



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 301 539**

51 Int. Cl.:
C12N 15/82 (2006.01)
C12N 15/29 (2006.01)
C07K 14/415 (2006.01)
A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01930493 .0**
86 Fecha de presentación : **13.04.2001**
87 Número de publicación de la solicitud: **1272651**
87 Fecha de publicación de la solicitud: **08.01.2003**

54 Título: **Control de la dehiscencia de la fruta en *Arabidopsis* por genes indehiscentes 1.**

30 Prioridad: **13.04.2000 US 548971**

45 Fecha de publicación de la mención BOPI:
01.07.2008

45 Fecha de la publicación del folleto de la patente:
01.07.2008

73 Titular/es:
**The Regents of the University of California
University of California San Diego
Technology Transfer and Intellectual Proper
La Jolla, California 92093-0910, US**

72 Inventor/es: **Liljegren, Sarah y
Yanofsky, Martin, F.**

74 Agente: **Ponti Sales, Adelaida**

ES 2 301 539 T3

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

DESCRIPCIÓN

Control de la dehiscencia de la fruta en *Arabidopsis* por genes indehiscentes 1.

5 Declaración con respecto a los derechos de las invenciones realizadas bajo la investigación y el desarrollo con apoyo federal.

La presente invención se realizó con el apoyo del Gobierno bajo el número de la National Science Foundation Grant IBN-9985530. El Gobierno tiene ciertos derechos sobre la presente invención.

10 **Campo de la invención**

La presente invención se refiere a la ingeniería genética de plantas. En particular, la presente invención se refiere a procedimientos y composiciones que modulan la dehiscencia de los frutos en plantas.

15 **Antecedentes de la invención**

20 La semilla de colza es uno de los cultivos de semillas oleaginosas después de la soja y la semilla del algodón, representando el 10% de la producción mundial de semillas oleaginosas en 1990. La semilla de colza contiene un 40% de aceite, que se prensa a partir de la semilla, dejando una masa de semilla elevada en proteínas de valor para la alimentación de animales y fertilizante con nitrógeno. El aceite de la semilla de colza, también conocido como aceite de canola, es un producto valioso, que representa el cuarto aceite vegetal más habitualmente comercializado en el mundo.

25 Desafortunadamente, el rendimiento de la semilla del aceite de colza y plantas relacionadas está limitado por la dehiscencia de las vainas, que es un proceso que aparece de forma tardía en el desarrollo del fruto, mediante el cual la vaina se abre y se liberan las semillas encerradas. La degradación y la separación de las paredes celulares a lo largo de una capa discreta de células dividiendo las dos mitades de la vaina, denominadas la "zona de dehiscencia", dan lugar a la separación de las dos mitades de la vaina y la liberación de las semillas contenidas. La zona de dehiscencia es una
30 región de sólo una a tres células de anchura que se extiende a lo largo a la longitud completa del límite válvula/replum (Meakin y Roberts, J. Exp. Botany 41: 995-1002 (1990)). A medida que las células en la zona de dehiscencia se separan entre sí, las válvulas se separan del replum, permitiendo la dispersión de las semillas. El "desgrane" de las semillas, mediante el cual las semillas son despojadas prematuramente a través de la dehiscencia antes de poder recoger la cosecha, es un problema significativo afrontado por los productores de semillas comerciales y representa una pérdida
35 de ingresos en la industria. Las condiciones climatológicas adversas exacerban el proceso de dehiscencia, dando lugar a más de un 50% en la pérdida de rendimiento de las semillas.

El fruto, una estructura compleja única de las plantas que florecen, media en la maduración y dispersión de las semillas. En la mayoría de las plantas que florecen el fruto consiste en el pericarpio, que deriva de la pared del ovario,
40 y las semillas, que se desarrollan a partir de óvulos fertilizados. La *Arabidopsis*, que es la habitual de más de 3000 especies de *Brassicaceae*, produce un fruto en el que las dos válvulas de carpelo (paredes del ovario) se unen al replum, una sutura visible que divide los dos carpelos.

La hormona etileno de las plantas se produce mediante el desarrollo de semillas y parece ser un importante regulador del proceso de dehiscencia. Una línea de evidencia que soporta un papel para el etileno en la regulación de la dehiscencia proviene de estudios de maduración de los frutos, que, como la dehiscencia de los frutos, es un proceso que implica la rotura de la materia de la pared celular. En la maduración del fruto, el etileno actúa en parte mediante la activación de las enzimas que degradan la pared celular, tales como poligalacturonasa (Theologis *et al.*, Develop. Genetics 14: 282-295 (1993)). Además, en plantas de tomate modificadas genéticamente en que se bloquea la respuesta del etileno, tales como plantas transgénicas de tomate que expresan poligalacturonasa antisentido, existe un retraso significativo en la maduración del fruto (Lanahan *et al.*, The Plant Cell 6: 521-530 (1994); Smith *et al.*, Nature 334: 724-726 (1988)).

En la dehiscencia, cambios estructurales que culminan en la degradación de la lamela media de las paredes celulares de la zona de dehiscencia debilitan las vainas de las semillas de colza y finalmente conducen al desgrane de la vaina. Como en la maduración del fruto, las enzimas hidrolíticas incluyendo las poligalacturonasas juegan un papel en esta rotura programada. Por ejemplo, en la colza oleaginosa, una endopoligalacturonasa específica, RDPG1, se sobreexpresa y se expresa exclusivamente en la zona de dehiscencia de forma tardía en el desarrollo de las vainas (Petersen *et al.*, Plant Mol. Biol. 31: 517-527 (1996), que se incorpora en la presente por referencia). El etileno puede regular
60 la actividad de enzimas hidrolíticas implicadas en el proceso de dehiscencia al igual que en la maduración del fruto (Meakin y Roberts, J. Exp. Botany 41: 1003-1011 (1990), que se incorpora en la presente por referencia). Sin embargo, hasta ahora, las proteínas que controlan el proceso de dehiscencia, tales como las que regulan las enzimas hidrolíticas pertinentes, no han sido identificadas.

65 Los intentos para resolver el problema del desgrane de la vaina y la dehiscencia prematura del fruto durante los últimos 20 años se han centrado en la reproducción de variedades resistentes al desgrane. Sin embargo, estos híbridos de plantas son frecuentemente estériles y pierden las características favorables que se deben recuperar mediante retrocruzamiento, que consume tiempo y es laborioso. Otras estrategias para paliar el desgrane de las vainas incluyen el

uso de productos químicos, tales como sellantes de vainas o técnicas mecánicas, tales como la envoltura para reducir el desgrane estimulado por el viento. Hasta la fecha, sin embargo, no se ha descrito un procedimiento sencillo para producir plantas modificadas genéticamente que no se abran y liberen sus semillas de manera prematura.

- 5 De este modo, existe la necesidad para identificar genes que regulen el proceso de dehiscencia y para desarrollar genéticamente variedades de plantas modificadas en las que se retrasa el proceso natural de dispersión de las semillas. La presente invención satisface esta necesidad y proporciona también ventajas relacionadas.

10 Descripción resumida de la invención

La presente invención proporciona una cassette de expresión que comprende un promotor unido operativamente a una secuencia de polinucleótidos IND1, o complemento de la misma, que codifica un polipéptido IND1 idéntica por lo menos aproximadamente en un 70% a la SEC ID No: 1. El cassette de expresión, por ejemplo, puede comprender un polinucleótido que codifica la SEC ID No: 2. En otra realización, el cassette de expresión, por ejemplo, puede comprender las posiciones desde aproximadamente 2765 hasta aproximadamente 3361 de la SEC ID No: 1. Por ejemplo, el cassette de expresión puede comprender la SEC ID No: 1. En algunas realizaciones, el cassette de expresión comprende un promotor. El promotor, por ejemplo, puede ser constitutivo o específico de tejido. En un aspecto de la presente invención, el promotor es un promotor específico de la zona de dehiscencia. En otro aspecto, el promotor puede comprender las posiciones desde aproximadamente 1 hasta aproximadamente 2764 o las posiciones desde aproximadamente 3362 hasta aproximadamente 3856 de la SEC ID No: 1.

La presente invención también proporciona una planta que comprende un cassette de expresión recombinante que comprende un promotor unido operativamente a una secuencia de polinucleótidos que codifica un polipéptido IND1 idéntica por lo menos aproximadamente en un 70% a la SEC ID No: 1. En un aspecto, la secuencia de polinucleótidos que codifica el polipéptido IND1 está unida operativamente al promotor en la orientación antisentido. En otro aspecto, la secuencia de polinucleótidos que codifica el polipéptido IND1 está unida operativamente al promotor en la orientación de sentido. La secuencia de polinucleótidos puede comprender además una segunda secuencia de polinucleótidos que codifica el polipéptido IND1, donde la segunda secuencia de polinucleótidos está unida operativamente a un segundo promotor en la orientación antisentido. En algunas realizaciones, la planta de la invención ha reducido la lignificación en células del margen de las válvulas. En otras realizaciones, el promotor es un elemento regulador selectivo de la zona de dehiscencia. En algunas de estas realizaciones, el elemento regulador comprende las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1.

La presente invención también proporciona un procedimiento de retraso de la dehiscencia del fruto en una planta que comprende la supresión de la expresión de un ácido nucleico de IND1 en la planta mediante la introducción en la planta de un cassette de expresión recombinante que comprende un promotor unido operativamente a una secuencia de polinucleótidos que codifica un polipéptido IND1 idéntica por lo menos en un 70% a la SEC ID No: 2. En algunas realizaciones, el polipéptido IND1 es la SEC ID No: 2. En otras realizaciones, el polinucleótido IND1 comprende las posiciones desde aproximadamente 2765 hasta aproximadamente 3361 de la SEC ID No: 1. En un aspecto de la presente invención, el polinucleótido IND1 comprende la SEC ID No: 1. El procedimiento puede incluir una secuencia de polinucleótidos que codifica el polipéptido IND1 unido operativamente al promotor en la orientación antisentido. En otra realización, el promotor está unido al promotor en la orientación de sentido. La secuencia de polinucleótidos puede comprender además una segunda secuencia de polinucleótidos que codifica el polipéptido IND1, donde la segunda secuencia de polinucleótidos está unida operativamente a un segundo promotor en la orientación antisentida. En algunas realizaciones, el procedimiento da lugar a una planta con una lignificación reducida en las células del margen de las válvulas. En otras realizaciones, el promotor es un elemento regulador selectivo de la zona de dehiscencia. En algunas de estas realizaciones, el elemento regulador comprende las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1. En un aspecto, el cassette de expresión recombinante se introduce en la planta con *Agrobacterium*.

La presente invención también proporciona un procedimiento de retraso de la dehiscencia del fruto en una planta que comprende la supresión de la expresión de un gen de IND1 en la planta mediante la introducción en la planta de un cassette de expresión recombinante que comprende una secuencia de polinucleótidos idéntica por lo menos aproximadamente en un 70% con las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1. En un aspecto de la presente invención, la secuencia de polinucleótidos comprende las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1. En algunos aspectos, la lignificación se reduce en las células del margen de las válvulas.

Definiciones

65 La expresión “ácido nucleico” se refiere a un polímero de cadena única o doble cadena de bases de desoxiribonucleótido o ribonucleótido leído desde el extremo 5’ al extremo 3’. Los ácidos nucleicos también pueden incluir nucleótidos modificados que permiten la lectura correcta a través de una polimerasa y no alteran la expresión de un polipéptido codificado por el ácido nucleico.

ES 2 301 539 T3

La expresión “secuencia de polinucleótidos” o “secuencia de ácidos nucleicos” incluye las cadenas de sentido y antisentido de un ácido nucleico como cadenas únicas individuales o en dobles cadenas. Incluye, pero no se limita a, plásmidos autoreplicantes, secuencias cromosómicas y polímeros infecciosos de ADN o ARN.

5 La expresión “secuencia de ácido nucleico que codifica” se refiere a un ácido nucleico que dirige la expresión de una proteína o péptido específico. Las secuencias de ácidos nucleicos incluyen tanto la secuencia de cadena de ADN que se transcribe en ARN y la secuencia de ARN que se traduce en proteína. Las secuencias de ácidos nucleicos incluyen las secuencias de ácidos nucleicos de longitud completa, así como las secuencias que no son de longitud completa derivadas de las secuencias de longitud completa. Debería entenderse además que la secuencia incluye los
10 codones degenerados de la secuencia o secuencias nativas que se pueden introducir para proporcionar una preferencia de codón en una célula huésped específica.

El término “promotor” o “elemento regulador” se refiere a una región o secuencia determinantes localizadas en dirección 5’ o dirección 3’ desde el inicio de la transcripción y que están implicadas en el reconocimiento y unión
15 de ARN polimerasa y otras proteínas para iniciar la transcripción. Un “promotor de planta” es un promotor capaz de iniciar la transcripción en células vegetales. Dichos promotores no necesitan tener un origen vegetal, por ejemplo, se pueden utilizar en la presente invención promotores derivados de virus de plantas, tales como el promotor CaMV35S.

Tal como se utiliza en la presente invención, el término “elemento regulador selectivo de la zona de dehiscencia”
20 se refiere a una secuencia de nucleótidos que, cuando está unida operativamente a una molécula de ácido nucleico, confiere una expresión selectiva sobre la molécula de ácido nucleico unida operativamente en un número limitado de tejidos vegetales, incluyendo el margen de las válvulas o zona de dehiscencia. El margen de las válvulas es el futuro sitio de la zona de dehiscencia y comprende los márgenes del replum externo, así como las células de las válvulas adyacentes al replum externo. La zona de dehiscencia, que se desarrolla en la región del margen de las válvulas, se
25 refiere al grupo de células que se separan durante el proceso de dehiscencia, permitiendo que las válvulas se separen del replum y las semillas atrapadas se liberen. De este modo, un elemento regulador selectivo de la zona de dehiscencia, tal como se define en la presente invención, confiere una expresión selectiva en la zona de dehiscencia madura, o confiere una expresión selectiva en el margen de las válvulas, que marca el futuro sitio de la zona de dehiscencia.

30 Un elemento regulador selectivo de la zona de dehiscencia puede conferir una expresión específica exclusivamente en células del margen de las válvulas o zona de dehiscencia o puede conferir una expresión selectiva en un número limitado de células vegetales tipos incluyendo las células del margen de las válvulas o zona de dehiscencia. Un elemento regulador *SHATTERPROOF1* o *SHATTERPROOF2* (*SHP1* y *SHP2*, también denominados como *AGL1* y
35 *AGL5*, respectivamente), por ejemplo, que confiere una expresión selectiva en óvulos y la placenta, así como en la zona de dehiscencia, es un elemento regulador selectivo de la zona de dehiscencia tal como se define en la presente invención. De manera similar, un elemento regulador de *IND1* también confiere una expresión selectiva en la zona de dehiscencia. Un elemento regulador selectivo de la zona de dehiscencia generalmente se distingue de otros elementos reguladores confiriendo una expresión selectiva en el margen de las válvulas o la zona de dehiscencia sin conferir
40 expresión a través de las válvulas de carpelo adyacentes.

Se entiende que se pueden realizar modificaciones limitadas sin destruir la función biológica de un elemento regulador y que dichas modificaciones limitadas pueden dar lugar a elementos reguladores selectivos de la zona de dehiscencia que tienen una función sustancialmente equivalente o mejorada en comparación con un elemento regulador
45 de *IND1* de tipo natural. Estas modificaciones pueden ser deliberadas, como a través de la mutagénesis dirigida de sitio, o pueden ser accidentales, como a través de la mutación en huéspedes que albergan el elemento regulador. Todas estas secuencias de nucleótidos modificadas están incluidas en la definición de un elemento regulador selectivo de la zona de dehiscencia siempre y cuando la capacidad para conferir una expresión selectiva en el margen de las válvulas o la zona de dehiscencia se mantenga sustancialmente.

50 El término “planta” incluye plantas completas, órganos/estructuras vegetativas de un brote (por ejemplo, hojas, tallos, y tubérculos), raíces, flores y órganos/estructuras florales (por ejemplo, brácteas, sépalos, pétalos, estambres, carpelos, anteras y óvulos), semilla (incluyendo embriones, endospermas y recubrimiento de semilla) y fruto (el ovario maduro), tejido vegetal (por ejemplo, tejido vascular, tejido fundamental y similares) y células (por ejemplo, células
55 de guarda, óvulos, tricomas y similares), y la progenie de los mismos. La clase de plantas que se pueden utilizar en procedimiento de la presente invención es generalmente tan amplia como la clase de plantas superiores e inferiores susceptibles para técnicas de transformación, incluyendo angiospermas (plantas monocotiledóneas y dicotiledóneas), gimnospermas, helechos y algas multicelulares. Incluye plantas de una variedad de niveles de ploidía, incluyendo aneuploide, poliploide, diploide, haploide y hemicigótico.

60 El término “planta con semilla” significa angiosperma o gimnosperma. Una angiosperma es una planta portadora de semillas, las cuales se transportan en un ovario maduro (fruto). Una angiosperma se reconoce habitualmente como una planta que florece. Las angiospermas se dividen en dos clases ampliar en base al número de cotiledóneas, que son las hojas de semilla que generalmente almacenan o absorben comida. De este modo, una angiosperma monocotiledónea es una angiosperma que tiene una única cotiledónea, mientras que una angiosperma dicotiledónea es una angiosperma que tiene dos cotiledóneas. Se conoce una serie de angiospermas que incluyen, por ejemplo, plantas de colza, plantas leguminosas, plantas con fruto, flores ornamentales, plantas de cereales y árboles de madera noble, cuyas clases generales no son necesariamente exclusivas. El técnico en la materia entenderá que los procedimientos

ES 2 301 539 T3

de la presente invención se pueden poner en práctica estas u otras angiospermas, según se desee. Una gimnosperma es una planta portadora de semillas con semillas no incluidas en un ovario.

La expresión “célula huésped” se refiere a una célula de cualquier organismo. Las células huésped preferidas derivan de plantas, bacterias, levaduras, hongos, insectos u otros animales. Los procedimientos para introducir secuencias de polinucleótidos en varios tipos de células huésped son bien conocidos en la técnica.

El término “retrasado”, tal y como se utiliza en la presente invención en referencia al tiempo para la dispersión de las semillas en un fruto producido por una planta no natural de la invención, significa un tiempo significativamente posterior de la dispersión de las semillas en comparación con el tiempo en que las semillas se dispersan normalmente de una planta correspondiente en la misma etapa de desarrollo que expresa niveles naturales de IND1. De este modo, el término “retrasado” se utiliza ampliamente para comprender tanto la dispersión de las semillas que se postpone significativamente en comparación con la dispersión de las semillas en una planta correspondiente, como cuando la dispersión de las semillas está totalmente descartada, de manera que los frutos nunca liberan sus semillas a menos que exista intervención humana o cualquier otro tipo de intervención.

Se entiende que puede haber una variación natural del tiempo de dispersión de las semillas en una especie o variedad de planta. Sin embargo, un “retraso” en el tiempo de la dispersión de las semillas en una planta no natural de la presente invención se puede identificar fácilmente mediante el muestreo de una población de las plantas no naturales y la determinación de que la distribución normal de los tiempos de dispersión de las semillas son significativamente posteriores, de promedio, que la distribución normal de los tiempos de dispersión de las semillas en una población de la especie o variedad de planta correspondiente que no contiene un polinucleótido IND1 exógeno. De este modo, la producción de plantas no naturales de la presente invención proporciona un medio para variar la distribución normal del tiempo de dispersión de las semillas de la polinización, de manera que las semillas se dispersan, de promedio, por lo menos aproximadamente un 1%, 2%, 5%, 10%, 30%, 50%, 100%, 200% ó 500% más tarde que la especie de planta correspondiente que no contiene una molécula de ácido nucleico exógeno que codifica un producto génico de IND1.

El término “suprimido” o “disminuido” comprende la ausencia de proteína IND1 en una planta, así como la expresión de la proteína que está presente, pero reducida, en comparación con el nivel de expresión de proteína IND1 en una planta de tipo natural. Además, el término “suprimido” se refiere a la expresión de la proteína IND1 que se reduce a través del dominio completo de la expresión de IND1 o a la expresión que se reduce en alguna parte del dominio de expresión de IND1, con la condición de que la planta resultante esté caracterizada por una dispersión de las semillas retardada. El término “suprimido” también comprende una cantidad de proteína IND1 que es equivalente a la expresión de IND1 de tipo natural, pero donde la proteína IND1 tiene un nivel de actividad reducido. Tal como se ha descrito anteriormente, cada IND1 contiene un dominio HLH conservado y básico; mutaciones puntuales o deleciones masivas con el dominio HLH que reducen la actividad de unión a ADN de IND1 pueden reducir o destruir la actividad de IND y, por lo tanto, “suprimir” la expresión de IND1 tal como se define en la presente invención. Un experto en la materia entenderá que, preferiblemente, la expresión de IND1 está esencialmente ausente en el margen de las válvulas de una planta o la proteína IND1 es esencialmente no funcional.

Una actividad o expresión de IND1 “incrementada” o “aumentada” de un gen *IND1* se refiere a un cambio por aumento en la actividad de IND1. Entre los ejemplos de dicha actividad o expresión incrementada se incluyen los siguientes: actividad o expresión de IND1 del gen *IND1* se incrementa por encima del nivel plantas de control no transgénicas de tipo natural (es decir, aumenta la cantidad de actividad o expresión de IND1 del gen *IND1*). La actividad o expresión de IND1 del gen *IND1* es un órgano, tejido o célula donde no se detecta normalmente en plantas de control no transgénicas de tipo natural (es decir, aumenta la distribución espacial de la actividad o expresión de IND1 del gen *IND1*). La actividad o expresión de IND1 se incrementa cuando la actividad o expresión de IND1 del gen *IND1* está presente en un órgano, tejido o célula durante un periodo más largo que en plantas de control no transgénicas de tipo natural (es decir, aumenta la duración de la actividad o expresión de IND1 del gen *IND1*).

Una secuencia de polinucleótidos es “heteróloga a” una segunda secuencia de polinucleótidos si se origina a partir de una especie foránea, o, si se origina a partir de la misma especie, es modificada por la acción humana a partir de su forma original. Por ejemplo, un promotor unido operativamente a una secuencia codificante heteróloga se refiere a una secuencia codificante de una especie diferente de aquella de la que deriva el promotor, o, si se origina de la misma especie, una secuencia codificante que es diferente de cualquier variante alélica natural.

Un polinucleótido “exógeno a” una planta individual es un polinucleótido que se introduce en la planta, o una generación predecesora de la planta, mediante cualquier medio diferente del cruzamiento sexual. Una molécula de ácido nucleico exógena puede tener una secuencia de nucleótidos natural o no natural y puede ser una molécula de ácido nucleico heteróloga derivada de especies de plantas diferentes de la planta en la que se introduce la molécula de ácido nucleico o puede ser una molécula de ácido nucleico derivada de la misma especie de planta que la planta en la que se introduce. A continuación se describen ejemplos de medios mediante los cuales esto se puede llevar a cabo e incluyen la transformación mediada por *Agrobacterium*, métodos biolísticos, electroporación, técnicas en plantas, y similares.

Un “polinucleótido *IND1*” es una secuencia de ácidos nucleicos que comprende (o consiste en) una región codificante de aproximadamente 50 a aproximadamente 4000 nucleótidos, algunas veces de aproximadamente 100 a aproximadamente 3000 nucleótidos y algunas veces de aproximadamente 200 a aproximadamente 600 nucleótidos,

ES 2 301 539 T3

que se hibrida la SEC ID No: 1 en condiciones astringentes (tal como se definen a continuación), o que codifica un polipéptido IND1 o fragmento de por lo menos 15 aminoácidos del mismo. Los polinucleótidos IND1 también se pueden identificar por su capacidad para hibridarse en condiciones de baja astringencia (por ejemplo, $T_m \sim 40^\circ\text{C}$) a sondas de ácidos nucleicos que tienen la secuencia de SEC ID No: 1. La SEC ID No: 1 es un ejemplo de un polinucleótido IND1.

Un “promotor de un gen *IND1*” o “promotor *IND1*” tendrá habitualmente de aproximadamente 500 a aproximadamente 3000 nucleótidos de longitud, normalmente de aproximadamente 750 a 2750. Las secuencias de promotores de ejemplo se muestran como SEC ID No: 3 y SEC ID No: 4. La SEC ID No: 3 representa la región no traducida 5’ del *IND1* y la SEC ID No: 4 representa la región no traducida 3’ de *IND1*. Un promotor *IND1* también se puede identificar por su capacidad para dirigir la expresión en el margen de las válvulas del fruto. En particular el promotor *IND1* dirige la expresión en el margen de las válvulas del gineceo en desarrollo justo antes de la fertilización (etapa 13) a través de la maduración del fruto (etapa 17). El promotor no proporciona una expresión significativa en el tejido de la hoja.

Un “polipéptido IND1” es una secuencia de aproximadamente 50 a aproximadamente 200, algunas veces de 100 a 190, y preferiblemente 198 residuos de aminoácidos codificados por un polinucleótido IND1. Los polipéptidos IND1 se caracterizan por la presencia de un dominio básico hélice-bucle-hélice (HLH) que se une a secuencias de polinucleótidos específicas. Por ejemplo, los residuos de aminoácidos ISDDPQTVVARRRRERISEKIRILKRIVPG GAKMDTASMLDEAIRY TKFLK representan el dominio HLH del polipéptido mostrado en la SEC ID No: 2. El dominio HLH es conocido en la técnica y es compartido por otros factores de transcripción incluyendo secuencias no caracterizadas representadas por el número de acceso del Banco de Genes E1283552 y 2262147 y el producto génico PIF3 (Ni *et al.* Cell 95: 657 (1998)). El dominio HLH de IND1 es por tanto un dominio de unión a ADN.

Tal como se utiliza en la presente invención, un homólogo de un gen IND1 particular (por ejemplo, la SEC ID No: 1) es un segundo gen en el mismo tipo de planta o en un tipo de planta diferente, que tiene una secuencia de polinucleótidos de por lo menos 50 nucleótidos contiguos que son sustancialmente idénticos (determinada tal y como se describe a continuación) a una secuencia en el primer gen. Se cree que, en general, los homólogos comparten un pasado evolutivo común.

Una “secuencia de polinucleótidos de” un gen particular es una subsecuencia o secuencia de polinucleótidos de longitud completa de un gen IND1 que, cuando está presente en una planta transgénica, tiene el efecto deseado. Por ejemplo, un efecto es la inhibición de la expresión del gen endógeno que conduce la expresión de un polinucleótido heterólogo. Una secuencia de longitud completa de un gen particular descrito en la presente invención puede contener aproximadamente un 95%, habitualmente por lo menos aproximadamente un 98% de una secuencia completa mostrada en el Listado de Secuencias posterior.

El término “tejidos reproductores” tal como se utiliza en la presente invención incluye fruto, óvulos, semillas, polen, pistilos, flores o cualquier tejido embrionario.

Un “cassette de expresión” se refiere a una construcción de ácidos nucleicos que, cuando se introduce en una célula huésped, da lugar a la transcripción y/o traducción de un ARN o polipéptido, respectivamente. Las construcciones antisentido o de sentido que no se traducen o no se pueden traducir están incluidos expresamente en esta definición.

En el caso de tanto la expresión de transgenes como la inhibición de genes endógenos (por ejemplo, por supresión antisentido o de sentido), un experto entenderá que la secuencia de polinucleótidos insertada no necesita ser idéntica y puede ser “sustancialmente idéntica” a una secuencia del gen del que deriva. Tal como se explica a continuación, estas variantes están específicamente cubiertas por este término.

En el caso en el que la secuencia de polinucleótidos insertada se transcribe y traduce para producir un polipéptido funcional, un experto entenderá que debido a la degeneración de codones una serie de secuencias de polinucleótidos codificarán el mismo polipéptido. Estas variantes están específicamente cubiertas por el término “secuencia de polinucleótidos de” un gen particular del margen de las válvulas, tal como *IND1*. Además, el término incluye específicamente las secuencias (por ejemplo, las secuencias de longitud completa) sustancialmente idénticas (determinadas tal como se describe a continuación) con una secuencia del gen IND1 y que codifican proteínas que retienen la función de un polinucleótido IND1.

