gaccggaagg catacaggca agaactgatc gacgcggggt 12540
tttccgccga ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa 12600
ccttccagtc cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg 12660
tgcaactggc tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt tcgcgtcgtc 12720
tcgaacagga ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga 12780
cgaccaagaa gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc 12840
aggccgcgtt gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg 12900
atattgcgcc gtggccggac acgatgcgag cgatgccaaa cgacacggcc cgctctgccc 12960
tgttcaccac gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc 13020
acgtcaacaa ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgacg 13080
aactggtgtg gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca 13140
ccttcacgtt ctacgagctt tgccaggacc tgggctggtc gatcaatggc cggtattaca 13200
cgaaggccga ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc 13260
gcgttgggca cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca 13320
agaaaacgtc ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg tttgctggcg 13380
accactacac gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga 13440
tgttcgacta tttcagctcg caccgggagc cgtacccgct caagctggaa accttccgcc 13500
tcatgtgcgg atcggattcc acccgcgtga agaagtggcg cgagcaggtc ggcgaagcct 13560
gcgaagagtt gcgaggcagc ggcctggtgg aacacgcctg ggtcaatgat gacctggtgc 13620
attgcaaacg ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt 13680
tactggcatt tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct cagtatcgct 13740

cgggacgcac ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt 13800

gattaaggct cagattcgac ggcttggagc ggccgacgtg caggatttcc gcgagatccg 13860

attgtcggcc ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa 13920

aaagcccatg gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg gcgcctacat 13980

cgacggcgag atcattgggc tgtcggtctt caaacaggag gacggcccca aggacgctca 14040

caaggcgcat ctgtccggcg ttttcgtgga gcccgaacag cgaggccgag gggtcgccgg 14100

tatgctgctg cgggcgttgc cggcgggttt attgctcgtg atgatcgtcc gacagattcc 14160

aacgggaatc tggtggatgc gcatcttcat cctcggcgca cttaatattt cgctattctg 14220

gagcttgttg tttatttcgg tctaccgcct gccgggcggg gtcgcggcga cggtaggcgc 14280

tgtgcagccg ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt 14340

gatggcggtc ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg tgttgacacc 14400

aaacgcagcg ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg tttccatggc 14460

gttcggaacc gtgctgaccc gcaagtggca acctcccgtg cctctgctca cctttaccgc 14520

ctggcaactg gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc 14580

aatcccgatg cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc 14640

gggtttaacc tacttccttt ggttccgggg gatctcgcga ctcgaaccta cagttgtttc 14700

cttactgggc tttctcagcc gggatggcgc taagaagcta ttgccgccga tcttcatatg 14760

cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct 14820

tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 14880

tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 14940

gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 15000

aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 15060

ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 15120

gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 15180

ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 15240

ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 15300

cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 15360

attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 15420

ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 15480

aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt 15540

gtttgcaagc agcagattac gcgcagaaaa aaaggatatc aagaagatcc tttgatcttt 15600

tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 15660

ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 15720

taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 15780

atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata 15840

actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca 15900
cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 15960
agtggtcctg caactttatc cgcctccatc cagtctatta aacaagtggc agcaacggat 16020
tcgcaaacct gtcacgcctt ttgtgccaaa agccgcgcca ggtttgcgat ccgctgtgcc 16080
aggcgttagg cgtcatatga agatttcggt gatccctgag caggtggcgg aaacattgga 16140
tgctgagaac catttcattg ttcgtgaagt gttcgatgtg cacctatccg accaaggctt 16200
tgaactatct accagaagtg tgagccccta ccggaaggat tacatctcgg atgatgactc 16260
tgatgaagac tctgcttgct atggcgcatt catcgaccaa gagcttgtcg ggaagattga 16320
actcaactca acatggaacg atctagcctc tatcgaacac attgttgtgt cgcacacgca 16380
ccgaggcaaa ggagtcgcgc acagtctcat cgaatttgcg aaaaagtggg cactaagcag 16440
acagctcctt ggcatacgat tagagacaca aacgaacaat gtacctgcct gcaatttgta 16500
cgcaaaatgt ggctttactc tcggcggcat tgacctgttc acgtataaaa ctagacctca 16560
agtctcgaac gaaacagcga tgtactggta ctggttctcg ggagcacagg atgacgccta 16620
acaattcatt caagccgaca ccgcttcgcg gcgcggctta attcaggagt taaacatcat 16680
gagggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga 16740
gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 16800
cctgaagcca cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac 16860
aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga 16920
gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta 16980
tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 17040
cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 17100
tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga 17160
tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg 17220
cgatgagcga aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa 17280
aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca 17340
gcccgtcata cttgaagcta ggcaggctta tcttggacaa gaagatcgct tggcctcgcg 17400
cgcagatcag ttggaagaat ttgttcacta cgtgaaaggc gagatcacca aggtagtcgg 17460
caaataatgt ctaacaattc gttcaagccg acgccgcttc gcggcgcggc ttaactcaag 17520
cgttagagag ctggggaaga ctatgcgcga tctgttgaag gtggttctaa gcctcgtact 17580
tgcgatggca tcggggcagg cacttgctga cctgccaatt gttttagtgg atgaagctcg 17640
tcttccctat gactactccc catccaacta cgacatttct ccaagcaact acgacaactc 17700
cataagcaat tacgacaata gtccatcaaa ttacgacaac tctgagagca actacgataa 17760
tagttcatcc aattacgaca atagtcgcaa cggaaatcgt aggcttatat atagcgcaaa 17820
tgggtctcgc actttcgccg gctactacgt cattgccaac aatgggacaa cgaacttctt 17880

ttccacatct ggcaaaagga tgttctacac cccaaaaggg gggcgcggcg tctatggcgg 17940

caaagatggg agcttctgcg gggcattggt cgtcataaat ggccaatttt cgcttgccct 18000

gacagataac ggcctgaaga tcatgtatct aagcaactag cctgctctct aataaaatgt 18060

taggagcttg gctgccattt ttggggtgag gccgttcgcg gccgaggggc gcagcccctg 18120

gggggatggg aggcccgcgt tagcgggccg ggagggttcg agaaggggg gcacccccct 18180

tcggcgtgcg cggtcacgcg ccagggcgca gccctggtta aaaacaaggt ttataaatat 18240

tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaaccctt 18300

gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt gcgcccctca 18360

tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc 18420

gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtccaca tcatctgtgg 18480

gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc 18540

cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca 18600

agtgtcaacg tccgcccctc atctgtcagt gagggccaag ttttccgcga ggtatccaca 18660

acgccggcgg ccggccgcgg tgtctcgcac acggcttcga cggcgtttct ggcgcgtttg 18720

cagggccata gacggccgcc agcccagcgg cgagggcaac cagcccggtg agcgtcggaa 18780

aggg 18784

<210> 27
<211> 18784
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 27
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60

atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120

gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420

atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480

ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540

aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600

aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660

gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720

agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780

ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840

4197737_1

aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900
cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960
cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020
cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080
cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140
gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200
catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260
cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320
gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380
ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440
ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500
cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560
gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620
cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680
ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740
tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800
acttgctttg aagacgtggt tggaacgtct tctttttcca cgatgctcct cgtgggtggg 1860
ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920
caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980
atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040
atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100
tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160
ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220
tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280
tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340
gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400
atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520
acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580
atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640
attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700
attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760
tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820
tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880

caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940

aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000

ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060

ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120

aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180

aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240

aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300

tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360

aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420

gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480

catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540

accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600

catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660

aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720

tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780

gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840

cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900

atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960

atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020

agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080

cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140

ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200

taaaaaaaac catggcatgt cattacggac aacagcaaca gactagagca cctcatcttc 4260

agcttcaacc tagagcacag agagttgtga aggctgctac tgctgttact gtaaatttct 4320

gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa tttcgtatat 4380

gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg tttgatcgtt 4440

agatatcatc ttaattctcg attagggttt catagatatc atccgatttg ttcaaataat 4500

ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc tggtgttagt 4560

ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta acaggctgga 4620

ggatctcttc ttgttctctc tggacttact ctcgctggaa ctgttatcgc tctcactatc 4680

gctacacctc ttctcgttat cttctctcct gttctcgttc ctgctgtgat cactatcttc 4740

cttctcggag ctggatttct tgcttctggt ggatttggag ttgctgctct ctctgttctc 4800

tcttggatct acagatacct cactggaaaa catcctccag gtgctgatca acttgagtct 4860

gctaagacta agctcgcttc taaggctaga gagatgaagg atagagcaga gcaattctct 4920

tgtcagcctg ttgctggatc tcagacttct taatgaacat atggtcctgc tttaatgaga 4980

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 5040
aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 5100
atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtac 5160
cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga ataggtttat 5220
agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta ttacaaatcc 5280
aattttaaaa aaagcggcag aaccggtcaa acctaaaaga ctgattacat aaatcttatt 5340
caaatttcaa aaggccccag gggctagtat ctacgacaca ccgagcggcg aactaataac 5400
gttcactgaa gggaactccg gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg 5460
agtattggcc gtccgctcta ccgaaagtta cgggcaccat tcaacccggt ccagcacggc 5520
ggccgggtaa ccgacttgct gccccgagaa ttatgcagca ttttttggt gtatgtgggc 5580
cccaaatgaa gtgcaggtca aaccttgaca gtgacgacaa atcgttgggc gggtccaggg 5640
cgaattttgc gacaacatgt cgaggctcag caggacctgc aggcatgcaa gctagcttac 5700
tagtgatgca tattctatag tgtcacctaa atcttcgacg aattaattcc aatcccacaa 5760
aaatctgagc ttaacagcac agttgctcct ctcagagcag aatcgggtat tcaacaccct 5820
catatcaact actacgttgt gtataacggt ccacatgccg gtatatacga tgactggggt 5880
tgtacaaagg cggcaacaaa cggcgttccc ggagttgcac acaagaaatt tgccactatt 5940
acagaggcaa gagcagcagc tgacgcgtac acaacaagtc agcaaacaga caggttgaac 6000
ttcatcccca aaggagaagc tcaactcaag cccaagagct ttgctaaggc cctaacaagc 6060
ccaccaaagc aaaaagccca ctggctcacg ctaggaacca aaaggcccag cagtgatcca 6120
gccccaaaag agatctcctt tgccccggag attacaatgg acgatttcct ctatctttac 6180
gatctaggaa ggaagttcga aggtgaaggt gacgacacta tgttcaccac tgataatgag 6240
aaggttagcc tcttcaattt cagaaagaat gctgacccac agatggttag agaggcctac 6300
gcagcaggtc tcatcaagac gatctacccg agtaacaatc tccaggagat caaatacctt 6360
cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta attgcatcaa gaacacagag 6420
aaagacatat ttctcaagat cagaagtact attccagtat ggacgattca aggcttgctt 6480
cataaaccaa ggcaagtaat agagattgga gtctctaaaa aggtagttcc tactgaatct 6540
aaggccatgc atggagtcta agattcaaat cgaggatcta acagaactcg ccgtgaagac 6600
tggcgaacag ttcatacaga gtcttttacg actcaatgac aagaagaaaa tcttcgtcaa 6660
catggtggag cacgacactc tggtctactc caaaaatgtc aaagatacag tctcagaaga 6720
ccaaagggct attgagactt ttcaacaaag gataatttcg ggaaacctcc tcggattcca 6780
ttgcccagct atctgtcact tcatcgaaag gacagtagaa aaggaaggtg gctcctacaa 6840
atgccatcat tgcgataaag gaaaggctat cattcaagat ctctctgccg acagtggtcc 6900
caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 6960
ttcaaagcaa gtggattgat gtgacatctc cactgacgta agggatgacg cacaatccca 7020

ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggacacg 7080

ggatccttgc tccgttaaaa aaaaccatgg ctatcctcga ttctgctggt gttactactg 7140

tgactgagaa tggtggtgga gagttcgttg atctcgatag actcagaaga agaaagtcta 7200

gatctgtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt ttcgttcgat 7260

cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga agacgatttt 7320

ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag atatcatccg 7380

atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg atttctatct 7440

agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct gagtttttct 7500

gattaacagg attcttctaa cggacttctc ctctctggat ctgataacaa ctctccttct 7560

gatgatgttg gtgctcctgc tgatgtgaga gatagaatcg attctgtggt gaacgatgat 7620

gctcaaggta ctgctaacct cgctggtgat aataacggtg gaggtgataa caatggtgga 7680

ggaagaggtg gaggtgaagg tagaggaaac gctgatgcta ctttcactta cagaccatct 7740

gtgcctgctc atagaagagc tagagagtct cctctctctt ctgatgctat cttcaagcag 7800

tctcacgctg gacttttcaa cctctgtgtg gtggttctta tcgctgtgaa ctctagactc 7860

atcatcgaga acctcatgaa gtacggatgg ctcatcagaa ctgatttctg gttctcttct 7920

agatctctca gagattggcc tcttttcatg tgctgcatct cactctcaat cttccctctc 7980

gctgctttta ctgttgagaa gctcgtgctc cagaagtata tcgctgaacc tgtggtgatc 8040

ttcctccaca tcatcatcac tatgactgag gttctctacc ctgtttacgt gactctcaga 8100

tgcgattctg ctttcctctc tggtgttact cttatgctcc tcacttgcat tgtgtggctt 8160

aagctcgtgt cttacgctca cacttcttac gatatcagat ctctcgctaa cgctgctgat 8220

aaggctaacc ctgaagtgtc ttactacgtg tctctcaagt ctctcgctta cttcatggtt 8280

gctcctacac tttgttacca gccatcttac cctagatctg cttgcattag aaagggatgg 8340

gtggcaagac aattcgctaa gttggtgatc ttcactggat tcatgggatt catcatcgag 8400

cagtacatca accctattgt gagaaactct aagcaccctc tcaagggtga tcttctctac 8460

gctatcgaga gagttcttaa gctctctgtg cctaaccttt atgtgtggct ctgcatgttc 8520

tactgtttct tccacctctg gcttaacatc cttgctgagt tgctttgctt cggagataga 8580

gagttctaca aggattggtg gaacgctaag tctgttggag attattggag aatgtggaac 8640

atgcctgtgc ataagtggat ggtgcgtcac atctacttcc cttgcctcag atctaagatc 8700

cctaagactc tcgctatcat tatcgctttc ctcgtgtctg ctgttttcca tgagttgtgt 8760

atcgctgttc cttgcagact tttcaagctt tgggctttcc tcggaatcat gttccaggtt 8820

ccactcgtgt tcatcactaa ctacctccaa gagagattcg gatctactgt tggaaacatg 8880

attttctggt tcattttctg catcttcgga cagcctatgt gcgttctcct ctactaccac 8940

gatctcatga acagaaaggg atctatgtct taatgaagga tccacccagc tttcttgtac 9000

aaagtggtga tgggttcgaa atcgataagc ttggatcctc tagagtcctg ctttaatgag 9060

atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa 9120

aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata 9180
tatcacccgt tactatcgta tttttatgaa taatattctc cgttcaattt actgattgta 9240
ccctactact tatatgtaca atattaaaat gaaaacaata tattgtgctg aataggttta 9300
tagcgacatc tatgatagag cgccacaata acaaacaatt gcgttttatt attacaaatc 9360
caattttaaa aaaagcggca gaaccggtca aacctaaaag actgattaca taaatcttat 9420
tcaaatttca aaaggcccca ggggctagta tctacgacac accgagcggc gaactaataa 9480
cgttcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt ccttgaagtt 9540
gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaacccgg tccagcacgg 9600
cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg tgtatgtggg 9660
ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg cgggtccagg 9720
gcgaattttg cgacaacatg tcgaggctca gcaggacctg caggcatgca agctagctta 9780
ctagtgatgc atattctata gtgtcaccta aatctgcggc cgcctgcagg tcgatatggg 9840
agagctccca acgcgttgga tgcatagctt gagtattcta tagtgtcacc taaatagctt 9900
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 9960
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 10020
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 10080
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggg ctgagtggct 10140
ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg gcttgtcccg cgtcatcggc 10200
gggggtcata acgtgactcc cttaattctc cgctcatgat cagattgtcg tttcccgcct 10260
tcagtttaaa ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag 10320
cgtttattag aataatcgga tatttaaaag ggcgtgaaaa ggtttatccg ttcgtccatt 10380
tgtatgtgca tgccaaccac agggttcccc tcgggagtgc ttggcattcc gtgcgataat 10440
gacttctgtt caaccaccca aacgtcggaa agcctgacga cggagcagca ttccaaaaag 10500
atcccttggc tcgtctgggt cggctagaag gtcgagtggg ctgctgtggc ttgatccctc 10560
aacgcggtcg cggacgtagc gcagcgccga aaaatcctcg atcgcaaatc cgacgctgtc 10620
gaaaatcgtg atctgcttgt cgctctttcg gccgacgtcc tggccagtca tcacgcgcca 10680
aagttccgtc acaggatgat ctggcgcgag ttgctggatc tcgccttcaa tccgggtctg 10740
tggcgggaac tccacgaaaa tatccgaacg cagcaagatg tcgacggatc ttttccgctg 10800
cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc tttttcgcac 10860
gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc caacggcgtc 10920
agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc ttcactgtcc 10980
cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg ccggctaccg 11040
ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc aggggtgatg 11100
ctgccaactt actgatttag tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt 11160

tctatcagct gtccctcctg ttcagctact gacggggtgg tgcgtaacgg caaaagcacc 11220

gccggacatc agcgctatct ctgctctcac tgccgtaaaa catggcaact gcagttcact 11280

tacaccgctt ctcaacccgg tacgcaccag aaaatcattg atatggccat gaatggcgtt 11340

ggatgccggg caacagcccg cattatgggc gttggcctca acacgatttt acgtcactta 11400

aaaaactcag gccgcagtcg gtaacctcgc gcatacagcc gggcagtgac gtcatcgtct 11460

gcgcggaaat ggacgaacag tggggctatg tcggggctaa atcgcgccag cgctggctgt 11520

tttacgcgta tgacagtctc cggaagacgg ttgttgcgca cgtattcggt gaacgcacta 11580

tggcgacgct ggggcgtctt atgagcctgc tgtcaccctt tgacgtggtg atatggatga 11640

cggatggctg gccgctgtat gaatcccgcc tgaagggaaa gctgcacgta atcagcaagc 11700

gatatacgca gcgaattgag cggcataacc tgaatctgag gcagcacctg gcacggctgg 11760

gacggaagtc gctgtcgttc tcaaaatcgg tggagctgca tgacaaagtc atcgggcatt 11820

atctgaacat aaaacactat caataagttg gagtcattac ccaaccagga agggcagccc 11880

acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa aggcggcggc 11940

ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca aaatcacggg 12000

cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc tgggccgcct 12060

gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt tcggtgatgc 12120

cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg gcaaggtcat 12180

gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa aacggccggg 12240

gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc gacttcgcgg 12300

agctggtatt cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg 12360

tctacgggac cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg 12420

ggtcaaatca ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag 12480

ggtgaatgaa tcggacgttt gaccggaagg catacaggca agaactgatc gacgcggggt 12540

tttccgccga ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa 12600

ccttccagtc cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg 12660

tgcaactggc tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt tcgcgtcgtc 12720

tcgaacagga ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga 12780

cgaccaagaa gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc 12840

aggccgcgtt gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg 12900

atattgcgcc gtggccggac acgatgcgag cgatgccaaa cgacacggcc cgctctgccc 12960

tgttcaccac gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc 13020

acgtcaacaa ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgacg 13080

aactggtgtg gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca 13140

ccttcacgtt ctacgagctt tgccaggacc tgggctggtc gatcaatggc cggtattaca 13200

cgaaggccga ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc 13260

gcgttgggca cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca 13320
agaaaacgtc ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg tttgctggcg 13380
accactacac gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga 13440
tgttcgacta tttcagctcg caccgggagc cgtacccgct caagctggaa accttccgcc 13500
tcatgtgcgg atcggattcc acccgcgtga agaagtggcg cgagcaggtc ggcgaagcct 13560
gcgaagagtt gcgaggcagc ggcctggtgg aacacgcctg ggtcaatgat gacctggtgc 13620
attgcaaacg ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt 13680
tactggcatt tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct cagtatcgct 13740
cgggacgcac ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt 13800
gattaaggct cagattcgac ggcttggagc ggccgacgtg caggatttcc gcgagatccg 13860
attgtcggcc ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa 13920
aaagcccatg gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg gcgcctacat 13980
cgacggcgag atcattgggc tgtcggtctt caaacaggag gacggcccca aggacgctca 14040
caaggcgcat ctgtccggcg ttttcgtgga gcccgaacag cgaggccgag gggtcgccgg 14100
tatgctgctg cgggcgttgc cggcgggttt attgctcgtg atgatcgtcc gacagattcc 14160
aacgggaatc tggtggatgc gcatcttcat cctcggcgca cttaatattt cgctattctg 14220
gagcttgttg tttatttcgg tctaccgcct gccgggcggg gtcgcggcga cggtaggcgc 14280
tgtgcagccg ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt 14340
gatggcggtc ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg tgttgacacc 14400
aaacgcagcg ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg tttccatggc 14460
gttcggaacc gtgctgaccc gcaagtggca acctcccgtg cctctgctca cctttaccgc 14520
ctggcaactg gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc 14580
aatcccgatg cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc 14640
gggtttaacc tacttccttt ggttccgggg gatctcgcga ctcgaaccta cagttgtttc 14700
cttactgggc tttctcagcc gggatggcgc taagaagcta ttgccgccga tcttcatatg 14760
cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct 14820
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 14880
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 14940
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 15000
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 15060
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 15120
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 15180
ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 15240
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 15300

cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 15360

attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 15420

ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 15480

aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggtttttt 15540

gtttgcaagc agcagattac gcgcagaaaa aaaggatatc aagaagatcc tttgatcttt 15600

tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 15660

ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 15720

taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 15780

atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata 15840

actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca 15900

cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 15960

agtggtcctg caactttatc cgcctccatc cagtctatta aacaagtggc agcaacggat 16020

tcgcaaacct gtcacgcctt ttgtgccaaa agccgcgcca ggtttgcgat ccgctgtgcc 16080

aggcgttagg cgtcatatga agatttcggt gatccctgag caggtggcgg aaacattgga 16140

tgctgagaac catttcattg ttcgtgaagt gttcgatgtg cacctatccg accaaggctt 16200

tgaactatct accagaagtg tgagccccta ccggaaggat tacatctcgg atgatgactc 16260

tgatgaagac tctgcttgct atggcgcatt catcgaccaa gagcttgtcg ggaagattga 16320

actcaactca acatggaacg atctagcctc tatcgaacac attgttgtgt cgcacacgca 16380

ccgaggcaaa ggagtcgcgc acagtctcat cgaatttgcg aaaaagtggg cactaagcag 16440

acagctcctt ggcatacgat tagagacaca aacgaacaat gtacctgcct gcaatttgta 16500

cgcaaaatgt ggctttactc tcggcggcat tgacctgttc acgtataaaa ctagacctca 16560

agtctcgaac gaaacagcga tgtactggta ctggttctcg ggagcacagg atgacgccta 16620

acaattcatt caagccgaca ccgcttcgcg gcgcggctta attcaggagt taaacatcat 16680

gagggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga 16740

gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 16800

cctgaagcca cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac 16860

aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga 16920

gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta 16980

tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 17040

cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 17100

tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga 17160

tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg 17220

cgatgagcga aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa 17280

aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca 17340

gcccgtcata cttgaagcta ggcaggctta tcttggacaa gaagatcgct tggcctcgcg 17400

4197737_1

cgcagatcag ttggaagaat ttgttcacta cgtgaaaggc gagatcacca aggtagtcgg 17460
caaataatgt ctaacaattc gttcaagccg acgccgcttc gcggcgcggc ttaactcaag 17520
cgttagagag ctggggaaga ctatgcgcga tctgttgaag gtggttctaa gcctcgtact 17580
tgcgatggca tcggggcagg cacttgctga cctgccaatt gttttagtgg atgaagctcg 17640
tcttccctat gactactccc catccaacta cgacatttct ccaagcaact acgacaactc 17700
cataagcaat tacgacaata gtccatcaaa ttacgacaac tctgagagca actacgataa 17760
tagttcatcc aattacgaca atagtcgcaa cggaaatcgt aggcttatat atagcgcaaa 17820
tgggtctcgc actttcgccg gctactacgt cattgccaac aatgggacaa cgaacttctt 17880
ttccacatct ggcaaaagga tgttctacac cccaaaaggg gggcgcggcg tctatggcgg 17940
caaagatggg agcttctgcg gggcattggt cgtcataaat ggccaatttt cgcttgccct 18000
gacagataac ggcctgaaga tcatgtatct aagcaactag cctgctctct aataaaatgt 18060
taggagcttg gctgccattt ttggggtgag gccgttcgcg gccgaggggc gcagccctg 18120
gggggatggg aggcccgcgt tagcgggccg ggagggttcg agaaggggg gcacccccct 18180
tcggcgtgcg cggtcacgcg ccagggcgca gccctggtta aaaacaaggt ttataaatat 18240
tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaaccctt 18300
gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt gcgcccctca 18360
tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc 18420
gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtccaca tcatctgtgg 18480
gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc 18540
cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca 18600
agtgtcaacg tccgccctc atctgtcagt gagggccaag ttttccgcga ggtatccaca 18660
acgccggcgg ccggccgcgg tgtctcgcac acggcttcga cggcgtttct ggcgcgtttg 18720
cagggccata gacggccgcc agcccagcgg cgagggcaac cagcccggtg agcgtcggaa 18780
aggg 18784

<210> 28
<211> 18784
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 28
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60
atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120
gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240
gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

4197737_1

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420

atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480

ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540

aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600

aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660

gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720

agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780

ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840

aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900

cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960

cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020

cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080

cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140

gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200

catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260

cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320

gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380

ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440

ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500

cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560

gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620

cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680

ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740

tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800

acttgctttg aagacgtggt tggaacgtct tcttttcca cgatgctcct cgtgggtggg 1860

ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920

caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980

atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040

atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100

tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160

ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220

tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280

tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340

gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400

4197737_1

atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520
acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580
atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640
attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700
attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760
tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820
tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880
caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940
aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000
ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060
ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120
aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180
aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240
aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300
tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360
aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420
gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480
catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540
accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600
catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660
aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720
tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780
gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840
cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900
atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960
atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020
agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080
cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140
ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200
taaaaaaaac catggcatgt cattacggac aacagcaaca gactagagca cctcatcttc 4260
agcttcaacc tagagcacag agagttgtga aggctgctac tgctgttact gtaaatttct 4320
gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa tttcgtatat 4380
gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg tttgatcgtt 4440

agatatcatc ttaattctcg attagggttt catagatatc atccgatttg ttcaaataat 4500

ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc tggtgttagt 4560

ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta acaggctgga 4620

ggatctcttc ttgttctctc tggacttact ctcgctggaa ctgttatcgc tctcactatc 4680

gctacacctc ttctcgttat cttctctcct gttctcgttc ctgctgtgat cactatcttc 4740

cttctcggag ctggatttct tgcttctggt ggatttggag ttgctgctct ctctgttctc 4800

tcttggatct acagatacct cactggaaaa catcctccag gtgctgattg tcttgagtct 4860

gctaagacta agctcgcttc ttgtgctaga gagatgaagg atagagcaga gcaattctct 4920

tgtcagcctg ttgctggatc tcagacttct taatgaacat atggtcctgc tttaatgaga 4980

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 5040

aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 5100

atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtac 5160

cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga ataggtttat 5220

agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta ttacaaatcc 5280

aattttaaaa aaagcggcag aaccggtcaa acctaaaaga ctgattacat aaatcttatt 5340

caaatttcaa aaggccccag gggctagtat ctacgacaca ccgagcggcg aactaataac 5400

gttcactgaa gggaactccg gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg 5460

agtattggcc gtccgctcta ccgaaagtta cgggcaccat tcaacccggt ccagcacggc 5520

ggccgggtaa ccgacttgct gccccgagaa ttatgcagca ttttttttggt gtatgtgggc 5580

cccaaatgaa gtgcaggtca aaccttgaca gtgacgacaa atcgttgggc gggtccaggg 5640

cgaattttgc gacaacatgt cgaggctcag caggacctgc aggcatgcaa gctagcttac 5700

tagtgatgca tattctatag tgtcacctaa atcttcgacg aattaattcc aatcccacaa 5760

aaatctgagc ttaacagcac agttgctcct ctcagagcag aatcgggtat tcaacaccct 5820

catatcaact actacgttgt gtataacggt ccacatgccg gtatatacga tgactggggt 5880

tgtacaaagg cggcaacaaa cggcgttccc ggagttgcac acaagaaatt tgccactatt 5940

acagaggcaa gagcagcagc tgacgcgtac acaacaagtc agcaaacaga caggttgaac 6000

ttcatcccca aaggagaagc tcaactcaag cccaagagct ttgctaaggc cctaacaagc 6060

ccaccaaagc aaaaagccca ctggctcacg ctaggaacca aaaggcccag cagtgatcca 6120

gccccaaaag agatctcctt tgccccggag attacaatgg acgatttcct ctatctttac 6180

gatctaggaa ggaagttcga aggtgaaggt gacgacacta tgttcaccac tgataatgag 6240

aaggttagcc tcttcaattt cagaaagaat gctgacccac agatggttag agaggcctac 6300

gcagcaggtc tcatcaagac gatctacccg agtaacaatc tccaggagat caaatacctt 6360

cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta attgcatcaa gaacacagag 6420

aaagacatat ttctcaagat cagaagtact attccagtat ggacgattca aggcttgctt 6480

cataaaccaa ggcaagtaat agagattgga gtctctaaaa aggtagttcc tactgaatct 6540

aaggccatgc atggagtcta agattcaaat cgaggatcta acagaactcg ccgtgaagac 6600
tggcgaacag ttcatacaga gtcttttacg actcaatgac aagaagaaaa tcttcgtcaa 6660
catggtggag cacgacactc tggtctactc caaaaatgtc aaagatacag tctcagaaga 6720
ccaaagggct attgagactt ttcaacaaag gataatttcg ggaaacctcc tcggattcca 6780
ttgcccagct atctgtcact tcatcgaaag gacagtagaa aaggaaggtg gctcctacaa 6840
atgccatcat tgcgataaag gaaaggctat cattcaagat ctctctgccg acagtggtcc 6900
caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 6960
ttcaaagcaa gtggattgat gtgacatctc cactgacgta agggatgacg cacaatccca 7020
ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggacacg 7080
ggatccttgc tccgttaaaa aaaaccatgg ctatcctcga ttctgctggt gttactactg 7140
tgactgagaa tggtggtgga gagttcgttg atctcgatag actcagaaga agaaagtcta 7200
gatctgtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt ttcgttcgat 7260
cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga agacgatttt 7320
ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag atatcatccg 7380
atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg atttctatct 7440
agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct gagtttttct 7500
gattaacagg attcttctaa cggacttctc ctctctggat ctgataacaa ctctccttct 7560
gatgatgttg gtgctcctgc tgatgtgaga gatagaatcg attctgtggt gaacgatgat 7620
gctcaaggta ctgctaacct cgctggtgat aataacggtg gaggtgataa caatggtgga 7680
ggaagaggtg gaggtgaagg tagaggaaac gctgatgcta ctttcactta cagaccatct 7740
gtgcctgctc atagaagagc tagagagtct cctctctctt ctgatgctat cttcaagcag 7800
tctcacgctg gacttttcaa cctctgtgtg gtggttctta tcgctgtgaa ctctagactc 7860
atcatcgaga acctcatgaa gtacggatgg ctcatcagaa ctgatttctg gttctcttct 7920
agatctctca gagattggcc tcttttcatg tgctgcatct cactctcaat cttccctctc 7980
gctgctttta ctgttgagaa gctcgtgctc cagaagtata tcgctgaacc tgtggtgatc 8040
ttcctccaca tcatcatcac tatgactgag gttctctacc ctgtttacgt gactctcaga 8100
tgcgattctg ctttcctctc tggtgttact cttatgctcc tcacttgcat tgtgtggctt 8160
aagctcgtgt cttacgctca cacttcttac gatatcagat ctctcgctaa cgctgctgat 8220
aaggctaacc ctgaagtgtc ttactacgtg tctctcaagt ctctcgctta cttcatggtt 8280
gctcctacac tttgttacca gccatcttac cctagatctg cttgcattag aaagggatgg 8340
gtggcaagac aattcgctaa gttggtgatc ttcactggat tcatgggatt catcatcgag 8400
cagtacatca accctattgt gagaaactct aagcaccctc tcaagggtga tcttctctac 8460
gctatcgaga gagttcttaa gctctctgtg cctaaccttt atgtgtggct ctgcatgttc 8520
tactgtttct tccacctctg gcttaacatc cttgctgagt tgctttgctt cggagataga 8580

gagttctaca aggattggtg gaacgctaag tctgttggag attattggag aatgtggaac 8640
atgcctgtgc ataagtggat ggtgcgtcac atctacttcc cttgcctcag atctaagatc 8700
cctaagactc tcgctatcat tatcgctttc ctcgtgtctg ctgttttcca tgagttgtgt 8760
atcgctgttc cttgcagact tttcaagctt tgggctttcc tcggaatcat gttccaggtt 8820
ccactcgtgt tcatcactaa ctacctccaa gagagattcg gatctactgt tggaaacatg 8880
attttctggt tcattttctg catcttcgga cagcctatgt gcgttctcct ctactaccac 8940
gatctcatga acagaaaggg atctatgtct taatgaagga tccacccagc tttcttgtac 9000
aaagtggtga tgggttcgaa atcgataagc ttggatcctc tagagtcctg ctttaatgag 9060
atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa 9120
aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata 9180
tatcacccgt tactatcgta tttttatgaa taatattctc cgttcaattt actgattgta 9240
ccctactact tatatgtaca atattaaaat gaaaacaata tattgtgctg aataggttta 9300
tagcgacatc tatgatagag cgccacaata acaaacaatt gcgttttatt attacaaatc 9360
caattttaaa aaaagcggca gaaccggtca aacctaaaag actgattaca taaatcttat 9420
tcaaatttca aaaggcccca ggggctagta tctacgacac accgagcggc gaactaataa 9480
cgttcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt ccttgaagtt 9540
gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaacccgg tccagcacgg 9600
cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg tgtatgtggg 9660
ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg cgggtccagg 9720
gcgaattttg cgacaacatg tcgaggctca gcaggacctg caggcatgca agctagctta 9780
ctagtgatgc atattctata gtgtcaccta aatctgcggc cgcctgcagg tcgatatggg 9840
agagctccca acgcgttgga tgcatagctt gagtattcta tagtgtcacc taaatagctt 9900
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 9960
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 10020
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 10080
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggg ctgagtggct 10140
ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg gcttgtcccg cgtcatcggc 10200
gggggtcata acgtgactcc cttaattctc cgctcatgat cagattgtcg tttcccgcct 10260
tcagtttaaa ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag 10320
cgtttattag aataatcgga tatttaaaag ggcgtgaaaa ggtttatccg ttcgtccatt 10380
tgtatgtgca tgccaaccac agggttcccc tcgggagtgc ttggcattcc gtgcgataat 10440
gacttctgtt caaccaccca aacgtcggaa agcctgacga cggagcagca ttccaaaaag 10500
atcccttggc tcgtctgggt cggctagaag gtcgagtggg ctgctgtggc ttgatccctc 10560
aacgcggtcg cggacgtagc gcagcgccga aaaatcctcg atcgcaaatc cgacgctgtc 10620
gaaaatcgtg atctgcttgt cgctctttcg gccgacgtcc tggccagtca tcacgcgcca 10680

aagttccgtc acaggatgat ctggcgcgag ttgctggatc tcgccttcaa tccgggtctg 10740
tggcgggaac tccacgaaaa tatccgaacg cagcaagatg tcgacggatc ttttccgctg 10800
cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc tttttcgcac 10860
gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc caacggcgtc 10920
agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc ttcactgtcc 10980
cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg ccggctaccg 11040
ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc aggggtgatg 11100
ctgccaactt actgatttag tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt 11160
tctatcagct gtccctcctg ttcagctact gacggggtgg tgcgtaacgg caaaagcacc 11220
gccggacatc agcgctatct ctgctctcac tgccgtaaaa catggcaact gcagttcact 11280
tacaccgctt ctcaacccgg tacgcaccag aaaatcattg atatggccat gaatggcgtt 11340
ggatgccggg caacagcccg cattatgggc gttggcctca acacgatttt acgtcactta 11400
aaaaactcag gccgcagtcg gtaacctcgc gcatacagcc gggcagtgac gtcatcgtct 11460
gcgcggaaat ggacgaacag tggggctatg tcggggctaa atcgcgccag cgctggctgt 11520
tttacgcgta tgacagtctc cggaagacgg ttgttgcgca cgtattcggt gaacgcacta 11580
tggcgacgct ggggcgtctt atgagcctgc tgtcaccctt tgacgtggtg atatggatga 11640
cggatggctg gccgctgtat gaatcccgcc tgaagggaaa gctgcacgta atcagcaagc 11700
gatatacgca gcgaattgag cggcataacc tgaatctgag gcagcacctg gcacggctgg 11760
gacggaagtc gctgtcgttc tcaaaatcgg tggagctgca tgacaaagtc atcgggcatt 11820
atctgaacat aaaacactat caataagttg gagtcattac ccaaccagga agggcagccc 11880
acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa aggcggcggc 11940
ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca aaatcacggg 12000
cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc tgggccgcct 12060
gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt tcggtgatgc 12120
cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg gcaaggtcat 12180
gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa aacggccggg 12240
gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc gacttcgcgg 12300
agctggtatt cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg 12360
tctacgggac cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg 12420
ggtcaaatca ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag 12480
ggtgaatgaa tcggacgttt gaccggaagg catacaggca agaactgatc gacgcggggt 12540
tttccgccga ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa 12600
ccttccagtc cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg 12660
tgcaactggc tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt tcgcgtcgtc 12720

tcgaacagga ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga 12780

cgaccaagaa gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc 12840

aggccgcgtt gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg 12900

atattgcgcc gtggccggac acgatgcgag cgatgccaaa cgacacggcc cgctctgccc 12960

tgttcaccac gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc 13020

acgtcaacaa ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgacg 13080

aactggtgtg gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca 13140

ccttcacgtt ctacgagctt tgccaggacc tgggctggtc gatcaatggc cggtattaca 13200

cgaaggccga ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc 13260

gcgttgggca cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca 13320

agaaaacgtc ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg tttgctggcg 13380

accactacac gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga 13440

tgttcgacta tttcagctcg caccgggagc cgtacccgct caagctggaa accttccgcc 13500

tcatgtgcgg atcggattcc acccgcgtga agaagtggcg cgagcaggtc ggcgaagcct 13560

gcgaagagtt gcgaggcagc ggcctggtgg aacacgcctg ggtcaatgat gacctggtgc 13620

attgcaaacg ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt 13680

tactggcatt tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct cagtatcgct 13740

cgggacgcac ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt 13800

gattaaggct cagattcgac ggcttggagc ggccgacgtg caggatttcc gcgagatccg 13860

attgtcggcc ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa 13920

aaagcccatg gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg gcgcctacat 13980

cgacggcgag atcattgggc tgtcggtctt caaacaggag gacggcccca aggacgctca 14040

caaggcgcat ctgtccggcg ttttcgtgga gcccgaacag cgaggccgag gggtcgccgg 14100

tatgctgctg cgggcgttgc cggcgggttt attgctcgtg atgatcgtcc gacagattcc 14160

aacgggaatc tggtggatgc gcatcttcat cctcggcgca cttaatattt cgctattctg 14220

gagcttgttg tttatttcgg tctaccgcct gccgggcggg gtcgcggcga cggtaggcgc 14280

tgtgcagccg ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt 14340

gatggcggtc ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg tgttgacacc 14400

aaacgcagcg ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg tttccatggc 14460

gttcggaacc gtgctgaccc gcaagtggca acctcccgtg cctctgctca cctttaccgc 14520

ctggcaactg gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc 14580

aatcccgatg cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc 14640

gggtttaacc tacttccttt ggttccgggg gatctcgcga ctcgaaccta cagttgtttc 14700

cttactgggc tttctcagcc gggatggcgc taagaagcta ttgccgccga tcttcatatg 14760

cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct 14820

tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 14880
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 14940
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 15000
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 15060
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 15120
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 15180
ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 15240
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 15300
cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 15360
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 15420
ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 15480
aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt 15540
gtttgcaagc agcagattac gcgcagaaaa aaaggatatc aagaagatcc tttgatcttt 15600
tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 15660
ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 15720
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 15780
atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata 15840
actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca 15900
cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 15960
agtggtcctg caactttatc cgcctccatc cagtctatta aacaagtggc agcaacggat 16020
tcgcaaacct gtcacgcctt ttgtgccaaa agccgcgcca ggtttgcgat ccgctgtgcc 16080
aggcgttagg cgtcatatga agatttcggt gatccctgag caggtggcgg aaacattgga 16140
tgctgagaac catttcattg ttcgtgaagt gttcgatgtg cacctatccg accaaggctt 16200
tgaactatct accagaagtg tgagccccta ccggaaggat tacatctcgg atgatgactc 16260
tgatgaagac tctgcttgct atggcgcatt catcgaccaa gagcttgtcg ggaagattga 16320
actcaactca acatggaacg atctagcctc tatcgaacac attgttgtgt cgcacacgca 16380
ccgaggcaaa ggagtcgcgc acagtctcat cgaatttgcg aaaaagtggg cactaagcag 16440
acagctcctt ggcatacgat tagagacaca aacgaacaat gtacctgcct gcaatttgta 16500
cgcaaaatgt ggctttactc tcggcggcat tgacctgttc acgtataaaa ctagacctca 16560
agtctcgaac gaaacagcga tgtactggta ctggttctcg ggagcacagg atgacgccta 16620
acaattcatt caagccgaca ccgcttcgcg gcgcggctta attcaggagt taaacatcat 16680
gagggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga 16740
gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 16800
cctgaagcca cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac 16860

