



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 292 234**

51 Int. Cl.:
C12N 15/82 (2006.01)
C12N 5/10 (2006.01)
C12N 9/02 (2006.01)
A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **99908362 .9**
86 Fecha de presentación : **23.02.1999**
87 Número de publicación de la solicitud: **1056875**
87 Fecha de publicación de la solicitud: **06.12.2000**

54 Título: **Promotores sintéticos.**

30 Prioridad: **24.02.1998 US 28819**

45 Fecha de publicación de la mención BOPI:
01.03.2008

45 Fecha de la publicación del folleto de la patente:
01.03.2008

73 Titular/es:
PIONEER HI-BRED INTERNATIONAL, Inc.
800 Capital Square, 400 Locust Street
Des Moines, Iowa 50309, US

72 Inventor/es: **Bowen, Benjamin, A.;**
Bruce, Wesley, B.;
Lu, Guihua;
Sims, Lynne, E. y
Tagliani, Laura, A.

74 Agente: **Carvajal y Urquijo, Isabel**

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

DESCRIPCIÓN

Promotores sintéticos.

5 **Campo de la invención**

Esta invención se relaciona generalmente con el campo de la biología molecular vegetal y en particular para mejorar la expresión de los genes estructurales deseados en ambas plantas monocotiledóneas y dicotiledóneas.

10 **Antecedentes de la invención**

La expresión génica abarca un número de etapas que originan de la plantilla del ADN finalmente a la proteína o el producto de la proteína final. El control y la regulación de la expresión génica pueden ocurrir a través de numerosos mecanismos. La iniciación de la transcripción de un gen se piensa generalmente como el control predominante de la expresión génica. Los controles transcripcionales (o promotores) se relegan generalmente a las secuencias relativamente cortas incrustadas en la región flanqueante-5' o del extremo 5' del gen transcrito. Existen secuencias de ADN que afectan la expresión génica en respuesta al estímulo ambiental, disponibilidad de nutrientes, o condiciones adversas que incluyen choque térmico, anaerobiosis o la presencia de metales pesados. También hay secuencias de ADN que controlan la expresión génica durante el desarrollo o en un tejido, o una forma específica del órgano.

Los promotores contienen las señales para que la polimerasa del ARN comience la transcripción de modo que la síntesis de la proteína pueda proceder. El ADN que une, las proteínas nucleares interactúa específicamente con estas secuencias cognadas de ADN del promotor para promover la formación del complejo transcripcional y eventualmente iniciar el proceso de la expresión génica.

Uno de los motivos más comunes de las secuencias presentes en los promotores de genes transcritos por un sistema II (polII) de la polimerasa de ARN eucariota es el elemento "TATA" que reside en el extremo 5' del inicio de la transcripción. Los promotores eucariotas son complejos y abarcan los componentes que incluyen una secuencia consenso de la caja TATA a aproximadamente 35 pares de bases 5' relativo al sitio inicial de transcripción o sitio remate que se define como +1. El motivo TATA es el sitio donde la proteína que une TATA (TBP) como parte de un complejo de varios polipéptidos (complejo TFIID) se une e interactúa productivamente (directa o indirectamente) con los factores unidos a otros elementos de las secuencias del promotor. Este complejo TFIID a su vez recluta el complejo de la polimerasa de ARN II que se colocará para iniciar la transcripción generalmente 25 a 30 pares de bases del extremo 3' del elemento TATA y promueve la elongación que produce las moléculas del ARN. Las secuencias más o menos al inicio de la transcripción (denominadas INR) de algunos genes polII parecen proporcionar un sitio de enlace alternativo para los factores que también reclutan los miembros del complejo TFIID y así "activan" la transcripción. Estas secuencias INR son particularmente relevantes en los promotores que carecen de elementos funcionales TATA proporcionando los sitios de enlace promotor central para la transcripción eventual. Se ha propuesto que los promotores que contienen ambos un TATA funcional y un motivo INR son los más eficientes en la actividad transcripcional. (Zenzie-Gregory *et al.* (1992) J. Biol.Chem. 267:2823-2830).

En la mayoría de los casos, los elementos de las secuencias con excepción del motivo TATA se requieren para la transcripción exacta. Tales elementos se localizan con frecuencia en el extremo 5' del motivo TATA y un subconjunto puede tener homología con la secuencia consenso CCAAT.

Se ha encontrado que otras secuencias de ADN elevan el nivel total de expresión de los genes cercanos. Uno de los elementos más comunes que ha sido descrito reside lejos del extremo 5' del sitio de iniciación y parece mostrar características de posición y orientación independientes. Estos elementos lejanos del extremo 5' se han denominado potenciadores.

Uno de los elementos menos comunes en virtud de sus especificidades son las secuencias que interactúan con factores que unen el ADN específico. Estos motivos de secuencias se conocen colectivamente como elementos del extremo 5' que son dependientes generalmente de la posición y la orientación.

Muchos elementos del extremo 5' se han identificado de un número de promotores de la planta basados inicialmente en la función y en segundo término en las homología de las secuencias. Estos elementos del extremo 5' del promotor oscilan ampliamente en el tipo de control: de respuestas ambientales como temperatura, humedad, lesiones, etc., señales de desarrollo, (germinación, maduración de la semilla, floración, etc.) para la información espacial (especificidad del tejido). Estos elementos también parecen mostrar modularidad en eso pueden ser intercambiados con otros elementos mientras que mantienen su control característico sobre la expresión génica.

Los promotores usualmente se sitúan 5' o en el extremo 5' relativo al inicio de la región codificante al gen correspondiente, y la región total que contiene todos los elementos auxiliares que afectan la regulación o los niveles absolutos de transcripción se pueden contener de menos de 100 pares de bases o tanto como 1 par de kilobases.

Un número de promotores que son activos en las células vegetales se han descrito en la literatura. Estos incluyen los promotores nopalina sintasa (NOS) y octopina sintasa (OCS) (que se llevan en el tumor que induce los plásmidos del *Agrobacterium tumefaciens*). También se incluyen, el virus mosaico de la coliflor (CaMV) 19S y los promotores

35S, el promotor leve-inducible a partir de la pequeña subunidad de carboxilasa bifosfato ribulosa (ssRUBISCO, un polipéptido vegetal muy abundante), y el promotor sacarosa sintasa. Todos estos promotores se han utilizado para crear varios tipos de construcciones de ADN que se han expresado en las plantas. (Ver por ejemplo la publicación PCT WO84/0291 Rogers, *et al*).

5 Dos promotores que se han utilizado ampliamente en las transformaciones de las células vegetales son aquellos de los genes que codifican el alcohol dehidrogenasa, AdhI y AdhII. Ambos genes se inducen después del principio de la anaerobiosis. El maíz AdhI ha sido clonado y secuenciado como ha sido el AdhII. La formación de un gen quimérico AdhI, El Adh-Cat que abarca el promotor AdhI conectado a la cloramfenicol acetiltransferasa (CAT) que
10 codifica las secuencias y la señal nopalina sintasa (NOS) 3' causó la expresión CAT a niveles aproximadamente 4-veces más altos, en concentraciones de oxígeno bajas, que bajo las condiciones control. Los elementos de la secuencia necesarios para la inducción anaeróbica del ADH-CAT quimérico también han sido identificados. Las existencias de elementos reguladores anaeróbicos (ARE) entre las posiciones -140 y -99 del promotor AdhI de maíz compuestos de al menos dos elementos de secuencias en las posiciones -133 a -124 y las posiciones -113 a 99 ambos de los cuales
15 se han encontrado necesarios y son suficientes para la baja expresión de oxígeno de la actividad del gen ADH-CAT. El promotor Adh sin embargo responde a la anaerobiosis y no es un promotor constitutivo limitando drásticamente su efectividad.

Otro promotor utilizado comúnmente es el promotor 35S del Virus del Mosaico de la Coliflor. El promotor 35S (CaMV) es un promotor de virus dicotiledóneo sin embargo dirige la expresión de los genes introducidos en los
20 protoplastos de dicotiledóneas y monocotiledónea. El promotor 35S es un promotor muy fuerte y esto explica su extenso uso para la expresión a alto nivel de rasgos en plantas transgénicas. El promotor CaMV 35S sin embargo también ha demostrado actividad relativamente baja en varias plantas gramíneas de importancia agrícola tales como el trigo. Mientras que todos estos promotores dan una expresión alta en las dicotiledóneas, pocos dan niveles altos de
25 expresión en monocotiledóneas. Existe una necesidad para los promotores sintéticos y otros elementos que inducen la expresión en células transformadas del protoplasto de las monocotiledóneas.

Resumen de la invención

30 Se proporcionan los métodos y composiciones para la expresión de secuencias heterólogas en las células huésped. Las composiciones encuentran particular uso en controlar la expresión de secuencias en las plantas. Las composiciones de la invención comprenden las secuencias promotor. En particular, se proporcionan, una molécula sintética novedosa del promotor central y los elementos reguladores útiles para controlar la expresión en las células diana. El promotor central comprende una caja TATA y un inicio de la transcripción. Además, la región "TATA para iniciar" es 64% o
35 más rica en GC. Los elementos reguladores incluyen un elemento novedoso del extremo 5' y las regiones que activan el extremo 5'. La región de activación del extremo 5' es diferente del elemento del extremo 5' sintético. Los elementos se pueden utilizar juntos o con otros elementos del promotor para controlar la expresión de las secuencias de interés.

Es un objeto de primer orden de la invención, proporcionar los elementos reguladores sintéticos que mejoren la
40 expresión de los genes introducidos en las células vegetales y los tejidos vegetales.

Es un objeto de la invención proporcionar una molécula del promotor recombinante que prevea los niveles altos confiables de la expresión de genes introducidos en las células diana. Es aún otro objeto de la invención proporcionar los elementos potenciadores del extremo 5' heterólogos que pueden mejorar la actividad de cualquier promotor.
45

Es otro objeto de la invención proporcionar plantas, células vegetales y tejidos vegetales que contienen cualquiera o ambos el promotor recombinante o el elemento del extremo 5' de la invención.

Es otro objeto de la invención proporcionar los vehículos para la transformación de las células vegetales que
50 incluyen vectores virales o plásmidos y casetes de expresión que incorporan el promotor sintético y los elementos del extremo 5' de la invención.

Es otro objeto de la invención proporcionar células bacterianas que comprenden tales vectores para el mantenimiento, y la transformación de la planta.
55

Otros objetos de la invención, serán evidentes a partir de la descripción de la invención que sigue. De acuerdo con la presente invención, así se proporciona una secuencia de ADN del promotor vegetal sintético, dicha secuencia que comprende:

- 60 un motivo TATA;
- un sitio inicial de transcripción;
- una región entre dicho motivo TATA y dicho sitio inicial que es al menos 64% rica en GC;
- 65 un elemento del extremo 5'; y
- una región de activación del extremo 5',

en donde dicha secuencia del promotor comprende una secuencia seleccionada del grupo que consiste de:

- a) una secuencia de nucleótidos publicada en la SEQ ID NO: 12, y
- b) una secuencia de nucleótidos que tiene al menos un 90% de identidad de la secuencia con a).

La invención también proporciona un casete de expresión que comprende:

una secuencia de ADN del promotor sintético de la invención;

un gen estructural ligado de manera operable a dicho promotor; y

una señal de poliadenilación del sitio de terminación de la transcripción.

La invención también proporciona un vector de ácidos nucleicos que comprende un promotor de la invención ligado de manera operable a un gen estructural.

La invención también proporciona una célula huésped procariota o eucariota transformada con un vector de ácidos nucleicos de la invención.

La invención también proporciona un método de control del nivel de expresión de una secuencia de nucleótidos transgénica en una célula vegetal, dicho método que comprende la transformación de dicha célula vegetal con un casete de expresión de la invención.

La invención también proporciona una planta transgénica que comprende una célula vegetal o progenitor de esta que ha sido transformada con un vector de ácidos nucleicos de la invención.

La invención también proporciona una planta transformada que contiene en su genoma un casete de expresión de la invención.

La invención también proporciona una planta transgénica que expresa un gen oxalato oxidasa exógeno a un nivel alto en donde dicho gen oxalato oxidasa esta bajo el control transcripcional de un promotor recombinante de la invención.

La invención también proporciona una semilla de un vegetal de la invención, una semilla que contiene en su genoma un promotor recombinante de la invención.

Descripción de las figuras

Figura 1 es una descripción de una configuración base del nucleótido típica de un promotor central que contiene las secuencias consenso de los motivos TATA e INR presentes en los promotores de la planta. A denomina +1 de la región transcrita.

Figura 2 es una descripción de la Secuencia del Promotor Core Syn II completa con un ejemplo de un promotor de la planta y ambos se alinean en el inicio principal de la transcripción (letra en negrilla). El motivo TATA se subraya. El promotor 35S CaMV se muestra con el porcentaje de las secuencias del contenido de GC en paréntesis.

Figura 3 es la secuencia de ADN del elemento Rsyn 7 del extremo 5'. Los motivos TGACG se indican en negrilla.

Figura 4 es un mapa del plásmido del promotor Central Syn II y los elementos Rsyn7 conducen un GUS que contiene la construcción.

Figura 5 describe varios esquemáticos de los promotores sintéticos probados en los transformables transitorios y estable.

Figura 6 es una descripción de datos de ensayo transitorio utilizando los plásmidos que incorporan las secuencias del promotor central descritos aquí.

Figura 7(A). Actividad de Rsyn7::GUS (PHP6086) para las plantas del maíz T0.

Figura 7(B) es un esquemático de la etapa VT de las plantas de maíz con los sitios de muestras del tejido indicado.

Figura 8 describe la actividad GUS en segmentos de raíz de una población segregada de los semilleros transgénicos T1 del maíz que contienen el Rsyn7::GUS (PHP6086) o la construcción UBI:GUS (PHP3953).

Figura 9 describe la expresión GUS de tres promotores sintéticos en plantas transgénicas del maíz T0 que incluyen las secuencias del promotor central descritos aquí como comparación.

Figura 10 muestra la comparación de las actividades del promotor Rsyn7, el promotor 35S CaMV, y el promotor 35S_{UR}syn7 en la expresión transitoria en cotiledones del girasol.

Figura 11 muestra el efecto de la región de activación del extremo 5' de *Ubi-l* en la potencia del promotor Rsyn-7 en la expresión transitoria en los cotiledones del girasol.

Figura 12 muestra que en los callos de girasol transformados estables, la expresión GUS detrás del control del 35S_{UR}syn7 es 20% más alta que cuando detrás del control del promotor 35S CaMV.

Figura 13 muestra el efecto de la región de activación del extremo 5' de *Ubi-l* en la actividad del promotor Rsyn7 en el ensayo de callos del girasol transgénicos.

Descripción detallada de la invención

En la descripción que sigue un número de términos se utilizan extensamente. Las siguientes definiciones se proporcionan con el fin de eliminar ambigüedades en el intento o alcance de sus usos en la especificación y las reivindicaciones, y para facilitar la comprensión de la invención.

Un gen estructural es una secuencia de ADN que se transcribe en el ARN mensajero (mARN) el cual luego se traduce en una secuencia de aminoácidos característicos de un polipéptido específico.

Un promotor es una secuencia de ADN que dirige la transcripción de un gen estructural. Usualmente un promotor se localiza en la región 5' de un gen, próximo al sitio inicial transcripcional de un gen estructural. El promotor de la invención comprende un promotor central según lo definido abajo. Adicionalmente, el promotor incluye los elementos del extremo 5'. Tales elementos incluyen UARs y opcionalmente, otras secuencias de ADN que afectan la transcripción de un gen estructural tal como un elemento del extremo 5' sintético.

Un promotor central o promotor mínimo contiene las secuencias de nucleótidos esenciales para la expresión de la secuencia codificante ligada de manera operable, que incluyen la caja TATA y el inicio de la transcripción. Por esta definición, un promotor central puede o no tener actividad detectable en la ausencia de secuencias específicas que pueden potenciar la actividad o conferir actividad específica del tejido. Por ejemplo, el promotor central del gen SGB6 del maíz consiste de aproximadamente 37 nucleótidos 5' del sitio inicial transcripcional del gen SGB6, mientras que el promotor central del Virus del Mosaico de la Coliflor (CaMV) 35S consiste de aproximadamente 33 nucleótidos 5' del sitio inicial transcripcional del genoma 35S.

ADH se refiere generalmente a un gen alcohol dehidrogenasa vegetal expresable y específicamente al gen alcohol dehidrogenasa del maíz.

ADH 1 UAR se refiere al fragmento de ADN que abarca la región entre las posiciones del nucleótido aproximadamente -1094 a aproximadamente -106 del gen alcohol dehidrogenasa 1 del maíz, o el fragmento homólogo que es equivalente funcionalmente. La secuencia se enumera con el inicio del sitio de transcripción, denominada como +1 de acuerdo con la corrección publicada por Ellis *et al.* (1987) *supra*.

"TATA para iniciar" significaría la secuencia entre el motivo TATA de primer orden y el inicio de la transcripción.

Un ADN sintético es una secuencia de ADN creada artificialmente que no se produce de modo natural, y se debe introducir a un organismo o a un progenitor de ese organismo para controlar o para ser expresado.

El elemento OCS se refiere al motivo TGACG identificado del gen octopina sintasa, los genes histona, genes de enzimas para la biosíntesis de la agropina, el gen mannopina sintasa, el gen CaMV 35S, gen H3 histona y gen nopalina sintasa. Según se utiliza aquí el término incluye cualquier secuencia capaz de unir el factor ASF-1 según lo identificado en Patente U.S. No. 4,990,607 por Katagiri.

UAR es usualmente un elemento dependiente de la posición u orientación que en primer lugar dirige el tejido, tipo de célula, o expresión regulada.

Un potenciador es un elemento regulador de ADN que puede incrementar la eficiencia de la transcripción a pesar de la distancia u orientación del potenciador relativo al sitio inicial de la transcripción.

El término expresión se refiere a la biosíntesis de un producto génico. En el caso de un gen estructural, la expresión involucra la transcripción del gen estructural en mARN y luego la traducción del mARN en uno o más polipéptidos.

Un vector de clonación es una molécula de ADN tal como un plásmido, cósmido o fago bacteriano que tiene la capacidad de replicación autónomamente en una célula huésped. Los vectores de clonación usualmente contienen uno o un número pequeño de sitios de reconocimiento de la restricción de la endonucleasa en los cuales las secuencias de ADN foráneo se pueden insertar de una manera determinable sin pérdida de función biológica individual del vector, así como un gen marcador que es apropiado para utilizar en la identificación y selección de las células transformadas.

ES 2 292 234 T3

con el vector de clonación. Los genes marcadores usualmente incluyen los genes que proporcionan la resistencia a la tetraciclina, resistencia a la higromicina o resistencia a la ampicilina.

Un vector de expresión es una molécula de ADN que comprende un gen que se expresa en una célula huésped. Usualmente la expresión génica se coloca bajo el control de ciertos elementos reguladores que incluyen promotores, elementos reguladores de tejido específico, y potenciadores. Tal gen se indica para ser “ligado de manera operable a” los elementos reguladores.

Un huésped recombinante puede ser cualquier célula procariota o eucariota que contiene ya sea un vector de clonación o un vector de expresión. Este término también incluye aquellas células procariotas o eucariotas que se han modificado genéticamente para contener los genes clonados en el cromosoma o en el genoma de la célula huésped.

Una planta transgénica es una planta que tiene una o más células vegetales que contienen un vector de expresión.

Será entendido que puede haber variaciones menores de la secuencia dentro de la secuencia o los fragmentos utilizados o revelados en esta aplicación. Por “variaciones menores” se pretende que las secuencias tengan al menos 80%, preferiblemente 90% de identidad de la secuencia. Estas variaciones se pueden determinar por técnicas estándar para permitir a aquellos de ordinaria habilidad en el oficio manipular y poner en utilidad las unidades funcionales de los elementos del promotor necesarios para dirigir la iniciación de la transcripción en el gen estructural seguido por una señal de terminación de transcripción (y quizás poliadenilación) expresable de la planta.

El tejido vegetal incluye tejidos o plantas diferenciadas y sin diferencias, que incluyen pero no limitan a raíces, tallos, brotes, hojas, polen, semillas, tejido tumoral y varias formas de células y cultivo tales como células simples, protoplasto, embriones, y tejido de callos. El tejido vegetal puede estar en plantas o en cultivo de órganos, tejidos o células.

Un elemento revelado aquí, el promotor central, se muestra en la SEQ ID NO: 1. El promotor central es capaz de conducir la expresión de una secuencia codificante en una célula diana, particularmente las células vegetales. El promotor central encuentra uso en conducir la expresión de las secuencias que solamente son necesarias a niveles mínimos en las células diana. También se revela un elemento novedoso del extremo 5', SEQ ID NO:2, que ayuda a potenciar la transcripción. El promotor central sintético se puede utilizar con combinaciones del potenciador, elementos del extremo 5', y/o secuencias que activan a partir de las regiones flanqueantes de 5' de los genes estructurales expresables de la planta. Del mismo modo el elemento del extremo 5' se puede utilizar en combinación con varias secuencias del promotor central de la planta. El promotor central y elemento del extremo 5' se pueden utilizar juntos para obtener una expresión diez-veces más alta de un gen marcador introducido en plantas transgénicas monocotiledóneas que se obtienen con el promotor ubiquitin I del maíz.

El promotor central comprende un motivo TATA y una región “TATA para iniciar la transcripción” rica en GC (64% o más del contenido de GC) que es generalmente característico de los promotores de animales. La secuencia se coloca en 5' de un gen estructural y estimulará la expresión constitutiva que es tejido no-específica en células de plantas transgénicas.

También se describe aquí un casete de expresión que comprende el elemento del extremo 5', el promotor central sintético, un gen estructural, la expresión que se desea en las células vegetales, y una señal de poliadenilación o de parada. El casete de expresión se puede incluir en vectores virales o plásmidos para la transformación de las células del protoplasto de la planta.