En el caso de polinucleótidos utilizados para inhibir la expresión de un gen endógeno, la secuencia introducida no necesita ser perfectamente idéntica a una secuencia del gen endógeno diana. La secuencia de polinucleótidos introducida habitualmente será por lo menos sustancialmente idéntica (tal como se determina a continuación) a la secuencia endógena diana.

Dos secuencias de ácidos nucleicos o polipéptidos se dice que son “idénticas” si la secuencia de nucleótidos o residuos de aminoácidos, respectivamente, en las dos secuencias es la misma cuando se alinea para la máxima correspondencia tal como se describe a continuación. El término “complementaria a” se utiliza en la presente invención para indicar que la secuencia es complementaria a toda o parte de una secuencia de polinucleótidos de referencia.

La alineación óptima de secuencias para la comparación se puede realizar mediante el algoritmo de homología local de Smith y Waterman Add. APL. Math. 2: 482 (1981), mediante el algoritmo de alineación por homología de

Needle man y Wunsch J. Mol. Biol. 48: 443 (1970), mediante el método de búsqueda de similitud de Pearson y Lipman Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444 (1988), mediante implementaciones informáticas de estos algoritmos (GAP, BESTFIT, BLAST, FASTA y TFASTA en el Paquete Informático de Genética de Wisconsin, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI) o mediante inspección.

5

El “porcentaje de identidad en la secuencia” se determina comparando dos secuencias alineadas de manera óptima sobre una ventana de comparación, donde la parte de la secuencia de polinucleótidos en la ventana de comparación puede comprender adiciones o deleciones (es decir, espacios) en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en que la base del ácido nucleico o el residuo de aminoácido idénticos aparecen en ambas secuencias para producir el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando el resultado por 100 para producir el porcentaje de identidad en la secuencia.

El término “identidad sustancial” de secuencias de polinucleótidos significa que un polinucleótido comprende una secuencia que tiene por lo menos una identidad en la secuencia del 25%. Alternativamente, el porcentaje de identidad puede ser cualquier número entero desde el 25% al 100%. Las realizaciones más preferidas incluyen por lo menos: 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 99%, en comparación con una secuencia de referencia utilizando los programas descritos en la presente invención; preferiblemente BLAST utilizando parámetros estándar, tal como se describe a continuación. Por consiguiente, las secuencias de IND1 de la presente invención incluyen secuencias de ácidos nucleicos que tienen una identidad sustancial con la SEC ID No: 1. Las secuencias de IND1 de la presente invención también incluyen secuencias de polipéptidos que tienen una identidad sustancial con la SEC ID No: 2. Un experto entenderá que estos valores se pueden ajustar de manera apropiada para determinar la identidad correspondiente de proteínas codificadas por dos secuencias de nucleótidos teniendo en cuenta la degeneración de codones, la similitud de aminoácidos, la posición del marco de lectura y similares. La identidad sustancial de secuencias de aminoácidos para estos objetivos significa normalmente la identidad en la secuencia de por lo menos un 40%. El porcentaje preferido de identidad de polipéptidos puede ser cualquier número entero desde el 40% al 100%. Las realizaciones más preferidas incluyen por lo menos un 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o 99%. Las realizaciones aún más preferidas incluyen por lo menos un 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74% y 75%. Los polipéptidos que son “sustancialmente similares” comparten secuencias tal como se ha indicado anteriormente a excepción de que las posiciones de los residuos que no son idénticos pueden diferir por cambios de aminoácidos conservativos. Las sustituciones de aminoácidos conservativos se refieren a la posibilidad de intercambio de residuos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales hidroxilalifáticas es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Los grupos de sustitución de aminoácidos conservativos preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina, ácido aspártico-ácido glutámico, y asparagina-glutamina.

Otra indicación de que las secuencias de nucleótidos son sustancialmente idénticas es si dos moléculas se hibridan entre sí, o a un tercer ácido nucleico, en condiciones astringentes. Las condiciones astringentes son dependientes de las secuencias y serán diferentes en circunstancias diferentes. En general, las condiciones astringentes se seleccionan para que sean aproximadamente 5°C inferior al punto de fusión térmico (T_f) para la secuencia específica a una fuerza iónica y pH definidos. La T_f es la temperatura (a una fuerza iónica y pH definidos) a la que el 50% de la secuencia diana se hibrida a una sonda perfectamente coincidente. Habitualmente, las condiciones astringentes serán aquellas en las que la concentración de sal es de aproximadamente 0,01 molar a pH 7 y la temperatura es por lo menos aproximadamente 60°C.

50

En la presente invención, el ARNm codificado por genes IND1 de la presente invención se pueden identificar en transferencias Northern en condiciones astringentes utilizando ADNc de la presente invención o fragmentos de por lo menos aproximadamente 100 nucleótidos. Para los objetivos de esta descripción, las condiciones astringentes para dichas hibridaciones ARN-ADN son aquellas que incluyen por lo menos un lavado en 0,2xSSC a 63°C durante 20 minutos, o condiciones equivalentes. El ADN genómico o ADNc que comprende genes de la presente invención se pueden identificar utilizando los mismos ADNc (o fragmentos de por lo menos aproximadamente 100 nucleótidos) en condiciones astringentes, que para los objetivos de esta descripción, incluyen por lo menos un lavado (habitualmente dos) en 0,2xSSC a una temperatura de por lo menos aproximadamente 50°C, normalmente aproximadamente 55°C, durante 20 minutos, o condiciones equivalentes.

60

Breve descripción de los dibujos

La figura I ilustra los tipos de células del fruto de Arabidopsis en la madurez.

65

Descripción detallada

I. Introducción

5 La presente invención proporciona procedimientos de modulación del desarrollo del fruto en plantas. En particular, la presente invención proporciona procedimientos para el retraso o prevención de la dehiscencia del fruto mediante la supresión de la expresión de un gen bHLH, tal como *IND1* en una planta. La presente invención también proporciona plantas transgénicas que comprenden varios polinucleótidos que codifican un polipéptidos bHLH, tal como *IND1*.

10 La presente invención se refiere al descubrimiento previo de que una planta con el doble mutante *agl1 aglj* tiene fenotipo de la dispersión de semillas retardada (Liljegren *et al.*, Nature 404: 766-770 (2000). Las mutaciones con pérdida de función en los genes *SHP1* y *SHP2* se produjeron mediante la inserción disruptiva de ADN-T y la recombinación homóloga. En las plantas resultantes con doble mutante *shp1* y *shp2*, la zona de dehiscencia no consiguió desarrollarse con normalidad y los frutos maduros no experimentaron dehiscencia. De este modo, se requiere la expresión de los genes *SHP1* o *SHP2* para el desarrollo de la zona de dehiscencia. Estos resultados indican que *SHP1* y *SHP2* regulan la dehiscencia de la vaina y que la manipulación de la expresión de *SHP1* y *SHP2* puede permitir el proceso de desgrane de la vaina a controlar.

15 La presente invención proporciona evidencia de que *IND1* está regulado por *SHP1* y *SHP2* y que la expresión de *IND1* modula la dehiscencia del fruto. La presente invención también proporciona procedimientos para el retraso de la dehiscencia del fruto mediante la supresión de la expresión de *IND1*.

20 Los genes *SHP1* y *SHP2* de *Arabidopsis* codifican las proteínas caja MADS con un 85% de identidad a nivel de aminoácidos. Los patrones de expresión de ADN de *SHP1* y *SHP2* también son notablemente similares. En particular, ambos ARN se expresan específicamente en flores, donde se acumulan en los carpelos en desarrollo. En particular, la fuerte expresión de estos genes se observa en el replum externo a lo largo del límite válvula/replum (Ma *et al.*, *supra*, 1991; Savidge *et al.*, The Plant Cell 7: 721-723 (1995); Flanagan *et al.*, The Plant Journal 10: 343-353 (1996)). De este modo, *SHP1* y *SHP2* se expresan en el margen de las válvulas, por lo menos en las células del replum externo.

25 En general, la nomenclatura y los procedimientos de laboratorio en la tecnología de ADN recombinante descritos a continuación son aquellos bien conocidos y utilizados normalmente en la técnica. Las técnicas estándar se utilizan para la clonación, aislamiento de ADN y ARN, amplificación y purificación. Las reacciones generalmente enzimáticas que implican la ADN ligasa, la ADN polimerasa, las endonucleasas de restricción y similares, se realizan según las especificaciones del fabricante. Estas técnicas y varias otras técnicas se realizan en general según Sambrook *et al.*, Molecular Cloning - A Laboratory Manual, 2ª Ed., Cold Spring Harbor Laboratory, Cold spring Harbor, Nueva Cork, (1989).

II. Aislamiento de los ácidos nucleicos de la invención

30 El aislamiento de secuencias de los genes de la presente invención se puede realizar mediante una serie de técnicas. Por ejemplo, las sondas de oligonucleótidos basadas en las secuencias descritas en la presente invención se pueden utilizar para identificar el gen deseado en una biblioteca de ADNc o ADN genómico de una especie de planta deseada. Para construir bibliotecas genómicas, se generan grandes segmentos de ADN genómico mediante fragmentación aleatoria, por ejemplo, utilizando endonucleasas de restricción, y se ligan con ADN vector para formar concatémicos que se pueden empaquetar en el vector apropiado. Para preparar una biblioteca de ADNc específicas de embriones, tales como *IND1*, se aísla el ARNm de los embriones y se prepara a partir del ARNm una biblioteca de ADNc que contiene los transcritos de genes.

35 EL ADNc o biblioteca genómica se puede entonces cribar utilizando una sonda basada en la secuencia de un gen *IND1* clonado, tal como los polinucleótidos descritos en la presente invención. Las sondas se pueden utilizar para hibridarse con secuencias de ADN genómico o ADNc para aislar genes homólogos en la misma o diferentes especies de plantas.

40 Alternativamente, los ácidos nucleicos de interés se pueden amplificar a partir de las muestras de ácido nucleico utilizando técnicas de amplificación. Por ejemplo, la tecnología de la reacción en cadena de la polimerasa (PCR) para amplificar las secuencias de los genes directamente de ARNm, de ADNc, de bibliotecas genómicas o bibliotecas de ADNc. La PCR y otros procedimientos de amplificación *in vitro* también pueden ser útiles, por ejemplo, para clonar secuencias de ácidos nucleicos que codifican para proteínas a expresar, para fabricar ácidos nucleicos a utilizar como sondas para detectar la presencia del ARNm deseado en muestras, para secuenciar ácidos nucleicos, o para otros propósitos.

45 Los cebadores y sondas apropiadas para identificar genes, tales como *IND1*, de tejidos vegetales se generan a partir de comparaciones de las secuencias proporcionadas en la presente invención. Para una visión general de la PCR, véase PCR Protocols: A Guide to Methods and Applications. (Innis, M, Gelfand, D., Sninsky, J. y White, T., eds.), Academic Press, San Diego (1990). Los cebadores apropiados para la amplificación de la región genómica de *IND1* o el ADNc de *IND1* incluyen los siguientes pares de cebadores: 5'-gatgaaatggaaaatggtatgtata-3' and 5'-gttcacagggttggagttgtg-3'. Las condiciones de amplificación son habitualmente las siguientes. Componentes de reacción: 10 mM de Tris-HCl, pH 8,3, 50 mM de cloruro potásico, 1,5 mM de cloruro magnésico, geltain al 0,001%, 200 μ M de dATP, 200 μ M de

ES 2 301 539 T3

dCTP, 200 μ M de dGTP, 200 μ M de dTTP, 0,4 μ M de cebadores, y 100 unidades por ml de Taq polimerasa. Programa: 96 C durante 3 min., 30 ciclos de 96 C durante 45 s, 50 C durante 60 s, 72 durante 60 s, seguido de 72 C durante 5 min.

Los polinucleótidos también se pueden sintetizar mediante técnicas bien conocidas tal como se describe en la literatura técnica. Véase, por ejemplo, Carruthers *et al.*, Cold Spring Harbor Symp. Quant. Biol. 47: 411-418 (1982), y Adams *et al.*, J. Am. Chem. Soc. 105: 661 (1983). A continuación, se pueden obtener fragmentos de ADN de doble cadena mediante la síntesis de la cadena complementaria y la hibridación de las cadenas en condiciones apropiadas, o mediante la adición de la cadena complementaria utilizando ADN polimerasa con una secuencia de cebadores adecuada.

El género de secuencias de ácidos nucleicos IND1 de la presente invención incluye genes y productos génicos identificados y caracterizados por análisis utilizando secuencias de ácidos nucleicos de la presente invención, incluyendo la SEC ID No: 1, y secuencias de proteínas de la presente invención, incluyendo la SEC ID No: 2. Las secuencias de IND1 de la presente invención incluyen secuencias de ácidos nucleicos que tienen una identidad sustancial con la SEC ID No: 1. Las secuencias de IND1 de la presente invención también incluyen secuencias de polipéptidos que tienen una identidad sustancial con la SEC ID No: 2.

III. Uso de los ácidos nucleicos de la invención

A. Uso de los ácidos nucleicos de la invención para inhibir o suprimir la expresión de genes

La presente invención proporciona procedimientos de modulación de la dehiscencia del fruto en una planta mediante la introducción en una planta de un cassette de expresión recombinante que comprende un elemento regulador unido operativamente a un polinucleótido HLH, tal como IND1. La presente invención también proporciona procedimientos para retrasar la dispersión de semillas en una planta mediante la supresión de la expresión de una molécula de ácido nucleico que codifica un producto génico IND1. En una planta transgénica de la presente invención, una molécula de ácido nucleico, o construcciones antisentido de la misma, que codifican un producto génico IND1 pueden estar unidos operativamente a un elemento regulador exógeno. La presente invención proporciona, por ejemplo, una planta transgénica caracterizada por una dispersión de semillas retardada que tiene una molécula de ácido nucleico expresada que codifica un producto génico IND1, o construcción antisentido del mismo, que está unido operativamente a un elemento regulador constitutivo exógeno. En una realización, la presente invención proporciona una planta transgénica que está caracterizada por la dispersión de semillas retardada debido a la supresión de una molécula de ácido nucleico que codifica un ortólogo de IND1. En algunas realizaciones preferidas, la supresión de IND1 da lugar a una menor lignificación en las células del margen de las válvulas. Véase, también, U.S. 6.410.826.

Las secuencias *IND1* preparadas tal como se describe en la presente invención, se pueden utilizar para preparar cassettes de expresión útiles en una serie de técnicas, incluyendo la inhibición o supresión de la expresión. Se pueden utilizar una serie de procedimientos para inhibir la expresión génica en plantas. Por ejemplo, se puede utilizar de manera conveniente la tecnología antisentido. Para llevar a cabo esto, un segmento de ácido nucleico del gen deseado se clona y se une operativamente a un promotor, de manera que se transcribirá la cadena antisentido de ARN. A continuación, el cassette de expresión se transforma en plantas y se produce la cadena antisentido de ARN. En las células vegetales, se ha sugerido que el ARN antisentido inhibe la expresión génica mediante la prevención de la acumulación de ARNm que codifica la enzima de interés, véase, por ejemplo, Sheehy *et al.*, Proc. Nat. Acad. Sci. USA, 85: 8805-8809 (1988); Pnueli *et al.*, The Plant Cell 6: 175-186 (1994); y Hiatt *et al.*, U.S. Patent No. 4.801.340.

La secuencia de ácido nucleico antisentido transformada en las plantas será sustancialmente idéntica a por lo menos una parte del gen o genes endógenos a reprimir. Sin embargo, la secuencia no tiene que ser perfectamente idéntica para inhibir la expresión. De este modo, una molécula de ácido nucleico antisentido o de sentido que codifica sólo una parte de IND1 puede ser útil para producir una planta en la que se suprime la expresión de IND1. Los vectores de la presente invención se pueden diseñar de manera que el efecto inhibitorio se aplica a otras proteínas en una familia de genes que exhiben una homología o una homología sustancial con el gen diana.

Para la supresión antisentido, la secuencia introducida tampoco necesita que sea de longitud completa con respecto al producto de transcripción primaria o el ARNm totalmente procesado. En general, se puede utilizar una homología más elevada para compensar el uso de una secuencia más corta. Además, la secuencia introducida no necesita tener el mismo patrón de intrones o exones, y la homología de segmentos no codificantes puede ser igualmente efectiva. Normalmente, debería utilizarse una secuencia entre aproximadamente 30 ó 40 nucleótidos y aproximadamente nucleótidos de longitud completa, se prefiere mediante una secuencia de por lo menos aproximadamente 100 nucleótidos, es más preferido una secuencia de por lo menos aproximadamente 200 nucleótidos, y es especialmente preferido una secuencia de por lo menos aproximadamente 500 nucleótidos.

Las moléculas de ARN catalíticas o ribozimas también se pueden utilizar para inhibir la expresión de genes IND1. Es posible diseñar ribozimas que se emparejan específicamente con prácticamente cualquier ARN diana y dividen el esqueleto fosfodiéster en un punto específico, inactivando funcionalmente de esta manera el ARN diana. Al realizar esta división, el ribozima no se altera y, de este modo, es capaz de reciclarse y dividir otras moléculas, convirtiéndole en una verdadera enzima. La inclusión de secuencias de ribozimas en los ARN antisentido les confiere actividad de división de ARN, aumentando así la actividad de las construcciones.

ES 2 301 539 T3

Se han identificado una serie de tipos de ribozimas. Un tipo de ribozimas deriva de una serie de ARNs circulares pequeños que son capaces de autodividirse y replicarse en plantas. Los ARN se replican solos (ARNs tiroides) o con un virus auxiliar (ARNs satélites). Entre los ejemplos se incluyen ARNs del viroide de la mancha del sol de aguacate y los ARNs satélites de virus del mosaico anular del tabaco, virus del rayado transitorio de la alfalfa, virus del mosaico del tabaco de terciopelo, virus del mosaico del solanum nodiflorum, y virus del mosaico del trébol subterráneo. El diseño y uso de ribozimas específicos de ARN diana se describe en Haseloff *et al.*, Nature, 334: 585-591 (1988).

Otro método de supresión es la supresión de sentido (también conocido como cosupresión). Se ha observado que la introducción de cassettes de expresión en los que un ácido nucleico está configurado en la orientación de sentido con respecto al promotor es un medio eficaz por el cual se bloquea la transcripción de los genes diana. Para un ejemplo del uso de este método para modular la expresión de genes endógenos, véase Napoli *et al.*, The Plant Cell 2: 279-289 (1990); Flavell, Proc. Natl. Acad. Sci., USA 91: 3490-3496 (1994); Kooter y Mol, Current Opin. Biol. 4: 166-171 (1993); y las Patentes de Estados Unidos 5.034.323, 5.231.020 y 5.283.184.

En general, donde se desea la inhibición de la expresión, aparece cierta transcripción de la secuencia introducida. El efecto puede aparecer donde la secuencia introducida no contiene secuencia codificante *per se*, sino solamente secuencias de intrones o no traducidas homólogas a las secuencias presentes en el transcrito primario de la secuencia endógena. La secuencia introducida en general será sustancialmente idéntica a la secuencia endógena que se pretende reprimir. Esta identidad mínima será habitualmente mayor de aproximadamente un 65%, pero una mayor identidad podría ejercer una represión más eficaz de la expresión de las secuencias endógenas. Se prefiere una identidad sustancialmente mayor de más de aproximadamente un 80%, aunque lo más preferido sería aproximadamente de un 95% a la identidad absoluta. Con respecto a la regulación antisentido, el efecto debería aplicarse a cualquier otra proteína en una familia similar de genes que exhiben homología u homología sustancial.

Para la supresión de sentido, la secuencia introducida en el cassette de expresión, que necesita menos que la identidad absoluta, tampoco necesita ser de longitud completa, en relación con el producto de transcripción primario o el ARNm totalmente procesado. Esto puede ser preferible para evitar la producción simultánea de algunas plantas que son sobreexpresadoras. Una mayor identidad en una secuencia más corta que la longitud completa compensa una secuencia más larga pero menos idéntica. Además, la secuencia introducida no necesita tener el mismo patrón de intrones o exones y la identidad de segmentos no codificantes será igualmente eficaz. Normalmente, se utiliza una secuencia de los rangos de tamaño indicados anteriormente para la regulación antisentido.

En una realización preferida, la expresión de un ácido nucleico de interés se puede suprimir por la expresión simultánea de las construcciones tanto de sentido como antisentido (Waterhouse *et al.*, Proc. Natl. Acad. Sci. USA 95: 13959-13964 (1998). Véase también Tabara *et al.* Science 282: 430-431 (1998).

Un experto en la materia entenderá que utilizando la tecnología basada en secuencias de nucleótidos específicas (por ejemplo, tecnología de supresión antisentido o de sentido), las familias de genes homólogos se pueden suprimir con un único transcrito de sentido o antisentido. Por ejemplo, si un transcrito de sentido o antisentido está diseñado para tener una secuencia que se conserva entre una familia de genes, entonces se pueden suprimir múltiples miembros de una familia de genes. En cambio, si el objetivo es solamente suprimir un miembro de una familia de genes homólogos, entonces el transcrito de sentido o antisentido debería dirigirse a secuencias con la mayor variación posible entre miembros de la familia.

Otro medio para inhibir la función de *IND1* en una planta es mediante la creación de mutaciones negativas dominantes. En esta estrategia, se introducen en la planta polipéptidos *IND1* mutantes no funcionales, que retienen la capacidad de interactuar con subunidades naturales. También se puede utilizar una construcción negativa dominante para suprimir la expresión *IND1* en una planta. Una construcción negativa dominante útil en la presente invención contiene en general una parte de la secuencia codificante de *IND1* completa suficiente, por ejemplo, para la unión a ADN o para una interacción proteína-proteína, tal como una interacción proteína-proteína homodimérica o heterodimérica, pero que carece de la actividad transcripcional de la proteína de tipo natural. Por ejemplo, un mutante por delección carboxi terminal de *AGAMOUS* se utilizó como construcción negativa dominante para suprimir la expresión del gen *AGAMOUS* de la caja *MADS* (Mizukami *et al.*, Plant Cell 8:831.844 (1996)). Un experto en la materia entiende que, de manera similar, una construcción *IND1* negativa dominante se puede utilizar para suprimir la expresión de *IND1* en una planta.

B. Uso de ácidos nucleicos de la invención para aumentar la expresión génica

Las secuencias aisladas preparadas tal como se describe en la presente invención también se pueden utilizar para preparar cassettes de expresión que aumentan o incrementan la expresión del gen *IND1* endógeno. Cuando se desea la sobreexpresión de un gen, el gen deseado de una especie diferente se puede utilizar para disminuir los efectos potenciales de la supresión de sentido. La expresión aumentada de polinucleótidos de *IND1* es útil, por ejemplo, para producir plantas con fruto pequeño.

Se pueden utilizar cualquiera de los medios conocidos en la técnica para aumentar la actividad de *IND1* en plantas. Cualquier órgano puede estar marcado, tal como órganos/estructuras vegetativas de un brote (por ejemplo, hojas, tallos y tubérculos), raíces, flores, y órganos/estructuras florales (por ejemplo, brácteas, sépalos, pétalos, estambres, carpelos, anteras y óvulos), semilla (incluyendo embriones, endospermas y recubrimiento de semilla) y fruto. Alternativamente,

uno o varios de los genes *IND1* se pueden expresar de manera constitutiva (por ejemplo, utilizando el promotor CaMV 35S).

Un experto entenderá que los polipéptidos codificados por los genes de la presente invención, como otras proteínas, tienen dominios diferentes que realizan diferentes funciones. De este modo, las secuencias de genes no necesitan ser de longitud completa, siempre y cuando se exprese el dominio funcional deseado de la proteína. Tal como se ha explicado anteriormente, los polipéptidos *IND1* portan un dominio bHLH, que es capaz de unirse a ADN. De este modo, sin estar unido a ninguna teoría o mecanismo, *IND1* actúa probablemente como un modulador transcripcional.

10 C. Modificación de genes *IND1* endógenos

Los procedimientos para introducir mutaciones genéticas en genes de plantas y seleccionar plantas con las características deseadas son bien conocidos. Por ejemplo, las semillas u otro material vegetal se pueden tratar con una sustancia química mutagénica, según las técnicas estándar. Entre dichas sustancias químicas se incluyen, pero no se limitan a, las siguientes: sulfato de dietilo, etilen imina, metanosulfonato de etilo y N-nitroso-N-etilurea. Alternativamente, se puede usar también la radiación ionizante de fuentes tales como rayos X o rayos gamma.

Las cadenas de proteínas modificadas también se pueden diseñar fácilmente utilizando varias técnicas de ADN recombinante bien conocidas por los expertos en la materia y se describen, por ejemplo, en Sambrook *et al.*, *supra*. La hidroxilamina también se puede utilizar para introducir mutaciones de bases individuales en la región codificante del gen (Sikorski, *et al.* (1991) *Meth. Enzymol.* 194: 302-318). Por ejemplo, las cadenas pueden variar de la secuencia natural a nivel de estructura primaria mediante sustituciones, adiciones, deleciones de aminoácidos, y similares. Estas modificaciones se pueden utilizar en una serie de combinaciones para producir la cadena de proteína final modificada.

Alternativamente, se puede utilizar una recombinación homóloga para inducir las modificaciones en el gen marcado marcando específicamente el gen *IND1* *in vivo* (ver, en general, Grewal y Klar, *Genetics* 146: 1221-1238 (1997) y Xu *et al.*, *Genes Dev.* 10: 2411-2422 (1996)). Se ha demostrado la recombinación homóloga en plantas (Puchta *et al.*, *Experiencia* 50: 277-284 (1994), Swoboda *et al.*, *EMBO J.* 13: 484-489 (1994); Offringa *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 7346-7350 (1993); y Kempin *et al.* *Nature* 389: 802-803 (1997)).

Al aplicar la tecnología de recombinación homóloga a los genes de la presente invención, se realizan mutaciones *in vitro* en partes seleccionadas de secuencias de un gen *IND1* (incluyendo regiones en dirección 5', regiones en dirección 3' e intragénicas), tales como las descritas en la presente invención y, a continuación, se introducen en la planta deseada utilizando técnicas estándar. Debido a que se sabe que la recombinación homóloga es dependiente de los vectores utilizados, se utilizan vectores dicistrónicos dirigidos a genes tal como se describen en Mountford *et al.*, *Proc. Natl. Acad. Sci. USA* 91: 4303-4307 (1994); y Vaultont *et al.*, *Transgenic Res.* 4: 247-255 (1995) de manera conveniente para aumentar la eficacia de la selección de la expresión del gen *IND1* alterada en plantas transgénicas. El gen mutado interactuará con el gen natural diana de manera que la recombinación homóloga y la sustitución dirigida del gen natural tendrá lugar en células de plantas transgénicas, dando lugar a la supresión de la actividad de *IND1*.

Alternativamente, se pueden utilizar oligonucleótidos compuestos de un tramo contiguo de residuos de ARN y ADN en una conformación de doble cadena con cabezas de horquillas dobles en los extremos. La secuencia de ARN/ADN se diseña para alinearse con la secuencia del gen *IND1* diana y contener el cambio de nucleótido deseado. La introducción del oligonucleótido quimérico en un plásmido de ADN-T extracromosómico da lugar a una conversión eficaz y específica del gen *IND1* dirigida por moléculas quiméricas en un pequeño número de células vegetales transformadas. Este método se describe en Cole-Strauss *et al.*, *Science* 273: 1386-1389 (1996) y Yoon *et al.*, *Proc. Natl. Acad. Sci. USA* 93: 2071-2076 (1996).

En otras realizaciones, los promotores derivados de los genes *IND1* de la presente invención se pueden utilizar para dirigir la expresión de genes heterólogos de manera específica en el margen de las válvulas. Entre los genes estructurales que podrían utilizarse para este objetivo se incluyen genes que codifican proteínas citotóxicas tal como se describen a continuación.