4197737_1

aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga 16920
gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta 16980
tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 17040
cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 17100
tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga 17160
tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg 17220
cgatgagcga aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa 17280
aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca 17340
gcccgtcata cttgaagcta ggcaggctta tcttggacaa gaagatcgct tggcctcgcg 17400
cgcagatcag ttggaagaat ttgttcacta cgtgaaaggc gagatcacca aggtagtcgg 17460
caaataatgt ctaacaattc gttcaagccg acgccgcttc gcggcgcggc ttaactcaag 17520
cgttagagag ctggggaaga ctatgcgcga tctgttgaag gtggttctaa gcctcgtact 17580
tgcgatggca tcggggcagg cacttgctga cctgccaatt gttttagtgg atgaagctcg 17640
tcttccctat gactactccc catccaacta cgacatttct ccaagcaact acgacaactc 17700
cataagcaat tacgacaata gtccatcaaa ttacgacaac tctgagagca actacgataa 17760
tagttcatcc aattacgaca atagtcgcaa cggaaatcgt aggcttatat atagcgcaaa 17820
tgggtctcgc actttcgccg gctactacgt cattgccaac aatgggacaa cgaacttctt 17880
ttccacatct ggcaaaagga tgttctacac cccaaaaggg gggcgcggcg tctatggcgg 17940
caaagatggg agcttctgcg gggcattggt cgtcataaat ggccaatttt cgcttgccct 18000
gacagataac ggcctgaaga tcatgtatct aagcaactag cctgctctct aataaaatgt 18060
taggagcttg gctgccattt ttggggtgag gccgttcgcg gccgaggggc gcagccctg 18120
gggggatggg aggcccgcgt tagcgggccg ggagggttcg agaaggggg gcacccccct 18180
tcggcgtgcg cggtcacgcg ccagggcgca gccctggtta aaaacaaggt ttataaatat 18240
tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaaccctt 18300
gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt gcgcccctca 18360
tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc 18420
gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtccaca tcatctgtgg 18480
gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc 18540
cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca 18600
agtgtcaacg tccgcccctc atctgtcagt gagggccaag ttttccgcga ggtatccaca 18660
acgccggcgg ccggccgcgg tgtctcgcac acggcttcga cggcgtttct ggcgcgtttg 18720
cagggccata gacggccgcc agcccagcgg cgagggcaac cagcccggtg agcgtcggaa 18780
aggg 18784

<210> 29
<211> 18784

4197737_1

<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 29
gtcgacatcg tcaacgttca cttctaaaga aatagcgcca ctcagcttcc tcagcggctt 60

tatccagcga tttcctatta tgtcggcata gttctcaaga tcgacagcct gtcacggtta 120

agcgagaaat gaataagaag gctgataatt cggatctctg cgaggagatg atatttgatc 180

acaggcagca acgctctgtc atcgttacaa tcaacatgct accctccgcg agatcatccg 240

tgtttcaaac ccggcagctt agttgccgtt cttccgaata gcatcgggta acatgagcaa 300

agtctgccgc cttacaacgg ctctcccgct gacgccgtcc cggactgatg ggctgcctgt 360

atcgagtggt gattttgtgc cgagctgccg gtcggggagc tgttggctgg ctggtggcag 420

gatatattgt ggtgtaaaca aattgacgct tagacaactt aataacacat tgcggacgtt 480

tttaatgtac tgaattcgcc gctcggtgtg tcgtagatac tagcccctgg ggcacttttg 540

aaatttgaat aagatttatg taatcagtct tttaggtttg accggttctg ccgctttttt 600

taaaattgga tttgtaataa taaaacgcaa ttgtttgtta ttgtggcgct ctatcataga 660

tgtcgctata aacctattca gcacaatata ttgttttcat tttaatattg tacatataag 720

tagtagggta caatcagtaa attgaacgga gaatattatt cataaaaata cgatagtaac 780

gggtgatata ttcattagaa tgaaccgaaa ccggcggtaa ggatctgagc tacacatgct 840

caggtttttt acaacgtgca caacagaatt gaaagcaaat atcatgcgat cataggcgtc 900

tcgcatatct cattaaacaa gtgaagattt gattcaaact ccattgagag ccctgactat 960

gcattcggtt tgacccttcc aggttgagag acgatagccc cctaccttaa ttaaggggcc 1020

ccccctcgag gtcgacggta tcgataagct tgatatcgaa ttcctgcagc caggggatc 1080

ccccgggtca tcagatctcg gtgacgggca ggaccggacg gggcggtacc ggcaggctga 1140

agtccagctg ccagaaaccc acgtcatgcc agttcccgtg cttgaagccg gccgcccgca 1200

gcatgccgcg gggggcatat ccgagcgcct cgtgcatgcg cacgctcggg tcgttgggca 1260

gcccgatgac agcgaccacg ctcttgaagc cctgtgcctc cagggacttc agcaggtggg 1320

tgtagagcgt ggagcccagt cccgtccgct ggtggcgggg ggagacgtac acggtcgact 1380

cggccgtcca gtcgtaggcg ttgcgtgcct tccaggggcc cgcgtaggcg atgccggcga 1440

cctcgccgtc cacctcggcg acgagccagg gatagcgctc ccgcagacgg acgaggtcgt 1500

ccgtccactc ctgcggttcc tgcggctcgg tacggaagtt gaccgtgctt gtctcgatgt 1560

agtggttgac gatggtgcag accgccggca tgtccgcctc ggtggcacgg cggatgtcgg 1620

ccgggcgtcg ttctgggctc atggtagatc ccctcgagag agatagattt gtagagagag 1680

actggtgatt tcagcgtgtc ctctccaaat gaaatgaact tccttatata gaggaaggtc 1740

ttgcgaagga tagtgggatt gtgcgtcatc ccttacgtca gtggagatat cacatcaatc 1800

cacttgcttt gaagacgtgg ttggaacgtc ttctttttcc acgatgctcc tcgtgggtgg 1860

gggtccatct ttgggaccac tgtcggcaga ggcatcttga acgatagcct ttcctttatc 1920

gcaatgatgg catttgtagg tgccaccttc cttttctact gtccttttga tgaagtgaca 1980

gatagctggg caatggaatc cgaggaggtt tcccgatatt acccttttgtt gaaaagtctc 2040

aatagccctt tggtcttctg agactgtatc tttgatattc ttggagtaga cgagagtgtc 2100

gtgctccacc atgttatcac atcaatccac ttgctttgaa gacgtggttg gaacgtcttc 2160

tttttccacg atgctcctcg tgggtggggg tccatctttg ggaccactgt cggcagaggc 2220

atcttgaacg atagcctttc ctttatcgca atgatggcat ttgtaggtgc caccttcctt 2280

ttctactgtc cttttgatga agtgacagat agctgggcaa tggaatccga ggaggtttcc 2340

cgatattacc ctttgttgaa aagtctcaat agccctttgg tcttctgaga ctgtatcttt 2400

gatattcttg gagtagacga gagtgtcgtg ctccaccatg ttgctgatag tgaccttagg 2460

cgacttttga acgcgcaata atggtttctg acgtatgtgc ttagctcatt aaactccaga 2520

aacccattaa cgcttacaat ttccattcgc cattcaggct gcgcaactgt tgggaagggc 2580

gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc 2640

gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 2700

aattgtaata cgactcacta tagggcgaat tgggcccgac gtcgcatgct cccggccgcc 2760

atggccgcgg gatatcacta gtgcggccgc tcgacgaatt aattccaatc ccacaaaaat 2820

ctgagcttaa cagcacagtt gctcctctca gagcagaatc gggtattcaa caccctcata 2880

tcaactacta cgttgtgtat aacggtccac atgccggtat atacgatgac tggggttgta 2940

caaaggcggc aacaaacggc gttcccggag ttgcacacaa gaaatttgcc actattacag 3000

aggcaagagc agcagctgac gcgtacacaa caagtcagca aacagacagg ttgaacttca 3060

tccccaaagg agaagctcaa ctcaagccca agagctttgc taaggcccta acaagcccac 3120

caaagcaaaa agcccactgg ctcacgctag gaaccaaaag gcccagcagt gatccagccc 3180

caaaagagat ctcctttgcc ccggagatta caatggacga tttcctctat ctttacgatc 3240

taggaaggaa gttcgaaggt gaaggtgacg acactatgtt caccactgat aatgagaagg 3300

ttagcctctt caatttcaga aagaatgctg acccacagat ggttagagag gcctacgcag 3360

caggtctcat caagacgatc tacccgagta acaatctcca ggagatcaaa taccttccca 3420

agaaggttaa agatgcagtc aaaagattca ggactaattg catcaagaac acagagaaag 3480

acatatttct caagatcaga agtactattc cagtatggac gattcaaggc ttgcttcata 3540

aaccaaggca agtaatagag attggagtct ctaaaaaggt agttcctact gaatctaagg 3600

ccatgcatgg agtctaagat tcaaatcgag gatctaacag aactcgccgt gaagactggc 3660

gaacagttca tacagagtct tttacgactc aatgacaaga agaaaatctt cgtcaacatg 3720

gtggagcacg acactctggt ctactccaaa aatgtcaaag atacagtctc agaagaccaa 3780

agggctattg agacttttca acaaaggata atttcgggaa acctcctcgg attccattgc 3840

ccagctatct gtcacttcat cgaaaggaca gtagaaaagg aaggtggctc ctacaaatgc 3900

catcattgcg ataaaggaaa ggctatcatt caagatctct ctgccgacag tggtcccaaa 3960

gatggacccc cacccacgag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca 4020
aagcaagtgg attgatgtga catctccact gacgtaaggg atgacgcaca atcccactat 4080
ccttcgcaag acccttcctc tatataagga agttcatttc atttggagag gacacgctcg 4140
aggaattcgg taccccatca caagtttgta caaaaaagca ggctgcggcc gcttgctccc 4200
ttaaaaaaaa ccatggcatg tcattacgga caacagcaac agacttgtgc acctcatctt 4260
cagcttcaac ctagagcatg tagagttgtg aaggctgcta ctgctgttac tgtaaatttc 4320
tgtgttcctt attctctcaa aatcttcgat tttgttttcg ttcgatccca atttcgtata 4380
tgttctttgg tttagattct gttaatctta gatcgaagac gattttctgg gtttgatcgt 4440
tagatatcat cttaattctc gattagggtt tcatagatat catccgattt gttcaaataa 4500
tttgagtttt gtcgaataat tactcttcga tttgtgattt ctatctagat ctggtgttag 4560
tttctagttt gtgcgatcga atttgtcgat taatctgagt ttttctgatt aacaggctgg 4620
aggatctctt cttgttctct ctggacttac tctcgctgga actgttatcg ctctcactat 4680
cgctacacct cttctcgtta tcttctctcc tgttctcgtt cctgctgtga tcactatctt 4740
ccttctcgga gctggatttc ttgcttctgg tggatttgga gttgctgctc tctctgttct 4800
ctcttggatc tacagatacc tcactggaaa acatcctcca ggtgctgatc aacttgagtc 4860
tgctaagact aagctcgctt ctaaggctag agagatgaag gatagagcag agcaattctc 4920
ttgtcagcct gttgctggat ctcagacttc ttaatgaaca tatggtcctg ctttaatgag 4980
atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa 5040
aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata 5100
tatcacccgt tactatcgta tttttatgaa taatattctc cgttcaattt actgattgta 5160
ccctactact tatatgtaca atattaaaat gaaaacaata tattgtgctg aataggttta 5220
tagcgacatc tatgatagag cgccacaata acaaacaatt gcgttttatt attacaaatc 5280
caattttaaa aaaagcggca gaaccggtca aacctaaaag actgattaca taaatcttat 5340
tcaaatttca aaaggcccca ggggctagta tctacgacac accgagcggc gaactaataa 5400
cgttcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt ccttgaagtt 5460
gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaacccgg tccagcacgg 5520
cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg tgtatgtggg 5580
ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg cgggtccagg 5640
gcgaattttg cgacaacatg tcgaggctca gcaggacctg caggcatgca agctagctta 5700
ctagtgatgc atattctata gtgtcaccta aatcttcgac gaattaattc caatcccaca 5760
aaaatctgag cttaacagca cagttgctcc tctcagagca gaatcgggta ttcaacaccc 5820
tcatatcaac tactacgttg tgtataacgg tccacatgcc ggtatatacg atgactgggg 5880
ttgtacaaag gcggcaacaa acggcgttcc cggagttgca cacaagaaat ttgccactat 5940
tacagaggca agagcagcag ctgacgcgta cacaacaagt cagcaaacag acaggttgaa 6000

cttcatcccc aaaggagaag ctcaactcaa gcccaagagc tttgctaagg ccctaacaag 6060

cccaccaaag caaaaagccc actggctcac gctaggaacc aaaaggccca gcagtgatcc 6120

agccccaaaa gagatctcct ttgccccgga gattacaatg gacgatttcc tctatcttta 6180

cgatctagga aggaagttcg aaggtgaagg tgacgacact atgttcacca ctgataatga 6240

gaaggttagc ctcttcaatt tcagaaagaa tgctgaccca cagatggtta gagaggccta 6300

cgcagcaggt ctcatcaaga cgatctaccc gagtaacaat ctccaggaga tcaaatacct 6360

tcccaagaag gttaaagatg cagtcaaaag attcaggact aattgcatca agaacacaga 6420

gaaagacata tttctcaaga tcagaagtac tattccagta tggacgattc aaggcttgct 6480

tcataaacca aggcaagtaa tagagattgg agtctctaaa aaggtagttc ctactgaatc 6540

taaggccatg catggagtct aagattcaaa tcgaggatct aacagaactc gccgtgaaga 6600

ctggcgaaca gttcatacag agtcttttac gactcaatga caagaagaaa atcttcgtca 6660

acatggtgga gcacgacact ctggtctact ccaaaaatgt caaagataca gtctcagaag 6720

accaaagggc tattgagact tttcaacaaa ggataatttc gggaaacctc ctcggattcc 6780

attgcccagc tatctgtcac ttcatcgaaa ggacagtaga aaaggaaggt ggctcctaca 6840

aatgccatca ttgcgataaa ggaaaggcta tcattcaaga tctctctgcc gacagtggtc 6900

ccaaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt ccaaccacgt 6960

cttcaaagca agtggattga tgtgacatct ccactgacgt aagggatgac gcacaatccc 7020

actatccttc gcaagaccct tcctctatat aaggaagttc atttcatttg gagaggacac 7080

gggatccttg ctccgttaaa aaaaaccatg gctatcctcg attctgctgg tgttactact 7140

gtgactgaga atggtggtgg agagttcgtt gatctcgata gactcagaag aagaaagtct 7200

agatctgtaa atttctgtgt tccttattct ctcaaaatct tcgattttgt tttcgttcga 7260

tcccaatttc gtatatgttc tttggtttag attctgttaa tcttagatcg aagacgattt 7320

tctgggtttg atcgttagat atcatcttaa ttctcgatta gggtttcata gatatcatcc 7380

gatttgttca aataatttga gttttgtcga ataattactc ttcgatttgt gatttctatc 7440

tagatctggt gttagtttct agtttgtgcg atcgaatttg tcgattaatc tgagtttttc 7500

tgattaacag gattcttcta acggacttct cctctctgga tctgataaca actctccttc 7560

tgatgatgtt ggtgctcctg ctgatgtgag agatagaatc gattctgtgg tgaacgatga 7620

tgctcaaggt actgctaacc tcgctggtga taataacggt ggaggtgata acaatggtgg 7680

aggaagaggt ggaggtgaag gtagaggaaa cgctgatgct actttcactt acagaccatc 7740

tgtgcctgct catagaagag ctagagagtc tcctctctct tctgatgcta tcttcaagca 7800

gtctcacgct ggacttttca acctctgtgt ggtggttctt atcgctgtga actctagact 7860

catcatcgag aacctcatga agtacggatg gctcatcaga actgatttct ggttctcttc 7920

tagatctctc agagattggc ctcttttcat gtgctgcatc tcactctcaa tcttccctct 7980

cgctgctttt actgttgaga agctcgtgct ccagaagtat atcgctgaac ctgtggtgat 8040

cttcctccac atcatcatca ctatgactga ggttctctac cctgtttacg tgactctcag 8100

4197737_1

atgcgattct gctttcctct ctggtgttac tcttatgctc ctcacttgca ttgtgtggct 8160
taagctcgtg tcttacgctc acacttctta cgatatcaga tctctcgcta acgctgctga 8220
taaggctaac cctgaagtgt cttactacgt gtctctcaag tctctcgctt acttcatggt 8280
tgctcctaca ctttgttacc agccatctta ccctagatct gcttgcatta gaaagggatg 8340
ggtggcaaga caattcgcta agttggtgat cttcactgga ttcatgggat tcatcatcga 8400
gcagtacatc aaccctattg tgagaaactc taagcaccct ctcaagggtg atcttctcta 8460
cgctatcgag agagttctta agctctctgt gcctaacctt tatgtgtggc tctgcatgtt 8520
ctactgtttc ttccacctct ggcttaacat ccttgctgag ttgctttgct tcggagatag 8580
agagttctac aaggattggt ggaacgctaa gtctgttgga gattattgga gaatgtggaa 8640
catgcctgtg cataagtgga tggtgcgtca catctacttc ccttgcctca gatctaagat 8700
ccctaagact ctcgctatca ttatcgcttt cctcgtgtct gctgttttcc atgagttgtg 8760
tatcgctgtt ccttgcagac ttttcaagct ttgggctttc ctcggaatca tgttccaggt 8820
tccactcgtg ttcatcacta actacctcca agagagattc ggatctactg ttggaaacat 8880
gattttctgg ttcattttct gcatcttcgg acagcctatg tgcgttctcc tctactacca 8940
cgatctcatg aacagaaagg gatctatgtc ttaatgaagg atccacccag ctttcttgta 9000
caaagtggtg atgggttcga aatcgataag cttggatcct ctagagtcct gctttaatga 9060
gatatgcgag acgcctatga tcgcatgata tttgctttca attctgttgt gcacgttgta 9120
aaaaacctga gcatgtgtag ctcagatcct taccgccggt ttcggttcat tctaatgaat 9180
atatcacccg ttactatcgt atttttatga ataatattct ccgttcaatt tactgattgt 9240
accctactac ttatatgtac aatattaaaa tgaaaacaat atattgtgct gaataggttt 9300
atagcgacat ctatgataga gcgccacaat aacaaacaat tgcgttttat tattacaaat 9360
ccaattttaa aaaaagcggc agaaccggtc aaacctaaaa gactgattac ataaatctta 9420
ttcaaatttc aaaaggcccc aggggctagt atctacgaca caccgagcgg cgaactaata 9480
acgttcactg aagggaactc cggttccccg ccggcgcgca tgggtgagat tccttgaagt 9540
tgagtattgg ccgtccgctc taccgaaagt tacgggcacc attcaacccg gtccagcacg 9600
gcggccgggt aaccgacttg ctgccccgag aattatgcag catttttttg gtgtatgtgg 9660
gccccaaatg aagtgcaggt caaaccttga cagtgacgac aaatcgttgg gcgggtccag 9720
ggcgaatttt gcgacaacat gtcgaggctc agcaggacct gcaggcatgc aagctagctt 9780
actagtgatg catattctat agtgtcacct aaatctgcgg ccgcctgcag gtcgatatgg 9840
gagagctccc aacgcgttgg atgcatagct tgagtattct atagtgtcac ctaaatagct 9900
tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 9960
acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 10020
tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 10080
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg gctgagtggc 10140

tccttcaacg ttgcggttct gtcagttcca aacgtaaaac ggcttgtccc gcgtcatcgg 10200

cgggggtcat aacgtgactc ccttaattct ccgctcatga tcagattgtc gtttcccgcc 10260

ttcagtttaa actatcagtg tttgacagga tatattggcg ggtaaaccta agagaaaaga 10320

gcgtttatta gaataatcgg atatttaaaa gggcgtgaaa aggtttatcc gttcgtccat 10380

ttgtatgtgc atgccaacca cagggttccc ctcgggagtg cttggcattc cgtgcgataa 10440

tgacttctgt tcaaccaccc aaacgtcgga aagcctgacg acggagcagc attccaaaaa 10500

gatcccttgg ctcgtctggg tcggctagaa ggtcgagtgg gctgctgtgg cttgatccct 10560

caacgcggtc gcggacgtag cgcagcgccg aaaaatcctc gatcgcaaat ccgacgctgt 10620

cgaaaatcgt gatctgcttg tcgctctttc ggccgacgtc ctggccagtc atcacgcgcc 10680

aaagttccgt cacaggatga tctggcgcga gttgctggat ctcgccttca atccgggtct 10740

gtggcgggaa ctccacgaaa atatccgaac gcagcaagat gtcgacggat cttttccgct 10800

gcataaccct gcttcggggt cattatagcg atttttcgg tatatccatc ctttttcgca 10860

cgatatacag gattttgcca aagggttcgt gtagactttc cttggtgtat ccaacggcgt 10920

cagccgggca ggataggtga agtaggccca cccgcgagcg ggtgttcctt cttcactgtc 10980

ccttattcgc acctggcggt gctcaacggg aatcctgctc tgcgaggctg gccggctacc 11040

gccggcgtaa cagatgaggg caagcggatg gctgatgaaa ccaagccaac caggggtgat 11100

gctgccaact tactgattta gtgtatgatg gtgtttttga ggtgctccag tggcttctgt 11160

ttctatcagc tgtccctcct gttcagctac tgacggggtg gtgcgtaacg gcaaaagcac 11220

cgccggacat cagcgctatc tctgctctca ctgccgtaaa acatggcaac tgcagttcac 11280

ttacaccgct tctcaacccg gtacgcacca gaaaatcatt gatatggcca tgaatggcgt 11340

tggatgccgg gcaacagccc gcattatggg cgttggcctc aacacgattt tacgtcactt 11400

aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc cgggcagtga cgtcatcgtc 11460

tgcgcggaaa tggacgaaca gtggggctat gtcggggcta aatcgcgcca gcgctggctg 11520

ttttacgcgt atgacagtct ccggaagacg gttgttgcgc acgtattcgg tgaacgcact 11580

atggcgacgc tggggcgtct tatgagcctg ctgtcaccct ttgacgtggt gatatggatg 11640

acggatggct ggccgctgta tgaatcccgc ctgaagggaa agctgcacgt aatcagcaag 11700

cgatatacgc agcgaattga gcggcataac ctgaatctga ggcagcacct ggcacggctg 11760

ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc atgacaaagt catcgggcat 11820

tatctgaaca taaaacacta tcaataagtt ggagtcatta cccaaccagg aagggcagcc 11880

cacctatcaa ggtgtactgc cttccagacg aacgaagagc gattgaggaa aaggcggcgg 11940

cggccggcat gagcctgtcg gcctacctgc tggccgtcgg ccagggctac aaaatcacgg 12000

gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat caatggcgac ctgggccgcc 12060

tgggcggcct gctgaaactc tggctcaccg acgacccgcg cacggcgcgg ttcggtgatg 12120

ccacgatcct cgccctgctg gcgaagatcg aagagaagca ggacgagctt ggcaaggtca 12180

tgatgggcgt ggtccgcccg agggcagagc catgactttt ttagccgcta aaacggccgg 12240

ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca tcaagaagag cgacttcgcg 12300
gagctggtat tcgtgcaggg caagattcgg aataccaagt acgagaagga cggccagacg 12360
gtctacggga ccgacttcat tgccgataag gtggattatc tggacaccaa ggcaccaggc 12420
gggtcaaatc aggaataagg gcacattgcc ccggcgtgag tcggggcaat cccgcaagga 12480
gggtgaatga atcggacgtt tgaccggaag gcatacaggc aagaactgat cgacgcgggg 12540
ttttccgccg aggatgccga aaccatcgca agccgcaccg tcatgcgtgc gccccgcgaa 12600
accttccagt ccgtcggctc gatggtccag caagctacgg ccaagatcga gcgcgacagc 12660
gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg ccgtggagcg ttcgcgtcgt 12720
ctcgaacagg aggcggcagg tttggcgaag tcgatgacca tcgacacgcg aggaactatg 12780
acgaccaaga agcgaaaaac cgccggcgag gacctggcaa aacaggtcag cgaggccaag 12840
caggccgcgt tgctgaaaca cacgaagcag cagatcaagg aaatgcagct ttccttgttc 12900
gatattgcgc cgtggccgga cacgatgcga gcgatgccaa acgacacggc ccgctctgcc 12960
ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc tgcaaaacaa ggtcattttc 13020
cacgtcaaca aggacgtgaa gatcacctac accggcgtcg agctgcgggc cgacgatgac 13080
gaactggtgt ggcagcaggt gttggagtac gcgaagcgca cccctatcgg cgagccgatc 13140
accttcacgt tctacgagct ttgccaggac ctgggctggt cgatcaatgg ccggtattac 13200
acgaaggccg aggaatgcct gtcgcgccta caggcgacgg cgatgggctt cacgtccgac 13260
cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct tccgcgtcct ggaccgtggc 13320
aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa tcgtcgtgct gtttgctggc 13380
gaccactaca cgaaattcat atgggagaag taccgcaagc tgtcgccgac ggcccgacgg 13440
atgttcgact atttcagctc gcaccgggag ccgtacccgc tcaagctgga aaccttccgc 13500
ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc gcgagcaggt cggcgaagcc 13560
tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct gggtcaatga tgacctggtg 13620
cattgcaaac gctagggcct tgtggggtca gttccggctg ggggttcagc agccagcgct 13680
ttactggcat ttcaggaaca agcgggcact gctcgacgca cttgcttcgc tcagtatcgc 13740
tcgggacgca cggcgcgctc tacgaactgc cgataaacag aggattaaaa ttgacaattg 13800
tgattaaggc tcagattcga cggcttggag cggccgacgt gcaggatttc cgcgagatcc 13860
gattgtcggc cctgaagaaa gctccagaga tgttcgggtc cgtttacgag cacgaggaga 13920
aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc cgtggcattc ggcgcctaca 13980
tcgacggcga gatcattggg ctgtcggtct tcaaacagga ggacggcccc aaggacgctc 14040
acaaggcgca tctgtccggc gttttcgtgg agcccgaaca gcgaggccga ggggtcgccg 14100
gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt gatgatcgtc cgacagattc 14160
caacgggaat ctggtggatg cgcatcttca tcctcggcgc acttaatatt tcgctattct 14220
ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg ggtcgcggcg acggtaggcg 14280

ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct gctaggtagc ccgatacgat 14340

tgatggcggt cctgggggct atttgcggaa ctgcgggcgt ggcgctgttg gtgttgacac 14400

caaacgcagc gctagatcct gtcggcgtcg cagcgggcct ggcgggggcg gtttccatgg 14460

cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt gcctctgctc acctttaccg 14520

cctggcaact ggcggccgga ggacttctgc tcgttccagt agctttagtg tttgatccgc 14580

caatcccgat gcctacagga accaatgttc tcggcctggc gtggctcggc ctgatcggag 14640

cgggtttaac ctacttcctt tggttccggg ggatctcgcg actcgaacct acagttgttt 14700

ccttactggg ctttctcagc cgggatggcg ctaagaagct attgccgccg atcttcatat 14760

gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcatcaggc gctcttccgc 14820

ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 14880

ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 14940

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 15000

taggctccgc cccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 15060

cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 15120

tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 15180

gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 15240

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 15300

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 15360

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 15420

cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 15480

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 15540

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat caagaagatc ctttgatctt 15600

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 15660

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 15720

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 15780

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat 15840

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 15900

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 15960

aagtggtcct gcaactttat ccgcctccat ccagtctatt aaacaagtgg cagcaacgga 16020

ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc aggtttgcga tccgctgtgc 16080

caggcgttag gcgtcatatg aagatttcgg tgatccctga gcaggtggcg gaaacattgg 16140

atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt gcacctatcc gaccaaggct 16200

ttgaactatc taccagaagt gtgagcccct accggaagga ttacatctcg gatgatgact 16260

ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca agagcttgtc gggaagattg 16320

aactcaactc aacatggaac gatctagcct ctatcgaaca cattgttgtg tcgcacacgc 16380

accgaggcaa aggagtcgcg cacagtctca tcgaatttgc gaaaaagtgg gcactaagca 16440
gacagctcct tggcatacga ttagagacac aaacgaacaa tgtacctgcc tgcaatttgt 16500
acgcaaaatg tggctttact ctcggcggca ttgacctgtt cacgtataaa actagacctc 16560
aagtctcgaa cgaaacagcg atgtactggt actggttctc gggagcacag gatgacgcct 16620
aacaattcat tcaagccgac accgcttcgc ggcgcggctt aattcaggag ttaaacatca 16680
tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc agaggtagtt ggcgtcatcg 16740
agcgccatct cgaaccgacg ttgctggccg tacatttgta cggctccgca gtggatggcg 16800
gcctgaagcc acacagtgat attgatttgc tggttacggt gaccgtaagg cttgatgaaa 16860
caacgcggcg agctttgatc aacgaccttt tggaaacttc ggcttcccct ggagagagcg 16920
agattctccg cgctgtagaa gtcaccattg ttgtgcacga cgacatcatt ccgtggcgtt 16980
atccagctaa gcgcgaactg caatttggag aatggcagcg caatgacatt cttgcaggta 17040
tcttcgagcc agccacgatc gacattgatc tggctatctt gctgacaaaa gcaagagaac 17100
atagcgttgc cttggtaggt ccagcggcgg aggaactctt tgatccggtt cctgaacagg 17160
atctatttga ggcgctaaat gaaaccttaa cgctatggaa ctcgccgccc gactgggctg 17220
gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg gtacagcgca gtaaccggca 17280
aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga gcgcctgccg gcccagtatc 17340
agcccgtcat acttgaagct aggcaggctt atcttggaca agaagatcgc ttggcctcgc 17400
gcgcagatca gttggaagaa tttgttcact acgtgaaagg cgagatcacc aaggtagtcg 17460
gcaaataatg tctaacaatt cgttcaagcc gacgccgctt cgcggcgcgg cttaactcaa 17520
gcgttagaga gctggggaag actatgcgcg atctgttgaa ggtggttcta agcctcgtac 17580
ttgcgatggc atcggggcag gcacttgctg acctgccaat tgttttagtg gatgaagctc 17640
gtcttcccta tgactactcc ccatccaact acgacatttc tccaagcaac tacgacaact 17700
ccataagcaa ttacgacaat agtccatcaa attacgacaa ctctgagagc aactacgata 17760
atagttcatc caattacgac aatagtcgca acggaaatcg taggcttata tatagcgcaa 17820
atgggtctcg cactttcgcc ggctactacg tcattgccaa caatgggaca acgaacttct 17880
tttccacatc tggcaaaagg atgttctaca ccccaaaagg ggggcgcggc gtctatggcg 17940
gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa tggccaattt tcgcttgccc 18000
tgacagataa cggcctgaag atcatgtatc taagcaacta gcctgctctc taataaaatg 18060
ttaggagctt ggctgccatt tttggggtga ggccgttcgc ggccgagggg cgcagcccct 18120
gggggggatgg gaggcccgcg ttagcgggcc gggagggttc gagaaggggg ggcacccccc 18180
ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt aaaaacaagg tttataaata 18240
ttggtttaaa agcaggttaa aagacaggtt agcggtggcc gaaaaacggg cggaaaccct 18300
tgcaaatgct ggattttctg cctgtggaca gcccctcaaa tgtcaatagg tgcgcccctc 18360
atctgtcagc actctgcccc tcaagtgtca aggatcgcgc ccctcatctg tcagtagtcg 18420

cgcccctcaa gtgtcaatac cgcagggcac ttatccccag gcttgtccac atcatctgtg 18480

ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga ggctggccag ctccacgtcg 18540

ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc gccgggtgag tcggcccctc 18600

aagtgtcaac gtccgcccct catctgtcag tgagggccaa gttttccgcg aggtatccac 18660

aacgccggcg gccggccgcg gtgtctcgca cacggcttcg acggcgtttc tggcgcgttt 18720

gcagggccat agacggccgc cagcccagcg gcgagggcaa ccagcccggt gagcgtcgga 18780

aagg 18784

<210> 30
<211> 18784
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 30
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60

atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120

gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420

atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480

ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540

aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600

aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660

gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720

agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780

ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840

aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900

cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960

cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020

cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080

cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140

gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200

catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260

cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320

gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380

4197737_1

ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440
ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500
cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560
gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620
cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680
ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740
tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800
acttgctttg aagacgtggt tggaacgtct tctttttcca cgatgctcct cgtgggtggg 1860
ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920
caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980
atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040
atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100
tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160
ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220
tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280
tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340
gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400
atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460
gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520
acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580
atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640
attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700
attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760
tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820
tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880
caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940
aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000
ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060
ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120
aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180
aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240
aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300
tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360
aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420

gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480

catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540

accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600

catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660

aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720

tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780

gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840

cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900

atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960

atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020

agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080

cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140

ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200

taaaaaaaac catggcatgt cattacggac aacagcaaca gacttgtgca cctcatcttc 4260

agcttcaacc tagagcatgt agagttgtga aggctgctac tgctgttact gtaaatttct 4320

gtgttcctta ttctctcaaa atcttcgatt ttgttttcgt tcgatcccaa tttcgtatat 4380

gttctttggt ttagattctg ttaatcttag atcgaagacg attttctggg tttgatcgtt 4440

agatatcatc ttaattctcg attagggttt catagatatc atccgatttg ttcaaataat 4500

ttgagttttg tcgaataatt actcttcgat ttgtgatttc tatctagatc tggtgttagt 4560

ttctagtttg tgcgatcgaa tttgtcgatt aatctgagtt tttctgatta acaggctgga 4620

ggatctcttc ttgttctctc tggacttact ctcgctggaa ctgttatcgc tctcactatc 4680

gctacacctc ttctcgttat cttctctcct gttctcgttc ctgctgtgat cactatcttc 4740

cttctcggag ctggatttct tgcttctggt ggatttggag ttgctgctct ctctgttctc 4800

tcttggatct acagatacct cactggaaaa catcctccag gtgctgattg tcttgagtct 4860

gctaagacta agctcgcttc ttgtgctaga gagatgaagg atagagcaga gcaattctct 4920

tgtcagcctg ttgctggatc tcagacttct taatgaacat atggtcctgc tttaatgaga 4980

tatgcgagac gcctatgatc gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa 5040

aaacctgagc atgtgtagct cagatcctta ccgccggttt cggttcattc taatgaatat 5100

atcacccgtt actatcgtat ttttatgaat aatattctcc gttcaattta ctgattgtac 5160

cctactactt atatgtacaa tattaaaatg aaaacaatat attgtgctga ataggtttat 5220

agcgacatct atgatagagc gccacaataa caaacaattg cgttttatta ttacaaatcc 5280

aattttaaaa aaagcggcag aaccggtcaa acctaaaaga ctgattacat aaatcttatt 5340

caaatttcaa aaggccccag gggctagtat ctacgacaca ccgagcggcg aactaataac 5400

gttcactgaa gggaactccg gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg 5460

agtattggcc gtccgctcta ccgaaagtta cgggcaccat tcaacccggt ccagcacggc 5520

4197737_1

ggccgggtaa ccgacttgct gccccgagaa ttatgcagca ttttttggt gtatgtgggc 5580
cccaaatgaa gtgcaggtca aaccttgaca gtgacgacaa atcgttgggc gggtccaggg 5640
cgaattttgc gacaacatgt cgaggctcag caggacctgc aggcatgcaa gctagcttac 5700
tagtgatgca tattctatag tgtcacctaa atcttcgacg aattaattcc aatcccacaa 5760
aaatctgagc ttaacagcac agttgctcct ctcagagcag aatcgggtat tcaacaccct 5820
catatcaact actacgttgt gtataacggt ccacatgccg gtatatacga tgactggggt 5880
tgtacaaagg cggcaacaaa cggcgttccc ggagttgcac acaagaaatt tgccactatt 5940
acagaggcaa gagcagcagc tgacgcgtac acaacaagtc agcaaacaga caggttgaac 6000
ttcatcccca aaggagaagc tcaactcaag cccaagagct ttgctaaggc cctaacaagc 6060
ccaccaaagc aaaaagccca ctggctcacg ctaggaacca aaaggcccag cagtgatcca 6120
gccccaaaag agatctcctt tgccccggag attacaatgg acgatttcct ctatctttac 6180
gatctaggaa ggaagttcga aggtgaaggt gacgacacta tgttcaccac tgataatgag 6240
aaggttagcc tcttcaattt cagaaagaat gctgacccac agatggttag agaggcctac 6300
gcagcaggtc tcatcaagac gatctacccg agtaacaatc tccaggagat caaatacctt 6360
cccaagaagg ttaaagatgc agtcaaaaga ttcaggacta attgcatcaa gaacacagag 6420
aaagacatat ttctcaagat cagaagtact attccagtat ggacgattca aggcttgctt 6480
cataaaccaa ggcaagtaat agagattgga gtctctaaaa aggtagttcc tactgaatct 6540
aaggccatgc atggagtcta agattcaaat cgaggatcta acagaactcg ccgtgaagac 6600
tggcgaacag ttcatacaga gtcttttacg actcaatgac aagaagaaaa tcttcgtcaa 6660
catggtggag cacgacactc tggtctactc caaaaatgtc aaagatacag tctcagaaga 6720
ccaaagggct attgagactt ttcaacaaag gataatttcg ggaaacctcc tcggattcca 6780
ttgcccagct atctgtcact tcatcgaaag gacagtagaa aaggaaggtg gctcctacaa 6840
atgccatcat tgcgataaag gaaaggctat cattcaagat ctctctgccg acagtggtcc 6900
caaagatgga cccccaccca cgaggagcat cgtggaaaaa gaagacgttc caaccacgtc 6960
ttcaaagcaa gtggattgat gtgacatctc cactgacgta agggatgacg cacaatccca 7020
ctatccttcg caagaccctt cctctatata aggaagttca tttcatttgg agaggacacg 7080
ggatccttgc tccgttaaaa aaaaccatgg ctatcctcga ttctgctggt gttactactg 7140
tgactgagaa tggtggtgga gagttcgttg atctcgatag actcagaaga agaaagtcta 7200
gatctgtaaa tttctgtgtt ccttattctc tcaaaatctt cgattttgtt ttcgttcgat 7260
cccaatttcg tatatgttct ttggtttaga ttctgttaat cttagatcga agacgatttt 7320
ctgggtttga tcgttagata tcatcttaat tctcgattag ggtttcatag atatcatccg 7380
atttgttcaa ataatttgag ttttgtcgaa taattactct tcgatttgtg atttctatct 7440
agatctggtg ttagtttcta gtttgtgcga tcgaatttgt cgattaatct gagtttttct 7500
gattaacagg attcttctaa cggacttctc ctctctggat ctgataacaa ctctccttct 7560

