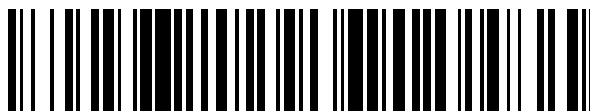


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 650 721**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C12P 19/34 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.03.2012 PCT/US2012/031334**

87 Fecha y número de publicación internacional: **11.10.2012 WO12138549**

96 Fecha de presentación y número de la solicitud europea: **30.03.2012 E 12767530 (4)**

97 Fecha y número de publicación de la concesión europea: **25.10.2017 EP 2694682**

54 Título: **Análisis de alto rendimiento de los bordes transgénicos**

30 Prioridad:

05.04.2011 US 201161471833 P

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

22.01.2018

73 Titular/es:

**DOW AGROSCIENCES LLC (100.0%)
9330 Zionsville Road
Indianapolis IN 46268-1054, US**

72 Inventor/es:

**CAO, ZEHUI;
NOVAK, STEPHEN y
ZHOU, NING**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 650 721 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

Análisis de alto rendimiento de los bordes transgénicos

Campo técnico

5 La presente invención se refiere en general a los campos de biología molecular y bioquímica de las plantas. La presente invención se refiere a un método modificado de Reacción en Cadena de Polimerasa (PCR) para analizar los bordes transgénicos y determinar la secuencia de cromosomas que flanquea el transgén.

Antecedentes

10 La determinación de la localización genómica y la secuencia flanqueante cromosómica adyacente al transgén insertado es técnicamente desafiante. Se han desarrollado varios métodos para superar la limitación de identificar las secuencias de DNA desconocidas que flanquean una secuencia de DNA conocida. Sin embargo, estos métodos tradicionales de PCR para la identificación de las secuencias cromosómicas genómicas que flanquean un transgén conocido tales como LM-PCR (también descrito como Genome Walking) y otros métodos que incluyen: PCR inversa (i-PCR), PCR entrelazada térmica asimétrica (TAIL-PCR), PCR anclada (a-PCR) y PCR de cebado aleatorio (rm-PCR) se ven obstaculizados por la baja sensibilidad de detección (requiriendo grandes cantidades de la plantilla de DNA) o baja especificidad debido a las pérdidas de DNA durante la preparación.

15 La reacción en cadena de polimerasa (PCR) es un método de biología molecular comúnmente empleado. El método se lleva a cabo mediante la desnaturalización de la plantilla de DNA bicatenario, hibridación de los cebadores de oligonucleótidos a la plantilla de DNA, y la extensión de una cadena de DNA mediante una DNA polimerasa. Los cebadores de oligonucleótidos están diseñados para hibridar las cadenas opuestas del DNA y posicionarse de modo que la cadena de DNA producida por la DNA polimerasa sirva como una cadena plantilla para el otro cebador. Cada ciclo se repite, dando como resultado una amplificación exponencial de un fragmento de DNA (Mullis y colaboradores. Documentos de Patente U.S. Pat. No. 4,683,195, 4,683,202, y 4,800,159). El uso de PCR por los expertos en la técnica es fundamental para la amplificación y aislamiento de los fragmentos de DNA para el posterior análisis.

20 El aislamiento y análisis de las plantillas de DNA mediante la reacción en cadena de polimerasa (PCR) requiere el conocimiento de las secuencias de DNA flanqueantes. Desafortunadamente, este requerimiento limita la amplificación por PCR a regiones de secuencias de DNA conocidas. El uso de las metodologías de PCR para identificar la localización de un sitio transgénico dentro de un genoma se ve obstaculizado por la inserción aleatoria del transgén en un sitio cromosómico desconocido dentro del genoma de un organismo. Los métodos para identificar las secuencias de DNA desconocidas que se localizan adyacentes a una secuencia de DNA conocida son necesarios para la identificación de un sitio transgénico dentro del cromosoma de un organismo. Además, tales métodos se pueden utilizar para identificar nuevas secuencias de genes para identificar nuevos rasgos, para determinar el sitio genómico de un transposón o la secuencia viral que se ha insertado en el genoma de un organismo, o para identificar el sitio cromosómico de las secuencias de polinucleótidos insertadas en el genoma mediante la mutagénesis de inserción.

25 Se han desarrollado varios métodos para superar la limitación de las secuencias de DNA desconocidas que flanquean una secuencia de DNA conocida. Un método de PCR mediado por ligación (LM PCR) en donde se genera una genoteca y se hibridan los adaptadores a los fragmentos de DNA para la amplificación por PCR se comercializa como GENOME WALKER UNIVERSAL KIT™ (véase documento de Patente U.S. Pat. No. 5,565,340 y documento de Patente U.S. Pat. No. 5,759,822). Otro método utilizado comúnmente es la reacción por PCR inversa (véase Silver y Keeridatte (1989), J. Virol., 63: 1924-1928), en donde el DNA se digiere con una enzima de restricción y se auto une dando como resultado un círculo contiguo. La amplificación por PCR utilizando cebadores de oligonucleótidos que se unen a secuencias conocidas da como resultado la amplificación y elucidación de las secuencias flanqueantes desconocidas. Desafortunadamente estos métodos son ineficaces y llevan mucho tiempo. Estos y otros métodos tradicionales de PCR (incluyendo PCR entrelazada térmica asimétrica [TAIL-PCR], PCR anclada [a-PCR] y PCR de cebado aleatorio [rm-PCR]) se ven obstaculizados por la baja sensibilidad de detección (que requiere grandes cantidades de plantilla de DNA) o por la baja especificidad debido a las pérdidas de DNA durante la preparación. El documento de Patente WO0023620 describe el uso de ligadura de adaptador para la toma de la huella dactilar del DNA. Rosenthal y colaboradores. Nucleic Acid Reserach, 19, pp. 5395-5401, 1991, describen el uso de adaptadores monocatenarios para PCR walking.

30 El desarrollo de un método que pueda mejorar la sensibilidad de la detección por purificación de los fragmentos de DNA cromosómico que contienen tanto las secuencias de DNA conocidas como desconocidas puede dar como resultado un método sensible para detectar y caracterizar regiones de DNA desconocidas que se localizan adyacentes a una secuencia de DNA conocida. El desarrollo del método de la reacción en cadena de polimerasa mediada por amplificación lineal (LAM PCR) logra estos objetivos. Véase el documento de Patente U.S. Pat. No 6,514,706. El método LAM PCR es particularmente adecuado para amplificar y analizar fragmentos de DNA, cuya secuencia solo se conoce en parte.

35 El desarrollo de un método que pueda mejorar la sensibilidad de la detección mediante la purificación de los

fragmentos de DNA cromosómico que contienen ambas secuencias de DNA conocidas y desconocidas puede dar como resultado un método sensible para la detección y caracterización de las regiones de DNA desconocidas que se localizan adyacentes a una secuencia de DNA conocida. El desarrollo del método LAM PCR losgra estos objetivos. LAM PCR es un método de PCR modificado que se utiliza para analizar secuencias cromosómicas flanqueantes desconocidas localizadas adyacentes a una secuencia de DNA conocida. El método LAM PCR se puede utilizar para identificar y/o secuenciar una secuencia de DNA o RNA desconocida que flanquea una región de DNA o RNA conocida.