Habitualmente, los promotores deseados se identifican mediante el análisis de las secuencias 5' de un clon genómico correspondiente a los genes *IND1* descritos en la presente invención. Las secuencias características de las secuencias de promotor se pueden utilizar para identificar el promotor. Las secuencias que controlan la expresión de genes eucariotas han sido ampliamente estudiadas. Por ejemplo, los elementos de la secuencia de promotor incluyen la secuencia de consenso de la caja TATA (TATAAT), que está habitualmente de 20 a 30 pares de bases en dirección 5' del punto de inicio de la transcripción. En la mayoría de los casos, se requiere la caja TATA para un inicio de la transcripción precisa. En las plantas, más adelante la dirección 5' de la caja TATA, en las posiciones -80 a -100, existe habitualmente un elemento promotor con una serie de adeninas que rodean el trinucleótido G (o T) N G. J. Messing *et al.*, in *GENETIC ENGINEERING IN PLANTS*, páginas 221-227 (Kosage, Meredith and Hollaender, eds. (1983)).

Existen una serie de métodos que son conocidos por los expertos en la materia para identificar y caracterizar las regiones de promotores en ADN genómico de planta (véase, por ejemplo, Jordano, *et al.*, *Plant Cell*, 1: 855-866 (1989); Bustos, *et al.*, *Plant Cell*, 1: 839-854 (1989); Green, *et al.*, *EMBO J.* 7, 4035-4044 (1988); Meier, *et al.*, *Plant Cell*, 3, 309-316 (1991); y Zhang, *et al.*, *Plant Physiology* 110: 1069-1079 (1996)).

IV. Preparación de vectores recombinantes

Para usar las secuencias aisladas en las técnicas anteriores, se preparan vectores de ADN recombinante adecuados para la transformación de células vegetales. Las técnicas para transformar una amplia variedad de especies de plantas superiores son bien conocidas y se describen en la literatura técnica y científica. Véase, por ejemplo, Weising *et al.* Ann Rev. Genet., 22: 421-477 (1988). Una secuencia codificante de ADN para el polipéptido deseado, por ejemplo, una secuencia de ADNc que codifica una proteína de longitud completa, se combinará preferiblemente con secuencias reguladoras de iniciación transcripcional y traduccional que dirigirán la transcripción de la secuencia del gen en los tejidos pretendidos de la planta transformada.

Por ejemplo, para la sobreexpresión, se puede utilizar un fragmento de promotor de planta que dirigirá la expresión del gen en todos los tejidos de una planta regenerada. A dichos promotores se hace referencia en la presente invención como promotores “constitutivos” y son activos en la mayoría de condiciones ambientales y estados de desarrollo o diferenciación celular. Entre los ejemplos de promotores constitutivos se incluyen la región de iniciación de la transcripción del virus del mosaico de la coliflor (CaMV) 35S, el promotor 1’ ó 2’ derivado de ADN-T de *Agrobacterium tumefaciens*, y otras regiones de iniciación de la transcripción de varios genes de planta conocidos por los expertos en la materia.

Alternativamente, el promotor de la planta puede dirigir la expresión del polinucleótido de la presente invención en un tejido específico (promotores específicos de tejido) o puede estar sinó bajo un control ambiental más preciso (promotores inducibles). Entre los ejemplos de promotores específicos de tejido bajo control de desarrollo se incluyen promotores que inician la transcripción sólo en ciertos tejidos, tales como fruto, semillas o flores. Tal como se ha indicado anteriormente, los promotores de los genes IND1 descritos en la presente invención son particularmente útiles para dirigir la expresión génica, de manera que un producto génico deseado se localiza en el margen de las válvulas de fruto. Otros promotores adecuados incluyen aquellos de genes tales como SHP1 o SHP2 (Savidge, B., Rounsley, S.D., y Yanofsky, M.F. (1995) Plant Cell 7: 721-733). Entre los ejemplos de condiciones ambientales que pueden afectar a la transcripción por promotores inducibles se incluyen condiciones anaeróbicas, temperatura elevada, o la presencia de luz.

Si se desea la correcta expresión del polipéptido, debería incluirse una región de poliadenilación en el extremo 3’ de la región codificante. La región de poliadenilación puede derivar del gen natural, de una variedad de otros genes de plantas o de ADN-T.

El vector que comprende las secuencias (por ejemplo, promotores o regiones codificantes) de genes de la presente invención comprenderá habitualmente un gen marcador que confiere un fenotipo seleccionable en células vegetales. Por ejemplo, el marcador puede codificar resistencia a biocida, particularmente resistencia a antibiótico, tal como resistencia a kanamicina, G418, bleomicina, higromicina o resistencia a herbicida, tal como resistencia a clorosulfuron o Basta.

Las secuencias de ácido nucleico *IND1* de la presente invención se expresan recombinantemente en células vegetales para aumentar e incrementar los niveles de polipéptidos IND1 endógenos. Alternativamente, las construcciones antisentido u otras construcciones IND1 (descritas anteriormente) se utilizan para suprimir los niveles de expresión de IND1. Se pueden preparar una serie de diferentes construcciones de expresión diferentes, tales como cassettes de expresión y vectores adecuados para la transformación de células vegetales. Las técnicas para transformar una amplia variedad de especies de plantas superiores son bien conocidas y se describen en la literatura técnica y científica. Véase, por ejemplo, Weising *et al.*, Ann. Rev. Genet. 22: 421-477 (1988). Una secuencia de ADN que codifica para un polipéptido IND1, por ejemplo una secuencia de ADNc que codifica una proteína de longitud completa, se puede combinar con secuencias reguladoras transcripcionales que actúan en cis (promotor) y actúan en trans (potenciador) para dirigir el transcurso, el tipo de tejido y los niveles de transcripción en los tejidos pretendidos de la planta transformada. También se pueden utilizar elementos de control traduccionales.

La presente invención proporciona un ácido nucleico IND1 unido operativamente a un promotor que, en una realización preferida, es capaz de dirigir la transcripción de la secuencia codificante de IND1 en plantas. El promotor puede derivar, por ejemplo, de fuentes vegetales o virales. El promotor puede ser, por ejemplo, constitutivamente activo, inducible o específico de tejido. En la construcción de cassettes de expresión, vectores, transgénicos recombinantes, de la presente invención, se pueden elegir y utilizar diferentes promotores para dirigir de manera diferencial de la expresión génica, por ejemplo, en algunos o todos los tejidos de una planta o animal. Habitualmente, tal como se ha descrito anteriormente, los promotores deseados se identifican analizando las secuencias 5’ de un clon genómico correspondiente a los genes IND1 descritos en la presente invención.

A. Promotores constitutivos

Se puede utilizar un fragmento de promotor que dirigirá la expresión del ácido nucleico IND1 en todas las células de tejido transformadas, por ejemplo, como las de una planta regenerada. El término “elemento regulador constitutivo” significa un elemento regulador que confiere un nivel de expresión sobre una molécula de ácido nucleico unida operativamente que es relativamente independientemente del tipo de célula o tejido en el que se expresa el elemento regulador constitutivo. Un elemento regulador constitutivo que se expresa en una planta generalmente se expresa ampliamente en un gran número de tipos de células y tejidos. A los promotores que dirigen la expresión de manera

ES 2 301 539 T3

continua en condiciones fisiológicas se hace referencia como promotores “constitutivos” y son activos bajo la mayoría de condiciones ambientales y estados de desarrollo o diferenciación celular.

5 En la técnica son conocidas una serie de elementos reguladores constitutivos útiles para la expresión ectópica en una planta transgénica. El promotor del virus del mosaico de la coliflor 35S (CaMV 35S), por ejemplo, es un elemento regulador constitutivo bien caracterizado que produce un nivel elevado de expresión en todos los tejidos de plantas (Odell *et al.*, Nature 313: 810-812 (1985)). El promotor de CaMV 35S puede ser particularmente útil debido a su actividad en numerosas especies de plantas diversas (Benfey y Chua, Science 250: 959-966 (1990); Futterer *et al.*, Physiol. Plant 79: 154 (1990); Odell *et al.*, *supra*, 1985). Un promotor de 35S en tandem, en el que el elemento
10 promotor intrínseco ha sido duplicado, confiere niveles de expresión más elevados en comparación con el promotor de 35S no modificado (Kay *et al.*, Science 236: 1299 (1987)). Entre otros elementos reguladores constitutivos útiles se incluyen, por ejemplo, el promotor de virus del mosaico de la coliflor 19S; el promotor del virus del mosaico de Figwort; y el promotor del gen de la nopalina sintasa (nos) (Singer *et al.*, Plant Mol. Biol. 14: 433 (1990); An, Plant Physiol. 81: 86 (1986)).

15 En la técnica se conocen elementos reguladores constitutivos adicionales que incluyen aquellos para la expresión eficaz en monocots, por ejemplo, el promotor pEmu y promotores basados en la región 5' Actin-1 del arroz (Last *et al.*, Theor. Appl. Genet. 81: 581 (1991); Mcelroy *et al.*, Mol. Gen. Genet. 231: 150 (1991); Mcelroy *et al.*, Plant Cell 2: 163 (1990)). Los elementos reguladores quiméricos, que combinan elementos de genes diferentes, también pueden ser
20 útiles para expresar de manera ectópica una molécula de ácido nucleico que codifica un polinucleótido IND1 (Comai *et al.*, Plant Mol. Biol. 15: 373 (1990)).

Otros ejemplos de promotores constitutivos incluyen el promotor 1' o 2' derivado de ADN-T de *Agrobacterium tumefaciens* (véase, por ejemplo, Mengiste (1997) *supra*; O'Grady (1995) Plant Mol. Biol. 29: 99-108); promotores de actina, tales como promotor del gen de actina de Arabidopsis (véase, por ejemplo, Huang (1997) Plant Mol. Biol. 1997
25 33: 125-139); promotores del gen de la alcohol deshidrogenasa (Adh) (véase, por ejemplo, Millar (1996) Plant Mol. Biol. 31: 897-904); *ACT11* de Arabidopsis (Huang *et al.* Plant Mol. Biol. 33: 125-139 (1996)), *Cat3* de Arabidopsis (Banco de Genes No. U43147, Zhong *et al.*, Mol. Gen. Genet. 251: 196-203 (1996)), el gen que codifica la proteína portadora esteroil-acilo desaturasa de *Brassica napus* (Banco de Genes No. X74782, Solcombe *et al.* Plant Physiol.
30 104: 1167-1176 (1994)), *Gpc1* de maíz (Banco de Genes No. X15596, Martinez *et al.* J. Mol. Biol. 208: 551-565 (1989)), *Gpc2* de maíz (Banco de Genes No. U45855, Manjunath *et al.*, Plant Mol. Biol. 33: 97-112 (1997)), otras regiones de iniciación de la transcripción de varios genes de plantas conocidos por los expertos en la materia. Véase también Holtorf Plant Mol. Biol. 29: 637-646 (1995).

35 B. Promotores inducibles

Alternativamente, un promotor de planta puede dirigir la expresión del ácido nucleico IND1 de la presente invención bajo la influencia del cambio en las condiciones ambientales o las condiciones de desarrollo. Entre los ejemplos de condiciones ambientales que pueden afectar a la transcripción mediante promotores inducibles se incluyen condi-
40 ciones anaeróbicas, temperatura elevada, SECuía, o la presencia de luz. A dichos promotores se les hace referencia en la presente invención como promotores “inducibles”. Por ejemplo, la presente invención incorpora el promotor inducible por la SECuía del maíz (Busk (1997) *supra*); el promotor inducible por el frío, la SECuía y la sal elevada de la patata (Kirch (1997) Plant Mol. Biol. 33: 897-909).

45 Alternativamente, los promotores de plantas que son inducibles tras la exposición a hormonas de plantas, tales como auxinas, se utilizan para expresar los ácidos nucleicos de la presente invención. Por ejemplo, la presente invención puede utilizar el fragmento de promotor de elementos de respuesta a auxina E1 (AuxREs) en la soja (*Glycine max* L.) (Liu (1997) Plant Physiol. 115: 397-407); el promotor GST6 de Arabidopsis sensible a auxina (también sensible a ácido salicílico y peróxido de hidrógeno) (Chen (1996) Plant J. 10: 955-966); el promotor parC inducible por auxina del
50 tabaco (Sakai (1996) 37: 906-913); un elemento de respuesta a la biotina de planta (Streit (1997) Mol. Plant Microbe Interact. 10: 933-937); y, el promotor sensible a la hormona del estrés, ácido abscísico (Sheen (1996) Science 274: 1900-1902).

Los promotores de plantas que son inducibles tras la exposición a reactivos químicos que se pueden aplicar a
55 la planta, tal como herbicidas o antibióticos, también se utilizan para expresar los ácidos nucleicos de la presente invención. Por ejemplo, se puede utilizar el promotor In2-2 del maíz, activado por el antídoto (“safener”) de herbicidas bencenosulfonamida, (De Veylder (1997) Plant Cell Physiol. 38: 568-577); la aplicación de diferentes antídotos de herbicidas induce diferentes patrones de expresión génica, incluyendo la expresión en la raíz, hidatodos, y el meristema apical del brote. La secuencia codificante de IND1 también puede estar bajo el control de, por ejemplo, un promotor
60 inducible por tetraciclina, por ejemplo, tal como se describe con plantas transgénicas del tabaco que contienen el gen de la arginina carboxilasa de la Avena sativa L. (avena) (Masgrau (1997) Plant J. 11: 465-473); o, un elemento sensible a ácido salicílico (Stange (1997) Plant J. 11: 1315-1324; Uknes *et al.*, Plant Cell 5: 159-169 (1993); Bi *et al.*, Plant J. 8: 235-245 (1995)).

65 Entre los elementos reguladores inducibles particularmente útiles se incluyen elementos reguladores inducibles por sobre (Mett *et al.*, Proc. Natl. Acad. Sci. USA 90: 4567-4571 (1993); Furst *et al.*, Cell 55: 705-717 (1988)); elementos reguladores inducibles por tetraciclina y cloro-tetraciclina (Gatz *et al.*, Plant J. 2: 397-404 (1992); Röder *et al.*, Mol. Gen. Genet. 243: 32-38 (1994); Gatz, Meth. Cell Biol. 50: 411-424 (1995)); elementos reguladores inducibles por

ecdisona (Christopherson *et al.*, Proc. Natl. Acad. Sci. USA 89: 6314-6318 (1992); Kreutzweiser *et al.*, Ecotoxicol. Environ. Safety 28: 14-24 (1994)); elementos reguladores inducibles por choque térmico (Takahashi *et al.*, Plant Physiol. 99: 383-390 (1992); Yabe *et al.*, Plant Cell Physiol. 35: 1207-1219 (1994); Ueda *et al.*, Mol. Gen. Genet. 250: 533-539 (1996)); y elementos del operón lac, que se utilizan en combinación con un represor lac expresado de manera constitutiva para conferir, por ejemplo, una expresión inducible por IPTG (Wilde *et al.*, EMBO J. 11: 1251-1259 (1992)). Un elemento regulador inducible útil en las plantas transgénicas de la presente invención también pueden ser, por ejemplo, un promotor inducible por nitrato derivado del gen de la nitrito reductasa de la espinaca (Back *et al.*, Plant Mol. Biol. 17: 9 (1991)) o un promotor inducible por la luz, tal como el asociado con la pequeña subunidad de la RuBP carboxilasa o las familias de genes LHCP (Feinbaum *et al.*, Mol. Gen. Genet. 226: 449 (1991); Lam y Chua, Science 248: 471 (1990)).

C. Promotores específicos de tejido

Alternativamente, el promotor de planta puede dirigir la expresión del polinucleótido de la presente invención en un tejido específico (promotores específicos de tejido). Los promotores específicos de tejido son elementos de control transcripcional que sólo son activos en células o tejidos concretos en momentos específicos durante el desarrollo de la planta, tal como en tejidos vegetativos o tejidos reproductores. Los promotores de los genes *IND1* de la presente invención son particularmente útiles para la dirección específica de tejido de la expresión génica, de manera que un producto génico deseado se genera sólo o preferencialmente en embriones o semillas, tal como se describe a continuación.

Entre los ejemplos de promotores específicos de tejido bajo el control del desarrollo se incluyen promotores que inician la transcripción sólo (o principalmente sólo) e ciertos tejidos, tales como tejidos vegetativos, por ejemplo, raíces u hojas, o tejidos reproductores, tales como fruto, óvulos, semillas, polen, pistilos, flores, o cualquier tejido embrionario. Los promotores reproductores específicos de tejido pueden ser, por ejemplo, específicos de óvulo, específicos de embrión, específicos de endosperma, específicos de integumento, específicos de semilla y cubierta de semilla, específicos de polen, específicos de pétalo, específicos de sépalo, o alguna combinación de los mismos.

La presente invención proporciona una planta transgénica que se caracteriza por una dispersión de semillas retardada debido a la expresión de una molécula de ácido nucleico que codifica un producto génico *IND1*, o una construcción antisentido del mismo, unido operativamente a un elemento regulador selectivo de la zona de dehiscencia. El elemento regulador selectivo de la zona de dehiscencia puede ser, por ejemplo, un elemento regulador *SHP1* o un elemento regulador *SHP2*. El elemento regulador *SHP1* puede derivar de la secuencia genómica *SHP1* de Arabidopsis descrita en la presente invención como SEC ID No: 5 y puede ser, por ejemplo, una secuencia reguladora 5' o un elemento regulador intrónico. De manera similar, el elemento regulador *SHP2* puede derivar de la secuencia genómica *SHP2* de Arabidopsis descrita en la presente invención como SEC ID No: 6 y puede ser, por ejemplo, una secuencia reguladora 5' o elemento regulador intrónico.

Un elemento regulador selectivo de la zona de dehiscencia puede derivar de un gen que es un ortólogo de *IND1* de Arabidopsis y se expresa selectivamente en el margen de las válvulas o la zona de dehiscencia de una planta con semilla. Un elemento regulador selectivo de la zona de dehiscencia puede derivar, por ejemplo, de un ortólogo de *IND1* de *Brassicaceae*, tal como ortólogo *IND1* de *Brassica napus*, *Brassica oleracea*, *Brassica campestris*, *Brassica juncea*, *Brassica nigra* o *Brassica carinata*. Un elemento regulador selectivo de la zona de dehiscencia puede derivar, por ejemplo, de un ortólogo de *IND1* de cánola. Un elemento regulador selectivo de la zona de dehiscencia también puede derivar, por ejemplo, de un ortólogo de *IND1* de leguminosa, tal como soja, guisante, garbanzo, aconitifolia, habas, alubia de riñón, alubia lima, lenteja, frijol de vara, habas secas, cacahuete, alfalfa, forraje, lotus corniculatus, trébol, stylosanthes, lotononis bainessii, u ortólogo *IND1* de sainfoin.

Un elemento regulador selectivo de la zona de dehiscencias también puede derivar de una variedad de otros genes que se expresan selectivamente en el margen de las válvulas o la zona de dehiscencia de una planta con semillas. Por ejemplo, el gen *RDPG1* de la semilla de colza se expresa selectivamente en la zona de dehiscencia (Petersen *et al.*, Plant Mol. Biol. 31: 517-527 (1996)). De este modo, el promotor *RDPG1* o un fragmento activo del mismo puede ser un elemento regulador selectivo de la zona de dehiscencia tal como se define en la presente invención. También se sabe que genes adicionales tales como el gen *SAC51* de la semilla de colza se expresan selectivamente en la zona de dehiscencia; el promotor *SAC51* o un fragmento activo del mismo también puede ser un elemento regulador selectivo de la zona de dehiscencia de la presente invención (Coupe *et al.*, Plant Mol. Biol. 23: 1223-1232 (1993)). El técnico en la materia entiende que un elemento regulador de cualquiera de los genes expresados selectivamente en células del margen de las válvulas o zona de dehiscencia puede ser un elemento regulador selectivo de la zona de dehiscencia tal como se define en la presente invención.

Los elementos reguladores selectivos de la zona de dehiscencia adicionales se pueden identificar y aislar utilizando metodología de rutina. Las estrategias de cribado diferencial utilizando, por ejemplo, ARN preparado de la zona de dehiscencia y ARN preparado de material de la vaina adyacente se puede utilizar para aislar ADNc expresado selectivamente en células de la zona de dehiscencia (Coupe *et al.*, *supra*, 1993); posteriormente, los genes correspondientes se aíslan utilizando la secuencia de ADNc como sonda.

Las estrategias de atrapamiento de potenciador o atrapamiento de genes también se pueden utilizar para identificar y aislar un elemento regulador selectivo de la zona de dehiscencia de la presente invención (Sundaresan *et al.*, *supra*,

1995; Koncz *et al.*, Proc. Natl. Acad. Sci. USA 86: 8467-8471 (1989); Kertbundit *et al.*, Proc. Natl. Acad. Sci. USA 88: 5212-5216 (1991); Topping *et al.*, Development 112: 1009-1019 (1991)). Los elementos de atrapamiento del potenciador incluyen un gen informador tal como GUS con un promotor débil o mínimo, mientras que los elementos de atrapamiento de genes carecen de una secuencia de promotor, dependiendo de la transcripción de un gen cromosómico
 5 flanqueante para la expresión de un gen informador. Los elementos transposables incluidos en las construcciones median las fusiones a loci endógenos; las construcciones expresadas selectivamente en el margen de las válvulas o la zona de dehiscencia se identifican por su patrón de expresión. Con el elemento insertado como una etiqueta, el elemento flanqueante regulador selectivo de la zona de dehiscencia se clona utilizando, por ejemplo, la metodología de reacción en cadena de la polimerasa inversa (véase, por ejemplo, Aarts *et al.*, Nature 363: 715-717 (1993); véase
 10 también, Ochman *et al.*, "Amplification of Flanking SEquences by Inverse PCR", in Innis *et al.*, *supra*, 1990). El sistema de transposición Ac/Ds de Sundaresan *et al.*, Genes. Devel. 9: 1797-1810 (1995), puede ser particularmente útil en la identificación y aislamiento de un elemento regulador selectivo de la zona de dehiscencia de la presente invención.

15 Los elementos reguladores selectivos de la zona de dehiscencia también se pueden aislar mediante la inserción de una biblioteca de fragmentos aleatorios de ADN genómico delante de un gen informador sin promotor y el cribado de plantas transgénicas transformadas por la biblioteca para la expresión del gen informador selectivo de la zona de dehiscencia. El vector sin promotor pROA97, que contiene el gen npt y el gen GUS, cada uno bajo el control el promotor de 35S mínimo, puede ser útil para dicho cribado. La biblioteca genómica puede ser, por ejemplo, fragmentos
 20 Sau3A de ADN genómico de *Arabidopsis thaliana* o ADN genómico de, por ejemplo, otra Brassicaceae de interés (Ott *et al.*, Mol. Gen. Genet. 223: 169-179 (1990); Claes *et al.*, The Plant Journal 1: 15-26 (1991)).

La expresión selectiva de la zona de dehiscencia de un elemento regulador de la presente invención puede demostrarse o confirmarse mediante técnicas rutinarias, por ejemplo, utilizando un gen informador y un análisis de expresión
 25 *in situ*. Los informadores GUS y luciferasa de luciérnaga son particularmente útiles para la localización *in situ* de la expresión génica de la planta (Jeferson *et al.*, EMBO J. 6: 3901 (1987); Ow *et al.*, Science 334: 856 (1986)), y los vectores sin promotor que contienen el cassette de expresión GUS están disponibles comercialmente, por ejemplo, en Clontech (Palo Alto, CA). Para identificar un elemento regulador selectivo de la zona de dehiscencia de interés tal como un elemento regulador *IND1*, se pueden generar una o más partes de nucleótidos del gen *IND1* utilizando
 30 metodología enzimática o basada en PCR (Glick y Thompson, *supra*, 1993; Innis *et al.*, *supra*, 1990); los segmentos resultantes se fusionan a un gen informador, tal como GUS, y se analizan tal como se ha descrito anteriormente.

Otros promotores específicos de tejido incluyen promotores de semillas. Los promotores específicos de semilla adecuados derivan de los siguientes genes: *MAC1* de maíz (Sheridan (1996) Genetics 142: 1009-1020); *Cat3* de maíz
 35 (Banco de Genes No. L05934, Abler (1993) Plant Mol. Biol. 22: 10131-1038); *viviparous-1* de *Arabidopsis* (Banco de Genes No. U93215); *atmyc1* de *Arabidopsis* (Urao (1996) Plant Mol. Biol. 32: 571-57; Conceicao (1994) Plant 5: 493-505); *napA* de *Brassica napus* (Banco de Genes No. J02798, Josefsson (1987) JBL 26: 12196-1301); y la familia de genes napin de *Brassica napus* (Sjodahl (1995) Planta 197: 264-271).

40 También se pueden utilizar una serie de promotores específicamente activos en tejidos vegetativos, tales como hojas, tallos, raíces y tubérculos, para expresar los ácidos nucleicos *IND1* de la presente invención. Por ejemplo, se pueden utilizar los promotores que controlan la patatina, la principal proteína de almacenamiento del tubérculo de la patata, véase, por ejemplo, Kim (1994) Plant Mol. Biol. 26: 603-615; Martin (1997) Plant J. 11: 53-62. También se puede utilizar el promotor ORF13 de *Agrobacterium rhizogenes* que muestra una gran actividad en raíces (Hansen
 45 (1997) Mol. Gen. Genet. 254: 337-343. Otros promotores vegetativos útiles específicos de tejido incluyen: el promotor tarin del gen que codifica una globulina de una familia principal de proteínas del bulbo del taro (Colocasia esculenta L. Schott), tarin (Bezerra (1995) Plant Mol. Biol. 28: 137-144); el promotor curculin activo durante el desarrollo del bulbo del taro (de Castro (1992) Plant Cell 4: 1549-1559) y el promotor para el gen TobRB7 específico de la raíz del tabaco, cuya expresión se localiza en el meristema de la raíz y las regiones inmaduras del cilindro central (Yamamoto
 50 (1991) Plant Cell 3: 371-382).

Se pueden utilizar promotores específicos de hojas, tales como promotores de ribulosa bifosfato carboxilasa (RBCS). Por ejemplo, los genes del tomate RBCS1, RBCS2 y RBCS3A se expresan en hojas y plantas de semi-
 55 lleros que crecen con la luz, sólo RBCS 1 y RBCS2 se expresan en frutos tomate en desarrollo (Meier (1997) FEBS Lett. 415: 91-95). Se pueden utilizar promotores de ribulosa bisfosfato carboxilasa expresados casi exclusivamente en células mesofílicas en láminas foliares y las vainas de las hojas a niveles elevados, descritos por Matsuoka (1994) Plant J. 6: 311-319. Otro promotor específico de hoja es el promotor del gen de la proteína de unión a la clorofila a/b fotorreceptora, véase, por ejemplo, Shiina (1997) Plant Physiol. 115: 477-483; Casal (1998) Plant Physiol. 116: 1533-1538. El promotor del gen relacionado con myb de *Arabidopsis thaliana* (*Atmyb5*) descrito por Li (1996) FEBS Lett.
 60 379: 117-121, es específico de hoja. El promotor *Atmyb5* se expresa en tricomas, estípulas, y células epidérmicas de hojas en desarrollo en los márgenes de hojas de rosetas jóvenes y hojas caulinares, y en semillas inmaduras. El ARNm de *Atmyb5* aparece entre la fertilización y la etapa de 16 células del desarrollo embrionario y persiste más allá de la etapa de corazón. También se puede utilizar un promotor de hoja identificado en el maíz por Busk (1997) Plant J. 11: 1285-1295.

65 Otra clase de promotores vegetativos útiles específicos de tejido son promotores meristemáticos (root punta de raíz y brote apical). Por ejemplo, se pueden utilizar los promotores "SHOOTMERISTEMLESS" y "SCARECROW", que son activos en el brote en desarrollo o los meristemas apicales de la raíz, descritos por Di Laurenzio (1996) Cell 86:

423-433; y, Long (1996) *Nature* 379: 66-69. Otro promotor útil es aquel que controla la expresión del gen de la 3-hidroxi-3-metilglutaril coenzima A reductasa HMG2, cuya expresión está limitada a tejidos meristemáticos y florales (zona secretora del estigma, granos de polen maduros, tejido vascular del gineceo, y óvulos fertilizados) (véase, por ejemplo, Enjuto (1995) *Plant Cell*. 7: 517-527). También son útiles los genes relacionados con *kn1* del maíz y otras especies que muestran una expresión específica en el meristema, véase, por ejemplo, Granger (1996) *Plant Mol. Biol.* 31: 373-378; Kerstetter (1994) *Plant Cell* 6: 1877-1887; Hake (1995) *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 350: 45-51. Por ejemplo, el promotor KNAT1 de *Arabidopsis thaliana*. En el brote apical, el transcrito KNAT1 se localiza principalmente en el meristema apical del brote; la expresión de KNAT1 en el meristema del brote disminuye durante la transición floral y está limitada al córtex del tallo con inflorescencia (véase, por ejemplo, Lincoln (1994) *Plant Cell* 6: 1859-1876).