La invención también abarca las células bacterianas transformadas para el mantenimiento y replicación del vector, así como las células monocotiledóneas o dicotiledóneas transformadas y finalmente las plantas transgénicas.

También se revela aquí un elemento del extremo 5' que se pueden utilizar en combinación con el promotor sintético o con otros promotores conocidos en el oficio. El elemento del extremo 5' comprende al menos 3 motivos (TGACG) unidos a OCS con una secuencia que interviene novedosa. Una modalidad se revela en la SEQ ID NO: 2 y se coloca 5' a una secuencia del promotor central para potenciar los niveles de transcripción del producto génico resultante. De esta manera describimos un casete de expresión que comprende el elemento del extremo 5' sintético de la invención, 5' en un promotor inducible de la planta que está en 5' en un gen estructural. Este casete de expresión se puede incorporar en los vectores y plásmidos según lo descrito previamente.

El elemento del extremo 5' sintético se puede utilizar en combinación con la secuencia del promotor central sintético para lograr la expresión constitutiva de tejido no-específico del producto génico que es un incremento de diez-veces del promotor *Ubi-1* del maíz.

La presente invención abarca una construcción del promotor que comprende el promotor central sintético descrito arriba y una región de activación del extremo 5'. La región de activación del extremo 5' es diferente del elemento del extremo 5' sintético. Preferiblemente la región de activación del extremo 5' es una región de activación del extremo 5' (UAR) que tiene considerable similitud de secuencia con la UAR de CaMV 35S. Las construcciones del promotor de la invención comprenden el promotor central sintético en combinación con una UAR y un elemento del extremo 5' sintético.

La construcción del promotor se puede incluir por conveniencia en un casete de expresión. Este casete de expresión se puede incluir en vectores de transformación.

La secuencia de la región de activación del extremo 5' (UAR) del gen Ubi-1 del maíz también se revela. Esta UAR se puede utilizar en combinación con cualquier promotor central para potenciar la actividad del promotor.

El promotor central como se ve en la SEQ ID NO: 1 y/o SEQ ID NO: 10 (promotor central modificado), se pueden utilizar para obtener niveles altos de la expresión de genes estructurales. Del mismo modo el elemento del extremo 5' de la SEQ ID NO: 2 se puede utilizar en combinación con otros promotores o el promotor de la invención para potenciar los niveles de la transcripción en las plantas modificadas genéticamente. La producción de un tejido vegetal modificado genéticamente que expresa un gen estructural bajo el control de los elementos reguladores de la invención combina las instrucciones de la presente divulgación con una variedad de técnicas y expedientes conocidos en el oficio. En la mayoría de los casos existen expedientes alternos para cada etapa del proceso completo. La elección de los expedientes depende de las variables tales como el sistema del vector plásmido seleccionado para la clonación e introducción de la molécula de ADN recombinante, las especies vegetales a ser modificadas, el gen estructural particular, los elementos del promotor y los elementos del extremo 5' utilizados. Las personas de habilidad en el oficio son capaces de seleccionar y utilizar alternativas apropiadas para lograr la funcionabilidad. Las condiciones de cultivo para expresar los genes estructurales deseados y las células cultivadas son conocidas en el oficio. También según se conoce en el oficio, se pueden obtener, un número de ambas especies vegetales monocotiledóneas y dicotiledóneas son transformables y regenerables tal que las plantas completas que contienen y que expresan los genes deseados bajo el control regulador de las moléculas promotor de la invención. Como se conoce por aquellos de habilidad en el oficio, la expresión en plantas transformadas puede ser específica al tejido y/o específica a ciertas etapas para el desarrollo o influencias ambientales. La selección del promotor truncada y la selección del gen estructural son otros parámetros que pueden ser optimizados para lograr la expresión vegetal deseada como se conoce por aquellos de habilidad en el oficio y se enseña aquí.

Las secuencias de nucleótidos de la invención se pueden introducir en cualquier vegetal. Los genes que se introducirán se pueden utilizar convenientemente en los casetes de expresión por la introducción y expresión en cualquier planta de interés.

Tales casetes de expresión comprenderán un promotor de la invención unido a una secuencia de nucleótidos de interés. Tal casete de expresión se proporciona con una pluralidad de los sitios de restricción por inserción del gen de interés para estar bajo la regulación transcripcional de las regiones reguladoras. El casete de expresión puede adicionalmente contener genes marcadores seleccionables.

El casete transcripcional incluirá en la dirección 5'-3' de la transcripción, una región de iniciación transcripcional y de traducción, una secuencia de ADN de interés, y una región de terminación transcripcional y de traducción funcional en plantas. La región de terminación puede ser nativa con la región de iniciación transcripcional, puede ser nativa con la secuencia de ADN de interés, o puede ser derivada de otra fuente. Las regiones de terminación convenientes están disponibles del Tiplasmid de *A. tumefaciens*, tales como las regiones de terminación octopina sintasa y nopalina sintasa. Ver también, Guerineau *et al.*, (1991) Mol. Gen. Genet. 262:141-144; Proudfoot (1991) Cell 64:671-674; Sanfacon *et al.* (1991) Genes Dev. 5:141-149; Mogen *et al.* (1990) Plant Cell 2:1261-1272; Munroe *et al.* (1990) Gene 91:151-158; Ballas *et al.* 1989) Nucleic Acids Res. 17:7891-7903; Joshi *et al.* (1987) Nucleic Acid Res. 15:9627-9639.

Los genes descritos aquí se proporcionan en casetes de expresión para la expresión en la planta de interés. El casete incluirá secuencias reguladoras 5' y 3' ligadas de manera operable al gen de interés. El casete puede adicionalmente contener al menos un gen adicional para ser cotransformado en el organismo. Alternativamente, el o los genes adicionales pueden ser proporcionados en otro casete de expresión. Donde es apropiado, el gen(s) puede ser optimizado para incrementar la expresión en la planta transformada. Es decir que, los genes se pueden sintetizar utilizando codones preferidos de plantas para mejorar la expresión. Los métodos están disponibles en el oficio para sintetizar los genes preferidos de las plantas. Ver, por ejemplo, Patente U.S. Nos. 5,380,831, 5,436, 391, y Murray *et al.* (1989) Nucleic Acids Res. 17:477-498.

Se conocen modificaciones de la secuencia adicionales para potenciar la expresión génica en un huésped celular. Estas incluyen eliminación de las secuencias que codifican las señales de poliadenilación falsas, señales del sitio de división y empalme exón-intrón, repeticiones semejantes a transposón, y otras secuencias bien caracterizadas que pueden ser nocivas a la expresión génica. El contenido G-C de la secuencia se puede ajustar a niveles promedios para un huésped celular dado, como se calcula por referencia a aquellos genes conocidos expresados en la célula huésped. Cuando sea posible, la secuencia se modifica para evitar las predichas estructuras de horquilla del mARN secundario.

La selección de un vector de expresión apropiado dependerá del método para introducir el vector de expresión en las células huésped. Usualmente un vector de expresión contiene (1) elemento de ADN procariota que codifica para un origen de una replicación bacteriana y un gen de resistencia a un antibiótico para proporcionar la amplificación y selección del vector de expresión en un huésped bacteriano; (2) elementos de ADN que controlan la iniciación de la transcripción tal como un promotor; (3) elementos de ADN que controlan el proceso de las transcripciones tales como los intrones, terminación de la transcripción/secuencia de poliadenilación; y (4) un gen indicador que se liga de manera operable a los elementos de ADN para controlar la iniciación de la transcripción. Los genes indicadores úti-

les incluyen β -glucoronidasa, β -galactosidasa, cloramfenicol acetil transferasa, luciferasa, proteína fluorescente verde (GFP) y similares. Preferiblemente el gen indicador es ya sea β -glucoronidasa (GUS), GFP o luciferasa. Las descripciones generales de vectores de expresión de las plantas y genes indicadores se pueden encontrar en Gruber, *et al.*, "Vectors for Plant Transformation, in Methods in Plant Molecular Biology & Biotechnology" en Glich *et al.*, (Eds. pp. 89-119, CRC Press, 1993). Adicionalmente los vectores de expresión GUS y casetes del gen GUS están disponibles de Clonetech Laboratories, Inc., Palo Alto, California mientras los vectores de expresión luciferasa y casetes del gen luciferasa están disponibles de Promega Corp. (Madison, Wisconsin).

Los vectores de expresión que contienen fragmentos genómicos o sintéticos se pueden introducir en los protoplastos o en los tejidos intactos o células aisladas. Preferiblemente los vectores de expresión se introducen en tejido intacto. Los métodos generales de cultivos de tejidos vegetales se proporcionan por ejemplo por Maki *et al.* "Procedures for Introducing Foreign ADN into Plants" en Methods in Plant Molecular Biology & Biotechnology, Glich *et al.* (Eds. pp. 67-88 CRC Press, 1993); y por Phillips *et al.* "Cell- Tissue Culture and *In-Vitro* Manipulation" en Corn & Corn Improvement, 3rd Edition Sprague *et al.* (Eds. pp. 345-387) American Society of Agronomy Inc. *et al.* 1988.

Los métodos para introducir los vectores de expresión en el tejido vegetal incluyen la infección directa o co-cultivo de la célula vegetal con *Agrobacterium tumefaciens*, Horsch *et al.*, Science, 227:1229 (1985). Las descripciones de los sistemas del vector *Agrobacterium* y los métodos para la transferencia del gen mediada por el *Agrobacterium* por Gruber, *et al. supra*.

Preferiblemente, los vectores de expresión se introducen en maíz u otros tejidos vegetales utilizando un método directo de transferencia del gen tales como entrega mediada por microproyectiles, inyección de ADN, electroporación y similares. Más preferiblemente los vectores de expresión se introducen en los tejidos vegetales utilizando la entrega del medio de microproyectiles con el dispositivo biolístico. Ver, por ejemplo, Tomes *et al.* "Direct DNA transfer into intact plant cells via microprojectiles bombardment" In: Gamborg and Phillips (Eds.) Plant Cell, Tissue and Organ Culture: Fundamental Methods, Springer-Verlag, Berlin (1995).

Los vectores de la invención no pueden solamente ser utilizados para la expresión de genes estructurales sino también pueden ser utilizados en la clonación exón-trampa, o procedimientos de promotor trampa para detectar la expresión génica diferencial en variedades de tejidos, K. Lindsey *et al.*, 1993 "Tagging Genomic Sequences That Direct Transgene Expression by Activation of a Promoter Trap in Plants", Transgenic Research 2:33-47. D. Auch & Reth, *et al.*, "Exon Trap Cloning: Using PCR to Rapidly Detect and Clone Exons from Genomic ADN fragments", Nucleic Acids Research, Vol. 18, No. 22, p. 6743.

La identificación del promotor central se basa en parte en el descubrimiento que una región "TATA para iniciar" rica en GC en un promotor de la planta actúa como un promotor central muy fuerte de tejido no-específico que induce la expresión constitutiva en las células vegetales. El elemento TATA de los promotores de la planta de los genes *polII* generalmente tienen la secuencia TATA(A/T)A(A/T)A, SEQ ID NO:3, mientras el consenso del inicio de la transcripción consiste de la secuencia 5'...TYYTCAT(A/C)AA..3'. SEQ ID NO.:3, donde la A designa la base inicial para la transcripción. La secuencia del promotor vegetal típica se describe en la Figura 1.

Se ha mostrado que las secuencias que intervienen el elemento TATA y el inicio de la transcripción juegan un papel significativo en la eficiencia de la activación transcripcional. La proteína que une TATA se ha demostrado que interactúa con el surco menor de la doble hélice unida al motivo TATA doblada hacia en lado del surco principal (Kim, *et al.* 1993, Nature, 365: 512-520). Siguen de esta manera las secuencias del extremo 3' del motivo TATA que impacta este descubrimiento afectarán la eficiencia de la formación transcripcional del complejo estable y finalmente la expresión. El reconocimiento de las regiones "TATA para iniciar" de los promotores de la planta muestran un nivel significativamente más alto de las secuencias ricas en AT conducen al potencial de la compresión menor del surco (Yaurawj *et al* Biological Abstracts Vol. 47, Issue 8, Ref. 144712, "Consensus Sequences for Plant Minimal Promoters" Annual Meeting of the American Society of Plant Physiologists, July 29-August 2, 1995, Plant Physiology 108 [2 Supp.] 1995, 114). Generalmente los promotores de animales muestran una secuencia rica en GC "TATA para iniciar" que conduce a una compresión del surco principal que sugiere que el promedio de los complejos del promotor central transcripcional de plantas y animales reconocen e interactúan con un TATA algo diferente para iniciar la estructura con la diferencia de la secuencia correspondiente. Muy sorprendentemente el aplicante ha encontrado que un promotor sintético tipo animal rico en GC trabaja muy bien en las plantas.

Mientras la invención no se limita por ninguna teoría, es posible que el motivo TATA rico en AT presente en una secuencia rica en GC pueda "por sí mismo presentar" más destacadamente al complejo enlace-TATA por una demarcación afilada del motivo TATA que podrían interactuar más estrechamente con el complejo enlace-TATA. Esto podría mejorar el inicio de la eficiencia de la transcripción, alternando el equilibrio del enlace a una forma más estabilizada, mientras que la versión "TATA no-unida", i.e. que tiene un nivel más alto de las secuencias ricas en AT que flanquean el motivo TATA, el complejo enlace-TATA podría potencialmente deslizarse o vacilar 5' o 3' al sitio inicial y reducir efectivamente la eficiencia de unir finalmente la transcripción reducida. Pocos datos respecto a esta región de promotores de la planta es disponible excepto de las delecciones crudas y algunas mutaciones punto. El diseño obvio de un promotor central sintético para la expresión de la planta podría incluir la secuencia "TATA para iniciar" rico en AT basado en la inspección de promotores de la planta conocidos. Sin embargo, basado en el mecanismo "ligado", se postula por el mecanismo de la invención que un promotor central más eficiente es un resultado de un motivo TATA incrustado en una secuencia rica en GC.

La Figura 2 describe la secuencia del promotor Central Syn II, SEQ ID NO: 1, con ejemplos de promotores centrales de la planta alineados al inicio principal de la transcripción. Otro ejemplo de un promotor de la planta 35S de CaMV (SEQ ID NO:4) se muestra con el porcentaje de las secuencias ricas en GC se mostraron a la derecha en el paréntesis. La secuencia Central Syn II no muestra ninguna homología de la secuencia significativa a las secuencias en las bases de datos públicos de la secuencia.

La secuencia del promotor sintético Coren Syn II muestra una secuencia "TATA para iniciar" 64% rica en GC diferente de la secuencia completa 40% rica en GC presente en los promotores tradicionales de la planta (CaMV35S por ejemplo). El promotor central UBI que ocurre de modo natural y aislado que potencia niveles muy altos de la actividad en las monocotiledóneas usualmente muestra una secuencia "TATA para iniciar" 64% rica en GC más similar a los promotores de animales. Tales ejemplos proveen el ímpetu para diseñar una elevada secuencia "TATA para iniciar" rica en GC para la transcripción eficiente en oposición al dogma corriente de los promotores centrales de la planta.

De esta manera el promotor de la invención comprende una secuencia del promotor central sintético de la planta que comprende un motivo TATA y una región "TATA para iniciar" que es 64% rica en GC o más. En una modalidad preferida, el promotor puede incluir la restricción de los sitios diana de la endonucleasa para facilitar la clonación. En la modalidad más preferida, la secuencia es aquella de la SEQ ID NO: 1. Como será apreciado por aquellos de habilidad en el oficio, varias transversiones base dentro de la SEQ ID NO: 1 puede ocurrir que mantendrán el porcentaje del contenido de GC y se pretenden dentro del alcance de esta invención. Por ejemplo las guaninas podrían ser reemplazadas con citocinas y *vice-versa* sin que afectan la eficacia completa del promotor, con la condición que el porcentaje del contenido de GC se mantenga.

También se describe un elemento del extremo 5' sintético situado en 5' en cualquier promotor que ocurre de modo natural o sintético para utilizar en plantas, particularmente la expresión génica del maíz.

De la actividad de numerosos promotores, los elementos básicos (sitios de enlace) se han definido. Estos incluyen por ejemplo regiones ricas en AT a partir de los promotores de choque térmico, y elementos (AS-1) del sitio de enlace ASF-1 presentes en octopina sintasa (OCS) y promotores del Virus del Mosaico de la Coliflor. AS-1 es uno de los mejores elementos conocidos del extremo 5' y su secuencia enlace (elemento OCS) está presente en muchos promotores constitutivos de la planta tales como los promotores histona del trigo CaMV35S, *A. tumefaciens*, NOS y OCS. El elemento OCS primero se aisló como un elemento potenciador en el promotor del gen OCS donde fue identificado como una secuencia palindrómica de 16-pares de bases (Ellis *et al.* (1987) EMBO J. 6:11-16), pero ha sido reducido a sus características esenciales como un motivo TGACG. Ver Patente U.S. No. 4,990,607. El elemento del extremo 5' presente en los promotores de la invención tiene una identidad del 71% al elemento potenciador del promotor se revela en la Patente U.S. 5,023,179 to Lam *et al.* Las dos secuencias son muy diferentes en sus secuencias flanqueadoras circundantes al motivo TGACG, las regiones que se han mostrado por impactar el mejoramiento del nivel de la transcripción. La actividad mejorada transcripcional del elemento OCS se correlaciona con el enlace *in-vitro* de un factor transcripcional. Los elementos similares también se identificaron en las regiones del promotor de seis otros genes de cADN involucrados en la síntesis de opina y tres promotores virales vegetales que incluyen el promotor 35S CaMV (Bouchez *et al.* 1989) *supra*. Estos elementos se mostraron que unen el factor de transcripción OCS *in-vitro* y potencian la transcripción en las células vegetales.

En tabaco un factor que une el ADN, TGA1, se muestra que interactúa específicamente con el elemento AS-1 ya sea solo o en conjunción con otros elementos promotor. (Katagiri *et al.* 1989, Nature 340:727-730). Este factor también se demostró que se expresa de una manera preferida en la raíz en las plantas de tabaco. Los promotores centrales con una o dos copias del elemento OCS del extremo 5' tienden a potenciar la expresión génica mientras 4 o más repeticiones de este elemento producen más o menos actividad constitutiva a pesar de bajas relativas a los promotores 35S intactos.

De esta manera también describimos un elemento del extremo 5' sintético que se puede utilizar con el promotor central descrito aquí u otros promotores centrales para potenciar la expresión génica. El elemento incorpora tres motivos como OCS y las secuencias novedosas que intervienen las cuales potencian la expresión génica.

La Figura 3, SEQ ID NO: 2 muestra la secuencia completa del elemento del extremo 5' sintético RSyn7 que incorpora al menos tres motivos como OSC de TGACG SEQ ID NO:5 que se indican en negrilla.

Las secuencias que flanquean muchos elementos tales como el motivo TGACG SEQ ID NO: 5 se han mostrado que tienen impactos profundos en las afinidades del enlace de los factores que unen el ADN y así juegan un papel importante como los motivos centrales de ellos mismos. (Burrows *et al.* 1992, Plant Molecular Biology 19:665-675, Shinder *et al.* 1992, Plant Cell 4:1309-1319, Foster *et al.* 1994, FASEBJ 8:192-200). Las secuencias novedosas que flanquean el motivo TGACGs en el promotor Rsyn7 se han determinado y establecen un claro mejoramiento de la actividad transcripcional con varios promotores, particularmente cuando se utilizan con el promotor Core Syn II.

El elemento Rsyn7 del extremo 5' ha sido clonado en el extremo 5' del promotor Core Syn II conduciendo una construcción GUS y tiene los niveles producidos de la actividad GUS en las plantas transgénicas del maíz aproximadamente diez-veces más alto que el promotor ubiquitin, el promotor del maíz más fuerte a la fecha.

En el promotor de la presente invención, una región de activación del extremo 5' (UAR), que es diferente del elemento del extremo 5' sintético, se liga de manera operable al promotor central sintético. También describimos que la UAR puede ser utilizada sola o en combinación con el elemento del extremo 5' sintético descrito aquí. Preferiblemente se utiliza, la región de activación del extremo 5' del promotor 35S del virus del mosaico de la coliflor (CaMV).
 5 Adicionalmente, las secuencias que tienen similitud de secuencia a esta UAR pueden ser utilizadas con tal que tales secuencias retengan la habilidad para potenciar la actividad del promotor. El mejoramiento se puede medir mediante un análisis de los niveles de transcripciones o alternativamente la producción de la proteína.

CaMV 35S UARs se han estudiado bien en el oficio. La secuencia completa de nucleótidos del ADN de doble
 10 cadena circular CaMV ha sido establecida en el oficio. Ver Guillely *et al.* (1980) Cell 21:285-294. El promotor 35S transcribe la transcripción del ARN 35S principal del genoma viral circular por el núcleo de la polimerasa de ARN II. Ver Guillely *et al.* (1982) Cell 30:763-773. Además, las 35S UARs pueden funcionar con un promotor heterólogo e incrementar la expresión de un gen de interés en células y plantas transgénicas. Shah *et al.* (1986) Science 233:478-481. Se han identificado elementos reguladores cis múltiples para la actividad del promotor 35S CaMV. Ver Odell *et al.* (1985) Nature 313:810-812; Fang *et al.* (1989) Plant Cell 1:141-150.