El método LAM PCR consiste en las siguientes etapas. Se realiza una reacción de extensión del cebador utilizando un DNA cromosómico como plantilla y un cebador de oligonucleótido que se une a una secuencia de DNA conocida dentro del DNA cromosómico. El cebador de oligonucleótido es complementario a una secuencia de repetición terminal larga (LTR), que es una secuencia característica de un retrovirus, y se marca con biotina en el extremo del cebador de oligonucleótido. El producto de DNA monocatenario de la PCR lineal se une a perlas magnéticas que tienen estreptavidina inmovilizada. Esta etapa sirve para aislar el fragmento de DNA monocatenario amplificado que contiene la secuencia LTR conocida y una secuencia desconocida derivada del cromosoma. El DNA monocatenario se convierte en DNA bicatenario sintetizando la cadena complementaria. El DNA bicatenario se escinde con una enzima de restricción que reconoce una secuencia y escinde el DNA bicatenario en la secuencia. Un DNA bicatenario llamado casete enlazador se une al extremo. Las reacciones PCR posteriores se llevan a cabo utilizando el producto de unión así obtenido como una plantilla, así como un cebador complementario al TLR y un cebador complementario al casete enlazador. Se amplifica un fragmento de DNA que contiene el LTR y el DNA cromosómico que flanquea la secuencia adyacente al LTR. Como resultado se puede determinar el sitio de integración del retrovirus previamente desconocido.

El método LAM PCR se considera actualmente como un sistema efectivo para analizar secuencias de DNA desconocidas adyacentes a una secuencia de DNA conocida. Las modificaciones y mejoras del método LAM PCR se han descrito en la técnica. Véase documento de Patente U.S. Pat. App US2007/0037139 y Harkey y colaboradores, (2007) *Stem Cells Dev.*, Junio; 16(3): 381-392.

El método LAM PCR se modificó en el documento de Patente U.S. Pat. App. US2007/0037139 para mejorar la detección de una muestra biológica que tiene un retrovirus integrado en varios sitios. Las condiciones de reacción del método LAM PCR tradicional produjeron resultados que no reflejaban el estado real de los clones existentes en la población celular de la muestra. Se desarrolló una modificación en la que se amplificaron más fragmentos de integración por PCR sin estar sesgados hacia un fragmento amplificado de un clon específico. La modificación del método LAM PCR permitió a los investigadores determinar la extensión de las células que tienen un gen integrado en la población y determinar la proporción de una célula específica en la población.

Además, Harkey y colaboradores, (2007) describen una modificación optimizada, múltiples brazo y alto rendimiento del método LAM PCR en donde la capacidad de detección se mejoró en un 90% con un muestreo exhaustivo. El protocolo modificado facilitó estimaciones precisas del tamaño total de la agrupación, proporcionando así un enfoque rápido, y rentable para generar grandes datos del sitio de inserción de las localizaciones genómicas preferidas para la integración de vectores.

La presente invención describe una modificación significativa adicional y resuelve varios problemas tradicionales del LAM PCR mediante la eliminación de las etapas de generación de un fragmento de DNA bicatenario después de digerir el fragmento de DNA bicatenario y desnaturalizar el fragmento de DNA bicatenario.

Breve compendio de la invención

La presente invención proporciona un método para encontrar una secuencia de polinucleótido desconocida adyacente a una secuencia de polinucleótido conocida en DNA vegetal aislado, que comprende digerir el DNA vegetal aislado que contiene una parte o la totalidad de una secuencia de polinucleótidos conocida y una secuencia de polinucleótidos desconocida adyacente con una o más enzimas de restricción adecuadas para producir una pluralidad de fragmentos de restricción del polinucleótidos digeridos; sintetizar una cadena complementaria de los fragmentos de restricción del polinucleótido digerido utilizando un cebador de polinucleótidos, diseñado para unirse específicamente a la secuencia de polinucleótidos conocida, teniendo dicha secuencia de oligonucleótidos una química de unión unida al extremo 5' de la secuencia de la secuencia del cebador de oligonucleótido; aislar la cadena complementaria mediante la unión de la química de unión a una matriz de aislamiento adecuada, desnaturalizando luego la cadena complementaria de los fragmentos de restricción del polinucleótido digerido; unir un adaptador monocatenario a la cadena complementaria aislada unida a la matriz de aislamiento para producir una cadena complementaria aislada ligada; llevar a cabo una primera amplificación por PCR de la cadena complementaria aislada ligada utilizando un primer cebador de PCR diseñado para unirse a la secuencia de polinucleótido conocida y un segundo cebador de PCR diseñado para unirse al adaptador monocatenario para producir un primer amplicón de PCR; llevar a cabo una segunda amplificación por PCR de dicho amplicón de PCR, en donde la segunda amplificación por PCR amplifica una secuencia interna de dicho primer amplicón de PCR para producir un segundo amplicón de PCR; y secuenciar el segundo amplicón de PCR para determinar la secuencia de la secuencia de polinucleótidos desconocida.

Una realización de la presente invención, descrita aquí, es un método para el aislamiento e identificación de secuencias de los bordes transgénicos. Una realización de la presente invención es un método que es fácilmente aplicable para aplicaciones de alto rendimiento para determinar el número de copias transgénicas y la localización cromosómica de un sitio de inserción genómica. Además, la presente invención se puede utilizar para la detección simultánea de múltiples sitios de inserción dentro de una reacción. La presente invención describe un método que tiene una sensibilidad y especificidad mejoradas para la detección de fragmentos de polinucleótidos desconocidos que flanquean un fragmento de polinucleótido conocido. Además, la presente invención puede desplegarse para detectar las secuencias de DNA desconocidas que se localizan adyacentes a cualquier secuencia objetivo, incluyendo secuencias virales y sitios de mutagénesis de inserción creados mediante mutagénesis del transposón o mutagénesis generada mediante integración de la cadena T.

Una realización de la presente invención se refiere en parte a la identificación de sucesos transgénicos utilizando dichas secuencias de flaqueo, unión e inserción. De acuerdo con la presente invención, un análisis de PCR modificado y los métodos de análisis de secuenciación de DNA que utilizan amplicones que abarcan el DNA transgénico insertado y sus bordes se pueden utilizar para detectar o identificar variedades o líneas de plantas transgénicas comercializadas derivadas de líneas de plantas transgénicas patentadas.

El borde transgénico y las secuencias flanqueantes cromosómicas adyacentes de la presente invención son diagnósticas para un suceso transgénico. Basándose en estas secuencias, las líneas de plantas transgénicas se pueden identificar en diferentes genotipos vegetales mediante el análisis de secuencias flanqueantes cromosómicas y transgénicas. Por lo tanto, una realización de la presente invención describe un método que se puede utilizar para identificar líneas de plantas transgénicas.

Las secuencias flanqueantes cromosómicas de la presente invención son especialmente útiles en conjunción con el fitomejoramiento, para determinar que plantas progenies comprende un suceso dado, después de cruzar una planta madre que comprende un suceso de interés con otra línea de planta en un esfuerzo por impartir uno o más rasgos de interés adicionales en los progenies. Una realización de la presente invención es la determinación de las secuencias de flaqueo/unión cromosómicas para beneficiar los programas de mejoramiento así como el control de calidad, especialmente para líneas de plantas transgénicas comercializadas.

Además, la identificación de secuencias flanqueantes cromosómicas se pueden utilizar para identificar específicamente la localización genómica de cada inserción transgénica. Esta información se puede utilizar para desarrollar sistemas de marcadores moleculares específicos para cada suceso. Estos sistemas de marcadores moleculares se pueden utilizar para estrategias de reproducción acelerada y para establecer la vinculación de los datos.

Además, la información de la secuencia flanqueante cromosómica se puede utilizar para estudiar y caracterizar los procesos de integración transgénica, las características del sitio de integración genómica, la clasificación de sucesos, la estabilidad de los transgenes y sus secuencias flanqueantes, y la expresión génica (especialmente relacionada con el silenciamiento de genes, patrones de metilación de transgenes, efectos de posición, y posibles elementos relacionados con la expresión tales como MARS (regiones de unión a la matriz), y similares).