Un experto en la material entenderá que un promotor específico de tejido puede provocar la expresión de secuencias unidas operativamente en tejidos diferentes del tejido diana. De este modo, tal como se utiliza en la presente invención un promotor específico de tejido es aquel que provoca la expresión preferencialmente en el tejido diana, pero también conduce a cierta expresión en otros tejidos.

En otra realización, un ácido nucleico *IND1* se expresa a través de un elemento trasposable. Esto permite la expresión constitutiva, aunque periódica e infrecuente, del polipéptido constitutivamente activo. La presente invención también se proporciona para el uso de promotores específicos de tejido derivados de virus que pueden incluir, por ejemplo, el promotor subgenómico de tobamovirus (Kumagai (1995) *Proc. Natl. Acad. Sci. USA* 92: 1679-1683; el virus baciliforme tungro del arroz (RTBV), que se replica sólo en células de floema en plantas de arroz infectadas, con su promotor que provoca la expresión fuerte del gen informador específico de floema; el promotor del virus de mosaico de la vena de la mandioca (CVMV), con la mayor actividad en los elementos vasculares, en células mesofílicas de la hoja, y en puntas de raíz (Verdaguer (1996) *Plant Mol. Biol.* 31: 1129-1139).

V. Producción de plantas transgénicas

Las construcciones de ADN de la invención pueden introducirse en el genoma de la planta huésped deseada mediante una serie de técnicas convencionales. Por ejemplo, puede introducirse la construcción de ADN directamente en el ADN genómico de la célula vegetal utilizando técnicas tales como la electroporación y la microinyección de protoplastos de célula vegetal, o pueden introducirse las construcciones de ADN directamente en el tejido vegetal utilizando métodos balísticos, tales como el bombardeo de partículas de ADN. Alternativamente, las construcciones de ADN pueden combinarse con regiones flanqueantes adecuadas de ADN-T e introducirse en un vector huésped de *Agrobacterium tumefaciens* convencional. Las funciones virulentas del huésped de *Agrobacterium tumefaciens* dirigirán la inserción de la construcción y el marcador contiguo en el ADN de la célula vegetal cuando se infecte la célula por las bacterias.

Las técnicas de microinyección son conocidas en la técnica y están bien descritas en la literatura científica y de patentes. La introducción de construcciones de ADN utilizando la precipitación de polietilenglicol está descrita en Paszkowski *et al* *Embo J.* 3:2717-2722 (1984). Las técnicas de electroporación se describen en Fromm *et al.* *Proc. Natl. Acad. Sci. USA* 82:5824 (1985). Las técnicas de transformación balística se describen en Klein *et al.* *Nature* 327:70-73 (1987).

Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, incluyendo el desarme y el uso de vectores binarios, se han descrito en la literatura científica. Ver, por ejemplo, Horsch *et al.* *Science* 233:496-498 (1984), y Fraley *et al.* *Proc. Natl. Acad. Sci. USA* 80:4803 (1983).

Las células vegetales transformadas que se derivan mediante cualquiera de las técnicas de transformación citadas anteriormente pueden cultivarse para regenerar una planta completa que posea el genotipo transformado y de este modo el fenotipo deseado tal como la falta de semillas. Dichas técnicas de regeneración dependen de la manipulación de ciertas fitohormonas en un medio de crecimiento de cultivo tisular, dependiendo habitualmente de un marcador biocida y/o herbicida que ha sido introducido junto con las secuencias de nucleótidos deseadas. La regeneración de plantas a partir de protoplastos cultivados se describe en Evans *et al.*, *Protoplasts Isolation and Culture*, Handbook of Plant Cell Culture, pág. 124-176, MacMillan Publishing Company, New York, 1983; y Binding, *Regeneration of Plants, Plant Protoplasts*, pág. 21-73, CRC Press, Boca Raton, 1985. La regeneración también puede obtenerse a partir de callo de planta, explantes, órganos o partes de la misma. Tales técnicas de regeneración se describen de forma general en Klee *et al.* *Ann. Rev. of Plant Phys.* 38:467-486 (1987).

Los ácidos nucleicos de la invención pueden utilizarse para conferir los rasgos deseados a esencialmente cualquier planta. De este modo, la invención puede usarse sobre un amplio rango de plantas, incluidas las especies de los géneros *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Cucumis*, *Cucurbita*, *Daucus*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Oryza*, *Panicum*, *Pannisetum*, *Persea*, *Pisum*, *Pyrus*, *Prunus*, *Raphanus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Trigonella*, *Triticum*, *Vitis*, *Vigna*, y, *Zea*. Una planta de semilla dehiscente puede ser una planta útil en la invención, y puede ser una planta particularmente útil un miembro de las *Brassicaceae*, como por ejemplo la semilla de colza, o un miembro de las *Fabaceae*, como por ejemplo una planta de soja, guisante, lenteja o alubia.

En una realización, la invención proporciona una planta de semilla dehiscente que se caracteriza por una dispersión de semilla retardada debida a la expresión suprimida de una molécula de ácido nucleico que codifica un producto génico IND1 en la planta de semilla dehiscente. Tal y como se utiliza en la presente invención, el término “planta de semilla dehiscente” significa una planta que produce un fruto dehiscente seco, que tiene paredes en el fruto que se abren para permitir la salida de las semillas contenidas en el mismo. Los frutos dehiscentes contienen normalmente varias semillas e incluyen los frutos conocidos, como por ejemplo, las legumbres, las cápsulas y las silicuas.

En una realización, la invención proporciona una planta que se caracteriza por una dispersión de semilla retardada debida a la expresión suprimida de una molécula de ácido nucleico que codifica un producto génico IND1, en la que la planta es un miembro de las *Brassicaceae*. Las *Brassicaceae*, comúnmente conocidas como las Brassicas, son un grupo diverso de plantas de cultivo con gran valor económico en el mundo entero (ver, por ejemplo, Williams y Hill, Science 232:1385-1389 (1986)). Las *Brassicaceae* producen aceites de semilla para la margarina, aceite para ensaladas, aceite para cocinar, usos plásticos e industriales; mostaza de condimento; verduras de hojas, almacenadas, elaboradas y en vinagre; forraje de animales y abonos verdes para el rejuvenecimiento de la tierra. La planta de la colza oleaginosa es una planta Brassica de la invención particularmente útil que no se encuentra de forma natural.

Hay seis especies de Brassica principales de importancia económica, cada una conteniendo un rango de formas de planta. Las *Brassica napus* incluyen plantas como la colza oleaginosa y el colinabo. Las *Brassica oleracea* son los cultivos de coles como por ejemplo la col, la coliflor, la col verde, el colirrábano, y las coles de Bruselas. La *Brassica campestris* (*Bressica rapa*) incluye: plantas como la col china, el nabo, y el pak choi. La *Brassica juncea* incluye una variedad de mostazas; la *Brassica nigra* es la mostaza negra; y la *Brassica carinata* es la mostaza etíope. El experto en la materia entiende que cualquier miembro de las *Brassicaceae* puede modificarse tal y como se describe en la presente invención para producir una planta Brassica que no se encuentra de forma natural caracterizada por una dispersión de semilla retardada.

En una segunda realización, la invención proporciona una planta que se caracteriza por una dispersión de semilla retardada debida a la expresión suprimida de una molécula de ácido nucleico que codifica un producto génico IND1, en la que la planta es un miembro de las *Fabaceae*. Las *Fabaceae*, que son comúnmente conocidas como miembros de la familia del guisante, son plantas que producen un fruto dehiscente seco característico conocido como una legumbre. La legumbre deriva de un único carpelo y se dehiscen a lo largo de la sutura de los márgenes del carpelo y a lo largo de la vena mediana. Las *Fabaceae* comprenden tanto las legumbres de grano como las legumbres de pasto. Las legumbres de grano incluyen, por ejemplo, la soja (glicina), el guisante, el garbanzo, aconitifolia, haba, alubia de riñón, alubia lima, lenteja, frijol de vara, habas secas y cacahuete. Las legumbres de pasto incluyen alfalfa, forraje, lotus cornilatus, el trébol, las especies del género *Stylosanthes*, el *lotoononis bainessii* y sainfoin. El experto en la técnica entenderá que cualquier miembro de las *Fabaceae* puede modificarse tal y como se explica en la presente invención para producir una planta de la invención que no se encuentra de forma natural caracterizada por una dispersión de semilla retardada.

Una planta de la invención que no se encuentra de forma natural caracterizada por una dispersión de semilla retardada también puede ser un miembro del género de plantas *Cuphea* (familia *Lythraceae*). Una planta *Cuphea* es particularmente válida ya que las semillas oleaginosas de la *Cuphea* contienen ácidos grasos de cadena media importantes industrial y nutricionalmente, especialmente el ácido láurico, que es suministrado actualmente sólo mediante aceites de coco y de nuez de palma.

Una planta de la invención que no se encuentra de forma natural también puede ser, por ejemplo, una de las hierbas monocotiledóneas, las cuales producen muchos de los cultivos válidos de cereal de grano pequeño del mundo. La supresión de la expresión de IND1 tal y como se ha descrito anteriormente, puede ser útil en la generación de una nueva planta de cereal de grano pequeño que no se encuentre de forma natural, como por ejemplo las plantas de la cebada, el trigo, la avena, el centeno, el dátilo, la hierba de guinea, el sorgo o la hierba de césped caracterizadas por la dispersión de semilla retardada.

VI. Modificaciones adicionales que modulan la dispersión de la semilla

Debe entenderse que una planta de la invención, que contiene un polinucleótido *IND1* exógeno, también puede contener una o más modificaciones adicionales, incluyendo las modificaciones naturales o no naturales, que pueden modular el retraso en la dispersión de la semilla. Por ejemplo, la hormona de la planta etileno fomenta la dehiscencia del fruto, y la expresión o la actividad modificada de reguladores positivos o negativos de la respuesta de etileno pueden incluirse en una planta de la invención (ver, de forma general, Meakin y Roberts, J. Exp. Botany 41:1003-1011 (1990); Ecker, Science 268:667-675 (1995); Chao *et al.*, Cell 89:1133-1144 (1997)).

Las mutaciones en reguladores positivos de la respuesta de etileno muestran una reducción o ausencia de la sensibilidad al tratamiento con etileno exógeno. Las mutaciones de *Arabidopsis* en reguladores positivos de la respuesta de etileno incluyen mutaciones en *etr*, que inactivan un receptor de etileno histidina quinasa (Bleeker *et al.*, Science 241:1086-1089 (1988); Schaller y Bleeker, Science 270:1809-1811 (1995)); *ers* (Hua y otros, Science 269:1712-1714 (1995)); *ein2* (Guzman y Ecker, Plant Cell 2:513 (1990)); *ein3* (Rothenberg y Ecker, Sem. Dev. Biol. Plant Dev. Genet. 4:3-13 (1993); Kieber y Ecker, Trends Genet. 9:356-362 (1993)); *ain1* (van der Straeten *et al.*, Plant Physiol. 102:401-408 (1993)); *eti* (Harpham *et al.*, An. Bot. 68:55 (1991)) y *ein4*, *ein5*, *ein6*, y *ein7* (Roman *et al.*, Genetics 139: 1393-1409 (1995)). En otras especies de plantas se encuentran funciones genéticas similares; por ejemplo, la mutación nunca-madura corresponde a *etr* y confiere insensibilidad al etileno en el tomate (Lanahan *et al.*, The Plant Cell 6:521-530

(1994); Wilkinson *et al.*, Science 270:1807-1809 (1995)). Una planta de la invención puede incluir una modificación que da lugar a una expresión o actividad alterada de cualquiera de los reguladores positivos de la respuesta del etileno. Puede incluirse, por ejemplo, una mutación en un regulador positivo en una planta de la invención y puede modificar el retraso en la dispersión de la semilla en dichas plantas, por ejemplo, mediante el aplazamiento adicional del retraso en la dispersión de la semilla.

Las mutaciones en reguladores negativos de la respuesta del etileno manifiestan sensibilidad del etileno en ausencia de etileno exógeno. Dichas mutaciones incluyen aquellas relativas a la sobreproducción del etileno, por ejemplo, los mutantes *eto1*, *eto2*, y *eto3*, y aquellas relativas a la activación constitutiva del mecanismo de señalización del etileno, por ejemplo, mutaciones en *CTR1*, un regulador negativo con similitud de secuencias con la familia Raf de proteína quinasas (Kieber *et al.*, Cell 72: 427-441 (1993)).

Una planta de la invención puede incluir una modificación que da lugar a la expresión o la actividad alterada de cualquiera de los reguladores negativos de la respuesta del etileno. Puede incluirse, por ejemplo, una mutación que da lugar a una sensibilidad del etileno en ausencia de etileno exógeno en una planta de la invención que no se encuentra de forma natural y puede modificar, por ejemplo, disminuir, el retraso en la dispersión de la semilla.

También puede incluirse mutaciones morfológicas del fruto en una planta de la invención. Tales mutaciones incluyen aquellas en los genes de identidad del carpelo como por ejemplo *AGAMOUS* (Bowman *et al.*, ver referencia anterior, 1989; Yanofsky *et al.*, ver referencia anterior, 1990) y en genes requeridos para el desarrollo normal del fruto como por ejemplo *ETTIN*, *CRABS CLAW*, *SPATULA*, *AGL8* y *TOUSLED* (Sessions *et al.*, Development 121:1519-1532 (1995); Alvarez y Smyth, Flowering Newsletter 23:12-17 (1997); y Roe *et al.*, Cell 75:939-950 (1993)). De este modo, se entiende que una planta de la invención puede incluir una o más modificaciones genéticas adicionales, las cuales pueden disminuir o aumentar el retraso en la dispersión de la semilla.

VII. Expresión de productos génicos citotóxicos

La presente invención también proporciona una molécula de ácido nucleico recombinante que incluye un elemento regulador selectivo de la zona dehiscente unido operativamente a una molécula de ácido nucleico que codifica un producto génico citotóxico. Adicionalmente en la presente invención se proporciona una planta de la invención que se caracteriza por una dispersión de la semilla retardada debida a la expresión de una molécula de ácido nucleico recombinante que tiene un elemento regulador selectivo de la zona dehiscente unido operativamente a una molécula de ácido nucleico que codifica un producto génico citotóxico.

Un producto génico citotóxico es un producto génico que causa la muerte de la célula en la cual se expresa y, preferiblemente, no da lugar a la muerte de células diferentes de la célula en la que se ha expresado. De este modo, la expresión de un producto génico citotóxico a partir de un elemento regulador selectivo de la zona dehiscente puede utilizarse para extirpar la zona dehiscente sin células vecinas perturbantes del replum o la válvula. En la técnica se conoce una variedad de productos génicos citotóxicos útiles en plantas de semillas incluyendo, por ejemplo, los polipéptidos de cadena A de la toxina de la difteria; RNasa T1; Barnasa RNasa; polipéptidos de cadena A de la toxina de la ricina; y productos génicos de la timidina quinasa (tk) del virus de herpes simplex. A pesar de que la cadena A de la toxina de la difteria, la RNasa T1 y la Barnasa RNasa son productos génicos citotóxicos preferentes, el experto en la materia entiende que estos, u otros productos génicos citotóxicos pueden utilizarse con un elemento regulador selectivo de la zona dehiscente para generar una planta de semilla que no se encuentra de forma natural caracterizada por la dispersión de la semilla retardada.

La toxina de la difteria es la toxina natural de la *Comebacterium diphtariae*, la cual cataliza la ADP-ribosilación del alargamiento del factor 2, dando lugar a la inhibición de la síntesis proteica y la consecuente muerte celular (Collier, Bacteriol. Rev. 39:54-85 (1975)). Una sola molécula de la toxina completamente activa es suficiente para matar una célula (Yamaizumi *et al.*, Cell 15:245-250 (1978)). La toxina de la difteria tiene dos subunidades: la cadena B de la toxina de la difteria dirige la internalización hacia la mayoría de células eucariotas a través de un receptor de membrana específico, mientras que la cadena A codifica el dominio catalítico tóxico. La cadena DT-A catalítica no incluye un péptido señal y no se secreta. Adicionalmente, cualquier DT-A liberado por células muertas en ausencia de la cadena B de la toxina de la difteria se excluye de la unión celular. De esta manera, la DT-A es autónoma respecto la célula y dirige la citólisis sólo de las células en las que se expresa sin daño aparente hacia las células vecinas. El cassette de expresión de DT-A de Palmiter *et al.*, que contiene los 193 residuos de la cadena A diseñada con un ATG sintético y con una carencia de la secuencia líder nativa, es particularmente útil en las plantas de semilla de la invención (Palmiter *et al.*, Cell 50:435-443 (1987); Greenfield *et al.*, Proc. Natl. Acad. Sci., USA 80: 6853-6857 (1983)).

La RNasa T1 de *Aspergillus oryzae* y la Barnasa RNasa de *Bacillus amylolique-faciens* también son productos génicos citotóxicos útiles en las plantas de semillas de la invención (Thorsness and Nasrallah, Methods in Cell Biology 50:439-448 (1995)). La Barnasa RNasa generalmente puede ser más tóxica para las plantas que la RNasa T1 y, de esta manera se la prefiere en los procedimientos de la invención.

La ricina, una proteína desactivadora de ribosoma producida por semillas de ricino, también es un producto génico citotóxico útil en una planta de semilla de la invención no natural. El polipéptido de la cadena A de la toxina de la ricina puede utilizarse para dirigir la ablación específica de la célula tal y como se describe, por ejemplo en Moffat *et al.*, Development 114:681-687 (1992). Los ribosomas de la planta son variablemente susceptibles a la toxina de la

ricina derivada de la planta. El experto entiende que la toxicidad de la ricina depende de si es variable y debe valorarse su toxicidad en las especies de la planta de semilla de interés (ver Olsnes y Phil, Molecular Action of Toxins and Viruses, págs. 51-105, Amsterdam: Elsevier Biomedical Press (1982)).

5 Ejemplos

Los siguientes ejemplos se ofrecen para ilustrar, pero no para limitar la invención reivindicada.

Ejemplo 1

10

El marcador del margen de la válvula GT140 (Sundaresan, V., *et al.* Genes Dev. 9, 1797-1810 (1995)) se expresa en el margen de la válvula del gineceo en desarrollo justo antes de la fertilización (etapa 13) y este patrón persiste en el fruto maduro (etapa 17). Como expresión de que este marcador está en gran parte ausente de los márgenes de la válvula de los frutos indehiscientes *shp1 shp2* (Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. Nature, en la prensa), se esperaba que el gen correspondiente a este marcador también podría estar involucrado en el desarrollo del margen de la válvula y ser requerido para la dehiscencia del fruto.

15

Para aislar la secuencia genómica flanqueante del sitio de inserción del marcador GT140, se realizó una TAIL/PCR tal y como se ha descrito anteriormente (Tsugeki, R., *et al.* Plant J. 10, 479-489 (1996)). La secuenciación posterior de los productos de la PCR aislados demostró que se corresponden a un BAC completamente secuenciado a partir del cromosoma 4, disponible en la base de datos pública como parte de la Arabidopsis Genome Initiative. La inserción de GT140 se localiza entre dos genes, uno que codifica un factor de transcripción hélice-bucle-hélice básico (bHLH) pronosticado y otro que representa un gen novedoso.

20

Mediante diferentes líneas de posterior investigación, se confirmó que el factor de transcripción bHLH (en la presente invención referido como *IND1* tal como se indica a continuación) era el gen relevante (SEC ID N°:1). Las fusiones promotor/potenciador::GUS del gen *IND1* se introdujeron en plantas de tipo natural y se descubrió que expresaban GUS en un patrón idéntico al de la línea del marcador de GT140. De modo interesante, aproximadamente el 25% de las líneas transgénicas no consiguieron expresar actividad de GUS significativa y se manifestó un fenotipo indehisciente. La explicación más probable de estos resultados es que las fusiones *IND1*::GUS, además del gen *IND1* endógeno, se cosuprimieron. La transferencia de ARN posterior confirmó una subregulación del gen *IND1* en estas líneas, y una transferencia de ARN adicional mostró, tal y como se esperaba, una disminución en la expresión del gen de *IND1* en los frutos *shp1 shp2*.

25

30

En paralelo a los estudios del marcador GT140 del margen de las válvulas descrito anteriormente, también se llevaron a cabo cribados para los mutantes de Arabidopsis que producen frutos indehiscientes. Además de la obtención de alelos adicionales de *SHP1* y *SHP2* mediante mutagénesis EMS de las existencias de semillas *shp2-1* y *shp1-1*, se obtuvieron también mutantes indehiscientes que no eran alélicos a *SHP1* o *SHP2*, respectivamente. Debido a que los estudios de GT140 sugirieron la posibilidad de que uno o más de esos mutantes indehiscientes pudiera corresponder al gen *IND1*, se clonó y se secuenció *IND1* de varios de estos mutantes. Cuatro alelos representan alelos mutantes independientes de *IND1*. El alelo más fuerte, *ind1-2*, contiene una única delección de nucleótido dentro del codón 55 que da lugar a un desplazamiento del marco de lectura y la producción de una proteína truncada de 64 en lugar de 198 aminoácidos. Los alelos *ind1-1* y *ind1-3* contienen sustituciones de nucleótidos en los codones 141 y 128 que cambian un aminoácido de leucina por uno de fenilalanina y una arginina por una histidina, respectivamente. Estos aminoácidos afectados se encuentran ambos en posiciones conservadas dentro del dominio de bHLH. El alelo *ind1-4* contiene una sustitución de nucleótido en el codón 92 que cambia una glutamina por un codón de parada, causando la producción de una proteína truncada de 91 aminoácidos. Dado que la inactivación de este factor de transcripción bHLH previene la dehiscencia del fruto, el gen es referido como INDEHISCENTE1 (*IND1*) y el mutante como, *ind1*. Hasta la fecha, *ind1* representa la única mutación génica individual descrita en Arabidopsis que bloquea específicamente la dehiscencia del fruto.

35

40

45

50

Referencias citadas en la descripción

Esta lista de referencias citadas por el solicitante se muestra únicamente para conveniencia del lector. No forma parte del documento de Patente Europea. Aunque se ha tenido una gran precaución a la hora de recopilar las referencias, no se pueden excluir errores u omisiones y la Oficina Europea de Patentes declina cualquier responsabilidad al respecto.

55

60 Documentos de patente citados en la descripción

- US 6410826 B [0058]
- US 4801340 A [0059]
- US 5034323 A [0064]
- US 5231020 A [0064]

65

ES 2 301 539 T3

- US 5283184 A [0064]
- US 0111967 W [0138]
- 5 • US 09548971 B [0138]

Documentos que no son patentes citados en la descripción

- 10 • **MEAKIN; ROBERTS.** *J. Exp. Botany*, 1990, vol. 41, 995-1002 [0004]
- **THEOLOGIS et al.** *Develop. Genetics*, 1993, vol. 14, 282-295 [0006]
- **LANAHAN et al.** *The Plant Cell*, 1994, vol. 6, 521-530 [0006] [0123]
- 15 • **SMITH et al.** *Nature*, 1988, vol. 334, 724-726 [0006]
- **PETERSEN et al.** *Plant Mol. Biol.*, 1996, vol. 31, 517-527 [0007] [0100]
- **MEAKIN; ROBERTS.** *J. Exp. Botany*, 1990, vol. 41, 1003-1011 [0007] [0122]
- 20 • **NI et al.** *Cell*, 1998, vol. 95, 657 [0032]
- **SMITH; WATERMAN.** *Add. APL. Math.*, 1981, vol. 2, 482 [0041]
- 25 • **NEEDLE MAN; WUNSCH.** *J. Mol. Biol.*, 1970, vol. 48, 443 [0041]
- **PEARSON; LIPMAN.** *Proc. Natl. Acad. Sci. (U.S.A.)*, 1988, vol. 85, 2444 [0041]
- **LILJEGREN et al.** *Nature*, 2000, vol. 404, 766-770 [0048]
- 30 • **SAVIDGE et al.** *The Plant Cell*, 1995, vol. 7, 721-723 [0050]
- **FLANAGAN et al.** *The Plant Journal*, 1996, vol. 10, 343-353 [0050]
- 35 • **SAMBROOK et al.** *Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory*, 1989 [0051]
- PCR Protocols: A Guide to Methods and Applications. *Academic Press*, 1990 [0055]
- **CARRUTHERS et al.** *Cold Spring Harbor Symp. Quant. Biol.*, 1982, vol. 47, 411-418 [0056]
- 40 • **ADAMS et al.** *J. Am. Chem. Soc.*, 1983, vol. 105, 661 [0056]
- **SHEEHY et al.** *Proc. Nat. Acad. Sci. USA*, 1988, vol. 85, 8805-8809 [0059]
- 45 • **PNUELI et al.** *The Plant Cell*, 1994, vol. 6, 175-186 [0059]
- **HASELOFF et al.** *Nature*, 1988, vol. 334, 585-591 [0063]
- **NAPOLI et al.** *The Plant Cell*, 1990, vol. 2, 279-289 [0064]
- 50 • **FLAVELL.** *Proc. Natl. Acad. Sci., USA*, 1994, vol. 91, 3490-3496 [0064]
- **KOOTER; MOL.** *Current Opin. Biol.*, 1993, vol. 4, 166-171 [0064]
- 55 • **WATERHOUSE et al.** *Proc. Natl. Acad. Sci. USA*, 1998, vol. 95, 13959-13964 [0067]
- **TABARA et al.** *Science*, 1998, vol. 282, 430-431 [0067]
- **MIZUKAMI et al.** *Plant Cell*, 1996, vol. 8, 831-844 [0069]
- 60 • **SIKORSKI et al.** *Meth. Enzymol.*, 1991, vol. 194, 302-318 [0074]
- **GREWAL; KLAR.** *Genetics*, 1997, vol. 146, 1221-1238 [0075]
- 65 • **XU et al.** *Genes Dev.*, 1996, vol. 10, 2411-2422 [0075]
- **PUCHTA et al.** *Experientia*, 1994, vol. 50, 277-284 [0075]