En la presente invención, un fragmento grande de las regiones que activan el extremo 5' (UARs) del promotor 35S CaMV se pueden utilizar para potenciar la actividad del promotor sintético central. El tamaño de la UAR puede, por ejemplo, oscilar aproximadamente de 15 pares de bases a aproximadamente 850 pares de bases, preferiblemente de
 20 aproximadamente 20 a aproximadamente 500, más preferiblemente de aproximadamente 20-25 a aproximadamente 50-200 pares de bases. Una región preferida del 35S CaMV del extremo región 5' incluye las secuencias aproximadamente de -421 a aproximadamente -90. Se reconoce que las modificaciones, en longitud y secuencia de nucleótidos se pueden hacer a la región y aún resultar en la actividad mejorada del promotor sintético central. Tales modificaciones se pueden analizar por el efecto en la actividad utilizando sistemas de expresión como se publica en la Sección Experi-
 25 mental de la presente aplicación. Los números en el diagrama de la secuencia UAR indican la posición del extremo 5' del sitio inicial de transcripción, o +1 posición del gen estructural 35S. Por ejemplo, -25 significa una posición de 25 pares de bases del extremo 5' a partir del sitio inicial de transcripción del gen estructural 35S.

La región de activación del extremo 5' del gen ubiquitin del maíz *Ubi-1* también se describe. La secuencia de la
 30 región reguladora de la transcripción del gen *Ubi-1* se revela en la Patente US No. 5,510,474. Ver también Christensen *et al.* (1992) Plant Mol. Biol. 18:675-689; Cornejo *et al.* (1993) Plant Mol. Biol. 23:567-581; Takimoto *et al.* (1994) Plant Mol. Biol. 26: 1007-1012; y Christensen *et al.* (1996) Transgenic Res. 5:213-218. El UAR del promotor del gen *Ubi-1* comprende preferiblemente de aproximadamente -867 a aproximadamente -54. Según lo indicado arriba para el 35S UAR, también se describen las modificaciones del *Ubi-1* UAR que funcionarían para potenciar la actividad del
 35 promotor central.

Mientras que la secuencia completa del promotor ubiquitin se ha publicado, esta es la primera divulgación de la Ubi UAR. De esta manera, revelamos la UAR del promotor ubiquitin así como la Ubi UAR en combinación con cualquier promotor. Adicionalmente, se describen los métodos para utilizar la Ubi UAR para potenciar la actividad de
 40 los promotores.

Las regiones que activan el extremo 5' según lo descrito aquí se pueden unir con el promotor central sintético y/o otros elementos del extremo 5' por cualquier método convencional que se conoce usualmente en el oficio con tal que un elemento operativo o promotor se construya. Las regiones que activan el extremo 5' generalmente se ligan de manera
 45 operable en el extremo 5' del promotor central. Cuando el elemento del extremo 5' sintético también está presente, la regiones que activan el extremo 5' se pueden unir a ya sea el extremo 5' del elemento del extremo 5' sintético, el extremo 3' del promotor sintético central o insertar entre el elemento del extremo 5' sintético y el promotor central sintético. En una modalidad preferida las regiones que activan el extremo 5' se ligan en la proximidad cercana al elemento del extremo 5' sintético y el promotor central sintético. Por proximidad cercana se pretende dentro de aproximadamente 1
 50 a aproximadamente 50 nucleótidos. Sin embargo, se reconoce que más de 50 nucleótidos pueden separar los elementos. Las regiones que activan el extremo 5' pueden estar en la dirección 5' a 3' o la dirección 3' a 5', pero preferiblemente en la dirección 5' a 3' en el extremo 5' del promotor central sintético o el elemento del extremo 5' sintético.

También describimos que se pueden utilizar, una o múltiples copias de las regiones que activan el extremo 5'. Cuando se utilizan múltiples copias, pueden ser repeticiones en tándem de una UAR o combinaciones de varias UARs. De esta manera, el nivel de expresión de una secuencia de nucleótidos de interés se puede controlar por el número de UARs presentes en la construcción del promotor desde que los resultados indican que se obtuvieron niveles de expresión incrementados, aumentando los números de los UARs. De esta manera, la invención proporciona los métodos para regular los niveles de expresión de un gen o una secuencia de nucleótidos de interés.
 60

Según lo indicado, describimos que múltiples copias de una UAR se pueden utilizar para potenciar la actividad del promotor ligado de manera operable. Como se registra, las múltiples copias de las UARs iguales o diferentes pueden ser utilizados.

Los promotores de esta invención que tienen una UAR según lo descrito arriba se pueden proporcionar en casetes de expresión y tales casetes contenidos en vectores virales o plásmidos. Tales vectores se pueden utilizar para la transformación de células bacterianas y vegetales. Las plantas transgénicas se pueden regenerar finalmente de tales células vegetales transformadas.
 65

Las UARs incorporadas en los promotores de la planta pueden sustancialmente potenciar la actividad de transcripción en plantas transgénicas.

Las construcciones del promotor de la invención se pueden ligar de manera operable a cualquier secuencia de nucleótidos o gen de interés. La construcción del promotor, por ejemplo, se puede utilizar para potenciar la expresión del gen oxalato oxidasa en plantas transgénicas. La oxalato oxidasa es una enzima vegetal implicada en los mecanismos de defensa de las plantas contra el ataque patógeno. La enzima degrada el compuesto químico ácido oxálico secretado por patógenos vegetales. *Ver por ejemplo*, PCT Publication No. WO 92/14824. Incrementando el nivel de oxalato oxidasa en plantas tales como el girasol conducirán al incremento de la resistencia vegetal a patógenos vegetales.

Los siguientes ejemplos son solamente para propósitos de ilustración y no pretenden de ninguna manera limitar el alcance o aplicación de la presente invención. Aquellos de habilidad en el oficio apreciarán que muchas permutaciones se pueden lograr y de hecho tienen la intención de estar dentro del alcance de la invención.

Ejemplo 1

Los plásmidos se designaron utilizando el sitio de clonación múltiple de pBlueScriptIIS+ de Stratagene. (Para facilitar la clonación de la diferente combinación de los elementos). Los oligonucleótidos que contienen las secuencias de los elementos se sintetizaron con la restricción de los sitios de la endonucleasa en los extremos. De esta manera los elementos podrían ser adicionados o retirados y reemplazados según sea necesario. GUS y Luciferasa se utilizaron para genes indicadores.

Para los ensayos transitorios, el ADN del plásmido se introdujo en semilleros de maíz de 3-días-de edad intactos, mediante el bombardeo de partículas. Continuando con 16 horas de incubación a 25EC en la oscuridad, la expresión se analizó midiendo la actividad de la enzima GUS en extractos de raíces y brotes para cada semillero para determinar si cualquier expresión de tejido preferido se demostró. La actividad GUS se midió utilizando un kit de ensayo GUS-Light de Tropix (47 Wiggins Avenue, Bedford, MA 01730).

Las construcciones que dieron niveles altos de la expresión se introdujeron en una línea celular para producir transformables estables. Estos transformables estables (TO) se analizaron por PCR para determinar la presencia del gen GUS por ensayo de MUG (4- metilumelliferil-glucuronida) para cuantificar el nivel de actividad de la proteína GUS que se produce. Cuando las plantas se alistaron para ser transferidas al invernadero se analizaron histoquímicamente con X-gluc para determinar donde el producto GUS fue sintetizado. Las plantas que demostraron los niveles de expresión preferidos se cultivaron en el invernadero para la etapa V6.

Ejemplo 2

Construcción de Plásmidos que Contienen el Promotor Core Syn II

Las técnicas de biología molecular estándar se realizaron de acuerdo con Maniantis *et al.* (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. Todos los plásmidos utilizados en la invención se pueden preparar de acuerdo con las direcciones de la especificación por alguien de ordinaria habilidad en el oficio sin la indebida experimentación empleando materiales disponibles fácilmente en el oficio.

Oligos N306 SEQ ID NO:6 5'-TCGACACTGC AGCTCTAGGG ATGGTAGCGC AGGGTGCGTA GGTACG TATT TATAGCCGCT CGAGTG-3' y N307 SEQ ID NO: 7 5'-GATCCACTCG AGCGGCTATA AATACGTACC TACGCACCCT GCGCTACCAT CCTAGAGCT GCAGTG-3' se sintetizaron de acuerdo con las direcciones en un sintetizador de ADN automatizado (tal como Applied Biosystems Inc. ADN Synthesizer (Model 380B)). Estos sintetizadores automatizados son disponibles comercialmente. Los oligos luego se ligaron al fragmento BamHI del plásmido pBlueScriptIIS+ que comprende del gen $\beta\beta$ -glucoronidasa interrumpido por la región 1 del intrón ADHI del maíz. Un mapa de un plásmido que incorpora ambos el promotor Core Syn II y el elemento del extremo 5' se revelan según la Figura 4. Varios otros plásmidos se describen en la Figura 5. Los números del plásmido se muestran a la derecha de cada diagrama del promotor con la leyenda correspondiente colocada debajo de los diagramas. El diagrama arriba muestra la unidad del transplante transcripcional completa con los diagramas subsiguientes enfocándose en las diferencias salientes entre 35S y el promotor Central Syn. La leyenda muestra el número y naturaleza de los varios subelementos del promotor, la secuencia si es relativamente corta, la fuente del elemento y posición relativa al inicio de la transcripción.

La secuencia del promotor central consiste de 35 pares de bases con los sitios de enzima del extremo 5' de una caja TATA y un inicio de la transcripción con 10 a 15 pares de bases del extremo 3'. Los elementos del extremo 5' (4-sitios de enlace Gal, Rsyn, AT-GBL etc.) se fundieron a la secuencia central con los genes marcadores diferentes y ADHI-intrón (LUC o GUS) y ambas se mostraron funcionales en los ensayos transitorios (Figura 6) y las plantas transformadas estables Rsyn (Figura 7).

ES 2 292 234 T3

Ejemplo 3

Construcción del Elemento del extremo 5' Rsyn7 Fundido al Promotor Core Syn II Resultando en los Plásmidos PHP5903 y PHP6086

Los oligos para la construcción del subelemento N del promotor Rsyn7 1965: (SEQ ID NO:8) GATCCTATGA CG TATGGTAT GACGTGTGTT CAAGATGATG ACTTCAAACC TACCTATGAC GTATGGTATG ACGTGTGTCG ACTGATGACT TA- 3' y N1966: (SEQ ID NO:9) GATCTAAGTC ATCAGTCGAC AGACGTCATA CCATACGT CA TAGGTAGGTT TGAAGTCATC ATCTTGAACA CACGTCATAC CATACTGCA TAG-3' se sintetizaron según lo descrito previamente. Los oligos se hibridaron y clonaron en un plásmido PHP3398 del extremo 5' de la secuencia Core Syn II y resultan en varias versiones de la secuencia Rsyn7 original debido a las delecciones espontáneas. La versión Rsyn7-2 involucró una delección de base única que resulta en un motivo reiterativo 3X TGACG del extremo 5' del promotor Core Syn II (Rsyn7::LUC, P5903). La secuencia LUC codificante se reemplazó por la secuencia codificante GUS para producir la construcción Rsyn7::GUS P6086. P6086 luego se introdujo en el maíz transgénico que resulta en niveles altos de actividad constitutiva en cuatro de los seis eventos activos examinados (Figura 7).

La progenie de las plantas T0 de varios eventos de transformación se examinaron y la actividad de GUS fluctuó de 1 a 400 PPM (microgramos de enzima GUS/GFW en tejido de raíz de un semillero de 7-días de edad) Figura 8. Estas plantas T0 y T1 generalmente produjeron una actividad GUS 4X-10X mayor que las plantas hospedadas en el gen indicador ubiquitin::GUS.

Ejemplo 4

Transformación y Expresión con el Promotor Core Syn II y/o Elemento del extremo 5' Rsyn7

Utilizando ensayos de bombardeo transitorios la secuencia Central Syn II del promotor se comparó contra la secuencia central 35S ya sea solo o en conjunción con numerosos elementos de activación. La Figura 6 es una descripción de los datos de ensayo transitorios utilizando los plásmidos que incorporan las secuencias del promotor central descritas aquí y muestra la actividad transitoria de GUS o LUC en raíces de maíz de tres-días de edad o callos BMS bombardeados con el promotor quimérico::GUS o construcciones LUC. El -33 CaMV35S en las versiones del promotor Core Syn II de las construcciones del promotor sintético::GUS (o LUC) se bombardearon en raíces de tres-días de edad (o callos BMS cultivados según lo descrito más adelante) y ensayados para la actividad de la enzima 20 horas después de los bombardeos. Los datos mostrados son las unidades de enzimas crudas de una recopilación de al menos tres experimentos y no han sido normalizados de cualquier manera debido a la variabilidad inherente de los ensayos transitorios. Los plásmidos control 1654 y 3537 son las construcciones LUC probadas en los callos BMS del maíz. Hay aproximadamente 4 a 20 veces de diferencia en actividad transitoria entre las versiones Core 35S y Syn II. El eje Y es en escala logarítmica. Ambos promotores centrales fueron conducidos en un GUS que contenía una construcción (Figuras 4 y 5) y generaron un nivel basal de actividad (Figura 6). Sin embargo cuando los elementos activadores se colocaron en el extremo 5' del motivo TATA, el Core Syn II proporcionaron generalmente los niveles más altos de actividad (2-4 veces mejores) en las células del maíz que cuando los elementos activadores se colocaron en el extremo 5' del centro 35S (Figura 6).

La secuencia Core Syn II se ha mostrado para potenciar la actividad en plantas estables transformadas. Además con ciertas secuencias activador del extremo 5' de los niveles de actividad del elemento TATA en las plantas estables transformadas de maíz alcanzan niveles diez veces más grandes que las construcciones ubiquitin del maíz que producen niveles sumamente altos de actividad.

Las Figuras 7 y 8 muestran los niveles de actividad GUS de tejidos aislados de la etapa VT de las plantas T0 y tejido de raíz de los semilleros T1, respectivamente. Estos datos demostraron que esta secuencia central puede participar en potenciar los niveles muy altos de actividad como un socio funcional para los promotores quiméricos activos. La Figura 7 muestra la actividad Rsyn7::GUS (6086) en las plantas del maíz T0. Las plantas de la etapa VT con espigas pos polinización de 3 a 8 días se disecaron y analizaron para la actividad GUS. 7A describe la expresión GUS en tejidos designados. 7B describe un esquemático de una planta de maíz indicando los sitios de medición. Se analizaron las plantas de los eventos T0 que demostraron una fluctuación de las actividades con el promotor Rsyn7. Nuevamente se registra la escala logarítmica. La actividad oscila para las plantas UBI::GUS se indica a la derecha de la gráfica para comparaciones. Estos datos demostraron que el promotor Rsyn7 puede incrementar la actividad a niveles diez-veces arriba del promotor ubiquitin aún muestra poca preferencia del tejido haciendo el Rsyn7 apropiado como un promotor constitutivo fuerte.

La Figura 8 describe la actividad GUS en segmentos de raíz de una población segregada de los semilleros transgénicos T1 del maíz que contiene la construcción Rsyn7::GUS (6086) o la UBI::GUS (3953). Los segmentos de 1 cm de raíz de semilleros transgénicos de maíz de seis a siete-días de edad se disecaron, pesaron y analizaron para GUS utilizando el kit GUS-light. La actividad se representa como partes por millón en peso fresco. La actividad de la raíz de varias plantas T1 que hospedan el promotor Rsyn7::GUS muestra una actividad más alta que muchos de los niveles de actividad producida por el promotor UBI. Esto es consistente con los datos de la planta transgénica T0. Los niveles de actividad en hojas jóvenes que contienen Rsyn7::GUS también son mucho más altos que los niveles de actividad de las hojas jóvenes que contienen UBI::GUS- (datos no mostrados). La secuencia Central Syn II se muestra que funciona

bien con una variedad de elementos del extremo 5' que incluyen sitios de enlace GAL 4, elementos Rsyn7, elementos GBL, etc.

La Figura 9 muestra la expresión GUS de tres promotores sintéticos en plantas T0 transgénicas del maíz. Los tejidos disecados (Ver Figura 7B) de las plantas T0 transgénicas de la etapa VT que hospedan Rsyn7 (Rsyn), Atsyn o las construcciones Syn II Core solo promotor (syn-core)::GUS se analizaron cuantitativamente para la actividad GUS. Cada círculo representa un promedio de la actividad del tejido de las plantas transgénicas del maíz de un evento de transformación único. El motivo TGACG corresponde a la secuencia Rsyn7 y el motivo "AT-com" se refiere al elemento combinado como AT consenso, Atcom, de W. Gurley, *et al.* 1993. In: Control of Plant Gene Expression. ed. by Desh Pal Verma. CRC press, Boca Raton, FL. pp. 103-123. Syn-core se refiere a la secuencia Central Syn II del promotor que contiene el elemento TATA y el inicio de la transcripción.

Ejemplo 5

Construcción del promotor Rsyn7 que tiene la regiones que activan el extremo 5' del gen CaMV 35S y el gen Ubi-1 del maíz

Para construir un vector de expresión que tiene 35SU (regiones que activan el extremo 5' del 35S gen)-promotor Rsyn7, PHP413 se digirió con BgIII y EcoRV. Los extremos alternados/cohesivos del vector linealizado se rellenaron mediante Klenow en presencia de dNTP. El fragmento CaMV 2x fue extremo como ligador en BamHI digerido PHP6086 después de rellenar los terminales BamHI. El fragmento CaMV 35S-Rsyn7 luego se recortó a partir de la nueva construcción por digestión con XbaI y Pst 1, y se ligó en el vector XbaI-Pst1 de 4 kb a partir de PHP9925 para formar el vector de expresión de 35SU-Rsyn7::GUS (PHP9778).

Para construir un vector de expresión que tiene UbiU (elementos que activan el extremo 5' del gen del maíz Ubi-1)-promotor Rsyn7, el fragmento XbaI-SpeI de PHP8277 se ligó en el sitio XbaI de PHP6086 para formar PHP10539, en el sitio XbaI de PHP10970 para formar PHP10971, y en el sitio XbaI de PHP10971 para formar el vector de expresión de Ubi-1-Rsyn7::GUS (PHP10972).

Aquellas secuencias no referenciadas de otra manera incluyen:

SEQ ID NO: 11 expone el 35S UAR.

SEQ ID NO: 12 expone la secuencia del promotor SCP1, (35S UAR ligado de manera operable al promotor central de la SEQ ID NO: 1).

SEQ ID NO: 13 expone el Ubi1 UAR.

SEQ ID NO: 14 expone SCP1 ligado de manera operable a la secuencia codificante oxalato oxidasa ligada de manera operable con el PinII terminador.

SEQ ID NO: 16 expone la secuencia del promotor UCP2 (2 copias de Ubi1 UAR operable con el promotor central).

SEQ ID NO: 18 expone la secuencia del promotor UCP4 (4 copias de Ubi1 UAR operable con el promotor central).

Ejemplo 6

Transformación y Expresión de construcciones del promotor

Los diversos fragmentos promotores::GUS se clonaron en un vector binario Bin9 que contiene ALS3::NPTII como un marcador de selección para generar callos de girasol transgénicos o *Arabidopsis*.

Para la expresión transitoria, semillas del girasol SMF3 se plantaron en el invernadero. Las semillas de 15-días-de edad después de la polinización se colectaron de las plantas y se utilizaron en el sistema de expresión transitoria. Después de remover el pericarpio, los cotiledones con la semilla cubierta se esterilizaron por la incubación en lejía al 20% a RT por 15 minutos, y se lavó cuatro veces con agua destilada esterilizada dos veces. Luego los cotiledones se incubaron en un filtro de 3MM empapado con medio MS durante la noche antes se bombardearon de acuerdo con el método revelado por Klein *et al.* (1989) Proc. Natl. Acad. Sci. USA 86:6681-6685. La actividad GUS se analizó 20 horas después del bombardeo utilizando el kit GUS-Light de ensayo de Tropix de acuerdo con el protocolo del fabricante.

Para la transformación del disco de la hoja, la hoja joven del girasol SMF3 expandida del girasol de 30-días-de edad se cosechó y esterilizó en lejía al 20% con un par de gotas de Tween 20 por 20 minutos. Los discos de la hoja se prepararon a partir de hojas estériles después de lavarlas con agua destilada esterilizada dos veces 4 veces. Los discos de la hoja luego se incubaron por 10 minutos en medio de inoculación (MES 12.5 mM, NH₄Cl 1 g/l, y C 0.3 g/l de MgSO₄) que contiene Agrobacterium (EHA105) transformado con las construcciones del vector que se probara a A600=0.75. Los discos de la hoja luego se cultivaron por 3 días en medio sin-selección y luego se transfirieron al medio de selección.

ES 2 292 234 T3

Para transformar la *Arabidopsis*, *Arabidopsis* se cultivó en el invernadero hasta la etapa cuando los brotes comienzan a emerger en 15 plantas/maceta. Los brotes emergentes se sujetaron completamente para impulsar el crecimiento de los brotes secundarios múltiples. Después de 7 días, las plantas se alistaron para la infiltración. El *Agrobacterium* (EHA105) que lleva la construcción que se probará se cultivó a 28°C hasta cuando A600 fue entre 0.65 y 0.8. Las células se cosecharon en medio de inoculación (4.3g/l de sal MS, 0.5 mg/l de ácido nicotínico, 0.5 mg/l de piridoxina-HCl, 1 mg/l de Tiamina-HCl, 0.1 g de mio-inositol, 1 g/l de casaminoácidos, 0.01 mg de BAP, 68.5 g/l de sacarosa, y 36 g/l de glucosa) a A600 de 7.5.

Las plantas sujetadas con grapas para ser transformadas se invirtieron en un 250 ml vaso de precipitados que contiene la solución anterior del *Agrobacterium*. El vaso de precipitados se colocó en una vasija de campana y se sometió a vacío hasta que se forman burbujas en la superficie de las hojas y del tallo. Después de 15 minutos de infiltración, el vacío se liberó y las plantas se retiraron del vaso de precipitados, acostadas en su lado en un plástico plano, y se cubrió con un manto plástico. Las plantas se colocaron de pie y crecieron en el invernadero por cuatro semanas antes de que las semillas se cosecharan. Las semillas transgénicas se seleccionaron sembrando las semillas en una placa de medio que contiene 65 µg/ml de Kanamicina.