Los métodos de esta invención se pueden utilizar para obtener y determinar la secuencia de un polinucleótido desconocido de un organismo transgénico. En cualquiera de los métodos de esta invención, la muestra puede ser DNA genómico y el organismo transgénico puede ser una planta transgénica. Las plantas transgénicas analizadas por cualquiera de los métodos de esta invención se pueden seleccionar de plantas que consisten en cebada, maíz, avena, sorgo, césped, caña de azúcar, trigo, alfalfa, plátano, brócoli, frijol, col, canola, zanahoria, tapioca, coliflor, apio, cítricos, algodón, cucurbitáceas, eucaliptos, lino, ajos, uvas, cebollas, lechugas, guisantes, cacahuets, pimientos, patatas, álamos, pinos, centeno, arroz, girasol, cártamo, soja, fresa, remolacha azucarera, batata, tabaco, tomate, ornamentales, arbustos, nueces, mijo y pasto.

Los métodos de esta invención se pueden utilizar para obtener y determinar la secuencia de un polinucleótido desconocido de un organismo no transgénico. En cualquiera de los métodos de esta invención, la muestra puede ser DNA genómico y el organismo no transgénico puede ser una planta. Las plantas analizadas por cualquiera de los métodos de esta invención se pueden seleccionar de plantas que consisten en cebada, maíz, avena, sorgo, césped, caña de azúcar, trigo, alfalfa, plátano, brócoli, frijol, col, canola, zanahoria, mandioca, coliflor, apio, cítricos, algodón, cucurbitáceas, eucaliptos, lino, ajos, uvas, cebollas, lechugas, guisantes, cacahuets, pimientos, patatas, álamos, pinos, centeno, arroz, girasol, cártamo, soja, fresa, remolacha azucarera, boniato, tabaco, tomate, ornamental, arbusto, nuez, mijo y pasto. En cualquiera de los métodos de la invención, la secuencia de polinucleótidos desconocida adyacente a una secuencia de polinucleótidos conocida puede ser un polinucleótido nativo de interés agronómico.

Breve descripción de las secuencias

SEQ ID NO:1 describe el cebador biotinilado 5' marcado como 4468-3PA01-2Btn.

SEQ ID NO: 2 describe el adaptador fosforilado 5' marcado como ZC-Adp-01.

SEQ ID NO:3 describe el cebador marcado como PAT-InvPriF

SEQ ID NO:4 describe el cebador marcado como Zn_Adt_PCR_01.

Descripción detallada de la invención

5 Como se usa aquí, los términos “comprende”, “que comprende”, “incluye”, “que incluye”, “tiene”, “que tiene”, “contiene” o “que contiene”, o cualquier otra variación de los mismos, están destinados a ser no exclusivos o abiertos. Por ejemplo, una composición, una mezcla, un proceso, un método, un artículo, o un aparato que comprende una lista de elementos no se limita necesariamente a solo esos elementos sino que puede incluir otros elementos no listados expresamente o inherentes a dicha composición, mezcla, proceso, método, artículo, o aparato. Además, a menos que se establezca expresamente lo contrario, “o” se refiere a un inclusive o y no a un
10 exclusive o. Por ejemplo, una condición A o B se satisface por una cualquiera de las siguientes: A es verdadero (o presente) y B es falso (o no presente), A es falso (o no presente) y B es verdadero (o presente), y tanto A como B son verdaderos (o presentes).

15 También, los artículos indefinidos “un” y “un” que preceden a un elemento o componente de la invención pretenden no ser restrictivos con respecto al número de instancias, es decir, ocurrencias del elemento o componente. Por lo tanto, “un” o “un” deben leerse para incluir uno o al menos uno, y la forma de palabra singular del elemento o componente también incluye el plural a menos que obviamente el número sea singular.

20 Los términos “ácido nucleico”, “polinucleótido”, “secuencia de polinucleótidos”, y “secuencia de nucleótidos” se utilizan para referirse a un polímero de nucleótidos (A, C, T, U, G, etc. o análogos de nucleótidos naturales o artificiales), por ejemplo, DNA o RNA, o una representación de los mismos, por ejemplo, una cadena de caracteres, etc, dependiendo del contexto pertinente. Los términos “ácido nucleico” y “polinucleótido” se utilizan aquí indistintamente; estos términos se utilizan en referencia a DNA, RNA, u otras moléculas de ácido nucleico nuevas de la invención, a menos que se indique lo contrario o se contradiga claramente por contexto. Un polinucleótido dado o un polinucleótido complementario se puede determinar a partir de una secuencia de nucleótidos específica. Un ácido nucleico puede estar en forma monocatenaria o bicatenaria.

25 El término “aislado”, se refiere a un material, tal como un ácido nucleico o una proteína, que está: (1) sustancialmente o esencialmente libre de componentes que normalmente acompañan o interaccionan con el material tal como se encuentra en su entorno natural o (2) si el material está en su entorno natural, el material ha sido alterado por una intervención humana deliberada a una composición y/o colocado en un locus en la célula que no sea el locus nativo del material.

30 El término “planta”, incluye plantas y partes de plantas que incluyen pero no se limitan a células vegetales y tejidos vegetales tales como hojas, tallos, raíces, flores, polen y semillas. La clase de plantas que se pueden utilizar en la presente invención es generalmente tan amplia como la clase de plantas superiores e inferiores susceptibles de mutagénesis incluyendo angiospermas (plantas monocotiledóneas y dicotiledóneas), gimnospermas, helechos y algas multicelulares.

35 El término “promotor”, se refiere típicamente a una secuencia de DNA que dirige la transcripción de un gen estructural para producir RNA. Típicamente, un promotor se localiza en una región anterior de 500 pares de bases de un gen, próximo al sitio de inicio de la transcripción. Si un promotor es un promotor inducible, entonces la velocidad de la transcripción aumenta o disminuye en respuesta a un agente inductor exógeno o endógeno. Por el contrario, la velocidad de transcripción se regula en menor grado por un agente inductor si el promotor es un
40 promotor constitutivo.

El término “planta transgénica”, se refiere a una planta o descendiente de la misma derivada de una célula vegetal o protoplasto transformada, en donde el DNA vegetal contiene una molécula de DNA exógena introducida no presente originalmente en una planta nativa, no transgénica de la misma.

45 El término “vector”, como se usa aquí se refiere a cualquier construcción de polinucleótidos recombinante que se puede utilizar con el propósito de la transformación, es decir, la introducción de DNA heterólogo en una célula huésped.

El término “cadena complementaria”, describe secuencias o moléculas de ácidos nucleicos en las que cada base en una molécula se combina con su base complementaria en la otra cadena, para formar una molécula bicatenaria helicoidal estable. Las cadenas individuales se denominan cadenas complementarias.

50 El término “cebador de oligonucleótidos”, es una secuencia de oligonucleótidos lineales de aproximadamente diez a aproximadamente quince nucleótidos de longitud que es complementaria de las secuencias de nucleótidos 5’ o 3’ a amplificar. Un par de cebadores de oligonucleótidos, en los que uno de los cebadores es complementario a una secuencia de nucleótidos 5’ del fragmento de polinucleótidos que se va a amplificar mientras que el otro cebador del par es complementario a una secuencia de nucleótidos 3’ localizada en el fragmento de polinucleótidos que se va a
55 amplificar se puede utilizar para amplificar una secuencia de polinucleótidos. Un experto en la técnica entiende que un par de cebadores de oligonucleótidos significa dos oligonucleótidos complementarios a cadenas opuestas de

ácido y que flanquean la secuencia de polinucleótidos a amplificar.