ES 2 301 539 T3

- SWOBODA *et al. EMBO J.*, 1994, vol. 13, 484-489 [0075]
- OFFRINGA *et al. Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 7346-7350 [0075]
- 5 • KEMPIN *et al. Nature*, 1997, vol. 389, 802-803 [0075]
- MOUNTFORD *et al. Proc. Natl. Acad. Sci. USA*, 1994, vol. 91, 4303-4307 [0076]
- VAULONT *et al. Transgenic Res.*, 1995, vol. 4, 247-255 [0076]
- 10 • COLE-STRAUSS *et al. Science*, 1996, vol. 273, 1386-1389 [0077]
- YOON *et al. Proc. Natl. Acad. Sci. USA*, 1996, vol. 93, 2071-2076 [0077]
- 15 • N G. J. MESSING *et al. GENETIC ENGINEERING IN PLANTS*. 1983, 221-227 [0079]
- JORDANO *et al. Plant Cell*, 1989, vol. 1, 855-866 [0080]
- BUSTOS *et al. Plant Cell*, 1989, vol. 1, 839-854 [0080]
- 20 • GREEN *et al. EMBO J.*, 1988, vol. 7, 4035-4044 [0080]
- MEIER *et al. Plant Cell*, 1991, vol. 3, 309-316 [0080]
- 25 • ZHANG *et al. Plant Physiology*, 1996, vol. 110, 1069-1079 [0080]
- WEISING *et al. Ann. Rev. Genet.*, 1988, vol. 22, 421-477 [0081] [0086]
- SAVIDGE, B.; ROUNSLEY, S.D.; YANOFSKY, M.F. *Plant Cell*, 1995, vol. 7, 721-733 [0083]
- 30 • ODELL *et al. Nature*, 1985, vol. 313, 810-812 [0089]
- BENFEY; CHUA. *Science*, 1990, vol. 250, 959-966 [0089]
- 35 • FUTTERER *et al. Physiol. Plant*, 1990, vol. 79, 154 [0089]
- KAY *et al. Science*, 1987, vol. 236, 1299 [0089]
- SINGER *et al. Plant Mol. Biol.*, 1990, vol. 14, 433 [0089]
- 40 • AN. *Plant Physiol.*, 1986, vol. 81, 86 [0089]
- LAST *et al. Theor. Appl. Genet.*, 1991, vol. 81, 581 [0090]
- 45 • MCELROY *et al. Mol. Gen. Genet.*, 1991, vol. 231, 150 [0090]
- MCELROY *et al. Plant Cell*, 1990, vol. 2, 163 [0090]
- COMAI *et al. Plant Mol. Biol.*, 1990, vol. 15, 373 [0090]
- 50 • O'GRADY. *Plant Mol. Biol.*, 1995, vol. 29, 99-108 [0091]
- HUANG. *Plant Mol. Biol.*, 1997, vol. 33, 125-139 [0091]
- 55 • MILLAR. *Plant Mol. Biol.*, 1996, vol. 31, 897-904 [0091]
- HUANG *et al. Plant Mol. Biol.*, 1996, vol. 33, 125-139 [0091]
- ZHONG *et al. Mol. Gen. Genet.*, 1996, vol. 251, 196-203 [0091]
- 60 • SOLOCOMBE *et al. Plant Physiol.*, 1994, vol. 104, 1167-1176 [0091]
- MARTINEZ *et al. J. Mol. Biol.*, 1989, vol. 208, 551-565 [0091]
- 65 • MANJUNATH *et al. Plant Mol. Biol.*, 1997, vol. 33, 97-112 [0091]
- HOLTORF. *Plant Mol. Biol.*, 1995, vol. 29, 637-646 [0091]

ES 2 301 539 T3

- **KIRCH**. *Plant Mol. Biol.*, 1997, vol. 33, 897-909 [0092]
- **LIU**. *Plant Physiol.*, 1997, vol. 115, 397-407 [0093]
- 5 • **CHEN**. *Plant J.*, 1996, vol. 10, 955-966 [0093]
- **STREIT**. *Mol. Plant Microbe Interact.*, 1997, vol. 10, 933-93.7 [0093]
- **SHEEN**. *Science*, 1996, vol. 274, 1900-1902 [0093]
- 10 • **DE VEYLDER**. *Plant Cell Physiol.*, 1997, vol. 38, 568-577 [0094]
- **MASGRAU**. *Plant J.*, 1997, vol. 11, 465-473 [0094]
- 15 • **STANGE**. *Plant J.*, 1997, vol. 11, 1315-1324 [0094]
- **UKNES et al.** *Plant Cell*, 1993, vol. 5, 159-169 [0094]
- **BI et al.** *Plant J.*, 1995, vol. 8, 235-245 [0094]
- 20 • **METT et al.** *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 4567-4571 [0095]
- **FURST et al.** *Cell*, 1988, vol. 55, 705-717 [0095]
- 25 • **GATZ et al.** *Plant J.*, 1992, vol. 2, 397-404 [0095]
- **RÖDER et al.** *Mol. Gen. Genet.*, 1994, vol. 243, 32-38 [0095]
- **GATZ**. *Meth. Cell Biol.*, 1995, vol. 50, 411-424 [0095]
- 30 • **CHRISTOPHERSON et al.** *Proc. Natl. Acad. Sci. USA*, 1992, vol. 89, 6314-6318 [0095]
- **KREUTZWEISER et al.** *Ecotoxicol. Environ. Safety*, 1994, vol. 28, 14-24 [0095]
- 35 • **TAKAHASHI et al.** *Plant Physiol.*, 1992, vol. 99, 383-390 [0095]
- **YABE et al.** *Plant Cell Physiol.*, 1994, vol. 35, 1207-1219 [0095]
- **UEDA et al.** *Mol. Gen. Genet.*, 1996, vol. 250, 533-539 [0095]
- 40 • **WILDE et al.** *EMBO J.*, 1992, vol. 11, 1251-1259 [0095]
- **BACK et al.** *Plant Mol. Biol.*, 1991, vol. 17, 9 [0095]
- 45 • **FEINBAUM et al.** *Mol. Gen. Genet.*, 1991, vol. 226, 449 [0095]
- **LAM; CHUA**. *Science*, 1990, vol. 248, 471 [0095]
- **COUPE et al.** *Plant Mol. Biol.*, 1993, vol. 23, 1223-1232 [0100]
- 50 • **KONCZ et al.** *Proc. Natl. Acad. Sci. USA*, 1989, vol. 86, 8467-8471 [0102]
- **KERTBUNDIT et al.** *Proc. Natl. Acad. Sci. USA*, 1991, vol. 88, 5212-5216 [0102]
- 55 • **TOPPING et al.** *Development*, 1991, vol. 112, 1009-1019 [0102]
- **AARTS et al.** *Nature*, 1993, vol. 363, 715-717 [0102]
- **SUNDARESAN et al.** *Genes. Devel.*, 1995, vol. 9, 1797-1810 [0102]
- 60 • **OTT et al.** *Mol. Gen. Genet.*, 1990, vol. 223, 169-179 [0103]
- **CLAES et al.** *The Plant Journal*, 1991, vol. 1, 15-26 [0103]
- 65 • **JEFIERSON et al.** *EMBO J.*, 1987, vol. 6, 3901 [0104]
- **OW et al.** *Science*, 1986, vol. 334, 856 [0104]

ES 2 301 539 T3

- **SHERIDAN.** *Genetics*, 1996, vol. 142, 1009-1020 [0105]
- **ABLER.** *Plant Mol. Biol.*, 1993, vol. 22, 10131-1038 [0105]
- 5 • **URAO.** *Plant Mol. Biol.*, 1996, vol. 32, 571-57 [0105]
- **CONCEICAO.** *Plant*, 1994, vol. 5, 493-505 [0105]
- **JOSEFSSON.** *JBL*, 1987, vol. 26, 12196-1301 [0105]
- 10 • **SJODAHL.** *Planta*, 1995, vol. 197, 264-271 [0105]
- **KIM.** *Plant Mol. Biol.*, 1994, vol. 26, 603-615 [0106]
- 15 • **MARTIN.** *Plant J.*, 1997, vol. 11, 53-62 [0106]
- **HANSEN.** *Mol. Gen. Genet.*, 1997, vol. 254, 337-343 [0106]
- **BEZERRA.** *Plant Mol. Biol.*, 1995, vol. 28, 137-144 [0106]
- 20 • **DE CASTRO.** *Plant Cell*, 1992, vol. 4, 1549-1559 [0106]
- **YAMAMOTO.** *Plant Cell*, 1991, vol. 3, 371-382 [0106]
- 25 • **MEIER.** *FEBS Lett.*, 1997, vol. 415, 91-95 [0107]
- **MATSUOKA.** *Plant J.*, 1994, vol. 6, 311-319 [0107]
- **SHIINA.** *Plant Physiol.*, 1997, vol. 115, 477-483 [0107]
- 30 • **CASAL.** *Plant Physiol.*, 1998, vol. 116, 1533-1538 [0107]
- **LI.** *FEBS Lett.*, 1996, vol. 379, 117-121 [0107]
- 35 • **BUSK.** *Plant J.*, 1997, vol. 11, 1285-1295 [0107]
- **DI LAURENZIO.** *Cell*, 1996, vol. 86, 423-433 [0108]
- **LONG.** *Nature*, 1996, vol. 379, 66-69 [0108]
- 40 • **ENJUTO.** *Plant Cell*, 1995, vol. 7, 517-527 [0108]
- **GRANGER.** *Plant Mol. Biol.*, 1996, vol. 31, 373-378 [0108]
- 45 • **KERSTETTER.** *Plant Cell*, 1994, vol. 6, 1877-1887 [0108]
- **HAKE.** *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 1995, vol. 350, 45-51 [0108]
- **LINCOLN.** *Cell*, 1994, vol. 6, 1859-1876 [0108]
- 50 • **KUMAGAI.** *Proc. Natl. Acad. Sci. USA*, 1995, vol. 92, 1679-1683 [0110]
- **VERDAGUER.** *Plant Mol. Biol.*, 1996, vol. 31, 1129-1139 [0110]
- 55 • **PASZKOWSKI et al.** *Embo J.*, 1984, vol. 3, 2717-2722 [0112]
- **FROMM et al.** *Proc. Natl. Acad. Sci. USA*, 1985, vol. 82, 5824 [0112]
- **KLEIN et al.** *Nature*, 1987, vol. 327, 70-73 [0112]
- 60 • **HORSCH et al.** *Science*, 1984, vol. 233, 496-498 [0113]
- **FRALEY et al.** *Proc. Natl. Acad. Sci. USA*, 1983, vol. 80, 4803 [0113]
- 65 • **EVANS et al.** *Protoplasts Isolation and Culture. Handbook of Plant Cell Culture*, 1983, 124-176 [0114]
- **BINDING.** *Regeneration of Plants, Plant Protoplasts. CRC Press*, 1985, 21-73 [0114]

ES 2 301 539 T3

- **KLEE** *et al. Ann. Rev. of Plant Phys.*, 1987, vol. 38, 467-486 [0114]
- **WILLIAMS; HILL**. *Science*, 1986, vol. 232, 1385-1389 [0117]
- 5 • **ECKER**. *Science*, 1995, vol. 268, 667-675 [0122]
- **CHAO** *et al. Cell*, 1997, vol. 89, 1133-1144 [0122]
- **BLEEKER** *et al. Science*, 1988, vol. 241, 1086-1089 [0123]
- 10 • **SCHALLER; BLEEKER**. *Science*, 1995, vol. 270, 1809-1811 [0123]
- **HUA** *et al. Science*, 1995, vol. 269, 1712-1714 [0123]
- 15 • **GUZMAN; ECKER**. *Plant Cell*, 1990, vol. 2, 513 [0123]
- **ROTHENBERG; ECKER**. *Sem. Dev. Biol. Plant Dev. Genet.*, 1993, vol. 4, 3-13 [0123]
- **KIEBER; ECKER**. *Trends Genet.*, 1993, vol. 9, 356-362 [0123]
- 20 • **VAN DER STRAETEN** *et al. Plant Physiol.*, 1993, vol. 102, 401-408 [0123]
- **HARPHAM** *et al. An. Bot.*, 1991, vol. 68, 55 [0123]
- 25 • **ROMAN** *et al. Genetics*, 1995, vol. 139, 1393-1409 [0123]
- **WILKINSON** *et al. Science*, 1995, vol. 270, 1807-1809 [0123]
- **KIEBER** *et al. Cell*, 1993, vol. 72, 427-441 [0124]
- 30 • **SESSIONS** *et al. Development*, 1995, vol. 121, 1519-1532 [0126]
- **ALVAREZ; SMYTH**. *Flowering Newsletter*, 1997, vol. 23, 12-17 [0126]
- 35 • **ROE** *et al. Cell*, 1993, vol. 75, 939-950 [0126]
- **COLLIER**. *Bacteriol. Rev.*, 1975, vol. 39, 54-85 [0129]
- **YAMAIZUMI** *et al. Cell*, 1978, vol. 15, 245-250 [0129]
- 40 • **PALMITER** *et al. Cell*, 1987, vol. 50, 435-443 [0129]
- **GREENFIELD** *et al. Proc. Natl. Acad. Sci., USA*, 1983, vol. 80, 6853-6857 [0129]
- 45 • **THORSNESS; NASRALLAH**. *Methods in Cell Biology*, 1995, vol. 50, 439-448 [0130]
- **MOFFAT** *et al. Development*, 1992, vol. 114, 681-687 [0131]
- **OLSNES; PIHL**. *Molecular Action of Toxins and Viruses. Elsevier Biomedical Press*, 1982, 51-105 [0131]
- 50 • **SUNDARESAN, V. et al. Genes Dev., 1995, vol. 9, 1797-1810 [0133]**
- **LILJEGREN, S.J.; DITTA, G.S.; ESHED, Y.; SAVIDGE, B.; BOWMAN, J.L.; YANOFSKY, M.F.** *Nature*
[0133]
- 55 • **TSUGEKI, R. et al. Plant J., 1996, vol. 10, 479-489 [0134]**

60

65

ES 2 301 539 T3

REIVINDICACIONES

- 5 1. Cassette de expresión que comprende un promotor unido operativamente a una secuencia de polinucleótidos IND1, o un complemento de la misma, que codifica un polipéptido IND1 idéntica por lo menos aproximadamente en un 70% a la SEC ID No: 2.
2. Cassette de expresión según la reivindicación 1, en el que el polipéptido IND1 comprende la SEC ID No: 2.
- 10 3. Cassette de expresión según la reivindicación 1, en el que el polinucleótido IND1 comprende las posiciones desde aproximadamente 2765 a aproximadamente 3361 de la SEC ID No: 1.
4. Cassette de expresión según la reivindicación 3, en el que el polinucleótido IND1 comprende la SEC ID No: 1.
- 15 5. Cassette de expresión que comprende un promotor heterólogo unido operativamente a un polinucleótido, o un complemento del mismo, en el que el polinucleótido es idéntico por lo menos en un 60% a por lo menos 200 nucleótidos contiguos de una secuencia que codifica la SEC ID No: 2.
6. Cassette de expresión según la reivindicación 5, en el que la secuencia es la SEC ID No: 1.
- 20 7. Cassette de expresión según la reivindicación 5, en el que el polinucleótido es idéntico por lo menos en un 60% a una secuencia de nucleótidos que codifica la SEC ID No: 2.
8. Cassette de expresión según cualquiera de las reivindicaciones 1 a 7, en el que el promotor es constitutivo.
- 25 9. Cassette de expresión según cualquiera de las reivindicaciones 1 a 7, en el que el promotor es específico de tejido.
10. Cassette de expresión según la reivindicación 9, en el que el promotor es un promotor específico de la zona de dehiscencia.
- 30 11. Cassette de expresión según la reivindicación 10, en el que el promotor comprende las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1.
- 35 12. Planta que comprende un cassette de expresión recombinante que comprende un promotor unido operativamente a una secuencia de polinucleótidos IND1 que codifica un polipéptido IND1 idéntica por lo menos aproximadamente en un 70% a la SEC ID No: 2.
- 40 13. Planta que comprende un cassette de expresión recombinante tal como se define en la reivindicación 5.
14. Procedimiento para retrasar la dehiscencia del fruto en una planta, comprendiendo el procedimiento la supresión de la expresión de un ácido nucleico IND1 en la planta mediante la introducción en la planta de un cassette de expresión recombinante que comprende un promotor unido operativamente a una secuencia de polinucleótidos que codifica un polipéptido IND1 idéntica por lo menos aproximadamente en un 70% a la SEC ID No: 2.
- 45 15. Planta según la reivindicación 12 ó 13, en la que la secuencia de polinucleótidos está unida operativamente al promotor en la orientación antisentido.
- 50 16. Procedimiento según la reivindicación 14, en el que la secuencia de polinucleótidos está unida operativamente al promotor en la orientación antisentido.
17. Planta según la reivindicación 12 ó 13, en la que la secuencia de polinucleótidos está unida operativamente al promotor en la orientación de sentido.
- 55 18. Procedimiento según la reivindicación 14, en la que la secuencia de polinucleótidos está unida operativamente al promotor en la orientación de sentido.
- 60 19. Planta según la reivindicación 17, cuando depende de la reivindicación 12, en la que la secuencia de polinucleótidos comprende además una segunda secuencia de polinucleótidos que codifica el polipéptido IND1 donde la segunda secuencia de polinucleótidos está unida operativamente a un segundo promotor en la orientación antisentido.
20. Procedimiento según la reivindicación 18, en el que la secuencia de polinucleótidos comprende además una segunda secuencia de polinucleótidos que codifica el polipéptido IND1 donde la segunda secuencia de polinucleótidos está unida operativamente a un segundo promotor en la orientación antisentido.
- 65 21. Planta según la reivindicación 17, cuando depende de la reivindicación 13, en la que la planta comprende además un segundo polinucleótido idéntico por lo menos en un 60% a por lo menos 200 nucleótidos contiguos de una

ES 2 301 539 T3

secuencia que codifica la SEC ID No: 2, donde la segunda secuencia de polinucleótidos está unida operativamente a un segundo promotor en la orientación antisentido.

5 22. Planta según la reivindicación 12 ó 13, en la que la lignificación se reduce en las células del margen de las válvulas.

23. Procedimiento según la reivindicación 14, en el que la lignificación se reduce en las células del margen de las válvulas.

10 24. Planta según la reivindicación 12, en la que el promotor es tal como se define en la reivindicación 10 o la reivindicación 11.

15 25. Procedimiento según la reivindicación 14, en el que el promotor es tal como se define en la reivindicación 10 o la reivindicación 11.

26. Planta según la reivindicación 13, en la que el promotor es tal como se ha definido en la reivindicación 8 ó 10.

20 27. Procedimiento según la reivindicación 14, en el que el cassette de expresión recombinante se introduce en la planta utilizando *Agrobacterium*.

28. Procedimiento según la reivindicación 14, en el que el polipéptido IND1 comprende la SEC ID No: 2.

25 29. Procedimiento según la reivindicación 14, en el que el polinucleótido IND1 comprende las posiciones desde aproximadamente 2765 a aproximadamente 3361 de la SEC ID No: 1.

30. Procedimiento según la reivindicación 14, en el que el polinucleótido IND1 comprende la SEC ID No: 1.

31. Planta según la reivindicación 13, en la que la secuencia es la SEC ID No: 1.

30 32. Planta según la reivindicación 13, en la que el polinucleótido es idéntico por lo menos en un 60% a una secuencia de nucleótidos que codifica la SEC ID No: 2.

35 33. Planta según la reivindicación 13, en la que la planta está **caracterizada** por una dehiscencia de la semilla retardada en comparación con una planta que no comprende el cassette de expresión.

34. Procedimiento para retrasar la dehiscencia del fruto en una planta, comprendiendo el procedimiento:

40 (i) suprimir la expresión de un gen IND1 o un polipéptido idéntico por lo menos en un 60% a la SEC ID No: 2 en la planta mediante la introducción en la planta de un cassette de expresión recombinante que comprende una secuencia de polinucleótidos idéntica por lo menos aproximadamente en un 70% a las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3586 de la SEC ID No: 1; y

45 (ii) seleccionar una planta con dehiscencia de fruto retardada en comparación con una planta en la que la expresión del gen IND1 o el polipéptido no está suprimida.

35. Procedimiento según la reivindicación 34, parte (i), en el que la secuencia de polinucleótidos comprende las posiciones desde aproximadamente 1 a aproximadamente 2764 o desde aproximadamente 3362 a aproximadamente 3856 de la SEC ID No: 1.

50 36. Procedimiento según la reivindicación 34, en el que la lignificación se reduce en las células del margen de las válvulas.

37. Procedimiento para retrasar la dehiscencia del fruto en una planta, comprendiendo el procedimiento las etapas de:

55 (i) tratar las semillas o material vegetal con una sustancia química mutagénica o con radiación ionizante;

60 (ii) identificar las plantas con un gen *IND1* mutado, en las que el gen *IND1*, antes de mutar, codifica un polipéptido que tiene una identidad en la secuencia de por lo menos un 60% con la SEC ID No: 2; y

(iii) seleccionar una planta con dehiscencia del fruto retardada en comparación con una planta en la que el gen IND1 no está mutado.

65 38. Procedimiento según la reivindicación 34, parte (i), en la que la etapa de supresión comprende la introducción en la planta de un cassette de expresión recombinante tal como se define en cualquiera de las reivindicaciones 5 a 7.

39. Procedimiento según la reivindicación 38, en el que el polinucleótido está unido operativamente al promotor en la orientación antisentido.

ES 2 301 539 T3

40. Procedimiento según la reivindicación 38, en el que el polinucleótido está unido operativamente al promotor en la orientación de sentido.

5 41. Procedimiento según la reivindicación 40, en el que el polinucleótido comprende además un segundo polinucleótido idéntico por lo menos en un 60% a por lo menos 200 nucleótidos contiguos de una secuencia que codifica la SEC ID No: 2, donde el segundo polinucleótido está unido operativamente a un segundo promotor en la orientación antisentido.

10 42. Procedimiento según la reivindicación 38, en el que el promotor es un elemento regulador selectivo de la zona de dehiscencia

43. Procedimiento según la reivindicación 38, en el que el cassette de expresión recombinante se introduce en la planta utilizando *Agrobacterium*.

15 44. Planta que comprende un polinucleótido modificado idéntico por lo menos en un 65% a la SEC ID No: 1, en la que la planta muestra una dehiscencia de fruto retardada en comparación con una planta que carece del polinucleótido modificado.

20 45. Planta según la reivindicación 44, en la que el polinucleótido modificado es idéntico por lo menos en un 70% a la SEC ID No: 1.

25 46. Procedimiento de creación de una planta con dehiscencia de fruto retardada, comprendiendo el procedimiento la introducción en dicha planta de un cassette de expresión recombinante que comprende un promotor unido operativamente a un polinucleótido que es idéntico por lo menos en un 60% a por lo menos 200 nucleótidos contiguos de la SEC ID No: 1 y la selección de una planta con dehiscencia de fruto retardada en comparación con una planta que carece del cassette de expresión recombinante.

30 47. Procedimiento según la reivindicación 46, que comprende la expresión en la planta de un polinucleótido de sentido idéntico por lo menos en un 65% a por lo menos 200 nucleótidos contiguos de la SEC ID No: 1 y la expresión en la planta de un polinucleótido antisentido idéntico por lo menos en un 65% a por lo menos 200 nucleótidos contiguos de la SEC ID No: 1.

35

40

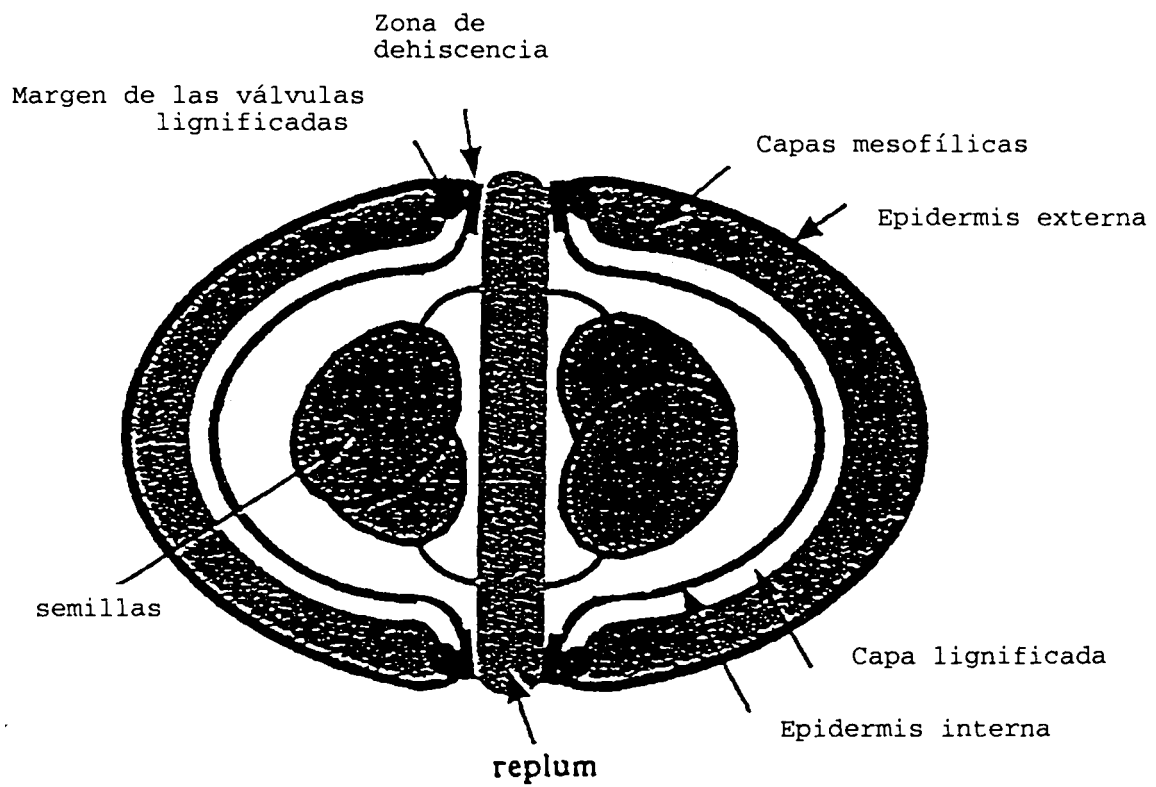
45

50

55

60

65



ES 2 301 539 T3

LISTA DE SECUENCIAS

SEC ID No: 1 INDI genómico

5 Rango de secuencia: 1 a 3856

60
CTCTAGACCA TCTACTATCC GGTGTGTGAC CCTTAAAGCT TTTGAAGACT ACTAGAATAA
10 120
TGCAAATACC ATATGTCCAT ATCCATCCTT TTCTTTTGTG TGAAGTGAAC ATTCTAATTT
180
TGTAAGAGAA AAAACCTTAT GTTAATATCA CCGTAGGCAA AAAAAATATC TCATCATATT
15 240
AAATTTTAT TATAAGATTA TACATTCTCT CGTTGTAAGA GTTACTCCAA TTGCAAGTGT
300
TGTATTAAC TATAAAAAGG ACGAAAGTAG GAAGCTTATA ATTAATTGAT GTTGCCATAGT
20 360
ACTGGTATAT TGTTGATGAA TATAACAAGT ATGAACATTA ATGCATGAAA CGGGGTATTT
420
TGTCTTGAAC TCATTAAAGG CAATGTGAAA AGAAGATGTG AGGTCTCATT TTGAAAATTT
25 480
ATCTTCTAGC TTTGTGCGATT TTAATCTAT GAAATGAACG CAACATATAG AAATTTCATG
540
TGGACAACGA CATTAGACG GTATCTTAAT TAGACCGATT AATTAGTAAT ATACTTATAT
30 600
ATATAATTAG TGGTGATTAT AAGTTACTT ATCCACTTGA GAATTTAAAC AATGGGCAAT
35 660
ACCTTAATGT CGAAAGAAGC CGTCCCACT TCGTGTAAATG AGTTATGGGG GAGAGATCCT
720
GTAAATCGT CAAATAAAC AACTTAAGAA CTAGAAATTG ACACCAAAA TCATAAAGAG
40 780
AACGTTGAAG AAGTCATTTA TCGTATCCAG CTCATATTTT CTAGCTAAGA TCAAATCAAG
840
GCCGTTGAAA GGGCTTGTA GAAAATGTCG AAGAAACCGT GGGGTTTAGA AGAAAGACAA
45 900
GAAATAGAAG AACAATGATG TTAATTTGCC TATTTTGGTG TATAGGAGTT GTCAAAAGAG
960
GAGAGAGAGA AGAAAATTAG GTCAAAATAA TGAGCACTAA AAATGGAGAC ATGTGTTGAG
50 1020
TAACTATTAC AAGAGCGACT TATGCTTCCT TATGGCAATG ATATCCAAC CAAAGTGCAA
55 1080
CGCTCCTTTT TTGCCCTAAT TTCGTAAAGT CTCTCTCCTT CTTCGTCCTT AGGAAAAACC
1140
CTAGAAATTT AATCCCTTGT TCTTGATCTT GCTTTTGGAG TAAACATGAT TTTGACCACA
60 1200
CACTATTTCT TCTATCTTTT GTGGTCTATA GGATTTTGCT TTATATGTGT TTCTTGTATT
1260
GCTCCGTACG TACGTATACG AATTTAATG GTTATAACAA GGTTTATATA AACTAGCACA
65

