

LOCUS GENÉTICOS NOVEDOSOS ASOCIADOS CON RESISTENCIA A LA ROYA EN LA SOJA

REFERENCIA CRUZADA A SOLICITUDES RELACIONADAS

Esta solicitud reivindica la prioridad de la Solicitud de Patente China No. 201910584420.1, presentada el 1 de julio de 2019, cuya divulgación se incorpora en el presente documento en su totalidad.

LISTADO DE SECUENCIAS

La presente solicitud contiene un Listado de Secuencias, que se ha presentado electrónicamente en formato ASCII y se incorpora en el presente documento por referencia en su totalidad. Dicha copia ASCII, creada el 22 de junio de 2020, se denomina 115479.000346_Sequence Listing_22June2020_ST25.txt y tiene un tamaño de 167,289 bytes.

CAMPO TÉCNICO

La invención se refiere al campo de la ingeniería genética de plantas, en particular a una proteína relacionada con la resistencia a la roya en la soja, un gen codificante y su uso.

ANTECEDENTES

La soja (*Glycine max*) es uno de los cuatro principales cultivos oleaginosos alrededor del mundo y uno de los cultivos más importantes para la producción de proteínas. La roya es una enfermedad importante en la producción de soja alrededor del mundo y el método principal para controlar la enfermedad es la aplicación de fungicidas foliares.

En la actualidad, no existen variedades comerciales con resistencia completa frente a la roya. El patógeno de la roya es *Phakopsora pachyrhizi*, y los huéspedes del patógeno de la roya incluyen una amplia gama de plantas leguminosas (al

menos 31 especies en 17 géneros; Slaminko *et al.*, (2008) *Plant Dis.*, 92: 797-771; y al menos 42 especies en 19 géneros; Frederick *et al.*, (2002) *Mycology*, 92: 217-227, respectivamente). Además, se han reconocido otras 152 especies como huéspedes potenciales de *Phakopsora pachyrhizi* (Bonde *et al.*, (2008) *Plant Dis.*, 92: 30-38; Goellner *et al.*, (2010) *Molecular Plant Pathology*, 11: 169-177; Ono *et al.*, (1992) *Mycol. Res.*, 96 (10): 825-850; y Slaminko *et al.*, (2008) *Plant Dis.*, 92: 797-771). La aplicación de fungicidas es actualmente el único método disponible para mitigar la roya.

Hay pocos recursos de soja que sean resistentes a *Phakopsora pachyrhizi*. La resistencia de los recursos de soja a *Phakopsora pachyrhizi* es específica para razas fisiológicas individuales, por lo tanto, cuando se usan recursos resistentes para el fitomejoramiento, si se ignora tal especificidad de resistencia, la resistencia en los recursos resistentes puede perderse debido a la incompatibilidad entre la resistencia del huésped y las razas fisiológica, lo que no es propicio para la utilización persistente de los recursos resistentes.

El cultivo de variedades resistentes a la roya es la forma más económica y eficaz de prevenir el daño de la roya. Al explorar nuevos genes de resistencia para enfermedades en el germoplasma de soja y otras plantas leguminosas y transferirlas como un solo gen, múltiples genes o en forma de múltiples casetes de genes a soja, estos genes de resistencia pueden proporcionar resistencia a *Phakopsora pachyrhizi* a través de la expresión homóloga o heteróloga. Por consiguiente, lo que se necesita son genes de resistencia novedosos a la roya que puedan introducirse en plantas de soja comerciales para controlar la resistencia a la soja.

SUMARIO DE LA INVENCION

Este compendio enumera varias realizaciones de la materia en cuestión que se divulga en la presente y en muchos casos enumera variaciones y permutaciones de estas realizaciones.

Por lo tanto, es un objeto de la materia en cuestión que se divulga en la presente proporcionar métodos para transmitir resistencia a la roya al germoplasma o líneas de plantas no resistentes de soja. Además, la materia en cuestión que se describe en la presente proporciona nuevas líneas de *Glycine max* que comprenden en su genoma un intervalo de cromosoma, locus y/o gen que se deriva de cepas de *Glycine max* SX6907 y que confiere, además, resistencia a la roya de la soja asiática (en la presente, "ASR") en dicha nueva línea de *Glycine max*.