El término “adaptador”, describe un segmento de polinucleótido de oligonucleótidos corto que se puede unir a una molécula de polinucleótido en un extremo recto o cohesivo. Los adaptadores pueden contener secuencias de reconocimiento de la enzima de restricción dentro del fragmento de polinucleótido. El tamaño del adaptador puede variar desde aproximadamente diez hasta aproximadamente ciento cincuenta nucleótidos de longitud. Los adaptadores pueden ser monocatenarios o bicatenarios.

El término “cadena complementaria aislada ligada”, se refiere a un fragmento de polinucleótido que comprende un adaptador unido a un segundo fragmento de DNA que contiene una parte o todo de la secuencia de polinucleótidos conocida y una secuencia de polinucleótidos desconocida adyacente mediante una reacción de ligación. Una “cadena complementaria aislada ligada” está flanqueada por un adaptador en un extremo y una secuencia de polinucleótidos conocida en el otro extremo.

Una reacción de ligación se completa mediante una enzima, generalmente se hace referencia como una ligasa que cataliza la formación de un enlace fosfodiéster entre los extremos 3'-OH y 5'-P adyacentes en el DNA.

El aislamiento de un DNA vegetal puede llevarse a cabo por métodos conocidos en la técnica. Generalmente, el aislamiento del DNA vegetal da como resultado la obtención del DNA vegetal purificado que está libre de lípidos, proteínas y otros residuos celulares. Los métodos preferidos de aislamiento del DNA vegetal incluyen: lisis, calentamiento, precipitación con alcohol, precipitación salina, extracción orgánica, extracción en fase sólida, extracción con membrana de gel de sílice, purificación con gradiente de CsCl, y cualquier combinación de las mismas. Un método más preferido de aislamiento del DNA vegetal es la tecnología de membrana de gel de sílice comercializada como el kit DNeasy (Qiagen, Valencia, CA) o el protocolo de aislamiento del DNA de bromuro de cetiltrimetilamonio (CTAB).

Las digestiones con enzimas de restricción, también referenciadas como digestiones con endonucleasa de restricción, se realizan cuando se utiliza una enzima nucleasa para escindir las secuencias de polinucleótidos. Hay numerosas enzimas de restricción disponibles para los expertos en la técnica. Como se describe en www.neb.com/nebecomm/tech_reference/restriction_enzymes/overview.asp, se utilizan cuatro clasificaciones para caracterizar las enzimas de restricción. Estas clasificaciones están hechas sobre la base de la composición de la subunidad, la posición de la escisión, la especificidad de la secuencia y los requisitos del cofactor.

Las enzimas del Tipo I cortan aleatoriamente el DNA en localizaciones que están a una distancia de la secuencia de reconocimiento/unión (> 1.000 pares de bases de distancia). Los sitios de reconocimiento que están unidos por una enzima de Tipo I son asimétricos. Como resultado estas enzimas no se utilizan para clonación génica porque estas enzimas no producen fragmentos de restricción discretos o patrones de bandas de gel. Las enzimas de Tipo I son multifuncionales y las diferentes subunidades que comprenden una enzima de restricción de Tipo I son responsables de las diferentes actividades (es decir, la subunidad HsdR codifica la restricción, la subunidad HsdM codifica la metilación del DNA, y la subunidad HsdS codifica la especificidad de la secuencia de reconocimiento).

Las enzimas de Tipo II digieren el DNA en las posiciones localizadas cerca de las secuencias de reconocimiento. Estas enzimas funcionan como un dímero, en donde una subunidad se une a la cadena con sentido y una segunda copia de la subunidad se une a la cadena antisentido en una secuencia palindrómica que tiene típicamente entre 4-8 nucleótidos de longitud. El dímero de Tipo II que se une al DNA puede ser un homodímero que se une a las secuencias de DNA simétricas, o un heterodímero que se une a las secuencias de DNA asimétricas. Las enzimas pueden reconocer tanto secuencias continuas como secuencias discontinuas. Las enzimas de Tipo II están comercialmente disponibles y se utilizan comúnmente para el análisis de DNA y la clonación génica. El uso generalizado de estas enzimas es el resultado de distintos fragmentos de restricción que se producen y se pueden resolver en un gel de agarosa.

Las enzimas de Tipo II son una colección de proteínas no relacionadas que son altamente divergentes en la similitud de la secuencia de aminoácidos. Las enzimas de Tipo II se han dividido en subcategorías que se marcan utilizando un sufijo. Las enzimas de restricción IIB son multímeros que contienen más de una subunidad. Estas enzimas cortan ambos lados de la secuencia de reconocimiento, lo que da como resultado la eliminación de la secuencia de reconocimiento. Las enzimas de restricción de Tipo IIE y Tipo IIF escinden el DNA después de la interacción con dos copias de sus secuencias de reconocimiento. Las enzimas de restricción de tipo IIG comprenden una única subunidad. La parte N-terminal de la enzima posee un dominio de escisión de DNA y un dominio de modificación de DNA. La parte C-terminal de estas enzimas posee un dominio de unión a la secuencia de DNA. Estas enzimas se escinden fuera de su secuencia de reconocimiento. Las enzimas de restricción de Tipo IIM reconocen y cortan el DNA metilado. Las enzimas de restricción de Tipo IIS funcionan como un dímero y escinden el DNA en una localización que está fuera de los sitios de reconocimiento asimétricos no palindrómicos. Estas enzimas comprenden dos dominios distintos, una para la unión al DNA y el otro para la escisión del DNA.

Las enzimas de Tipo III son enzimas de restricción y modificación de combinación. Estas enzimas reconocen dos secuencias no palindrómicas separadas y se escinden fuera de sus secuencias de reconocimiento. Las enzimas de tipo III requieren dos secuencias de reconocimiento en orientaciones opuestas dentro de la misma molécula de DNA

para lograr la escisión.

Las enzimas de tipo IV reconocen el DNA metilado. Los ejemplos incluyen los sistemas McrBC y Mrr de *E. coli*.

Se conocen otros métodos en la técnica para escindir los polinucleótidos y pueden utilizarse en lugar de digerir el polinucleótido con una enzima de restricción, uno cualquiera del grupo que consiste en: lisis, un agente de escisión específico de una secuencia, un agente de escisión no específico de una secuencia, sonicación, esfuerzo cortante, prensa francesa, radiación UV, radiación ionizante, y DNasa. Además, para las enzimas de restricción descritas anteriormente, se podrían utilizar endonucleasas homing o endonucleasas Flap o cualquier combinación de estas enzimas para digerir el DNA aislado. Un método preferido para digerir el DNA vegetal aislado es el uso de una enzima de restricción de Tipo II que se sabe que corta fuera de la secuencia transgénica que se transforma en la planta. Otro método preferido para digerir el DNA vegetal aislado es el uso de una enzima de restricción de Tipo II que se sabe que corta en un sitio que está muy cerca del extremo de la secuencia transgénica.

Las reacciones de extensión del cebador se utilizan para producir cadenas de DNA y RNA que contienen una secuencia de polinucleótido conocida y una secuencia de polinucleótido adyacente desconocida. Las metodologías de extensión del cebador dan como resultado la producción de una cadena complementaria de DNA o RNA que contiene una secuencia de polinucleótido desconocida. La cadena complementaria de DNA o RNA se produce por una polimerasa que se extiende a lo largo de una cadena de plantilla de DNA o RNA después de complejarse con un cebador de oligonucleótido que se une a la cadena de plantilla conocida de DNA o RNA. El cebador de oligonucleótido está diseñado para unirse específicamente a una secuencia de DNA o RNA conocida dentro de la cadena de plantilla de DNA o RNA. Numerosos tipos de polimerasas están comercialmente disponibles para la reacción de extensión; polimerasa T4, polimerasa TAQ, polimerasa PFU, o Transcriptasa inversa son unos pocos ejemplos no limitantes de polimerasas utilizadas comúnmente. Cada polimerasa tiene requisitos de tampón especiales y funciona a una temperatura específica para unas condiciones de reacción óptimas. Una reacción de extensión del cebador preferida es el uso de la polimerasa TAQ comercializada como el kit Platinum Taq.