ES 2 301 539 T3

1320
 AATGAGTCCA TGAAATTTGT TAGCGAAAAA GGTAGAAATA TATTGAGTCT TTAACCGGCA
 5
 1380
 ATATATATAA TTTTGCTGCA AAACCTAGCT TTAATCATGA TCTAATGATA TTTTCTTTAA
 1440
 TTTCCTTTGC CAAATTAATC ACATGCACGG ATTTTTGGCA AGTTATGTGT CGAATTCCTC
 10
 1500
 CATTACACACA AACTAAACT TAATTAGAAC TCTAGGAAAT ATTTTAAAAT GACAACITTA
 1560
 TCGAAAAAAA TTTAGTTATG AAAACAATTC CAGAATTTAA CATGAGCTAT ATAATTTAAG
 1620
 ATAAAATGAA GTAATATTGA TATGTATGTA ATAACATATC TGATTGCGGT AAAAAAAAC
 1680
 ATATCTGATT AAATTGTTCA TGCAGGCCCA TGTCACTATG ATGTCATCAC GTTTTATTT
 20
 1740
 TCACAATAAC TAATATATAT TCAAAAAAAT AGTTTTGTCA GATTAAATTT TTTTGGTGG
 1800
 TCAGCTTTCT CCAACCTACT AAACCTAGTT GGAATGTTCT CTTCTTTATT TTTCTTTTTC
 25
 1860
 TTGATTCTT ATGTTTTTTA TTTATGGAAT TTTAAGACGG ATTGTTTAGG TCGTTTCTCT
 1920
 CTTTTCTGT TTTCTAAAGT TACTTTTGTA AACTCATCTC CTCCCAATTA GACAGTCAAT
 30
 1980
 CATATAGTTA TCTTTTAATA TATGTCTAGT TGATAAAAAA AATGAAAAA TACTGGTGGT
 2040
 AGTTCTACTA ATGTTTGTGT AAAAAATCTG ATATTATGAA TCTAATCAAT TTCTTTGATC
 35
 2100
 GTATAATGTG GGTAAATTT AGTAATTTTT TACATAAATA AGAACTGTAA TGTTGATGTA
 2160
 TATGGGGAA TCAGTATATT AGCTTGGGTA ACTATACTTC TGGAAATACT TGAAGATTTA
 40
 2220
 ACTATTTGCA AAATTATAAT TTAGTCCCGA AAAATACAGA CGACGGGACA CGACAACATA
 2280
 TAAGCAGGTT TGAATCTTGG AAAATTTTGT ATACATAACC TATATAAATA CTAATGTTCT
 45
 2340
 GGTGGGTTT AAAAGCCTTT TCAAAAGTTC CATTTTTTAA ATTCAAGGAC ATTTTACATA
 2400
 GGAAATAAGT TGAGTCATAA AAAATAATGG TTATTTTGTA AGGTTTTTTT TTTGATTAAA
 50
 2460
 ACGCACATAT TAAGAAGTTA GTTTTTTTTC ACTACCAAAT ATCAATTAAT TTAACCACAT
 2520
 GCAACCATTC ATAAAACAAT ACTATTAAAG AATATAAATA ATCACAAT ATTAATATCA
 55
 2580
 CTTAAAATTT ACATATAAAT TTACAAAACA TCTAATTAAT TGAACAGAA AGGAAAAGGT
 60
 2640

ES 2 301 539 T3

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

AAAAATATATC ATAAAATGAG ACATATATCC TATAAAAAA AAATGAGGCA TATGAAGTAA
2700
ATAATAAGAG ACATGCATGT AAGCATTCCG TTAATTAATC GAGTCAAAGA TATATATCAG
2760
TAAATACATA TGTGTATATT TCTGGAAAAA GAATATATAT ATTGAGAAAT AAGAAAAGAT
2820
GAAAATGGAA AATGGTATGT ATAAAAGAA AGGAGTGTGC GACTCTTGTG TCTCGTCCAA
M E N G M Y K K K G V C D S C V S S K>
2880
AAGCAGATCC AACCCAGCC CCAAAGAAG CATGATGGAG CCTCAGCCTC ACCATCTCC
S R S N H S P K R S M M E P Q P H H L L>
2940
CATGGATTGG AACAAAGCTA ATGATCTTCT CACACAAGAA CAGCAGCTT TTCTCAATGA
M D W N K A N D L L T Q E H A A F L N D>
3000
TCCTCACCAT CTCATGTTAG ATCCACCTCC CGAAACCCTA ATTCACCTGG ACGAAGACGA
P H H L M L D P P P E T L I H L D E D E>
3060
AGAGTACGAT GAAGACATGG ATGCGATGAA GGAGATGCAG TACATGATCG CCGTCAATGA
E Y D E D M D A M K E M Q Y M I A V M Q>
3120
GCCCGTAGAC ATCGACCCTG CCACGGTCCC TAAGCCGAAC CGCCGTAACG TAAGGATAAG
P V D I D P A T V P K P N R R N V R I S>
3180
CGACGATCCT CAGACGGTGG TTGCTCGTCG GCGTCGGGAA AGGATCAGCG AGAAGATCCG
D D P Q T V V A R R R R R E R I S E K I R>
3240
AATTCTCAAG AGGATCGTGC CTGGTGGTGC GAAGATGGAC ACAGCTTCCA TGCTCGACGA
I L K R I V P G G A K M D T A S M L D E>
3300
AGCCATACGT TACACCAAGT TCTTGAACG GCAGGTGAGG ATTCTTCAGC CTCACTCTCA
A I R Y T K F L K R Q V R I L Q P H S Q>
3360
GATTGGAGCT CCTATGGCTA ACCCCTCTTA CCTTTGTTAT TACCACAACCT CCCAACCCCTG
I G A P M A N P S Y L C Y Y H N S Q P >
3420
ATGAACTACA CAGAAGCTCG CTAGCTAGAC ATTTGGTGTG ATCCTCTCAA CCTTTTTCAT
3480
GTTGATATAT TATATATAGA TGCATAAAGA TTCGATCCAA GATTGTATGG GTGTTTAAAT
3540
ATTATTATTC TAAGATATAT GATGTACAAT TGTGTACCAA GTTCTTTTAT CTTGATATCA
3600
TATGCATAAA TAATTGGTGA ATAAAAGAA GATATTGATT GTAACAATAA AAAAGAAGAT
3660
ATTGATTGTT AATTAGGETT TGATCATTCT GTATGAAAGC TTTGGCCTGC AAATTAATTT
3720
TCGATATATA TATATATATA TGGAGAATAT ATATCAAATA CTTTTTTAAAT TTGACTATAA
3780
TTTGTATCAA TTATCTGAAT CTGATGAGTG TAGGTTATAT ATGGATTAGC AAAAAAGAAA
3840
ACAACCATTA TTACGCACCT ACATTAAAAA TCATCCACCA AAGAAGAAC CATCCTCAAG
AGGGTTCCCT CTAGAG

ES 2 301 539 T3

SEC ID No: 2 Proteína IND1

5 M E N G M Y K K K G V C D S C V S S K
 S R S N H S P K R S M M E P Q P H H L L
 10 M D W N K A N D L L T Q E H A A F L N D
 P H H L M L D P P P E T L I H L D E D E
 E Y D E D M D A M K E M Q Y M I A V M Q
 15 P V D I D P A T V P K P N R R N V R I S
 D D P Q T V V A R R R R E R I S E K I R
 20 I L K R I V P G G A K M D T A S M L D E
 A I R Y T K F L K R Q V R I L Q P H S Q
 25 I G A P M A N P S Y L C Y Y H N S Q P *

SEC ID No: 3 Promotor 5' de IND1

30 60 CTCTAGACCA TCTACTATCC GGTGTGGAC CCTTAAAGCT TTTGAAGACT ACTAGAAATA
 120
 35 TGCAAATACC ATATGTCCAT ATCCATCCTT TTCTTTTGGT TGAAGTGAAC ATTCTAATTT
 180
 TGTAAGAAGAA AAAACCTTAT GTTAATATCA CCGTAGGCAA AAAAAATATC TCATCATATT
 240
 40 AAATTTTAT TATAAGATTA TACATTCTCT CGTTGTAAGA GTTACTCCAA TTGCAAGTGT
 300
 TGTATTAAGT AATAAAAAGG ACGAAAGTAG GAAGCTTATA ATTAATTGAT GTTGCAATAGT
 360
 45 ACTGGTATAT TGTGATGAA TATAACAAGT ATGAACATTA ATGCATGAAA CGGGGTATTT
 420
 TGTCTTGAAC TCATTAAAGG CAATGTGAAA AGAAGATGTG AGGTCTCATT TTGAAAATTT
 480
 50 ATCTTCTAGC TTTGTGATTT TTAATCTAT GAAATGAACG CAACATATAG AAATTTTATG
 540
 55 TGGACAACGA CATTAGACG GATCTTAAAT TAGACCGATT AATTAGTAAT ATACTTATAT
 600
 ATATAATTAG TGGTGATTAT AAGTTTACTT ATCCAATTGA GAATTTAAAC AATGGGCAAT
 660
 60 ACCTTAATGT CGAAGAAGC CGTCCCCACT TCGTGTAATG AGTTATGGGG GAGAGATCCT
 720
 GTTAAATCGT CAAATAAAC AACTTAAGAA CTAGAAATG ACACCAAAAA TCATAAAGAG

ES 2 301 539 T3

780
 AACGTTGAAG AAGTCATTTA TCGTATCCAG CTCATATTTT CTAGCTAAGA TCAAATCAAG
 840
 5 GCCGTTGAAA GGGCTTGTA GAAATGTGCG AAGAAACCGT GGGGTTTAGA AGAAAGACAA
 900
 GAAATAGAAG AACAATGATG TAAATTGCC TATTTTGGTG TATAGGAGTT GTCAAAAGAG
 960
 10 GAGAGAGAGA AGAAAATTAG GTCAAAATAA TGAGCACTAA AAATGGAGAC ATGTGTTGAG
 1020
 TAACTATTAC AAGAGCGACT TATGCTTCCT TATGGCAATG ATATCCAAAC CAAAGTGCAA
 1080
 15 CGCTCCTTTT TTGCCCTAAT TTCGTAAAGT CTCTCTCCTT CTTGTCCTT AGGAAAAACC
 1140
 CTAGAAATTT AATCCCTTGT TCTTGATCTT GCTTTTTGAG TAACCATGAT TTTGACCACA
 1200
 20 CACTATTCTT TCTATCTTTT GTGGTCTATA GGATTTTGCT TTATATGTGT TTCTTGATTT
 1260
 25 GCTCCGTACG TACGTATACG AATTTAAATG GTTATAACAA GGTTTATATA AACTAGCACA
 1320
 AATGAGTCCA TGAAATTTGT TAGCGAAAAA GGTAGAAATA TATTGAGTCT TTAACGGCA
 1380
 30 ATATATATAA TTTTGCTGCA AAACCTAGCT TTAATCATGA TCTAATGATA TTTTCTTTAA
 1440
 TTTCTTTTGC CAAATTAATC ACATGCACCG ATTTTGGCA AGTTATGTGT CGAATCTTC
 1500
 35 CATTACACA AACTAAACT TAATTAGAAC TCTAGGAAAT ATTTTAAAT GACAACTTTA
 1560
 TCGAAAAAAA TTTAGTTATG AAAACAATTC CAGAATTAAT CATGAGCTAT ATAATTTAAG
 1620
 40 ATAAAAATGAA GTAATATTGA TATGTATGTA ATAACATATC TGATTGCGGT AAAAAAATC
 1680
 ATATCTGATT AAATTGTTCA TGCAGGCCCA TGCACTATG ATGTCATCAC GTTTTATTTT
 1740
 45 TCACAATAAC TAATATATAT TCAAAAAAAT AGTTTGTCA GATTAAATTT TTTTGGTGG
 1800
 TCAGCTTTCT CCAACCTACT AAACCTAGTTT GGAATGTTCT CTTCTTTATTT TTTCTTTTTC
 1860
 50 TTGATTTCTT ATGTTTTTTA TTTATGGAAT TTTAAGACGG ATTGTTTAGG TCGTTTCTCT
 1920
 55 CTTTTCTTGT TTTCTAAAGT TACTTTTGTA AACTCATCTC CTCCCAATTA GACAGTCAAT
 1980
 CATATAGTTA TCTTTTAATA TATGTCTAGT TGATAAAAAA AATGAAAAAA TACTGGTGGT
 2040
 60 AGTTCTACTA ATGTTTGTGT AAAAAATCTG ATATTATGAR TCTAATCAAT TTCTTTGATC
 2100

65

ES 2 301 539 T3

GTATAATGTG GGTAAATTT AGTAATTTT TACATAAATA AGAACTGTAA TGTTGATGTA
2160
5 TATTGGGGAA TCAGTATATT AGCTTGGGTA ACTATACTTC TGGAAATACT TGAAGATTTA
2220
ACTATTTGCA AAATTATAAT TTAGTCCCGA AAAATACAGA CGACGGGACA CGACAACATA
2280
10 TAAGCAGGTT TGAATCTTGG AAAATTTTGT ATACATAACC TATATAAATA CTAATGTTCT
2340
GGTTGGGTTT AAAAGCCTTT TCAAAAGTTC CATTITTTTAA ATTCAAGGAC ATTTTACATA
2400
15 GGAAATAAGT TGAGTCATAA AAAATAATGG TTATTTTGTGTA AGGTTTTTTT TTTGATTAA
2460
20 ACGCACATAT TAAGAAGTTA GTTTTTTTTC ACTACCAAAT ATCAATTAAT TTAAAACCAT
2520
GCAACCATTTC ATAAAACAAT ACTATTAAAG AATATAAATA ATCACAAAAT ATTAATACA
2580
25 CTTAAAATTT ACATATAAAT TTACAAAACA TCTAATTAAT TGAAACAGAA AGGAAAAGGT
2640
30 AAAATATATC ATAAAATGAG ACATATATCC TATAAAAAA AAATGAGGCA TATGAAGTAA
2700
ATAATAAGAG ACATGCATGT AAGCATTCCG TTAATTAATC GAGTCAAAGA TATATATCAG
2760
35 TAAATACATA TGTGTATATT TCTGGAAAAA GAATATATAT ATTGAGAAAT AAGAAAAGAT
GAAAA

40 SEC ID NO: 4 Promotor 3' de IND1

3420
45 ATGAACTACA CAGAAGCTCG CTAGCTAGAC ATTTGGTGTG ATCCTCTCAA CCTTTTTTCAT
3480
GTTGATATAT TATATATAGA TGCATAAAGA TTCGATCCAA GATTGTATGG GTGTTTAAAT
3540
50 ATTATTATTC TAAGATATAT GATGTACAAT TGTGTACCAA GTTCTTTTAT CTTGATATCA
3600
TATGCATAAA TAATTGGTGA ATAAAAGAA GATATTGATT GTAACAAAA AAAAGAAGAT
3660
55 ATTGATTGTT AATTAGGTT TGATCATTCT GTATGAAAGC TTTGGCCTGC AAATTAATT
3720
TCGATATATA TATATATATA TGGAGAATAT ATATCAAATA CTTTTTAAAT TTGACTATAA
3780
60 TTTGTATCAA TTATCTGAAT CTGATGAGTG TAGGTTATAT ATGGATTAGC AAAAAAGAAA
3840
65 ACAACCATTG TTACGCACCT ACATTAATAA TCATCCACCA AAGAAGAAAC CATCCTCAAG
AGGGTCCCT CTAGAG

ES 2 301 539 T3

SEC ID NO: 5 SHP1 genómico

```

60
5  AGATCTGCAA CAGTGAAAAG AGAAAACAAA ATGGACTTGA AGAGGTTTTG ACAATGCCAG
120
AGATAATGCT TATTCCCTAA TATGTTGCCA GCCAAGTGTC AAATTGGCTT TTTAAATATG
180
10 GATTTCTGTA TCAGTGGTCA TATTTGTGGA TCCAACGTAT TCATCATCAA GTTCTCAAGT
240
TTGCTTTCAG TGCAATTCTA ATTACACCGT TTAACHTTAA CATGCATGTC ATTATAATTA
300
15 CTTCTTCACT AAGACACAAT ACGGCAAACC TTTCAGATTA TATTAATCTC CATAAATGAA
360
ATAATTAACC TCATAATCAA GATTCAATGT TTCTAAATAT ATATGGACAA AATTTACACG
420
20 GAAGATTAGA TACGTATATT AGTAGATTTA GTCTTTTCGTT TGTGCGATAA GATTAACCAC
480
25 CTCATAGATA GTAATATCAT TGTCAAATTC CTCTCGGTTT AGTCGCTAAA TTGTATCTTT
540
TTTAAGCCTA AAAGTAGTGT ATTGCGATAT GACTTATCGT CCTAACTTTT TTTTAAATTA
600
30 ACAAAAAAAT CGAAAAGAAA ATAATCTGTT AAATATTTTT TAAGTACTCC ATTAAGTTTA
660
GTTTCTATTT AAAAAATGCT TGAAATTTGA CAGTTATGTT CAACAATTTT GAATCATGAG
720
35 CGATGTCTAG ATACTCAGAA TTTAATCAAG ATGTCTTATC AAATTTGTTG TCACTCGAGG
780
ACCCACGCAA AAGAAAAGAC TAATATGATT TTTATTTGGT CTGGATATTT TTGTAGAGGA
840
40 TGAAACTAAG AGAGTGAAAAG ATTCGAAATC CACAATGTTT AAGAGAGCTC AAAGCAAAAA
900
45 GAAAAATGAA GATGAAGGAC TAAAGAACAA TAAGCAACTA CTTATACCCT ATTTCCATAA
960
AGGATTCAGG TACTAGGAGA AGTTGAGGCA AGTTNNNNNN NATTGATTCA AATTTTCATT
1020
50 TATTTTTACA ATTTAATTCA CCTAAGTTAT TATGCATTC TCATCATTGG TACATTTTCT
1080
GTATAGCGTA TTTACATATA TGAAATAAAT TAAATATGTC CTCACGTTGC AAGTAGTTAA
1140
55 TGAATGTCCC CACGCAAAAA AAAATCCCTC CAAATATGTC CACCTTTTCT TTTCTTTTTA
1200
ATTCCAAAAT TACCATAAAC TTTTGGTTTA CAAAAGATTT CTAGAAATG AGGAAGATAT
1260
CCTAAATGAT TCATGAATCC TTCAATAATC TGAAGTTTGC GATATTTTCG ATTTTCTTCA
1320
65 AGAGTTGCGA TATTTGTAAT TTGGTGACCT TAAACTTTTT TTGATAAAGA GTAAACGTTT

```

ES 2 301 539 T3

1380
 TTTCTTAAAA GTAAAACCTG ATTTTATGTT TTAGGGTTCT AGCTCAACTT TGTATTATAT
 1440
 5 TTCTTGCAAA AAGAGTTCGT TAACTGCATT CTTCAACACT ATAAAGTGAT TATCAAAAC
 1500
 ATCTTCATGA ACATTAAGAA AAACAATATT TGGTTTCGGT TAGAGCTTGG TTTTGCTTGG
 1560
 10 CTTGATTCAC ATACCCATTC TAGACTTTGG CATAAATTG ATACGATAGA GAGTATCTAA
 1620
 TGGTAATGCA GAAGGGTAAA AAAAGGAAGA GAGAAAAGGT GAGAAAGATT ACCAAAAATA
 1680
 15 AGGAGTTTCA AAAGATGGTT CTGATGAGAA ACAGAGCCCA TCCCTCTCCT TTTCCCCTTC
 1740
 CCATGAAAGA AATCGGATGG TCCTCCTTCA ATGTCCTCCA CCTACTCTTC TCTTCTTTCT
 1800
 20 TTTTTTCTTT CTTATTATTA ACCATTAAAT TAATTTCCCC TTCAATTTC A GTTCTIAGTT
 1860
 25 CTGTAAAAAG AAAATACACA TCTCACTTAT AGATATCCAT ATCTATTTAT ATGCATGTAT
 1920
 AGAGAATAAA AAAGTGTGAG TTTCTAGGTA TGTTGAGTAT GTGCTGTTTG GACAATTGTT
 1980
 30 AGATGATCTG TCCATTTTTT TCTTTTTTCT TCTGTGTATA AATATATTTG AGCACAAGA
 2040
 AAAACTAATA ACCTTCTGTT TTCAGCAACT AGGGTCTTAT AACCTTCAA GAAATATTC
 2100
 35 TTCAATTGAA AACCCATAAA CCAAATAGA TATTACAAA GGAAAGAGAG ATATTTTCAA
 2160
 GAACAACATA ATTAGAAAAG CAGAAGCAGC AGTTAAGTGG TACTGAGATA AATGATATAG
 2220
 40 TTTCTCTTCA AGAACAGTTT CTCATTACCC ACCTTCTCCT TTTTGCTGAT CTATCGTAAT
 2280
 CTTGAGAACT CAGGTAAGGT TGTGAATATT ATGCACCATT CATTAAACCCT AAAAATAAGA
 2340
 45 GATTTAAAAT AAATGTTTCT TCTTTCTCTG ATTCTTGTGT AACCAATTCA TGGGTTTGAT
 2400
 ATGTTTCTTG GTTATTGCTT ATCAACAAAG AGATTTGATC ATTATAAAGT AGATTAATAA
 2460
 50 CTCTTAAACA CACAAAGTTT CTTTATTTTT TAGTTACATC CCTAATTCTA GACCAGAACA
 2520
 TGGATTTGAT CTATTTCTTG GTTATGTATC TTGATCAGGA AAAGGGATTI GATCATCAAG
 2580
 55 ATTAGCCTTC TCTCTCTCTC TCTAGATATC TTTCTTGAAT TTAGAAATCT TTATTTAATT
 2640
 60 ATTTGGTGAT GTCATATATG GATCAATGGA GGAAGGTGGG AGTAGTCACG ACCGAGAGAG
 2700
 TAGCAAGAAA CTAGGGAGAG GGAAAATAGA GATAAAGAGG ATAGAGAACA CAACAATCG

65

ES 2 301 539 T3

2760
 TCAAGTACT TTCTGCAAC GACGCAATGG TCTTCTCAAG AAAGCTTATG AACTCTCTGT
 2820
 5 CTTGTGTGAT GCCGAAGTTG CCTTCGTCAAT CTTCTCCACT CGTGGCCGTC TCTATGAGTA
 2880
 CGCCAACAAC AGGTACGCTT CTCCTACTCT ATTTCTTGAT CTTGTTTTCT TAATTTAAC
 2940
 10 TAAACAAGAT CCTAGTTCAA ATGATAACAA AGTGGGGATT GAGAGCCAAG ATTAGGGTTT
 3000
 GGTAAATTTA GAAAACCAGA TTCACTTGTI TGATACATTT AATATCTCTC TAGCTAGATT
 3060
 15 TAGTACTCTC TCCTCTATAT ATGTGTGGGT GTGTGTGTAA GTGTGTATAT GSTATGCAAAT
 3120
 GCAAGAAGAA GAAGAAAAG TTATCTTGTI TTCTCAAATT CTGATCAGCT TTGACCTTAG
 3180
 20 TTCACTCTT TTTTCTGCAA ATCATTGAA CCTGATGCAT GTCAGTTTCT ACAATACACT
 3240
 25 TTAAATTTTG ACGGCCCATC AAATTTCTTA GGGTTACTT CAGTGAACAA AATTGGGTTT
 3300
 TTGACACGAT TTAGCATGTA TATATAAAA TAGGGGATGA TCAAGACTTA TGTAACCTCT
 3360
 30 GTCTGGTGAA ACTAGGGACA AAGTCTACTG ATGAGTTGTC ACTAGGGATC CATTGTATCA
 3420
 TTAAATCCCA ACAAATGA AACAAATTT TGAGAATTTA TATGCTGAAG TTTTCAACC
 3480
 35 CTCTTTTTTA AATAACTTTA TATTATGTAG ATTTGTATTT AGGGTAATTT GTCCAACACTAG
 3540
 AAGTCCTAAA AATCAATAAA CACACGGATG ACTTTGTCTA ACATTGTATC AGTCATCAA
 3600
 40 TGTAATTTG TACAATAAT GAAATTAAG ATTTAGTCTC TTTTATTTT TTTGTTTAGG
 3660
 GTGTATATAT ATATATATAT GSTATTTGTI TGCAATGATA TATCAATGAG AGGGAGAGAA
 3720
 45 CTCAGAGAAG TGTCGGAAAT TAAAATGGTA CGAGCCAATT GGAATCTCTG GCATTCTGAG
 3780
 CTTCAATTTGT TTGTTATTAG AAAAAAAAAA AAAAAATCCT TTAAGATAC CTTCAATGATG
 3840
 50 ACATTGAATC ATGTAATATA CACGATACAT GGTCTAATTC CTCCTCAAC CCTAATTACC
 3900
 AATTTGAAA CCATAATATT TACTAGTATG TTTATATATC CTTACTTTAA GACATTGTTT
 3960
 55 GTTTATAATA CCTTGTGAAT TAAGAAAAA AAAAAAAC TTTGGGATCT ATTCAAGCCA
 4020
 60 TGTGTTAGAA TAAATTTATA AATTTCTCC TCGTACTGGT CAGATATTGG TCCAACTCC
 4080

65

ES 2 301 539 T3

AAAGCCTTCC CTTTTCAAGG AAAAAACAT TTCGAAATTA ACTCTAATTA ATCAAGAATT
 TCCTACAATG TATACATCTA ATGTTTTTTC CGCGATCTTA CTTATTAGTG TGAGGGGTAC 4140
 AATTGAAAGG TACAAGAAAG CTGTTCCEGA TGCCGTCAAC CCTCCTTCCG TCACCGAAGC 4200
 TAATACTCAG GTACCAATTT ATATTGTTTG ATTCTCTTTG TTTTATCTTC TTCTTTTCAT 4260
 TATATATATG ATCAACAAA AATATAACCT ACAAAAAGAG AGAGTTC AAG GAAATGCATT 4320
 GAAACGGTTT CGTTATGGTG TTTGAATACA TGGATTTTTG AAGTACTATC AGCAAGAAGC 4380
 CTCTAAGCTT CGGAGGCAGA TTCGAGATAT TCAGAATCA AATAGGTAAT TCATTAACTT 4440
 TTCATGAACT CTTCGATTTG GTATTAGGTC ACTTAATTG GTGTCGGTCC AAAAGTCCGC 4500
 TTGTAGTTTT CTTTAGAAGT TGTTTTGTTC AATGTTTATG TTTACAAATT GAAGGCATAT 4560
 TGTGGGGAA TCACTTGGTT CCTTGAACCT CAAGGAACTC AAAAACCTAG AAGGACGTCT 4620
 TGAAAAGGA ATCAGCCGTG TCCGCTCCAA AAAGGTAAAA TCTACGTTGC TCTCTCTCTG 4680
 TGTCTCTGTC TCTCTCTCTA TATATAGTCC CTTAGTTTAT ATAGTTCATC ACCCTTTTGT 4740
 GAGAATTTTG CAGAATGAGC TGTTAGTGGC AGAGATAGAG TATATGCAGA AGAGGGTAAG 4800
 AACGTTTCTC CCATTCCAAG TAATTAGATC TTTCTTCGTC TTTGTGAGGG TTTGAGTTTT 4860
 CCCATAAATC ATGTGTAGGA AATGGAGTTG CAACACAATA ACATGTACCT GCGAGCAAAG 4920
 GTTAGCCACG TTCTGTCCA AATCTTAATC TCAATATCTA CTCTTTTCTT CATTGTATAA 4980
 CTAAGATAAC GTGAATAACA AGAAAACTTT TGTTTTTGGG TTTAATAGAT AGCCGAAAGC 5040
 GCCAGATTGA ATCCGGACCA GCAGGAATCG AGTGTGATAC AAGGGACGAC AGTTTACGAA 5100
 TCCGGTGTAT CTCTCATGA CCACTGCGAG CATTATAATC GGAAGTATAT TCCGGTGAAC 5160
 CTCTTGAAC CGAATCAGCA ATTCTCCGGC CAAGACCAAC CTCTCTTCA ACTTGTGTAA 5220
 CTCAAAACAT GATAACTTGT TTCTTCCCT CATAACGATT AAGAGAGAGA CGAGAGAGTT 5280
 CATTTTATAT TTATAACGGC ACTGTGTATT CATAGTTTAG GTTCTAATAA TGATAATAAC 5340
 AAAACTGTTG TTTCTTTGCT TAATTACATC AACATTTAAA TCCAAAGTTC TAAAACACGT 5400