La presente invención proporciona intervalos cromosómicos derivados de la cepa de *Glycine max* SX6907 que cuando se introducen en una planta (por ejemplo, una soja tal como la cepa de *Glycine max* Williams 82) son suficientes para conferir una resistencia aumentada a la roya, tal como por ejemplo resistencia a la roya de la soja asiática ("ASR"), en comparación con una planta de control que no comprende dicho intervalo cromosómico. La invención también proporciona proteínas y ácidos nucleicos derivados de la cepa de *Glycine max* SX6907 que confiere resistencia a la roya.

También se proporcionan composiciones y métodos para identificar, seleccionar y producir plantas *Glycine* (incluyendo *Glycines* salvajes (p. ej. líneas de *Glycine tomentella* y *Glycine max*)) con resistencia potenciada a la roya. También se proporcionan plantas y germoplasmas de soja resistentes a la roya.

En algunas realizaciones, se proporcionan métodos de identificación de una planta o germoplasma de soja resistente a la roya. Métodos de este tipo pueden comprender detectar, en la planta o el germoplasma de soja, locus genéticos o un marcador molecular (por ejemplo, SNP o locus de rasgos cuantitativos (QTL)) asociados con resistencia potenciada a enfermedades, en particular resistencia a la ASR. En algunas realizaciones, los locus genéticos o el marcador molecular está asociado con la presencia de un intervalo cromosómico que comprende la secuencia de nucleótidos o una parte de la misma de las SEQ ID NO: 11, 12 o 13, o una parte de la misma donde la parte de la misma está asociada con resistencia a la ASR. En otra realización, los locus genético o marcador molecular se asocian con la presencia de nucleótidos de SEQ ID NO: 2 o una parte de los mismos asociada con resistencia a la ASR. Aún en otra realización, los locus genético o marcador molecular se asocian con la presencia de nucleótido que codifica para la secuencia de aminoácidos de SEQ ID NO: 1 o una parte de los mismos asociada con resistencia a la ASR.

En algunas realizaciones, se proporcionan métodos de producción de plantas de soja resistentes a ASR. Dichos métodos pueden comprender la detección, en una planta de soja o germoplasma, la presencia de locus genéticos y/o un marcador genético asociado con resistencia a patógenos potenciada (por ejemplo, ASR) y la producción de una planta de la descendencia a partir de dicho germoplasma de soja. En algunas realizaciones, los métodos se usan para generar cepas de *Glycine max* resistentes a ASR novedosas.

Otras realizaciones incluyen métodos de selección de un germoplasma o planta de soja resistente a enfermedades. Dichos métodos pueden incluir el cruce de una primera planta o germoplasma de soja con una segunda planta o germoplasma de

soja, donde la primera planta o germoplasma de soja comprende unos locus genéticos derivados de cepas de *Glycine max* SX6907 o una planta de la descendencia de los mismos que comprende una cualquiera de las SEQ ID NO: 2, 11, 12, o 13 o una parte de las mismas asociada con resistencia y/o tolerancia potenciada a enfermedades y/o ASR, y la selección de una planta o germoplasma de la descendencia que posee los locus genéticos.

En algunas realizaciones, se proporcionan métodos de introgresar locus genéticos derivados de cepas de soja SX6907 asociados con resistencia a la roya potenciada en una planta o germoplasma de soja. Tales métodos comprenden cruzar una primera planta de soja o germoplasma que comprende un intervalo cromosómico (por ejemplo, SEQ ID No 11, 12, o 13, o una parte de la misma) derivado de cepas de soja SX6907 asociadas con resistencia a la roya mejorada (ASR) con una segunda planta de soja o germoplasma que carece de dichos locus genéticos y opcionalmente que retrocruza repetidamente plantas de progenie que comprenden dicho alelo genético con la segunda planta de soja o germoplasma para producir una planta de soja (por ejemplo, *Glycine max*) o germoplasma con resistencia a patógeno potenciada que comprende el intervalo cromosómico derivado de las cepas de soja SX6907 y/o ZRYCR1 asociadas con resistencia a potenciada. En una realización, el intervalo cromosómico comprende SEQ ID NO: 2. En otra realización, el intervalo cromosómico comprende SEQ ID NO: 1.