En ensayos de expresión transitoria, el promotor Rsyn7 en PHP10464 tiene el 15% de la actividad del promotor 35S (PHP9925), mientras que el promotor-35SU Rsyn7 (PHP9778) (más adelante promotor SCP1) tiene aproximadamente 107% de la actividad del promotor 35S (Figura 10). De esta manera, la región de activación del extremo 5' del gen CaMV 35S incrementó la actividad de Rsyn7 por aproximadamente 6 veces.

El elemento del extremo 5' (UbiU) del maíz *Ubi-l* tiene efectos similares en el promotor Rsyn7 en ensayos transitorios. Cuando la región de activación del extremo 5' (UAR) del maíz *Ubi-l* se fusionó a Rsyn7, la actividad de la enzima GUS incrementó con el número de copias del UAR. En la presencia de tres copias del UbiU, la actividad GUS incrementó aproximadamente 4-veces. Este efecto aditivo de UbiU no se observó cuando se coloca en el contexto del promotor ubiquitin del maíz (PHP 11974). Esto sugiere que el reemplazo del promotor central del maíz Ubi1 con Rsyn7 puede convertir el promotor Ubi1 monocotiledóneo en un promotor altamente activo en las plantas dicotiledóneas (Figura 11).

En los callos estables del girasol transformados, la expresión GUS es 20% más alta detrás del control del promotor 35SURsyn7 (SCP1 promotor) que cuando detrás del control del promotor 35S CaMV (Figura 12).

Los resultados del ensayo de los callos transgénicos se dan en la Figura 13. El promotor Rsyn7 que contiene una sola copia de UbiU (PHP10991) (de ahora en adelante promotor UCP 1) (SEQ ID NO: 15) mostró una actividad del promotor de aproximadamente 3 veces que la del promotor 35SU-Rsyn7 (SCP1) (PHP 10940). Tres copias de UbiU (PHP 10993) incrementaron la actividad del promotor Rsyn7 a aproximadamente 7 veces que la del 35SU-promotor Rsyn7. El 3xUbiU-Rsyn7 (UCP3 promotor) (SEQ ID NO: 17) es con mucho el promotor más fuerte en tejidos del girasol.

Para determinar la actividad y especificidad del tejido de los promotores Rsyn7 mejorados, el girasol transformado estable y *Arabidopsis* se generaron para la transformación mediada por el *Agrobacterium*. La tinción histoquímica de la expresión GUS en *Arabidopsis* T1 transgénica indica que 35SU-Rsyn7 (SCP1 promotor) (PHP10940) tiene idéntica especificidad del tejido y actividad similar que el promotor 35S CaMV (PHP 10989). Ambos promotores expresan GUS en la hoja, tallo, peciolo, y partes florales. UbiU-Rsyn7 (USCP1 promotor) (PHP10991) muestra actividad más alta que el promotor *Ubi-l* del maíz (PHP11031) en tejidos de tallo y hoja de *Arabidopsis*.

Todas las publicaciones y aplicaciones de las patentes mencionadas en la especificación son indicativas del nivel de aquellos de habilidad en el oficio para los cuales esta invención concierne.

Aunque la invención precedente se ha descrito en un cierto detalle por medio de ilustración y ejemplo para propósitos de claridad de comprensión, será obvio que ciertos cambios y modificaciones se pueden practicar dentro del alcance de las reivindicaciones anexas.

Referencias citadas en la descripción

Esta lista de referencias citadas por el aspirante es únicamente para conveniencia del lector. No forma parte del documento de la Patente Europea. Aún cuando se ha tomado gran cuidado en recopilar las referencias, los errores u omisiones no se pueden excluir y la EPO desconoce toda responsabilidad en este respecto.

Documentos de patentes citados en la descripción

- WO 840291 A, Rogers [0010]
- US 4990607 A, Katagiri [0037] [0074]
- US 5380831 A [0060]

- US 5436391 A [0060]
- US 5023179 A [0074]
- 5 • US 5510474 A [0083]
- WO 9214824 A [0090]
- US 9903863 W [0121]

Literatura no-patente citados en la descripción

- **ZENZIE-GREGORY** *et al. J. Biol. Chem.*, 1992, vol. 267, 2823-2830 [0004]
- 15 • **GUERINEAU** *et al. Mol. Gen. Genet.*, 1991, vol. 262, 141-144 [0059]
- **PROUDFOOT** . *Cell*, 1991, vol. 64, 671-674 [0059]
- **SANFACON** *et al. Genes Dev.*, 1991, vol. 5, 141-149 [0059]
- 20 • **MOGEN** *et al. Plant Cell*, 1990, vol. 2, 1261-1272 [0059]
- **MUNROE** *et al. Gene*, 1990, vol. 91, 151-158 [0059]
- 25 • **BALLAS** *et al. Nucleic Acids Res.*, 1989, vol. 17, 7891-7903 [0059]
- **JOSHI** *et al. Nucleic Acid Res.*, 1987, vol. 15, 9627-9639 [0059]
- **MURRAY** *et al. Nucleic Acids Res.*, 1989, vol. 17, 477-498 [0060]
- 30 • **GRUBER** *et al. Vectors for Plant Transformation*, in *Methods in Plant Molecular Biology & Biotechnology*. *CRC Press*, 1993, 89-119 [0062]
- Procedures for Introducing Foreign DNA into Plants. MAKI *et al. Methods in Plant Molecular Biology & Bio-*
35 *technology. CRC Press*, 1993, 67-88 [0063]
- Cell-Tissue Culture and *In-Vitro* Manipulation. PHILLIPS *et al. Corn & Corn Improvement*. American Society
of Agronomy Inc, 1988, 345-387 [0063]
- 40 • Direct DNA transfer into intact plant cells via microprojectile bombardment. TOMES *et al. Plant Cell, Tissue*
and Organ Culture: Fundamental Methods. Springer-Verlag, 1995 [0065]
- K. LINDSEY *et al. Tagging Genomic Sequences That Direct Transgene Expression by Activation of a Promoter*
Trap in Plants. *Transgenic Research*, 1993, vol. 2, 33-47 [0066]
- 45 • D. AUCH; RETH *et al. Exon Trap Cloning: Using PCR to Rapidly Detect and Clone Exons from Genomic*
DNA fragments. *Nucleic Acids Research*, vol. 18 (22), 6743 [0066]
- KIM *et al. Nature*, 1993, vol. 365, 512-520 [0068]
- 50 • YAURAWJ *et al. Biological Abstracts*, vol. 47 [0068]
- Consensus Sequences for Plant Minimal Promoters. *Annual Meeting of the American Society of Plant Physiolo-*
gists, 29 July 1995 [0068]
- 55 • *Plant Physiology*, 1995, vol. 108, 114 [0068]
- ELLIS *et al. EMBO J.*, 1987, vol. 6, 11-16 [0074]
- 60 • KATAGIRI *et al. Nature*, 1989, vol. 340, 727-730 [0075]
- BURROWS *et al. Plant Molecular Biology*, 1992, vol. 19, 665-675 [0078]
- SHINDER *et al. Plant Cell*, 1992, vol. 4, 1309-1319 [0078]
- 65 • FOSTER *et al. FASEBJ*, 1994, vol. 8, 192-200 [0078]
- GUILLEY *et al. Cell*, 1980, vol. 21, 285-294 [0081]

ES 2 292 234 T3

- **GUILLEY** *et al.* *Cell*, 1982, vol. 30, 763-773 [0081]
- **SHAH** *et al.* *Science*, 1986, vol. 233, 478-481 [0081]
- 5 • **ODELL** *et al.* *Nature*, 1985, vol. 313, 810-812 [0081]
- **FANG** *et al.* *Plant Cell*, 1989, vol. 1, 141-150 [0081]
- **CHRISTENSEN** *et al.* *Plant Mol. Biol.*, 1992, vol. 18, 675-689 [0083]
- 10 • **CORNEJO** *et al.* *Plant Mol. Biol.*, 1993, vol. 23, 567-581 [0083]
- **TAKIMOTO** *et al.* *Plant Mol. Biol.*, 1994, vol. 26, 1007-1012 [0083]
- 15 • **CHRISTENSEN** *et al.* *Transgenic Res.*, 1996, vol. 5, 213-218 [0083]
- **MANIANTIS** *et al.* *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory*, 1982 [0095]
- **W. GURLEY** *et al.* *Control of Plant Gene Expression*, 1993, 103-123 [0104]
- 20 • **KLEIN** *et al.* *Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 6681-6685 [0109]

REIVINDICACIONES

1. Una secuencia de ADN del promotor vegetal sintético, la citada secuencia comprende:

un motivo TATA;

un sitio inicial de transcripción;

una región entre dicho motivo TATA y dicho sitio inicial que es al menos 64% rico en GC; un elemento del extremo 5'; y una región de activación del extremo 5', en donde dicha secuencia del promotor comprende una secuencia seleccionada del grupo que consiste de:

a) una secuencia de nucleótidos publicada en la SEQ ID NO: 12, y

b) una secuencia de nucleótidos que tiene al menos 90% de identidad con la secuencia a).

2. Una secuencia de ADN del promotor vegetal sintético de acuerdo con la reivindicación 1 que comprende una secuencia de nucleótidos publicada en la SEQ ID NO: 12.

3. Una secuencia de ADN del promotor vegetal sintético de acuerdo con la reivindicación 1, dicha secuencia que contiene en el enlace operable:

una secuencia del promotor sintético central que comprende la secuencia presentada en la SEQ ID NO: 1 o la SEQ ID NO: 10;

un elemento del extremo 5' sintético ligado de manera operable a dicho promotor sintético central, dicho elemento del extremo 5' sintético que comprende la secuencia presentada en la SEQ ID NO: 2; y una región de activación del extremo 5'.

4. Una secuencia de ADN del promotor vegetal sintético de acuerdo con la reivindicación 3, en donde dicha región de activación del extremo 5' comprende una región de activación del extremo 5' de CaMV 35S.

5. Una secuencia de ADN del promotor vegetal sintético de acuerdo con la reivindicación 4, en donde dicha región de activación del extremo 5' de CaMV 35S comprende del nucleótido -430 al nucleótido -90 del gen CaMV 35S.

6. Una secuencia de ADN del promotor vegetal sintético de acuerdo con la reivindicación 4, en donde dicha región de activación del extremo 5' de CaMV 35S comprende la secuencia de nucleótidos presentada en la SEQ ID NO: 11.

7. Un casete de expresión que comprende:

una secuencia de ADN del promotor sintético de acuerdo con cualquiera de las precedentes reivindicaciones;

un gen estructural ligado de manera operable a dicho promotor; y

una señal de poliadenilación del sitio de terminación de la transcripción.

8. Un casete de expresión de acuerdo con la reivindicación 7, en donde dicho gen estructural es un gen oxalato oxidasa.

9. Un vector de ácidos nucleicos que comprende el promotor de cualquiera de las reivindicaciones 1 a 6 ligado de manera operable a un gen estructural.

10. El vector de la reivindicación 9, en donde dicho vector es un vector de clonación.

11. El vector de la reivindicación 9 o 10, en donde dicho vector es un vector de expresión.

12. El vector de cualquiera de las reivindicaciones 9 a 11, que además comprende un gen marcador para la selección de las células transformadas.

13. El vector de la reivindicación 12, en donde dicho gen marcador es un gen de resistencia a un antibiótico.

14. El vector de cualquiera de las reivindicaciones 9 a 13, que adicionalmente comprende una señal de poliadenilación.

15. Un célula huésped procariota o eucariota transformada con el vector de ácidos nucleicos de cualquiera de las reivindicaciones 9 a 14.

ES 2 292 234 T3

16. Un método de control del nivel de expresión de una secuencia de nucleótidos transgénica en una célula vegetal, dicho método que comprende la transformación de dicha célula vegetal con un casete de expresión de acuerdo con la reivindicación 7 u 8.

5 17. Una planta transgénica que comprende una célula vegetal o un progenitor de esta, que ha sido transformado con el vector de cualquiera de las reivindicaciones 9 a 14.

18. Una planta transformada que contiene en su genoma el casete de expresión de la reivindicación 7 o 8.

10 19. La planta de la reivindicación 17 o 18, en donde dicha planta es una monocotiledónea.

20. La planta de la reivindicación 17 o 18, en donde dicha planta es una dicotiledónea.

15 21. La planta de la reivindicación 20, en donde dicha planta es un girasol.

22. Una planta transgénica que expresa un gen oxalato oxidasa exógeno en un nivel alto en donde dicho gen oxalato oxidasa está bajo el control transcripcional del promotor recombinante de cualquiera de las reivindicaciones 1 a 6.

20 23. La planta transgénica de la reivindicación 22, en donde dicha planta es una dicotiledónea.

24. La planta transgénica de la reivindicación 23, en donde dicha dicotiledónea es un girasol.

25 25. Semillas de la planta de cualquiera de las reivindicaciones 17 a 24, semillas que contienen en su genoma el promotor recombinante de cualquiera de las reivindicaciones 1 a 6.

30

35

40

45

50

55

60

65

(SEQ ID NO:3)
5...TATA(AT)A(AT)A...25bp.....TYYTCAT(AC)AA...3'

FIG. 1.

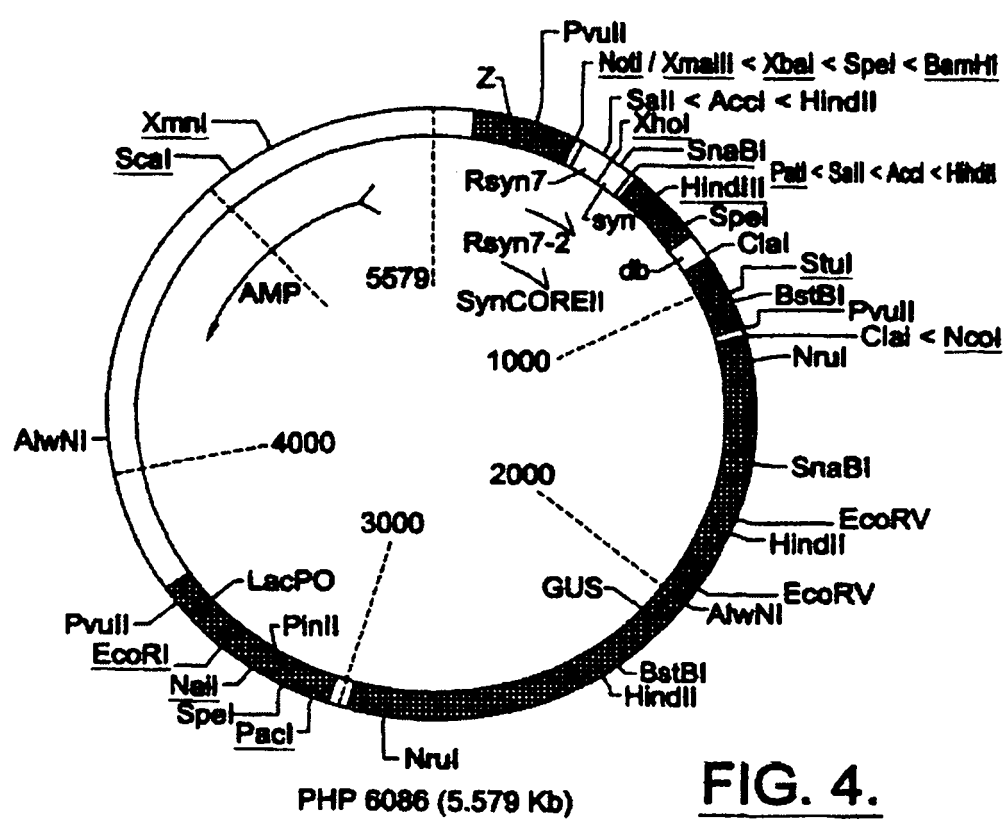
SYNII CORE (64%)
(SEQ ID NO:1)
...GGATCCACTCGAGCGGCTATAAATACGTACCTACGCAAGCTGGGCTACCATCCCGAGCACTGCAGTGTGAC...

CAM35S (40%)
(SEQ ID NO:4)
...CCTCTATATAAGCAAGTTCATTTCATTGGAGAGGAAACG...

FIG. 2.

RSYN7 (SEQ ID NO:2)
5' GGATCCTATGCGTATGGTATGACGTGTGTTCAAGATGATGACTTCAAACCTACCTATGACGTAT
GGTATGACGTGTGTGCTGACTGATGACTTAGATC 3'

FIG. 3.



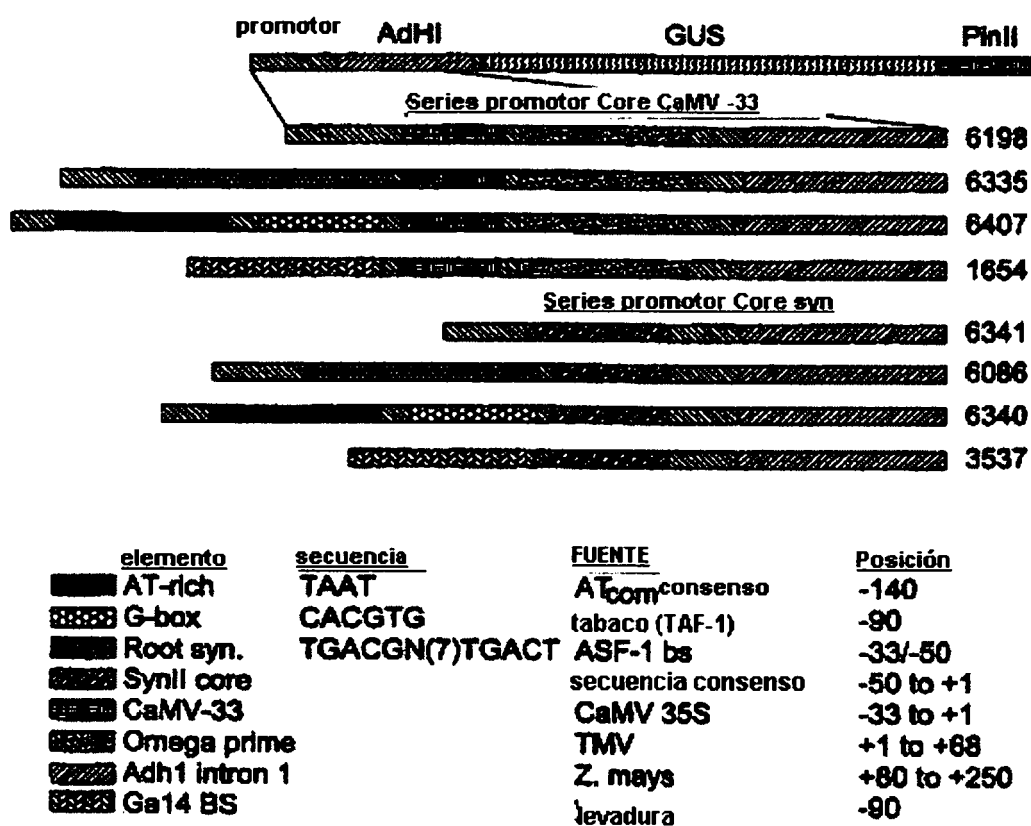


FIG. 5.

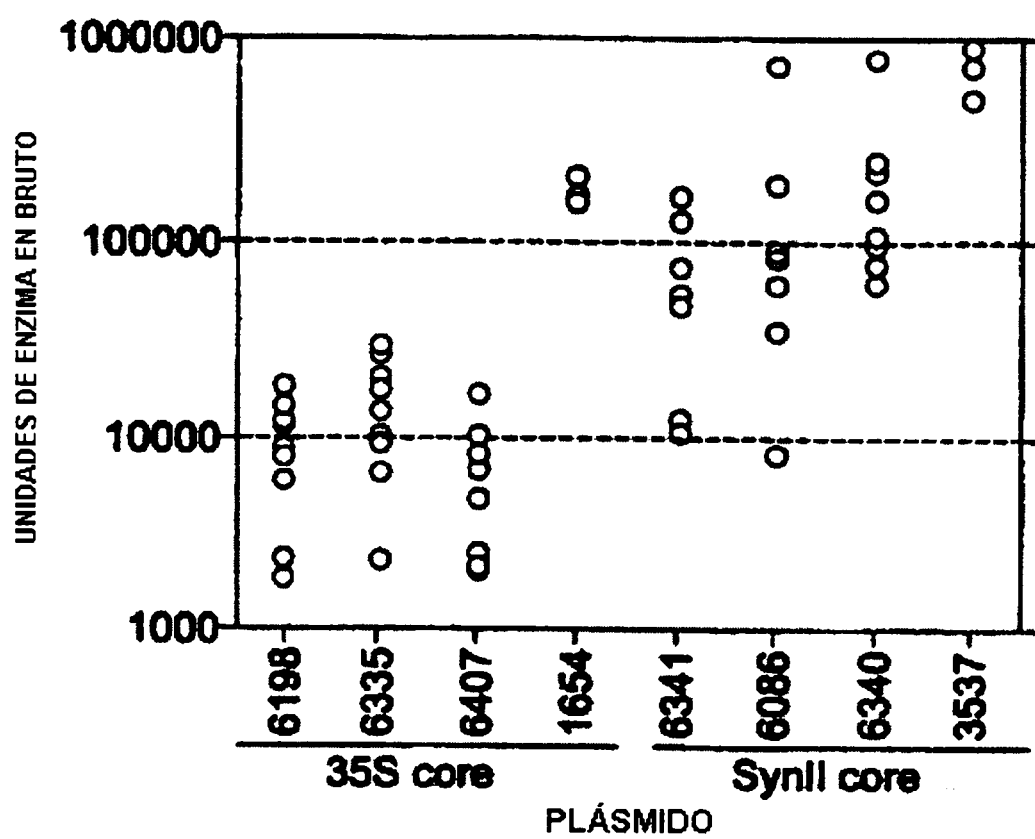


FIG. 6.