Las químicas de fijación unidas a una matriz de aislamiento tales como sistemas basados en perlas magnéticas, se utilizan para aislar el DNA monocatenario producido mediante la reacción de extensión del cebador. La cadena de DNA que se produce por la reacción de amplificación del cebador se puede purificar a partir del DNA genómico mediante una interacción de streptavidina – biotina. La biotilación se utiliza ampliamente para permitir el aislamiento, separación, concentración y procesamiento y análisis posteriores de biomoléculas (por ejemplo, los métodos descritos en los documentos de Patente U.S. No. 5,948,624, U.S. No. 5,972,693, y U.S. No. 5,512,439). Hay una variedad de reactivos de biotilación comercialmente disponibles que se dirigen a diferentes grupos funcionales como aminas primarias, sulfhidrilos, carboxilos, carbohidratos, cadenas laterales de tirosina e histidina y bases de cianidina y citosina. El uso de cebadores cortos de oligonucleótidos específicos de secuencias funcionalizadas con biotina (o el equivalente, por ejemplo digoxigenina) y perlas magnéticas para separar las secuencias de DNA específicas a partir del genoma para el posterior análisis tiene múltiples usos. El aislamiento utilizando el método basado en perlas permite el enriquecimiento de una población de DNA para una secuencia particular, lo que permite realizar un análisis posterior que no se pudo realizar en presencia de todo el complemento genómico de DNA. Dichos métodos basados en perlas son apropiados para una automatización de alto rendimiento.

Aunque la interacción biotina – estreptavidina es el mejor par de unión descrito, se conocen otras moléculas que tienen una fuerte afinidad mutua. Las fijaciones químicas que se pueden incluir en un cebador de oligonucleótidos incluyen: ACRYDITE™ un fijador químico basado en una fosoramidita acrílica que se puede añadir a los oligonucleótidos como una modificación en 5', y reacciona covalentemente con superficies modificadas con tiol; Modificaciones de alquinos que reaccionan con grupos funcionales marcados con azida para formar enlaces estables a través de la reacción de cicloadición de Huisgen azida-alquino (también denominada como reacción Click); y modificaciones de tiol que pueden acoplarse e interaccionar con alta afinidad a un ligando o superficie correspondiente (tal como una superficie de oro). Estas moléculas se pueden utilizar para purificar o enriquecer las secuencias de DNA. En donde, un cebador se marca con una primera molécula y la segunda molécula se une a una matriz que puede inmovilizar la primera molécula (por ejemplo, las perlas magnéticas). Una cadena de DNA producida a partir del cebador marcado con la primera molécula se puede aislar haciendo pasar el DNA sobre una columna que contiene la matriz inmovilizada (por ejemplo, las perlas magnéticas) marcada con la segunda molécula. Como resultado de la afinidad por la segunda molécula, se aíslan las secuencias de DNA amplificadas que contienen el cebador marcado con la primera molécula. Las fijaciones químicas preferidas incluyen interacciones acrílico – tiol, interacciones alquino – azida, e interacciones tiol – ligando. Una fijación química más preferida es la interacción estreptavidina – biotina.

Como se utiliza aquí, el término matriz de aislamiento se refiere a una superficie a la que se puede unir una molécula de cualquier clase. Preferiblemente, una matriz de aislamiento es un material insoluble al que se puede unir una molécula de modo que dicha molécula se pueda separar fácilmente de otros componentes en una reacción. Matrices de aislamiento preferidas pueden incluir, pero no se limitan a, un filtro, una resina cromatográfica, una perla, una partícula magnética, o composiciones que comprenden vidrio, plástico, metal, uno o más polímeros y combinaciones de los mismos. Una matriz de aislamiento más preferida es el sistema basado en perlas magnéticas.

Los adaptadores se pueden unir a un DNA monocatenario inmovilizado mediante una ligasa monocatenaria.

Tradicionalmente, las ligasas disponibles comercialmente solo estaban disponibles para unir fragmentos de DNA bicatenario. Recientemente, se ha demostrado que se puede utilizar una ligasa de RNA para unir fragmentos de DNA monocatenarios (Zhang y Chiang (1995) *Nucleic Acids Research*, 24(5); 990-991). Las ligasas monocatenarias preferidas están comercialmente disponibles y se comercializan como CIRCLIGASE™ (Epicentre Biotechnologies, Madison, WI), T4 RNA Ligase 1 y T4 RNA Ligase 2 (New England Biolabs, Ipswich, MA), y Single Strand DNA Ligase (Wako Chemicals, Richmond, VA). Una ligasa de DNA monocatenario preferida es Thermostable RNA Ligase (TRL) de Epicentre Biotechnologies (Madison, WI).

Como describen Brautigam y colaboradores, 2010, el análisis de la secuencia de DNA se puede utilizar para determinar la secuencia de nucleótidos del fragmento aislado y amplificado. Los fragmentos amplificados se pueden aislar y sub-clonar en un vector y secuenciarse utilizando el método de terminación de cadena (también denominado como secuenciación de Sanger) o secuenciación de terminación de colorante. Además, el amplicón se puede secuenciar con la secuenciación de nueva generación. Las tecnologías NGS no requieren la etapa de sub-clonación, y se pueden completar múltiples lecturas de secuenciación en una única reacción. Están comercialmente disponibles tres plataformas NGS, Genome Sequencer FLX de 454 Life Sciences/Roche, Illumina Genome Analyser de Solexa y Applied Biosystems' SOLiD (acrónimo de "Secuenciación por unión y detección de Oligo"). Además, hay dos métodos de secuenciación de moléculas individuales que se están desarrollando en la actualidad. Estos incluyen la verdadera secuenciación de moléculas individuales (tSMS) de Helicos Bioscience y la secuenciación en tiempo real de moléculas individuales (SMRT) de Pacific Biosciences.

El Genome Sequencer FLX que está comercializado por 454 Life Sciences/Roche es un NGS de lectura larga, que utiliza una PCR de emulsión y pirosecuenciación para generar lecturas de secuenciación. Se pueden utilizar fragmentos de DNA de 300 – 800 pares de bases o librerías que contienen fragmentos de 3 – 20 kbp. Las reacciones pueden producir alrededor de un millón de lecturas de aproximadamente 250 a 400 bases por carrera para un rendimiento total de 250 a 400 megabases. Esta tecnología produce las lecturas más largas pero la salida de la secuencia total por carrera es baja comparado con otras tecnologías NGS.

El Illumina Genome Analyser que está comercializado por Solexa es un NGS de lectura corta que utiliza el método de síntesis por secuenciación con nucleótidos terminadores reversibles marcados con colorante y se basa en la PCR de puente en fase sólida. Se puede utilizar la construcción de bibliotecas de secuencias de terminación emparejadas que contienen fragmentos de DNA de hasta 10 kb. Las reacciones producen alrededor de 100 millones de lecturas cortas que son de 35 – 76 bases de longitud. Este dato puede producir de 3 – 6 gigabases por carrera.