La descendencia que comprende el intervalo cromosómico asociado con resistencia potenciada a patógenos puede identificarse detectando, en sus genomas, la presencia de un marcador asociado con o ligado genéticamente a dicho intervalo cromosómico derivado de las cepas de número de acceso de soja SX6907 y/o ZRYCR1, donde dicho intervalo

cromosómico comprende las SEQ ID NO: 11, 12 o 13 o una parte de las mismas y el marcador puede ser cualquiera de los alelos favorables descritos en la Tabla 1.

También se proporcionan plantas y/o germoplasmas de soja identificados, producidos o seleccionados por los métodos de esta invención, así como cualquier descendencia y/o semillas derivadas de una planta o germoplasma de soja identificado, producido o seleccionado por estos métodos. En una realización, se pueden utilizar marcadores moleculares que se asocian con la presencia de intervalos cromosómicos representados en una cualquiera de SEQ ID NO 11, 12 o 13 para identificar o seleccionar líneas de plantas resistentes a la ASR. Además, dichos marcadores moleculares pueden estar ubicados dentro de 20 cM, 10 cM, 5 cM, 4 cM, 3 cM, 2 cM y 1 cM de dicho intervalo cromosómico o de cualquier alelo favorable respectivo asociado con resistencia a la ASR como se representa en la Tabla 1. En otra realización, dichos marcadores moleculares pueden estar ubicados dentro de 20 cM, 10 cM, 5 cM, 4 cM, 3 cM, 2 cM, 1 cM de cualquiera de los marcadores de SNP asociados con ASR, tal como se representa en la Tabla 1.

También se proporcionan semillas de soja que no se producen de manera natural, plantas y/o germoplasmas que comprenden uno o locus genéticos derivados de las cepas SX6907 y/o ZRYCR1 asociados con resistencia a la roya potenciada. En realizaciones específicas dichos locus genéticos comprenden una cualquiera de las SEQ ID NO: 11, 12, o 13, o una parte de las mismas y/o cualquier alelo tal como se representa en la Tabla 1. En otras realizaciones, los locus genéticos comprenden la secuencia de ácido nucleico de la SEQ ID NO: 2 o un ácido nucleico que codifica para la proteína de SEQ ID NO: 1.

Un marcador asociado con resistencia potenciada a la roya (ASR) puede comprender, consistir esencialmente en o consistir en un único alelo o una combinación de alelos a uno o más locus genéticos derivados de las cepas SX6907 y/o ZRYCR1 que se asocian con resistencia potenciada a patógenos (ASR). En una realización, el marcador está dentro de un intervalo cromosómico según se describe por las SEQ ID NO: 11, 12, o 13. En otra realización, el marcador está dentro de la SEQ ID NO: 2. En otra realización, el marcador es uno cualquiera de los alelos favorables representados en la Tabla 1.

También se proporcionan composiciones y métodos adicionales para producir plantas glicina que tiene resistencia a la enfermedad potenciada. En un aspecto de la invención, se proporciona un constructo de ADN que comprende un promotor que funciona en células vegetales unidas operativamente a una molécula de ADN que codifica para una proteína que tiene al menos el 80% el 100% de homología con la SEQ ID NO: 1. La divulgación actual también se refiere a moléculas de ADN. Las moléculas de ADN a modo de ejemplo incluyen (B1) una molécula de ADN mostrada en SEQ ID NO: 2; (B2) una molécula de ADN que hibrida con la molécula de ADN definida en (B1) en una condición estricta y que codifica para la proteína; (B3) una molécula de ADN que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con las secuencias de ADN definidas en (B1) y (B2) y que codifica para la proteína.

En otro aspecto de la invención se refiere a un casete de expresión, un vector recombinante, una bacteria recombinante, o una línea celular transgénica que comprende la molécula de ácido nucleico. La invención también se refiere a un método de preparar una planta transgénica

fértil, que comprende proporcionar un casete de expresión de plantas que comprende al menos uno de un gen *RG21* y un gen *RG22* y poner en contacto células vegetales receptoras con el casete de expresión de plantas bajo condiciones que permiten la absorción del casete de expresión de plantas por parte de las células receptoras; seleccionar las células vegetales receptoras que contienen el casete de expresión de plantas; regenerar plantas de las células vegetales receptoras seleccionadas; e identificar una planta transgénica fértil que es resistente a patógenos de la soja, particularmente ASR.