ES 2 301 539 T3

CGAGATCCAA AGTTTGT CAT ACAAGATTAG ACGCATAAC GATCAGTTAA TAGATTTTAA 5460

5

GTGCCTTTTA ATATTTACAT ATAGTTGCAG CTTCGATTAG ATCATGTCCA CCAAACACTC 5520

10

ACAATTAGAG ACAAGCAAAA CTATAAACAT TGATCATAAA ATGATTACAA CATGTCCATA 5580
AATTAATTAT GGATTACAAA AATAAAAACT TACAAAAGAT CT

SEC ID No: 6 SHP2 genómico

15

60

GAATTCGTAA CAGAATTTAG TGAATAATAT TGTAATTACC AGGCAAGGAC TCTCCAAACG

20

GATAGCTCGA ATATCGTTAT TAAAGAGTAA ATGATCCAAT ATGTAAGCCA TTGTTGATCA 120

25

TCTAACATTG TTGGACTCTC TATTGCTCGA AATGATGCAT ACCTAATCAT TTATTCAGTT 180

AACTATCAAG TTGCATTTGT AAAAACCAAA CATTAAATT CAGATTTGAT ATCACTTACA 240

30

GAGGATAGAG AAGCATGACT CCAGGCCTGC ATGCAACAAG AAAAAGGAAG AAAATAATGT 300

TAAAAATTTG ACAAATATAG TGTTTATTTT TATTATATGA GACAGAATTT GAATAAATC 360

35

CTACCCAAC T AGAGCATCAA AACGTTTTGC AATCGCAATA ATGAAACCCA TTTTCTTTTT 420

GAGTTTTTAC TCTTCTTTCA ACAGAACTT TCTCAAACGT CTTTAGCACT GTGACGTTAG 480

40

ATATATACAC AAAAGCTTGA AATTCTTCA AGCAAAAGAA TCTTTGTGGG AGTTAAGGCA 540

ACAAGCCAGG TAAAGAATCT CCAACGCATT GTTACGTTT CATGAACCTA TTTATTATAT 600

45

GTTCTAAGAA AGAAAAAAT ATCTCAAAGT AAACGTTGGA AATTTTCTGA TGAAGGGAAA 660

TCCAAAGTCT TGGGTTTAGT ATCCCTATGA ATGGTATTTG GAATATGTTT TCGTCAAAC 720

50

AAAAGATTCT TTTCTTTTC ACAAGAGTTA GTGATCAATA ACTTATGCAC TAATTAATGA 780

GATTGGACGT ATACACAATT TGATTATGAT ACTTGAGTAA AAATCACCTG TCCTTTAATT 840

55

TGGAATCTC TCTTTCTTAC CCATTATAT ACTACTTCTT TTCATTAAAA TTAATTTCA 900

ATTATCAATC ATCGTTCAAT TTGATAAAGA TTTAACATTT TTTGTACAG GGCTAGTAAA 960

60

AGCAATCTTT ACATAATTCA TCTTTCTTAC ATATATATAT TACCTTTTTC TTCATTAGTA 1020

65

TTCTATTTGA TTATGATTAT TTTGTACATAA AGCTAGTAAA TTAACACTC GATATGAGAA 1080

ES 2 301 539 T3

1140
 TTATATTACT TCACGCTAAT TAACTCTTAA CACAACAAGA ACTAGTGCAT ATTCAACTTT
 1200
 5 CAAAGCATAT ACTATATATT GAGAATATAG ACCACGAAAG TCAATCAAAA GACCTACCAG
 1260
 CTCTCATCAA GTTCTTTCTT GAAATGATTT TGCAGAATTT CCAACTTAAT TAATTGACA
 1320
 10 TGAATGTGAA AATGTGTGTT GCTCGTTAAG AAAATTGAAT AGAAGTACAA TGAAAATGAT
 1380
 GAGGAATGGG CAAAACACAA AAGAGTTTCC TTTCGTAAC ACAAATTAATT AATGCAAAATC
 1440
 15 TGAGAAAGGG TTCATGGATA ATGACTACAC ACATGATTAG TCATTCCCCG TGGGCTCTCT
 1500
 GCTTTCATTT ACTTTATTAG TTTCATCTTC TCTAATTATA TTGTGCATA TATGATGCAG
 1560
 20 TTCTTTTGTG TAAATTACGT AATATGATGT AATTAATAT CAAAATAAAT ATTCAAATG
 1620
 25 CCGTTGGACT AACCTAATGT CCAAGATTAA GACTTGAACA TARGAATTTT GGAAAACTA
 1680
 AACCAGTIAT AATATATACT CTAAATTGC CATTCTGAA CACAACCAA TAATAATATA
 1740
 30 TACTATTTAC AGTTTTTTTT AATTGGCAAG AACACTGAAA TCTTATTCAT TGTCTCGCTT
 1800
 GGTAGTTGAC AAGTTATAAC ACTCATATTC ATATAACCC CATTCTAACGT TGACGACGAA
 1860
 35 CACTCATATA AACCAACCAA ATTCTTAGCA TATTAGCTAA ATATTGGTTT AATTGGAAAT
 1920
 ATTTTTTTTA TATATAAAAT GCCAGGTAAA TATTAACGAC ATGCAATGTA TATAGGAGTA
 1980
 40 GGGCAATAAA AAGAAAAGGA GAATAAAAAG GGATTACCAA AAAAGGAAAG TTTCCAAAAG
 2040
 GTGATTCTGA TGAGAAACAG AGCCCATACC TCTCTTTTTT CCTCTAAACA TGAAAGAAAA
 2100
 45 ATFGGATGGT CCTCCTTCAA TGCTCTCTCC CCACCCAATC CAAACCCAAC TGTCTTCTTT
 2160
 CTTTCTTTTT TCTTCTTCT AATTGATAT TTTCTACCAC TTAATTCCAA TCAATTTCAA
 2220
 50 ATTTCAATCT AAATGTATGC ATATAGAAT TAATTAAAAG AATTAGGTGT GTGATATTTG
 2280
 55 AGAAAATGTT AGAAGTAATG GTCCATGTTT TTTCTTTCTT TTTCTTCTA TAACACTTCA
 2340
 GTTTGAAAAA AAACCTACCA ACCTTCTGTT TTCTGCAAAT GGGTTTTTAA ATACTTCCAA
 2400
 60 AGAAATATTC CTCTAAAAGA AATTATAAAC CAAAACAGAA ACCAAAAACA AAAAATAAAG
 2460

65

ES 2 301 539 T3

TTGAAGCAGC AGTTAAGTGG TACTGAGATA ATAAGAATAG TATCTTTAGG CCAATGAACA
 2520
 5 AATTAACTCT CTCATAATTC ATCTTCCCAT CCTCACTTCT CTTTCTTTCT GATATAATTA
 2580
 ATCTTGCTAA GCCAGGTATG GTTATTGATG ATTTACACTT TTTTTTAAAA GTTTCCTCCT
 2640
 10 TTTCTCCAAT CAAATTCTTC AGTTAATCCT TATAAACCAT TTCTTTAATC CAAGGTGTTT
 2700
 GAGTGCAAAA GGATTTGATC TATTTCTCTT GTGTTTATAC TTCAGCTAGG GCTTATAGAA
 2760
 15 ATGGAGGGTG GTGCGAGTAA TGAAGTAGCA GAGAGCAGCA AGAAGATAGG GAGAGGGAAG
 2820
 ATAGAGATAA AGAGGATAGA GAACACTACG AATCGTCAAG TCACTTTCTG CAAACGACGC
 2880
 20 AATGGTTTAC TCAAGAAAGC TTATGAGCTC TCTGTCTTGT GTGACGCTGA GGTGCTCTT
 2940
 GTCATCTTCT CCACTCGAGG CCGTCTCTAC GAGTAGGCCA ACAACAGGTA CACATCTTTT
 3000
 25 AGCTAGATCT TGATTTTGTI GAATTTTTTT TCTAGAATAA AGTTTCGACT CTTCTGGTGG
 3060
 30 GTTTTTCAAT CTTTATGGTC TCTTIATAGT TTTTTTCCTT AGTTTCTCTG AAGCTCAAAT
 3120
 CTCTTTAAAA ATCCCCAAA TTAGGGTTTG TTTAAAATA GGGAACCTTA CTTTAACTTC
 3180
 35 TTCTCTTAG TAAAAAAGCA GTGAGGGTCT TCTCTGATCA TTAATTAGCA TCCCCATAC
 3240
 CTGTTCAG TCACTTTTTT TCCACAAATC CTTATAACAG TATCTATATA TGTATCTATT
 3300
 40 TATGTCAGTT TGTACAAGAC ACTTCGATCA ATTTGATGAC CCATCAAGTT TTATTCTGTC
 3360
 AGATTGATCA TTAGGTTTCC ATCATAGTAA TGAAAAAGTA GGGTTCTTGA TAAAAATATA
 3420
 45 ATAATATATA TTATTTGGCT ATATAAAAAA GCTATGTAGA TTCCTTAAAA ATTGATTAC
 3480
 TAGGGAGAGA CTAGTAGGTG TTTGTCTTCT GACACTTCTC TAATCTTTTG GTGAATCCTT
 3540
 50 TTGTTAAATC AAGAAAATGA ATCAGGGACA AAGCTTATTG TTGAGTCACT TAATTAATCA
 3600
 TCCGATCCAT CAATCAAGAA AAATAACGAA ACAGAAAATT TTGATTTTTG ATTGTTATTT
 3660
 55 TCTCCACTTC AAGTTGGGGA CTTGTCAATT CCGTTTTTCT ATACGTTTCC AGCTATTAAC
 3720
 60 AGCTCATGTT CATTTCACCA TTTTGATTAT TTGTCTGCTT TTTAAAGATA AATGTTTCA
 3780
 AAAATATTGT TTTTATTTGC TTGGCTAGTT AATACTATAA TTGAGGTTGA TGTATGACTA

65

ES 2 301 539 T3

3840
 TAATCTATAA GTCAAGTCTC ATATCATGGA TCTAAGTTAA AACTAGTAAA TTTGTAGTTT
 3900
 5 CAATGTGAAC TTTCAACAAC ACTAAAGAAC TGATCTGAAG TTTATAATGG ACATGACTAA
 3960
 TTTGATTAAC AAAAGAGGAA TGCATTATGT ATGTAGAAAC ATGTGATATA TATATGTTTC
 4020
 10 TATTATCAA AGTGTAGTTA ACTTCTTAT TTCAAACACC CTCATGCTTT AGTAGTATCT
 4080
 TACTTTTGAC ATTTCTCAAC TTCAGCTTTC CATTATACAA CAGCACAATG TAAATTACTT
 4140
 15 GTATATGAAT ATGAAAGCAT AACGTTATGC AAAGATTTCT AGCTTTTCTT TTTCTGTTTT
 4200
 GCAAAAGATT TACAAATATC ATGTTCTTGG TAAAAACATA CTTGCCTCAG CCACATATGC
 4260
 20 ATGTAAATGT AATGTTCAA TATTAATTCA GGAAAAACA AGAAGAAGCA AAATTAGCTT
 4320
 CTAGAGTAGG GAATCTATTG ACTTGACCTG AAAATCACTT CTTTTCTTA AAGCCTAGTA
 4380
 25 GTGAATTTT TAATCTAAT AGGCCAAAT ATATACTAGC CTAATATA ATTTGGATT
 4440
 TGTGTCGTAC ATAAATTGGG ACCAATTCCA ATTAACCTAAG AGCATATGCA ATTCAATTC
 4500
 30 TTTTATTTT CTCTCCGAT TTGCTACTTC TTTCTTTGT ATGTTTTCAA ATTAGGATTA
 4560
 CACTTTTTG GGAAGTACA CATTAGGGTC TTCTCGAACT TTGATTATAC ATATATATAT
 4620
 35 ATATATATAT ATATACTTT GTGAGATGTC ACTGTTAATA GATAATAGGC AATAACAATA
 4680
 ATATCCAAA AAGAAGGCGC AAACAAATCA TATACTATAT GGTACTGGTC CATTCACTAT
 4740
 TTTGTGGTT GAATTTAAGG TTTGGCGTAC AAACTTTGT TCAAACCTTT ATTATTCCGT
 4800
 45 CTTTCTGTGT GTTTTGATA TCCAGAAGAT AAAAATATCA ATTTCTTTAA CGACTTCATA
 4860
 TATATATATA TATATATATA TATATATATT TTTCTCTTCT GGTTTTAGTG TTTGAATCCA
 4920
 50 ACAGTTATAG TTTGTTGTTT CTTGTTTTA CTTGTGGTGG TTTAAGTTTG AGATTTTCAC
 4980
 CGATTGCATC TATTTACATA TATAGCTACC ACAAAGAAGA TTGCATTTA AATCTTTTC
 5040
 55 CTTTGTGTA ATGTTGATGA AGTGTGAGAG GAACAATAGA AAGGTACAAG AAGCTTGCT
 5100
 CCGACGCCGT TAACCCTCCG ACCATCACCG AAGCTAATAC TCAGGTAGC TTTAATTAA
 5160
 60 TACACCTAGC TAGCTAGTTC GTTAATFACT TAATTTCTTC TTCTTTTAGT TATCTGACCT
 65

ES 2 301 539 T3

5220
TTTTTTCACC TCTTGTAACA ATGATGGGAT CGAAATTGAT GAAGTACTAT CAGCAAGAGG

5 5280
CGTCTAAACT CCGGAGACAG ATTCGGGACA TTCAGAATTT GAACAGACAC ATTCTTGGTG

10 5340
AATCTCTTGG TTCCTTGAAC TTTAAGGAAC TCAAGAACCT TGAAAGTAGG CTTGAGAAAAG

5400
GAATCAGTCG TGTCCGATCC AAGAAGGTAC ATCACTAACT CTCCATCAAT CTCCTTATCA

15 5460
TTGAATATAT ATCCATCTGA TTCTTGCCCG TTATATTTGG TTTTCTCTC CAGCACGAGA

5520
TGTTAGTTGC AGAGATTGAA TACATGCAA AAAGGGTAAA AGTAAAACCT ATCTTCCTTC

20 5580
ACAATGAACT ACCCCTACTT TATTAGCAAC TTCTCTTTCT GATGATCATC TTTTTTATTT

25 5640
TCTGTTGTCG CTTGCATTGT AGGAAATCGA GCTGCAAAC GATAACATGT ATCTCCGCTC

5700
CAAGGTTTIA TACATAACTC TTTTGGCAT TTTTGATCAT CATTTTTTTC CGGTAGACAA

30 5760
TCTCTTGATG TGCAAATTCT AAATATCTCT GCAGATTACT GAAAGAACAG GTCTACAGCA

35 5820
ACAAGAATCG AGTGTGATAC ATCAAGGGAC AGTTTACGAG TCGGGTGTTA CTTCTTCTCA

5980
CCAGTCGGGG CAGTATAACC GGAATTATAT TGCGGTTAAC CTTCTTGAAC CGAATCAGAA

40 5940
TTCTCCAAC CAAGACCAAC CACCTCTGCA ACTTGTTTGA TTCAGTCTAA CATAAGCTTC

45 6000
TTTCTCAGC CTGAGATCGA TCTATAGTGT CACCTAAATG CGGCCGCGTC CCTCAACATC

6060
TAGTOGCAAG CTGAGGGGAA CCACTAGTGT CATACGAACC TCCAAGAGAC GGTACACAA

50 6120
ACGGGTACAT TGTGATGTC ATGTATGACA ATCGCCCAAG TAAGTATCCA GCTGTGTTCA

55 GAACGTACGT CCGAATTC

60 LISTADO DE SECUENCIAS

<110> Liljegren, Sarah
Yanofsky, Martin F.
The Regents of the University of California

65 <120> Control la dehiscencia del Fruto en Arabidopsis mediante genes INDEHISCENTES 1

ES 2 301 539 T3

<130> 19452A-000700US

<140> WO PCT/US01/11967

5 <141> 2001-04-13

<150> US 09/548,971

<151> 2000-04-13

10

<160> 9

15

<170> PatentIn Ver. 2.1

<210> 1

<211> 3856

20

<212> ADN

<213> *Arabidopsis thaliana*

<220>

25

<223> INDEHISCENTE1 (IND1) genómico

<220>

<221> CDS

30

<222> (2765) .. (3361)_

<223> IND1

<400> 1

35

ctctagacca tctactatcc ggttggtgac ccttaaagct tttgaagact actagaataa 60

tgcaaatacc atatgtccat atccatcctt ttcttttgtt tgaactgaac attctaattt 120

40

tgtaaaagaa aaaaccttat gttaatatca ccgtaggcaa aaaaaatata tcatcatatt 180

aaatttttat tataagatta tacattctct cgttgtaaga gttactccaa ttgcaagtgt 240

45

tgtattaact aataaaaagg acgaaagtag gaagcttata attaattgat gttgcatagt 300

actggtatat tgttgatgaa tataacaagt atgaacatta atgcatgaaa cgggggtattt 360

tgtcttgaac tcattaaagg caatgtgaaa agaagatgtg aggtctcatt ttgaaaattt 420

50

atcttctagc tttgtcgatt ttaaactctat gaaatgaacg caacatatag aaatttcatg 480

tggacaacga catttagacg gtatcttaat tagaccgatt aattagtaat atacttatat 540

55

atataattag tggtgattat aagtttactt atccacttga gaatttaaac aatgggcaat 600

accttaatgt cgaaagaagc cgtccccact tcgtgtaatg agttatgggg gagagatcct 660

60

gttaaatcgt caaataaaac aacttaagaa ctagaattg acaccaaaaa tcataaagag 720

aacggtgaag aagtcattta tcgtatccag ctcatatttc cttagctaaga tcaaatcaag 780

65

ES 2 301 539 T3

gccgttgaag gggcttgtaa gaaaatgtcg aagaaaccgt ggggtttaga agaaagacaa 840
 gaaatagaag aacaatgatg ttaaattgcc tttttgggtg tataggagtt gtcaaaagag 900
 5 gagagagaga agaaaattag gtcaaaataa tgagcactaa aaatggagac atgtgttgag 960
 taactattac aagagcgact tatgcttctt tatggcaatg atatccaaac caaagtgcaa 1020
 10 cgctcctttt ttgccctaatt ttcgtaaagt ctctctctct cttcgtcctt aggaaaaacc 1080
 ctagaatttt aatcccttgt tcttgatctt gctttttgag taacctgat tttgaccaca 1140
 cactatttct tctatctttt gtggcttata ggattttgct ttatatgtgt ttcttgtatt 1200
 15 gctccgtacg tacgtatacg aatttaaagt gttataacaa ggtttatata aactagcaca 1260
 aatgagtcga tgaaatttgt tagcgaaaaa ggtagaataa tattgagctt ttaaaccggca 1320
 20 atatatataa ttttgctgca aaacttagct ttaatcatga tctaatgata ttttctttaa 1380
 tttcctttgc caaattaatc acatgcacgg atttttggca agttatgtgt cgaattcttc 1440
 cattcacaca aactaaact taattagaac tctaggaat attttaaaat gacaacttta 1500
 25 tcgaaaaaaaa tttagttatg aaaacaattc cagaattaa catgagctat ataatttaag 1560
 ataaaatgaa gtaatattga tatgtatgta ataacatctc tgattgcggt aaaaaaaaaac 1620
 atatctgatt aaattgttca tgcaggccca tgcactatg atgtcatcac gtttttattt 1680
 30 tcacaataac taatataat tcaaaaaaat agttttgtca gattaaattt tttttgggtg 1740
 tcagctttct ccaacctact aaactagttt ggaatgttct cttctttatt tttcttttc 1800
 35 ttgatttctt atgtttttta tttatggaat ttttaagacgg attgttttagg tegtctctct 1860
 cttttcttgt tttctaaagt tacttttgta aactcatctc ctccaatta gacagtcaat 1920
 catatagtta tcttttaata tatgtctagt tgataaaaaa aatgaaaaaa tactggtggt 1980
 40 agttctacta atgtttgtgt aaaaaatctg atattatgaa tctaatcaat ttttttgatc 2040
 gtataatgtg ggttaaattt agtaattttt tacataaata agaactgtaa tgttgatgta 2100
 45 tattggggaa tcagtatatt agcttgggta actatacttc tggaaatact tgaagattta 2160
 actatttgca aaattataat ttagtcccga aaaatacaga cgacgggaca cgacaacata 2220
 taagcaggtt tgaatcttgg aaaattttgt atacataacc tatataaata ctaatgttct 2280
 50 ggttgggttc aaaagccttt tcaaaagttc cattttttta attcaaggac attttacata 2340
 ggaaataagt tgagtcataa aaaataatgg ttattttgta aggttttttt tttgattaaa 2400
 55 acgcacatat taagaagtta gttttttttc actaccaaat atcaattaat ttaaaacat 2460
 gcaaccattc ataaaacaat actattaag aatataaata atcacaaaat attaaataca 2520
 cttaaaattt acatataaat ttacaaaaca tctaattaat tgaaacagaa aggaaaagggt 2580
 60 aaaatatact ataaaatgag acatataatc tataaaaaaa aaatgaggca tatgaagtaa 2640

65

ES 2 301 539 T3

ataataagag acatgcatgt aagcattcgg ttaattaatc gagtcaaaga tatatatcag 2700
 taaatacata tgtgtatatt tctggaaaaa gaatatatat attgagaaat aagaaaagat 2760
 5 gaaa atg gaa aat ggt atg tat aaa aag aaa gga gtg tgc gac tct tgt 2809
 Met Glu Asn Gly Met Tyr Lys Lys Lys Gly Val Cys Asp Ser Cys
 1 5 10 15
 10 gtc tgg tcc aaa agc aga tcc aac cac agc ccc aaa aga agc atg atg 2857
 Val Ser Ser Lys Ser Arg Ser Asn His Ser Pro Lys Arg Ser Met Met
 20 25 30
 15 gag cct cag cct cac cat ctc ctc atg gat tgg aac aaa gct aat gat 2905
 Glu Pro Gln Pro His His Leu Leu Met Asp Trp Asn Lys Ala Asn Asp
 35 40 45
 20 ctt ctc aca caa gaa cac gca gct ttt ctc aat gat cct cac cat ctc 2953
 Leu Leu Thr Gln Glu His Ala Ala Phe Leu Asn Asp Pro His His Leu
 50 55 60
 25 atg tta gat cca cct ccc gaa acc cta att cac ttg gac gaa gac gaa 3001
 Met Leu Asp Pro Pro Pro Glu Thr Leu Ile His Leu Asp Glu Asp Glu
 65 70 75
 30 gag tac gat gaa gac atg gat gcg atg aag gag atg cag tac atg atc 3049
 Glu Tyr Asp Glu Asp Met Asp Ala Met Lys Glu Met Gln Tyr Met Ile
 80 85 90 95
 35 gcc gtc atg cag ccc gta gac atc gac cct gcc acg gtc cct aag ccg 3097
 Ala Val Met Gln Pro Val Asp Ile Asp Pro Ala Thr Val Pro Lys Pro
 100 105 110
 40 aac cgc cgt aac gta agg ata agc gac gat cct cag acg gtg gtt gct 3145
 Asn Arg Arg Asn Val Arg Ile Ser Asp Asp Pro Gln Thr Val Val Ala
 115 120 125
 45 cgt cgg cgt cgg gaa agg atc agc gag aag atc cga att ctc aag agg 3193
 Arg Arg Arg Arg Glu Arg Ile Ser Glu Lys Ile Arg Ile Leu Lys Arg
 130 135 140
 50 atc gtg cct ggt ggt gcg aag atg gac aca gct tcc atg ctc gac gaa 3241
 Ile Val Pro Gly Gly Ala Lys Met Asp Thr Ala Ser Met Leu Asp Glu
 145 150 155
 55 gcc ata cgt tac acc aag ttc ttg aaa cgg cag gtg agg att ctt cag 3289
 Ala Ile Arg Tyr Thr Lys Phe Leu Lys Arg Gln Val Arg Ile Leu Gln
 160 165 170 175
 60 cct cac tct cag att gga gct cct atg gct aac ccc tct tac ctt tgt 3337
 Pro His Ser Gln Ile Gly Ala Pro Met Ala Asn Pro Ser Tyr Leu Cys
 180 185 190
 65 tat tac cac aac tcc caa ccc tga tgaactacac agaagctcgc tagctagaca 3391
 Tyr Tyr His Asn Ser Gln Pro
 195
 tttggtgtca tctctcaac ctttttcatg ttgatattt atatatagat gcataaagat 3451
 tcgatccaag attgtatggg tgttttaata ttattattct aagatatatg atgtacaatt 3511
 gtgtaccaag tttctttatc ttgatattcat atgcataaat aattgggtgaa taaaaagaag 3571

ES 2 301 539 T3

atattgattg taaacaaaaa aaagaagata ttgattgta attagggttt gatcattctg 3631
 tatgaaagct ttggcctgca aattaatttt cgatatatat atatatatat ggagaatata 3691
 5 tatcaaatac ttttttaatt tgactataat ttgtatcaat tatctgaatc tgatgagtgt 3751
 aggttatata tggattagca aaaaagaaaa caaccattat tacgcaccta cattaataat 3811
 10 catccaccaa agaagaaacc atcctcaaga gggttccctc tagag 3856

<210> 2

15 <211> 198

<212> PRT

<213> *Arabidopsis thaliana*

20 <223> proteína INDEHISCENTE1 (IND1)

<400> 2

25	Met	Glu	Asn	Gly	Met	Tyr	Lys	Lys	Lys	Gly	Val	Cys	Asp	Ser	Cys	Val
	1				5					10					15	
	Ser	Ser	Lys	Ser	Arg	Ser	Asn	His	Ser	Pro	Lys	Arg	Ser	Met	Met	Glu
			20						25					30		
	Pro	Gln	Pro	His	His	Leu	Leu	Met	Asp	Trp	Asn	Lys	Ala	Asn	Asp	Leu
			35					40					45			
30	Leu	Thr	Gln	Glu	His	Ala	Ala	Phe	Leu	Asn	Asp	Pro	His	His	Leu	Met
	50					55					60					
	Leu	Asp	Pro	Pro	Pro	Glu	Thr	Leu	Ile	His	Leu	Asp	Glu	Asp	Glu	Glu
	65					70					75					80
35	Tyr	Asp	Glu	Asp	Met	Asp	Ala	Met	Lys	Glu	Met	Gln	Tyr	Met	Ile	Ala
					85					90					95	
	Val	Met	Gln	Pro	Val	Asp	Ile	Asp	Pro	Ala	Thr	Val	Pro	Lys	Pro	Asn
				100					105					110		
40	Arg	Arg	Asn	Val	Arg	Ile	Ser	Asp	Asp	Pro	Gln	Thr	Val	Val	Ala	Arg
			115					120					125			
	Arg	Arg	Arg	Glu	Arg	Ile	Ser	Glu	Lys	Ile	Arg	Ile	Leu	Lys	Arg	Ile
	130					135						140				
	Val	Pro	Gly	Gly	Ala	Lys	Met	Asp	Thr	Ala	Ser	Met	Leu	Asp	Glu	Ala
	145					150					155					160
45	Ile	Arg	Tyr	Thr	Lys	Phe	Leu	Lys	Arg	Gln	Val	Arg	Ile	Leu	Gln	Pro
					165					170					175	
	His	Ser	Gln	Ile	Gly	Ala	Pro	Met	Ala	Asn	Pro	Ser	Tyr	Leu	Cys	Tyr
				180					185					190		
50	Tyr	His	Asn	Ser	Gln	Pro										
			195													

<210> 3

55 <211> 2765

<212> ADN

<213> *Arabidopsis thaliana*

60 <220>

<223> región no traducida 5' del promotor de IND1

65

ES 2 301 539 T3

<400> 3

5 ctctagacca tctactatcc ggttggtgac ccttaaagct tttgaagact actagaataa 60
 tgcaaatacc atatgtccat atccatcctt ttcttttgtt tgaactgaac attctaattt 120
 tgtaaaagaa aaaaccttat gttaatatca ccgtaggcaa aaaaaatata tcatcatatt 180
 aaatttttat tataagatta tacattctct cgttgtaaga gttactccaa ttgcaagtgt 240
 tgtattaact aataaaaagg acgaaagtag gaagcttata attaattgat gttgcatagt 300
 actggtatat tgttgatgaa tataacaagt atgaacatta atgcatgaaa cggggatttt 360