gatgatgttg gtgctcctgc tgatgtgaga gatagaatcg attctgtggt gaacgatgat 7620

gctcaaggta ctgctaacct cgctggtgat aataacggtg gaggtgataa caatggtgga 7680

ggaagaggtg gaggtgaagg tagaggaaac gctgatgcta ctttcactta cagaccatct 7740

gtgcctgctc atagaagagc tagagagtct cctctctctt ctgatgctat cttcaagcag 7800

tctcacgctg gacttttcaa cctctgtgtg gtggttctta tcgctgtgaa ctctagactc 7860

atcatcgaga acctcatgaa gtacggatgg ctcatcagaa ctgatttctg gttctcttct 7920

agatctctca gagattggcc tcttttcatg tgctgcatct cactctcaat cttccctctc 7980

gctgctttta ctgttgagaa gctcgtgctc cagaagtata tcgctgaacc tgtggtgatc 8040

ttcctccaca tcatcatcac tatgactgag gttctctacc ctgtttacgt gactctcaga 8100

tgcgattctg ctttcctctc tggtgttact cttatgctcc tcacttgcat tgtgtggctt 8160

aagctcgtgt cttacgctca cacttcttac gatatcagat ctctcgctaa cgctgctgat 8220

aaggctaacc ctgaagtgtc ttactacgtg tctctcaagt ctctcgctta cttcatggtt 8280

gctcctacac tttgttacca gccatcttac cctagatctg cttgcattag aaagggatgg 8340

gtggcaagac aattcgctaa gttggtgatc ttcactggat tcatgggatt catcatcgag 8400

cagtacatca accctattgt gagaaactct aagcaccctc tcaagggtga tcttctctac 8460

gctatcgaga gagttcttaa gctctctgtg cctaaccttt atgtgtggct ctgcatgttc 8520

tactgtttct tccacctctg gcttaacatc cttgctgagt tgctttgctt cggagataga 8580

gagttctaca aggattggtg gaacgctaag tctgttggag attattggag aatgtggaac 8640

atgcctgtgc ataagtggat ggtgcgtcac atctacttcc cttgcctcag atctaagatc 8700

cctaagactc tcgctatcat tatcgctttc ctcgtgtctg ctgttttcca tgagttgtgt 8760

atcgctgttc cttgcagact tttcaagctt tgggctttcc tcggaatcat gttccaggtt 8820

ccactcgtgt tcatcactaa ctacctccaa gagagattcg gatctactgt tggaaacatg 8880

attttctggt tcattttctg catcttcgga cagcctatgt gcgttctcct ctactaccac 8940

gatctcatga acagaaaggg atctatgtct taatgaagga tccacccagc tttcttgtac 9000

aaagtggtga tgggttcgaa atcgataagc ttggatcctc tagagtcctg ctttaatgag 9060

atatgcgaga cgcctatgat cgcatgatat ttgctttcaa ttctgttgtg cacgttgtaa 9120

aaaacctgag catgtgtagc tcagatcctt accgccggtt tcggttcatt ctaatgaata 9180

tatcacccgt tactatcgta tttttatgaa taatattctc cgttcaattt actgattgta 9240

ccctactact tatatgtaca atattaaaat gaaaacaata tattgtgctg aataggttta 9300

tagcgacatc tatgatagag cgccacaata acaaacaatt gcgttttatt attacaaatc 9360

caattttaaa aaaagcggca gaaccggtca aacctaaaag actgattaca taaatcttat 9420

tcaaatttca aaaggcccca ggggctagta tctacgacac accgagcggc gaactaataa 9480

cgttcactga agggaactcc ggttccccgc cggcgcgcat gggtgagatt ccttgaagtt 9540

gagtattggc cgtccgctct accgaaagtt acgggcacca ttcaacccgg tccagcacgg 9600

cggccgggta accgacttgc tgccccgaga attatgcagc atttttttgg tgtatgtggg 9660

ccccaaatga agtgcaggtc aaaccttgac agtgacgaca aatcgttggg cgggtccagg 9720
gcgaattttg cgacaacatg tcgaggctca gcaggacctg caggcatgca agctagctta 9780
ctagtgatgc atattctata gtgtcaccta aatctgcggc cgcctgcagg tcgatatggg 9840
agagctccca acgcgttgga tgcatagctt gagtattcta tagtgtcacc taaatagctt 9900
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 9960
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 10020
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 10080
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggg ctgagtggct 10140
ccttcaacgt tgcggttctg tcagttccaa acgtaaaacg gcttgtcccg cgtcatcggc 10200
gggggtcata acgtgactcc cttaattctc cgctcatgat cagattgtcg tttcccgcct 10260
tcagtttaaa ctatcagtgt ttgacaggat atattggcgg gtaaacctaa gagaaaagag 10320
cgtttattag aataatcgga tatttaaaag ggcgtgaaaa ggtttatccg ttcgtccatt 10380
tgtatgtgca tgccaaccac agggttcccc tcgggagtgc ttggcattcc gtgcgataat 10440
gacttctgtt caaccaccca aacgtcggaa agcctgacga cggagcagca ttccaaaaag 10500
atcccttggc tcgtctgggt cggctagaag gtcgagtggg ctgctgtggc ttgatccctc 10560
aacgcggtcg cggacgtagc gcagcgccga aaaatcctcg atcgcaaatc cgacgctgtc 10620
gaaaatcgtg atctgcttgt cgctctttcg gccgacgtcc tggccagtca tcacgcgcca 10680
aagttccgtc acaggatgat ctggcgcgag ttgctggatc tcgccttcaa tccgggtctg 10740
tggcgggaac tccacgaaaa tatccgaacg cagcaagatg tcgacggatc ttttccgctg 10800
cataaccctg cttcggggtc attatagcga ttttttcggt atatccatcc tttttcgcac 10860
gatatacagg attttgccaa agggttcgtg tagactttcc ttggtgtatc caacggcgtc 10920
agccgggcag gataggtgaa gtaggcccac ccgcgagcgg gtgttccttc ttcactgtcc 10980
cttattcgca cctggcggtg ctcaacggga atcctgctct gcgaggctgg ccggctaccg 11040
ccggcgtaac agatgagggc aagcggatgg ctgatgaaac caagccaacc aggggtgatg 11100
ctgccaactt actgatttag tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt 11160
tctatcagct gtccctcctg ttcagctact gacggggtgg tgcgtaacgg caaaagcacc 11220
gccggacatc agcgctatct ctgctctcac tgccgtaaaa catggcaact gcagttcact 11280
tacaccgctt ctcaacccgg tacgcaccag aaaatcattg atatggccat gaatggcgtt 11340
ggatgccggg caacagcccg cattatgggc gttggcctca acacgatttt acgtcactta 11400
aaaaactcag gccgcagtcg gtaacctcgc gcatacagcc gggcagtgac gtcatcgtct 11460
gcgcggaaat ggacgaacag tggggctatg tcggggctaa atcgcgccag cgctggctgt 11520
tttacgcgta tgacagtctc cggaagacgg ttgttgcgca cgtattcggt gaacgcacta 11580
tggcgacgct ggggcgtctt atgagcctgc tgtcaccctt tgacgtggtg atatggatga 11640
cggatggctg gccgctgtat gaatcccgcc tgaagggaaa gctgcacgta atcagcaagc 11700

gatatacgca gcgaattgag cggcataacc tgaatctgag gcagcacctg gcacggctgg 11760

gacggaagtc gctgtcgttc tcaaaatcgg tggagctgca tgacaaagtc atcgggcatt 11820

atctgaacat aaaacactat caataagttg gagtcattac ccaaccagga agggcagccc 11880

acctatcaag gtgtactgcc ttccagacga acgaagagcg attgaggaaa aggcggcggc 11940

ggccggcatg agcctgtcgg cctacctgct ggccgtcggc cagggctaca aaatcacggg 12000

cgtcgtggac tatgagcacg tccgcgagct ggcccgcatc aatggcgacc tgggccgcct 12060

gggcggcctg ctgaaactct ggctcaccga cgacccgcgc acggcgcggt tcggtgatgc 12120

cacgatcctc gccctgctgg cgaagatcga agagaagcag gacgagcttg gcaaggtcat 12180

gatgggcgtg gtccgcccga gggcagagcc atgacttttt tagccgctaa aacggccggg 12240

gggtgcgcgt gattgccaag cacgtcccca tgcgctccat caagaagagc gacttcgcgg 12300

agctggtatt cgtgcagggc aagattcgga ataccaagta cgagaaggac ggccagacgg 12360

tctacgggac cgacttcatt gccgataagg tggattatct ggacaccaag gcaccaggcg 12420

ggtcaaatca ggaataaggg cacattgccc cggcgtgagt cggggcaatc ccgcaaggag 12480

ggtgaatgaa tcggacgttt gaccggaagg catacaggca agaactgatc gacgcggggt 12540

tttccgccga ggatgccgaa accatcgcaa gccgcaccgt catgcgtgcg ccccgcgaaa 12600

ccttccagtc cgtcggctcg atggtccagc aagctacggc caagatcgag cgcgacagcg 12660

tgcaactggc tccccctgcc ctgcccgcgc catcggccgc cgtggagcgt tcgcgtcgtc 12720

tcgaacagga ggcggcaggt ttggcgaagt cgatgaccat cgacacgcga ggaactatga 12780

cgaccaagaa gcgaaaaacc gccggcgagg acctggcaaa acaggtcagc gaggccaagc 12840

aggccgcgtt gctgaaacac acgaagcagc agatcaagga aatgcagctt tccttgttcg 12900

atattgcgcc gtggccggac acgatgcgag cgatgccaaa cgacacggcc cgctctgccc 12960

tgttcaccac gcgcaacaag aaaatcccgc gcgaggcgct gcaaaacaag gtcattttcc 13020

acgtcaacaa ggacgtgaag atcacctaca ccggcgtcga gctgcgggcc gacgatgacg 13080

aactggtgtg gcagcaggtg ttggagtacg cgaagcgcac ccctatcggc gagccgatca 13140

ccttcacgtt ctacgagctt tgccaggacc tgggctggtc gatcaatggc cggtattaca 13200

cgaaggccga ggaatgcctg tcgcgcctac aggcgacggc gatgggcttc acgtccgacc 13260

gcgttgggca cctggaatcg gtgtcgctgc tgcaccgctt ccgcgtcctg gaccgtggca 13320

agaaaacgtc ccgttgccag gtcctgatcg acgaggaaat cgtcgtgctg tttgctggcg 13380

accactacac gaaattcata tgggagaagt accgcaagct gtcgccgacg gcccgacgga 13440

tgttcgacta tttcagctcg caccgggagc cgtacccgct caagctggaa accttccgcc 13500

tcatgtgcgg atcggattcc acccgcgtga agaagtggcg cgagcaggtc ggcgaagcct 13560

gcgaagagtt gcgaggcagc ggcctggtgg aacacgcctg ggtcaatgat gacctggtgc 13620

attgcaaacg ctagggcctt gtggggtcag ttccggctgg gggttcagca gccagcgctt 13680

tactggcatt tcaggaacaa gcgggcactg ctcgacgcac ttgcttcgct cagtatcgct 13740

cgggacgcac ggcgcgctct acgaactgcc gataaacaga ggattaaaat tgacaattgt 13800

gattaaggct cagattcgac ggcttggagc ggccgacgtg caggatttcc gcgagatccg 13860
attgtcggcc ctgaagaaag ctccagagat gttcgggtcc gtttacgagc acgaggagaa 13920
aaagcccatg gaggcgttcg ctgaacggtt gcgagatgcc gtggcattcg gcgcctacat 13980
cgacggcgag atcattgggc tgtcggtctt caaacaggag gacggcccca aggacgctca 14040
caaggcgcat ctgtccggcg ttttcgtgga gcccgaacag cgaggccgag gggtcgccgg 14100
tatgctgctg cgggcgttgc cggcgggttt attgctcgtg atgatcgtcc gacagattcc 14160
aacgggaatc tggtggatgc gcatcttcat cctcggcgca cttaatattt cgctattctg 14220
gagcttgttg tttatttcgg tctaccgcct gccgggcggg gtcgcggcga cggtaggcgc 14280
tgtgcagccg ctgatggtcg tgttcatctc tgccgctctg ctaggtagcc cgatacgatt 14340
gatggcggtc ctgggggcta tttgcggaac tgcgggcgtg gcgctgttgg tgttgacacc 14400
aaacgcagcg ctagatcctg tcggcgtcgc agcgggcctg gcgggggcgg tttccatggc 14460
gttcggaacc gtgctgaccc gcaagtggca acctcccgtg cctctgctca cctttaccgc 14520
ctggcaactg gcggccggag gacttctgct cgttccagta gctttagtgt ttgatccgcc 14580
aatcccgatg cctacaggaa ccaatgttct cggcctggcg tggctcggcc tgatcggagc 14640
gggtttaacc tacttccttt ggttccgggg gatctcgcga ctcgaaccta cagttgtttc 14700
cttactgggc tttctcagcc gggatggcgc taagaagcta ttgccgccga tcttcatatg 14760
cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct 14820
tcctcgctca ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac 14880
tcaaaggcgg taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga 14940
gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat 15000
aggctccgcc cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac 15060
ccgacaggac tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct 15120
gttccgaccc tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg 15180
ctttctcaat gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg 15240
ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt 15300
cttgagtcca acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg 15360
attagcagag cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac 15420
ggctacacta gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga 15480
aaaagagttg gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt 15540
gtttgcaagc agcagattac gcgcagaaaa aaaggatatc aagaagatcc tttgatcttt 15600
tctacggggt ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga 15660
ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc 15720
taaagtatat atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct 15780
atctcagcga tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata 15840

actacgatac gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca 15900

cgctcaccgg ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga 15960

agtggtcctg caactttatc cgcctccatc cagtctatta aacaagtggc agcaacggat 16020

tcgcaaacct gtcacgcctt ttgtgccaaa agccgcgcca ggtttgcgat ccgctgtgcc 16080

aggcgttagg cgtcatatga agatttcggt gatccctgag caggtggcgg aaacattgga 16140

tgctgagaac catttcattg ttcgtgaagt gttcgatgtg cacctatccg accaaggctt 16200

tgaactatct accagaagtg tgagccccta ccggaaggat tacatctcgg atgatgactc 16260

tgatgaagac tctgcttgct atggcgcatt catcgaccaa gagcttgtcg ggaagattga 16320

actcaactca acatggaacg atctagcctc tatcgaacac attgttgtgt cgcacacgca 16380

ccgaggcaaa ggagtcgcgc acagtctcat cgaatttgcg aaaaagtggg cactaagcag 16440

acagctcctt ggcatacgat tagagacaca aacgaacaat gtacctgcct gcaatttgta 16500

cgcaaaatgt ggctttactc tcggcggcat tgacctgttc acgtataaaa ctagacctca 16560

agtctcgaac gaaacagcga tgtactggta ctggttctcg ggagcacagg atgacgccta 16620

acaattcatt caagccgaca ccgcttcgcg gcgcggctta attcaggagt taaacatcat 16680

gagggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga 16740

gcgccatctc gaaccgacgt tgctggccgt acatttgtac ggctccgcag tggatggcgg 16800

cctgaagcca cacagtgata ttgatttgct ggttacggtg accgtaaggc ttgatgaaac 16860

aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga 16920

gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcgtta 16980

tccagctaag cgcgaactgc aatttggaga atggcagcgc aatgacattc ttgcaggtat 17040

cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca 17100

tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccggttc ctgaacagga 17160

tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg 17220

cgatgagcga aatgtagtgc ttacgttgtc ccgcatttgg tacagcgcag taaccggcaa 17280

aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca 17340

gcccgtcata cttgaagcta ggcaggctta tcttggacaa gaagatcgct tggcctcgcg 17400

cgcagatcag ttggaagaat ttgttcacta cgtgaaaggc gagatcacca aggtagtcgg 17460

caaataatgt ctaacaattc gttcaagccg acgccgcttc gcggcgcggc ttaactcaag 17520

cgttagagag ctggggaaga ctatgcgcga tctgttgaag gtggttctaa gcctcgtact 17580

tgcgatggca tcggggcagg cacttgctga cctgccaatt gttttagtgg atgaagctcg 17640

tcttccctat gactactccc catccaacta cgacatttct ccaagcaact acgacaactc 17700

cataagcaat tacgacaata gtccatcaaa ttacgacaac tctgagagca actacgataa 17760

tagttcatcc aattacgaca atagtcgcaa cggaaatcgt aggcttatat atagcgcaaa 17820

tgggtctcgc actttcgccg gctactacgt cattgccaac aatgggacaa cgaacttctt 17880

ttccacatct ggcaaaagga tgttctacac cccaaaaggg gggcgcggcg tctatggcgg 17940

caaagatggg agcttctgcg gggcattggt cgtcataaat ggccaatttt cgcttgccct 18000

gacagataac ggcctgaaga tcatgtatct aagcaactag cctgctctct aataaaatgt 18060

taggagcttg gctgccattt ttggggtgag gccgttcgcg gccgaggggc gcagcccctg 18120

gggggatggg aggcccgcgt tagcgggccg ggagggttcg agaaggggg gcacccccct 18180

tcggcgtgcg cggtcacgcg ccagggcgca gccctggtta aaaacaaggt ttataaatat 18240

tggtttaaaa gcaggttaaa agacaggtta gcggtggccg aaaaacgggc ggaaaccctt 18300

gcaaatgctg gattttctgc ctgtggacag cccctcaaat gtcaataggt gcgcccctca 18360

tctgtcagca ctctgcccct caagtgtcaa ggatcgcgcc cctcatctgt cagtagtcgc 18420

gcccctcaag tgtcaatacc gcagggcact tatccccagg cttgtccaca tcatctgtgg 18480

gaaactcgcg taaaatcagg cgttttcgcc gatttgcgag gctggccagc tccacgtcgc 18540

cggccgaaat cgagcctgcc cctcatctgt caacgccgcg ccgggtgagt cggcccctca 18600

agtgtcaacg tccgcccctc atctgtcagt gagggccaag ttttccgcga ggtatccaca 18660

acgccggcgg ccggccgcgg tgtctcgcac acggcttcga cggcgtttct ggcgcgtttg 18720

cagggccata gacggccgcc agcccagcgg cgagggcaac cagcccggtg agcgtcggaa 18780

aggg 18784

<210> 31
<211> 18802
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 31
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60

atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120

gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420

atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480

ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540

aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600

aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660

gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720

agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780

ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840

4197737_1

aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900

cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960

cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020

cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080

cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140

gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200

catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260

cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320

gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380

ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440

ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500

cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560

gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620

cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680

ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740

tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800

acttgctttg aagacgtggt tggaacgtct tctttttcca cgatgctcct cgtgggtggg 1860

ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920

caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980

atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040

atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100

tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160

ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220

tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280

tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340

gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400

atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460

gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520

acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580

atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640

attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700

attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760

tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820

tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880

caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940

aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000
ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060
ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120
aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180
aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240
aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300
tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360
aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420
gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480
catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540
accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600
catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660
aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720
tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780
gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840
cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900
atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960
atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020
agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080
cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140
ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200
taaaaaaaac catggcatgt cattacggat gtggacaaca gcaacagact tgtgcacctc 4260
atctttgtca gcttcaacct agagcatgta gagttgtgaa ggctgctact gctgttactg 4320
taaatttctg tgttccttat tctctcaaaa tcttcgattt tgttttcgtt cgatcccaat 4380
ttcgtatatg ttctttggtt tagattctgt taatcttaga tcgaagacga ttttctgggt 4440
ttgatcgtta gatatcatct taattctcga ttagggtttc atagatatca tccgatttgt 4500
tcaaataatt tgagttttgt cgaataatta ctcttcgatt tgtgatttct atctagatct 4560
ggtgttagtt tctagtttgt gcgatcgaat ttgtcgatta atctgagttt ttctgattaa 4620
caggctggag gatctcttct tgttctctct ggacttactc tcgctggaac tgttatcgct 4680
ctcactatcg ctacacctct tctcgttatc ttctctcctg ttctcgttcc tgctgtgatc 4740
actatcttcc ttctcggagc tggatttctt gcttctggtg gatttggagt tgctgctctc 4800
tctgttctct cttggatcta cagatacctc actggatgta aacatcctcc aggtgctgat 4860
tgtcttgagt ctgcttgtaa gactaagctc gcttcttgtg ctagagagat gaaggattgt 4920
agagcagagc aattctcttg tcagcctgtt gctggatctc agacttctta atgaacatat 4980

ggtcctgctt taatgagata tgcgagacgc ctatgatcgc atgatatttg ctttcaattc 5040

tgttgtgcac gttgtaaaaa acctgagcat gtgtagctca gatccttacc gccggtttcg 5100

gttcattcta atgaatatat cacccgttac tatcgtattt ttatgaataa tattctccgt 5160

tcaatttact gattgtaccc tactacttat atgtacaata ttaaaatgaa aacaatatat 5220

tgtgctgaat aggtttatag cgacatctat gatagagcgc cacaataaca aacaattgcg 5280

ttttattatt acaaatccaa ttttaaaaaa agcggcagaa ccggtcaaac ctaaaagact 5340

gattacataa atcttattca aatttcaaaa ggccccaggg gctagtatct acgacacacc 5400

gagcggcgaa ctaataacgt tcactgaagg gaactccggt tccccgccgg cgcgcatggg 5460

tgagattcct tgaagttgag tattggccgt ccgctctacc gaaagttacg ggcaccattc 5520

aacccggtcc agcacggcgg ccgggtaacc gacttgctgc cccgagaatt atgcagcatt 5580

tttttggtgt atgtgggccc caaatgaagt gcaggtcaaa ccttgacagt gacgacaaat 5640

cgttgggcgg gtccagggcg aattttgcga caacatgtcg aggctcagca ggacctgcag 5700

gcatgcaagc tagcttacta gtgatgcata ttctatagtg tcacctaaat cttcgacgaa 5760

ttaattccaa tcccacaaaa atctgagctt aacagcacag ttgctcctct cagagcagaa 5820

tcgggtattc aacaccctca tatcaactac tacgttgtgt ataacggtcc acatgccggt 5880

atatacgatg actggggttg tacaaaggcg gcaacaaacg gcgttcccgg agttgcacac 5940

aagaaatttg ccactattac agaggcaaga gcagcagctg acgcgtacac aacaagtcag 6000

caaacagaca ggttgaactt catccccaaa ggagaagctc aactcaagcc caagagcttt 6060

gctaaggccc taacaagccc accaaagcaa aaagcccact ggctcacgct aggaaccaaa 6120

aggcccagca gtgatccagc cccaaaagag atctcctttg ccccggagat tacaatggac 6180

gatttcctct atctttacga tctaggaagg aagttcgaag gtgaaggtga cgacactatg 6240

ttcaccactg ataatgagaa ggttagcctc ttcaatttca gaaagaatgc tgacccacag 6300

atggttagag aggcctacgc agcaggtctc atcaagacga tctacccgag taacaatctc 6360

caggagatca aataccttcc caagaaggtt aaagatgcag tcaaaagatt caggactaat 6420

tgcatcaaga acacagagaa agacatattt ctcaagatca gaagtactat tccagtatgg 6480

acgattcaag gcttgcttca taaaccaagg caagtaatag agattggagt ctctaaaaag 6540

gtagttccta ctgaatctaa ggccatgcat ggagtctaag attcaaatcg aggatctaac 6600

agaactcgcc gtgaagactg gcgaacagtt catacagagt cttttacgac tcaatgacaa 6660

gaagaaaatc ttcgtcaaca tggtggagca cgacactctg gtctactcca aaaatgtcaa 6720

agatacagtc tcagaagacc aaagggctat tgagactttt caacaaagga taatttcggg 6780

aaacctcctc ggattccatt gcccagctat ctgtcacttc atcgaaagga cagtagaaaa 6840

ggaaggtggc tcctacaaat gccatcattg cgataaagga aaggctatca ttcaagatct 6900

ctctgccgac agtggtccca aagatggacc cccacccacg aggagcatcg tggaaaaaga 6960

agacgttcca accacgtctt caaagcaagt ggattgatgt gacatctcca ctgacgtaag 7020

ggatgacgca caatcccact atccttcgca agacccttcc tctatataag gaagttcatt 7080

4197737_1

tcatttggag aggacacggg atccttgctc cgttaaaaaa aaccatggct atcctcgatt 7140
ctgctggtgt tactactgtg actgagaatg gtggtggaga gttcgttgat ctcgatagac 7200
tcagaagaag aaagtctaga tctgtaaatt tctgtgttcc ttattctctc aaaatcttcg 7260
attttgtttt cgttcgatcc caatttcgta tatgttcttt ggtttagatt ctgttaatct 7320
tagatcgaag acgattttct gggtttgatc gttagatatc atcttaattc tcgattaggg 7380
tttcatagat atcatccgat ttgttcaaat aatttgagtt ttgtcgaata attactcttc 7440
gatttgtgat ttctatctag atctggtgtt agtttctagt ttgtgcgatc gaatttgtcg 7500
attaatctga gtttttctga ttaacaggat tcttctaacg gacttctcct ctctggatct 7560
gataacaact ctccttctga tgatgttggt gctcctgctg atgtgagaga tagaatcgat 7620
tctgtggtga acgatgatgc tcaaggtact gctaacctcg ctggtgataa taacggtgga 7680
ggtgataaca atggtggagg aagaggtgga ggtgaaggta gaggaaacgc tgatgctact 7740
ttcacttaca gaccatctgt gcctgctcat agaagagcta gagagtctcc tctctcttct 7800
gatgctatct tcaagcagtc tcacgctgga cttttcaacc tctgtgtggt ggttcttatc 7860
gctgtgaact ctagactcat catcgagaac ctcatgaagt acggatggct catcagaact 7920
gatttctggt tctcttctag atctctcaga gattggcctc ttttcatgtg ctgcatctca 7980
ctctcaatct tccctctcgc tgcttttact gttgagaagc tcgtgctcca gaagtatatc 8040
gctgaacctg tggtgatctt cctccacatc atcatcacta tgactgaggt tctctaccct 8100
gtttacgtga ctctcagatg cgattctgct ttcctctctg gtgttactct tatgctcctc 8160
acttgcattg tgtggcttaa gctcgtgtct tacgctcaca cttcttacga tatcagatct 8220
ctcgctaacg ctgctgataa ggctaaccct gaagtgtctt actacgtgtc tctcaagtct 8280
ctcgcttact tcatggttgc tcctacactt tgttaccagc catcttaccc tagatctgct 8340
tgcattagaa agggatgggt ggcaagacaa ttcgctaagt tggtgatctt cactggattc 8400
atgggattca tcatcgagca gtacatcaac cctattgtga gaaactctaa gcaccctctc 8460
aagggtgatc ttctctacgc tatcgagaga gttcttaagc tctctgtgcc taacctttat 8520
gtgtggctct gcatgttcta ctgtttcttc cacctctggc ttaacatcct tgctgagttg 8580
ctttgcttcg gagatagaga gttctacaag gattggtgga acgctaagtc tgttggagat 8640
tattggagaa tgtggaacat gcctgtgcat aagtggatgg tgcgtcacat ctacttccct 8700
tgcctcagat ctaagatccc taagactctc gctatcatta tcgctttcct cgtgtctgct 8760
gttttccatg agttgtgtat cgctgttcct tgcagacttt tcaagctttg ggctttcctc 8820
ggaatcatgt tccaggttcc actcgtgttc atcactaact acctccaaga gagattcgga 8880
tctactgttg gaaacatgat tttctggttc attttctgca tcttcggaca gcctatgtgc 8940
gttctcctct actaccacga tctcatgaac agaaagggat ctatgtctta atgaaggatc 9000
cacccagctt tcttgtacaa agtggtgatg ggttcgaaat cgataagctt ggatcctcta 9060
gagtcctgct ttaatgagat atgcgagacg cctatgatcg catgatattt gctttcaatt 9120

ctgttgtgca cgttgtaaaa aacctgagca tgtgtagctc agatccttac cgccggtttc 9180
ggttcattct aatgaatata tcacccgtta ctatcgtatt tttatgaata atattctccg 9240
ttcaatttac tgattgtacc ctactactta tatgtacaat attaaaatga aaacaatata 9300
ttgtgctgaa taggtttata gcgacatcta tgatagagcg ccacaataac aaacaattgc 9360
gttttattat tacaaatcca attttaaaaa aagcggcaga accggtcaaa cctaaaagac 9420
tgattacata aatcttattc aaatttcaaa aggccccagg ggctagtatc tacgacacac 9480
cgagcggcga actaataacg ttcactgaag ggaactccgg ttccccgccg gcgcgcatgg 9540
gtgagattcc ttgaagttga gtattggccg tccgctctac cgaaagttac gggcaccatt 9600
caacccggtc agcacggcg gccgggtaac cgacttgctg ccccgagaat tatgcagcat 9660
ttttttggtg tatgtgggcc ccaaatgaag tgcaggtcaa accttgacag tgacgacaaa 9720
tcgttgggcg ggtccagggc gaattttgcg acaacatgtc gaggctcagc aggacctgca 9780
ggcatgcaag ctagcttact agtgatgcat attctatagt gtcacctaaa tctgcggccg 9840
cctgcaggtc gatatgggag agctcccaac gcgttggatg catagcttga gtattctata 9900
gtgtcaccta aatagcttgg cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta 9960
tccgctcaca attccacaca acatacgagc cggaagcata aagtgtaaag cctggggtgc 10020
ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt tccagtcggg 10080
aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag gcggtttgcg 10140
tattggggct gagtggctcc ttcaacgttg cggttctgtc agttccaaac gtaaaacggc 10200
ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 10260
gattgtcgtt tcccgccttc agtttaaact atcagtgttt gacaggatat attggcgggt 10320
aaacctaaga gaaaagagcg tttattagaa taatcggata tttaaaaggg cgtgaaaagg 10380
tttatccgtt cgtccatttg tatgtgcatg ccaaccacag ggttcccctc gggagtgctt 10440
ggcattccgt gcgataatga cttctgttca accacccaaa cgtcggaaag cctgacgacg 10500
gagcagcatt ccaaaaagat cccttggctc gtctgggtcg gctagaaggt cgagtgggct 10560
gctgtggctt gatccctcaa cgcggtcgcg gacgtagcgc agcgccgaaa aatcctcgat 10620
cgcaaatccg acgctgtcga aaatcgtgat ctgcttgtcg ctctttcggc cgacgtcctg 10680
gccagtcatc acgcgccaaa gttccgtcac aggatgatct ggcgcgagtt gctggatctc 10740
gccttcaatc cgggtctgtg gcgggaactc cacgaaaata tccgaacgca gcaagatgtc 10800
gacggatctt ttccgctgca taaccctgct tcggggtcat atagcgatt ttttcggtat 10860
atccatcctt tttcgcacga tatacaggat tttgccaaag ggttcgtgta gactttcctt 10920
ggtgtatcca acggcgtcag ccgggcagga taggtgaagt aggcccaccc gcgagcgggt 10980
gttccttctt cactgtccct tattcgcacc tggcggtgct caacgggaat cctgctctgc 11040
gaggctggcc ggctaccgcc ggcgtaacag atgagggcaa gcggatggct gatgaaacca 11100
agccaaccag gggtgatgct gccaacttac tgatttagtg tatgatggtg tttttgaggt 11160
gctccagtgg cttctgtttc tatcagctgt ccctcctgtt cagctactga cggggtggtg 11220

cgtaacggca aaagcaccgc cggacatcag cgctatctct gctctcactg ccgtaaaaca 11280
tggcaactgc agttcactta caccgcttct caacccggta cgcaccagaa aatcattgat 11340
atggccatga atggcgttgg atgccgggca acagcccgca ttatgggcgt tggcctcaac 11400
acgattttac gtcacttaaa aaactcaggc cgcagtcggt aacctcgcgc atacagccgg 11460
gcagtgacgt catcgtctgc gcggaaatgg acgaacagtg gggctatgtc ggggctaaat 11520
cgcgccagcg ctggctgttt tacgcgtatg acagtctccg gaagacggtt gttgcgcacg 11580
tattcggtga acgcactatg gcgacgctgg ggcgtcttat gagcctgctg tcacccttg 11640
acgtggtgat atggatgacg gatggctggc cgctgtatga atcccgcctg aagggaaagc 11700
tgcacgtaat cagcaagcga tatacgcagc gaattgagcg gcataacctg aatctgaggc 11760
agcacctggc acggctggga cggaagtcgc tgtcgttctc aaaatcggtg gagctgcatg 11820
acaaagtcat cgggcattat ctgaacataa aacactatca ataagttgga gtcattaccc 11880
aaccaggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 11940
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 12000
gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 12060
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 12120
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 12180
cgagcttggc aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gacttttta 12240
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtccccatg cgctccatca 12300
agaagagcga cttcgcggag ctggtattcg tgcagggcaa gattcggaat accaagtacg 12360
agaaggacgg ccagacggtc tacgggaccg acttcattgc cgataaggtg gattatctgg 12420
acaccaaggc accaggcggg tcaaatcagg aataagggca cattgccccg gcgtgagtcg 12480
gggcaatccc gcaaggaggg tgaatgaatc ggacgtttga ccggaaggca tacaggcaag 12540
aactgatcga cgcggggttt tccgccgagg atgccgaaac catcgcaagc cgcaccgtca 12600
tgcgtgcgcc ccgcgaaacc ttccagtccg tcggctcgat ggtccagcaa gctacggcca 12660
agatcgagcg cgacagcgtg caactggctc cccctgccct gcccgcgcca tcggccgccg 12720
tggagcgttc gcgtcgtctc gaacaggagg cggcaggttt ggcgaagtcg atgaccatcg 12780
acacgcgagg aactatgacg accaagaagc gaaaaaccgc cggcgaggac ctggcaaaac 12840
aggtcagcga ggccaagcag gccgcgttgc tgaaacacac gaagcagcag atcaaggaaa 12900
tgcagctttc cttgttcgat attgcgccgt ggccggacac gatgcgagcg atgccaaacg 12960
acacggcccg ctctgccctg ttcaccacgc gcaacaagaa aatcccgcgc gaggcgctgc 13020
aaaacaaggt cattttccac gtcaacaagg acgtgaagat cacctacacc ggcgtcgagc 13080
tgcgggccga cgatgacgaa ctggtgtggc agcaggtgtt ggagtacgcg aagcgcaccc 13140
ctatcggcga gccgatcacc ttcacgttct acgagctttg ccaggacctg ggctggtcga 13200
tcaatggccg gtattacacg aaggccgagg aatgcctgtc gcgcctacag gcgacggcga 13260

tgggcttcac gtccgaccgc gttgggcacc tggaatcggt gtcgctgctg caccgcttcc 13320

gcgtcctgga ccgtggcaag aaaacgtccc gttgccaggt cctgatcgac gaggaaatcg 13380

tcgtgctgtt tgctggcgac cactacacga aattcatatg ggagaagtac cgcaagctgt 13440

cgccgacggc ccgacggatg ttcgactatt tcagctcgca ccgggagccg tacccgctca 13500

agctggaaac cttccgcctc atgtgcggat cggattccac ccgcgtgaag aagtggcgcg 13560

agcaggtcgg cgaagcctgc gaagagttgc gaggcagcgg cctggtggaa cacgcctggg 13620

tcaatgatga cctggtgcat tgcaaacgct agggccttgt ggggtcagtt ccggctgggg 13680

gttcagcagc cagcgcttta ctggcatttc aggaacaagc gggcactgct cgacgcactt 13740

gcttcgctca gtatcgctcg ggacgcacgg cgcgctctac gaactgccga taaacagagg 13800

attaaaattg acaattgtga ttaaggctca gattcgacgg cttggagcgg ccgacgtgca 13860

ggatttccgc gagatccgat tgtcggccct gaagaaagct ccagagatgt tcgggtccgt 13920

ttacgagcac gaggagaaaa agcccatgga ggcgttcgct gaacggttgc gagatgccgt 13980

ggcattcggc gcctacatcg acggcgagat cattgggctg tcggtcttca aacaggagga 14040

cggccccaag gacgctcaca aggcgcatct gtccggcgtt ttcgtggagc ccgaacagcg 14100

aggccgaggg gtcgccggta tgctgctgcg ggcgttgccg gcgggtttat tgctcgtgat 14160

gatcgtccga cagattccaa cgggaatctg gtggatgcgc atcttcatcc tcggcgcact 14220

taatatttcg ctattctgga gcttgttgtt tatttcggtc taccgcctgc cgggcggggt 14280

cgcggcgacg gtaggcgctg tgcagccgct gatggtcgtg ttcatctctg ccgctctgct 14340

aggtagcccg atacgattga tggcggtcct gggggctatt tgcggaactg cgggcgtggc 14400

gctgttggtg ttgacaccaa acgcagcgct agatcctgtc ggcgtcgcag cgggcctggc 14460

gggggcggtt tccatggcgt tcggaaccgt gctgacccgc aagtggcaac ctcccgtgcc 14520

tctgctcacc tttaccgcct ggcaactggc ggccggagga cttctgctcg ttccagtagc 14580

tttagtgttt gatccgccaa tcccgatgcc tacaggaacc aatgttctcg gcctggcgtg 14640

gctcggcctg atcggagcgg gtttaaccta cttcctttgg ttccggggga tctcgcgact 14700

cgaacctaca gttgtttcct tactgggctt tctcagccgg gatggcgcta agaagctatt 14760

gccgccgatc ttcatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc 14820

atcaggcgct cttccgcttc ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg 14880

cgagcggtat cagctcactc aaaggcggta atacggttat ccacagaatc aggggataac 14940

gcaggaaaga acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg 15000

ttgctggcgt ttttccatag gctccgcccc cctgacgagc atcacaaaaa tcgacgctca 15060

agtcagaggt ggcgaaaccc gacaggacta taaagatacc aggcgtttcc ccctggaagc 15120

tccctcgtgc gctctcctgt tccgaccctg ccgcttaccg gatacctgtc cgcctttctc 15180

ccttcgggaa gcgtggcgct ttctcaatgc tcacgctgta ggtatctcag ttcggtgtag 15240

gtcgttcgct ccaagctggg ctgtgtgcac gaaccccccg ttcagcccga ccgctgcgcc 15300

ttatccggta actatcgtct tgagtccaac ccggtaagac acgacttatc gccactggca 15360

4197737_1

gcagccactg gtaacaggat tagcagagcg aggtatgtag gcggtgctac agagttcttg 15420
aagtggtggc ctaactacgg ctacactaga aggacagtat ttggtatctg cgctctgctg 15480
aagccagtta ccttcggaaa aagagttggt agctcttgat ccggcaaaca aaccaccgct 15540
ggtagcggtg gttttttgt ttgcaagcag cagattacgc gcagaaaaaa aggatatcaa 15600
gaagatcctt tgatcttttc tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa 15660
gggattttgg tcatgagatt atcaaaaagg atcttcacct agatcctttt aaattaaaaa 15720
tgaagtttta aatcaatcta aagtatatat gagtaaactt ggtctgacag ttaccaatgc 15780
ttaatcagtg aggcacctat ctcagcgatc tgtctatttc gttcatccat agttgcctga 15840
ctccccgtcg tgtagataac tacgatacgg gagggcttac catctggccc cagtgctgca 15900
atgataccgc gagacccacg ctcaccggct ccagatttat cagcaataaa ccagccagcc 15960
ggaagggccg agcgcagaag tggtcctgca actttatccg cctccatcca gtctattaaa 16020
caagtggcag caacggattc gcaaacctgt cacgcctttt gtgccaaaag ccgcgccagg 16080
tttgcgatcc gctgtgccag gcgttaggcg tcatatgaag atttcggtga tccctgagca 16140
ggtggcggaa acattggatg ctgagaacca tttcattgtt cgtgaagtgt tcgatgtgca 16200
cctatccgac caaggctttg aactatctac cagaagtgtg agcccctacc ggaaggatta 16260
catctcggat gatgactctg atgaagactc tgcttgctat ggcgcattca tcgaccaaga 16320
gcttgtcggg aagattgaac tcaactcaac atggaacgat ctagcctcta tcgaacacat 16380
tgttgtgtcg cacacgcacc gaggcaaagg agtcgcgcac agtctcatcg aatttgcgaa 16440
aaagtgggca ctaagcagac agctccttgg catacgatta gagacacaaa cgaacaatgt 16500
acctgcctgc aatttgtacg caaaatgtgg ctttactctc ggcggcattg acctgttcac 16560
gtataaaact agacctcaag tctcgaacga aacagcgatg tactggtact ggttctcggg 16620
agcacaggat gacgcctaac aattcattca agccgacacc gcttcgcggc gcggcttaat 16680
tcaggagtta aacatcatga gggaagcggt gatcgccgaa gtatcgactc aactatcaga 16740
ggtagttggc gtcatcgagc gccatctcga accgacgttg ctggccgtac atttgtacgg 16800
ctccgcagtg gatggcggcc tgaagccaca cagtgatatt gatttgctgg ttacggtgac 16860
cgtaaggctt gatgaaacaa cgcggcgagc tttgatcaac gacctttttgg aaacttcggc 16920
ttcccctgga gagagcgaga ttctccgcgc tgtagaagtc accattgttg tgcacgacga 16980
catcattccg tggcgttatc cagctaagcg cgaactgcaa tttggagaat ggcagcgcaa 17040
tgacattctt gcaggtatct tcgagccagc cacgatcgac attgatctgg ctatcttgct 17100
gacaaaagca agagaacata gcgttgcctt ggtaggtcca gcggcggagg aactctttga 17160
tccggttcct gaacaggatc tatttgaggc gctaaatgaa accttaacgc tatggaactc 17220
gccgcccgac tgggctggcg atgagcgaaa tgtagtgctt acgttgtccc gcatttggta 17280
cagcgcagta accggcaaaa tcgcgccgaa ggatgtcgct gccgactggg caatggagcg 17340
cctgccggcc cagtatcagc ccgtcatact tgaagctagg caggcttatc ttggacaaga 17400