El sistema de secuenciación por unión y detección de oligo (SOLiD) comoercializado por Applied Biosystem es una tecnología de lectura corta. Esta tecnología NGS utiliza DNA bicatenario fragmentado que tiene hasta 10 kbp de longitud. El sistema utiliza la secuenciación por unión de cebadores de oligonucleótidos marcados con colorante y la PCR de emulsión para generar un billón de lecturas cortas que dan como resultado una salida de secuencia total de hasta 30 gigabases por carrera.

tSMS de Helicos Bioscience y SMRT de Pacific Biosciences aplican un método diferente que utiliza moléculas de DNA individuales para las reacciones de secuencias. El sistema de Helicos tSMS produce hasta 800 millones de lecturas cortas que dan como resultado 21 gigabases por carrera. Estas reacciones se completan utilizando nucleótidos terminadores virtuales marcados con colorante que se describen como un método de "secuenciación por síntesis".

El sistema de secuenciación de nueva generación SMRT comercializado por Pacific Biosciences utiliza una secuenciación por síntesis en tiempo real. Esta tecnología puede producir lecturas de hasta 1.000 pares de bases de longitud como resultado de no estar limitado por terminadores reversibles. El rendimiento de lectura sin procesar que es equivalente a la cobertura de un pliegue de un genoma humano diploide se puede producir por día utilizando esta tecnología.

Los siguientes ejemplos describen un método desarrollado para aislar e identificar las secuencias flanqueantes genómicas de una inserción transgénica. Además, el método se puede utilizar para determinar el número de copias del transgén y la localización genómica de un transgén para un suceso transgénico.

Realizaciones de la presente invención se definen además en los siguientes Ejemplos.

Ejemplo 1

Un plásmido que contiene un gen de un casete de expresión de interés y un casete de expresión de un gen marcador de selección se utilizó para transformar el tejido vegetal Hi-II de *Zea mays* mediante la pistola génica de Biorad. La producción de maíz transgénico a partir de un callo de Tipo II bombardeado: efecto del tamaño de partícula de oro y la morfología del callo en la eficacia de la transformación. *In Vitro Cell. Dev. Biol-Plant.* 36:21-29). Se modificó el protocolo: componentes de medios, agentes de selección y temporización se optimizaron para mejorar la eficacia del proceso de transformación. Un fragmento linealizado *Fsp I* del plásmido se utilizó para la transformación. Las transformaciones resultantes produjeron plantas de maíz transgénicas que contenían un gen de un casete de expresión de interés que estaba unido al casete de expresión del gen marcador de selección de la planta.

Ejemplo 2

El DNA genómico se aisló de tres acontecimientos de maíz diferentes (3)-001, (3)-008, y (3)-009 y controles de maíz intransformados. Se emplearon varios métodos para aislar el gDNA, tal como el kit DNeasy (Qiagen, Valencia, CA) o el protocolo tradicional de aislamiento de DNA CTAB. Las concentraciones de DNA se determinaron utilizando Nanodrop (Thermo Scientific, Wilmington, DE). Un total de 250 ng de gDNA se digirieron con la enzima de restricción *TaqI*. La reacción de digestión se purificó después utilizando el kit MinElute Reaction Cleanup (Qiagen, Valencia, CA).

Ejemplo 3

Se completaron las reacciones de extensión del cebador que utilizan el gDNA aislado y purificado. Un cebador marcado con doble biotina se sintetizó por Integrated DNA Technologies Inc. (Coralville, IA) y se utilizó para la reacción (SEQ ID NO:1 (4468-3PA01-2Btn) 5'-\Dual Biotin\ -GGACAGAGCCACAAACACCACAAGA-3'). El kit Platinum Taq (Invitrogen, Carlsbad, CA) se utilizó para sintetizar una cadena de DNA mediante la extensión del cebador. Los siguientes reactivos: 2µL, tampón TAQ 10X platino; 1,25 µL, MgCl₂ 50mM; 0,8 µL, dNTP 10mM; 0,1 µL, 4468-3PA01-2Btn 10 µM; 0,1µL TAQ de platino; 14,75µL H₂O; 1 µL gDNA se mezclaron en un tubo. La amplificación se completó utilizando las siguientes condiciones de reacción: 1) 94°C 3 minutos; 2) 98°C 10 segundos; 3) 63°C 1 minuto; 4) 72°C 5 minutos; 5) repetir las etapas 2 – 4 15 veces; 6) 72°C 3 minutos; 7) mantener a 4°C.

Ejemplo 4

Una reacción de captura se completó con 2,5 µL de perlas magnéticas de estreptavidina Dynabeads M-280 (Invitrogen, Carlsbad, CA). Las perlas se lavaron en un imán con tampón PBST (solución salina tamponada con fosfato y tween 20) una vez y con tampón PBS (solución salina tamponada con fosfato) dos veces. Después de que el sobrenadante se eliminase del imán, se añadieron 20 µL de PBS a las perlas y las perlas se mezclaron y se resuspendieron. Esta solución se añadió a una reacción de extensión de un cebador individual a una concentración 1:1, 20 µL de perlas se mezclaron con 20 µL de reacción de extensión del cebador. La solución resultante se incubó durante 1 hora con pipeteo suave a temperatura ambiente. Las perlas se lavaron después sobre un imán con PBST dos veces, PBS dos veces, y H₂O una vez. Todas las soluciones de lavado se separaron de las perlas.

Ejemplo 5

Un adaptador monocatenario se unió al gDNA objetivo capturado monocatenario de los sucesos (3)-001, (3)-008, y (3)-009. El adaptador monocatenario (SEQ ID NO:2 (ZC-Adp-01) 5'-\5Phos\ATTGGATTCTCTGACGGTCGGACGC\36-FAM\ -3'), que se sintetizó en Integrated DNA Technologies (Coralville, IA), se unió con Thermostable RNA Ligase (TRL) de Epicentre Technologies (Madison, WI). La siguiente reacción se utilizó para unir el adaptador al DNA monocatenario: 0,125µL de ZC-Adp-01 100uM; 5,0µL de PEG 8000 al 50% (peso/volumen en H₂O); 1,0µL de DMSO; y 1,875µL de H₂O. El coctel se mezcló y se desnaturizó en un termociclo a 94°C durante 5 minutos enfriándose después a temperatura ambiente. Después, 1µL de tampón 10X TRL, 0,5µL de ATP 1 mM, y 0,5µL de TRL se añadieron y se mezclaron en la solución. La solución resultante se añadió a las perlas lavadas de la reacción de captura y se incubó en un termociclo a 60°C durante 1 hora y después a 4°C. Las perlas se lavaron en un imán con tampón 0,1X TE varias veces, y todos los líquidos se separaron de las perlas.

Ejemplo 6

Las reacciones de PCR se completaron utilizando el kit Takara LA TAG HS PCR (Milipore, Billerica, Ma). Los siguientes cebadores se utilizaron para amplificar el suceso y la secuencia flanqueante: el cebador específico transgénico, SEQ ID NO:3 (PAT-InvPriF) 5'-CGCTTACGATTGGACAGTTGAGAGTACTG-3') y el cebador del adaptador, SEQ ID NO:4 (Zn_Adt_PCR_01) 5'-GTCCGACCGTCAGAGAATCCAAT-3'). Los siguientes reactivos se utilizaron en la reacción de PCR: 5µL, tampón 10X LA TAQ HS; 8 µL, dNTP 2,5mM, 1µL, de cebador específico transgénico 10µM; 1µL, del cebador específico del adaptador 10µM; 0,5µL de polimerasa LA Taq HS; y 34,5µL, H₂O. El coctel se añadió a las perlas lavadas de la reacción de unión y se amplificó utilizando las siguientes condiciones de la Tabla 1.