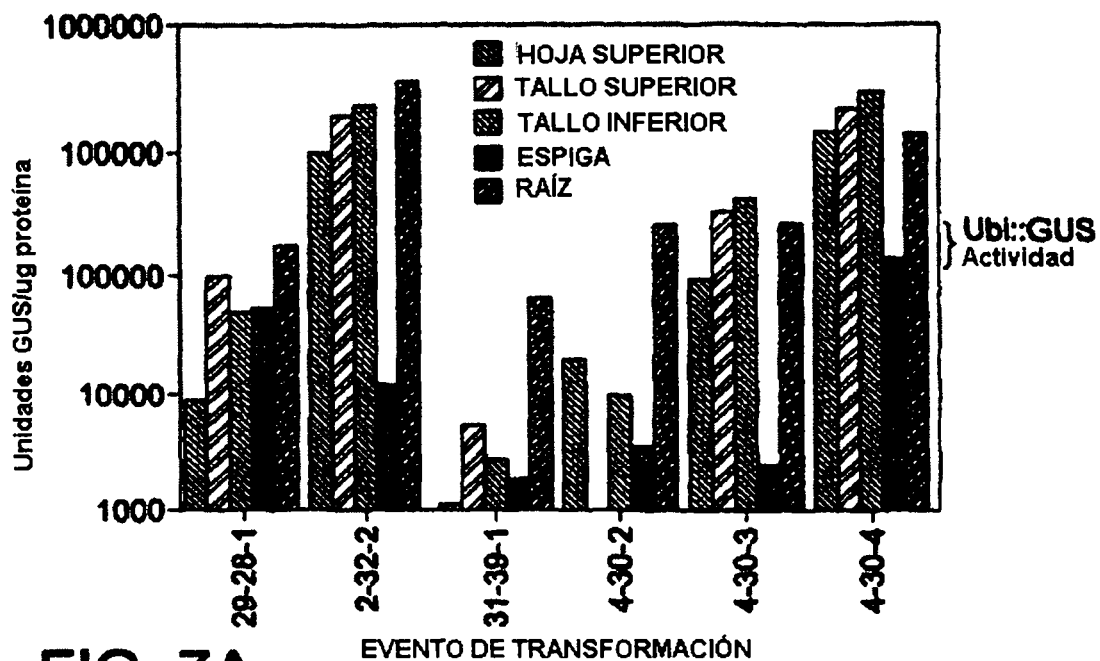


FIG. 7A.

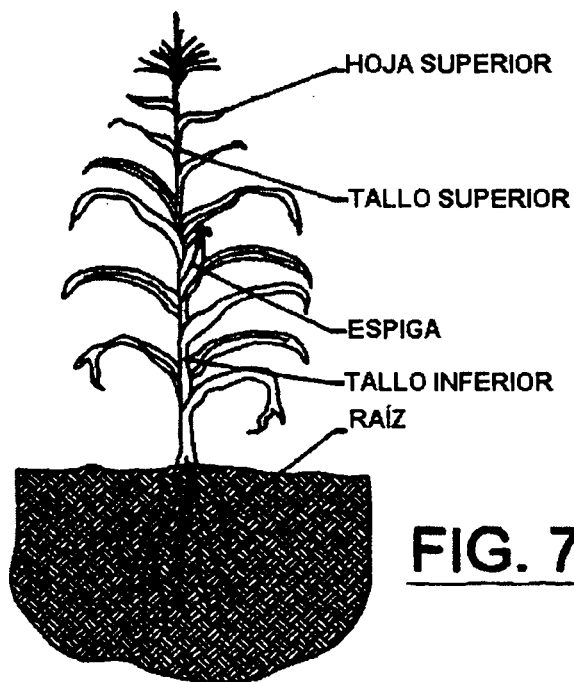


FIG. 7B.

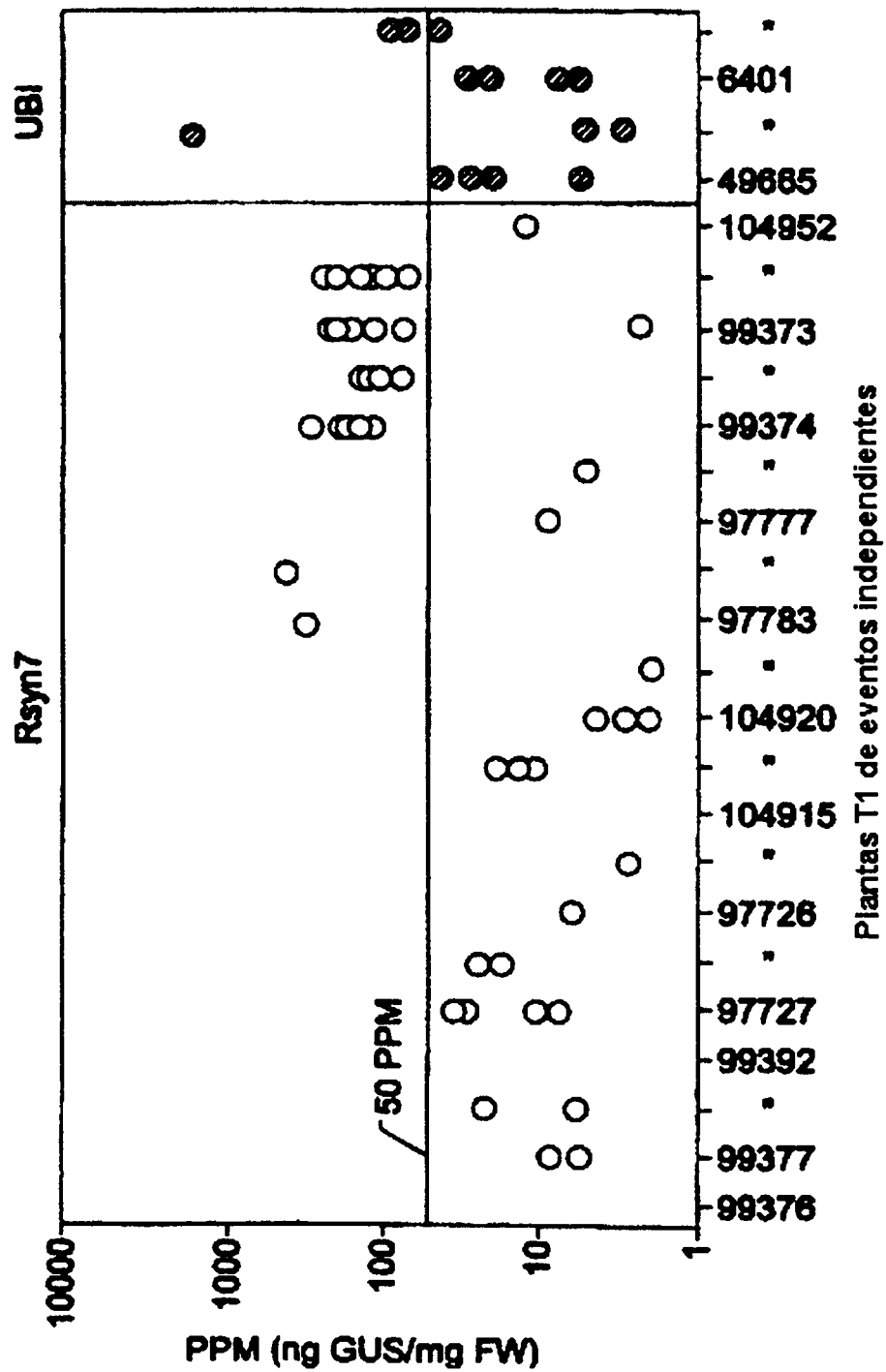
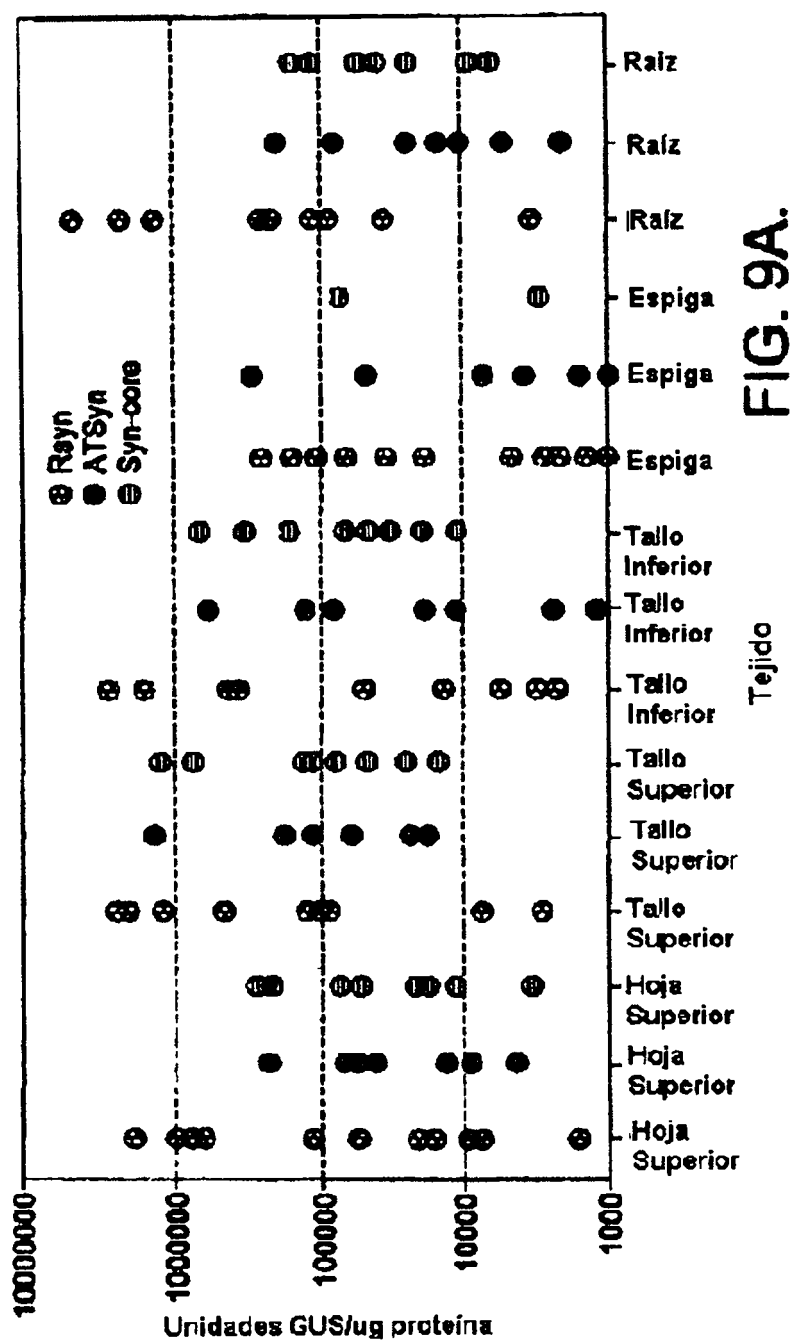


FIG. 8.



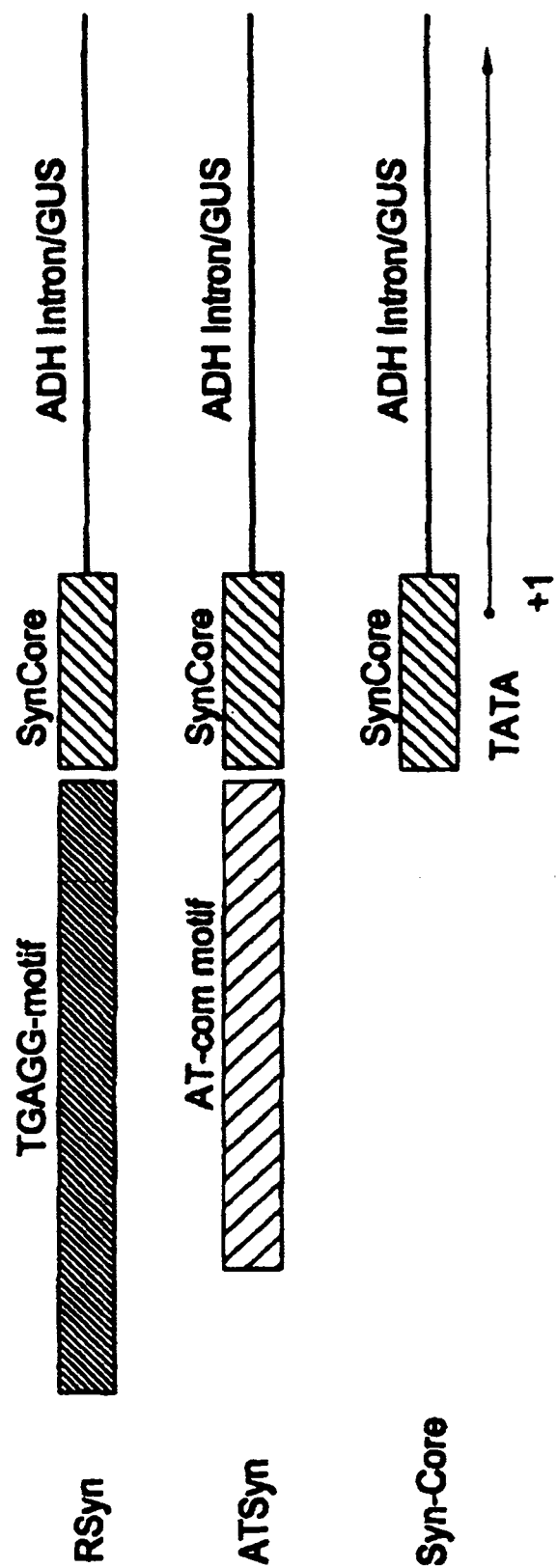


FIG. 9B.

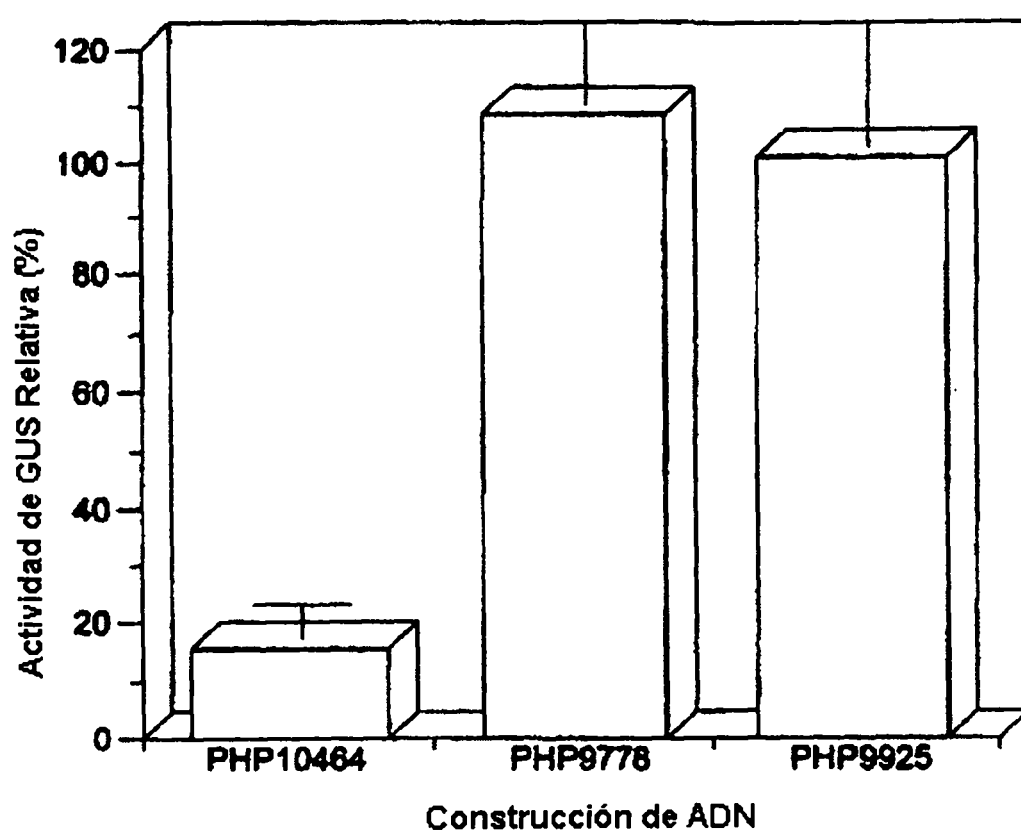


FIG. 10.

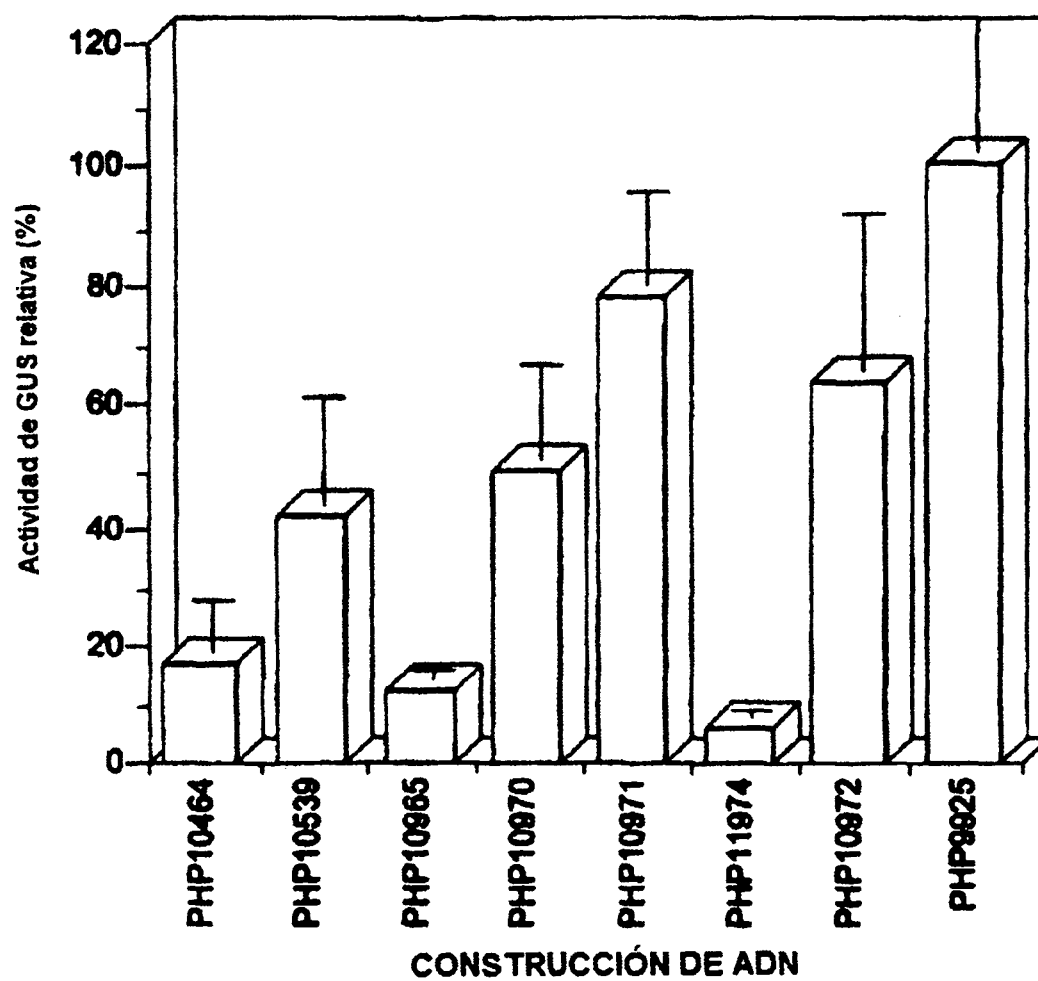


FIG. 11

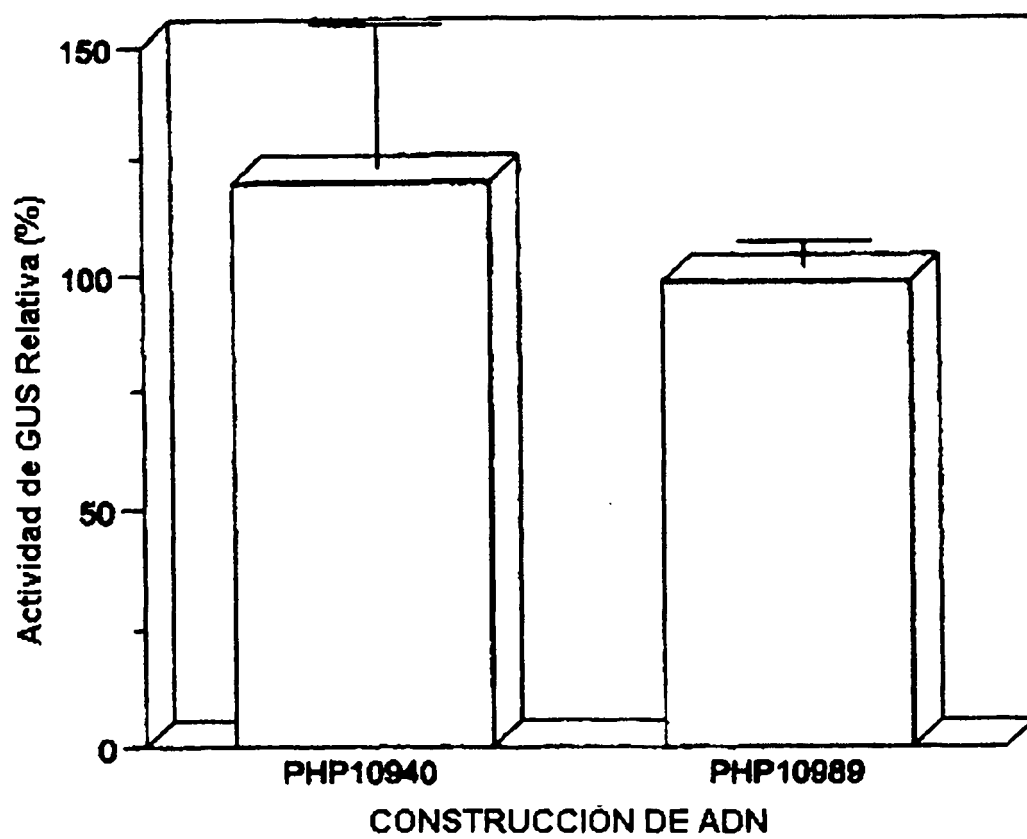


FIG. 12.