10

15 tgtcttgaac tcattaaagg caatgtgaaa agaagatgtg aggtctcatt ttgaaaattt 420
 atcttctagc tttgtcgatt ttaaactctat gaaatgaacg caacatatag aaatttcatg 480
 tggacaacga cathtagacg gtatcttaat tagaccgatt aattagtaat atactttatat 540
 atataattag tgggtgattat aagtttactt atccacttga gaatttaaac aatgggcaat 600
 accttaatgt cgaagaagc cgccccact tcgtgtaatg agttatgggg gagagatcct 660
 20 gttaaatcgt caaataaaa aacttaagaa ctagaattg acaccaaaaa tcataaagag 720
 aacggtgaag aagtcattta tcgtatccag ctcatatttc cttagtaaga tcaaatcaag 780
 gccgttgaaa gggcttgtaa gaaaatgtcg aagaaaccgt ggggtttaga agaaagacaa 840
 gaaatagaag aacaatgatg ttaaattgcc tattttgggtg tataggagt gtcaaaagag 900
 gagagagaga agaaaattag gtcaaaataa tgagcactaa aaatggagac atgtggtgag 960
 25 taactattac aagagcgact tatgcttcct tatggcaatg atatccaaac caaagtgcaa 1020
 cgctcctttt ttgccctaata ttcgtaaagt ctctctcctt ctctcgtcctt aggaaaaacc 1080
 ctagaatttt aatcccttgt tcttgatcct gctttttgag taaccatgat tttgaccaca 1140
 cactatttct tctatctttt gtggtctata ggattttgct ttatatgtgt ttcttgtatt 1200
 gctccgtacg tacgtatacg aatttaaatg gttataacaa ggtttatata aactagcaca 1260
 30 atagagtcca tgaaatttgt tagcgaaaaa ggtagaataa tattgagtct ttaaaccgca 1320
 atatatataa ttttgctgca aaacttagct ttaatcatga tctaatagata ttttctttaa 1380
 tttcctttgc caaattaatc acatgcacgg atttttggca agttatgtgt cgaattcttc 1440
 cattcacaca aactaaact taattagaac tctaggaat attttaaat gacaacttta 1500
 35 tcgaaaaaaaa tttagttatg aaaacaatc cagaattaaa catgagctat ataatttaag 1560
 ataaaatgaa gtaatatgta tatgtatgta ataacatata tgattgagggt aaaaaaaaaac 1620
 atatctgatt aaattgttca tgcaggcca tgctactatg atgtcatcac gtttttattt 1680
 tcacaataac taatatatat tcaaaaaaat agttttgtca gattaaattt tttttgggtg 1740
 tcagctttct ccaacctact aaactagttt ggaatgttct ctctcttatt tttctttttc 1800
 40 ttgatttctt atgtttttta tttatggaat ttttaagacgg attgtttagg tcgtttctct 1860
 cttttcttgt tttctaaagt tacttttgta aactcatctc ctccaatta gacagtcaat 1920
 catatagtta tcttttaata tatgtctagt tgataaaaa aatgaaaaaa tactggtggt 1980
 agttctacta atgtttgtgt aaaaaatctg atattatgaa tctaatacaat ttctttgatc 2040
 45 gtataatgtg ggttaaattt agtaattttt tacataaata agaactgtaa tgttgatgta 2100
 tattggggaa tcagtatatt agctgggta actatacttc tggaaatact tgaagattta 2160
 actatttgca aaattataat ttagtcccga aaaatacaga cgacgggaca cgacaacata 2220
 taagcagggt tgaatcttgg aaaattttgt atacataacc tatataaata ctaatgttct 2280
 ggttgggttc aaaagccttt tcaaaagttc cattttttta attcaaggac attttacata 2340
 50 ggaaataagt tgagtcataa aaaataatgg ttattttgta aggttttttt tttgattaaa 2400
 acgcacatat taagaagtta gttttttttc actaccaaat atcaattaat ttaaaacct 2460
 gcaaccattc ataaaacaat actattaag aatataaata atcacaaaat attaaataca 2520
 cttaaaattt acatataaat ttacaaaaca tctaattaat tgaacagaa aggaaaaggt 2580
 55 aaaatatatc ataaaatgag acatataatc tataaaaaaa aaatgaggca tatgaagtaa 2640
 ataataagag acatgcatgt aagcattcgg ttaattaatc gagtcaaaga tatatatcag 2700
 taaatacata tgtgtatatt tctggaaaaa gaatatatat attgagaaat aagaaaagat 2760
 gaaaa 2765

60 <210> 4

<211> 496

<212> ADN

65 <213> *Arabidopsis thaliana*

ES 2 301 539 T3

<220>

<223> región no traducida 3' del promotor de IND1

5 <400> 4

```
10      atgaactaca cagaagctcg ctagctagac atttgggtgc atcctctcaa cctttttcat 60
      gttgatatat tatatataga tgcataaaga ttcgatccaa gattgtatgg gtgttttaat 120
      attattattc taagatatat gatgtacaat tgtgtaccaa gtttctttat cttgatatca 180
      tatgcataaa taattgggtga ataaaaagaa gatattgatt gtaaacaaaa aaaagaagat 240
      attgattggt aattaggggt tgatcattct gtatgaaagc tttggcctgc aaattaattt 300
      tcgatatata tatatatata tggagaatat atatcaaata cttttttaat ttgactataa 360
15      tttgtatcaa ttatctgaat ctgatgagtg taggttatat atggattagc aaaaaagaaa 420
      acaaccatta ttacgcacct acattaaaaa tcatccacca aagaagaaac catcctcaag 480
      agggttccct ctagag                                     496
```

20

<210> 5

<211> 5622

25 <212> ADN

<213> *Arabidopsis thaliana*

<220>

30 <223> SHATTERPROOF1 (SHP1) genómico

<220>

<221> base modificada

35 <222> (935).._(941)_

<223> n = cualquier nucleótido

40

45

50

55

60

65

ES 2 301 539 T3

<400> 5

5 agatctgcaa cagtgaaaag agaaaacaaa atggacttga agaggttttg acaatgccag 60
 agataatgct tattccctaa tatggtgcca gccaaagtgc aaattggctt tttaaatatg 120
 gatttctgta tcagtggtca tatttgtgga tccaacgtat tcatcatcaa gttctcaagt 180
 ttgctttcag tgcaattcta attcacacgt ttaactttaa catgcatgtc attataatta 240
 10 cttcttcaact aagacacaat acggcaaac tttcagatta tattaatctc cataaatgaa 300
 ataattaacc tcataatcaa gattcaatgt ttctaaatat atatggacaa aatttacacg 360
 gaagattaga tacgtatatt agtagattta gtctttcggt tgtgcgataa gattaaccac 420
 ctcatagata gtaatatcat tgtcaaattc ctctcggttt agtcgctaaa ttgtatcttt 480
 tttaaagccta aaagtagtgt attcgcataat gacttatcgt cctaactttt tttttaatta 540
 acaaaaaaat cgaaaagaaa ataatctggt aaatatTTTT taagtactcc attaagttta 600
 15 gtttctattt aaaaaatgct tgaaatttga cagttatggt caacaatttt gaatcatgag 660
 cgatgtctag atactcagaa tttaatcaag atgtcttata aattttgttg tcaactcagg 720
 acccacgcaa agaaaaagac taatatgatt tttatttggg ctggatattt ttgtagagga 780
 tgaaaactaag agagtgaag atctgaaatc cacaatggtc aagagagctc aaagcaaaaa 840
 gaaaaatgaa gatgaaggac taaagaacaa taagcaacta cttataccct atttcataa 900
 20 aggattcagg tactaggaga agttgaggca agttnnnnnn nattgattca aattttcatt 960
 tatttttaca atttaattca cctaagttat tatgcatttc tcatcattgg tacattttct 1020
 gtatagcgta tttacatata tgaaataaat taaatatgct ctacggttgc aagtagttaa 1080
 tgaatgtccc cacgcaaaaa aaaatccctc caaatatgct caccttttct tttcttttta 1140
 attccaaaaat taccataaac ttttggttta caaaagattt ctagaatttg aggaagatat 1200
 25 cctaaatgat tcatgaaatcc ttcaataatc tgaagtttgc gatattttcg attttcttca 1260
 agagttgcca tatttghtaatt ttggtgacct taaacttttt ttgataaaga gtaaactgtt 1320
 tttcttaaaa gtaaaacttg attttatggt ttagggttct agctcaactt tgtattatat 1380
 ttcttgcaaa aagagttcgt taactgcatt cttcaactat ataaagtgat tatcaaaaaac 1440
 atcttcatga acattaagaa aaacaatatt tgggtttcgg tagagcttgg ttttgcttgg 1500
 30 cttgatccac ataccattc tagactttgg cataaatttg atacgataga gagtatctaa 1560
 tggtaatgca gaagggtaaa aaaaggaaga gagaaaagg gagaaagatt accaaaaata 1620
 aggagtttca aaagatgggt ctgatgagaa acagagccca tcctctcct tttcccctc 1680
 ccatgaaaga aatcggatgg tctccttca atgtcctcca cctactctc tcttcttct 1740
 tttttcttt cttattatta accatttaat taatttcccc ttcaatttca gtttctagtt 1800
 35 ctgtaaaaaag aaaatacaca tttcacttat agatattcat atctatttat atgcatttat 1860
 agagaataaaa aaagtgtgag tttctaggta tggttgagat gtgctgtttg gacaattgtt 1920
 agatgatctg tccatttttt tcttttttct tctgtgtata aatatatttg agcacaaga 1980
 aaaactaata accttctgtt ttcagcaact agggctctat aaccttcaa gaaatattcc 2040
 40 ttcaattgaa aaccataaa ccaaaataga tattacaaaa ggaagagag atattttcaa 2100
 gaacaacata attagaaaag cagaagcagc agttaagtgg tactgagata aatgatatag 2160
 tttctcttca agaacagttt ctcatccc acctctcct ttttgctgat ctatcgtaat 2220
 cttgagaact caggttaagg tgtgaatatt atgcaccatt catlaaccct aaaaaataga 2280
 gatttaaaat aaatgtttct tctttctctg attcctgtgt aaccaattca tgggtttgat 2340
 45 atgtttcttg gttattgctt atcaacaaag agatttgatc attataaagt agattaataa 2400
 ctcttaaaaa cacaaagttt ctttattttt tagttacatc cctaattcta gaccagaaca 2460
 tggatttgat ctatttcttg gttatgtatc ttgatcagga aaagggattt gatcatcaag 2520
 attagccttc tctctctctc tctagatata tttcttgaat ttagaaatct ttatttaatt 2580
 atttgggtgat gtcatatatg gatcaatgga ggaaggtggg agtagtcaag acgcagagag 2640
 50 tagcaagaaa ctagggagag ggaaaataga gataaagagg atagagaaca caacaaatcg 2700
 tcaagttact ttctgcaaac gacgcaatgg tcttctcaag aaagcttatg aactctctgt 2760
 cttgtgtgat gccgaagttg ccctcgtcat cttctccact cgtggccgtc tctatgagta 2820

55

60

65

ES 2 301 539 T3

cgccaacaac aggtacgctt ctctactct atttcttgat cttgttttct taattttaac 2880
 taaacaagat cctagttcaa atgataacaa agtggggatt gagagccaag attagggttt 2940
 ggtaattta gaaaaccaga tttcacttgg tgatacattt aatatctctc tagctagatt 3000
 5 tagtactctc tctctatat atgtgtgggt gtgtgtgtaa gtgtgtatat gtatgcaaat 3060
 gcaagaagaa gaagaaaaag ttatcttgtc ttctcaaatt ctgatcagct ttgaccttag 3120
 tttcactctt ttttctgcaa atcatttgaa cctgatgcat gtcagtttct acaatacact 3180
 ttttaattttg acggcccatc aaatttccca gggtttactt cagtgaacaa aattgggttc 3240
 ttgacacgat ttagcatgta tatataaaaa taggggatga tcaagactta tgtaacctct 3300
 10 gtctggtgaa actagggaca aagtctactg atgagttgtc actagggatc catttgatca 3360
 tttaatccca acaaaaatga acaaaaattt tgagaattta tatgctgaag tttttcaacc 3420
 ctctttttta aataacttta tattatgtag atttgtattt agggtaattt gtccaactag 3480
 aagtcctaaa aatcaataaa cacacggatg actttgtcta acattgtatc agtcatcaaa 3540
 tgtaaaattg tacaataaat gaaattaaag atttagtctc ttttattttt tttgttttag 3600
 15 gtgtatata atatatatgt gtataattgt tgcattgata tatcaatgag agggagagaa 3660
 ctgagagaag tgcgggaaat taaaatggtc gcagccaatt ggaatctctg gcatctctag 3720
 ctctatttgt ttgttattag aaaaaaaaaa aaaaaatcct ttaaagatac ctctcatgatg 3780
 acattgaatc atgtaatata cacgatacat ggtctaattc ctctcaaac cctaattacc 3840
 aatttcgaaa ccataatatt tactagtatg tttatatatc ctactttaa gacattgttt 3900
 20 gtttataata ccttgtgaat taagaaaaaa aaaaaaaaaa ttgtggatct attcaagcca 3960
 tgtgttagaa taaatttata aattttctcc tcgtactggg cagatattgg tccaaactcc 4020
 aaagccttcc cttttcagga aaaaaaacat ttcgaaatta actctaatta atcaagaatt 4080
 tcttacaatg tatacatcta atgttttttc cgcatctta ctattagtg tgaggggtac 4140
 25 aattgaaagg tacaagaaag cttgttccga tgccgtcaac cctccttccg tcaccgaagc 4200
 taatactcag gtaccaattt atattgtttg attctctttg ttttatcttc ttttttcat 4260
 tatatatgt atacaacaaa aatataacct acaaaaagag agagtcaag gaaatgcatt 4320
 gaaacggttt cgttatggtg tttgaataca tggatttttg aagtactatc agcaagaagc 4380
 ctctaagctt cggaggcaga ttcgagatat tcagaattca aataggtaat tcattaactt 4440
 30 ttcatgaact cttcgatttg gtattaggtc acttaatttg gtgtcgggtc aaaagtccgc 4500
 ttgtagtttt ctttagaagt tgttttgttt aatgttcatg tttacaaatt gaaggcatat 4560
 tgttggggaa tcacttggtt ccttgaactt caaggaaactc aaaaacctag aaggacgtct 4620
 tgaaaaagga atcagccgtg tccgctccaa aaaggtaaaa tctacgttgc tctctctctg 4680
 35 tgtctctgtc tctctctcta tatatagtec cttagtttat atagttcatc acccttttgt 4740
 gagaattttg cagaatgagc tgtttagtggc agagatagag tatatgcaga agagggtaag 4800
 aacgtttctc ccattccaag taattagatc tttctctgtc tttgtgaggg tttgagtttt 4860
 cccataaatac atgtgttaga aatggagtgt caacacaata acatgtacct gcgagcaaa 4920
 40 gttatgccag ttctgttcca aatcttaate tcaatatcta ctcttttctt cattgtataa 4980
 ctaagataac gtgaataaca agaaaacttt tgtttttggg tttaatagat agccgaaggc 5040
 gccagattga atccggacca gcaggaatcg agtgtgatac aagggacgac agtttacgaa 5100
 tccggtgtat cttctcatga ccagtcgcag cattataate ggaactatat tccggtgaac 5160
 cttcttgaac cgaatcagca attctccggc caagaccaac ctctcttca acttgtgtaa 5220
 ctcaaaacat gataacttgt ttcttcccct cataacgatt aagagagaga cgagagagt 5280
 45 cattttatat ttataacgcg actgtgtatt catagtttag gttctaataa tgataataac 5340
 aaaactgttg tttctttgct taattacatc aacatttaaa tccaaagttc taaaacacgt 5400
 cgagatccaa agtttgtcat acaagattag acgcatacac gatcagttaa tagattttaa 5460
 gtgcctttta atatttacat atagttgcag cttcgattag atcatgtcca ccaaacactc 5520
 50 acaattagag acaagcaaaa ctataaacat tgatcataaa atgattacaa catgtccata 5580
 aattaattat ggattacaaa aataaaaact taaaaaagat ct 5622

<210> 6

55 <211> 6138

<212> ADN

<213> *Arabidopsis thaliana*

60 <220>

<223> SHATTERPROOF2 (SHP2) genómico

65

ES 2 301 539 T3

<400> 6

5 gaattcgtaa cagaatttag tgaataatat tgtaattacc aggcaaggac tctccaaacg 60
gatagctcga atatcgttat taaagagtaa atgatccaat atgtaagcca ttgttgatca 120
tctaacattg ttggactctc tattgctcga aatgatgcat acctaatacat ttattcagtt 180

10 aactatcaag ttgcatttgt aaaaacccaa catttaaat cagattttgat atcacttaca 240
gaggatagag aagcatgact ccaggcctgc atgcaacaag aaaaaggaag aaaataatgt 300
taaaaatttg acaaatatag tgtttattct tattatatga gacagaattt gaataaaatc 360
ctaccaact agagcatcaa acggttttgc aatcgcaata atgsaaccca tttctttttt 420
gagtttttac tcttctttca acagaaact tctcaaacgt ctttagcact gtgacgtag 480
atatatacac aaaagcttga aatttcttca agcaaaagaa tctttgtggg agttaaggca 540
acaagccagg taaagaatct ccaacgcatt gttacgtttt catgaacctt tttattatat 600
gttctaagaa agaaaaaat atctcaagat aaacgtttgg aattttctga tgaagggaaa 660
15 tccaaagtct tgggttttagt atccctatga atgggtattg gaatatgttt tctgcaaaac 720
aaaagattct tttctttttc acaagagtta gtgatcaata acttatgcac taattaatga 780
gattggacgt atacacaatt tgattatgat acttgagtaa aaatcacctg tcctttaatt 840
tgggaattct tcttctttac ccattttctt ttcattttaa ttaaatttca 900
attatcaatc atcgttcaat ttgataaaga tttaacattt tttgtcacag ggctagtaaa 960
agcaatcttt acataattca tcttctttac atatatatat tacctttttc ttcattagta 1020
20 ttctatttga ttatgattat ttgtcataa agctagtaaa ttaaacactc gatattagaa 1080
ttatattact tcaacgtaat taactcttaa cacaacnaga actagtgcac attcaacttt 1140
caaagcatal actatatatt gagaatatag accacgaaag tcaatcaaaa gacctaccag 1200
ctctcatcaa gttctttctt gaaatgattt tgcagaattt ccaacttaat taattcgaca 1260
tgaatgtgaa aatgtgtggt gctcgtttagt aaaaattgaa agaagtacaa tgaattatgt 1320
25 gaggaaatggg caaaacacaa aagagtttcc tttcgttaact acaatttaatt aatgcaaatc 1380
tgagaaaggg tcatgggata atgactacac acatgattag tcaatccccg tgggctctct 1440
gctttcattt actttattag tttcatcttc tctaattata ttgtcgcata tatgatgcag 1500
tcttttgtc taattacgt aatatgattg aatttaattat caaaataaat attcaaatgt 1560
ccgttggact aacctaatgt ccaagattaa gacttgaaca taagaatttt ggaaaaacta 1620
aaccagttat aatataact cttaaatgtc catttctgaa cacaccaaa taataatata 1680
30 tactatttac agtttttttt aattggcaag aacactgaaa tcttattcat tgcctcgtt 1740
ggtagttgac aagttataac actcatattc atataacccc attctaactg tgacgacgaa 1800
cactcatata aaccacccaa attcttagca tattagctaa atattggttt aattggaaat 1860
atftttttta tatataaaat gccaggtaaa tattaacgac atgcaatgta tataggagta 1920
gggcaataaa agaaaaagga gaataaaaaa ggattaccaa aaaaggaag tttccaaaag 1980
35 gtgattctga tgagaaacag agccatacc tctctttttt cctctaama tgaagaaaa 2040
attggatggt cctccttcaa tgcctctctc ccacccaatc caaacccaac tgccttcttt 2100
ctttcttttt tcttctttct aatttgatat tttctaccac ttaattccaa tcaattttca 2160
atftcaatct aaatgtatgc atatagaatt taattaaaag aattaggtgt gtgatatttg 2220
agaaaatggt agaagtaatg gtccatgttc tttctttctt tttccttcta taacacttca 2280
gtttgaaaaa aaactaccaa accttctgtt ttctgcaaat gggtttttaa atacttccaa 2340
40 agaataatc ctctaaaaga aattataaac caaacagaaa accaaaaaca aaaaataaag 2400
ttgaagcagc agttaagtgg tactgagata ataagaatag ccaatgaaca gatataatta 2460
aatcaactc ctcataatc atcttcccat cctcacttct ctctctttct gatataatta 2520
atcttgctaa gccaggatg gttattgatg atttacactt ttttttaaaa gtttcttct 2580
ttctccaat caaattcttc agttaatcct tataaacat tctttaaate caagggtgtt 2640
gagtgcaaaa ggatttgatc tatttctctt gtgtttatac ttcagctagg gcttatagaa 2700
45 atggaggggtg gtgogagtaa tgaagtagca gagagcagca agaagatagg gagaggggaa 2760
atagagataa agaggataga gaacactacg aatcgtcaag tcaactttctg caaacgacgc 2820
aatggtttac tcaagaaagc ttatgagctc tctgtcttgt gtgacgctga ggttgcctct 2880
gtcatcttct ccaactgagg ccgtctctac gagtacgcca acaacaggta cacatctttt 2940
agctagatct tgattttggt gaattttttt tctagaataa agtttgcact ctctgggtgg 3000
gtttttcaat ctttatgggt tctttatagt tttttctctt agtttctctg aagctcaaat 3060
50 ctctttaaaa atccccaaaa ttagggtttg tttaaaacta gggaaacctt ctttaacttc 3120
tttctcttag taaaaaagca gtgagggtct tctctgatca ttaattagca tccccatc 3180
cttgttccag tcaacttttc tccacaaatc cttataacag tatctatata tgtatctatt 3240
tatgtcagtt tgtacaagac acttcgatca atttgatgac ccatcaagtt ttatttctgc 3300
agattgatca ttagggtttc atcatagtaa tgaaaaagta gggttcttga taaaattata 3360
55 ataatatata ttatttggct atataaaaa gctatgtaga ttccttaaaa attgattcac 3420
tagggagaga ctagtaggtg tttgtcttct gacacttctc taatcttttg gtgaatcctt 3480
ttgttaaatc aagaaaaatga atcagggaca aagcttattg ttgagtcact taattaatca 3540
tccgatccat caatcaagaa aaataacgaa acagaaaatt ttgatttttg atgttattc 3600
tctccacttc aagtggggga ctgtcattt cegtttttct atacgtttcc agctattaac 3660
agctcatggt catttcacca ttttgattat ttgtctgctt ttaaaagata aatgttttca 3720
60 aaaaatttgt ttttatttgc ttggctagt aatactataa ttgaggttga tgtatgacta 3780
taatctataa gtcaagctc atatcatgga tctaagttaa aactagtaaa tttgtagttt 3840

ES 2 301 539 T3

caatgtgaac tttcacaacg actaaagaac tgatctgaag tttataatgg acatgactaa 3900
 tttgattaac aaaagaggaa tgcattatgt atgtagaac atgtagatata tatatgttc 3960
 tattatcaaa agtgtagtta actttcttat ttcaaacacc ctcatgcttt agtagtatct 4020
 5 tacttttgac atttctcaac ttcagctttc cattatacaa cagcacaatg taaattactt 4080
 gtatatgaat atgaaagcat aacgttatgc aaagatttct agcttttctt tttctgtttt 4140
 gcaaaagatt tacaaatatac atgttcttgg taaaaacata cttgcctcag ccacatatgc 4200
 atgtaaagt aatgttcaaa tattaattca ggaaaaacaa agaagaagca aaattagctt 4260
 ctagagtagg gaatctattg acttgacctg aaaatcactt ctttttctta aagcctagta 4320
 10 gtgaattttt taatctaatt aggcctaaat atatactagc ctaaaatata atttggattt 4380
 tgtgtcgtac ataaaattggg accaattcca attaactaag agcatatgca attcaaattc 4440
 tttttatttt cttctccgat ttgctacttc tttcttttgt atgttttcaa attaggatta 4500
 cacttttttg gggaggtaca cattagggtc ttctcgaact ttgattatac atatatatat 4560
 atatatatat atataacttt gtgagatgtc actgttaata gataataggc aataacaata 4620
 15 atatccaaaa aagaaggcgc aaacaaatca tatactatat ggtactggtc cattcactat 4680
 tttgtcgggt gaatttaagg tttggcgtac aaactttgtt tcaaactttt attattccgt 4740
 ctttctgtgt gttttgtata tccagaagat aaaaatatca atttctttaa cgacttcata 4800
 tatatatata tatatatata tatatatatt tttctcttct ggtttttagtg tttgaatcca 4860
 20 acagttatag tttcgtgtgt ctttgtttta cttgtgggtgg ttttaagttt agattttcac 4920
 cgattgcac tatttacata tatagctacc acaaaaaaga ttgcatttta aaatcttttc 4980
 ctttgtgtga atgttgatga agtgtgagag gaacaataga aaggtaacaag aaagcttgct 5040
 ccgacgccgt taaccctccg accatcaccg aagctaatac tcaggttagc ttttaattaa 5100
 tacacctagc tagctagttc gtttaattact taatttcttc ttcttttagt tatctgacct 5160
 25 ttttttcacc tcttgtaaca atgatgggat cgaaattgat gaagtactat cagcaagagg 5220
 cgtctaaact ccggagacag attcgggaca ttcagaattt gaacagacac attcctgggtg 5280
 aatctcttgg ttctttgaa ttttaaggaac tcaagaacct tgaagtagg cttgagaaag 5340
 gaatcagtcg tgtccgatcc aagaaggtaac atcactaact ctccatcaat ctcttatca 5400
 30 ttgaatata atccatctga ttcttggccg ttatatattg tttttctctc cagcacgaga 5460
 tgttagttgc agagattgaa tacatgcaaa aaagggtaaa agtaaaacct atcttcttc 5520
 acaatgaact accctactt tattagcaac ttctctttct gatgatcacc ttttttattt 5580
 tctgttgcg cttgcattgt aggaaatcga gctgcaaaac gataacatgt atctccgctc 5640
 caaggtttta tacataactc tttttggcat ttttgatcat ctttttttc cggtagacaa 5700
 35 tctcttgatg tgcaaattct aaatatctct gcagattact gaaagaacag gtctacagca 5760
 acaagaatcg agtgtgatac atcaagggac agtttacgag tcgggtgta cttcttctca 5820
 ccagtcgggg cagtataacc ggaattatat tgcgggtaac cttcttgaac cgaatcagaa 5880
 ttctccaac caagaccaac cacctctgca acttgtttga ttcagtcctaa cataagcttc 5940
 40 tttctcagc ctgagatcga tctatagtgt cacctaaatg cggccgcgct cctcaacatc 6000
 tagtcgcaag ctgaggggaa ccactagtgt catacgaacc tccaagagac ggttacacaa 6060
 acgggtacat tgttgatgtc atgtatgaca atcgcccaag taagtatcca gctgtgttca 6120
 gaacgtacgt ccgaattc 6138

45 <210> 7

<211> 51

<212> PRT

50 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: dominio de unión a ADN hélice-bucle-hélice (bHLH) básico de IND1

55

<400> 7

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

ES 2 301 539 T3

<210> 8
<211> 26
<212> ADN
5 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: cebador para la amplificación de región genómica IND1 de ADNc
10
<400> 8

gatgaaaatg gaaaatggta tgtata 26

15
<210> 9
<211> 23
<212> ADN
20 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: cebador para la amplificación de región genómica IND1 de ADNc
25
<400> 9

gttcatcagg gttgggagtt gtg 23

30

35

40

45

50

55

60

65