agatcgcttg gcctcgcgcg cagatcagtt ggaagaattt gttcactacg tgaaaggcga 17460

gatcaccaag gtagtcggca aataatgtct aacaattcgt tcaagccgac gccgcttcgc 17520

ggcgcggctt aactcaagcg ttagagagct ggggaagact atgcgcgatc tgttgaaggt 17580

ggttctaagc ctcgtacttg cgatggcatc ggggcaggca cttgctgacc tgccaattgt 17640

tttagtggat gaagctcgtc ttccctatga ctactcccca tccaactacg acatttctcc 17700

aagcaactac gacaactcca taagcaatta cgacaatagt ccatcaaatt acgacaactc 17760

tgagagcaac tacgataata gttcatccaa ttacgacaat agtcgcaacg gaaatcgtag 17820

gcttatatat agcgcaaatg ggtctcgcac tttcgccggc tactacgtca ttgccaacaa 17880

tgggacaacg aacttctttt ccacatctgg caaaaggatg ttctacaccc caaaaggggg 17940

gcgcggcgtc tatggcggca aagatgggag cttctgcggg gcattggtcg tcataaatgg 18000

ccaattttcg cttgccctga cagataacgg cctgaagatc atgtatctaa gcaactagcc 18060

tgctctctaa taaaatgtta ggagcttggc tgccattttt ggggtgaggc cgttcgcggc 18120

cgaggggcgc agcccctggg gggatgggag gcccgcgtta gcgggccggg agggttcgag 18180

aagggggggc acccccccttc ggcgtgcgcg gtcacgcgcc agggcgcagc cctggttaaa 18240

aacaaggttt ataaatattg gtttaaaagc aggttaaaag acaggttagc ggtggccgaa 18300

aaacgggcgg aaacccttgc aaatgctgga ttttctgcct gtggacagcc cctcaaatgt 18360

caataggtgc gcccctcatc tgtcagcact ctgcccctca agtgtcaagg atcgcgcccc 18420

tcatctgtca gtagtcgcgc ccctcaagtg tcaataccgc agggcactta tccccaggct 18480

tgtccacatc atctgtggga aactcgcgta aaatcaggcg ttttcgccga tttgcgaggc 18540

tggccagctc cacgtcgccg gccgaaatcg agcctgcccc tcatctgtca acgccgcgcc 18600

gggtgagtcg gcccctcaag tgtcaacgtc cgcccctcat ctgtcagtga gggccaagtt 18660

ttccgcgagg tatccacaac gccggcggcc ggccgcggtg tctcgcacac ggcttcgacg 18720

gcgtttctgg cgcgtttgca gggccataga cggccgccag cccagcggcg agggcaacca 18780

gcccggtgag cgtcggaaag gg 18802

<210> 32
<211> 18823
<212> DNA
<213> Artificial Sequence

<220>
<223> Artificial

<400> 32
tcgacatcgt caacgttcac ttctaaagaa atagcgccac tcagcttcct cagcggcttt 60

atccagcgat ttcctattat gtcggcatag ttctcaagat cgacagcctg tcacggttaa 120

gcgagaaatg aataagaagg ctgataattc ggatctctgc gaggagatga tatttgatca 180

caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 240

gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcgggtaa catgagcaaa 300

gtctgccgcc ttacaacggc tctcccgctg acgccgtccc ggactgatgg gctgcctgta 360

tcgagtggtg attttgtgcc gagctgccgg tcggggagct gttggctggc tggtggcagg 420
atatattgtg gtgtaaacaa attgacgctt agacaactta ataacacatt gcggacgttt 480
ttaatgtact gaattcgccg ctcggtgtgt cgtagatact agcccctggg gcacttttga 540
aatttgaata agatttatgt aatcagtctt ttaggtttga ccggttctgc cgctttttt 600
aaaattggat ttgtaataat aaaacgcaat tgtttgttat tgtggcgctc tatcatagat 660
gtcgctataa acctattcag cacaatatat tgttttcatt ttaatattgt acatataagt 720
agtagggtac aatcagtaaa ttgaacggag aatattattc ataaaaatac gatagtaacg 780
ggtgatatat tcattagaat gaaccgaaac cggcggtaag gatctgagct acacatgctc 840
aggtttttta caacgtgcac aacagaattg aaagcaaata tcatgcgatc ataggcgtct 900
cgcatatctc attaaacaag tgaagatttg attcaaactc cattgagagc cctgactatg 960
cattcggttt gacccttcca ggttgagaga cgatagcccc ctaccttaat taaggggccc 1020
cccctcgagg tcgacggtat cgataagctt gatatcgaat tcctgcagcc agggggatcc 1080
cccgggtcat cagatctcgg tgacgggcag gaccggacgg ggcggtaccg gcaggctgaa 1140
gtccagctgc cagaaaccca cgtcatgcca gttcccgtgc ttgaagccgg ccgcccgcag 1200
catgccgcgg ggggcatatc cgagcgcctc gtgcatgcgc acgctcgggt cgttgggcag 1260
cccgatgaca gcgaccacgc tcttgaagcc ctgtgcctcc agggacttca gcaggtgggt 1320
gtagagcgtg gagcccagtc ccgtccgctg gtggcggggg gagacgtaca cggtcgactc 1380
ggccgtccag tcgtaggcgt tgcgtgcctt ccaggggccc gcgtaggcga tgccggcgac 1440
ctcgccgtcc acctcggcga cgagccaggg atagcgctcc cgcagacgga cgaggtcgtc 1500
cgtccactcc tgcggttcct gcggctcggt acggaagttg accgtgcttg tctcgatgta 1560
gtggttgacg atggtgcaga ccgccggcat gtccgcctcg gtggcacggc ggatgtcggc 1620
cgggcgtcgt tctgggctca tggtagatcc cctcgagaga gatagatttg tagagagaga 1680
ctggtgattt cagcgtgtcc tctccaaatg aaatgaactt ccttatatag aggaaggtct 1740
tgcgaaggat agtgggattg tgcgtcatcc cttacgtcag tggagatatc acatcaatcc 1800
acttgctttg aagacgtggt tggaacgtct tcttttttcca cgatgctcct cgtgggtggg 1860
ggtccatctt tgggaccact gtcggcagag gcatcttgaa cgatagcctt tcctttatcg 1920
caatgatggc atttgtaggt gccaccttcc ttttctactg tccttttgat gaagtgacag 1980
atagctgggc aatggaatcc gaggaggttt cccgatatta ccctttgttg aaaagtctca 2040
atagcccttt ggtcttctga gactgtatct ttgatattct tggagtagac gagagtgtcg 2100
tgctccacca tgttatcaca tcaatccact tgctttgaag acgtggttgg aacgtcttct 2160
ttttccacga tgctcctcgt gggtgggggt ccatctttgg gaccactgtc ggcagaggca 2220
tcttgaacga tagcctttcc tttatcgcaa tgatggcatt tgtaggtgcc accttccttt 2280
tctactgtcc ttttgatgaa gtgacagata gctgggcaat ggaatccgag gaggtttccc 2340
gatattaccc tttgttgaaa agtctcaata gccctttggt cttctgagac tgtatctttg 2400

atattcttgg agtagacgag agtgtcgtgc tccaccatgt tgctgatagt gaccttaggc 2460

gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 2520

acccattaac gcttacaatt tccattcgcc attcaggctg cgcaactgtt gggaagggcg 2580

atcggtgcgg gcctcttcgc tattacgcca gctggcgaaa gggggatgtg ctgcaaggcg 2640

attaagttgg gtaacgccag ggttttccca gtcacgacgt tgtaaaacga cggccagtga 2700

attgtaatac gactcactat agggcgaatt gggcccgacg tcgcatgctc ccggccgcca 2760

tggccgcggg atatcactag tgcggccgct cgacgaatta attccaatcc cacaaaaatc 2820

tgagcttaac agcacagttg ctcctctcag agcagaatcg ggtattcaac accctcatat 2880

caactactac gttgtgtata acggtccaca tgccggtata tacgatgact ggggttgtac 2940

aaaggcggca acaaacggcg ttcccggagt tgcacacaag aaatttgcca ctattacaga 3000

ggcaagagca gcagctgacg cgtacacaac aagtcagcaa acagacaggt tgaacttcat 3060

ccccaaagga gaagctcaac tcaagcccaa gagctttgct aaggccctaa caagcccacc 3120

aaagcaaaaa gcccactggc tcacgctagg aaccaaaagg cccagcagtg atccagcccc 3180

aaaagagatc tcctttgccc cggagattac aatggacgat ttcctctatc tttacgatct 3240

aggaaggaag ttcgaaggtg aaggtgacga cactatgttc accactgata atgagaaggt 3300

tagcctcttc aatttcagaa agaatgctga cccacagatg gttagagagg cctacgcagc 3360

aggtctcatc aagacgatct acccgagtaa caatctccag gagatcaaat accttcccaa 3420

gaaggttaaa gatgcagtca aaagattcag gactaattgc atcaagaaca cagagaaaga 3480

catatttctc aagatcagaa gtactattcc agtatggacg attcaaggct tgcttcataa 3540

accaaggcaa gtaatagaga ttggagtctc taaaaaggta gttcctactg aatctaaggc 3600

catgcatgga gtctaagatt caaatcgagg atctaacaga actcgccgtg aagactggcg 3660

aacagttcat acagagtctt ttacgactca atgacaagaa gaaaatcttc gtcaacatgg 3720

tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca gaagaccaaa 3780

gggctattga gacttttcaa caaaggataa tttcgggaaa cctcctcgga ttccattgcc 3840

cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc tacaaatgcc 3900

atcattgcga taaaggaaag gctatcattc aagatctctc tgccgacagt ggtcccaaag 3960

atggaccccc acccacgagg agcatcgtgg aaaaagaaga cgttccaacc acgtcttcaa 4020

agcaagtgga ttgatgtgac atctccactg acgtaaggga tgacgcacaa tcccactatc 4080

cttcgcaaga cccttcctct atataaggaa gttcatttca tttggagagg acacgctcga 4140

ggaattcggt accccatcac aagtttgtac aaaaaagcag gctgcggccg cttgctccct 4200

taaaaaaaac catggcatgt cattacggaa tggcatgtca ttacggatgt ggacaacagc 4260

aacagacttg tgcacctcat ctttgtcagc ttcaacctag agcatgtaga gttgtgaagg 4320

ctgctactgc tgttactgta aatttctgtg ttccttattc tctcaaaatc ttcgattttg 4380

ttttcgttcg atcccaattt cgtatatgtt ctttggttta gattctgtta atcttagatc 4440

gaagacgatt ttctgggttt gatcgttaga tatcatctta attctcgatt agggtttcat 4500

agatatcatc cgatttgttc aaataatttg agttttgtcg aataattact cttcgatttg 4560
tgatttctat ctagatctgg tgttagtttc tagtttgtgc gatcgaattt gtcgattaat 4620
ctgagttttt ctgattaaca ggctggagga tctcttcttg ttctctctgg acttactctc 4680
gctggaactg ttatcgctct cactatcgct acacctcttc tcgttatctt ctctcctgtt 4740
ctcgttcctg ctgtgatcac tatcttcctt ctcggagctg gatttcttgc ttctggtgga 4800
tttggagttg ctgctctctc tgttctctct tggatctaca gatacctcac tggatgtaaa 4860
catcctccag gtgctgattg tcttgagtct gcttgtaaga ctaagctcgc ttcttgtgct 4920
agagagatga aggattgtag agcagagcaa ttctcttgtc agcctgttgc tggatgttct 4980
cagacttctt aatgaacata tggtcctgct ttaatgagat atgcgagacg cctatgatcg 5040
catgatattt gctttcaatt ctgttgtgca cgttgtaaaa aacctgagca tgtgtagctc 5100
agatccttac cgccggtttc ggttcattct aatgaatata tcacccgtta ctatcgtatt 5160
tttatgaata atattctccg ttcaatttac tgattgtacc ctactactta tatgtacaat 5220
attaaaatga aaacaatata ttgtgctgaa taggtttata gcgacatcta tgatagagcg 5280
ccacaataac aaacaattgc gttttattat tacaaatcca attttaaaaa aagcggcaga 5340
accggtcaaa cctaaaagac tgattacata aatcttattc aaatttcaaa aggccccagg 5400
ggctagtatc tacgacacac cgagcggcga actaataacg ttcactgaag ggaactccgg 5460
ttccccgccg gcgcgcatgg gtgagattcc ttgaagttga gtattggccg tccgctctac 5520
cgaaagttac gggcaccatt caacccggtc agcacggcg gccgggtaac cgacttgctg 5580
ccccgagaat tatgcagcat ttttttggtg tatgtgggcc ccaaatgaag tgcaggtcaa 5640
accttgacag tgacgacaaa tcgttgggcg ggtccagggc gaattttgcg acaacatgtc 5700
gaggctcagc aggacctgca ggcatgcaag ctagcttact agtgatgcat attctatagt 5760
gtcacctaaa tcttcgacga attaattcca atcccacaaa aatctgagct taacagcaca 5820
gttgctcctc tcagagcaga atcgggtatt caacaccctc atatcaacta ctacgttgtg 5880
tataacggtc cacatgccgg tatatacgat gactggggtt gtacaaaggc ggcaacaaac 5940
ggcgttcccg gagttgcaca caagaaattt gccactatta cagaggcaag agcagcagct 6000
gacgcgtaca caacaagtca gcaaacagac aggttgaact tcatccccaa aggagaagct 6060
caactcaagc ccaagagctt tgctaaggcc ctaacaagcc caccaaagca aaaagcccac 6120
tggctcacgc taggaaccaa aaggcccagc agtgatccag ccccaaaaga gatctccttt 6180
gccccggaga ttacaatgga cgatttcctc tatctttacg atctaggaag gaagttcgaa 6240
ggtgaaggtg acgacactat gttcaccact gataatgaga aggttagcct cttcaatttc 6300
agaaagaatg ctgacccaca gatggttaga gaggcctacg cagcaggtct catcaagacg 6360
atctacccga gtaacaatct ccaggagatc aaataccttc ccaagaaggt taaagatgca 6420
gtcaaaagat tcaggactaa ttgcatcaag aacacagaga aagacatatt tctcaagatc 6480
agaagtacta ttccagtatg gacgattcaa ggcttgcttc ataaaccaag gcaagtaata 6540

gagattggag tctctaaaaa ggtagttcct actgaatcta aggccatgca tggagtctaa 6600

gattcaaatc gaggatctaa cagaactcgc cgtgaagact ggcgaacagt tcatacagag 6660

tcttttacga ctcaatgaca agaagaaaat cttcgtcaac atggtggagc acgacactct 6720

ggtctactcc aaaaatgtca aagatacagt ctcagaagac caaagggcta ttgagacttt 6780

tcaacaaagg ataatttcgg gaaacctcct cggattccat tgcccagcta tctgtcactt 6840

catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa tgccatcatt gcgataaagg 6900

aaaggctatc attcaagatc tctctgccga cagtggtccc aaagatggac ccccacccac 6960

gaggagcatc gtggaaaaag aagacgttcc aaccacgtct tcaaagcaag tggattgatg 7020

tgacatctcc actgacgtaa gggatgacgc acaatcccac tatccttcgc aagacccttc 7080

ctctatataa ggaagttcat ttcatttgga gaggacacgg gatccttgct ccgttaaaaa 7140

aaaccatggc tatcctcgat tctgctggtg ttactactgt gactgagaat ggtggtggag 7200

agttcgttga tctcgataga ctcagaagaa gaaagtctag atctgtaaat ttctgtgttc 7260

cttattctct caaaatcttc gattttgttt tcgttcgatc ccaatttcgt atatgttctt 7320

tggtttagat tctgttaatc ttagatcgaa gacgattttc tgggtttgat cgttagatat 7380

catcttaatt ctcgattagg gtttcataga tatcatccga tttgttcaaa taatttgagt 7440

tttgtcgaat aattactctt cgatttgtga tttctatcta gatctggtgt tagtttctag 7500

tttgtgcgat cgaatttgtc gattaatctg agtttttctg attaacagga ttcttctaac 7560

ggacttctcc tctctggatc tgataacaac tctccttctg atgatgttgg tgctcctgct 7620

gatgtgagag atagaatcga ttctgtggtg aacgatgatg ctcaaggtac tgctaacctc 7680

gctggtgata ataacggtgg aggtgataac aatggtggag gaagaggtgg aggtgaaggt 7740

agaggaaacg ctgatgctac tttcacttac agaccatctg tgcctgctca tagaagagct 7800

agagagtctc ctctctcttc tgatgctatc ttcaagcagt ctcacgctgg acttttcaac 7860

ctctgtgtgg tggttcttat cgctgtgaac tctagactca tcatcgagaa cctcatgaag 7920

tacggatggc tcatcagaac tgatttctgg ttctcttcta gatctctcag agattggcct 7980

cttttcatgt gctgcatctc actctcaatc ttccctctcg ctgcttttac tgttgagaag 8040

ctcgtgctcc agaagtatat cgctgaacct gtggtgatct tcctccacat catcatcact 8100

atgactgagg ttctctaccc tgtttacgtg actctcagat gcgattctgc tttcctctct 8160

ggtgttactc ttatgctcct cacttgcatt gtgtggctta agctcgtgtc ttacgctcac 8220

acttcttacg atatcagatc tctcgctaac gctgctgata aggctaaccc tgaagtgtct 8280

tactacgtgt ctctcaagtc tctcgcttac ttcatggttg ctcctacact ttgttaccag 8340

ccatcttacc ctagatctgc ttgcattaga aagggatggg tggcaagaca attcgctaag 8400

ttggtgatct tcactggatt catgggattc atcatcgagc agtacatcaa ccctattgtg 8460

agaaactcta agcaccctct caagggtgat cttctctacg ctatcgagag agttcttaag 8520

ctctctgtgc ctaaccttta tgtgtggctc tgcatgttct actgtttctt ccacctctgg 8580

cttaacatcc ttgctgagtt gctttgcttc ggagatagag agttctacaa ggattggtgg 8640

4197737_1

aacgctaagt ctgttggaga ttattggaga atgtggaaca tgcctgtgca taagtggatg 8700
gtgcgtcaca tctacttccc ttgcctcaga tctaagatcc ctaagactct cgctatcatt 8760
atcgctttcc tcgtgtctgc tgttttccat gagttgtgta tcgctgttcc ttgcagactt 8820
ttcaagcttt gggctttcct cggaatcatg ttccaggttc cactcgtgtt catcactaac 8880
tacctccaag agagattcgg atctactgtt ggaaacatga ttttctggtt cattttctgc 8940
atcttcggac agcctatgtg cgttctcctc tactaccacg atctcatgaa cagaaaggga 9000
tctatgtctt aatgaaggat ccacccagct ttcttgtaca aagtggtgat gggttcgaaa 9060
tcgataagct tggatcctct agagtcctgc tttaatgaga tatgcgagac gcctatgatc 9120
gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa aaacctgagc atgtgtagct 9180
cagatcctta ccgccggttt cggttcattc taatgaatat atcacccgtt actatcgtat 9240
ttttatgaat aatattctcc gttcaattta ctgattgtac cctactactt atatgtacaa 9300
tattaaaatg aaaacaatat attgtgctga ataggtttat agcgacatct atgatagagc 9360
gccacaataa caaacaattg cgttttatta ttacaaatcc aattttaaaa aaagcggcag 9420
aaccggtcaa acctaaaaga ctgattacat aaatcttatt caaatttcaa aaggccccag 9480
gggctagtat ctacgacaca ccgagcggcg aactaataac gttcactgaa gggaactccg 9540
gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg agtattggcc gtccgctcta 9600
ccgaaagtta cgggcaccat tcaacccggt ccagcacggc ggccgggtaa ccgacttgct 9660
gccccgagaa ttatgcagca ttttttggt gtatgtgggc cccaaatgaa gtgcaggtca 9720
aaccttgaca gtgacgacaa atcgttgggc gggtccaggg cgaattttgc gacaacatgt 9780
cgaggctcag caggacctgc aggcatgcaa gctagcttac tagtgatgca tattctatag 9840
tgtcacctaa atctgcggcc gcctgcaggt cgatatggga gagctcccaa cgcgttggat 9900
gcatagcttg agtattctat agtgtcacct aaatagcttg gcgtaatcat ggtcatagct 9960
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 10020
aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc 10080
actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 10140
cgcggggaga ggcggtttgc gtattggggc tgagtggctc cttcaacgtt gcggttctgt 10200
cagttccaaa cgtaaaacgg cttgtcccgc gtcatcggcg ggggtcataa cgtgactccc 10260
ttaattctcc gctcatgatc agattgtcgt ttcccgcctt cagtttaaac tatcagtgtt 10320
tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 10380
atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaaccaca 10440
gggttcccct cgggagtgct tggcattccg tgcgataatg acttctgttc aaccacccaa 10500
acgtcggaaa gcctgacgac ggagcagcat tccaaaaaga tcccttggct cgtctgggtc 10560
ggctagaagg tcgagtgggc tgctgtggct tgatccctca acgcggtcgc ggacgtagcg 10620
cagcgccgaa aaatcctcga tcgcaaatcc gacgctgtcg aaaatcgtga tctgcttgtc 10680

gctctttcgg ccgacgtcct ggccagtcat cacgcgccaa agttccgtca caggatgatc 10740

tggcgcgagt tgctggatct cgccttcaat ccgggtctgt ggcgggaact ccacgaaaat 10800

atccgaacgc agcaagatgt cgacggatct tttccgctgc ataaccctgc ttcggggtca 10860

ttatagcgat tttttcggta tatccatcct ttttcgcacg atatacagga ttttgccaaa 10920

gggttcgtgt agactttcct tggtgtatcc aacggcgtca gccgggcagg ataggtgaag 10980

taggcccacc cgcgagcggg tgttccttct tcactgtccc ttattcgcac ctggcggtgc 11040

tcaacgggaa tcctgctctg cgaggctggc cggctaccgc cggcgtaaca gatgagggca 11100

agcggatggc tgatgaaacc aagccaacca ggggtgatgc tgccaactta ctgatttagt 11160

gtatgatggt gtttttgagg tgctccagtg gcttctgttt ctatcagctg tccctcctgt 11220

tcagctactg acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctatctc 11280

tgctctcact gccgtaaaac atggcaactg cagttcactt acaccgcttc tcaacccggt 11340

acgcaccaga aaatcattga tatggccatg aatggcgttg gatgccgggc aacagcccgc 11400

attatgggcg ttggcctcaa cacgatttta cgtcacttaa aaaactcagg ccgcagtcgg 11460

taacctcgcg catacagccg ggcagtgacg tcatcgtctg cgcggaaatg gacgaacagt 11520

ggggctatgt cggggctaaa tcgcgccagc gctggctgtt ttacgcgtat gacagtctcc 11580

ggaagacggt tgttgcgcac gtattcggtg aacgcactat ggcgacgctg gggcgtctta 11640

tgagcctgct gtcacccttt gacgtggtga tatggatgac ggatggctgg ccgctgtatg 11700

aatcccgcct gaagggaaag ctgcacgtaa tcagcaagcg atatacgcag cgaattgagc 11760

ggcataacct gaatctgagg cagcacctgg cacggctggg acggaagtcg ctgtcgttct 11820

caaaatcggt ggagctgcat gacaaagtca tcgggcatta tctgaacata aaacactatc 11880

aataagttgg agtcattacc caaccaggaa gggcagccca cctatcaagg tgtactgcct 11940

tccagacgaa cgaagagcga ttgaggaaaa ggcggcggcg gccggcatga gcctgtcggc 12000

ctacctgctg gccgtcggcc agggctacaa aatcacgggc gtcgtggact atgagcacgt 12060

ccgcgagctg gcccgcatca atggcgacct gggccgcctg ggcggcctgc tgaaactctg 12120

gctcaccgac gacccgcgca cggcgcggtt cggtgatgcc acgatcctcg ccctgctggc 12180

gaagatcgaa gagaagcagg acgagcttgg caaggtcatg atgggcgtgg tccgcccgag 12240

ggcagagcca tgacttttt agccgctaaa acggccgggg ggtgcgcgtg attgccaagc 12300

acgtccccat gcgctccatc aagaagagcg acttcgcgga gctggtattc gtgcagggca 12360

agattcggaa taccaagtac gagaaggacg gccagacggt ctacgggacc gacttcattg 12420

ccgataaggt ggattatctg gacaccaagg caccaggcgg gtcaaatcag gaataagggc 12480

acattgcccc ggcgtgagtc ggggcaatcc cgcaaggagg gtgaatgaat cggacgtttg 12540

accggaaggc atacaggcaa gaactgatcg acgcggggtt ttccgccgag gatgccgaaa 12600

ccatcgcaag ccgcaccgtc atgcgtgcgc cccgcgaaac cttccagtcc gtcggctcga 12660

tggtccagca agctacggcc aagatcgagc gcgacagcgt gcaactggct ccccctgccc 12720

tgcccgcgcc atcggccgcc gtggagcgtt cgcgtcgtct cgaacaggag gcggcaggtt 12780

4197737_1

tggcgaagtc gatgaccatc gacacgcgag gaactatgac gaccaagaag cgaaaaaccg 12840
ccggcgagga cctggcaaaa caggtcagcg aggccaagca ggccgcgttg ctgaaacaca 12900
cgaagcagca gatcaaggaa atgcagcttt ccttgttcga tattgcgccg tggccggaca 12960
cgatgcgagc gatgccaaac gacacggccc gctctgccct gttcaccacg cgcaacaaga 13020
aaatcccgcg cgaggcgctg caaaacaagg tcattttcca cgtcaacaag gacgtgaaga 13080
tcacctacac cggcgtcgag ctgcgggccg acgatgacga actggtgtgg cagcaggtgt 13140
tggagtacgc gaagcgcacc cctatcggcg agccgatcac cttcacgttc tacgagcttt 13200
gccaggacct gggctggtcg atcaatggcc ggtattacac gaaggccgag gaatgcctgt 13260
cgcgcctaca ggcgacggcg atgggcttca cgtccgaccg cgttgggcac ctggaatcgg 13320
tgtcgctgct gcaccgcttc cgcgtcctgg accgtggcaa gaaaacgtcc cgttgccagg 13380
tcctgatcga cgaggaaatc gtcgtgctgt ttgctggcga ccactacacg aaattcatat 13440
gggagaagta ccgcaagctg tcgccgacgg cccgacggat gttcgactat ttcagctcgc 13500
accgggagcc gtacccgctc aagctggaaa ccttccgcct catgtgcgga tcggattcca 13560
cccgcgtgaa gaagtggcgc gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg 13620
gcctggtgga acacgcctgg gtcaatgatg acctggtgca ttgcaaacgc tagggccttg 13680
tggggtcagt tccggctggg ggttcagcag ccagcgcttt actggcattt caggaacaag 13740
cgggcactgc tcgacgcact tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta 13800
cgaactgccg ataaacagag gattaaaatt gacaattgtg attaaggctc agattcgacg 13860
gcttggagcg gccgacgtgc aggatttccg cgagatccga ttgtcggccc tgaagaaagc 13920
tccagagatg ttcgggtccg tttacgagca cgaggagaaa aagcccatgg aggcgttcgc 13980
tgaacggttg cgagatgccg tggcattcgg cgcctacatc gacggcgaga tcattgggct 14040
gtcggtcttc aaacaggagg acggccccaa ggacgctcac aaggcgcatc tgtccggcgt 14100
tttcgtggag cccgaacagc gaggccgagg ggtcgccggt atgctgctgc gggcgttgcc 14160
ggcgggttta ttgctcgtga tgatcgtccg acagattcca acgggaatct ggtggatgcg 14220
catcttcatc ctcggcgcac ttaatatttc gctattctgg agcttgttgt ttatttcggt 14280
ctaccgcctg ccgggcgggg tcgcggcgac ggtaggcgct gtgcagccgc tgatggtcgt 14340
gttcatctct gccgctctgc taggtagccc gatacgattg atggcggtcc tgggggctat 14400
ttgcggaact gcgggcgtgg cgctgttggt gttgacacca aacgcagcgc tagatcctgt 14460
cggcgtcgca gcgggcctgg cgggggcggt ttccatggcg ttcggaaccg tgctgacccg 14520
caagtggcaa cctcccgtgc ctctgctcac ctttaccgcc tggcaactgg cggccggagg 14580
acttctgctc gttccagtag ctttagtgtt tgatccgcca atcccgatgc ctacaggaac 14640
caatgttctc ggcctggcgt ggctcggcct gatcggagcg ggtttaacct acttcctttg 14700
gttccggggg atctcgcgac tcgaacctac agttgtttcc ttactgggct ttctcagccg 14760
ggatggcgct aagaagctat tgccgccgat cttcatatgc ggtgtgaaat accgcacaga 14820

tgcgtaagga gaaaataccg catcaggcgc tcttccgctt cctcgctcac tgactcgctg 14880

cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 14940

tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 15000

aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 15060

catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 15120

caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 15180

ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctcacgctgt 15240

aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 15300

gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 15360

cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 15420

ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 15480

tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 15540

tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 15600

cgcagaaaaa aaggatatca agaagatcct ttgatctttt ctacggggtc tgacgctcag 15660

tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 15720

tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 15780

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 15840

cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 15900

ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 15960

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 16020

gcctccatcc agtctattaa acaagtggca gcaacggatt cgcaaacctg tcacgccttt 16080

tgtgccaaaa gccgcgccag gtttgcgatc cgctgtgcca ggcgttaggc gtcatatgaa 16140

gatttcggtg atccctgagc aggtggcgga aacattggat gctgagaacc atttcattgt 16200

tcgtgaagtg ttcgatgtgc acctatccga ccaaggcttt gaactatcta ccagaagtgt 16260

gagcccctac cggaaggatt acatctcgga tgatgactct gatgaagact ctgcttgcta 16320

tggcgcattc atcgaccaag agcttgtcgg gaagattgaa ctcaactcaa catggaacga 16380

tctagcctct atcgaacaca ttgttgtgtc gcacacgcac cgaggcaaag gagtcgcgca 16440

cagtctcatc gaatttgcga aaaagtgggc actaagcaga cagctccttg gcatacgatt 16500

agagacacaa acgaacaatg tacctgcctg caatttgtac gcaaaatgtg gctttactct 16560

cggcggcatt gacctgttca cgtataaaac tagacctcaa gtctcgaacg aaacagcgat 16620

gtactggtac tggttctcgg gagcacagga tgacgcctaa caattcattc aagccgacac 16680

cgcttcgcgg cgcggcttaa ttcaggagtt aaacatcatg agggaagcgg tgatcgccga 16740

agtatcgact caactatcag aggtagttgg cgtcatcgag cgccatctcg aaccgacgtt 16800

gctggccgta catttgtacg gctccgcagt ggatggcggc ctgaagccac acagtgatat 16860

tgatttgctg gttacggtga ccgtaaggct tgatgaaaca acgcggcgag ctttgatcaa 16920

4197737_1

cgacctttttg gaaacttcgg cttcccctgg agagagcgag attctccgcg ctgtagaagt 16980
caccattgtt gtgcacgacg acatcattcc gtggcgttat ccagctaagc gcgaactgca 17040
atttggagaa tggcagcgca atgacattct tgcaggtatc ttcgagccag ccacgatcga 17100
cattgatctg gctatcttgc tgacaaaagc aagagaacat agcgttgcct tggtaggtcc 17160
agcggcggag gaactctttg atccggttcc tgaacaggat ctatttgagg cgctaaatga 17220
aaccttaacg ctatggaact cgccgcccga ctgggctggc gatgagcgaa atgtagtgct 17280
tacgttgtcc cgcatttggt acagcgcagt aaccggcaaa atcgcgccga aggatgtcgc 17340
tgccgactgg gcaatggagc gcctgccggc ccagtatcag cccgtcatac ttgaagctag 17400
gcaggcttat cttggacaag aagatcgctt ggcctcgcgc gcagatcagt tggaagaatt 17460
tgttcactac gtgaaaggcg agatcaccaa ggtagtcggc aaataatgtc taacaattcg 17520
ttcaagccga cgccgcttcg cggcgcggct taactcaagc gttagagagc tggggaagac 17580
tatgcgcgat ctgttgaagg tggttctaag cctcgtactt gcgatggcat cggggcaggc 17640
acttgctgac ctgccaattg ttttagtgga tgaagctcgt cttccctatg actactcccc 17700
atccaactac gacatttctc caagcaacta cgacaactcc ataagcaatt acgacaatag 17760
tccatcaaat tacgacaact ctgagagcaa ctacgataat agttcatcca attacgacaa 17820
tagtcgcaac ggaaatcgta ggcttatata tagcgcaaat gggtctcgca ctttcgccgg 17880
ctactacgtc attgccaaca atgggacaac gaacttcttt tccacatctg gcaaaaggat 17940
gttctacacc ccaaaagggg ggcgcggcgt ctatggcggc aaagatggga gcttctgcgg 18000
ggcattggtc gtcataaatg gccaattttc gcttgccctg acagataacg gcctgaagat 18060
catgtatcta agcaactagc ctgctctcta ataaaatgtt aggagcttgg ctgccatttt 18120
tggggtgagg ccgttcgcgg ccgaggggcg cagcccctgg ggggatggga ggcccgcgtt 18180
agcgggccgg gagggttcga gaaggggggg cacccccctt cggcgtgcgc ggtcacgcgc 18240
cagggcgcag ccctggttaa aaacaaggtt tataaatatt ggtttaaaag caggttaaaa 18300
gacaggttag cggtggccga aaaacgggcg gaaacccttg caaatgctgg attttctgcc 18360
tgtggacagc ccctcaaatg tcaataggtg cgcccctcat ctgtcagcac tctgccctc 18420
aagtgtcaag gatcgcgccc ctcatctgtc agtagtcgcg cccctcaagt gtcaataccg 18480
cagggcactt atccccaggc ttgtccacat catctgtggg aaactcgcgt aaaatcaggc 18540
gttttcgccg atttgcgagg ctggccagct ccacgtcgcc ggccgaaatc gagcctgccc 18600
ctcatctgtc aacgccgcgc cgggtgagtc ggcccctcaa gtgtcaacgt ccgcccctca 18660
tctgtcagtg agggccaagt tttccgcgag gtatccacaa cgccggcggc cggccgcggt 18720
gtctcgcaca cggcttcgac ggcgtttctg gcgcgtttgc agggccatag acggccgcca 18780
gcccagcggc gagggcaacc agcccggtga gcgtcggaaa ggg 18823

<210> 33
<211> 520
<212> PRT

4197737_1

<213> Artificial Sequence

<220>
<223> Artificial

<400> 33

```
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 Ala Glu Pro Val
        195                 200                 205

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

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
```