En otro aspecto de la invención se proporciona una planta transgénica fértil que comprende un casete de expresión de plantas tal como se describió anteriormente en donde la planta es resistente a patógenos de la soja, particularmente la ASR.

En otro aspecto de la invención se proporciona un método de controlar la ASR en un campo, que comprende la etapa de plantar la semilla de una planta que comprende un casete de expresión tal como se describe en el presente documento.

Como un aspecto adicional hay semillas que producen las plantas transgénicas de la invención y semillas producidas por las plantas transgénicas de la invención.

También se proporcionan productos recolectados derivados de las plantas transgénicas de la invención, donde el producto recolectado comprende opcionalmente una secuencia de nucleótidos, casete de expresión, vector y/o al menos una de una proteína o molécula de ADN de la invención. Además se proporcionan productos procesados derivados de los productos recolectados de la invención, donde el producto recolectado comprende opcionalmente una secuencia de nucleótidos, casete

de expresión, vector y/o al menos una de una proteína o molécula de ADN de la invención.

Aun adicionalmente, la descripción proporciona, como un aspecto adicional, un método de producir una planta transgénica con resistencia aumentada a un patógeno de la soja. En realizaciones, el método puede comprender el aumento del nivel de expresión y/o la actividad de una proteína que tiene al menos el 80%-100% de homología con la SEQ ID NO: 1. En otra realización, la descripción se refiere a métodos para fitomejorar una variedad de plantas con resistencia mejorada frente a la roya, que comprende la etapa de aumentar el nivel de expresión y/o la actividad de una proteína que tiene al menos el 80%-100% de homología con la SEQ ID NO: 1 en una planta receptora.

Las composiciones de la invención también incluye sondas y pares de cebadores para la detección de los genes de resistencia novedosos dados a conocer en el presente documento.

Los objetos y aspectos precedentes y otros de la presente invención se explican más detalladamente en los dibujos y la memoria descriptiva expuestos más adelante.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

La figura 1 es un mapa plasmídico de un vector recombinante pB2GW7-*RppRC1* con el gen *RppRC1* de resistencia a la roya.

La figura 2 es una figura de detección de PCR de generación T1 de plantas transgénicas *RppRC1*. M: marcador. L1-1, L1-2 y L1-3 son plantas individuales parciales de las plantas T1 del acontecimiento de transformación L1, y L2-1, L2-2 y L2-3 son plantas individuales parciales de las plantas T1 del acontecimiento de transformación L2, y el negativo:

control negativo Tianlong No. 1, positivo: control positivo SX6907.

La figura 3 muestra la identificación de RT-PCR de la expresión del gen RppRC1 en la generación T1 de plantas transgénicas. L1-2 es una planta individual de las plantas T1 del acontecimiento de transformación L1, y L2-1 es una planta individual de las plantas T1 del acontecimiento de transformación L2.

Las figuras 4A y 4B muestran el análisis de Southern de plantas transgénicas. La figura 4A muestra análisis de Southern de plantas transgénicas RppRC1. L1-1, L1-2, L1-3, L1-4 y L1-5 son plantas individuales parciales de las plantas T1 del acontecimiento de transformación L1, respectivamente. L2-1 y L2-2 son plantas individuales parciales de las plantas T1 del acontecimiento de transformación L2, mientras que Tianlong No.1 es el control negativo. La figura 4B muestra el análisis de Southern de plantas transgénicas con vector vacío. L3-1, L3-2, L3-3, L3-4, L3-5 y L3-6 eran plantas individual parciales de las plantas T1 del acontecimiento de transformación del vector vacío L3. CK es Tianlong No.1.

La figura 5 muestra el fenotipo para identificación de la resistencia de las plantas transgénicas T0 12 días tras la inoculación. SX6907 es un control de resistencia y muestra inmunidad; el acontecimiento de transformación de RppRC1 L2 muestra inmunidad; el acontecimiento de transformación del vector vacío L3 muestra susceptibilidad; Tianlong No.1 no transgénico muestra susceptibilidad.