Tabla 1

Condiciones de amplificación de PCR

1 ciclo	94°C, 2 min
2 ciclos	98°C, 10s
	66°C, 1 min
	68°C, 5 min

28 ciclos	98°C, 10s
	64°C, 30s
	68°C, 2,5min
1 ciclo	72°C, 4 min
1 ciclo	4°C, ∞

Ejemplo 7

- 5 Los productos resultantes de PCR, de tamaños superiores a ~850 pares de bases, se clonaron en un plásmido pCR2.1 (Invitrogen, Carlsbad, CA). Las colonias se aislaron y el plásmido pCR2.1 se confirmó que contenían un amplicón de PCR. Los vectores se secuenciaron utilizando cebadores delateros M13 e inversos M13. Los resultados de la secuenciación se esperaba que contuviesen la secuencia de nucleótidos de la secuencia flanqueante genómica 3' del maíz además de los elementos genéticos presentes del plásmido. La inserción transgénica 3' y las secuencias flanqueantes genómicas del maíz del clon #4 del suceso (3)-001, clon #10 del suceso (3)-008, clon #13 del suceso (3)-008, y (3)-009 se aislaron e identificaron utilizando la técnica descrita anteriormente.
- 10 La caracterización de los inserciones genómicas indicaban que el suceso (3)-001 contiene múltiples copias del transgen. Varias inserciones únicas se identificaron dentro de este suceso. El clon #4 del suceso (3)-001 posee una región flanqueante individual, que se aisló en secuencias adicionales flanqueantes de una segunda y tercera inserción que se reordenan (datos no descritos). Las regiones flanqueantes únicas indican que tres copias del transgen se insertaron en localizaciones individuales del genoma *Zea mays*.
- 15 El suceso (3)-008 contiene dos copias del transgen. Las secuencias flanqueantes del clon #10 del suceso (3)-008 y clon #13 del suceso (3)-008 son individuales y diferentes, lo que indica que dos copias del transgen se insertaron en localizaciones individuales del genoma de *Zea mays*.
- El suceso (3)-009 solo contiene una copia del transgen. La región flanqueante aislada se utilizó para identificar la localización cromosómica del inserto de transgen dentro del genoma de *Zea mays*.
- 20 Las secuencias flanqueantes genómicas del maíz identificadas se destruyeron frente a The Maize Genoma Sequencing Consortium, base de datos genómica B73 de *Zea mays* (Arizona Genomics Institute, University of Arizona) para identificar la localización cromosómica de la inserción del transgen. La secuencia flanqueante del clon #13 del suceso (3)-008 se mapeó en el cromosoma #5. La secuencia flanqueante del suceso (3)-009 se mapeó en el cromosoma #3.
- 25 **Lista de secuencias**
- <110> Cao, Zehui
Novak, Stephen
Zhou, Ning
- <120> Análisis de alto rendimiento de los bordes transgénicos
- 30 <130> 69891
- <160> 8
- <170> PatentIn version 3.5
- <210> 1
- <211> 25
- 35 <212> DNA
- <213> Secuencia Artificial
- <220>
- <223> Cebador de 4468-3PA01-2Btn
- <400> 1
- 40 ggacagagcc acaaacacca caaga 25
- <210> 2
- <211> 25
- <212> DNA

<213> Secuencia Artificial
 <220>
 <223> Adaptador de ZC-Adp-01
 <400> 2
 5 attggattct ctgacggctg gacgc 25
 <210> 3
 <211> 29
 <212> DNA
 <213> Secuencia Artificial
 10 <220>
 <223> Cebador de PAT-InvPriF
 <400> 3
 cgcttacgat tggacagtg agagtactg 29
 <210> 4
 15 <211> 23
 <212> DNA
 <213> Secuencia Artificial
 <220>
 <223> Cebador de Zn_Adt_PCR_01
 20 <400> 4
 gtccgaccgt cagagaatcc aat 23
 <210> 5
 <211> 806
 <212> DNA
 25 <213> Secuencia Artificial
 <220>
 <223> clon #4 de pDAS102408 (3)-001
 <220>
 <221> Secuencia Transgénica
 30 <222> (1)..(728)
 <220>
 <221> caract_misc
 <222> (89) .. (89)
 <223> n e s a , c , g , o t
 35 <220>
 <221> Secuencia Flanqueante Genómica
 <222> (729)..(770)
 <400> 5

ES 2 650 721 T3

gtgtgctgga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 60
gcgaactact tactctagct tcccggcanc aattaataga ctggatggag gcggataaag 120
ttgcaggacc acttctgctc tcggcccttc cggctggctg gtttattgct gataaatctg 180
gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagcctt 240
cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 300
agatcgctga gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact 360
catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga 420
tcctttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt 480
cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttctg cgcgtaactc 540
gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc 600
taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc 660
ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg octacatacc 720
tcgctctggg aaaaagagtt ggtagcaatt ggattctctg acggtcggac aagggcgaat 780
tccagcacac tggcggccgt tactag 806

<210> 6

<211> 836

<212> DNA

5 <213> Secuencia Artificial

<220>

<223> clon #10 de pDAS102408 (3)-008

<220>

<221> Secuencia Transgénica

10 <222> (1)..(731)

<220>

<221> Secuencia Flanqueante Genómica

<222> (732)..(785)

<220>

15 <221> caract_misc

<222> (822)..(822)

<223> n es a, c, g, o t

<220>

<221> caract_misc

20 <222> (826)..(826)

<223> n es a, c, g, o t

<400> 6

ES 2 650 721 T3

	tcggtcgcag cgtgtgcgtg tccgtcgtac gttctggccg gccgggcctt gggcgcgcga	60
	tcagaaagcg ttgcgttggc gtgtgtgtgc ttctggtttg cttaatttt accaagtttg	120
	tttcaaggtg gatcgcgtgg tcaaggcccg tgtgctttaa agaccaccg gcactggcag	180
	tgagtgttg tgcttgtgta ggctttggta cgtatgggct ttatttgctt ctggatgttg	240
	tgtactactt gggtttgtt aattattatg agcagttgcg tattgtaatt cagctgggct	300
	acctggacat tgttatgtat taataaatgc tttgctttct tctaaagatc ttttaagtgt	360
	gaattcatat ttaaatcaat tcaactggccg tcgttttaca acgtcgtgac tgggaaaacc	420
	ctggcgttac ccaacttaat cgccttgca caccatcccc ttccgcccagc tggcgtaata	480
	gcgaagaggg ccgcaccgat cgcccttccc aacagttgcg cagcctgaat ggcgaatggc	540
	gcctgatgcy gtattttctc cttacgcac tgtgcgggat tccacaccgc atatggtgca	600
	ctctcagtac aatctgctct gatgccgat agttaagcca gccccgacac ccgccaacac	660
	ccgctgacgc gccctgacgg gottgtctgc tcccggcacc cgcttacaga caagctgtga	720
	ccgtctccgg gtgttacttc tgcaggttac ccccggggtc gaattggatt ctctgacggt	780
	cggacaaggg cgaattccag cacactggcg gccgttacta gngganccga gctcgt	836
	<210> 7	
	<211> 944	
	<212> DNA	
5	<213> Secuencia Artificial	
	<220>	
	<223> pDAS102408 (3)-008 clone #13	
	<220>	
	<221> Secuencia Transgénica	
10	<222> (153)..(587)	
	<220>	
	<221> Secuencia Flanqueante Genómica	
	<222> (588)..(908)	
	<400> 7	
15	ttaagttggg taacgccagg gttttcccag tcacgacggt gtaaaacgac ggccagtgaa	60