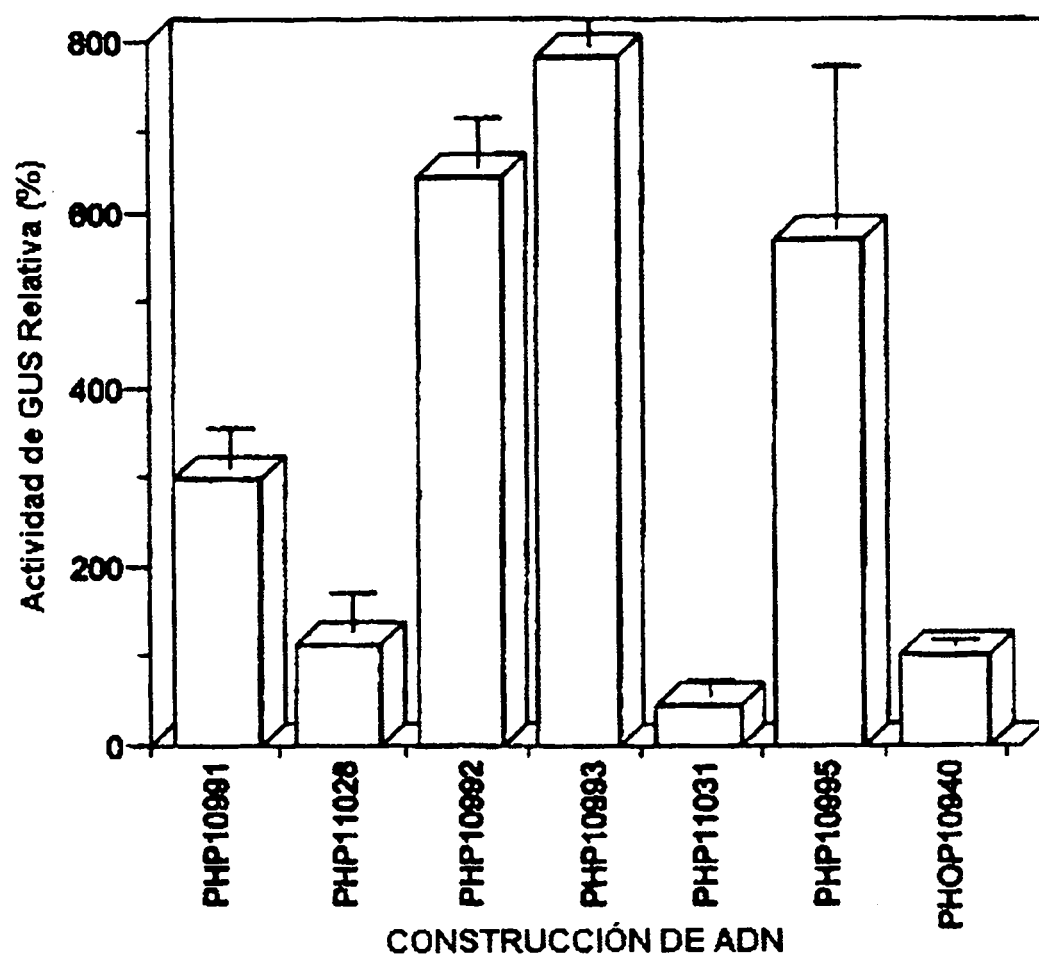


FIG. 13.

ES 2 292 234 T3

LISTA DE SECUENCIAS

(1) INFORMACIÓN GENERAL:

- 5 (i) APLICANTE: Bruce, Wesley Bowen, Benjamin A. Sims, Lynne Tagliani, Laura A. Lu, Guihua
- (ii) TÍTULO DE LA INVENCION: PROMOTORES SINTÉTICOS
- (iii) NÚMERO DE SECUENCIAS: 18
- 10 (iv) DIRECCIÓN DE CORRESPONDENCIA:
- (A) DESTINATARIO: W. Murray Spruill (Alston & Bird, LLP)
- (B) CALLE: 3605 Glenwood Ave. Suite 310
- 15 (C) CIUDAD: Raleigh
- (D) ESTADO: NC
- (E) PAÍS: USA
- (F) ZIP: 27622
- 20 (v) FORMA LEGIBLE DEL ORDENADOR:
- (A) TIPO MEDIO: Disco flexible
- (B) ORDENADOR: IBM PC compatible
- 25 (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
- (D) SOFTWARE: Patente en Edición #1.0, Versión #1.30
- (vi) DATOS DE APLICACIÓN ACTUAL:
- 30 (A) NÚMERO DE APLICACIÓN:
- (B) FECHA DE ARCHIVO:
- (C) CLASIFICACIÓN:
- 35 (viii) INFORMACIÓN DE AGENTE/ABOGADO:
- (A) NOMBRE: Spruill, W. Murray
- (B) NÚMERO DE REGISTRO: 32,943
- 40 (C) NÚMERO DE REFERENCIA/ETIQUETA: 5718-20
- (ix) INFORMACIÓN DE TELECOMUNICACIÓN:
- (A) TELÉFONO: 919 420 2202
- 45 (B) TELEFAX: 919 881 3175

(2) INFORMACIÓN PARA LA SEQ ID NO:1:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- 50 (A) LONGITUD: 72 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- 55 (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADN (genomic)
- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:1:
- 60 **GGATCCACTC GAGCGGCTAT AAATACGTAC CTACGCACGC TCGCTACCA TCCCGAGCAC 60**
- TGCAGTCTCG AC 72**

(2) INFORMACIÓN PARA LA SEQ ID NO:2:

- 65 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 96 pares de bases

ES 2 292 234 T3

- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

- (ii) TIPO DE MOLÉCULA: otro ácido nucleico
- (A) DESCRIPCIÓN: /desc = "Ácido nucleico sintético"

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:2:

GGATCCTATG CGTATGGTAT GACGTGTGTT CAAGATGATG ACTTCAAACC TACCTATGAC 60

GTATGGTATG ACGTGTGTCG ACTGATGACT TAGATC 96

- (2) INFORMACIÓN PARA LA SEQ ID NO:3:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 18 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

- (ii) TIPO DE MOLÉCULA: ADN (genómico)

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:3:

TATAWAWATY YTCATMAA 18

- (2) INFORMACIÓN PARA LA SEQ ID NO:4:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 40 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

- (ii) TIPO DE MOLÉCULA: ADN (genómico)

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:4:

CCTCTATATA AGCAAGTTCA TTTCATTTGG AGAGGAAACG 40

- (2) INFORMACIÓN PARA LA SEQ ID NO:5:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 5 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

- (ii) TIPO DE MOLÉCULA: ADN (genómico)

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:5:

TGAGC 5

- (2) INFORMACIÓN PARA LA SEQ ID NO:6:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:

ES 2 292 234 T3

5 (A) LONGITUD: 66 pares de bases
 (B) TIPO: ácido nucleico
 (C) DISPOSICIÓN EN HEBRAS: única
 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico
 (A) DESCRIPCIÓN: /desc = "Oligonucleótido sintético"

10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:6:

TCGACACTGC AGCTCTAGGG ATGGTAGCGC AGGGTGCGTA GGTACGTATT TATAGCCGCT 60

15 **CGAGTG 66**

(2) INFORMACIÓN PARA LA SEQ ID NO:7:

20 (i) CARACTERÍSTICAS DE LA SECUENCIA:
 (A) LONGITUD: 66 pares de bases
 (B) TIPO: ácido nucleico
 (C) DISPOSICIÓN EN HEBRAS: única
 (D) TOPOLOGÍA: lineal

25 (ii) TIPO DE MOLÉCULA: otro ácido nucleico
 (A) DESCRIPCIÓN: /desc = "Oligonucleótido sintético"

30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:7:

GATCCACTCG AGCGGCTATA AATACGTACC TACGCACCCT GCGCTACCAT CCCTAGAGCT 60

35 **GCAGTG 66**

(2) INFORMACIÓN PARA LA SEQ ID NO:8:

40 (i) CARACTERÍSTICAS DE LA SECUENCIA:
 (A) LONGITUD: 92 pares de bases
 (B) TIPO: ácido nucleico
 (C) DISPOSICIÓN EN HEBRAS: única
 (D) TOPOLOGÍA: lineal

45 (ii) TIPO DE MOLÉCULA: otro ácido nucleico
 (A) DESCRIPCIÓN: /desc = "Oligonucleótido sintético"

50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:8:

GATCCTATGA CGTATGGTAT GACGTGTGTT CAAGATGATG ACTTCAAACC TACCTATGAC 60

55 **GTATGGTATG ACGTGTGTCG ACTGATGACT TA 92**

(2) INFORMACIÓN PARA LA SEQ ID NO:9:

60 (i) CARACTERÍSTICAS DE LA SECUENCIA:
 (A) LONGITUD: 92 pares de bases
 (B) TIPO: ácido nucleico
 (C) DISPOSICIÓN EN HEBRAS: única
 (D) TOPOLOGÍA: lineal

65

ES 2 292 234 T3

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(A) DESCRIPCIÓN: /desc = "Oligonucleótido sintético"

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:9:

GATCTAAGTC ATCAGTCGAC ACACGTCATA CCATACGTCA TAGGTAGGTT TGAAGTCATC 60

ATCTTGAACA CACGTCATAC CATACGTCAT AG 92

(2) INFORMACIÓN PARA LA SEQ ID NO:10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 72 pares de bases

(B) TIPO: ácido nucleico

(C) DISPOSICIÓN EN HEBRAS: única

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(A) DESCRIPCIÓN: /desc = "Promotor central de ácido nucleico sintético con transversiones G/C"

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:10:

GGATCCACTC GAGCGGCTAT AAATASSTAS STASSSSASSS TSSSSTASSA TCCCAGGCAC 60

TGCAGTGTCTG AC 72

(2) INFORMACIÓN PARA LA SEQ ID NO: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 332 pares de bases

(B) TIPO: ácido nucleico

(C) DISPOSICIÓN EN HEBRAS: única

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 11:

CGTCAACATG GTGGAGCAGC ACACTCTCGT CTA CTCCAAG AATATCAAAG ATACAGTCTC 60

AGAAGACCAA AGGGCTATTG AGACTTTTCA ACAAAGGGTA ATATCGGGAA ACCTCCTCGG 120

ATTCCATTGC CCAGCTATCT GTCAC TTCAT CAAAAGGACA GTAGAAAAGG AAGGTGGCAC 180

CTACAAATGC CATCAT TGCG ATAAAGGAAA GGCTATCGTT CAAGATGCCT CTGCCGACAG 240

TGGTCCCCAA GATGGACCCC CACCCACGAG GAGCATCGTG GAAAAAGAAG ACGTTCCAAC 300

CACGTCTTCA AAGCAAGTGG ATTGATGTGA TG 332

(2) INFORMACIÓN PARA LA SEQ ID NO: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 499 pares de bases

(B) TIPO: ácido nucleico

(C) DISPOSICIÓN EN HEBRAS: única

(D) TOPOLOGÍA: lineal

ES 2 292 234 T3

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 12:

5	CGTCAACATG GTGGAGCACG ACACTCTCGT CTA	CTACTCCAAG AATATCAAAG ATACAGTCTC	60
	AGAAGACCAA AGGGCTATTG AGACTTTTCA ACAAAGGGTA	ATATCGGGAA ACCTCCTCGG	120
	ATTCCATTGC CCAGCTATCT GTCAC TTCAT CAAAAGGACA	G TAGAAAAGG AAGGTGGCAC	180
10	CTACAAATGC CATCATTGCG ATAAAGGAAA GGCTATCGTT	CAAGATGCCT CTGCCGACAG	240
	TGGTCCCAA GATGGACCCC CACCCACGAG GAGCATCGTG	GAAAAAGAAG ACGTTCCAAC	300
15	CACGTCTTCA AAGCAAGTGG ATTGATGTGA TGATCCTATG	CGTATGGTAT GACGTGTGTT	360
	CAAGATGATG ACTTCAAACC TACCTATGAC GTATGGTATG	ACGTGTGTCG ACTGATGACT	420
	TAGATCCACT CGAGCGGCTA TAAATACGTA CCTACGCACC	CTGCGCTACC ATCCCTAGAG	480
20	CTGCATGCTT ATTTTACA		499

(2) INFORMACIÓN PARA LA SEQ ID NO: 13:

25 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 813 pares de bases

(B) TIPO: ácido nucleico

30 (C) DISPOSICIÓN EN HEBRAS: única

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

35 (vi) FUENTE ORIGINAL:

(A) ORGANISMO: *Zea mays*

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:13:

40	TCTAGAGATA ATGAGCATTG CATGTCTAAG TTATAAAAAA	TTACCACATA TTTTTTTTGT	60
	CACACTTGTT TGAAGTGCAG TTTATCTATC TTTATACATA	TATTTAACT TTA	120
45	TACTCTACG AATAATATAA TCTATAGTAC TACAATAATA	TCAGTGT TTT AGAGAATCAT ATAAATGAAC	180
	AGTTAGACAT GGTCTAAAGG ACAATTGAGT ATTTTGACAA	CAGGACTCTA CAGTTTTATC	240
	TTTTTAGTGT GCATGTG TTCCTTTTTT TTGCAAATAG	CTTCACCTAT ATAATACTTC	300
50	ATCCATTTTA TTAGTACATC CATTIAGGGT TTAGGGTTAA	TGGTTTTTAT AGACTAATTT	360
	TTTTAGTACA TCTATTTTAT TCTATTTTAG CCTCTAAATT	AAGAAAATA AACTCTATT	420
55	TTAGTTTTTT TATTTAATAA TTTAGATATA AAATAGAATA	AAATAAAGTG ACTAAAAATT	480
	AAACAAATAC CCTTTAAGAA ATTAATAAAAA CTAAGGAAAC	ATTTTCTTG TTTCGAGTAG	540
	ATAATGCCAG CCTGTAAAC GCCGTCGACG AGTCTAACGG	ACACCAACCA GCGAACCAGC	600
60	AGCGTCGCGT CGGGCCAAGC GAAGCAGACG GCACGGCATC	TCTGTCGCTG CCTCTGGACC	660
	CCTCTCGAGA GTTCCGCTCC ACCGTTGGAC TTGCTCCGCT	GTCGGCATCC AGAAATTGCC	720
65	TGGCGGAGCG GCAGACGTGA GCCGGCACGG CAGGCGGCCT	CCTCCTCCTC TCACGGCACG	780
	GCAGCTACGG GGGATTCCTT TCCCACCGCT CCT		813

ES 2 292 234 T3

(2) INFORMACIÓN PARA LA SEQ ID NO: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1600 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 14:

```

15  GATCTGAGTC TAGAAATCCG TCAACATGGT GGAGCACGAC ACTCTCGTCT ACTCCAAGAA      60
    TATCAAAGAT ACAGTCTCAG AAGACCAAAG GGCTATTGAG ACTTTTCAAC AAAGGGTAAT      120
    ATCGGGAAAC CTCCTCGGAT TCCATTGCCC AGCTATCTGT CACTTCATCA AAAGGACAGT      180
20  AGAAAAGGAA GGTGGCACCT ACAAATGCCA TCATTGCGAT AAAGGAAAGG CTATCGTTCA      240
    AGATGCCTCT GCCGACAGTG GTCCCAAAGA TGGACCCCCA CCCACGAGGA GCATCGTGGA      300
25  AAAAGAAGAC GTTCCAACCA CGTCTTCAAA GCAAGTGGAT TGATGTGATG ATCCTATGCG      360
    TATGGTATGA CGTGTGTTCA AGATGATGAC TTCAAACCTA CCTATGACGT ATGGTATGAA      420
    CGTGTGTCGA CTGATGACTT AGATCCACTC GAGCGGCTAT AAATACGTAC CTACGCACCC      480
30  TCGGCTACCA TCCCTAGAGC TGCAGTTTAT TTTTACAACA ATTACCAACA ACAACAAACA      540
    ACAACAACA TTACAATTAC TATTTACAAT TACAGTCGAC CCGGGATCCA TGGGGTACTC      600
35  CAAAACCCTA GTAGCTGGCC TGTTGCAAT GCTGTTACTA GCTCCGGCCG TCTTGGCCAC      660
    CGACCCAGAC CCTCTCCAGG ACTTCTGTGT CGCCGACCTC GACGGCAAGG CGGTCTCGGT      720
    GAACGGGCAC ACGTGCAAGC CCATGTCGGA GGCCGGCGAC GACTTCCTCT TCTCGTCCAA      780
40  GTTGGCCAAG GCCGGCAACA CGTCCACCCC GAACGGCTCC GCCGTGACGG AGCTCGACGT      840
    GGCCGAGTGG CCCGGTACCA ACAAGCTGGG TGGTGTGATG AACC GCGTGG ATTTTGGTCC      900
    CGGAGGGACC AACCCACCAC ACATCCACCC GCGTGCCACC GAGATCGGCA TCGTGATGAA      960
45  AGGTGAGCTT CTCGTGGGAA TCCTTGGCAG CCTCGACTCC GGGAACAAGC TCTACTCGAG     1020
    GGTGTGCGC GCCGGAGAGA CGTTCCTCAT CCCACGGGGC CTCATGCACT TCCAGTTCAA     1080
50  CGTCGGTAAG ACCGAGGCCT CCATGGTCGT CTCCTTCAAC AGCCAGAACC CCGGCATTGT     1140
    CTTCTGTCCC CTCACGCTCT TCGGCTCCAA CCCGCCATC CCAACGCCGG TGCTACCAA     1200
    GGCACGCCG GTGGAGGCCA GGGTCGTGGA ACTTCTCAAG TCCAAGTTG CCGCTGGGTT     1260
55  TTAATTTCTA GGATCCTCTA GAGTCGAACC TAGACTTGTC CATCTTCTGG ATTGGCCAAC     1320
    TTAATTAATG TATGAAATAA AAGGATGCAC ACATAGTGAC ATGCTAATCA CTATAATGTG     1380
    GGCATCAAAG TTGTGTGTTA TGTGTAATTA CTAGTTATCT GAATAAAAGA GAAAGAGATC     1440
60  ATCCATATTT CTTATCCTAA ATGAATGTCA CGTGTCTTTA TAATTCTTTG ATGAACCAGA     1500
    TGCATTTTAT TAACCAAATC CATATACATA TAAATATTAA TCATATATAA TTAATATCAA     1560
65  TTGGGTTAGC AAAACAAATC TAGTCTAGGT GTGTTTTGCC     1600

```

(2) INFORMACIÓN PARA LA SEQ ID NO: 15:

ES 2 292 234 T3

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 994 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 15:

	TCTAGAGATA ATGAGCATTG CATGTCTAAG TTATAAAAAA TTACCACATA TTTTTTTTGT	60
15	CACACTTGTT TGAAGTGCAG TTTATCTATC TTTATACATA TATTTAAACT TTACTCTACG	120
	AATAATATAA TCTATAGTAC TACAATAATA TCAGTGTITT AGAGAATCAT ATAAATGAAC	180
20	AGTTAGACAT GGTCTAAAGG ACAATTGAGT ATTTTGACAA CAGGACTCTA CAGTTTTATC	240
	TTTTTAGTGT GCATGTGTTT TCCTTTTTTT TTGCAAATAG CTTCACCTAT ATAATACTTC	300
25	ATCCATTTTA TTAGTACATC CATTTAGGGT TTAGGGTTAA TGGTTTTTAT AGACTAATTT	360
	TTTTTAGTACA TCTATTTTAT TCTATTTTAT CCTCTAAAT AAGAAACTA AAACCTCTATT	420
	TTAGTTTTTT TATTTAATAA TTTAGATATA AAATAGAATA AAATAAAGTG ACTAAAAATT	480
30	AAACAAATAC CCTTTAAGAA ATTAAAAAAA CTAAGGAAAC ATTTTCTTG TTTTCGAGTAG	540
	ATAATGCCAG CCTGTAAAC GCCGTCGACG AGTCTAACGG ACACCAACCA GCGAACCAGC	600
35	AGCGTCGCGT CGGGCCAAGC GAAGCAGACG GCACGGCATC TCTGTCGCTG CCTCTGGACC	660
	CCTCTCGAGA GTTCCGCTCC ACCGTTGGAC TTGCTCCGCT GTCGGCATCC AGAAATTGCG	720
	TGGCGGAGCG GCAGACGTGA GCCGGCACGG CAGGCGGCCT CCTCCTCCTC TCACGGCAGC	780
40	GCAGCTACGG GGGATTCCTT TCCCACCGCT CCTACTAGAA CTAGTGGATC CTATGCGTAT	840
	GGTATGACGT GTGTTCAAGA TGATGACTTC AAACCTACCT ATGACGTATG GTATGACGTG	900
	TGTCGACTGA TGAAGTAGAT CCACTCGAGC GGCTATAAAT ACGTACCTAC GCACCCTGCG	960
45	CTACCATCCC TAGAGCTGCA TGCTTATTTT TACA	994