<210> 34
<211> 501
<212> DNA
<213> S. indicum

<400> 34
atggcggacc gcgaccgtcc acacccccac caaattcaag tccaccctca acatccgcac 60

cgctatgaag gtggcgtcaa gtctctcctc cctcaaaagg gcccctccac cacccagatt 120

ctcgccataa tcaccctcct tcccatcagc ggcacgcttc tttgcctagc tgggatcacg 180

ctcgtcggga ccctcatcgg acttgcagtc gccaccccag tcttcgtgat cttcagccct 240

gttctggttc ccgcagccat actgatagcc ggcgcggtca cggcgttttt gacgtccggg 300

gcttttgggc tgacggggct ttcgtcgctt tcttgggttc tgaattcatt cagacgggcg 360

acggggcagg ggccgttgga gtacgcgaag cgaggcgtgc aggaggggac tttgtatgtg 420

ggagagaaga cgaagcaagc gggcgaagcg attaagagca cagccaagga aggagggcga 480

gaagggactg cacggacttg a 501

<210> 35
<211> 166
<212> PRT
<213> S. indicum

<400> 35

Met Ala Asp Arg Asp Arg Pro His Pro His Gln Ile Gln Val His Pro
1 5 10 15

Gln His Pro His Arg Tyr Glu Gly Gly Val Lys Ser Leu Leu Pro Gln
20 25 30

Lys Gly Pro Ser Thr Thr Gln Ile Leu Ala Ile Ile Thr Leu Leu Pro
35 40 45

Ile Ser Gly Thr Leu Leu Cys Leu Ala Gly Ile Thr Leu Val Gly Thr
50 55 60

Leu Ile Gly Leu Ala Val Ala Thr Pro Val Phe Val Ile Phe Ser Pro
65 70 75 80

Val Leu Val Pro Ala Ala Ile Leu Ile Ala Gly Ala Val Thr Ala Phe
85 90 95

Leu Thr Ser Gly Ala Phe Gly Leu Thr Gly Leu Ser Ser Leu Ser Trp
100 105 110

Val Leu Asn Ser Phe Arg Arg Ala Thr Gly Gln Gly Pro Leu Glu Tyr
115 120 125

Ala Lys Arg Gly Val Gln Glu Gly Thr Leu Tyr Val Gly Glu Lys Thr
130 135 140

4197737_1

Lys Gln Ala Gly Glu Ala Ile Lys Ser Thr Ala Lys Glu Gly Gly Arg
145 150 155 160

Glu Gly Thr Ala Arg Thr
165

<210> 36
<211> 559
<212> DNA
<213> S. indicum

<400> 36
ggcacgagcg ccgtccccat ggcggacgaa ccccacgacc agcgccccac cgacgtcatc 60

aagagctacc tccccgaaaa gggtccctcc acctctcaag tcctcgccgt cgtgaccctc 120

ttccccctcg gcgccgtcct cctctgccta gccggtctca ttcttaccgg gaccatcatc 180

ggcctcgccg tcgccacccc gctcttcgtc atcttcagcc ccatcttggt ccccgccgcc 240

ctaaccatcg ccctagccgt caccggtttc ttgacctccg gagctttcgg catcaccgcc 300

ctgtcctcga tttcgtggtt gctgaactac gttaggcgaa tgcgggggag cttgccagag 360

cagctggatc atgcacggcg gcgcgtgcag gagacggtgg gccagaagac aagggaggcg 420

gggcagagaa gccaagatgt aataagaccg tgaggttttt ggatattaga tgttggttaa 480

tttgtgtgtt taatgtatat atgaggggtt gaataagtta ataaaattgc ggatttggta 540

caaaaaaaaa aaaaaaaaa 559

<210> 37
<211> 144
<212> PRT
<213> S. indicum

<400> 37

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

Tyr Leu Pro Glu Lys Gly Pro Ser Thr Ser Gln Val Leu Ala Val Val
20 25 30

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

Leu Thr Gly Thr Ile Ile Gly Leu Ala Val Ala Thr Pro Leu Phe Val
50 55 60

Ile Phe Ser Pro Ile Leu Val Pro Ala Ala Leu Thr Ile Ala Leu Ala
65 70 75 80

Val Thr Gly Phe Leu Thr Ser Gly Ala Phe Gly Ile Thr Ala Leu Ser
85 90 95

Ser Ile Ser Trp Leu Leu Asn Tyr Val Arg Arg Met Arg Gly Ser Leu
100 105 110

4197737_1

Pro Glu Gln Leu Asp His Ala Arg Arg Arg Val Gln Glu Thr Val Gly
115 120 125

Gln Lys Thr Arg Glu Ala Gly Gln Arg Ser Gln Asp Val Ile Arg Pro
130 135 140

<210> 38
<211> 1800
<212> DNA
<213> A. thaliana

<400> 38
ccatggctat acccaacctc ggtcttggtc acaccaggaa ctctctggta agctagctcc 60

actccccaga aacaaccggc gccaaattgc cggaattgct gacctgaaga cggaacatca 120

tcgtcgggtc cttgggcgat tgcggcggaa gatgggtcag cttgggcttg aggacgagac 180

ccgaatcgag tctgttgaaa ggttgttcat tgggatttgt atacggagat tggtcgtcga 240

gaggtttgag ggaaaggaca aatgggtttg gctctggaga aagagagtgc ggctttagag 300

agagaattga gaggtttaga gagagatgcg gcggcgatga cgggaggaga gacgacgagg 360

acctgcatta tcaaagcagt gacgtggtga aatttggaac ttttaagagg cagatagatt 420

tattatttgt atccattttc ttcattgttc tagaatgtcg cggaacaaat tttaaaacta 480

aatcctaaat ttttctaatt ttgttgccaa tagtggatat gtgggccgta tagaaggaat 540

ctattgaagg cccaaaccca tactgacgag cccaaaggtt cgttttgcgt tttatgtttc 600

ggttcgatgc caacgccaca ttctgagcta ggcaaaaaac aaacgtgtct ttgaatagac 660

tcctctcgtt aacacatgca gcggctgcat ggtgacgcca ttaacacgtg gcctacaatt 720

gcatgatgtc tccattgaca cgtgacttct cgtctccttt cttaatatat ctaacaaaca 780

ctcctacctc ttccaaaata tatacacatc tttttgatca atctctcatt caaaatctca 840

ttctctctag taaacaagaa caaaaaaatg gcggatacag ctagaggaac ccatcacgat 900

atcatcggca gagaccagta cccgatgatg ggccgagacc gagaccagta ccagatgtcc 960

ggacgaggat ctgactactc caagtctagg cagattgcta aagctgcaac tgctgtcaca 1020

gctggtggtt ccctccttgt tctctccagc cttacccttg ttggaactgt catagctttg 1080

actgttgcaa cacctctgct cgttatcttc agcccaatcc ttgtcccggc tctcatcaca 1140

gttgcactcc tcatcaccgg ttttctttcc tctggagggt ttggcattgc cgctataacc 1200

gttttctctt ggatttacaa gtaagcacac atttatcatc ttacttcata attttgtgca 1260

atatgtgcat gcatgtgttg agccagtagc tttggatcaa ttttttggt cgaataacaa 1320

atgtaacaat aagaaattgc aaattctagg gaacatttgg ttaactaaat acgaaatttg 1380

acctagctag cttgaatgtg tctgtgtata tcatctatat aggtaaaatg cttggtatga 1440

tacctattga ttgtgaatag gtacgcaacg ggagagcacc cacagggatc agacaagttg 1500

gacagtgcaa ggatgaagtt gggaagcaaa gctcaggatc tgaaagacag agctcagtac 1560

tacggacagc aacatactgg tggggaacat gaccgtgacc gtactcgtgg tggccagcac 1620

4197737_1

actacttaag ttaccccact gatgtcatcg tcatagtcca ataactccaa tgtcggggag 1680

ttagtttatg aggaataaag tgtttagaat ttgatcaggg ggagataata aaagccgagt 1740

ttgaatcttt ttgttataag taatgtttat gtgtgtttct atatgttgtc aaatggtacc 1800

<210> 39
<211> 173
<212> PRT
<213> A. thaliana

<400> 39

Met Ala Asp Thr Ala Arg Gly Thr His His Asp Ile Ile Gly Arg Asp
1 5 10 15

Gln Tyr Pro Met Met Gly Arg Asp Arg Asp Gln Tyr Gln Met Ser Gly
20 25 30

Arg Gly Ser Asp Tyr Ser Lys Ser Arg Gln Ile Ala Lys Ala Ala Thr
35 40 45

Ala Val Thr Ala Gly Gly Ser Leu Leu Val Leu Ser Ser Leu Thr Leu
50 55 60

Val Gly Thr Val Ile Ala Leu Thr Val Ala Thr Pro Leu Leu Val Ile
65 70 75 80

Phe Ser Pro Ile Leu Val Pro Ala Leu Ile Thr Val Ala Leu Leu Ile
85 90 95

Thr Gly Phe Leu Ser Ser Gly Gly Phe Gly Ile Ala Ala Ile Thr Val
100 105 110

Phe Ser Trp Ile Tyr Lys Tyr Ala Thr Gly Glu His Pro Gln Gly Ser
115 120 125

Asp Lys Leu Asp Ser Ala Arg Met Lys Leu Gly Ser Lys Ala Gln Asp
130 135 140

Leu Lys Asp Arg Ala Gln Tyr Tyr Gly Gln Gln His Thr Gly Gly Glu
145 150 155 160

His Asp Arg Asp Arg Thr Arg Gly Gly Gln His Thr Thr
165 170

<210> 40
<211> 450
<212> DNA
<213> A. thaliana

<400> 40
atggcggatc atcaacaaca tcagcaacaa caacaaccaa taatgaggag tctccatgaa 60

tcatcaccat cgactcggca gatagtgaga ttcgtaacgg cagctacgat cggcctatca 120

ctcctcgtgc tctcaggact aacactaacc ggaacggtga tcggtttgat cgtagcgacg 180

ccgttgatgg ttctgttcag cccggtgttg gtaccggcag tgataacgat agggcttctg 240

acgatgggat tcctattctc cggtggttgt ggggtggcag cagctacggc gttaacgtgg 300

atttataagt acgttaccgg aaaacacccg atgggagcgg ataaggtgga ttacgcgagg 360

atgaggatag cggagaaagc caaagagttg ggacattata cgcactcgca gccacaacaa 420

acacaccaaa ccacaacaac tactcattag 450

<210> 41
<211> 149
<212> PRT
<213> A. thaliana

<400> 41

Met Ala Asp His Gln Gln His Gln Gln Gln Gln Gln Pro Ile Met Arg
1 5 10 15

Ser Leu His Glu Ser Ser Pro Ser Thr Arg Gln Ile Val Arg Phe Val
20 25 30

Thr Ala Ala Thr Ile Gly Leu Ser Leu Leu Val Leu Ser Gly Leu Thr
35 40 45

Leu Thr Gly Thr Val Ile Gly Leu Ile Val Ala Thr Pro Leu Met Val
50 55 60

Leu Phe Ser Pro Val Leu Val Pro Ala Val Ile Thr Ile Gly Leu Leu
65 70 75 80

Thr Met Gly Phe Leu Phe Ser Gly Gly Cys Gly Val Ala Ala Ala Thr
85 90 95

Ala Leu Thr Trp Ile Tyr Lys Tyr Val Thr Gly Lys His Pro Met Gly
100 105 110

Ala Asp Lys Val Asp Tyr Ala Arg Met Arg Ile Ala Glu Lys Ala Lys
115 120 125

Glu Leu Gly His Tyr Thr His Ser Gln Pro Gln Gln Thr His Gln Thr
130 135 140

Thr Thr Thr Thr His
145

<210> 42
<211> 625
<212> DNA
<213> H. annuus

<400> 42
actaccacca cttacgaccg tcatttcacc accacccaac cccactaccg ccaagacgac 60

aggtcccgct acgaccagca aacccattcc cagtccacca gcaggacact cgccatcatc 120

4197737_1

gccctacttc ctgtcggcgg aatcttactc ggcttagccg ctctcacatt catcgggacg 180

cttatcgggc tcgccctcgc caccccgctt ttcgtcatat tcagcccgat catcgtgccg 240

gccgttctaa caatcgggct tgctgttaca ggctttttgg cgtcggggac gttcgggttg 300

acgggtttga gctcattgtc gtatttgttc aatatggtta ggcagacggc tgggtcggtg 360

cccgagtcct tggattatgt taaggggacg ttgcaggatg ccggtgagta tgccgggcag 420

aagacgaagg atttcgggca gaagattcag agcacggctc atgagatggg tgatcagggg 480

caggttggtg ttcatgctca agttggtggc gggaaagaag ggcgaaaaag cggtgatcgg 540

acttgaggat tcaaggttga tattgtggaa taataatgtt gatgtaagtt tttagtgtta 600

tcaaagcttt gtttgtttgt ttgta 625

<210> 43
<211> 181
<212> PRT
<213> H. annuus

<400> 43

Thr Thr Thr Thr Tyr Asp Arg His Phe Thr Thr Thr Gln Pro His Tyr
1 5 10 15

Arg Gln Asp Asp Arg Ser Arg Tyr Asp Gln Gln Thr His Ser Gln Ser
20 25 30

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

Leu Leu Gly Leu Ala Ala Leu Thr Phe Ile Gly Thr Leu Ile Gly Leu
50 55 60

Ala Leu Ala Thr Pro Leu Phe Val Ile Phe Ser Pro Ile Ile Val Pro
65 70 75 80

Ala Val Leu Thr Ile Gly Leu Ala Val Thr Gly Phe Leu Ala Ser Gly
85 90 95

Thr Phe Gly Leu Thr Gly Leu Ser Ser Leu Ser Tyr Leu Phe Asn Met
100 105 110

Val Arg Gln Thr Ala Gly Ser Val Pro Glu Ser Leu Asp Tyr Val Lys
115 120 125

Gly Thr Leu Gln Asp Ala Gly Glu Tyr Ala Gly Gln Lys Thr Lys Asp
130 135 140

Phe Gly Gln Lys Ile Gln Ser Thr Ala His Glu Met Gly Asp Gln Gly
145 150 155 160

Gln Val Gly Val His Ala Gln Val Gly Gly Gly Lys Glu Gly Arg Lys
165 170 175

Ser Gly Asp Arg Thr
180

<210> 44
<211> 737
<212> DNA
<213> B. napus

<400> 44
gaattccatt caaactagta aacaatgggg atactcagga agaaaaaaca cgagcgaaag 60

ccgtcgttta agagtgtttt aaccgcaata ttagctacac atgccgcaac attcctcttg 120

ttgatcgccg gtgtatccct cgccggcaca gccgccgcat ttatcgctac catgccacta 180

ttcgtagtat tcagtccgat tctcgtacca gctggtatta ccactggttt actgactacg 240

ggtttagcag ccgccggtgg cgccggcgcg actgctgtca ccatcatcct gtggctctac 300

aagcgagcaa cgggcaaggc gccgccaaaa gtcctagaaa aagtcttgaa aaagataata 360

ccaggtgctg cagctgcacc agcagccgct ccaggagccg ctccagcagc ggcgccagca 420

gccgcaccag ctgtggcgcc agcagccgca ccagctgctg cgccagcacc taagccagca 480

gccccaccag cacctaagcc agcagccgca ccgagtatat gaaaagaagt ggtgggcatg 540

agtaaaggtt gatatggaaa actggataca tagaaaaaag agtaatccaa cttttaaaaa 600

ataaataaca acttcacgtg gggatagaaa aattttcaaa tattatttta ctaatggatg 660

tcgcggtaca aaataataac aaatgtaagc ctttttattg tatagtattt taagaacgaa 720

gctatgtagc gttgaca 737

<210> 45
<211> 165
<212> PRT
<213> B. napus

<400> 45

Met Gly Ile Leu Arg Lys Lys Lys His Glu Arg Lys Pro Ser Phe Lys
1 5 10 15

Ser Val Leu Thr Ala Ile Leu Ala Thr His Ala Ala Thr Phe Leu Leu
20 25 30

Leu Ile Ala Gly Val Ser Leu Ala Gly Thr Ala Ala Ala Phe Ile Ala
35 40 45

Thr Met Pro Leu Phe Val Val Phe Ser Pro Ile Leu Val Pro Ala Gly
50 55 60

Ile Thr Thr Gly Leu Leu Thr Thr Gly Leu Ala Ala Ala Gly Gly Ala
65 70 75 80

Gly Ala Thr Ala Val Thr Ile Ile Leu Trp Leu Tyr Lys Arg Ala Thr
85 90 95

4197737_1

Gly Lys Ala Pro Pro Lys Val Leu Glu Lys Val Leu Lys Lys Ile Ile
100 105 110

Pro Gly Ala Ala Ala Ala Pro Ala Ala Ala Pro Gly Ala Ala Pro Ala
115 120 125

Ala Ala Pro Ala Ala Ala Pro Ala Val Ala Pro Ala Ala Ala Pro Ala
130 135 140

Ala Ala Pro Ala Pro Lys Pro Ala Ala Pro Pro Ala Pro Lys Pro Ala
145 150 155 160

Ala Ala Pro Ser Ile
165

<210> 46
<211> 1153
<212> DNA
<213> Z. mays

<400> 46
aactactccg tcactttgtt tgcaaagctc ctcctcgatc catcgatcac tgcaccggcc 60

ggcggcaccg cgctcgcagg ggctagccaa cgagacggca gcaatggcgg accgtgaccg 120

cagcggcatc tacggcggcg cccacgccac ctacgggcag cagcagcagc agggaggagg 180

cgggcgcccg atgggtgagc aggtgaaggg catgctccac gacaaggggc cgacggcgtc 240

gcaggcgctg acggtggcga cgctgttccc gctgggcggg ctgctgctgg tgctgtcggg 300

gctggcgctg acggcctccg tggtggggct ggccgtggcc acgccggtgt tcctgatctt 360

cagccccgtg ctggtccccg ccgcgctgct catcgggacg gccgtcatgg ggttcctcac 420

gtcgggcgcg ctggggctcg ggggcctgtc ctcgctcacg tgcctcgcca acacggcgcg 480

gcaggcgttc cagcgcaccc cggactacgt ggaggaggcg caccgcagga tggcggaggc 540

cgcggcgcac gcgggccaca agaccgcgca ggcaggccag gccatccagg gcagggcgca 600

ggaggccggc gccgggggag gtgcaggtgc cggcgctggc ggcggcggca gggcttcctc 660

gtaagcaagt catccatgca tggattatgg atagatgcgc gcgtgcgtgt ctatcagtat 720

cagcagccag cagggtcgtc gcggaatgct gtgttcctgt acgtgtgggt gaccgtcctt 780

ccgtccttcg tctttctccc cccgagtgtg tgttacgtat gtcctggtgt tcgtcgtgtg 840

tgttcatcgc cgctccagtt gaattccggt gtctgttcat cgccgctcca ggtcgtagat 900

gtgaatatac tttgctaggg gaataagtga taagtctgtc tggaaggtaa tgtttgagct 960

ttgctagtgt ggctgggcac tctggtcact ggttgtgttg tgcatgcatc agctgtatga 1020

tcgtcgtctg ttgtggaaaa ttggtcaatg tattctcttg ctgaataatt tgtgacatct 1080

aattgttatg tatcgtctct ttgctgaata atcagtttct gatttatctt gcattaaaaa 1140

aaaaaaaaaa aaa 1153

<210> 47
<211> 186

4197737_1

<212> PRT
<213> Z. mays

<400> 47

Met Ala Asp Arg Asp Arg Ser Gly Ile Tyr Gly Gly Ala His Ala Thr
1 5 10 15

Tyr Gly Gln Gln Gln Gln Gly Gly Gly Gly Arg Pro Met Gly Glu
20 25 30

Gln Val Lys Gly Met Leu His Asp Lys Gly Pro Thr Ala Ser Gln Ala
35 40 45

Leu Thr Val Ala Thr Leu Phe Pro Leu Gly Gly Leu Leu Leu Val Leu
50 55 60

Ser Gly Leu Ala Leu Thr Ala Ser Val Val Gly Leu Ala Val Ala Thr
65 70 75 80

Pro Val Phe Leu Ile Phe Ser Pro Val Leu Val Pro Ala Ala Leu Leu
85 90 95

Ile Gly Thr Ala Val Met Gly Phe Leu Thr Ser Gly Ala Leu Gly Leu
100 105 110

Gly Gly Leu Ser Ser Leu Thr Cys Leu Ala Asn Thr Ala Arg Gln Ala
115 120 125

Phe Gln Arg Thr Pro Asp Tyr Val Glu Glu Ala His Arg Arg Met Ala
130 135 140

Glu Ala Ala Ala His Ala Gly His Lys Thr Ala Gln Ala Gly Gln Ala
145 150 155 160

Ile Gln Gly Arg Ala Gln Glu Ala Gly Ala Gly Gly Gly Ala Gly Ala
165 170 175

Gly Ala Gly Gly Gly Gly Arg Ala Ser Ser
180 185

<210> 48
<211> 447
<212> DNA
<213> O. sativa

<400> 48
atgggtgatc agcacagagg agtgatcggc ggtggcggct acggtgaccg tggtggccag 60

gagcagcagg agaagcagcc cttcatgatg acggctctga agacggtgac cgcggcgacg 120

gccgggggct cgattctggt gctgtccggg ctgatcctgg ccgggaccgt catcgcgctc 180

acggtggcca ccccggtgtt ggtcatcttc agccccgtgc tggtaccggc ggccatcgcg 240

ctggcgctca tggcggccgg gttcgtcacc tcggttggtc tcggtgtagc cgcgctctcc 300

gttttctcgt ggatgtacaa gtacctcacc gggaagcacc cgccgggcgc cgaccatctg 360

gaccacacca aggcgagggt cgcgtccaag ctccgcgaca tcaaggaggc ggcgcatcac 420

ctcatcgacc aggcgcaggc gtcttag 447

<210> 49
<211> 148
<212> PRT
<213> O. sativa

<400> 49

Met Gly Asp Gln His Arg Gly Val Ile Gly Gly Gly Gly Tyr Gly Asp
1 5 10 15

Arg Gly Gly Gln Glu Gln Gln Glu Lys Gln Pro Phe Met Met Thr Ala
20 25 30

Leu Lys Thr Val Thr Ala Ala Thr Ala Gly Gly Ser Ile Leu Val Leu
35 40 45

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

Pro Val Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Ala Ile Ala
65 70 75 80

Leu Ala Leu Met Ala Ala Gly Phe Val Thr Ser Val Gly Leu Gly Val
85 90 95

Ala Ala Leu Ser Val Phe Ser Trp Met Tyr Lys Tyr Leu Thr Gly Lys
100 105 110

His Pro Pro Gly Ala Asp His Leu Asp His Thr Lys Ala Arg Val Ala
115 120 125

Ser Lys Leu Arg Asp Ile Lys Glu Ala Ala His His Leu Ile Asp Gln
130 135 140

Ala Gln Ala Ser
145

<210> 50
<211> 709
<212> DNA
<213> B. oleracea

<400> 50
gccgtttttt cagaatgttc tcttttatct tcccattgct gaacgttata aagcttatta 60

tagcttccgt gacctcctta gtctgcttag cgttttcttg tgtgacactc ggtggttcag 120

ccgtggcatt aatcgtatcc acaccacttt tcatcatatt tagtccaatt ctcgtacctg 180

ccactattgc cactaccctc ctagccagtg ggctcatggc gggtaccacc ctcggactga 240

ccggcatagg tctcatcacg gggctcgtta ggacggcagg aggagttaca ttggccgaat 300

4197737_1

caccgataag aagaattata ataaatagaa ttaaagcaag acttgggggt ggcggcggtt 360

cacgtctggc aatgctcaaa aaaattctgg gactcattaa aaagttgcgt ggtatgtctt 420

caggtggagc agcacctgcg ctgaagcagc accagcagct gcgcccgcgg atggagctgc 480

acccgcggca cctgcaccga cctaacaaag aacgttggtt catgctgttc caatatgtag 540

cacataaaaa ttgtgtaata attaacttaa gaatttatga ttcggaaact aaaaagaaaa 600

tagccctttt actatctttt atacaatata gttttctatg taataatgtt taatttgctt 660

ataactataa aagactcatg catagttgat taggaaaaaa aaaaaaaaa 709

<210> 51
<211> 216
<212> PRT
<213> B. oleracea

<400> 51

Arg Phe Phe Arg Met Phe Ser Phe Ile Phe Pro Leu Leu Asn Val Ile
1 5 10 15

Lys Leu Ile Ile Ala Ser Val Thr Ser Leu Val Cys Leu Ala Phe Ser
20 25 30

Cys Val Thr Leu Gly Gly Ser Ala Val Ala Leu Ile Val Ser Thr Pro
35 40 45

Leu Phe Ile Ile Phe Ser Pro Ile Leu Val Pro Ala Thr Ile Ala Thr
50 55 60

Thr Leu Leu Ala Ser Gly Leu Met Ala Gly Thr Thr Leu Gly Leu Thr
65 70 75 80

Gly Ile Gly Leu Ile Thr Gly Leu Val Arg Thr Ala Gly Gly Val Thr
85 90 95

Leu Ala Glu Ser Pro Ile Arg Arg Ile Ile Ile Asn Arg Ile Lys Ala
100 105 110

Arg Leu Gly Gly Gly Gly Gly Ser Arg Leu Ala Met Leu Lys Lys Ile
115 120 125

Leu Gly Leu Ile Lys Lys Leu Arg Gly Met Ser Ser Gly Gly Ala Ala
130 135 140

Pro Ala Leu Lys Gln His Gln Gln Leu Arg Pro Arg Met Glu Leu His
145 150 155 160

Pro Arg His Leu His Arg Pro Asn Lys Glu Arg Trp Phe Met Leu Phe
165 170 175

Gln Tyr Val Ala His Lys Asn Cys Val Ile Ile Asn Leu Arg Ile Tyr
180 185 190

Asp Ser Glu Thr Lys Lys Lys Ile Ala Leu Leu Leu Ser Phe Ile Gln
195 200 205

Tyr Ser Phe Leu Cys Asn Asn Val
210 215

<210> 52
<211> 447
<212> DNA
<213> C. arabica

<400> 52
atggctgagc actaccagct gcagcaacgc cccacagagg ccgtcaaaag cttccttcct 60

cagaagggtc catcaacttc acatgtgtta gcagttgtca cgctcctccc agttgcggga 120

gtcctgctag gcctttccgg gctgattctc gtcggaacgg tcatcggtct ggcggtgaca 180

accccgcttt tcgttatctt tagccccatt ttggtcccag ctgtatttgc cctagggctg 240

gccctggccg ggttcttgac ctccggtgct ttcgggatca ctgcacttgc ttcattgtcg 300

tggatgctga actacatccg actcatgaag gcgtcttccc aggagcaaat ggacctcgca 360

aagtggcgcg tgcaggacac tgccggccaa gttggtcaga aagcgagaga cgtgggccag 420

agaactcaag atgtagccag agcatga 447

<210> 53
<211> 148
<212> PRT
<213> C. arabica

<400> 53

Met Ala Glu His Tyr Gln Leu Gln Gln Arg Pro Thr Glu Ala Val Lys
1 5 10 15

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

Val Thr Leu Leu Pro Val Ala Gly Val Leu Leu Gly Leu Ser Gly Leu
35 40 45

Ile Leu Val Gly Thr Val Ile Gly Leu Ala Val Thr Thr Pro Leu Phe
50 55 60

Val Ile Phe Ser Pro Ile Leu Val Pro Ala Val Phe Ala Leu Gly Leu
65 70 75 80

Ala Leu Ala Gly Phe Leu Thr Ser Gly Ala Phe Gly Ile Thr Ala Leu
85 90 95

Ala Ser Leu Ser Trp Met Leu Asn Tyr Ile Arg Leu Met Lys Ala Ser
100 105 110

Ser Gln Glu Gln Met Asp Leu Ala Lys Trp Arg Val Gln Asp Thr Ala

115 120 125

Gly Gln Val Gly Gln Lys Ala Arg Asp Val Gly Gln Arg Thr Gln Asp
130 135 140

Val Ala Arg Ala
145

<210> 54
<211> 348
<212> PRT
<213> S. indicum

<400> 54

Met Asp Leu Ile His Thr Phe Leu Asn Leu Ile Ala Pro Pro Phe Thr
1 5 10 15

Phe Phe Phe Leu Leu Phe Phe Leu Pro Pro Phe Gln Ile Phe Lys Phe
20 25 30

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

Val Val Val Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ser Leu Ala
50 55 60

Tyr Glu Tyr Ala Lys Arg Gly Ala Cys Leu Val Leu Ala Ala Arg Arg
65 70 75 80

Glu Arg Ser Leu Gln Glu Val Ala Glu Arg Ala Arg Asp Leu Gly Ser
85 90 95

Pro Asp Val Val Val Val Arg Ala Asp Val Ser Lys Ala Glu Asp Cys
100 105 110

Arg Lys Val Val Asp Gln Thr Met Asn Arg Phe Gly Arg Leu Asp His
115 120 125

Leu Val Asn Asn Ala Gly Ile Met Ser Val Ser Met Leu Glu Glu Val
130 135 140

Glu Asp Ile Thr Gly Tyr Arg Glu Thr Met Asp Ile Asn Phe Trp Gly
145 150 155 160

Tyr Val Tyr Met Thr Arg Phe Ala Ala Pro Tyr Leu Arg Asn Ser Arg
165 170 175

Gly Arg Ile Val Val Leu Ser Ser Ser Ser Ser Trp Met Pro Thr Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Ile Ser Gln Phe Phe
195 200 205

```
Glu Thr Leu Arg Val Glu Phe Gly Pro Asp Ile Gly Ile Thr Leu Val
    210                 215                 220

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

Ala Gly Glu Arg Val Ile Asp Gln Asp Met Arg Asp Val Gln Val Ser
                245                 250                 255

Thr Thr Pro Ile Leu Arg Val Glu Ser Ala Ala Arg Ser Ile Val Arg
            260                 265                 270

Ser Ala Ile Arg Gly Glu Arg Tyr Val Thr Glu Pro Ala Trp Phe Arg
        275                 280                 285

Val Thr Tyr Trp Trp Lys Leu Phe Cys Pro Glu Val Met Glu Trp Val
    290                 295                 300

Phe Arg Leu Met Tyr Leu Ala Ser Pro Gly Glu Pro Glu Lys Glu Thr
305                 310                 315                 320

Phe Gly Lys Lys Val Leu Asp Tyr Thr Gly Val Lys Ser Leu Leu Tyr
                325                 330                 335

Pro Glu Thr Val Gln Val Pro Glu Pro Lys Asn Asp
            340                 345
```

<210> 55
<211> 348
<212> PRT
<213> S. indicum

<400> 55

```
Met Asp Leu Ile His Thr Phe Leu Asn Leu Ile Ala Pro Pro Phe Thr
1               5                   10                  15

Phe Phe Phe Leu Leu Phe Phe Leu Pro Pro Phe Gln Ile Phe Lys Phe
            20                  25                  30

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

Val Val Val Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ser Leu Ala
    50                  55                  60

Tyr Glu Tyr Ala Lys Arg Gly Ala Cys Leu Val Leu Ala Ala Arg Arg
65                  70                  75                  80

Glu Arg Ser Leu Gln Glu Val Ala Glu Arg Ala Arg Asp Leu Gly Ser
                85                  90                  95

Pro Asp Val Val Val Val Arg Ala Asp Val Ser Lys Ala Glu Asp Cys
```

4197737_1

100 105 110

Arg Lys Val Val Asp Gln Thr Met Asn Arg Phe Gly Arg Leu Asp His
115 120 125

Leu Val Asn Asn Ala Gly Ile Met Ser Val Ser Met Leu Glu Glu Val
130 135 140

Glu Asp Ile Thr Gly Tyr Arg Glu Thr Met Asp Ile Asn Phe Trp Gly
145 150 155 160

Tyr Val Tyr Met Thr Arg Phe Ala Ala Pro Tyr Leu Arg Asn Ser Arg
165 170 175

Gly Arg Ile Val Val Leu Ser Ser Ser Ser Ser Trp Met Pro Thr Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Ile Ser Gln Phe Phe
195 200 205

Glu Thr Leu Arg Val Glu Phe Gly Pro Asp Ile Gly Ile Thr Leu Val
210 215 220

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

Ala Gly Glu Arg Val Ile Asp Gln Asp Met Arg Asp Val Gln Val Ser
245 250 255

Thr Thr Pro Ile Leu Arg Val Glu Ser Ala Ala Arg Ser Ile Val Arg
260 265 270

Ser Ala Ile Arg Gly Glu Arg Tyr Val Thr Glu Pro Ala Trp Phe Arg
275 280 285

Val Thr Tyr Trp Trp Lys Leu Phe Cys Pro Glu Val Met Glu Trp Val
290 295 300

Phe Arg Leu Met Tyr Leu Ala Ser Pro Gly Glu Pro Glu Lys Glu Thr
305 310 315 320

Phe Gly Lys Lys Val Leu Asp Tyr Thr Gly Val Lys Ser Leu Leu Tyr
325 330 335

Pro Glu Thr Val Gln Val Pro Glu Pro Lys Asn Asp
340 345

<210> 56
<211> 1266
<212> DNA
<213> B. napus

<400> 56

4197737_1

ggtcgacggc cattacggcc gggatcccgt taatgtgtat tgtgttttga tcgaagatgg 60

aattgataaa cgatttcctc aatctaactg cacctttctt tactttcttt ggtctttgct 120

tcttcttgcc ccctttttat ttcttcaagt tcgtgcagtc tatcttctcg acaattttct 180

ctgaaaacgt atatgggaaa gtggttctca tcactggtgc ttcctccggg atcggcgagc 240

aattggcata tgagtacgca agtaaaggtg catgtttggc tctgactgcc cgaaggaaga 300

accgtctaga ggaagtggca gagattgctc gtgaagttgg atcccctaat gttgtcacag 360

ttcacgctga tgtttccaaa cctgatgatt gtagacgaat cgtcgatgag accatctccc 420

attttggcag attggatcat cttgtaaaca atgctggaat aatgcaaatt tcaatgttcg 480

aaaacattga agaaataact aggacaagag cagttatgga tactaacttt tggggagcgg 540

tttatacaac tcgtgctgcg cttccgtacc ttcgacaaag caatggtaag attgtggcta 600

tgtcgtcctc tgcggcatgg ctaaccgccc caaggatgag cttttataat gctagcaaag 660

cagctttgtt gaacttcttc gagacgttga ggattgagct tggtagcgat gtacacatta 720

caatcgtcac acctggttat attgaatctg aactcacaca aggcaagtac gtctctggtg 780

aaggcgagct agtagtcaac caagacatta gagatgttca aattggagca tttccggtaa 840

cgtcagtatc aggtcgtgcc aaggggatag tgaaaggtgt gtgtaggaaa gagagatacg 900

tgaccgaacc atcgtggttt aaggtgacgt acctttggaa agtgttttgt ccggaactga 960

tcgagtgggg ttgcagattg atgttcttgt ccggacatgg tacgccggag gaaaatgcac 1020

tcaacaagaa gatcctggac atacctggtg tacgtagtgc tctataccct gaacctatca 1080

gaacgccaga aatcaagtcg gagtagagtg aggttgatac ttaataagtg tctcataaag 1140

tggagccatg ttttgtaaat ggactttcta ttatgcacat gttactatga tgtatctgtt 1200

tgtttatgtg tataagaata agtgaacttt ggagctcaaa aaaaaaaaaa aaaaaaaaaa 1260

aaaaaa 1266

<210> 57
<211> 349
<212> PRT
<213> B. napus

<400> 57

Met Glu Leu Ile Asn Asp Phe Leu Asn Leu Thr Ala Pro Phe Phe Thr
1 5 10 15

Phe Phe Gly Leu Cys Phe Phe Leu Pro Pro Phe Tyr Phe Phe Lys Phe
20 25 30

Val Gln Ser Ile Phe Ser Thr Ile Phe Ser Glu Asn Val Tyr Gly Lys
35 40 45

Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Gln Leu Ala
50 55 60

Tyr Glu Tyr Ala Ser Lys Gly Ala Cys Leu Ala Leu Thr Ala Arg Arg

65 70 75 80

Lys Asn Arg Leu Glu Glu Val Ala Glu Ile Ala Arg Glu Val Gly Ser
85 90 95

Pro Asn Val Val Thr Val His Ala Asp Val Ser Lys Pro Asp Asp Cys
100 105 110

Arg Arg Ile Val Asp Glu Thr Ile Ser His Phe Gly Arg Leu Asp His
115 120 125

Leu Val Asn Asn Ala Gly Ile Met Gln Ile Ser Met Phe Glu Asn Ile
130 135 140

Glu Glu Ile Thr Arg Thr Arg Ala Val Met Asp Thr Asn Phe Trp Gly
145 150 155 160

Ala Val Tyr Thr Thr Arg Ala Ala Leu Pro Tyr Leu Arg Gln Ser Asn
165 170 175

Gly Lys Ile Val Ala Met Ser Ser Ser Ala Ala Trp Leu Thr Ala Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Leu Leu Asn Phe Phe
195 200 205