La figura 6 muestra el fenotipo para identificación de la resistencia de las plantas transgénicas T1 12 días tras la inoculación. SX6907 es un control de resistencia y muestra inmunidad; la planta individual L2-1 de las plantas T1 del acontecimiento de transformación de RppRC1 L2 muestra

inmunidad; Tianlong No. 1 de control negativo muestra susceptibilidad.

Con referencia a la figura 1-6, el mismo acontecimiento de transformación se marca con la misma etiqueta y la misma planta individual se marca con la misma etiqueta.

DESCRIPCIÓN DETALLADA DE REALIZACIONES ILUSTRATIVAS

Diversas publicaciones, artículos y patentes se citan o se describen en los antecedentes y por toda la memoria descriptiva; cada una de estas referencias se incorpora en el presente documento por referencia en su totalidad. La discusión de documentos, actos, materiales, dispositivos, artículos o similares que se han incluido en la presente memoria descriptiva tiene el propósito de proporcionar contexto para la invención. Tal discusión no es una admisión de que alguno o todos estos temas forman parte de la técnica anterior con respecto a cualquier invención dada a conocer o reivindicada.

Todas las patentes, solicitudes de patente publicadas , y publicaciones citadas en el presente documento se incorporan por referencia como si se establecieran completamente en el presente documento.

A menos que se defina lo contrario, todos los términos técnicos y científicos empleados en la presente, a no ser que se definan de otra forma, tienen el mismo significado que el que sobreentienden habitualmente los expertos en la técnica a la que pertenece esta invención. De lo contrario, determinados términos usados en el presente documento tienen los significados establecidos en la memoria descriptiva.

La presente solicitud se refiere a nuevos genes que codifican para las proteínas para la resistencia a la roya y su uso para proporcionar resistencia a la roya en plantas, en

particular en la soja. En una realización de la invención, el gen se deriva de la soja (*Glycine max*) SX6907.

Por consiguiente, en un aspecto, la presente solicitud proporciona proteínas relacionadas con la resistencia a la roya en una planta, un gen codificante y uso de las mismas. En la invención, la resistencia frente a la roya roya de la soja transgénica obtenida mediante la transformación del gen *RppRC1* en la variedad de soja Tianlong No.1 es significativamente mayor que la de la soja de tipo natural, lo que indica que *RppRC1* y el gen codificante del mismo puede regular y controlar la resistencia de plantas leguminosas frente a la roya, y mejorar la resistencia a la roya de plantas después de la sobreexpresión. *RppRC1* y el gen codificante del mismo pueden usarse para mejorar la resistencia a enfermedades de cultivos leguminosos y son de gran importancia para el fitomejoramiento de nuevas variedades con resistencia a enfermedades.

Otros aspectos de la invención incluyen métodos para transmitir resistencia a la roya en germoplasma de soja no resistente o líneas de plantas. Además, la materia en cuestión que se describe en la presente proporciona nuevas líneas de *Glycine max* que comprenden en su genoma un intervalo de cromosoma, locus y/o gen que se deriva de *Glycine max* SX6907 y que confiere, además, resistencia a la roya de la soja en dicha nueva línea de *Glycine max*.

Definiciones

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado entendido comúnmente por un experto en la materia a la que pertenece el contenido que se divulga en el presente documento.

Aunque se considera que los siguientes términos son de uso común para los expertos en la materia, las siguientes definiciones se exponen para facilitar la comprensión de la materia en cuestión que se describe en la presente.

A menos que se indique explícitamente lo contrario en el contexto, las formas singulares "un," "una," y "el/la" tal como se usa en el presente documento incluyen múltiples referencias. Por tanto, por ejemplo, las referencias a "una célula" incluyen una pluralidad de tales células, y referencias a "la proteína" incluyen referencias a una o más proteínas y sus equivalentes conocidos por los expertos en la técnica, y así sucesivamente. A menos que se indique explícitamente lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen los mismos significados en general entendidos por los expertos habituales en la técnica a la que pertenece la presente invención.

Tal como se usa en la memoria descriptiva y reivindicaciones, el término "comprende" y variaciones gramaticales del mismo pueden incluir aspectos de "consiste en" y "consiste sustancialmente en". "Comprender" y variaciones gramaticales del mismo también pueden significar "comprende, pero no se limita a".