(2) INFORMACIÓN PARA LA SEQ ID NO:16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1807 pares de bases
- (B) TIPO: ácido nucleico
- (C) DISPOSICIÓN EN HEBRAS: única
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:16:

ES 2 292 234 T3

	TCTAGAGATA ATGAGCATTG CATGTCTAAG TTATAAAAAA TTACCACATA TTTTTTTTGT	60
	CACACTTGTT TGAAGTGCAG TTTATCTATC TTTATACATA TATTTAACT TTA CTCTACG	120
5	AATAATATAA TCTATAGTAC TACAATAATA TCAGTGTTTT AGAGAATCAT ATAAATGAAC	180
	AGTTAGACAT GGTCTAAAGG ACAATTGAGT ATTTTGACAA CAGGACTCTA CAGTTTTATC	240
10	TTTTAGTGT GCATGTGTTT TCCTTTTTTT TTGCAAATAG CTTACCTAT ATAATACTTC	300
	ATCCATTTTA TTAGTACATC CATTTAGGGT TTAGGGTTAA TGGTTTTTAT AGACTAATTT	360
	TTTTAGTACA TCTATTTTAT TCTATTTTAG CCTCTAAAT AAGAAACTA AACTCTATT	420
15	TTAGTTTTTT TATTTAATAA TTTAGATATA AAATAGAATA AAATAAAGTG ACTAAAAATT	480
	AAACAAATAC CCTTTAAGAA ATTAACAAAAA CTAAGGAAAC ATTTTCTTG TTTCGAGTAG	540
20	ATAATGCCAG CCTGTTAAAC GCCGTCGACG AGTCTAACGG ACACCAACCA GCGAACCAGC	600
	AGCGTCGCGT CGGGCCAAGC GAAGCAGACG GCACGGCATC TCTGTCGCTG CCTCTGGACC	660
	CCTCTCGAGA GTTCCGCTCC ACCGTTGGAC TTGCTCCGCT GTCGGCATCC AGAAATTGCG	720
25	TGGCGGAGCG GCAGACGTGA GCCGGCACGG CAGGCGGCCT CCTCCTCCTC TCACGGCAGC	780
	GCAGCTACGG GGGATTCCCT TCCCACCGCT CCTACTAGAG ATAATGAGCA TTGCATGTCT	840
30	AAGTTATAAA AAATTACCAC ATATTTTTTT TGTCACTT GTTGAAGTG CAGTTTATCT	900
	ATCTTTATAC ATATATTTAA ACTTTACTCT ACGAATAATA TAATCTATAG TACTACAATA	960
	ATATCAGTGT TTTAGAGAAT CATATAAATG AACAGTTAGA CATGGTCTAA AGGACAATTG	1020
35	AGTATTTTGA CAACAGGACT CTACAGTTTT ATCTTTTTAG TGTGCATGTG TTCTCCTTTT	1080
	TTTTTGCAA TAGCTTCACC TATATAATAC TTCATCCATT TTATTAGTAC ATCCATTTAG	1140
	GGTTTAGGGT TAATGGTTTT TATAGACTAA TTTTTTAGT ACATCTATTT TATTCTATTT	1200
40	TAGCCTCTAA ATTAAGAAAA CTAAACTCT ATTTTAGTTT TTTATTTAA TAATTTAGAT	1260
	ATAAAATAGA ATAAATATA GTGACTAAAA ATTAAACAAA TACCCTTTAA GAAATTAAAA	1320
	AAACTAAGGA AACATTTTTC TTGTTTCGAG TAGATAATGC CAGCCTGTTA AACGCCGTCG	1380
45	ACGAGTCTAA CGGACACCAA CCAGCGAACC AGCAGCGTCG CGTCGGGCCA AGCGAAGCAG	1440
	ACGGCACGGC ATCTCTGTCG CTGCCTCTGG ACCCCTCTCG AGAGTTCCGC TCCACCGTTG	1500
50	GACTTGCTCC GCTGTCGGCA TCCAGAAATT GCGTGGCGGA GCGGCAGACG TGAGCCGGCA	1560
	CGGCAGGCGG CCTCCTCCTC CTCTCACGGC ACGGCAGCTA CGGGGGATTCTTTCCCACC	1620
	GCTCCTACTA GAACTAGTGG ATCCTATGCG TATGGTATGA CGTGTGTTCA AGATGATGAC	1680
55	TTCAAACCTA CCTATGACGT ATGGTATGAC GTGTGTCGAC TGATGACTTA GATCCACTCG	1740
	AGCGGCTATA AATACGTACC TACGCACCCT GCGCTACCAT CCCTAGAGCT GCATGCTTAT	1800
60	TTTTACA	1807

(2) INFORMACIÓN PARA LA SEQ ID NO:17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 65
- (A) LONGITUD: 2620 pares de bases
 - (B) TIPO: ácido nucleico

ES 2 292 234 T3

(C) DISPOSICIÓN EN HEBRAS: única

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:17:

```

TCTAGAGATA ATGAGCATTG CATGTCTAAG TTATAAAAAA TTACCACATA TTTTTTTTGT      60
CACACTTGTT TGAAGTGCAG TTTATCTATC TTTATACATA TATTTAAACT TTA CTCTACG      120
AATAATATAA TCTATAGTAC TACAATAATA TCAGTGTTTT AGAGAATCAT ATAAATGAAC      180
AGTTAGACAT GGTCTAAAGG ACAATTGAGT ATTTTGACAA CAGGACTCTA CAGTTTTATC      240
TTTTTAGTGT GCATGTGTTT TCCTTTTTTT TTGCAAATAG CTTACCTAT ATAATACTTC      300
ATCCATTTTA TTAGTACATC CATTTAGGGT TTAGGGTTAA TGGTTTTTAT AGACTAATTT      360
TTTTTAGTACA TCTATTTTAT TCTATTTTAT CCTCTAAATT AAGAAACTA AAACCTCTATT      420
TTAGTTTTTT TATTTAATAA TTTAGATATA AAATAGAATA AAATAAAGTG ACTAAAAATT      480
AAACAAATAC CCTTTAAGAA ATTAAAAAAA CTAAGGAAAC ATTTTCTTG TTTGAGTAG      540
ATAATGCCAG CCTGTAAAC GCCGTCGACG AGTCTAACGG ACACCAACCA GCGAACCAGC      600
AGCGTCGCGT CGGGCCAAGC GAAGCAGACG GCACGGCATC TCTGTCGCTG CCTCTGGACC      660
CCTCTCGAGA GTTCGCTCC ACCGTTGGAC TTGCTCCGCT GTCGGCATCC AGAAATTGCG      720
TGGCGGAGCG GCAGACGTGA GCCGGCACGG CAGGCGGCCT CCTCCTCCTC TCACGGCACG      780
GCAGCTACGG GGGATTCTTT TCCACCGCT CCTACTAGAG ATAATGAGCA TTGCATGTCT      840
AAGTTATAAA AAATTACCAC ATATTTTTTT TGTCACACTT GTTTGAAGTG CAGTTTATCT      900
ATCTTTATAC ATATATTTAA ACTTTACTCT ACGAATAATA TAATCTATAG TACTACAATA      960
ATATCAGTGT TTAGAGAAT CATATAATG AACAGTTAGA CATGGTCTAA AGGACAATTG      1020
AGTATTTTGA CAACAGGACT CTACAGTTT ATCTTTTTAG TGTGCATGTG TTCTCCTTTT      1080
TTTTTGCAA TAGCTTCACC TATATAATAC TTCATCCATT TTATTAGTAC ATCCATTTAG      1140
GGTTTAGGGT TAATGGTTTT TATAGACTAA TTTTTTTAGT ACATCTATTT TATTCTATTT      1200
TAGCCTCTAA ATTAAGAAAA CTAAACTCT ATTTTAGTTT TTTTATTTAA TAATTTAGAT      1260
ATAAAATAGA ATAAATAAAA GTGACTAAAA ATTAACAAA TACCCTTAA GAAATTAAAA      1320
AAACTAAGGA AACATTTTTC TTGTTTCGAG TAGATAATGC CAGCCTGTTA AACGCCGTCG      1380
ACGAGTCTAA CGGACACCAA CCAGCGAACC AGCAGCGTCG CGTCGGGCCA AGCGAAGCAG      1440
ACGGCACGGC ATCTCTGTCG CTGCCTCTGG ACCCTCTCG AGAGTCCGC TCCACCGTTG      1500
GACTTGCTCC GCTGTCGGCA TCCAGAAATT GCGTGGCGGA GCGGCAGACG TGAGCCGGCA      1560
CGGCAGGCGG CCTCCTCCTC CTCTACGGC ACGGCAGCTA CGGGGGATTC CTTCCCACC      1620
GCTCCTACTA GAGATAATGA GCATTGCATG TCTAAGTTAT AAAAAATTAC CACATATTTT      1680
TTTTGTCACA CTTGTTTGAA GTGCAGTTTA TCTATCTTTA TACATATATT TAACTTTTAC      1740
TCTACGAATA ATATAATCTA TAGTACTACA ATAATATCAG TGTTTTAGAG AATCATATAA      1800
ATGAACAGTT AGACATGGTC TAAAGGACAA TTGAGTATTT TGACAACAGG ACTCTACAGT      1860

```

ES 2 292 234 T3

	TTTATCTTTT TAGTGTGCAT GTGTTCTCCT TTTTTTTTGC AAATAGCTTC ACCTATATAA	1920
	TACTTCATCC ATTTTATTAG TACATCCATT TAGGGTTTAG GGTTAATGGT TTTTATAGAC	1980
5	TAATTTTTTT AGTACATCTA TTTTATTCTA TTTTAGCCTC TAAATTAAGA AAACATAAAC	2040
	TCTATTTTAG TTTTTTTATT TAATAATTTA GATATAAAAT AGAATAAAAT AAAGTGACTA	2100
10	AAAATTAAC AAATACCCTT TAAGAAATTA AAAAACTAA GGAAACATTT TTCTTGTTTC	2160
	GAGTAGATAA TGCCAGCCTG TTAACGCCG TCGACGAGTC TAACGGACAC CAACAGCGA	2220
	ACCAGCAGCG TCGCGTCGGG CCAAGCGAAG CAGACGGCAC GGCATCTCTG TCGCTGCCTC	2280
15	TGGACCCCTC TCGAGAGTTC CGCTCCACCG TTGGACTTGC TCCGCTGTCG GCATCCAGAA	2340
	ATTGCGTGGC GGAGCGGCAG ACGTGAGCCG GCACGGCAGG CGGCCTCCTC CTCCTCTCAC	2400
	GGCACGGCAG CTACGGGGGA TTCCTTTCCC ACCGCTCCTA CTAGAACTAG TGGATCCTAT	2460
20	GCGTATGGTA TGACGTGTGT TCAAGATGAT GACTTCAAAC CTACCTATGA CGTATGGTAT	2520
	GACGTGTGTC GACTGATGAC TTAGATCCAC TCGAGCGGCT ATAAATACGT ACCTACGCAC	2580
25	CCTGCGCTAC CATCCCTAGA GCTGCATGCT TATTTTTACA	2620

(2) INFORMACIÓN PARA LA SEQ ID NO: 18:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 3433 pares de bases
 - (B) TIPO: ácido nucleico
 - (C) DISPOSICIÓN EN HEBRAS: única
 - (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: otro ácido nucleico

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 18:

	TCTAGAGATA ATGAGCATTG CATGTCTAAG TTATAAAAAA TTACCACATA TTTTTTTTGT	60
	CACACTTGTT TGAAGTGCAG TTTATCTATC TTTATACATA TATTTAAACT TTACTCTACG	120
45	AATAATATAA TCTATAGTAC TACAATAATA TCAGTGTTTT AGAGAATCAT ATAAATGAAC	180
	AGTTAGACAT GGTCTAAAGG ACAATTGAGT ATTTTGACAA CAGGACTCTA CAGTTTTATC	240
	TTTTTAGTGT GCATGTGTTT TCCTTTTTTT TTGCAAATAG CTTACCTAT ATAATACTTC	300
50	ATCCATTTTA TTAGTACATC CATTTAGGGT TTAGGGTTAA TGGTTTTTAT AGACTAATTT	360
	TTTTAGTACA TCTATTTTAT TCTATTTTAG CCTCTAAATT AAGAAACTA AAACCTCTAT	420
55	TTAGTTTTTT TATTTAATAA TTTAGATATA AAATAGAATA AAATAAAGTG ACTAAAAATT	480
	AAACAAATAC CCTTTAAGAA ATTAACAAAAA CTAAGGAAAC ATTTTCTTG TTTGAGTAG	540
	ATAATGCCAG CCTGTAAAC GCCGTCGACG AGTCTAACGG ACACCAACCA GCGAACCAGC	600

ES 2 292 234 T3

	AGCGTCGCGT CGGGCCAAGC GAAGCAGACG GCACGGCATC TCTGTCGCTG CCTCTGGACC	660
	CCTCTCGAGA GTTCCGCTCC ACCGTTGGAC TTGCTCCGCT GTCGGCATCC AGAAATTGCG	720
5	TGGCGGAGCG GCAGACGTGA GCCGGCACGG CAGGCGGCCT CCTCCTCCTC TCACGGCACG	780
	GCAGCTACGG GGGATTCTTT TCCCACCGCT CCTACTAGAG ATAATGAGCA TTGCATGTCT	840
10	AAGTTATAAA AAATTACCAC ATATTTTTTT TGTCACTT GTTGAAGTG CAGTTTATCT	900
	ATCTTTATAC ATATATTTAA ACTTTACTCT ACGAATAATA TAATCTATAG TACTACAATA	960
	ATATCAGTGT TTTAGAGAAT CATATAAATG AACAGTTAGA CATGGTCTAA AGGACAATTG	1020
15	AGTATTTTGA CAACAGGACT CTACAGTTTT ATCTTTTTAG TGTGCATGTG TTCTCCTTTT	1080
	TTTTTGCAA TAGCTTCACC TATATAATAC TTCATCCATT TTATTAGTAC ATCCATTTAG	1140
	GGTTTAGGGT TAATGGTTTT TATAGACTAA TTTTTTTAGT ACATCTATTT TATTCTATTT	1200
20	TAGCCTCTAA ATTAAGAAAA CTAAACTCT ATTTTAGTTT TTTTATTTAA TAATTTAGAT	1260
	ATAAAATAGA ATAAATAAA GTGACTAAAA ATTAAACAAA TACCCTTTAA GAAATTAAAA	1320
	AAACTAAGGA AACATTTTTT TTGTTTCGAG TAGATAATGC CAGCCTGTTA AACGCCGTCG	1380
25	ACGAGTCTAA CGGACACCAA CCAGCGAACC AGCAGCGTCG CGTCGGGCCA AGCGAAGCAG	1440
	ACGGCACGGC ATCTCTGTCG CTGCCTCTGG ACCCCTCTCG AGAGTTCCGC TCCACCGTTG	1500
30	GA CT TGCTCC GCTGTCGGCA TCCAGAAAT GCGTGGCGGA GCGGCAGACG TGAGCCGGCA	1560
	CGGCAGGCGG CCTCCTCCTC CTCTCACGGC ACGGCAGCTA CGGGGGATT CTTTCCCACC	1620
	GCTCCTACTA GAGATAATGA GCATTGCATG TCTAAGTTAT AAAAAATTAC CACATATTTT	1680
35	TTTTGTCACA CTTGTTTGAA GTGCAGTTTA TCTATCTTTA TACATATATT TAACTTTTAC	1740
	TCTACGAATA ATATAATCTA TAGTACTACA ATAATATCAG TGTTTTAGAG AATCATATAA	1800
	ATGAACAGTT AGACATGGTC TAAAGGACAA TTGAGTATTT TGACAACAGG ACTCTACAGT	1860
40	TTTATCTTTT TAGTGTGCAT GTGTTCTCCT TTTTTTTTGC AAATAGCTTC ACCTATATAA	1920
	TACTTCATCC ATTTTATTAG TACATCCATT TAGGGTTTAG GGTTAATGGT TTTTATAGAC	1980
45	TAATTTTTTT AGTACATCTA TTTTATTCTA TTTTAGCCTC TAAATTAAGA AAATAAAAC	2040
	TCTATTTTAG TTTTTTTATT TAATAATTTA GATATAAAAT AGAATAAAAT AAAGTACTA	2100
	AAAATTAAAC AAATACCCTT TAAGAAATTA AAAAACTAA GGAAACATTT TTCTTGTTTC	2160
50	GAGTAGATAA TGCCAGCCTG TTAAACGCCG TCGACGAGTC TAACGGACAC CAACCAGCGA	2220
	ACCAGCAGCG TCGCGTCGGG CCAAGCGAAG CAGACGGCAC GGCATCTCTG TCGCTGCCTC	2280
	TGGACCCCTC TCGAGAGTTC CGCTCCACCG TTGGACTTGC TCCGCTGTCG GCATCCAGAA	2340
55	ATTGCGTGGC GGAGCGGCAG ACGTGAGCCG GCACGGCAGG CGGCCTCCTC CTCCTCTCAC	2400
	GGCACGGCAG CTACGGGGGA TTCCTTTCCC ACCGCTCCTA CTAGAGATAA TGAGCATTGC	2460
60		
65		

ES 2 292 234 T3

	ATGTCTAAGT TATAAAAAAT TACCACATAT TTTTTTTGTC ACACTTGTTT GAAGTGCAGT	2520
	TTATCTATCT TTATACATAT ATTTAAACTT TACTCTACGA ATAATATAAT CTATAGTACT	2580
5	ACAATAATAT CAGTGTTTTA GAGAATCATA TAAATGAACA GTTAGACATG GTCTAAAGGA	2640
	CAATTGAGTA TTTTGACAAC AGGACTCTAC AGTTTTATCT TTTTAGTGTG CATGTGTTCT	2700
	CCTTTTTTTT TGCAAATAGC TTCACCTATA TAATACTTCA TCCATTTTAT TAGTACATCC	2760
10	ATTTAGGGTT TAGGGTTAAT GGTTTTTATA GACTAATTTT TTTAGTACAT CTATTTTATT	2820
	CTATTTTAGC CTCTAAATTA AGAAAACATA AACTCTATTT TAGTTTTTTT ATTTAATAAT	2880
15	TTAGATATAA AATAGAATAA AATAAAGTGA CTAATAATTA AACAAATACC CTTTAAGAAA	2940
	TTAAAAAAC TAAGGAAACA TTTTCTTGT TTCGAGTAGA TAATGCCAGC CTGTAAACG	3000
	CCGTCGACGA GTCTAACGGA CACCAACCAG CGAACCAGCA GCGTCGCGTC GGGCCAAGCG	3060
20	AAGCAGACGG CACGGCATCT CTGTCGCTGC CTCTGGACCC CTCTCGAGAG TTCCGCTCCA	3120
	CCGTTGGAAT TGCTCCGCTG TCGGCATCCA GAAATTGCGT GGCGGAGCGG CAGACGTGAG	3180
	CCGGCACGGC AGGCGGCCTC CTCCTCCTCT CACGGCACGG CAGCTACGGG GGATTCCTTT	3240
25	CCCACCGCTC CTACTAGAAC TAGTGGATCC TATGCGTATG GTATGACGTG TGTCAAGAT	3300
	GATGACTTCA AACCTACCTA TGACGTATGG TATGACGTGT GTCGACTGAT GACTTAGATC	3360
30	CACTCGAGCG GCTATAAATA CGTACCTACG CACCCTGCGC TACCATCCCT AGAGCTGCAT	3420
	GCTTATTTTT ACA	3433

35 LISTADO DE SECUENCIAS

	<110> Wesley, Bruce
40	<120> Promotores sintéticos
	<130> 035718/176191 5718-20A-1-pc
	<140> PCT/US99/03863
45	<141> 1999-02-23
	<150> 08/661,601
	<151> 1996-06-11
	<160> 18
50	<170> Patente En Ver. 2.0
	<210> 1
	<211> 72
	<212> ADN
55	<213> Secuencia Artificial
	<220>
	<223> Descripción de la Secuencia Artificial: secuencia del promotor central sintético
60	<400> 1

	ggatccactc gagcggctat aaatacgtac ctacgcacgc tgcgctacca tcccagcac	60
65	tgcagtgtcg ac	72

<210> 2

ES 2 292 234 T3

<212> ADN
 <213> Secuencia Artificial
 <220>
 5 <223> Descripción de la Secuencia Artificial: OCS elemento
 <400> 5
 10 tgacg 5
 <210> 6
 <211> 66
 15 <212> ADN
 <213> Secuencia Artificial
 <220>
 20 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético
 <400> 6
 25 **tcgacactgc agctctaggg atggtagcgc aggggtgcgta ggtacgtatt tatagccgct 60**
 cgagtg **66**
 <210> 7
 30 <211> 66
 <212> ADN
 <213> Secuencia Artificial
 <220>
 35 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético
 <400> 7
 40 **gatccactcg agcggctata aatacgtacc tacgcaccct gcgctaccat ccctagagct 60**
 gcagtg **66**
 45 <210> 8
 <211> 92
 <212> ADN
 <213> Secuencia Artificial
 50 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético
 <400> 8
 55 **gatacctatga cgtatgggtat gacgtgtggt caagatgatg acttcaaacc tacctatgac 60**
 gtatgggtatg acgtgtgtcg actgatgact ta **92**
 60 <210> 9
 <211> 92
 <212> ADN
 65 <213> Secuencia Artificial
 <220>
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

ES 2 292 234 T3

<400> 9

5 **gatctaagtc atcagtcgac acacgtcata ccatacgtca taggtagggt tgaagtcac 60**
 atcttgaaca cacgtcatat catacgtcat ag 92

<210> 10

<211> 72

10 <212> ADN

<213> Secuencia Artificial

<220>

15 <223> Descripción de la Secuencia Artificial: promotor central del ácido nucleico sintético con transversiones G/C

<220>

<223> Los nucleótidos en las posiciones 26, 27, 30, 31, 34, 35, 36, 38, 39, 40, 42, 43, 44, 45, 48, y 49, pueden ser g o c.