Glu Thr Leu Arg Ile Glu Leu Gly Ser Asp Val His Ile Thr Ile Val
210 215 220

Thr Pro Gly Tyr Ile Glu Ser Glu Leu Thr Gln Gly Lys Tyr Val Ser
225 230 235 240

Gly Glu Gly Glu Leu Val Val Asn Gln Asp Ile Arg Asp Val Gln Ile
245 250 255

Gly Ala Phe Pro Val Thr Ser Val Ser Gly Arg Ala Lys Gly Ile Val
260 265 270

Lys Gly Val Cys Arg Lys Glu Arg Tyr Val Thr Glu Pro Ser Trp Phe
275 280 285

Lys Val Thr Tyr Leu Trp Lys Val Phe Cys Pro Glu Leu Ile Glu Trp
290 295 300

Gly Cys Arg Leu Met Phe Leu Ser Gly His Gly Thr Pro Glu Glu Asn
305 310 315 320

Ala Leu Asn Lys Lys Ile Leu Asp Ile Pro Gly Val Arg Ser Ala Leu
325 330 335

Tyr Pro Glu Pro Ile Arg Thr Pro Glu Ile Lys Ser Glu
340 345

<210> 58
<211> 1188
<212> DNA
<213> Z. mays

<400> 58
attgcgcaca ccttgatgct gggcatgagc aggacggggt tggccggagc ggcgctgcgg 60

gtggcgctga cggcgctgct gccgctggtc ctgccggcgt actacgtgta caagctgacc 120

acgtacctcc tcggcgccgt cttccccgag gacgtcgccg gcaaggtggt actcatcacc 180

ggcgcctcct ccggcatcgg cgagcacctg gcctatgagt acgcgaagcg gggagcctac 240

ctggcgctgg tggcgaggag ggaggcgagc ctccgcgagg tcggcgacgt cgcgctgggg 300

ctgggctcgc cgggcgtcct cgtcctcccg gctgacgtct ccaagcctcg ggactgcgag 360

ggcttcatcg acgacacgat tagctacttc ggtagactgg atcacctggt gaacaacgcg 420

tccatctggc aagtgtgcaa gttcgaagag atccaggacg tcaggcactt gagagccctg 480

atggacatca acttctgggg ccacgtgtac ccaacccggc tcgccatccc tcacctcagg 540

agaagccgtg gccggatcgt gggcgtcacc tccaactcgt cctacatatt catcgggagg 600

aacaccttct acaatgccag caaggcggcg gcgctcagct tctacgacac cctgaggatg 660

gagctgggca gcgacatccg catcaccgag gtggtgccag gcgtggtgga gtctgagatc 720

accaagggga agatgctcac caagggaggc gagatgaagg tggaccagga cgaaagagac 780

gccatcctgg ggccgacgcc ggccgagccc gtgggcgact tcgccaggac cgtggtgcgc 840

gacgtgtgcc ggggcgcgag gtacgtgttc gagcccaggt ggtacatggg cgtctacttg 900

ctgcgggcct gcctcccgga agtcctggcc tggaactccc gcctgctcac tgtcgacacg 960

gtcggcgcgt ccaccacgga cacgctcgga aagtggctgg tcgagctgcc cggcgtgcgc 1020

cgcgtcgtgc agccgccgtc gctccgctcg ccggagatca aggactagtg acggtgatcg 1080

tgtacgttct gtggccatgg atagcactag ctgtatgaga ccgaagttcc tttagacatc 1140

gacacgatta ataaaggact caattatctt aaaaaaaaaa aaaaaaaa 1188

<210> 59
<211> 350
<212> PRT
<213> Z. mays

<400> 59

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

Ala Leu Thr Ala Leu Leu Pro Leu Val Leu Pro Ala Tyr Tyr Val Tyr
20 25 30

Lys Leu Thr Thr Tyr Leu Leu Gly Ala Val Phe Pro Glu Asp Val Ala
35 40 45

Gly Lys Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu His

50 55 60

Leu Ala Tyr Glu Tyr Ala Lys Arg Gly Ala Tyr Leu Ala Leu Val Ala
65 70 75 80

Arg Arg Glu Ala Ser Leu Arg Glu Val Gly Asp Val Ala Leu Gly Leu
85 90 95

Gly Ser Pro Gly Val Leu Val Leu Pro Ala Asp Val Ser Lys Pro Arg
100 105 110

Asp Cys Glu Gly Phe Ile Asp Asp Thr Ile Ser Tyr Phe Gly Arg Leu
115 120 125

Asp His Leu Val Asn Asn Ala Ser Ile Trp Gln Val Cys Lys Phe Glu
130 135 140

Glu Ile Gln Asp Val Arg His Leu Arg Ala Leu Met Asp Ile Asn Phe
145 150 155 160

Trp Gly His Val Tyr Pro Thr Arg Leu Ala Ile Pro His Leu Arg Arg
165 170 175

Ser Arg Gly Arg Ile Val Gly Val Thr Ser Asn Ser Ser Tyr Ile Phe
180 185 190

Ile Gly Arg Asn Thr Phe Tyr Asn Ala Ser Lys Ala Ala Ala Leu Ser
195 200 205

Phe Tyr Asp Thr Leu Arg Met Glu Leu Gly Ser Asp Ile Arg Ile Thr
210 215 220

Glu Val Val Pro Gly Val Val Glu Ser Glu Ile Thr Lys Gly Lys Met
225 230 235 240

Leu Thr Lys Gly Gly Glu Met Lys Val Asp Gln Asp Glu Arg Asp Ala
245 250 255

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

Val Val Arg Asp Val Cys Arg Gly Ala Arg Tyr Val Phe Glu Pro Arg
275 280 285

Trp Tyr Met Gly Val Tyr Leu Leu Arg Ala Cys Leu Pro Glu Val Leu
290 295 300

Ala Trp Asn Ser Arg Leu Leu Thr Val Asp Thr Val Gly Ala Ser Thr
305 310 315 320

Thr Asp Thr Leu Gly Lys Trp Leu Val Glu Leu Pro Gly Val Arg Arg
325 330 335

Val Val Gln Pro Pro Ser Leu Arg Ser Pro Glu Ile Lys Asp
340 345 350

<210> 60
<211> 890
<212> DNA
<213> B. napus

<400> 60
gtttccaaac ctgatgattg tagacgaatc gtcgatgaga ccatctccca ttttggcaga 60

ttggatcatc ttgtaaacaa tgctggaata atgcagattt caatgttcga aaacattgaa 120

gaaataacta ggacaagagc agttatggat actaactttt ggggatcggt ttatacaact 180

cgtgctgcgc ttccgtacct tcgacaaagc aatggtaaga ttgtggccat gtcgtcctct 240

gcggcatggc taaccgcccc aaggatgagc ttttataatg ctagcaaagc cgcattgttg 300

aacttcttcg agacgttgag gattgagctt ggcagcgatg tacacattac aatcgtcaca 360

cctggttata ttgaatctga actcacacaa ggcaagtact tctctggtga aggcgagcta 420

gtagtcaacc aagacattag agatgttcaa attggagcat ttccggtaac gtcggtatca 480

ggttgtgcca aggggatagt gaaaggtgtg tgtaggaaac agagatacgt gaccgaacca 540

tcgtggttta aggtgacgta cctttggaaa gtgttttgtc cggaactgat cgagtggggt 600

tgcaggttgc tgttcttgtc cggacatggt acgtcggaga aaaatgcact caacaagaag 660

atcttggaca tacctggtgt acgtagtgct ctataccctg aatctatcag aacgccagaa 720

atcaagtcgg agtagagtga ggttgatact taataagtgt ctcatatagt ggagccatgt 780

tttgtaaatg gactttctat tatgcacatg ttactatgat gtatccgttt gtttatgtgt 840

ataagaataa gtgaactttg gagctccaaa aaaaaaaaaa aaaaaaaaaa 890

<210> 61
<211> 244
<212> PRT
<213> B. napus

<400> 61

Val Ser Lys Pro Asp Asp Cys Arg Arg Ile Val Asp Glu Thr Ile Ser
1 5 10 15

His Phe Gly Arg Leu Asp His Leu Val Asn Asn Ala Gly Ile Met Gln
20 25 30

Ile Ser Met Phe Glu Asn Ile Glu Glu Ile Thr Arg Thr Arg Ala Val
35 40 45

Met Asp Thr Asn Phe Trp Gly Ser Val Tyr Thr Thr Arg Ala Ala Leu
50 55 60

Pro Tyr Leu Arg Gln Ser Asn Gly Lys Ile Val Ala Met Ser Ser Ser
65 70 75 80

4197737_1

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

Ala Ala Leu Leu Asn Phe Phe Glu Thr Leu Arg Ile Glu Leu Gly Ser
100 105 110

Asp Val His Ile Thr Ile Val Thr Pro Gly Tyr Ile Glu Ser Glu Leu
115 120 125

Thr Gln Gly Lys Tyr Phe Ser Gly Glu Gly Glu Leu Val Val Asn Gln
130 135 140

Asp Ile Arg Asp Val Gln Ile Gly Ala Phe Pro Val Thr Ser Val Ser
145 150 155 160

Gly Cys Ala Lys Gly Ile Val Lys Gly Val Cys Arg Lys Gln Arg Tyr
165 170 175

Val Thr Glu Pro Ser Trp Phe Lys Val Thr Tyr Leu Trp Lys Val Phe
180 185 190

Cys Pro Glu Leu Ile Glu Trp Gly Cys Arg Leu Leu Phe Leu Ser Gly
195 200 205

His Gly Thr Ser Glu Lys Asn Ala Leu Asn Lys Lys Ile Leu Asp Ile
210 215 220

Pro Gly Val Arg Ser Ala Leu Tyr Pro Glu Ser Ile Arg Thr Pro Glu
225 230 235 240

Ile Lys Ser Glu

<210> 62
<211> 1030
<212> DNA
<213> S. indicum

<400> 62
ggcacgagag agaaaaaagg tgattttgtc aagggaaata tggcaactca tgttttggct 60

gctgcggcgg agagaaatgc tgcgttggcg ccggacgccc cgcttgctcc ggtgactatg 120

gagcgcccag tgcgcactga cttggagact tcgatcccga agccctatat ggcaagagga 180

ttggttgcac ctgatatgga tcaccccaac ggaacaccag gccatgtgca tgataatttg 240

agtgtgctgc aacagcattg tgctttcttt gatcaggatg ataacggaat catctatcca 300

tgggagactt actctggact tcgccaaatt ggtttcaatg tgatagcttc ccttataatg 360

gctatcgtca ttaatgtggc gctgagttat cctactctcc cgggttggat tccttctcct 420

tttttcccca tatatttgta caacatacac aaggccaaac atggaagcga ctccggaacc 480

tatgatactg aaggaaggta cctacctatg aattttgaga acctgttcag caagcatgcc 540

4197737_1

cggacaatgc ccgataggct cactctaggg gagctatgga gcatgactga agctaacaga 600

gaagcatttg acattttcgg ctggatcgca agcaaaatgg agtggactct cctctacatt 660

cttgcaagag accaggacgg tttcctgtcg aaagaagcca tcaggcggtg ttacgatggc 720

agtttgttcg agtactgtgc aaagatgcaa aggggagccg aggacaagat gaaatgaagg 780

aaatcggcta tcgcggtagg tgtaagttat gatgtggtgt gtatgatgga ttgaaagtgc 840

cagtgcttaa gttgtgtggc agagtcttgt gtaataacct ttgtgtacag atttaaggtc 900

tcggaattgg tgtaactgtg gagaagatgt tgactcctgt ttttgttcaa taagtccaac 960

tcttgacatt tggttggttt gcagggaaag atggggaatt ttgttttccg aaaaaaaaaa 1020

aaaaaaaaaa 1030

<210> 63
<211> 245
<212> PRT
<213> S. indicum

<400> 63

Met Ala Thr His Val Leu Ala Ala Ala Ala Glu Arg Asn Ala Ala Leu
1 5 10 15

Ala Pro Asp Ala Pro Leu Ala Pro Val Thr Met Glu Arg Pro Val Arg
20 25 30

Thr Asp Leu Glu Thr Ser Ile Pro Lys Pro Tyr Met Ala Arg Gly Leu
35 40 45

Val Ala Pro Asp Met Asp His Pro Asn Gly Thr Pro Gly His Val His
50 55 60

Asp Asn Leu Ser Val Leu Gln Gln His Cys Ala Phe Phe Asp Gln Asp
65 70 75 80

Asp Asn Gly Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Gln
85 90 95

Ile Gly Phe Asn Val Ile Ala Ser Leu Ile Met Ala Ile Val Ile Asn
100 105 110

Val Ala Leu Ser Tyr Pro Thr Leu Pro Gly Trp Ile Pro Ser Pro Phe
115 120 125

Phe Pro Ile Tyr Leu Tyr Asn Ile His Lys Ala Lys His Gly Ser Asp
130 135 140

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

Asn Leu Phe Ser Lys His Ala Arg Thr Met Pro Asp Arg Leu Thr Leu
165 170 175

4197737_1

Gly Glu Leu Trp Ser Met Thr Glu Ala Asn Arg Glu Ala Phe Asp Ile
180 185 190

Phe Gly Trp Ile Ala Ser Lys Met Glu Trp Thr Leu Leu Tyr Ile Leu
195 200 205

Ala Arg Asp Gln Asp Gly Phe Leu Ser Lys Glu Ala Ile Arg Arg Cys
210 215 220

Tyr Asp Gly Ser Leu Phe Glu Tyr Cys Ala Lys Met Gln Arg Gly Ala
225 230 235 240

Glu Asp Lys Met Lys
245

<210> 64
<211> 985
<212> DNA
<213> G. max

<400> 64
atacgagaga gtgagtaatc gagaaagaga agaagcatat atatcaaaat ggctgcagag 60

atggagaggg agtcattgat aactgaagct cctaatgcac cagttactgc acagagaagg 120

gtcagaaatg acttagaaaa ttctctacca aaaccatact tgccaagagc attgaaagct 180

cctgatacgg gtcacccaaa tggaacagca ggccacaggc accacaactt atctgttctt 240

cagcagcatt gtgctttttt tgatcaagat gacaatggaa tcatttaccc ttgggaaact 300

tacatggggc tgcgttctat tggatttaat gttgttgcat ctgttattat ggctattgtt 360

atcaatgttg gattgagtta ccccactcta cctaattggt tcccttctct ccttttttcct 420

atctacatac acaacataca caaagcaaag catgggagtg actctggagt ttatgacaca 480

gaaggacgtt atgtgccagc aaatattgag aacatattca gtaagtatgc tcgtacagta 540

cctgacaagc tcacacttgg ggagctctgg gacttgacag agggaaaccg aaatgctttt 600

gacatatttg gctggcttgc agcaaaattt gaatgggggg ttctgtacat tctggcaagg 660

gatgaggaag gtttcctgtc taaagaagct gttagaagat gctttgatgg gagcttattt 720

gaatactgtg ctaaaatgca tactactagt gatgccaaga tgagttgaaa ttggattatg 780

cagtgtgcat agttacgtgt gatgtgtcgt ttctgttaga catgttagtg gtgtaatgtg 840

aaaaaatagt aaaaaaataa acttattagg aactcgtggt ctgggattga tcagcataat 900

aaagagttta ctgtctctag catatgttgt tctcgagatg taatcctttc ctcacgtgtg 960

gaaagaagtt actctggttc ctaca 985

<210> 65
<211> 239
<212> PRT
<213> G. max

<400> 65

4197737_1

Met Ala Ala Glu Met Glu Arg Glu Ser Leu Ile Thr Glu Ala Pro Asn
1 5 10 15

Ala Pro Val Thr Ala Gln Arg Arg Val Arg Asn Asp Leu Glu Asn Ser
20 25 30

Leu Pro Lys Pro Tyr Leu Pro Arg Ala Leu Lys Ala Pro Asp Thr Gly
35 40 45

His Pro Asn Gly Thr Ala Gly His Arg His His Asn Leu Ser Val Leu
50 55 60

Gln Gln His Cys Ala Phe Phe Asp Gln Asp Asp Asn Gly Ile Ile Tyr
65 70 75 80

Pro Trp Glu Thr Tyr Met Gly Leu Arg Ser Ile Gly Phe Asn Val Val
85 90 95

Ala Ser Val Ile Met Ala Ile Val Ile Asn Val Gly Leu Ser Tyr Pro
100 105 110

Thr Leu Pro Asn Trp Phe Pro Ser Leu Leu Phe Pro Ile Tyr Ile His
115 120 125

Asn Ile His Lys Ala Lys His Gly Ser Asp Ser Gly Val Tyr Asp Thr
130 135 140

Glu Gly Arg Tyr Val Pro Ala Asn Ile Glu Asn Ile Phe Ser Lys Tyr
145 150 155 160

Ala Arg Thr Val Pro Asp Lys Leu Thr Leu Gly Glu Leu Trp Asp Leu
165 170 175

Thr Glu Gly Asn Arg Asn Ala Phe Asp Ile Phe Gly Trp Leu Ala Ala
180 185 190

Lys Phe Glu Trp Gly Val Leu Tyr Ile Leu Ala Arg Asp Glu Glu Gly
195 200 205

Phe Leu Ser Lys Glu Ala Val Arg Arg Cys Phe Asp Gly Ser Leu Phe
210 215 220

Glu Tyr Cys Ala Lys Met His Thr Thr Ser Asp Ala Lys Met Ser
225 230 235

<210> 66
<211> 1326
<212> DNA
<213> Z. mays

<400> 66
atcccagcag caggtagcag cagcgcatga ctttgtatct gcaactcgtt tcgactagcc 60

tgcacgccgg gagccctcgt ctcgcccttc tccacctccg aaggtagggt acctgatact 120

4197737_1

tccctgcccc gcgcatcgca tccgcggttt tagtttgctg gcctcggttc cgatcgattc 180

ctatcctccc tccgggcggg cgtccaccga cttcatcatc cactcgccgg ccgaaccgat 240

ttgtttcgtt tgccgttcga tccgcagatc ctgcgagttc aacccgcgag tgagcatgtc 300

gtcctactcc ccgccgccgc cgccgccgcg ggaccagtcc atggacaccg aggcacccaa 360

cgcgcccatc accagggagc ggaggctcaa ccccgatctg caggagcagc tccccaagcc 420

atatctcgcg agagctctcg aggcggtgga cccgagccac ccgcagggga ccaaggggcg 480

cgacccccgc ggcatgagcg tgcttcagca gcacgccgcc ttcttcgacc gcaatggcga 540

cggcgtcatc tacccctggg agacgtttca aggactgcga gcgataggat gtggactcac 600

tgtatcattc gcgttctcca tactgatcaa cctcttcctc agttatccca ctcagccggg 660

atggttacct tctcctttgc tgtccatccg tatagacaac atccacaagg gtaagcacgg 720

gagtgattct gaaacctacg acactgaagg gaggtttgat ccatcaaagt tcgatgctat 780

attcagtaag tacggtcgaa cccatcctaa tgctataaca agagacgagt taagctcaat 840

gcttcaagga aaccgcaata cgtacgattt ccttggctgg ttggccgctg ccggtgaatg 900

gctcttactc tacagcttgg cgaaagacaa ggatggcctc ttgcagcgcg aaactgtccg 960

tggtctattt gatgggagcc tatttgagcg actggaagac gacaacaaca agaagaaatc 1020

gtcatgaatg ctgagcagcc ttgtacagct cagggaagtg ctgtcagtac aaaactacca 1080

gatataccat tggtcgtgtt caaataacaa atgcttcggc tttgttcatc cgtcattaac 1140

tatgagtgct gggatttgtt tgtatgtgtg tcgtgctacc agtttcttct cctgtcgtct 1200

cacacaggta ctgaattacg catgtgtttt ctagtgttcg tgcggaagct gtattataag 1260

ctgaaaaatg tgcgtttgaa atttatgggc aaaactgtct tcttggtctt aaaaaaaaaa 1320

aaaaaa 1326

<210> 67
<211> 243
<212> PRT
<213> Z. mays

<400> 67

Met Ser Ser Tyr Ser Pro Pro Pro Pro Pro Pro Arg Asp Gln Ser Met
1 5 10 15

Asp Thr Glu Ala Pro Asn Ala Pro Ile Thr Arg Glu Arg Arg Leu Asn
20 25 30

Pro Asp Leu Gln Glu Gln Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu
35 40 45

Glu Ala Val Asp Pro Ser His Pro Gln Gly Thr Lys Gly Arg Asp Pro
50 55 60

Arg Gly Met Ser Val Leu Gln Gln His Ala Ala Phe Phe Asp Arg Asn
65 70 75 80

4197737_1

Gly Asp Gly Val Ile Tyr Pro Trp Glu Thr Phe Gln Gly Leu Arg Ala
85 90 95

Ile Gly Cys Gly Leu Thr Val Ser Phe Ala Phe Ser Ile Leu Ile Asn
100 105 110

Leu Phe Leu Ser Tyr Pro Thr Gln Pro Gly Trp Leu Pro Ser Pro Leu
115 120 125

Leu Ser Ile Arg Ile Asp Asn Ile His Lys Gly Lys His Gly Ser Asp
130 135 140

Ser Glu Thr Tyr Asp Thr Glu Gly Arg Phe Asp Pro Ser Lys Phe Asp
145 150 155 160

Ala Ile Phe Ser Lys Tyr Gly Arg Thr His Pro Asn Ala Ile Thr Arg
165 170 175

Asp Glu Leu Ser Ser Met Leu Gln Gly Asn Arg Asn Thr Tyr Asp Phe
180 185 190

Leu Gly Trp Leu Ala Ala Ala Gly Glu Trp Leu Leu Leu Tyr Ser Leu
195 200 205

Ala Lys Asp Lys Asp Gly Leu Leu Gln Arg Glu Thr Val Arg Gly Leu
210 215 220

Phe Asp Gly Ser Leu Phe Glu Arg Leu Glu Asp Asp Asn Asn Lys Lys
225 230 235 240

Lys Ser Ser

<210> 68
<211> 1058
<212> DNA
<213> B. napus

<400> 68
cttcccattc tcctctagag caaaaaagag cgagagatga gtacggcgac tgagataatg 60

gagagagacg caatggctgc ggtggctccc tacgctccgg tcacctttca ccgccgtgct 120

cgtgttgaca tggatgatag acttcctaaa ccttatatgc caagagcact gcaagcaccc 180

gacagagagc atccgtatgg aaccccaggc cataagaatt atggacttag tgttcttcaa 240

cagcacgtcg ccttcttcga tttagatgat aatggaatta tctacccttg ggagacctac 300

tctggactgc gaatgcttgg tttcaatatc attgtatcgc ttatcgcagc cgctgtaatc 360

aatttggccc ttagctatgc tactcttccg ggatggtttc cttcgccgtt cttcccaata 420

tacatacaca atatacacaa gtcaaagcat gggagcgact caagaacata tgacaatgaa 480

gggaggttta tgcctgtgaa tcttgagttg atatttagca aatatgcgaa aacattgcca 540

4197737_1

gacaagttga gtcttggaga attatgggag atgacacaag gacaacgtga cgcatgggac 600

atcttcggat ggttcgcaag caaaatagag tgggggttgt tgtacttgct agcgagggat 660

gaagaagggt ttctgtcaaa agaagcgatt aggaggtgtt ttgacgggag cttgttcgag 720

tattgtgcca agatatacgc aggtatcaat gaagacaaga cagcctacta ctaaaagtaa 780

atggtagagg agctttaggc tgataatcgt cgatgtgaat gtaacttgtg tctaaagcag 840

agtccatgtg tttgttatgt tatgtcaaaa tctgtaaggt agaagtatga tcagttgcag 900

ctggtataga aaaacttcta tgttgaataa tagtatgttt tgttgttgtg tttgtgtttg 960

tatcaaccct tttttagtta tttccagttc aaatgtaatt tttcattgtt gttacttggt 1020

attgagaatc atttcttgta taaaaaaaaa aaaaaaaa 1058

<210> 69
<211> 245
<212> PRT
<213> B. napus

<400> 69

Met Ser Thr Ala Thr Glu Ile Met Glu Arg Asp Ala Met Ala Ala Val
1 5 10 15

Ala Pro Tyr Ala Pro Val Thr Phe His Arg Arg Ala Arg Val Asp Met
20 25 30

Asp Asp Arg Leu Pro Lys Pro Tyr Met Pro Arg Ala Leu Gln Ala Pro
35 40 45

Asp Arg Glu His Pro Tyr Gly Thr Pro Gly His Lys Asn Tyr Gly Leu
50 55 60

Ser Val Leu Gln Gln His Val Ala Phe Phe Asp Leu Asp Asp Asn Gly
65 70 75 80

Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Met Leu Gly Phe
85 90 95

Asn Ile Ile Val Ser Leu Ile Ala Ala Ala Val Ile Asn Leu Ala Leu
100 105 110

Ser Tyr Ala Thr Leu Pro Gly Trp Phe Pro Ser Pro Phe Phe Pro Ile
115 120 125

Tyr Ile His Asn Ile His Lys Ser Lys His Gly Ser Asp Ser Arg Thr
130 135 140

Tyr Asp Asn Glu Gly Arg Phe Met Pro Val Asn Leu Glu Leu Ile Phe
145 150 155 160

Ser Lys Tyr Ala Lys Thr Leu Pro Asp Lys Leu Ser Leu Gly Glu Leu
165 170 175

4197737_1

Trp Glu Met Thr Gln Gly Gln Arg Asp Ala Trp Asp Ile Phe Gly Trp
180 185 190

Phe Ala Ser Lys Ile Glu Trp Gly Leu Leu Tyr Leu Leu Ala Arg Asp
195 200 205

Glu Glu Gly Phe Leu Ser Lys Glu Ala Ile Arg Arg Cys Phe Asp Gly
210 215 220

Ser Leu Phe Glu Tyr Cys Ala Lys Ile Tyr Ala Gly Ile Asn Glu Asp
225 230 235 240

Lys Thr Ala Tyr Tyr
245

<210> 70
<211> 1013
<212> DNA
<213> C. revoluta

<400> 70
acgcgggggc agttgaggtg atcctcaata tcgtagcgaa gatcgcccca aatggcttcc 60

gtcgagtctc ttcaaactac ggctctaagg gcaccagtta cactggaacg tagagtgaat 120

cccaatttgg atgatgaaat tccaaagcca tttctaccaa gagcgcttgt agctgtagat 180

acagagcatt tagatggatc cccgggacac cagcataaca acatgagtgt tcttcagcag 240

catgttgcat ttttcgatcg aaatcatgac ggaattgtgt atccttggga aacatacgaa 300

ggcttccgtg ccataggatt taacatcgtc atttccttaa tgtctgcact gttcatcaac 360

atagccttga gctatcttac tcttcctgga tggataccgt cactgctctt tccaatacat 420

ataaacagga ttcatcgagc aaaacatgga agtgattcag aggtgtatga caccgagggg 480

aggtttgttc catcgaaatt cgaggaaatt ttcactaaat atgcacgtgt acgtccagat 540

aggctcacat tctctgaaat actgttggcc ttgacggagg ccaatagaaa cgccaatgac 600

cctttcggat ggttaaccag caaggccaag tggggcctcc tctaccttct tgccaaggat 660

gatcaaggtt ttctgccgaa agaggcagtc agaggagttt atgatggcag tttgtttgag 720

ttattggaaa agcaacgatc atcccgtaag cagaaataaa atggcctttt gaattagacc 780

ctgaaagagc atgagtgaga aataaactgc acataaatgg cccactagcc catctatgct 840

gtgaaaaata tgatgtcatt agtcattgtc agttagcata ttttaacaag catgactttg 900

tacatattgc atgccgtatc atgaattatg ttcttggttg tgtgatctag actccatttc 960

ttaatgaaat gtcttcctta tttcataaaa aaaaaaaaaa aaaaaaaaaa aaa 1013

<210> 71
<211> 235
<212> PRT
<213> C. revoluta

<400> 71

4197737_1

Met Ala Ser Val Glu Ser Leu Gln Thr Thr Ala Leu Arg Ala Pro Val
1 5 10 15

Thr Leu Glu Arg Arg Val Asn Pro Asn Leu Asp Asp Glu Ile Pro Lys
20 25 30

Pro Phe Leu Pro Arg Ala Leu Val Ala Val Asp Thr Glu His Leu Asp
35 40 45

Gly Ser Pro Gly His Gln His Asn Asn Met Ser Val Leu Gln Gln His
50 55 60

Val Ala Phe Phe Asp Arg Asn His Asp Gly Ile Val Tyr Pro Trp Glu
65 70 75 80

Thr Tyr Glu Gly Phe Arg Ala Ile Gly Phe Asn Ile Val Ile Ser Leu
85 90 95

Met Ser Ala Leu Phe Ile Asn Ile Ala Leu Ser Tyr Leu Thr Leu Pro
100 105 110

Gly Trp Ile Pro Ser Leu Leu Phe Pro Ile His Ile Asn Arg Ile His
115 120 125

Arg Ala Lys His Gly Ser Asp Ser Glu Val Tyr Asp Thr Glu Gly Arg
130 135 140

Phe Val Pro Ser Lys Phe Glu Glu Ile Phe Thr Lys Tyr Ala Arg Val
145 150 155 160

Arg Pro Asp Arg Leu Thr Phe Ser Glu Ile Leu Leu Ala Leu Thr Glu
165 170 175

Ala Asn Arg Asn Ala Asn Asp Pro Phe Gly Trp Leu Thr Ser Lys Ala
180 185 190

Lys Trp Gly Leu Leu Tyr Leu Leu Ala Lys Asp Asp Gln Gly Phe Leu
195 200 205

Pro Lys Glu Ala Val Arg Gly Val Tyr Asp Gly Ser Leu Phe Glu Leu
210 215 220

Leu Glu Lys Gln Arg Ser Ser Arg Lys Gln Lys
225 230 235

<210> 72
<211> 720
<212> DNA
<213> C. sativus

<400> 72
atgtctttga cgatggaaaa cgaagccacc gccaccgagg cccgattcgc gcccgtcact 60

4197737_1

cgggagcgaa gggttcgcaa cgacctcgaa accacgctcc ccaagccata tttggcgagg 120

gctttggtgg cggcggactc caaccacccc acggggacgg tggggcataa acactatgga 180

atgacagtgc ttcaacagca tgttgctttc tttgaccaag atgataatgg tatcgtttac 240

ccttgggaga cttatgttgg gctaagagcg attggattca acatgataat gtccctagtt 300

atggccgtta ttataaattt cgccatgagc tatcgcactc aacaaggatg gattccatca 360

cctttctttc ccatctatat ttacaatatt cacagagaca aacatggcag cgacactgga 420

acctatgaca ctgaaggaag gtatatagcg gcgaatttcg agaacatgtt cagcaagtac 480

gctagaacac aaccggacaa gttgtcgctt ggcgagatat gggacatgac tgaagccaac 540

cgtctagcgt tcgaccccta cggatggata gcggcgaaac tggaatggtt catattatac 600

gtgctggcta gggacgagga cgggtatctg tcgaaagagg ccgtaagaag atgttatgat 660

ggaagtttgt tcgagtattg tgcaaagatg aatatgagtg ctcaatacaa gatgtattga 720

<210> 73
<211> 239
<212> PRT
<213> C. sativus

<400> 73

Met Ser Leu Thr Met Glu Asn Glu Ala Thr Ala Thr Glu Ala Arg Phe
1 5 10 15

Ala Pro Val Thr Arg Glu Arg Arg Val Arg Asn Asp Leu Glu Thr Thr
20 25 30

Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu Val Ala Ala Asp Ser Asn
35 40 45

His Pro Thr Gly Thr Val Gly His Lys His Tyr Gly Met Thr Val Leu
50 55 60

Gln Gln His Val Ala Phe Phe Asp Gln Asp Asp Asn Gly Ile Val Tyr
65 70 75 80

Pro Trp Glu Thr Tyr Val Gly Leu Arg Ala Ile Gly Phe Asn Met Ile
85 90 95

Met Ser Leu Val Met Ala Val Ile Ile Asn Phe Ala Met Ser Tyr Arg
100 105 110

Thr Gln Gln Gly Trp Ile Pro Ser Pro Phe Phe Pro Ile Tyr Ile Tyr
115 120 125

Asn Ile His Arg Asp Lys His Gly Ser Asp Thr Gly Thr Tyr Asp Thr
130 135 140

Glu Gly Arg Tyr Ile Ala Ala Asn Phe Glu Asn Met Phe Ser Lys Tyr
145 150 155 160

4197737_1

Ala Arg Thr Gln Pro Asp Lys Leu Ser Leu Gly Glu Ile Trp Asp Met
165 170 175

Thr Glu Ala Asn Arg Leu Ala Phe Asp Pro Tyr Gly Trp Ile Ala Ala
180 185 190

Lys Leu Glu Trp Phe Ile Leu Tyr Val Leu Ala Arg Asp Glu Asp Gly
195 200 205

Tyr Leu Ser Lys Glu Ala Val Arg Arg Cys Tyr Asp Gly Ser Leu Phe
210 215 220

Glu Tyr Cys Ala Lys Met Asn Met Ser Ala Gln Tyr Lys Met Tyr
225 230 235

<210> 74
<211> 2074
<212> DNA
<213> A. thaliana

<400> 74
tgaatccttt ttcctttctt cttcttcttc tcttcagaga aaactttgct tctctttcta 60

taaggaacca gacacgaatc ccattcccac cgatttctta gcttcttcct tcaatccgct 120

ctttccctct ccattagatt ctgtttcctc tttcaatttc ttctgcatgc ttctcgattc 180

tctctgacgc ctcttttctc ccgacgctgt ttcgtcaaac gcttttcgaa atggcgattt 240

tggattctgc tggcgttact acggtgacgg agaacggtgg cggagagttc gtcgatcttg 300

ataggcttcg tcgacggaaa tcgagatcgg attcttctaa cggacttctt ctctctggtt 360

ccgataataa ttctccttcg gatgatgttg gagctcccgc cgacgttagg gatcggattg 420

attccgttgt taacgatgac gctcagggaa cagccaattt ggccggagat aataacggtg 480

gtggcgataa taacggtggt ggaagaggcg gcggagaagg aagaggaaac gccgatgcta 540

cgtttacgta tcgaccgtcg gttccagctc atcggagggc gagagagagt ccacttagct 600

ccgacgcaat cttcaaacag agccatgccg gattattcaa cctctgtgta gtagttctta 660

ttgctgtaaa cagtagactc atcatcgaaa atcttatgaa gtatggttgg ttgatcagaa 720

cggatttctg gtttagttca agatcgctgc gagattggcc gcttttcatg tgttgtatat 780

ccctttcgat ctttcctttg gctgccttta cggttgagaa attggtactt cagaaataca 840

tatcagaacc tgttgtcatc tttcttcata ttattatcac catgacagag gttttgtatc 900

cagtttacgt caccctaagg tgtgattctg ctttttttatc aggtgtcact ttgatgctcc 960

tcacttgcat tgtgtggcta aagttggttt cttatgctca tactagctat gacataagat 1020

ccctagccaa tgcagctgat aaggccaatc ctgaagtctc ctactacgtt agcttgaaga 1080

gcttggcata tttcatggtc gctcccacat tgtgttatca gccaagttat ccacgttctg 1140

catgtatacg gaagggttgg gtggctcgtc aatttgcaaa actggtcata ttcaccggat 1200

tcatgggatt tataatagaa caatatataa atcctattgt caggaactca aagcatcctt 1260

tgaaaggcga tcttctatat gctattgaaa gagtgttgaa gctttcagtt ccaaatttat 1320

atgtgtggct ctgcatgttc tactgcttct tccacctttg gttaaacata ttggcagagc 1380

ttctctgctt cggggatcgt gaattctaca aagattggtg gaatgcaaaa agtgtgggag 1440

attactggag aatgtggaat atgcctgttc ataaatggat ggttcgacat atatacttcc 1500

cgtgcttgcg cagcaagata ccaaagacac tcgccattat cattgctttc ctagtctctg 1560

cagtctttca tgagctatgc atcgcagttc cttgtcgtct cttcaagcta tgggcttttc 1620

ttgggattat gtttcaggtg cctttggtct tcatcacaaa ctatctacag gaaaggtttg 1680

gctcaacggt ggggaacatg atcttctggt tcatcttctg cattttcgga caaccgatgt 1740

gtgtgcttct ttattaccac gacctgatga accgaaaagg atcgatgtca tgaaacaact 1800

gttcaaaaaa tgactttctt caaacatcta tggcctcgtt ggatctccgt tgatgttgtg 1860

gtggttctga tgctaaaacg acaaatagtg ttataaccat tgaagaagaa aagaaaatta 1920

gagttgttgt atctgcaaaa attttggtag agacacgcga acccgtttgg attttgttat 1980

ggtgtaaaga aatttcaatc aaaaaactgt tgtaataatt gttaccaaaa agaaatgctt 2040

ttctggaaac gaggggaaaa atagtagttt tgtt 2074

<210> 75
<211> 520
<212> PRT
<213> A. thaliana

<400> 75

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

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

4197737_1

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

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

<210> 76
<211> 2090
<212> DNA
<213> T. majus

<400> 76
acgcggggag ttttcaaaat catattatgc tttttcttca ctactgcatg aactttcttt 60

ctacttcttg caactgattt gtaatcctta cacatgtttc tagttttctc catataaaaa 120

aaatattctc tgagcttctc gattctctag agagagaagg ccaaaaaaaa atggcggtgg 180

cagagtcgtc acagaacacg acaaccatga gtggtcacgg cgactcggat ctcaacaatt 240

tccgtagaag gaaaccgagt tcctccgtga ttgaaccttc gtcgtccggt tttacatcca 300

ccaatggcgt accggcgact ggccacgtgg ctgagaatcg tgaccaggat cgggtagggg 360

ctatggagaa cgcaacagga tcggtcaact taattggaaa tggtggaggc gtggttatcg 420

ggaatgaaga gaaacaggta ggggagactg atatacgatt cacttaccgg ccttcgtttc 480

cggctcatcg gagggtgagg gagagtcctc ttagctctga tgcaatcttc aaacagagcc 540

atgcgggttt attcaacttg tgtatagtag tgctcattgc agtaaacagt aggcttatca 600

tcgaaaatct tatgaagtat ggttggttga tcgatactgg tttctggttt agctcaagat 660

cactgggtga ttggtccatc tttatgtgct gtcttacact cccaattttc ccacttgctg 720

cttttattgt tgaaaagctg gtgcagcgaa atcatatatc tgaacttgtt gctgttctcc 780

ttcatgtaat cgtttctacc gctgcagttt tatatccagt tattgtgatc ttaacgtgtg 840

attcggtgta tatgtctggt gtggtattga tgctctttgg ttgcattatg tggttgaagc 900

tggtgtcata tgcacatact agttctgata ttagaacact ggccaaatct ggctataagg 960

gggatgcgca ccccaattca accattgtga gttgctcata tgatgttagc ttgaagagtt 1020

4197737_1

tggcatactt catggttgct ccgacattat gttaccagcc tagctatcct cgttcgtcgt 1080

gtatccgcaa gggttgggtt gttcgtcaat ttgtcaaact aatagttttc ataggactca 1140

tggggttcat tatagaacaa tatattaatc ctatcgttcg aaattccaaa cacccattga 1200

aaggagattt tttatatgca atagaaagag ttttgaagct ttcagttcca aatctatatg 1260

tttggctttg catgttctac tctttttttcc acctctggtt gaacatactg gctgagcttc 1320

ttcgctttgg tgatcgtgaa ttctacaaag attggtggaa tgcaaaaact gttgcggagt 1380

attggaaaat gtggaatatg cctgttcata gatggatggt tcgtcatcta tattttccct 1440

gtttgaggaa tgggataccc aaggaaggtg ccattattat cgcgttctta gtttctggtg 1500

ctttccatga gctctgcatt gcagttcctt gccacgtatt caagttatgg gcctttatag 1560

gcattatgtt tcaggttccc ttggtattga ttacgaatta tctacaagaa aagttcagta 1620

attctatggt gggcaatatg atcttctggt tcatcttctg catacttggc caacctatgt 1680

gtgtccttct atattaccat gacctgataa atctaaagga aaagtgaaaa aatggaagtt 1740

gcctatgctc agagtattcc tatcccaatg cacacattat atggttctgt acaatctgtg 1800

ccccttcat cctttacacg tacccatgct ggttcctgca cgatgatttg ccttttgttt 1860

gtaagcaata tttggagaga gtccaattta ggaagtgact agtgtggctt atatcttgta 1920

tactaccttt agtcatgggg gggtttttat attactagta ccaaaagtca agttgtatat 1980

gatttacggt ttagtttctt tcatgttttt tgtttttgtg taaatatacg tttcatatat 2040

cactgttttt tcaaagtaaa atcaataata ccccatagat gttgaaactg 2090

<210> 77
<211> 518
<212> PRT
<213> T. majus

<400> 77

Met Ala Val Ala Glu Ser Ser Gln Asn Thr Thr Thr Met Ser Gly His
1 5 10 15

Gly Asp Ser Asp Leu Asn Asn Phe Arg Arg Arg Lys Pro Ser Ser Ser
20 25 30

Val Ile Glu Pro Ser Ser Ser Gly Phe Thr Ser Thr Asn Gly Val Pro
35 40 45

Ala Thr Gly His Val Ala Glu Asn Arg Asp Gln Asp Arg Val Gly Ala
50 55 60

Met Glu Asn Ala Thr Gly Ser Val Asn Leu Ile Gly Asn Gly Gly Gly
65 70 75 80

Val Val Ile Gly Asn Glu Glu Lys Gln Val Gly Glu Thr Asp Ile Arg
85 90 95

```
Phe Thr Tyr Arg Pro Ser Phe Pro Ala His Arg Arg Val Arg Glu Ser
            100                 105                 110

Pro Leu Ser Ser Asp Ala Ile Phe Lys Gln Ser His Ala Gly Leu Phe
        115                 120                 125

Asn Leu Cys Ile Val Val Leu Ile Ala Val Asn Ser Arg Leu Ile Ile
    130                 135                 140

Glu Asn Leu Met Lys Tyr Gly Trp Leu Ile Asp Thr Gly Phe Trp Phe
145                 150                 155                 160

Ser Ser Arg Ser Leu Gly Asp Trp Ser Ile Phe Met Cys Cys Leu Thr
                165                 170                 175

Leu Pro Ile Phe Pro Leu Ala Ala Phe Ile Val Glu Lys Leu Val Gln
            180                 185                 190

Arg Asn His Ile Ser Glu Leu Val Ala Val Leu Leu His Val Ile Val
        195                 200                 205

Ser Thr Ala Ala Val Leu Tyr Pro Val Ile Val Ile Leu Thr Cys Asp
    210                 215                 220

Ser Val Tyr Met Ser Gly Val Val Leu Met Leu Phe Gly Cys Ile Met
225                 230                 235                 240

Trp Leu Lys Leu Val Ser Tyr Ala His Thr Ser Ser Asp Ile Arg Thr
                245                 250                 255

Leu Ala Lys Ser Gly Tyr Lys Gly Asp Ala His Pro Asn Ser Thr Ile
            260                 265                 270

Val Ser Cys Ser Tyr Asp Val Ser Leu Lys Ser Leu Ala Tyr Phe Met
        275                 280                 285

Val Ala Pro Thr Leu Cys Tyr Gln Pro Ser Tyr Pro Arg Ser Ser Cys
    290                 295                 300

Ile Arg Lys Gly Trp Val Val Arg Gln Phe Val Lys Leu Ile Val Phe
305                 310                 315                 320

Ile Gly Leu Met Gly Phe Ile Ile Glu Gln Tyr Ile Asn Pro Ile Val
                325                 330                 335

Arg Asn Ser Lys His Pro Leu Lys Gly Asp Phe Leu Tyr Ala Ile Glu
            340                 345                 350

Arg Val Leu Lys Leu Ser Val Pro Asn Leu Tyr Val Trp Leu Cys Met
        355                 360                 365

Phe Tyr Ser Phe Phe His Leu Trp Leu Asn Ile Leu Ala Glu Leu Leu
```

4197737_1

370 375 380

Arg Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp Trp Asn Ala Lys Thr
385 390 395 400

Val Ala Glu Tyr Trp Lys Met Trp Asn Met Pro Val His Arg Trp Met
405 410 415

Val Arg His Leu Tyr Phe Pro Cys Leu Arg Asn Gly Ile Pro Lys Glu
420 425 430

Gly Ala Ile Ile Ile Ala Phe Leu Val Ser Gly Ala Phe His Glu Leu
435 440 445

Cys Ile Ala Val Pro Cys His Val Phe Lys Leu Trp Ala Phe Ile Gly
450 455 460

Ile Met Phe Gln Val Pro Leu Val Leu Ile Thr Asn Tyr Leu Gln Glu
465 470 475 480

Lys Phe Ser Asn Ser Met Val Gly Asn Met Ile Phe Trp Phe Ile Phe
485 490 495

Cys Ile Leu Gly Gln Pro Met Cys Val Leu Leu Tyr Tyr His Asp Leu
500 505 510

Ile Asn Leu Lys Glu Lys
515

<210> 78
<211> 1485
<212> DNA
<213> Z. mays

<400> 78
atggccccgc ccccctccat gcctgccgcc tccgatcgcg ccggccctgg ccgcgacgcg 60

ggcgactcgt cctcccttcg cctccgccgc gccccctcag ccgacgccgg cgaccttgcc 120

ggcgattcct cgggaggctt gcgggagaac ggcgagccgc aatcgccgac gaatccgccg 180

ccgcaggagc agcagcagca cgagatgcta tactaccgcg cgtcggcgcc cgcccaccgc 240

cgcgtcaagg agagcccccc cagctctgac gccatcttcc ggcagagcca tgctggtctt 300

ctgaatctat gcattgttgt tctgatcgca gtgaacagca gactcattat tgagaattta 360

atgaagtatg gcctgttgat aagagctgga ttttggttta gtgcaagatc gctgggtgac 420

tggccccttc taatgtgctg cctcactcta ccagttttcc cactagttgc actcatggct 480

gagaagctga tcacaagaaa gctcattggt gaacatgtgg ttattctact ccatatcatt 540

attacaacat ctgccattgt ctatccagtt gttgtgactc ttaagtgtga ctcagcagta 600

ctatctggat ttgtgctaat gtttcttgcg agcatcatgt ggatgaagct tgtctcttat 660

gcacatacaa attatgatat aagggtattg tccaaaagta ctgaaaaggg tgctgcatat 720

ggaaattatg tcgatcctga gaatatgaaa gatccaacct ttaaaagtct agtgtacttt 780

atgttggccc caacactttg ttaccagcca acttatcctc aaactacatg tattagaaag 840

ggttgggtga cccagcaact cataaagtgc gtggttttta caggcttgat gggcttcata 900

attgagcaat atataaaccc aattgtgaag aattccaaac atccactgaa agggaatttt 960

ttgaatgcta tagaaagagt cttaaaactc tcagtgccaa cattatatgt atggctttgc 1020

atgttctatt gctttttca tttatggctg aacattgtag ctgaactcct ctgtttcggt 1080

gaccgtgaat tctataagga ctggtggaat gccaaaactg ttgaagagta ctggaggatg 1140

tggaacatgc ctgttcataa gtggatcatc agacacatat attttccatg tataaggaaa 1200

ggcttttcca ggggtgtagc tattctaatc tcgtttctgg tttcagctgt atttcatgag 1260

atatgtattg cggtgccttg ccacattttc aaattctggg cattttctgg gatcatgttt 1320

cagataccct tggtattctt gacaagatat ctccatgcta cgttcaagca tgtaatggtg 1380

ggcaacatga tattttggtt cttcttcagt atagtcggac agccgatgtg tgtccttcta 1440

tactaccatg acgtcatgaa caggcaggcc caggcaagta gatag 1485

<210> 79
<211> 494
<212> PRT
<213> Z. mays

<400> 79

Met Ala Pro Pro Pro Ser Met Pro Ala Ala Ser Asp Arg Ala Gly Pro
1 5 10 15

Gly Arg Asp Ala Gly Asp Ser Ser Ser Leu Arg Leu Arg Arg Ala Pro
20 25 30

Ser Ala Asp Ala Gly Asp Leu Ala Gly Asp Ser Ser Gly Gly Leu Arg
35 40 45

Glu Asn Gly Glu Pro Gln Ser Pro Thr Asn Pro Pro Pro Gln Glu Gln
50 55 60

Gln Gln His Glu Met Leu Tyr Tyr Arg Ala Ser Ala Pro Ala His Arg
65 70 75 80

Arg Val Lys Glu Ser Pro Leu Ser Ser Asp Ala Ile Phe Arg Gln Ser
85 90 95

His Ala Gly Leu Leu Asn Leu Cys Ile Val Val Leu Ile Ala Val Asn
100 105 110

Ser Arg Leu Ile Ile Glu Asn Leu Met Lys Tyr Gly Leu Leu Ile Arg
115 120 125

Ala Gly Phe Trp Phe Ser Ala Arg Ser Leu Gly Asp Trp Pro Leu Leu
130 135 140

```
Met Cys Cys Leu Thr Leu Pro Val Phe Pro Leu Val Ala Leu Met Ala
145                 150                 155                 160

Glu Lys Leu Ile Thr Arg Lys Leu Ile Gly Glu His Val Val Ile Leu
                165                 170                 175

Leu His Ile Ile Ile Thr Thr Ser Ala Ile Val Tyr Pro Val Val Val
            180                 185                 190

Thr Leu Lys Cys Asp Ser Ala Val Leu Ser Gly Phe Val Leu Met Phe
        195                 200                 205

Leu Ala Ser Ile Met Trp Met Lys Leu Val Ser Tyr Ala His Thr Asn
    210                 215                 220

Tyr Asp Ile Arg Val Leu Ser Lys Ser Thr Glu Lys Gly Ala Ala Tyr
225                 230                 235                 240

Gly Asn Tyr Val Asp Pro Glu Asn Met Lys Asp Pro Thr Phe Lys Ser
                245                 250                 255

Leu Val Tyr Phe Met Leu Ala Pro Thr Leu Cys Tyr Gln Pro Thr Tyr
            260                 265                 270

Pro Gln Thr Thr Cys Ile Arg Lys Gly Trp Val Thr Gln Gln Leu Ile
        275                 280                 285

Lys Cys Val Val Phe Thr Gly Leu Met Gly Phe Ile Ile Glu Gln Tyr
    290                 295                 300

Ile Asn Pro Ile Val Lys Asn Ser Lys His Pro Leu Lys Gly Asn Phe
305                 310                 315                 320

Leu Asn Ala Ile Glu Arg Val Leu Lys Leu Ser Val Pro Thr Leu Tyr
                325                 330                 335

Val Trp Leu Cys Met Phe Tyr Cys Phe Phe His Leu Trp Leu Asn Ile
            340                 345                 350

Val Ala Glu Leu Leu Cys Phe Gly Asp Arg Glu Phe Tyr Lys Asp Trp
        355                 360                 365

Trp Asn Ala Lys Thr Val Glu Glu Tyr Trp Arg Met Trp Asn Met Pro
    370                 375                 380

Val His Lys Trp Ile Ile Arg His Ile Tyr Phe Pro Cys Ile Arg Lys
385                 390                 395                 400

Gly Phe Ser Arg Gly Val Ala Ile Leu Ile Ser Phe Leu Val Ser Ala
                405                 410                 415
```