ES 2 650 721 T3

```

ttgtaatacg actcactata gggcgaattg ggcctctag atgcatgctc gagcggccgc 120
cagtgtgatg gatatctgca gaattcgccc ttcgcttacg attggacagt tgagagtact 180
gtttacgtgt cacataggca tcaaagggtg ggcctaggat ccacattgta cacacatttg 240
cttaagtcta tggaggcgca aggttttaag tctgtggttg ctggtatagg ccttccaaac 300
gatccatctg ttaggttgca tgaggctttg ggatacacag cccgggggtac attgcgcgca 360
gctggataca agcatggttg atggcatgat gttggttttt ggcaaaggga ttttgagttg 420
ccagctcctc caaggccagt taggccagtt acccagatct gaggtaccct gagctcggtc 480
gcagctgtg cgtgtccgct gtacgttctg gccggccggg ccttgggctc gcgatcagaa 540
gcgttgcgtt gccgtgtgtg tgcttctggt ttgctttaat tttaccacgc cagccccgac 600
acccgccaac acccgctgac gcgccctgac gggcttgtct gctcccggca tccgcttaca 660
gacaagctgt gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga 720
aacgcgcgag acgaaagggc ctcgtgatac gccccacatc atcacaacca agacaaaaaa 780
agcatgaaaa gatgaccgca caaacaagtg cacgcatat attgaaataa aggaaaaggg 840
caaaccaaac cctatgcaac gaaacaaaaa aaatcatgaa atcgaattgg attctctgac 900
ggtcggacaa gggcgaattc cagcacactg gcggccgcta ctag 944

```

<210> 8

<211> 1024

<212> DNA

5 <213> Secuencia Artificial

<220>

<223> pDAS102408 (3)-009

<220>

<221> caract_misc

10 <222> (14)..(14)

<223> n es a, c, g, o t

<220>

<221>Secuencia Flanqueante Genómica

<222> (51)..(132)

15 <220>

<221> Secuencia Transgénica

<222> (133)..(839)

<220>

<221> caract_misc

20 <222> (982)..(982)

<223> n es a, c, g, o t

<220>

<221> caract_misc

<222> (992)..(994)

25 <223> n es a, c, g, o t

<400> 8

ES 2 650 721 T3

cgagctcgga tccnctagta acggccgcca gtgtgctgga attcgccctt gtccgaccgt	60
cagagaatcc aattogagct cgtttaaact gagggcactg aagtcgcttg atgtgtatct	120
tggtatcgcg catgaaaag tggcaaattt cctcagcact taaagatctt tagaagaaag	180
caaagcattt attaatacat aacaatgtcc aggtagccca gctgaattac aatacgcaac	240
tgctcataat aattcaaca acccaagtag tacacaacat ccagaagcaa ataaagccca	300
tacgtaccaa agcctacaca agcagcaaca ctcaactgcca gtgccggtgg gtctttaaac	360
cacacgggcc ttgaccacgc gatccacctt gaaacaaact tggtaaaatt aaagcaaacc	420
agaagcacac acacgccaac gcaacgctt tgatcgcgcg cccaaggccc ggccggccag	480
aacgtacgac ggacacgcac acgctgcgac cgagctcagg gtacctcaga tctgggtaac	540
tggcctaact ggccctggag gagctggcaa ctcaaaatcc ctttgccaaa aaccaacatc	600
atgccatcca ccatgcttgt atccagctgc gcgcaatgta ccccggtctg tgtatcccaa	660
agcctcatgc aacctaacag atggatcgtt tggaaggcct ataacagcaa ccacagactt	720
aaaaccttgc gcctccatag acttaagcaa atgtgtgtac aatgtggatc ctaggcccaa	780
cctttgatgc ctatgtgaca cgtaaacagt actctcaact gtccaatcgt aagcgaaggg	840
cgaattctgc agatatccat cacactggcg gccgctcgag catgcatcta gagggcccaa	900
ttcgccctat agtgagtcgt attacaattc actgggcccgt cgttttacaa cgtcgtgact	960
gggaaaacc tggcgttacc cnaacttaac gnnngcagca catccccctt ttcgcagctg	1020
gcgt	1024

REIVINDICACIONES

1. Un método para encontrar una secuencia de polinucleótidos desconocida adyacente a una secuencia de polinucleótidos conocida en un DNA aislado, que comprende:
 - 5 a) digerir el DNA aislado que contiene una parte o toda la secuencia de polinucleótidos conocida y una secuencia de polinucleótidos desconocida adyacente con uno o más enzimas de restricción adecuados para producir una pluralidad de fragmentos de restricción de polinucleótidos digeridos;
 - 10 b) sintetizar una cadena complementaria de los fragmentos de restricción de polinucleótidos digeridos utilizando un cebador de oligonucleótidos, diseñado para unirse específicamente a la secuencia de polinucleótidos conocida, teniendo dicha secuencia del cebador de oligonucleótidos una fijación química unida al extremo 5' de la secuencia del cebador de oligonucleótidos;
 - 15 c) aislar la cadena complementaria por unión de la fijación química a una matriz de aislamiento adecuada, desnaturalizando después la cadena complementaria de los fragmentos de restricción del polinucleótido digerido;
 - d) ligar un adaptador monocatenario a la cadena complementaria aislada unida a la matriz de aislamiento para producir una cadena complementaria aislada ligada;
 - e) realizar una primera amplificación de PCR de la cadena complementaria aislada ligada utilizando un primer cebador de PCR diseñado para unir la secuencia de polinucleótidos conocida y un segundo cebador de PCR diseñado para unir el adaptador monocatenario para producir un primer amplicón de PCR;
 - 20 f) realizar una segunda amplificación de PCR de dicho primer amplicón de PCR, en donde la segunda amplificación de PCR amplifica una secuencia interna de dicho primer amplicón de PCR para producir un segundo amplicón de PCR; y,
 - g) secuenciar el segundo amplicón de PCR para determinar la secuencia de la secuencia de polinucleótidos desconocida.
2. El método de la reivindicación 1, en donde el DNA aislado es un DNA genómico vegetal.
- 25 3. El método de la reivindicación 1, en donde la secuencia de polinucleótidos desconocida es una secuencia de bordes transgénicos.
4. El método de la reivindicación 1, en donde la secuencia de polinucleótidos desconocida es una secuencia cromosómica que flanquea una secuencia de polinucleótidos conocida.
5. El método de la reivindicación 1, en donde la secuencia de polinucleótidos desconocida es una secuencia 30 génica nativa que codifica un rasgo.
6. El método de la reivindicación 1, en donde la secuencia de polinucleótidos conocida es una secuencia viral de polinucleótidos conocida.
7. El método de la reivindicación 1, en donde la secuencia de polinucleótidos conocida es una secuencia del transgén de polinucleótidos conocida.
- 35 8. El método de la reivindicación 1, en donde la secuencia de polinucleótidos conocida es una secuencia de transposón de polinucleótidos conocida.
9. El método de la reivindicación 1, en donde la secuencia de polinucleótidos conocida es una secuencia génica que codifica un rasgo.
- 40 10. El método de la reivindicación 1, en donde el método se utiliza para identificar una localización cromosómica de una secuencia de polinucleótidos conocida insertada en un DNA vegetal aislado mediante una mutagénesis de inserción, en donde preferiblemente dicha mutagénesis de inserción se selecciona del grupo que consiste en mutagénesis de transposón, o mutagénesis de integración de la cadena T.
- 45 11. El método de la reivindicación 1, en donde el método se utiliza para caracterizar una secuencia de polinucleótidos desconocida que consiste en una secuencia cromosómica que flanquea una secuencia de polinucleótidos conocida, en donde preferiblemente dicha caracterización de un sitio de inserción del transgén identifica una secuencia de polinucleótidos que consiste en reorganizaciones, inserciones, deleciones o inversiones dentro de la secuencia de polinucleótidos desconocida que consiste en una secuencia cromosómica.
12. El método de la reivindicación 1, en donde el método se utiliza para determinar el número de copia del transgén.
- 50 13. El método de la reivindicación 1, en donde el método se utiliza para identificar líneas vegetales transgénicas.