20

<400> 10

25 **ggatccactc gagcggctat aaatasstas stasssassy tsssstassa tcccgagcac 60**
 tgcagtgtcg ac 72

<210> 11

<211> 332

30 <212> ADN

<213> Virus del mosaico de la coliflor

<400> 11

35

cgtcaacatg gtggagcacg acactctcgt ctactccaag aatatcaaag atacagtctc 60
 agaagaccaa agggctattg agacttttca acaaagggtg atacgggaa acctcctcgg 120
 attccattgc ccagctatct gtcacttcat caaaaggaca gtagaaaagg aagggtggcac 180
 ctacaaatgc catcattgcg ataaaggaaa ggctatcgtt caagatgcct ctgccgacag 240
 tggtcccaaa gatggacccc caccacgag gagcatcgtg gaaaaagaag acgttccaac 300
 cacgtcttca aagcaagtgg attgatgtga tg 332

45

<210> 12

<211> 499

<212> ADN

50 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: 35S UAR del virus del mosaico de la coliflor ligado de manera operable a la secuencia del promotor central

55

60

65

ES 2 292 234 T3

<400> 12

```

5      cgtaacatg gtggagcacg acactctcgt ctactccaag aatatcaaag atacagtctc 60
      agaagaccac agggctattg agacttttca acaaagggtg atatcgggaa acctcctcgg 120
      attccattgc ccagctatct gtcacttcat caaaaggaca gtagaaaagg aagggtggcac 180
      ctacaaatgc catcattgcg ataaaggaaa ggctatcgtt caagatgcct ctgccgacag 240
10     tgggtcccaa gatggacccc caccacgag gagcatcgtg gaaaaagaag acgttccaac 300
      cacgtcttca aagcaagtgg attgatgtga tgatcctatg cgtatggtat gacgtgtgtt 360
      caagatgatg acttcaaacc tacctatgac gtatgggatg acgtgtgtcg actgatgact 420
      tagatccact cgagcggcta taaatacgta cctacgcacc ctgcgctacc atccctagag 480
15     ctgcatgctt atttttaca                                     499

```

<210> 13

<211> 813

20 <212> ADN

<213> *Zea mays*

<400> 13

```

25     tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt 60
      cacacttggt tgaagtgcag tttatctatc tttatacata tatttaaaact ttactctacg 120
      aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataaatgaac 180
30     agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc 240
      tttttagtgt gcatgtgttc tccttttttt ttgcaaatag cttcacctat ataatacttc 300
      atccatttta ttagtacatc catttagggg ttaggggttaa tgggttttat agactaattt 360
      ttttagtaca tctattttat tctatttttag cctctaaatt aagaaaacta aaactctatt 420
35     ttagtttttt tatttaataa tttagatata aaatagaata aaataaagtg actaaaaatt 480
      aaacaaatac cctttaagaa attaaaaaaa ctaaggaaac atttttcttg tttcgagtag 540
      ataatgccag cctgttaaac gccgtcgacg agtctaacgg acaccaacca gcgaaccagc 600
40     agcgtcgcgt cgggccaagc gaagcagacg gcacggcatc tctgtcgtg cctctggacc 660
      cctctcgaga gttccgctcc accgttgac ttgctcgtg gtcggcatcc agaaattgcg 720
      tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcacggcacg 780
      gcagctacgg gggatttcct tcccaccgct cct                                     813

```

<210> 14

<211> 1600

50 <212> ADN

<213> Secuencia Artificial

<220>

55 <223> Descripción de la Secuencia Artificial: 35S UAR del virus del mosaico de la coliflor ligado de manera operable al promotor central ligado de manera operable a la secuencia oxalato oxidasa codificante ligado de manera operable al terminador PinII.

<400> 14

```

60     gatctgagtc tagaaatccg tcaacatggt ggagcacgac actctcgtct actccaagaa 60
      tatcaaagat acagtctcag aagaccaaag ggctattgag acttttcaac aaagggtaat 120

```

65

ES 2 292 234 T3

```

atcgggaaac ctctcggat tccattgccc agctatctgt cacttcatca aaaggacagt 180
agaaaaggaa ggtggcacct acaaatgcca tcattgcatg aaaggaaagg ctatcggtca 240
agatgcctct gccgacagtg gtcccaaaga tggacccccca cccacgagga gcatcggtga 300
5   aaaagaagac gttccaacca cgtcttcaaa gcaagtggat tgatgtgatg atcctatgcg 360
tatggtatga cgtgtgttca agatgatgac ttcaaacctc cctatgacgt atggtatgaa 420
cgtgtgtcga ctgatgactt agatccactc gagcggctat aaatacgtag ctacgcaccc 480
tgcgctacca tccctagagc tgcagcttat ttttacaaca attaccaaca acaacaaca 540
10  acaacaaca ttacaattac tatttacaat tacagtcgac ccgggatcca tggggtagtc 600
caaaacccta gtagctggcc tgttcgcaat gctgttacta gctccggcgg tcttgccac 660
cgacccagac cctctccagg acttctgtgt cgccgacctc gacggcaagg cggctcgggt 720
gaacgggcac acgtgcaagc ccatgtcgga ggccggcgac gacttcctct tctcgtccaa 780
15  gttggccaag gccggcaaca cgtccacccc gaacgggtcc gccgtgacgg agtcgacgt 840
ggccgagtg ccgggtacca acaagctggg tgggtgcatg aaccgcgtgg attttgtcc 900
cggaggggacc aaccaccac acatccaccc gcgtgccacc gagatcggca tctgtatgaa 960
aggtgagctt ctcgtgggaa tcttggcag cctcgactcc gggacaaga tctactcgag 1020
20  ggtggtgccc gccggagaga cgttcctcat cccacggggc ctcatgactc tccagttcaa 1080
cgtcggtaa accgagacct ccatggtcgt ctccttcaac agccagaacc ccggcattgt 1140
cttcgtgccc ctcacgctct tcggctccaa cccgccctc ccaacgccgg tgcaccaa 1200
ggcactccgg gtggaggcca ggtcgtgga acttctcaag tccaagttg ccgctgggtt 1260
25  ttaatttcta ggatcctcta gagtcgaacc tagacttgct catcttctgg attggccaac 1320
ttaattaatg tatgaaataa aaggatgcac acatagtgac atgctaata ctataatgtg 1380
ggcatcaaag ttgtgtgta tgtgtaatta ctagttatct gaataaaga gaaagagatc 1440
atccatattt cttatcctaa atgaatgtca cgtgtcttta taattctttg atgaaccaga 1500
30  tgcatttcat taaccaaata catatacata taaatattaa tcatatataa ttaatatcaa 1560
ttgggttagc aaaacaaatc tagtctaggt gtgttttgcc 1600

```

<210> 15

<211> 994

<212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: promotor Rsyn7 que contiene 1 copia del elemento del extremo 5' ubi-1 de *zea mays* (UbiU)

<400> 15

```

tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt 60
cacacttggt tgaagtgcag tttatctatc tttatacata tatttaaact ttactctacg 120
aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataatgaac 180
agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc 240
50  tttttagtgt gcatgtgttc tccctttttt ttgcaaatag cttcacctat ataatacttc 300
atccatttta ttagtacatc catttagggg ttagggttaa tgggttttat agactaattt 360
ttttagtaca tctattttat tctatttttag cctctaaatt aagaaaacta aaactctatt 420
ttagtttttt tatttaataa tttagatata aaatagaata aaataaagtg actaaaaatt 480
55  aaacaaatac ctttaagaa attaaaaaaa ctaaggaaac atttttcttg tttcgagtag 540
ataatgccag cctgttaaac gccgtcgacg agtctaacgg acaccaacca gcgaaccagc 600
agcgtcgggt cgggccaaag gaagcagacg gcacggcatc tctgtcgtg cctctggacc 660
cctctcgaga gttccgctcc accgttgga cttgtccgct gtcggcatcc agaaattgag 720
60  tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcacggcacg 780
gcagctacgg gggattcctt tcccaccgct cctactagaa ctagtggatc ctatgcgtat 840
ggtatgacgt gtgttcaaga tgatgacttc aaacctacct atgacgtatg gtatgacgtg 900
tgtcgactga tgacttagat ccactcgagc ggctataaat acgtacctac gcaccctgag 960
65  ctaccatccc tagagctgca tgcttatttt taca 994

```

ES 2 292 234 T3

<210> 16

<211> 1807

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: secuencia del promotor UCP2 que contiene 2 copias del ubil UAR de *zea mays* ligado de manera operable con el promotor central

10

<400> 16

15 tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt 60
cacacttggt tgaagtgcag tttatctatc tttatacata tatttaaact ttactctacg 120
aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataaatgaac 180
agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc 240
20 ttttttagtgt gcatgtgttc tccttttttt ttgcaaatag cttcacctat ataatacttc 300
atccatttta ttagtacatc catttagggg ttaggggttaa tggtttttat agactaattt 360
tttttagtaca tctattttat tctatttttag cctctaaatt aagaaaacta aaactctatt 420
ttagtttttt tatttaataa tttagatata aaatagaata aaataaagtg actaaaaatt 480
aaacaaatac cctttaagaa attaaaaaaa ctaaggaaac atttttcttg tttcgagtag 540
25 ataatgccag cctgttaaac gccgtcgacg agtctaacgg acaccaacca gcgaaccagc 600
agcgtcgcgt cgggccaagc gaagcagacg gcacggcatc tctgtcgcgt cctctggacc 660
cctctcgaga gttccgctcc accgttgga cttgtccgct gtcggcatcc agaaattgag 720
30 tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcacggcacg 780
gcagctacgg gggattcctt tcccaccgct cctactagag ataagagca ttgcatgtct 840
aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagtg cagtttatct 900
atctttatac atatatatta actttactct acgaataata taatctatag tactacaata 960
35 atatcagtggt tttagagaat catataaatg aacagttaga catggtctaa aggacaattg 1020
agtattttga caacaggact ctacagtttt atcttttttag tgtgcatgtg ttctcctttt 1080
tttttgcaaa tagcttcacc tatataatac ttcattccatt ttatttagtac atccatttag 1140
ggtttagggt taatgggttt tatagactaa ttttttttagt acatctattt tattctattt 1200
40 tagcctctaa attaagaaaa ctaaaactct attttaggtt ttttatttaa taatttagat 1260
ataaaataga ataaaataaa gtgactaaaa attaaacaaa taccctttaa gaaattaaaa 1320
aaactaagga aacatttttc ttgtttcgag tagataatgc cagcctgtta aacgccgtcg 1380
acgagtctaa cggacaacca ccagcgaacc agcagcgtcg cgtcgggcca agcgaagcag 1440
45 acggcacggc atctctgtcg ctgcctctgg acccctctcg agagtccgc tccaccgttg 1500
gacttgctcc gctgtcggca tccagaaatt gcgtggcgga gcggcagacg tgagccggca 1560
cggcagggcg cctcctctc ctctcacggc acggcagcta cgggggattc ctttccacc 1620
gctcctaact gaactagtgg atcctatgcg tatggtatga cgtgtgttca agatgatgac 1680
50 ttcaaacta cctatgacgt atggtatgac gtgtgtcgac tgatgactta gatccactcg 1740
agcggctata aatacgtacc tacgcacct gcgctaccat ccctagagct gcatgcttat 1800

55 ttttaca 1807

<210> 17

<211> 2620

60 <212> ADN

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: promotor Rsyn7 que contiene 3 copias del elemento del extremo 5' UbiU de *zea mays*

65

ES 2 292 234 T3

<400> 17

```

5      tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt 60
      cacacttggt tgaagtgcag tttatctatc tttatacata tattttaaact ttactctacg 120
      aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataatgaac 180
      agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc 240
      tttttagtggt gcatgtgttc tccttttttt ttgcaaatag cttcacctat ataatacttc 300
10     atccatttta ttagtacatc catttagggg ttaggggttaa tgggtttttat agactaattt 360
      ttttagtaca tctatttttat tctatttttag cctctaaatt aagaaaacta aaactctatt 420
      ttagtttttt tatttaataa tttagatata aaatagaata aaataaagtg actaaaaatt 480
      aaacaaatac cctttaagaa attaaaaaaa ctaaggaaac atttttcttg tttcgagtag 540
15     ataatgccag cctgttaaac gccgtcgagc agtctaacgg acaccaacca gcgaaccagc 600
      agcgtcgcgt cgggccaaagc gaagcagacg gcacggcacc tctgtcgcgt cctctggacc 660
      cctctcgaga gttccgctcc accgttggac ttgctccgct gtcggcatcc agaaattgag 720
      tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcacggcacg 780
20     gcagctacgg gggattcctt tcccaccgct cctactagag ataatgagca ttgcatgtct 840
      aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagtg cagtttatct 900
      atctttatac atatatattaa actttactct acgaataata taatctatag tactacaata 960
      atatcagtggt tttagagaat catataaatg aacagttaga catggtctaa aggacaattg 1020
25     agtattttga caacaggact ctacagtttt atcttttttag tgtgcatgtg ttctcctttt 1080
      tttttgcaaa tagcttcacc tatataatac ttcatccatt ttattagtac atccatttag 1140
      ggttttaggg taatggtttt tatagactaa ttttttagt acatctattt tattctattt 1200
      tagcctctaa attaagaaaa ctaaaactct attttagttt ttttatttaa taatttagat 1260
30     ataaaataga ataaaataaa gtgactaaaa attaaacaaa taccctttaa gaaattaaaa 1320
      aaactaagga aacatttttt ttgttcgag tagataatgc cagcctgtta aacgcgcgtc 1380
      acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg cgtcgggcca agcgaagcag 1440
      acggcacggc atctctgtcg ctgcctctgg accctctcg agagtccgc tccaccgttg 1500
35     gacttgctcc gctgtcggca tccagaaatt gcgtggcgga gcggcagacg tgagccggca 1560
      cggcaggcgg cctcctctc ctctcacggc acggcagcta cgggggattc ctttcccacc 1620
      gctcctacta gagataatga gcattgcatg tctaagttat aaaaaattac cacatatttt 1680
      ttttgtcaca cttgtttgaa gtgcagttta tctatcttta tacatatatt taaactttac 1740
40     tctacgaata atataatcta tagtactaca ataatatcag tgttttagag aatcatataa 1800
      atgaacagtt agacatggc taaaggacaa ttgagtattt tgacaacagg actctacagt 1860
      ttatcttttt tagtgtgcat gtgttctcct tttttttgac aaatagcttc acctatataa 1920
      tacttcatcc attttattag tacatccatt tagggtttag ggtaaatggt ttttatagac 1980
45     taattttttt agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaaac 2040
      tctatttttag tttttttatt taataattta gatataaaat agaataaaat aaagtgacta 2100

      aaaattaaac aaataccctt taagaaatta aaaaaactaa ggaaacattt ttcttggttc 2160
50     gagtagataa tgccagcctg ttaaagcccg tcgacgagtc taacggacac caaccagcga 2220
      accagcagcg tcgctcgagg ccaagcgaag cagacggcac ggcattctctg tcgctgcctc 2280
      tggacccttc tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcatccagaa 2340
55     attgctggc ggagcggcag acgtgagccg gcacggcagg cggcctctc ctctctcac 2400
      ggcacggcag ctacggggga ttcctttccc accgctccta ctagaactag tggatcctat 2460
      gcgtatggtg tgacgtgtgt tcaagatgat gacttcaaac ctacctatga cgtatggtat 2520
      gacgtgtgtc gactgatgac ttagatccac tcgagcggct ataaatacgt acctacgcac 2580
60     cctgcgtac catccctaga gctgcatgct tatttttaca 2620

```

<210> 18

<211> 3433

65 <212> ADN

<213> Secuencia Artificial

ES 2 292 234 T3

<220>

<223> Descripción de la Secuencia Artificial: secuencia del promotor UCP4 que contiene 4 copias del ubil UAR ligado de manera operable con el promotor central.

5

<400> 18

```

tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata ttttttttgt 60
cacacttggt tgaagtgcag tttatctatc tttatacata tatttaaact ttactctacg 120
aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataatgaac 180
agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc 240
tttttagtgt gcatgtgttc tccttttttt ttgcaaatag cttcacctat ataatacttc 300
atccatttta ttagtacatc catttagggg ttaggggtaa tgggttttat agactaattt 360
ttttagtaca tctattttat tctatttttag cctctaaatt aagaaaacta aaactctatt 420
ttagtttttt tatttaataa tttagatata aaatagaata aaataaagt actaaaaatt 480
aaacaaatac cctttaagaa attaaaaaaa ctaaggaaac atttttcttg tttcgagtag 540
ataatgccag cctgttaaac gccgtcgacg agtctaacgg acaccaacca gcgaaccagc 600
agcgtcgcg tgggccaagc gaagcagacg gcacggcatc tctgtcgctg cctctggacc 660
cctctcgaga gttccgctcc accgttggaac ttgctccgct gtcggcatcc agaaattgag 720
tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcacggcacg 780
gcagctacgg gggattcctt tcccacgct cctactagag ataatgagca ttgcatgtct 840
aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagt cagtttatct 900
atctttatac atatatattaa actttactct acgaataata taatctatag tactacaata 960
atatcagtgt tttagagaat catataaatg aacagttaga catggtctaa aggacaattg 1020
agtattttga caacaggact ctacagtttt atcttttttag tgtgcatgtg ttctcctttt 1080
tttttgcaaa tagcttcacc tatataatac ttcattccatt ttattagtag atccatttag 1140
ggtttagggg taatgggttt tatagactaa tttttttagt acatctattt tattctattt 1200
tagcctctaa attaagaaaa ctaaaactct attttagttt ttttatttaa taatttagat 1260
ataaaataga ataaaataaa gtgactaaaa attaaacaaa taccctttaa gaaattaaaa 1320
aaactaagga aacatttttc ttgtttcgag tagataatgc cagcctgtta aacgccgtcg 1380
acgagtctaa cggacaccaa ccagcgaacc agcagcgtcg cgtcgggcca agcgaagcag 1440
acggcacggc atctctgtcg ctgcctctgg acccctctcg agagtccgc tccaccgttg 1500
gacttgctcc gctgtcggca tccagaaatt gcgtggcgga gcggcagacg tgagccggca 1560
cggcaggcgg cctcctctc ctctcacggc acggcagcta cgggggattc ctttcccacc 1620

```

45

50

55

60

65

ES 2 292 234 T3

gctcctacta gagataatga gcattgcatg tctaagttat aaaaaattac cacatatattt 1680
 ttttgtcaca cttgtttgaa gtgcagttta tctatcttta tacatatattt taaactttac 1740
 tctacgaata atataatcta tagtactaca ataatatcag tgttttagag aatcatataa 1800
 atgaacagtt agacatggc taaaggacaa ttgagtattt tgacaacagg actctacagt 1860
 tttatctttt tagtgtgcat gtgttctcct tttttttgc aaatagcttc acctatataa 1920
 tacttcatcc attttattag tacatccatt tagggtttag ggtaaatggt ttttatagac 1980
 taattttttt agtacatcta ttttattcta ttttagcctc taaattaaga aaactaaac 2040
 tctatttttag tttttttatt taataattta gatataaaat agaataaaat aaagtgacta 2100
 aaaattaaac aaataccctt taagaaatta aaaaaactaa ggaacattt ttcttgtttc 2160
 gagtagataa tgccagcctg ttaaacgccg tcgacgagtc taacggacac caaccagcga 2220
 accagcagcg tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg tcgctgcctc 2280
 tggaccctc tcgagagttc cgctccaccg ttggacttgc tccgctgtcg gcatccagaa 2340
 attgcgtggc ggagcggcag acgtgagccg gcacggcagg cggcctcctc ctctctcac 2400
 ggcacggcag ctacggggga ttctttccc accgctccta ctagagataa tgagcattgc 2460
 atgtctaagt tataaaaaat taccacatat ttttttgc acacttgtt gaagtgcagt 2520
 ttatctatct ttatacatat atttaaactt tactctacga ataataaat ctatagtact 2580
 acaataatat cagtgtttta gagaatcata taaatgaaca gttagacatg gtctaaagga 2640
 caattgagta ttttgacaac aggactctac agttttatct ttttagtgtg catgtgttct 2700
 ctttttttt tgcaaatagc ttacacctata taatacttca tccattttat tagtacatcc 2760
 atttaggggt tagggttaat ggtttttata gactaatttt tttagtacat ctattttatt 2820
 ctatttttagc ctctaaatta agaaaaactaa aactctattt tagttttttt atttaataat 2880
 ttagatataa aatagaataa aataaagtga ctaaaaatta aacaaatacc ctttaagaaa 2940
 ttaaaaaaac taaggaaaca tttttcttgt ttcgagtaga taatgccagc ctgttaaacy 3000
 ccgtcgacga gtctaacgga caccaaccag cgaaccagca gcgtcgcgtc gggccaagcg 3060
 aagcagacgg cagggcatct ctgtcgtcgc ctctggaccc ctctcgagag ttccgctcca 3120
 ccgttgagct tgctccgctg tcggcatcca gaaattgcgt ggcggagcgg cagacgtgag 3180
 ccggcacggc aggcggcctc ctctcctctc cagggcacgg cagctacggg ggattccttt 3240
 cccaccgctc ctactagaac tagtgatcc tatgcgtatg gtatgacgtg tgttcaagat 3300
 gatgacttca aacctaccta tgacgtatgg tatgacgtgt gtcgactgat gacttagatc 3360
 cactcgagcg gctataaata cgtacctacg caccctgcgc taccatccct agagctgcat 3420
 gcttattttt aca 3433