4197737_1

Val Phe His Glu Ile Cys Ile Ala Val Pro Cys His Ile Phe Lys Phe
420 425 430

Trp Ala Phe Ser Gly Ile Met Phe Gln Ile Pro Leu Val Phe Leu Thr
435 440 445

Arg Tyr Leu His Ala Thr Phe Lys His Val Met Val Gly Asn Met Ile
450 455 460

Phe Trp Phe Phe Phe Ser Ile Val Gly Gln Pro Met Cys Val Leu Leu
465 470 475 480

Tyr Tyr His Asp Val Met Asn Arg Gln Ala Gln Ala Ser Arg
485 490

<210> 80
<211> 1330
<212> DNA
<213> A. thaliana

<400> 80
agttaaaaga ttggttattt gggctctgca ctcaagtgag agagaagata gatagatctg 60

agtagaatct tcgattcatt attcgttgtc gtcgttcatc tgtgagaagc ggacaaacca 120

aagaatccac cggagctagt gatatgggtg gttccagaga gttccgagct gaggaacatt 180

caaatcaatt ccactctatc atcgccatgg ccatctggct tggcgccatt cacttcaacg 240

tcgctcttgt tctctgttct ctcattttcc ttcctccttc tctatctctc atggtcttgg 300

gcttgctctc tctgtttatc tttatcccaa tcgatcatcg tagcaaatat ggtcgtaagc 360

tcgctaggta catatgcaag cacgcgtgta attatttccc cgtctctctg tacgtcgagg 420

attacgaagc tttccagcct aatcgtgcct atgtctttgg ttatgaacca cattcggtgc 480

taccgattgg agttgttgct ctttgtgatc tcacagggtt tatgcctatt cctaacatta 540

aagttcttgc aagtagtgct atattctaca ctccctttct aaggcatata tggacatggt 600

tagggctcac cgctgcttct aggaagaatt tcacttccct tttggattct ggctacagtt 660

gtgttcttgt acctggtggt gtgcaggaga cttttcatat gcaacatgat gctgagaatg 720

tcttcctttc aaggagaaga ggatttgtgc gcatagccat ggaacagggg agccctctgg 780

ttccagtatt ctgctttggt caggcacgcg tgtacaaatg gtggaagccg gattgtgatc 840

tctatcttaa actatctaga gcaatcagat tcaccccgat ctgcttctgg ggagtttttg 900

gatcaccatt accgtgtcga cagcctatgc atgtggtcgt tggtaaacca atagaagtca 960

caaaaactct gaagccaact gacgaagaga ttgctaagtt tcatggccag tatgtggaag 1020

cgcttaggga tctgtttgag aggcacaagt cccgagtcgg ctatgatctt gagctgaaaa 1080

ttctttgaac aaaatctcca atggaaataa ttacttgtgt gtatccttca ttaattgtta 1140

ccttggagct ggatttggac ttaatataaa tgactacatc atgtagtcta catgtattgc 1200

atgtctttag catcgactgt tgaagtaatg gaatacgttt ataaagcctg taaattacat 1260

gtcgtcttgc acaagagtat gtggtaataa taacatttga cccaaaaata atactagtta 1320

aatttttcct 1330

<210> 81
<211> 314
<212> PRT
<213> A. thaliana

<400> 81

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

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

4197737_1

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> 82
<211> 954
<212> DNA
<213> B. napus

<400> 82
atgggcaaag tcagagactt tggagctgag gatcatatcc catcaaacat attccatgca 60

gtgaccgcca tatccatctg cctcagcgcc atttacttga acctcgctct tgttctcatc 120

tccctcttct tcctcccaac ttctctctcc ctcctggtct tgggcctgct ctctctgttt 180

atcatcatcc ctatagatga tcgtagcaag tacggtctta agctggctag gtacatatgc 240

aagcacgcgg ctagttactt ccccgttact ctgcatgtcg aagactacga agctttcaag 300

cctgatcgct cctatgtatt tggttatgaa ccacactcgg tgtggcccat tggagctgtt 360

gcacttgttg atctggcagg gtttatgcct cttcctaaca tcaaacttct tgcaagcaat 420

gctatattct acacgccgtt tctaaggcac atgtgggcat ggttagggct cgcctctgct 480

tctaggaaga gtttctcttc tcttctggag tctggctata gttgtatcct tgtacctggt 540

ggtgtgcagg aaacatttca cttgcaacat gatgttgaga acgtcttcct ttcatcgaga 600

agaggatttg tgcgcatcgc catggaacaa ggggcacctc ttgttccagt tttctgcttt 660

ggtcagtccc gtgcgtacaa gtggtggaag ccggattgtg acctttattt taaactagca 720

agagcgatca ggtttactcc tatctgtttc tggggagttt tcggatcccc aataccatat 780

agacacccta ttcatgtggt ggttggtaaa ccaatacaag ttgcaaagtc tctgcagcca 840

actgatgaag agattgatga gttgcatggc cagtttgtgg aagcgcttaa ggatctgttt 900

gagaggcaca aggccggagc aggctactct gatctgcagt tgaacattct ttga 954

<210> 83
<211> 317
<212> PRT
<213> B. napus

<400> 83

Met Gly Lys Val Arg Asp Phe Gly Ala Glu Asp His Ile Pro Ser Asn

1 5 10 15

Ile Phe His Ala Val Thr Ala Ile Ser Ile Cys Leu Ser Ala Ile Tyr
20 25 30

Leu Asn Leu Ala Leu Val Leu Ile Ser Leu Phe Phe Leu Pro Thr Ser
35 40 45

Leu Ser Leu Leu Val Leu Gly Leu Leu Ser Leu Phe Ile Ile Ile Pro
50 55 60

Ile Asp Asp Arg Ser Lys Tyr Gly Leu Lys Leu Ala Arg Tyr Ile Cys
65 70 75 80

Lys His Ala Ala Ser Tyr Phe Pro Val Thr Leu His Val Glu Asp Tyr
85 90 95

Glu Ala Phe Lys Pro Asp Arg Ser Tyr Val Phe Gly Tyr Glu Pro His
100 105 110

Ser Val Trp Pro Ile Gly Ala Val Ala Leu Val Asp Leu Ala Gly Phe
115 120 125

Met Pro Leu Pro Asn Ile Lys Leu Leu Ala Ser Asn Ala Ile Phe Tyr
130 135 140

Thr Pro Phe Leu Arg His Met Trp Ala Trp Leu Gly Leu Ala Ser Ala
145 150 155 160

Ser Arg Lys Ser Phe Ser Ser Leu Leu Glu Ser Gly Tyr Ser Cys Ile
165 170 175

Leu Val Pro Gly Gly Val Gln Glu Thr Phe His Leu Gln His Asp Val
180 185 190

Glu Asn Val Phe Leu Ser Ser Arg Arg Gly Phe Val Arg Ile Ala Met
195 200 205

Glu Gln Gly Ala Pro Leu Val Pro Val Phe Cys Phe Gly Gln Ser Arg
210 215 220

Ala Tyr Lys Trp Trp Lys Pro Asp Cys Asp Leu Tyr Phe Lys Leu Ala
225 230 235 240

Arg Ala Ile Arg Phe Thr Pro Ile Cys Phe Trp Gly Val Phe Gly Ser
245 250 255

Pro Ile Pro Tyr Arg His Pro Ile His Val Val Val Gly Lys Pro Ile
260 265 270

Gln Val Ala Lys Ser Leu Gln Pro Thr Asp Glu Glu Ile Asp Glu Leu
275 280 285

4197737_1

His Gly Gln Phe Val Glu Ala Leu Lys Asp Leu Phe Glu Arg His Lys
290 295 300

Ala Gly Ala Gly Tyr Ser Asp Leu Gln Leu Asn Ile Leu
305 310 315

<210> 84
<211> 1637
<212> DNA
<213> A. hypogaea

<400> 84
aatgaacttg acataaagtg gttgtttgta acaccccatt tagtgttttg cttagatgtt 60

gagagttcta taaacttttg tactatttgg taccccgtaa ttaatagaaa tagaaatgtg 120

ataatggttc tatgtttcat tccagaaaaa aattgtcatt ttaaaaagtt ttcttaaatt 180

ctgaatggga atgatgatca ttcagatcaa taaggttaac acttttttat atgatatttt 240

atgtaatctg attaattttt ttttggtgac aaaaaactcg tgccgaattc ggcacgaggt 300

caaaacctca gaagagagaa aaggagaatt tggttcctaa ttaattctca ccatcaacga 360

tggaggtttc aggcgccgtt ctaaggaatg tcacgtgccc ttccttttct gtgcacgtga 420

gttcccgtcg tcgtggtggt gatagttgtg ttacagtgcc ggtgaggatg agaaaaaagg 480

cggtggtgcg ttgttgctgc gggttcagtg attcggggca tgtgcagtat tacggggacg 540

agaagaagaa ggagaatgga accgctatgt tgagcaccaa gaagaagctc aagatgctga 600

agaaacgtgt ccttttcgat gatcttcaag gaaacctgac ttgggatgct gctatggttt 660

tgatgaagca gctagagcaa gtaagggcag aggagaagga attgaagaaa aaaaggaagc 720

aagagaagaa ggaggcaaaa ctcaaagcct ctaagatgaa caccaatcct gattgcgaat 780

cgtcatcgtc atcgtcatca tctgaatctg aatctgaatc aagtgagagt gaatgtgaca 840

atgaggtggt tgacatgaag aagaacatta aggttggtgt tgccgttgct gttgccgatt 900

ccccacgaaa ggcggaaacc atgattctat acacctccct tgttgcccga gatgttagtg 960

ctaatcatca tcatcataat gccgtggaat tattctctag aaacaatgac atatcagttg 1020

gaagcattaa tggtggcctt aagaatgaga atactgcggt tattaccact gaagctattc 1080

ctcagaagag gattgaggta tgcatgggaa acaagtgcaa gaaatccgga tctattgcat 1140

tgttgcaaga atttgagaga gtggttggtg ctgaaggagg tgctgctgct gcagttgttg 1200

gatgcaagtg catggggaag tgcaagagtg cacctaatgt gaggattcag aactctactg 1260

cagataaaat agctgagggg ttcaatgatt cagttaaggt tccagctaac cctctttgca 1320

ttggggttgc atggaggatg ttgaaaccat tgtggcttag attcttgggc gagaatcagg 1380

aaagtactaa tgaataattt gctggtatgc tgtttggaaa attgtatata cgtagtgcca 1440

gaacctatca gattgttgtt ttattttata taaacataga ctgcatattg ttgtgagatt 1500

cgatttcctc atttattgga acttccagag cctgatttgt gtccattcga gctcgactca 1560

aagatttaca tggcctgctc aatctatgaa ttcaaatttg agggccctgt ttggcattaa 1620

4197737_1

tattaatata ttaatat 1637

<210> 85
<211> 345
<212> PRT
<213> A. hypogaea

<400> 85

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

Ser Val His Val Ser Ser Arg Arg Arg Gly Gly Asp Ser Cys Val Thr
20 25 30

Val Pro Val Arg Met Arg Lys Lys Ala Val Val Arg Cys Cys Cys Gly
35 40 45

Phe Ser Asp Ser Gly His Val Gln Tyr Tyr Gly Asp Glu Lys Lys Lys
50 55 60

Glu Asn Gly Thr Ala Met Leu Ser Thr Lys Lys Lys Leu Lys Met Leu
65 70 75 80

Lys Lys Arg Val Leu Phe Asp Asp Leu Gln Gly Asn Leu Thr Trp Asp
85 90 95

Ala Ala Met Val Leu Met Lys Gln Leu Glu Gln Val Arg Ala Glu Glu
100 105 110

Lys Glu Leu Lys Lys Lys Arg Lys Gln Glu Lys Lys Glu Ala Lys Leu
115 120 125

Lys Ala Ser Lys Met Asn Thr Asn Pro Asp Cys Glu Ser Ser Ser Ser
130 135 140

Ser Ser Ser Ser Glu Ser Glu Ser Glu Ser Ser Glu Ser Glu Cys Asp
145 150 155 160

Asn Glu Val Val Asp Met Lys Lys Asn Ile Lys Val Gly Val Ala Val
165 170 175

Ala Val Ala Asp Ser Pro Arg Lys Ala Glu Thr Met Ile Leu Tyr Thr
180 185 190

Ser Leu Val Ala Arg Asp Val Ser Ala Asn His His His His Asn Ala
195 200 205

Val Glu Leu Phe Ser Arg Asn Asn Asp Ile Ser Val Gly Ser Ile Asn
210 215 220

Gly Gly Leu Lys Asn Glu Asn Thr Ala Val Ile Thr Thr Glu Ala Ile
225 230 235 240

4197737_1

```
Pro Gln Lys Arg Ile Glu Val Cys Met Gly Asn Lys Cys Lys Lys Ser
                245                 250                 255

Gly Ser Ile Ala Leu Leu Gln Glu Phe Glu Arg Val Val Gly Ala Glu
            260                 265                 270

Gly Gly Ala Ala Ala Ala Val Val Gly Cys Lys Cys Met Gly Lys Cys
        275                 280                 285

Lys Ser Ala Pro Asn Val Arg Ile Gln Asn Ser Thr Ala Asp Lys Ile
    290                 295                 300

Ala Glu Gly Phe Asn Asp Ser Val Lys Val Pro Ala Asn Pro Leu Cys
305                 310                 315                 320

Ile Gly Val Ala Trp Arg Met Leu Lys Pro Leu Trp Leu Arg Phe Leu
                325                 330                 335

Gly Glu Asn Gln Glu Ser Thr Asn Glu
            340                 345
```

<210> 86
<211> 2811
<212> DNA
<213> A. thaliana

<400> 86

```
cccaagtgta atattgttat taatacatgg gctatactaa aagccccacg aaaagtttac      60
tgaactattt gaggcccaac aagagcctat cggattaacg cctactgcag aagaaaatct     120
gtctgcactc cacccaagaa aacgcagact aattaatgaa atcaacgaaa cggataggtc     180
gggtctaagg ttgaccatga accgcaacct gaaccaggag caaagtggtc aagttttgcc     240
atccggtccg agtcccttgg aggaataata ccagaacaga aaaaaacaga aaagtcgaca     300
ataaacaaaa gagacaaatt tgatttgatt ggttccagaa attcgcagag aaacagctct     360
ttgtctctct cgactgatct aacaatccct aatctgtgtt ctaaattcct ggacgagatt     420
tgacaaagtc cgtatagctt aacctggttt aatttcaagt gacagatatg ccccttattc     480
atcggaaaaa gccgacggag aaaccatcga cgccgccatc tgaagaggtg gtgcacgatg     540
aggattcgca aaagaaacca cacgaatctt ccaaatccca ccataagaaa tcgaacggag     600
gagggaagtg gtcgtgcatc gattcttgtt gttggttcat tgggtgtgtg tgtgtaacct     660
ggtggtttct tctcttcctt tacaacgcaa tgcctgcgag cttccctcag tatgtaacgg     720
agcgaatcac gggtcctttg cctgacccgc ccggtgttaa gctcaaaaaa gaaggtctta     780
aggcgaaaca tcctgttgtc ttcattcctg ggattgtcac cggtgggctc gagctttggg     840
aaggcaaaca atgcgctgat ggtttattta gaaaacgttt gtggggtgga acttttggtg     900
aagtctacaa aaggcctcta tgttgggtgg aacacatgtc acttgacaat gaaactgggt     960
tggatccagc tggtattaga gttcgagctg tatcaggact cgtggctgct gactactttg    1020
```

4197737_1

ctcctggcta ctttgtctgg gcagtgctga ttgctaacct tgcacatatt ggatatgaag 1080
agaaaaatat gtacatggct gcatatgact ggcggctttc gtttcagaac acagaggtac 1140
gtgatcagac tcttagccgt atgaaaagta atatagagtt gatggtttct accaacggtg 1200
gaaaaaaagc agttatagtt ccgcattcca tgggggtctt gtattttcta cattttatga 1260
agtgggttga ggcaccagct cctctgggtg gcgggggtgg gccagattgg tgtgcaaagt 1320
atattaaggc ggtgatgaac attggtggac catttcttgg tgttccaaaa gctgttgcag 1380
ggcttttctc tgctgaagca aaggatgttg cagttgccag agcgattgcc ccaggattct 1440
tagacaccga tatatttaga cttcagacct tgcagcatgt aatgagaatg acacgcacat 1500
gggactcaac aatgtctatg ttaccgaagg gaggtgacac gatatggggc gggcttgatt 1560
ggtcaccgga gaaaggccac acctgttgtg ggaaaaagca aaagaacaac gaaacttgtg 1620
gtgaagcagg tgaaaacgga gtttccaaga aaagtcctgt taactatgga aggatgatat 1680
cttttgggaa agaagtagca gaggctgcgc catctgagat taataatatt gattttcgag 1740
gtgctgtcaa aggtcagagt atcccaaatc acacctgtcg tgacgtgtgg acagagtacc 1800
atgacatggg aattgctggg atcaaagcta tcgctgagta taaggtctac actgctggtg 1860
aagctataga tctactacat tatgttgctc ctaagatgat ggcgcgtggt gccgctcatt 1920
tctcttatgg aattgctgat gatttggatg acaccaagta tcaagatccc aaatactggt 1980
caaatccgtt agagacaaaa ttaccgaatg ctcctgagat ggaaatctac tcattatacg 2040
gagtggggat accaacggaa cgagcatacg tatacaagct taaccagtct cccgacagtt 2100
gcatcccctt tcagatattc acttctgctc acgaggagga cgaagatagc tgtctgaaag 2160
caggagttta caatgtggat ggggatgaaa cagtacccgt cctaagtgcc gggtacatgt 2220
gtgcaaaagc gtggcgtggc aagacaagat tcaacccttc cggaatcaag acttatataa 2280
gagaatacaa tcactctccg ccggctaacc tgttggaagg gcgcgggacg cagagtggtg 2340
cccatgttga tatcatggga aactttgctt tgatcgaaga tatcatgagg gttgccgccg 2400
gaggtaacgg gtctgatata ggacatgacc aggtccactc tggcatattt gaatggtcgg 2460
agcgtattga cctgaagctg tgaatatcat gatctcttta agctgtcctg tcagcttatg 2520
tgaatccaat actttgaaag agagatcatc atcaattcat catcatcgtc atcatcatga 2580
tgctcaactc acaaagaagc ctgagaatga tactttggtg cgaaattctc aatacctctt 2640
taatattctt attgaatgta aattatacaa tcctatctaa tgtttgaacg ataacacaaa 2700
acttgctgcg ccatgtttgt ttgtcttgtc aaaagcatca atttgtgggt tatacgtagt 2760
gtagaggatg attcaaattt gtgataaatt tggtaatcaa agttaattct g 2811

<210> 87
<211> 671
<212> PRT
<213> A. thaliana

<400> 87

4197737_1

```
Met Pro Leu Ile His Arg Lys Lys Pro Thr Glu Lys Pro Ser Thr Pro
1               5                  10                  15

Pro Ser Glu Glu Val Val His Asp Glu Asp Ser Gln Lys Lys Pro His
            20                  25                  30

Glu Ser Ser Lys Ser His His Lys Lys Ser Asn Gly Gly Gly Lys Trp
        35                  40                  45

Ser Cys Ile Asp Ser Cys Cys Trp Phe Ile Gly Cys Val Cys Val Thr
    50                  55                  60

Trp Trp Phe Leu Leu Phe Leu Tyr Asn Ala Met Pro Ala Ser Phe Pro
65                  70                  75                  80

Gln Tyr Val Thr Glu Arg Ile Thr Gly Pro Leu Pro Asp Pro Pro Gly
                85                  90                  95

Val Lys Leu Lys Lys Glu Gly Leu Lys Ala Lys His Pro Val Val Phe
            100                 105                 110

Ile Pro Gly Ile Val Thr Gly Gly Leu Glu Leu Trp Glu Gly Lys Gln
        115                 120                 125

Cys Ala Asp Gly Leu Phe Arg Lys Arg Leu Trp Gly Gly Thr Phe Gly
    130                 135                 140

Glu Val Tyr Lys Arg Pro Leu Cys Trp Val Glu His Met Ser Leu Asp
145                 150                 155                 160

Asn Glu Thr Gly Leu Asp Pro Ala Gly Ile Arg Val Arg Ala Val Ser
                165                 170                 175

Gly Leu Val Ala Ala Asp Tyr Phe Ala Pro Gly Tyr Phe Val Trp Ala
            180                 185                 190

Val Leu Ile Ala Asn Leu Ala His Ile Gly Tyr Glu Glu Lys Asn Met
        195                 200                 205

Tyr Met Ala Ala Tyr Asp Trp Arg Leu Ser Phe Gln Asn Thr Glu Val
    210                 215                 220

Arg Asp Gln Thr Leu Ser Arg Met Lys Ser Asn Ile Glu Leu Met Val
225                 230                 235                 240

Ser Thr Asn Gly Gly Lys Lys Ala Val Ile Val Pro His Ser Met Gly
                245                 250                 255

Val Leu Tyr Phe Leu His Phe Met Lys Trp Val Glu Ala Pro Ala Pro
            260                 265                 270

Leu Gly Gly Gly Gly Gly Pro Asp Trp Cys Ala Lys Tyr Ile Lys Ala
```

275 280 285

Val Met Asn Ile Gly Gly Pro Phe Leu Gly Val Pro Lys Ala Val Ala
290 295 300

Gly Leu Phe Ser Ala Glu Ala Lys Asp Val Ala Val Ala Arg Ala Ile
305 310 315 320

Ala Pro Gly Phe Leu Asp Thr Asp Ile Phe Arg Leu Gln Thr Leu Gln
325 330 335

His Val Met Arg Met Thr Arg Thr Trp Asp Ser Thr Met Ser Met Leu
340 345 350

Pro Lys Gly Gly Asp Thr Ile Trp Gly Gly Leu Asp Trp Ser Pro Glu
355 360 365

Lys Gly His Thr Cys Cys Gly Lys Lys Gln Lys Asn Asn Glu Thr Cys
370 375 380

Gly Glu Ala Gly Glu Asn Gly Val Ser Lys Lys Ser Pro Val Asn Tyr
385 390 395 400

Gly Arg Met Ile Ser Phe Gly Lys Glu Val Ala Glu Ala Ala Pro Ser
405 410 415

Glu Ile Asn Asn Ile Asp Phe Arg Gly Ala Val Lys Gly Gln Ser Ile
420 425 430

Pro Asn His Thr Cys Arg Asp Val Trp Thr Glu Tyr His Asp Met Gly
435 440 445

Ile Ala Gly Ile Lys Ala Ile Ala Glu Tyr Lys Val Tyr Thr Ala Gly
450 455 460

Glu Ala Ile Asp Leu Leu His Tyr Val Ala Pro Lys Met Met Ala Arg
465 470 475 480

Gly Ala Ala His Phe Ser Tyr Gly Ile Ala Asp Asp Leu Asp Asp Thr
485 490 495

Lys Tyr Gln Asp Pro Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu
500 505 510

Pro Asn Ala Pro Glu Met Glu Ile Tyr Ser Leu Tyr Gly Val Gly Ile
515 520 525

Pro Thr Glu Arg Ala Tyr Val Tyr Lys Leu Asn Gln Ser Pro Asp Ser
530 535 540

Cys Ile Pro Phe Gln Ile Phe Thr Ser Ala His Glu Glu Asp Glu Asp
545 550 555 560

Ser Cys Leu Lys Ala Gly Val Tyr Asn Val Asp Gly Asp Glu Thr Val
565 570 575

Pro Val Leu Ser Ala Gly Tyr Met Cys Ala Lys Ala Trp Arg Gly Lys
580 585 590

Thr Arg Phe Asn Pro Ser Gly Ile Lys Thr Tyr Ile Arg Glu Tyr Asn
595 600 605

His Ser Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly
610 615 620

Ala His Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Met
625 630 635 640

Arg Val Ala Ala Gly Gly Asn Gly Ser Asp Ile Gly His Asp Gln Val
645 650 655

His Ser Gly Ile Phe Glu Trp Ser Glu Arg Ile Asp Leu Lys Leu
660 665 670

<210> 88
<211> 2473
<212> DNA
<213> R. communis

<400> 88
cttgaatttt gtttcgatta ctttaaaaaa aattgccttt ttctgaaagc gcttctagtt 60

tgatcttaat ggattattta cttttcagaa attagtagct caaacaaaat taagagagaa 120

aaagagatcg tgggtttttt caaagaaaaa gtttcaaaca gaaagcacaa actttccgat 180

gtggcttgcg aggaagttgg ggatgacata aattcttctg ccagtccgaa gttttggata 240

taaagcagtg gtttttgtat tttccttttt ccttttctgt atcgtttagt aaagtcacat 300

ttggcttatt gggtttgttt tatttcctct gtatttgctt tctgtacaaa gactatcaat 360

aattagttaa taagcttaac aaatttaaaa atcttatttt tctttaaacc cagaaatgcc 420

tgtaattcgg aggaaaaaac ccacttctga acccaacaaa aattcagcat cagactcaaa 480

aacgccaagc gaggaagagg aacatgaaca agaacaggaa caagaagaag ataaaaataa 540

caaaaagaaa tacccaaaga agaagagcag tgaaatcaat gcaaaaaaat ggtcatgcat 600

agacagctgt tgttggtttg ttggttgcat ctgcgtgacg tggtgggttt tactatttct 660

ttacaatgca gtgcctgcgt ctttgcctca atacgtaact gaggcaatca cgggtccttt 720

acccgatcca cctggtgtta agctgaaaaa agagggatta acagcaaagc atccagtggt 780

ttttgtacct gggattgtta ccgcggggct tgaattgtgg gaaggccatc agtgtgctga 840

tgggctgttt aggaaacggc tctggggtgg aacttttgga gaagtttata agaggcctct 900

ctgctgggta gagcatatgt ctctagacaa tgaaactgga ttggatcctc ctggtataag 960

ggtcaggcca gtctctggac ttgtggctgc tgattacttt gctccaggct attttgtgtg 1020

4197737_1

ggctgttctg attgctaatt tggcacgcat tggatatgag gagaaaacaa tgttcatggc 1080

ctcatacgat tggagacttt catttcagaa cactgaggtc cgtgaccaaa cattaagccg 1140

gatgaagagt aatatagaac ttatggtttc tatcaatggt ggaaataagg cagttattgt 1200

tccacattcc atgggtgttt tgtactttct gcattttatg aagtgggttg aggcaccagc 1260

tccaatggga ggaggtggtg gaccagattg gtgtgctaag catatcaagg cagtcatgaa 1320

cattggtggt ccatttttag gtgttcccaa agctgttgct gggcttttct cggctgaagc 1380

aagagatatt gcagttgcca gggccatagc accaggtttc ttagataatg atatgttccg 1440

cctacaaaca ttgcaacaca tgatgaggat gtctcgcaca tgggattcga ccatgtcaat 1500

gataccaaga ggtggggaca ctatctgggg cgatcttgat tggtcacctg aagaaggtta 1560

cattcctaga aagaaaaggc agagaaataa tgcaactgat aatgtaaacg aaggtggggc 1620

cgaaagtgag atttctcaaa gaaagattgt tagatatgga agaatgatat catttgggaa 1680

aaatatagca gaggcacctt catatgatat tgaaaggatt gactttaggg atgctgttaa 1740

aggtcgtagt gtggcaaata atacctgcct tgatgtgtgg actgaatacc atgaaatggg 1800

attcggaggt attaaagccg ttgcagagta taaggtctac actgctggat ctactataga 1860

gctgcttcag tttgtcgccc caaaaatgat ggagcgtggt agtgctcatt tttcttatgg 1920

aattgctgac aatttggagg acccaaaata tgagcactac aaatactggt caaatcccct 1980

ggagacaaag ttacctaatg ctccagaaat ggaaatattt tccatgtatg gagttggcat 2040

accaacagaa agagcttatg tttatgagtt ttctcctgct gctgagtgct acattccatt 2100

tcagattgat acatcagcta atgatggcga tgaagatggc tgtctgaaag atggagtcta 2160

tactgttgat ggggatgaga ctgttcctgt tttaagtgca ggcttcatgt gtgctaaagc 2220

ttggcgtggg aaaaccagat ttaatccttc aggaagtcga acatacatta gagagtacga 2280

tcattctcct ccagctaatt tgctagaggg ccgaggcacc caaagtggtg cccatgttga 2340

tataatgggt aattttgctt taatcgagga tattatgagg gtggcagccg gggctacagg 2400

agaagatttg ggaggcgatc aagtgtattc agatatcttt aagtggtctc agaagatcaa 2460

attaccactg taa 2473

<210> 89
<211> 685
<212> PRT
<213> R. communis

<400> 89

Met Pro Val Ile Arg Arg Lys Lys Pro Thr Ser Glu Pro Asn Lys Asn
1 5 10 15

Ser Ala Ser Asp Ser Lys Thr Pro Ser Glu Glu Glu Glu His Glu Gln
20 25 30

Glu Gln Glu Gln Glu Glu Asp Lys Asn Asn Lys Lys Lys Tyr Pro Lys
35 40 45

```
Lys Lys Ser Ser Glu Ile Asn Ala Lys Lys Trp Ser Cys Ile Asp Ser
     50                  55                  60

Cys Cys Trp Phe Val Gly Cys Ile Cys Val Thr Trp Trp Val Leu Leu
65                  70                  75                  80

Phe Leu Tyr Asn Ala Val Pro Ala Ser Leu Pro Gln Tyr Val Thr Glu
                85                  90                  95

Ala Ile Thr Gly Pro Leu Pro Asp Pro Pro Gly Val Lys Leu Lys Lys
            100                 105                 110

Glu Gly Leu Thr Ala Lys His Pro Val Val Phe Val Pro Gly Ile Val
        115                 120                 125

Thr Ala Gly Leu Glu Leu Trp Glu Gly His Gln Cys Ala Asp Gly Leu
    130                 135                 140

Phe Arg Lys Arg Leu Trp Gly Gly Thr Phe Gly Glu Val Tyr Lys Arg
145                 150                 155                 160

Pro Leu Cys Trp Val Glu His Met Ser Leu Asp Asn Glu Thr Gly Leu
                165                 170                 175

Asp Pro Pro Gly Ile Arg Val Arg Pro Val Ser Gly Leu Val Ala Ala
            180                 185                 190

Asp Tyr Phe Ala Pro Gly Tyr Phe Val Trp Ala Val Leu Ile Ala Asn
        195                 200                 205

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

Asp Trp Arg Leu Ser Phe Gln Asn Thr Glu Val Arg Asp Gln Thr Leu
225                 230                 235                 240

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

Asn Lys Ala Val Ile Val Pro His Ser Met Gly Val Leu Tyr Phe Leu
            260                 265                 270

His Phe Met Lys Trp Val Glu Ala Pro Ala Pro Met Gly Gly Gly Gly
        275                 280                 285

Gly Pro Asp Trp Cys Ala Lys His Ile Lys Ala Val Met Asn Ile Gly
    290                 295                 300

Gly Pro Phe Leu Gly Val Pro Lys Ala Val Ala Gly Leu Phe Ser Ala
305                 310                 315                 320
```

```
Glu Ala Arg Asp Ile Ala Val Ala Arg Ala Ile Ala Pro Gly Phe Leu
                325                 330                 335

Asp Asn Asp Met Phe Arg Leu Gln Thr Leu Gln His Met Met Arg Met
            340                 345                 350

Ser Arg Thr Trp Asp Ser Thr Met Ser Met Ile Pro Arg Gly Gly Asp
        355                 360                 365

Thr Ile Trp Gly Asp Leu Asp Trp Ser Pro Glu Glu Gly Tyr Ile Pro
    370                 375                 380

Arg Lys Lys Arg Gln Arg Asn Asn Ala Thr Asp Asn Val Asn Glu Gly
385                 390                 395                 400

Gly Ala Glu Ser Glu Ile Ser Gln Arg Lys Ile Val Arg Tyr Gly Arg
                405                 410                 415

Met Ile Ser Phe Gly Lys Asn Ile Ala Glu Ala Pro Ser Tyr Asp Ile
            420                 425                 430

Glu Arg Ile Asp Phe Arg Asp Ala Val Lys Gly Arg Ser Val Ala Asn
        435                 440                 445

Asn Thr Cys Leu Asp Val Trp Thr Glu Tyr His Glu Met Gly Phe Gly
    450                 455                 460

Gly Ile Lys Ala Val Ala Glu Tyr Lys Val Tyr Thr Ala Gly Ser Thr
465                 470                 475                 480

Ile Glu Leu Leu Gln Phe Val Ala Pro Lys Met Met Glu Arg Gly Ser
                485                 490                 495

Ala His Phe Ser Tyr Gly Ile Ala Asp Asn Leu Glu Asp Pro Lys Tyr
            500                 505                 510

Glu His Tyr Lys Tyr Trp Ser Asn Pro Leu Glu Thr Lys Leu Pro Asn
        515                 520                 525

Ala Pro Glu Met Glu Ile Phe Ser Met Tyr Gly Val Gly Ile Pro Thr
    530                 535                 540

Glu Arg Ala Tyr Val Tyr Glu Phe Ser Pro Ala Ala Glu Cys Tyr Ile
545                 550                 555                 560

Pro Phe Gln Ile Asp Thr Ser Ala Asn Asp Gly Asp Glu Asp Gly Cys
                565                 570                 575

Leu Lys Asp Gly Val Tyr Thr Val Asp Gly Asp Glu Thr Val Pro Val
            580                 585                 590
```

Leu Ser Ala Gly Phe Met Cys Ala Lys Ala Trp Arg Gly Lys Thr Arg
595 600 605

Phe Asn Pro Ser Gly Ser Arg Thr Tyr Ile Arg Glu Tyr Asp His Ser
610 615 620

Pro Pro Ala Asn Leu Leu Glu Gly Arg Gly Thr Gln Ser Gly Ala His
625 630 635 640

Val Asp Ile Met Gly Asn Phe Ala Leu Ile Glu Asp Ile Met Arg Val
645 650 655

Ala Ala Gly Ala Thr Gly Glu Asp Leu Gly Gly Asp Gln Val Tyr Ser
660 665 670

Asp Ile Phe Lys Trp Ser Gln Lys Ile Lys Leu Pro Leu
675 680 685

<210> 90
<211> 380
<212> PRT
<213> B. oleraceae

<400> 90

Met Arg Asn Glu Ile Gln Asn Glu Thr Ala Gln Thr Asp Gln Thr Gln
1 5 10 15

Gly Ser Met Phe Ser Phe Phe Asp Leu Phe Pro Phe Leu Leu Pro Met
20 25 30

Phe Glu Val Ile Lys Met Val Val Ala Ser Val Ala Ser Val Val Tyr
35 40 45

Leu Gly Phe Ala Gly Val Thr Leu Ser Gly Ser Ala Val Ala Leu Ala
50 55 60

Val Ser Thr Pro Leu Phe Ile Ile Phe Ser Pro Ile Leu Leu Pro Ala
65 70 75 80

Ile Ala Ala Thr Thr Val Leu Ala Ala Gly Leu Gly Ser Lys Lys Val
85 90 95

Ala Ala Ala Pro Ala Ala Ser Pro Ser Leu Ser Leu Leu Gly Ile Pro
100 105 110

Glu Ser Ile Lys Pro Ser Asn Val Ile Pro Glu Ser Ile Lys Pro Ser
115 120 125

Asn Ile Ile Pro Glu Ser Ile Lys Pro Ser Asn Ile Ile Pro Glu Ser
130 135 140

Val Lys Pro Ser Asn Ile Lys Asp Lys Ile Lys Asp Thr Ile Gly Lys
145 150 155 160

Val Lys Asn Lys Ile Asn Ala Lys Lys Glu Glu Lys Ser Lys Gly Lys
165 170 175

Ser Glu Asp Ser Ser Lys Gly Lys Gly Lys Ser Lys Gly Glu Asp Thr
180 185 190

Thr Thr Asp Glu Asp Lys Pro Gly Ser Gly Gly Lys His Gly Lys Gly
195 200 205

Glu Ser Lys His Gly Lys Gly Glu Ser Thr His Gly Lys Gly Gly Lys
210 215 220

His Gly Ser Glu Gly Ser Ser Met Asp Glu Gly Lys His Gly Gly Lys
225 230 235 240

His Gly Ser Gly Gly Ser Pro Met Gly Val Gly Lys His Gly Ser Gly
245 250 255

Gly Lys His Glu Ser Gly Gly Ser Pro Met Gly Gly Gly Lys His Gly
260 265 270

Ser Gly Gly Lys His Glu Ser Gly Gly Ala Ser Met Gly Gly Gly Lys
275 280 285

His Gly Ser Gly Gly Arg His Glu Gly Gly Gly Ser Ala Met Gly Gly
290 295 300

Gly Lys His Gly Ser Gly Gly Lys His Gly Ser Glu Gly Lys His Gly
305 310 315 320

Gly Glu Gly Ser Ser Met Gly Lys Asn Ser Leu Ser Lys Asn Lys Lys
325 330 335

Glu Phe His Tyr Arg Asp Gln Ala Met Asp Ala Ser Ser Thr Ser Glu
340 345 350

Ser Ser Asp Gly Ser Ser Asp Gly Ser Ser Ser Asp Gly Ser Ser Ser
355 360 365

Asp Gly Ser Ser His Gly Ser Gly Gly Lys His Ile
370 375 380

<210> 91
<211> 186
<212> PRT
<213> Z. mays

<400> 91

Met Ala Asp Arg Asp Arg Ser Gly Ile Tyr Gly Gly Ala His Ala Thr
1 5 10 15

Tyr Gly Gln Gln Gln Gln Gly Gly Gly Gly Arg Pro Met Gly Glu
20 25 30

Gln Val Lys Gly Met Leu His Asp Lys Gly Pro Thr Ala Ser Gln Ala
35 40 45

Leu Thr Val Ala Thr Leu Phe Pro Leu Gly Gly Leu Leu Leu Val Leu
50 55 60

Ser Gly Leu Ala Leu Thr Ala Ser Val Val Gly Leu Ala Val Ala Thr
65 70 75 80

Pro Val Phe Leu Ile Phe Ser Pro Val Leu Val Pro Ala Ala Leu Leu
85 90 95

Ile Gly Thr Ala Val Met Gly Phe Leu Thr Ser Gly Ala Leu Gly Leu
100 105 110

Gly Gly Leu Ser Ser Leu Thr Cys Leu Ala Asn Thr Ala Arg Gln Ala
115 120 125

Phe Gln Arg Thr Pro Asp Tyr Val Glu Glu Ala His Arg Arg Met Ala
130 135 140

Glu Ala Ala Ala His Ala Gly His Lys Thr Ala Gln Ala Gly Gln Ala
145 150 155 160

Ile Gln Gly Arg Ala Gln Glu Ala Gly Ala Gly Gly Gly Ala Gly Ala
165 170 175

Gly Ala Gly Gly Gly Gly Arg Ala Ser Ser
180 185

<210> 92
<211> 148
<212> PRT
<213> O. sativa

<400> 92

Met Gly Asp Gln His Arg Gly Val Ile Gly Gly Gly Gly Tyr Gly Asp
1 5 10 15

Arg Gly Gly Gln Glu Gln Gln Glu Lys Gln Pro Phe Met Met Thr Ala
20 25 30

Leu Lys Thr Val Thr Ala Ala Thr Ala Gly Gly Ser Ile Leu Val Leu
35 40 45

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

Pro Val Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Ala Ile Ala
65 70 75 80

```
Leu Ala Leu Met Ala Ala Gly Phe Val Thr Ser Val Gly Leu Gly Val
                85                  90                  95

Ala Ala Leu Ser Val Phe Ser Trp Met Tyr Lys Tyr Leu Thr Gly Lys
            100                 105                 110

His Pro Pro Gly Ala Asp His Leu Asp His Thr Lys Ala Arg Val Ala
        115                 120                 125

Ser Lys Leu Arg Asp Ile Lys Glu Ala Ala His His Leu Ile Asp Gln
    130                 135                 140

Ala Gln Ala Ser
145
```

<210> 93
<211> 245
<212> PRT
<213> S. indicum

<400> 93

```
Met Ala Thr His Val Leu Ala Ala Ala Ala Glu Arg Asn Ala Ala Leu
1               5                   10                  15

Ala Pro Asp Ala Pro Leu Ala Pro Val Thr Met Glu Arg Pro Val Arg
            20                  25                  30

Thr Asp Leu Glu Thr Ser Ile Pro Lys Pro Tyr Met Ala Arg Gly Leu
        35                  40                  45

Val Ala Pro Asp Met Asp His Pro Asn Gly Thr Pro Gly His Val His
    50                  55                  60

Asp Asn Leu Ser Val Leu Gln Gln His Cys Ala Phe Phe Asp Gln Asp
65                  70                  75                  80

Asp Asn Gly Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Gln
                85                  90                  95

Ile Gly Phe Asn Val Ile Ala Ser Leu Ile Met Ala Ile Val Ile Asn
            100                 105                 110

Val Ala Leu Ser Tyr Pro Thr Leu Pro Gly Trp Ile Pro Ser Pro Phe
        115                 120                 125

Phe Pro Ile Tyr Leu Tyr Asn Ile His Lys Ala Lys His Gly Ser Asp
    130                 135                 140

Ser Gly Thr Tyr Asp Thr Glu Gly Arg Tyr Leu Pro Met Asn Phe Glu
145                 150                 155                 160
```

4197737_1

Asn Leu Phe Ser Lys His Ala Arg Thr Met Pro Asp Arg Leu Thr Leu
165 170 175

Gly Glu Leu Trp Ser Met Thr Glu Ala Asn Arg Glu Ala Phe Asp Ile
180 185 190

Phe Gly Trp Ile Ala Ser Lys Met Glu Trp Thr Leu Leu Tyr Ile Leu
195 200 205

Ala Arg Asp Gln Asp Gly Phe Leu Ser Lys Glu Ala Ile Arg Arg Cys
210 215 220

Tyr Asp Gly Ser Leu Phe Glu Tyr Cys Ala Lys Met Gln Arg Gly Ala
225 230 235 240

Glu Asp Lys Met Lys
245

<210> 94
<211> 239
<212> PRT
<213> G. max

<400> 94

Met Ala Ala Glu Met Glu Arg Glu Ser Leu Ile Thr Glu Ala Pro Asn
1 5 10 15

Ala Pro Val Thr Ala Gln Arg Arg Val Arg Asn Asp Leu Glu Asn Ser
20 25 30

Leu Pro Lys Pro Tyr Leu Pro Arg Ala Leu Lys Ala Pro Asp Thr Gly
35 40 45

His Pro Asn Gly Thr Ala Gly His Arg His His Asn Leu Ser Val Leu
50 55 60

Gln Gln His Cys Ala Phe Phe Asp Gln Asp Asp Asn Gly Ile Ile Tyr
65 70 75 80

Pro Trp Glu Thr Tyr Met Gly Leu Arg Ser Ile Gly Phe Asn Val Val
85 90 95

Ala Ser Val Ile Met Ala Ile Val Ile Asn Val Gly Leu Ser Tyr Pro
100 105 110

Thr Leu Pro Asn Trp Phe Pro Ser Leu Leu Phe Pro Ile Tyr Ile His
115 120 125

Asn Ile His Lys Ala Lys His Gly Ser Asp Ser Gly Val Tyr Asp Thr
130 135 140

Glu Gly Arg Tyr Val Pro Ala Asn Ile Glu Asn Ile Phe Ser Lys Tyr
145 150 155 160

```
Ala Arg Thr Val Pro Asp Lys Leu Thr Leu Gly Glu Leu Trp Asp Leu
                165                 170                 175

Thr Glu Gly Asn Arg Asn Ala Phe Asp Ile Phe Gly Trp Leu Ala Ala
            180                 185                 190

Lys Phe Glu Trp Gly Val Leu Tyr Ile Leu Ala Arg Asp Glu Glu Gly
        195                 200                 205

Phe Leu Ser Lys Glu Ala Val Arg Arg Cys Phe Asp Gly Ser Leu Phe
    210                 215                 220

Glu Tyr Cys Ala Lys Met His Thr Thr Ser Asp Ala Lys Met Ser
225                 230                 235
```

<210> 95
<211> 243
<212> PRT
<213> Z. mays

<400> 95

```
Met Ser Ser Tyr Ser Pro Pro Pro Pro Pro Pro Arg Asp Gln Ser Met
1               5                   10                  15

Asp Thr Glu Ala Pro Asn Ala Pro Ile Thr Arg Glu Arg Arg Leu Asn
            20                  25                  30

Pro Asp Leu Gln Glu Gln Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu
        35                  40                  45

Glu Ala Val Asp Pro Ser His Pro Gln Gly Thr Lys Gly Arg Asp Pro
    50                  55                  60

Arg Gly Met Ser Val Leu Gln Gln His Ala Ala Phe Phe Asp Arg Asn
65                  70                  75                  80

Gly Asp Gly Val Ile Tyr Pro Trp Glu Thr Phe Gln Gly Leu Arg Ala
                85                  90                  95

Ile Gly Cys Gly Leu Thr Val Ser Phe Ala Phe Ser Ile Leu Ile Asn
            100                 105                 110

Leu Phe Leu Ser Tyr Pro Thr Gln Pro Gly Trp Leu Pro Ser Pro Leu
        115                 120                 125

Leu Ser Ile Arg Ile Asp Asn Ile His Lys Gly Lys His Gly Ser Asp
    130                 135                 140

Ser Glu Thr Tyr Asp Thr Glu Gly Arg Phe Asp Pro Ser Lys Phe Asp
145                 150                 155                 160
```

4197737_1

```
Ala Ile Phe Ser Lys Tyr Gly Arg Thr His Pro Asn Ala Ile Thr Arg
                165                 170                 175

Asp Glu Leu Ser Ser Met Leu Gln Gly Asn Arg Asn Thr Tyr Asp Phe
            180                 185                 190

Leu Gly Trp Leu Ala Ala Ala Gly Glu Trp Leu Leu Leu Tyr Ser Leu
        195                 200                 205

Ala Lys Asp Lys Asp Gly Leu Leu Gln Arg Glu Thr Val Arg Gly Leu
    210                 215                 220

Phe Asp Gly Ser Leu Phe Glu Arg Leu Glu Asp Asp Asn Asn Lys Lys
225                 230                 235                 240

Lys Ser Ser
```

<210> 96
<211> 348
<212> PRT
<213> S. indicum

<400> 96

```
Met Asp Leu Ile His Thr Phe Leu Asn Leu Ile Ala Pro Pro Phe Thr
1               5                   10                  15

Phe Phe Phe Leu Leu Phe Phe Leu Pro Pro Phe Gln Ile Phe Lys Phe
            20                  25                  30

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

Val Val Val Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ser Leu Ala
    50                  55                  60

Tyr Glu Tyr Ala Lys Arg Gly Ala Cys Leu Val Leu Ala Ala Arg Arg
65                  70                  75                  80

Glu Arg Ser Leu Gln Glu Val Ala Glu Arg Ala Arg Asp Leu Gly Ser
                85                  90                  95

Pro Asp Val Val Val Val Arg Ala Asp Val Ser Lys Ala Glu Asp Cys
            100                 105                 110

Arg Lys Val Val Asp Gln Thr Met Asn Arg Phe Gly Arg Leu Asp His
        115                 120                 125

Leu Val Asn Asn Ala Gly Ile Met Ser Val Ser Met Leu Glu Glu Val
    130                 135                 140

Glu Asp Ile Thr Gly Tyr Arg Glu Thr Met Asp Ile Asn Phe Trp Gly
145                 150                 155                 160
```

```
Tyr Val Tyr Met Thr Arg Phe Ala Ala Pro Tyr Leu Arg Asn Ser Arg
                165                 170                 175

Gly Arg Ile Val Val Leu Ser Ser Ser Ser Ser Trp Met Pro Thr Pro
            180                 185                 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Ile Ser Gln Phe Phe
        195                 200                 205

Glu Thr Leu Arg Val Glu Phe Gly Pro Asp Ile Gly Ile Thr Leu Val
    210                 215                 220

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

Ala Gly Glu Arg Val Ile Asp Gln Asp Met Arg Asp Val Gln Val Ser
                245                 250                 255

Thr Thr Pro Ile Leu Arg Val Glu Ser Ala Ala Arg Ser Ile Val Arg
            260                 265                 270

Ser Ala Ile Arg Gly Glu Arg Tyr Val Thr Glu Pro Ala Trp Phe Arg
        275                 280                 285

Val Thr Tyr Trp Trp Lys Leu Phe Cys Pro Glu Val Met Glu Trp Val
    290                 295                 300

Phe Arg Leu Met Tyr Leu Ala Ser Pro Gly Glu Pro Glu Lys Glu Thr
305                 310                 315                 320

Phe Gly Lys Lys Val Leu Asp Tyr Thr Gly Val Lys Ser Leu Leu Tyr
                325                 330                 335

Pro Glu Thr Val Gln Val Pro Glu Pro Lys Asn Asp
            340                 345
```

<210> 97
<211> 349
<212> PRT
<213> Brassica napus

<400> 97

```
Met Glu Leu Ile Asn Asp Phe Leu Asn Leu Thr Ala Pro Phe Phe Thr
1               5                   10                  15

Phe Phe Gly Leu Cys Phe Phe Leu Pro Pro Phe Tyr Phe Phe Lys Phe
            20                  25                  30

Val Gln Ser Ile Phe Ser Thr Ile Phe Ser Glu Asn Val Tyr Gly Lys
        35                  40                  45
```

```
Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Gln Leu Ala
    50                  55                  60

Tyr Glu Tyr Ala Ser Lys Gly Ala Cys Leu Ala Leu Thr Ala Arg Arg
65                  70                  75                  80

Lys Asn Arg Leu Glu Glu Val Ala Glu Ile Ala Arg Glu Val Gly Ser
                85                  90                  95

Pro Asn Val Val Thr Val His Ala Asp Val Ser Lys Pro Asp Asp Cys
            100                 105                 110

Arg Arg Ile Val Asp Glu Thr Ile Ser His Phe Gly Arg Leu Asp His
        115                 120                 125

Leu Val Asn Asn Ala Gly Ile Met Gln Ile Ser Met Phe Glu Asn Ile
    130                 135                 140

Glu Glu Ile Thr Arg Thr Arg Ala Val Met Asp Thr Asn Phe Trp Gly
145                 150                 155                 160

Ala Val Tyr Thr Thr Arg Ala Ala Leu Pro Tyr Leu Arg Gln Ser Asn
                165                 170                 175

Gly Lys Ile Val Ala Met Ser Ser Ser Ala Ala Trp Leu Thr Ala Pro
            180                 185                 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Leu Leu Asn Phe Phe
        195                 200                 205

Glu Thr Leu Arg Ile Glu Leu Gly Ser Asp Val His Ile Thr Ile Val
    210                 215                 220

Thr Pro Gly Tyr Ile Glu Ser Glu Leu Thr Gln Gly Lys Tyr Val Ser
225                 230                 235                 240

Gly Glu Gly Glu Leu Val Val Asn Gln Asp Ile Arg Asp Val Gln Ile
                245                 250                 255

Gly Ala Phe Pro Val Thr Ser Val Ser Gly Arg Ala Lys Gly Ile Val
            260                 265                 270

Lys Gly Val Cys Arg Lys Glu Arg Tyr Val Thr Glu Pro Ser Trp Phe
        275                 280                 285

Lys Val Thr Tyr Leu Trp Lys Val Phe Cys Pro Glu Leu Ile Glu Trp
    290                 295                 300

Gly Cys Arg Leu Met Phe Leu Ser Gly His Gly Thr Pro Glu Glu Asn
305                 310                 315                 320

Ala Leu Asn Lys Lys Ile Leu Asp Ile Pro Gly Val Arg Ser Ala Leu
```

325 330 335

Tyr Pro Glu Pro Ile Arg Thr Pro Glu Ile Lys Ser Glu
340 345

<210> 98
<211> 350
<212> PRT
<213> Zea mays

<400> 98

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

Ala Leu Thr Ala Leu Leu Pro Leu Val Leu Pro Ala Tyr Tyr Val Tyr
20 25 30

Lys Leu Thr Thr Tyr Leu Leu Gly Ala Val Phe Pro Glu Asp Val Ala
35 40 45

Gly Lys Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu His
50 55 60

Leu Ala Tyr Glu Tyr Ala Lys Arg Gly Ala Tyr Leu Ala Leu Val Ala
65 70 75 80

Arg Arg Glu Ala Ser Leu Arg Glu Val Gly Asp Val Ala Leu Gly Leu
85 90 95

Gly Ser Pro Gly Val Leu Val Leu Pro Ala Asp Val Ser Lys Pro Arg
100 105 110

Asp Cys Glu Gly Phe Ile Asp Asp Thr Ile Ser Tyr Phe Gly Arg Leu
115 120 125

Asp His Leu Val Asn Asn Ala Ser Ile Trp Gln Val Cys Lys Phe Glu
130 135 140

Glu Ile Gln Asp Val Arg His Leu Arg Ala Leu Met Asp Ile Asn Phe
145 150 155 160

Trp Gly His Val Tyr Pro Thr Arg Leu Ala Ile Pro His Leu Arg Arg
165 170 175

Ser Arg Gly Arg Ile Val Gly Val Thr Ser Asn Ser Ser Tyr Ile Phe
180 185 190

Ile Gly Arg Asn Thr Phe Tyr Asn Ala Ser Lys Ala Ala Ala Leu Ser
195 200 205

Phe Tyr Asp Thr Leu Arg Met Glu Leu Gly Ser Asp Ile Arg Ile Thr
210 215 220

```
Glu Val Val Pro Gly Val Val Glu Ser Glu Ile Thr Lys Gly Lys Met
225                 230                 235                 240

Leu Thr Lys Gly Gly Glu Met Lys Val Asp Gln Asp Glu Arg Asp Ala
                245                 250                 255

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

Val Val Arg Asp Val Cys Arg Gly Ala Arg Tyr Val Phe Glu Pro Arg
        275                 280                 285

Trp Tyr Met Gly Val Tyr Leu Leu Arg Ala Cys Leu Pro Glu Val Leu
    290                 295                 300

Ala Trp Asn Ser Arg Leu Leu Thr Val Asp Thr Val Gly Ala Ser Thr
305                 310                 315                 320

Thr Asp Thr Leu Gly Lys Trp Leu Val Glu Leu Pro Gly Val Arg Arg
                325                 330                 335

Val Val Gln Pro Pro Ser Leu Arg Ser Pro Glu Ile Lys Asp
            340                 345                 350
```

<210> 99
<211> 380
<212> PRT
<213> Brassica oleracea

<400> 99

```
Met Arg Asn Cys Ile Gln Asn Glu Thr Ala Gln Thr Cys Gln Thr Gln
1               5                   10                  15

Gly Ser Met Phe Ser Phe Phe Cys Leu Phe Pro Phe Leu Leu Pro Met
            20                  25                  30

Phe Glu Val Ile Lys Met Val Val Ala Ser Val Ala Ser Val Val Tyr
        35                  40                  45

Leu Gly Phe Ala Gly Val Thr Leu Ser Gly Ser Ala Val Ala Leu Ala
    50                  55                  60

Val Ser Thr Pro Leu Phe Ile Ile Phe Ser Pro Ile Leu Leu Pro Ala
65                  70                  75                  80

Ile Ala Ala Thr Thr Val Leu Ala Ala Gly Leu Gly Ser Lys Lys Val
                85                  90                  95

Ala Ala Ala Pro Ala Ala Ser Pro Ser Leu Ser Leu Leu Gly Ile Pro
            100                 105                 110

Glu Ser Ile Lys Pro Ser Asn Val Ile Pro Glu Ser Ile Lys Pro Ser
```

```
        115                 120                 125

Asn Ile Ile Pro Glu Ser Ile Lys Pro Ser Asn Ile Ile Pro Glu Ser
    130                 135                 140

Val Lys Pro Ser Asn Ile Lys Asp Lys Ile Lys Asp Thr Ile Gly Lys
145                 150                 155                 160

Val Lys Asn Lys Ile Asn Ala Lys Lys Glu Glu Lys Ser Lys Gly Lys
                165                 170                 175

Ser Glu Asp Ser Ser Lys Gly Lys Gly Lys Ser Lys Gly Glu Asp Thr
            180                 185                 190

Thr Thr Asp Glu Asp Lys Pro Gly Ser Gly Gly Lys His Gly Lys Gly
        195                 200                 205

Glu Ser Lys His Gly Lys Gly Glu Ser Thr His Gly Lys Gly Gly Lys
    210                 215                 220

His Gly Ser Glu Gly Ser Ser Met Asp Glu Gly Lys His Gly Gly Lys
225                 230                 235                 240

His Gly Ser Gly Gly Ser Pro Met Gly Val Gly Lys His Gly Ser Gly
                245                 250                 255

Gly Lys His Glu Ser Gly Gly Ser Pro Met Gly Gly Gly Lys His Gly
            260                 265                 270

Ser Gly Gly Lys His Glu Ser Gly Gly Ala Ser Met Gly Gly Gly Lys
        275                 280                 285

His Gly Ser Gly Gly Arg His Glu Gly Gly Gly Ser Ala Met Gly Gly
    290                 295                 300

Gly Lys His Gly Ser Gly Gly Lys His Gly Ser Glu Gly Lys His Gly
305                 310                 315                 320

Gly Glu Gly Ser Ser Met Gly Lys Asn Ser Leu Ser Lys Asn Lys Lys
                325                 330                 335

Glu Phe His Tyr Arg Asp Gln Ala Met Cys Ala Ser Ser Thr Ser Glu
            340                 345                 350

Ser Ser Asp Gly Ser Ser Cys Gly Ser Ser Ser Asp Gly Ser Ser Ser
        355                 360                 365

Cys Gly Ser Ser His Gly Ser Gly Gly Lys His Ile
    370                 375                 380

<210>  100
<211>  186
```

<212> PRT
<213> Zea mays

<400> 100

Met Ala Asp Arg Cys Arg Ser Gly Ile Tyr Gly Gly Ala His Ala Thr
1 5 10 15

Tyr Gly Gln Gln Gln Gln Gly Gly Gly Gly Arg Pro Met Gly Cys
20 25 30

Gln Val Lys Gly Met Leu His Cys Lys Gly Pro Thr Ala Ser Gln Ala
35 40 45

Leu Thr Val Ala Thr Leu Phe Pro Leu Gly Gly Leu Leu Leu Val Leu
50 55 60

Ser Gly Leu Ala Leu Thr Ala Ser Val Val Gly Leu Ala Val Ala Thr
65 70 75 80

Pro Val Phe Leu Ile Phe Ser Pro Val Leu Val Pro Ala Ala Leu Leu
85 90 95

Ile Gly Thr Ala Val Met Gly Phe Leu Thr Ser Gly Ala Leu Gly Leu
100 105 110

Gly Gly Leu Ser Ser Leu Thr Cys Leu Ala Asn Thr Ala Arg Gln Ala
115 120 125

Phe Gln Arg Thr Pro Asp Tyr Val Cys Glu Ala His Arg Arg Met Ala
130 135 140

Cys Ala Ala Ala His Ala Gly His Lys Thr Ala Gln Ala Gly Gln Ala
145 150 155 160

Ile Gln Gly Arg Ala Gln Cys Ala Gly Ala Gly Gly Gly Ala Gly Ala
165 170 175

Gly Ala Gly Gly Gly Gly Arg Ala Ser Ser
180 185

<210> 101
<211> 148
<212> PRT
<213> Oryza sativa

<400> 101

Met Gly Asp Gln His Cys Gly Val Ile Gly Gly Gly Gly Tyr Gly Cys
1 5 10 15

Arg Gly Gly Gln Glu Gln Gln Cys Lys Gln Pro Phe Met Met Thr Ala
20 25 30

Leu Lys Thr Val Thr Ala Ala Thr Ala Gly Gly Ser Ile Leu Val Leu

35 40 45

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

Pro Val Leu Val Ile Phe Ser Pro Val Leu Val Pro Ala Ala Ile Ala
65 70 75 80

Leu Ala Leu Met Ala Ala Gly Phe Val Thr Ser Val Gly Leu Gly Val
85 90 95

Ala Ala Leu Ser Val Phe Ser Trp Met Tyr Lys Tyr Leu Thr Gly Lys
100 105 110

His Pro Pro Gly Ala Asp His Leu Cys His Thr Lys Ala Arg Val Ala
115 120 125

Ser Lys Leu Arg Cys Ile Lys Glu Ala Ala His His Leu Ile Cys Gln
130 135 140

Ala Gln Ala Ser
145

<210> 102
<211> 245
<212> PRT
<213> S. indicum

<400> 102

Met Ala Thr His Val Leu Ala Ala Ala Ala Cys Arg Asn Ala Ala Leu
1 5 10 15

Ala Pro Cys Ala Pro Leu Ala Pro Val Thr Met Cys Arg Pro Val Arg
20 25 30

Thr Asp Leu Glu Thr Ser Ile Pro Lys Pro Tyr Met Ala Arg Gly Leu
35 40 45

Val Ala Pro Asp Met Asp His Pro Asn Gly Thr Pro Gly His Val His
50 55 60

Asp Asn Leu Ser Val Leu Gln Gln His Cys Ala Phe Phe Asp Gln Asp
65 70 75 80

Asp Asn Gly Ile Ile Tyr Pro Trp Glu Thr Tyr Ser Gly Leu Arg Gln
85 90 95

Ile Gly Phe Asn Val Ile Ala Ser Leu Ile Met Ala Ile Val Ile Asn
100 105 110

Val Ala Leu Ser Tyr Pro Thr Leu Pro Gly Trp Ile Pro Ser Pro Phe
115 120 125

```
Phe Pro Ile Tyr Leu Tyr Asn Ile His Lys Ala Lys His Gly Ser Asp
    130                 135                 140

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

Asn Leu Phe Ser Lys His Ala Arg Thr Met Pro Asp Arg Leu Thr Leu
                165                 170                 175

Gly Glu Leu Trp Ser Met Thr Glu Ala Asn Arg Glu Ala Phe Asp Ile
            180                 185                 190

Phe Gly Trp Ile Ala Ser Lys Met Glu Trp Thr Leu Leu Tyr Ile Leu
        195                 200                 205

Ala Arg Cys Gln Asp Gly Phe Leu Ser Cys Glu Ala Ile Arg Arg Cys
    210                 215                 220

Tyr Asp Gly Ser Leu Phe Glu Tyr Cys Ala Lys Met Gln Arg Gly Ala
225                 230                 235                 240

Glu Asp Lys Met Cys
                245
```

<210> 103
<211> 239
<212> PRT
<213> G. soja

<400> 103

```
Met Ala Ala Cys Met Glu Arg Glu Ser Leu Ile Thr Cys Ala Pro Asn
1               5                   10                  15

Ala Pro Val Thr Ala Gln Arg Cys Val Arg Asn Asp Leu Glu Asn Ser
            20                  25                  30

Leu Pro Lys Pro Tyr Leu Pro Arg Ala Leu Lys Ala Pro Asp Thr Gly
        35                  40                  45

His Pro Asn Gly Thr Ala Gly His Arg His His Asn Leu Ser Val Leu
    50                  55                  60

Gln Gln His Cys Ala Phe Phe Asp Gln Asp Asp Asn Gly Ile Ile Tyr
65                  70                  75                  80

Pro Trp Glu Thr Tyr Met Gly Leu Arg Ser Ile Gly Phe Asn Val Val
                85                  90                  95

Ala Ser Val Ile Met Ala Ile Val Ile Asn Val Gly Leu Ser Tyr Pro
            100                 105                 110

Thr Leu Pro Asn Trp Phe Pro Ser Leu Leu Phe Pro Ile Tyr Ile His
```

115 120 125

Asn Ile His Lys Ala Lys His Gly Ser Asp Ser Gly Val Tyr Asp Thr
130 135 140

Glu Gly Arg Tyr Val Pro Ala Asn Ile Glu Asn Ile Phe Ser Lys Tyr
145 150 155 160

Ala Arg Thr Val Pro Asp Lys Leu Thr Leu Gly Glu Leu Trp Asp Leu
165 170 175

Thr Glu Gly Asn Arg Asn Ala Phe Asp Ile Phe Gly Trp Leu Ala Ala
180 185 190

Lys Phe Glu Trp Gly Val Leu Tyr Ile Leu Ala Arg Cys Glu Glu Gly
195 200 205

Phe Leu Ser Lys Cys Ala Val Arg Arg Cys Phe Asp Gly Ser Leu Phe
210 215 220

Glu Tyr Cys Ala Lys Met His Thr Thr Ser Asp Ala Cys Met Ser
225 230 235

<210> 104
<211> 243
<212> PRT
<213> Z. mays

<400> 104

Met Ser Ser Tyr Ser Pro Pro Pro Pro Pro Pro Arg Asp Gln Ser Met
1 5 10 15

Asp Thr Cys Ala Pro Asn Ala Pro Ile Thr Cys Glu Arg Arg Leu Asn
20 25 30

Pro Asp Leu Gln Cys Gln Leu Pro Lys Pro Tyr Leu Ala Arg Ala Leu
35 40 45

Glu Ala Val Asp Pro Ser His Pro Gln Gly Thr Lys Gly Arg Asp Pro
50 55 60

Arg Gly Met Ser Val Leu Gln Gln His Ala Ala Phe Phe Asp Arg Asn
65 70 75 80

Gly Asp Gly Val Ile Tyr Pro Trp Glu Thr Phe Gln Gly Leu Arg Ala
85 90 95

Ile Gly Cys Gly Leu Thr Val Ser Phe Ala Phe Ser Ile Leu Ile Asn
100 105 110

Leu Phe Leu Ser Tyr Pro Thr Gln Pro Gly Trp Leu Pro Ser Pro Leu
115 120 125

```
Leu Ser Ile Arg Ile Asp Asn Ile His Lys Gly Lys His Gly Ser Asp
    130                 135                 140

Ser Glu Thr Tyr Asp Thr Glu Gly Arg Phe Asp Pro Ser Lys Phe Asp
145                 150                 155                 160

Ala Ile Phe Ser Lys Tyr Gly Arg Thr His Pro Asn Ala Ile Thr Arg
                165                 170                 175

Asp Glu Leu Ser Ser Met Leu Gln Gly Asn Arg Asn Thr Tyr Asp Phe
            180                 185                 190

Leu Gly Trp Leu Ala Ala Ala Gly Glu Trp Leu Leu Leu Tyr Ser Leu
        195                 200                 205

Ala Lys Asp Lys Asp Gly Leu Leu Gln Arg Cys Thr Val Arg Gly Leu
    210                 215                 220

Phe Asp Gly Ser Leu Phe Cys Arg Leu Glu Asp Asp Asn Asn Lys Lys
225                 230                 235                 240

Cys Ser Ser
```

<210> 105
<211> 348
<212> PRT
<213> S. indicum

<400> 105

```
Met Asp Leu Ile Cys Thr Phe Leu Asn Leu Ile Ala Pro Pro Phe Thr
1               5                   10                  15

Phe Phe Phe Leu Leu Phe Phe Leu Pro Pro Phe Cys Ile Phe Lys Phe
            20                  25                  30

Phe Leu Cys Ile Leu Gly Thr Leu Phe Ser Glu Asp Val Ala Gly Lys
        35                  40                  45

Val Val Val Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ser Leu Ala
    50                  55                  60

Tyr Glu Tyr Ala Lys Arg Gly Ala Cys Leu Val Leu Ala Ala Arg Arg
65                  70                  75                  80

Glu Arg Ser Leu Gln Glu Val Ala Glu Arg Ala Arg Asp Leu Gly Ser
                85                  90                  95

Pro Asp Val Val Val Val Arg Ala Asp Val Ser Lys Ala Glu Asp Cys
            100                 105                 110

Arg Lys Val Val Asp Gln Thr Met Asn Arg Phe Gly Arg Leu Asp His
```

115 120 125

Leu Val Asn Asn Ala Gly Ile Met Ser Val Ser Met Leu Glu Glu Val
130 135 140

Glu Asp Ile Thr Gly Tyr Arg Glu Thr Met Asp Ile Asn Phe Trp Gly
145 150 155 160

Tyr Val Tyr Met Thr Arg Phe Ala Ala Pro Tyr Leu Arg Asn Ser Arg
165 170 175

Gly Arg Ile Val Val Leu Ser Ser Ser Ser Ser Trp Met Pro Thr Pro
180 185 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Ile Ser Gln Phe Phe
195 200 205

Glu Thr Leu Arg Val Glu Phe Gly Pro Asp Ile Gly Ile Thr Leu Val
210 215 220

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

Ala Gly Glu Arg Val Ile Asp Gln Asp Met Arg Asp Val Gln Val Ser
245 250 255

Thr Thr Pro Ile Leu Arg Val Glu Ser Ala Ala Arg Ser Ile Val Arg
260 265 270

Ser Ala Ile Arg Gly Glu Arg Tyr Val Thr Glu Pro Ala Trp Phe Arg
275 280 285

Val Thr Tyr Trp Trp Lys Leu Phe Cys Pro Glu Val Met Glu Trp Val
290 295 300

Phe Arg Leu Met Tyr Leu Ala Ser Pro Gly Cys Pro Glu Lys Cys Thr
305 310 315 320

Phe Gly Lys Lys Val Leu Cys Tyr Thr Gly Val Lys Ser Leu Leu Tyr
325 330 335

Pro Glu Thr Val Gln Val Pro Glu Pro Lys Asn Asp
340 345

<210> 106
<211> 349
<212> PRT
<213> Brassica napus

<400> 106

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

```
Phe Phe Gly Leu Cys Phe Phe Leu Pro Pro Phe Tyr Phe Phe Cys Phe
            20                  25                  30

Val Gln Cys Ile Phe Ser Thr Ile Phe Ser Glu Asn Val Tyr Gly Lys
        35                  40                  45

Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Gln Leu Ala
    50                  55                  60

Tyr Glu Tyr Ala Ser Lys Gly Ala Cys Leu Ala Leu Thr Ala Arg Arg
65                  70                  75                  80

Lys Asn Arg Leu Glu Glu Val Ala Glu Ile Ala Arg Glu Val Gly Ser
                85                  90                  95

Pro Asn Val Val Thr Val His Ala Asp Val Ser Lys Pro Asp Asp Cys
            100                 105                 110

Arg Arg Ile Val Asp Glu Thr Ile Ser His Phe Gly Arg Leu Asp His
        115                 120                 125

Leu Val Asn Asn Ala Gly Ile Met Gln Ile Ser Met Phe Glu Asn Ile
    130                 135                 140

Glu Glu Ile Thr Arg Thr Arg Ala Val Met Asp Thr Asn Phe Trp Gly
145                 150                 155                 160

Ala Val Tyr Thr Thr Arg Ala Ala Leu Pro Tyr Leu Arg Gln Ser Asn
                165                 170                 175

Gly Lys Ile Val Ala Met Ser Ser Ser Ala Ala Trp Leu Thr Ala Pro
            180                 185                 190

Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Leu Leu Asn Phe Phe
        195                 200                 205

Glu Thr Leu Arg Ile Glu Leu Gly Ser Asp Val His Ile Thr Ile Val
    210                 215                 220

Thr Pro Gly Tyr Ile Glu Ser Glu Leu Thr Gln Gly Lys Tyr Val Ser
225                 230                 235                 240

Gly Glu Gly Glu Leu Val Val Asn Gln Asp Ile Arg Asp Val Gln Ile
                245                 250                 255

Gly Ala Phe Pro Val Thr Ser Val Ser Gly Arg Ala Lys Gly Ile Val
            260                 265                 270

Lys Gly Val Cys Arg Lys Glu Arg Tyr Val Thr Glu Pro Ser Trp Phe
        275                 280                 285
```

4197737_1

Lys Val Thr Tyr Leu Trp Lys Val Phe Cys Pro Glu Leu Ile Glu Trp
290 295 300

Gly Cys Arg Leu Met Phe Leu Ser Gly His Gly Thr Pro Cys Glu Asn
305 310 315 320

Ala Leu Asn Lys Lys Ile Leu Asp Ile Pro Gly Val Arg Ser Ala Leu
325 330 335

Tyr Pro Cys Pro Ile Arg Thr Pro Cys Ile Lys Ser Glu
340 345

<210> 107
<211> 349
<212> PRT
<213> Z. mays

<400> 107

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

Phe Phe Gly Leu Cys Phe Phe Leu Pro Pro Phe Tyr Phe Phe Cys Phe
20 25 30

Val Gln Cys Ile Phe Ser Thr Ile Phe Ser Glu Asn Val Tyr Gly Lys
35 40 45

Val Val Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Gln Leu Ala
50 55 60

Tyr Glu Tyr Ala Ser Lys Gly Ala Cys Leu Ala Leu Thr Ala Arg Arg
65 70 75 80

Lys Asn Arg Leu Glu Glu Val Ala Glu Ile Ala Arg Glu Val Gly Ser
85 90 95

Pro Asn Val Val Thr Val His Ala Asp Val Ser Lys Pro Asp Asp Cys
100 105 110

Arg Arg Ile Val Asp Glu Thr Ile Ser His Phe Gly Arg Leu Asp His
115 120 125

Leu Val Asn Asn Ala Gly Ile Met Gln Ile Ser Met Phe Glu Asn Ile
130 135 140

Glu Glu Ile Thr Arg Thr Arg Ala Val Met Asp Thr Asn Phe Trp Gly
145 150 155 160

Ala Val Tyr Thr Thr Arg Ala Ala Leu Pro Tyr Leu Arg Gln Ser Asn
165 170 175

Gly Lys Ile Val Ala Met Ser Ser Ser Ala Ala Trp Leu Thr Ala Pro
180 185 190

```
Arg Met Ser Phe Tyr Asn Ala Ser Lys Ala Ala Leu Leu Asn Phe Phe
        195                 200                 205

Glu Thr Leu Arg Ile Glu Leu Gly Ser Asp Val His Ile Thr Ile Val
    210                 215                 220

Thr Pro Gly Tyr Ile Glu Ser Glu Leu Thr Gln Gly Lys Tyr Val Ser
225                 230                 235                 240

Gly Glu Gly Glu Leu Val Val Asn Gln Asp Ile Arg Asp Val Gln Ile
                245                 250                 255

Gly Ala Phe Pro Val Thr Ser Val Ser Gly Arg Ala Lys Gly Ile Val
            260                 265                 270

Lys Gly Val Cys Arg Lys Glu Arg Tyr Val Thr Glu Pro Ser Trp Phe
        275                 280                 285

Lys Val Thr Tyr Leu Trp Lys Val Phe Cys Pro Glu Leu Ile Glu Trp
    290                 295                 300

Gly Cys Arg Leu Met Phe Leu Ser Gly His Gly Thr Pro Cys Glu Asn
305                 310                 315                 320

Ala Leu Asn Lys Lys Ile Leu Asp Ile Pro Gly Val Arg Ser Ala Leu
                325                 330                 335

Tyr Pro Cys Pro Ile Arg Thr Pro Cys Ile Lys Ser Glu
            340                 345
```

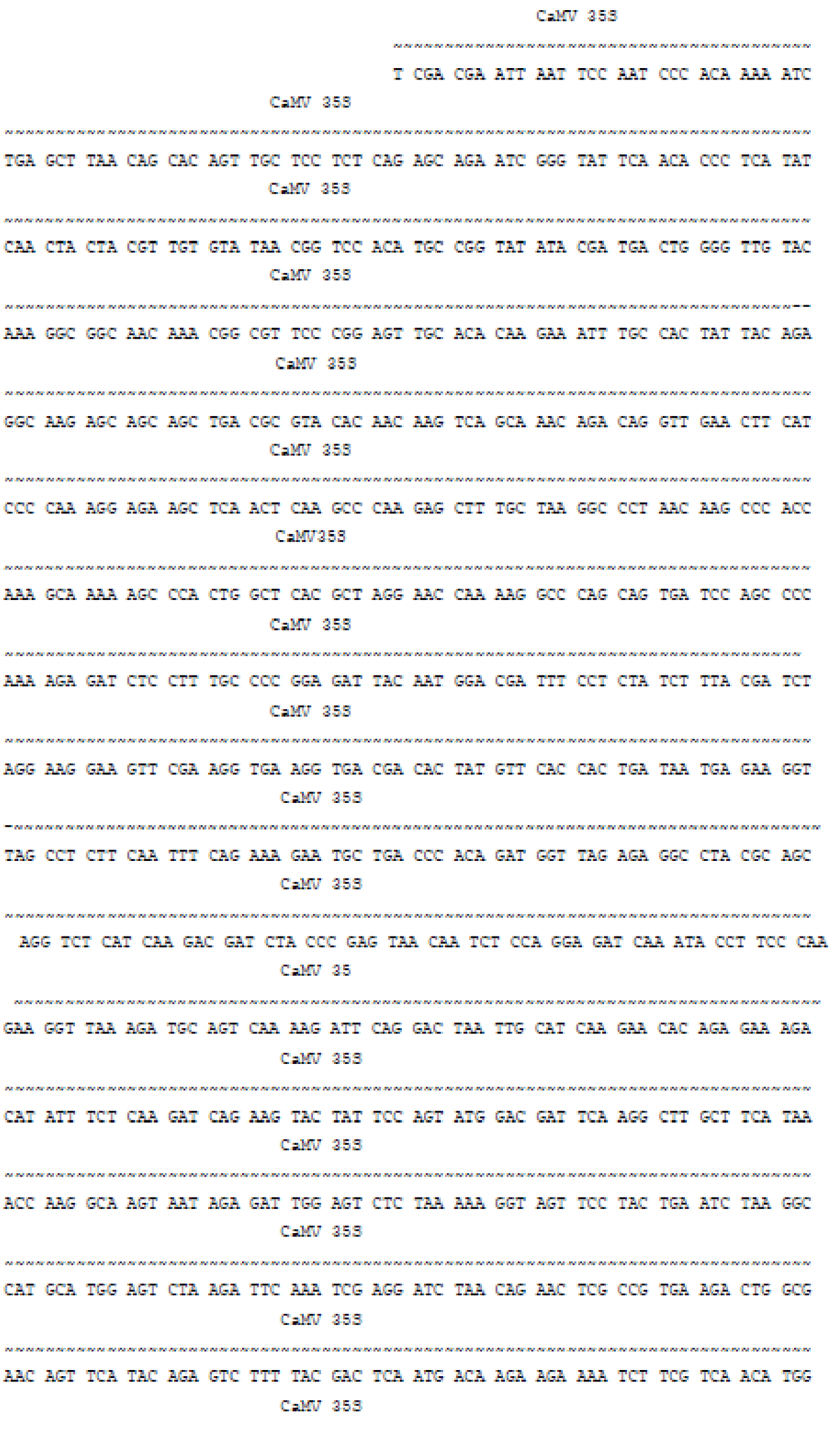
MÉTODO PARA INCREMENTAR LA ASIMILACIÓN DE CO_2 Y LA PRODUCCIÓN DE ACEITE EN ORGANISMOS NO FOTOSINTÉTICOS

RESUMEN – OBJETO Y FINALIDAD DE LA INVENCIÓN

El invento provee métodos para acrecentar la tasa de asimilación de CO_2 en células y plantas fotosintéticas. El invento comprende reducir o evitar el reciclado de lípidos y/o expresar oleosinas modificadas con residuos de cisteína introducidos artificialmente en las células y plantas fotosintéticas. El invento provee también métodos para acrecentar producción de aceite en plantas, mediante la expresión de oleosinas modificadas con residuos de cisteína introducidos artificialmente en los tejidos/órganos no fotosintéticos de las plantas. El método incluye también opcionalmente el paso de extraer el aceite de los tejidos/órganos no fotosintéticos de la planta, o procesar los tejidos/órganos no fotosintéticos ricos en aceite para formar alimentos para animales o materia prima de biocombustibles.

SECTOR: Ingeniería química y de petróleos.

Figura 1


CaMV 35S
T CGA CGA ATT AAT TCC AAT CCC ACA AAA ATC
CaMV 35S
TGA GCT TAA CAG CAC AGT TGC TCC TCT CAG AGC AGA ATC GGG TAT TCA ACA CCC TCA TAT
CaMV 35S
CAA CTA CTA CGT TGT GTA TAA CGG TCC ACA TGC CGG TAT ATA CGA TGA CTG GGG TTG TAC
CaMV 35S
AAA GGC GGC AAC AAA CGG CGT TCC CGG AGT TGC ACA CAA GAA ATT TGC CAC TAT TAC AGA
CaMV 35S
GGC AAG AGC AGC AGC TGA CGC GTA CAC AAC AAG TCA GCA AAC AGA CAG GTT GAA CTT CAT
CaMV 35S
CCC CAA AGG AGA AGC TCA ACT CAA GCC CAA GAG CTT TGC TAA GGC CCT AAC AAG CCC ACC
CaMV35S
AAA GCA AAA AGC CCA CTG GCT CAC GCT AGG AAC CAA AAG GCC CAG CAG TGA TCC AGC CCC
CaMV 35S
AAA AGA GAT CTC CTT TGC CCC GGA GAT TAC AAT GGA CGA TTT CCT CTA TCT TTA CGA TCT
CaMV 35S
AGG AAG GAA GTT CGA AGG TGA AGG TGA CGA CAC TAT GTT CAC CAC TGA TAA TGA GAA GGT
CaMV 35S
TAG CCT CTT CAA TTT CAG AAA GAA TGC TGA CCC ACA GAT GGT TAG AGA GGC CTA CGC AGC
CaMV 35S
AGG TCT CAT CAA GAC GAT CTA CCC GAG TAA CAA TCT CCA GGA GAT CAA ATA CCT TCC CAA
CaMV 35
GAA GGT TAA AGA TGC AGT CAA AAG ATT CAG GAC TAA TTG CAT CAA GAA CAC AGA GAA AGA
CaMV 35S
CAT ATT TCT CAA GAT CAG AAG TAC TAT TCC AGT ATG GAC GAT TCA AGG CTT GCT TCA TAA
CaMV 35S
ACC AAG GCA AGT AAT AGA GAT TGG AGT CTC TAA AAA GGT AGT TCC TAC TGA ATC TAA GGC
CaMV 35S
CAT GCA TGG AGT CTA AGA TTC AAA TCG AGG ATC TAA CAG AAC TCG CCG TGA AGA CTG GCG
CaMV 35S
AAC AGT TCA TAC AGA GTC TTT TAC GAC TCA ATG ACA AGA AGA AAA TCT TCG TCA ACA TGG
CaMV 35S