Tal como se usa en el presente documento, la palabra "o" se refiere a cualquier elemento de una lista particular y también comprende cualquier combinación de elementos de la lista.

Tal como se utiliza en la presente, el término "y/o" contempla y se refiere a todas y cada una de las combinaciones posibles de uno o más de los elementos enumerados asociados, así como también a la carencia de

combinaciones cuando se interpreta en la forma alternativa, "o".

El intervalo puede expresarse en la presente invención como desde "aproximadamente" un valor específico y/o hasta "aproximadamente" otro valor específico. Cuando se expresa tales intervalos, otros aspectos incluyen desde un valor específico y/o hasta otros valores específicos. De manera similar, cuando se expresa el valor como una aproximación, debe entenderse que al usar el antecedente "aproximadamente", el valor específico forma otro aspecto. También debe entenderse que los puntos finales de cada uno de los rangos están significativamente relacionados e independientes del otro punto final. También debe entenderse que hay valores múltiples dados a conocer en la presente invención y, además del propio valor, cada valor también se da a conocer en el presente documento en forma de "aproximadamente" el valor específico. Por ejemplo, si se da a conocer el valor "10", también se da a conocer "aproximadamente 10" . También debe entenderse que también se da a conocer cada unidad entre dos unidades específicas. Por ejemplo, si se dan a conocer 10 y 15, también se dan a conocer 11, 12, 13 y 14.

La expresión "consiste esencialmente en" (y variantes gramaticales de la misma), tal como se aplica a una secuencia de polinucleótidos de esta invención, significa una secuencia de polinucleótidos que consiste tanto en la secuencia reseñada (p. ej., SEQ ID NO) como en un total de diez o menos (p. Ej.,, 1, 2, 3, 4, 5, 6, 7, 8, 9, o 10) nucleótidos adicionales en los extremos 5' y/o 3' de la secuencia reseñada, de modo que no se altera materialmente la función del polinucleótido. El total de diez o menos nucleótidos adicionales incluye el número total de nucleótidos adicionales en ambos extremos añadidos juntos. La expresión

"materialmente alterados", tal como se aplica a polinucleótidos de la invención, se refiere a un incremento o una disminución en la capacidad de expresar la secuencia de polinucleótidos de al menos aproximadamente un 50 % o más en comparación con el nivel de expresión de una secuencia de polinucleótidos que consiste en la secuencia reseñada.

El término "introducido", tal como se utiliza en la presente, en relación a una planta, significa realizado de cualquier manera, incluyendo, pero no limitado a; introgresión, transgénico, modificación de Repeticiones Palindrómicas Cortas Agrupadas y Regularmente Interespaciadas (CRISPR), nucleasas efectoras de tipo activador de la transcripción (TALENs) (Feng et al. 2013, Joung y Sander 2013), meganucleasas o nucleasas con dedos de zinc (ZFNs).

Tal como se usa en el presente documento, el término "soja" se refiere a soja y cualquier variedad de planta fitomejorada o cultivada usando soja incluyendo plantas de "glicina salvaje".

Tal como se usa en el presente documento, el término "glicina salvaje" se refiere a una planta glicina perenne, por ejemplo una cualquiera de *G. canescens*, *G. argyrea*, *G. clandestine*, *G. latrobeana*, *G. albicans*, *G. aphyonota*, *G. arenaria*, *G. curvata*, *G. cyrtoloba*, *G. dolichocarpa*, *G. falcate*, *G. gracei*, *G. hirticaulis*, *G. lactovirens*, *G. latifolia*, *G. microphylla*, *G. montis-douglas*, *G. peratosa*, *G. pescadrensis*, *G. pindanica*, *G. pullenii*, *G. rubiginosa*, *G. stenophita*, *G. syndetika*, o *G. tomentella*.

En la presente invención, "ácido nucleico" se refiere a un polímero de desoxirribonucleótidos o de ribonucleótidos en forma monocatenaria o bicatenaria y, a menos que se limite lo contrario, abarca análogos conocidos (por ejemplo, ácidos nucleicos peptídicos) que tienen las propiedades básicas de

nucleótidos naturales en los siguientes aspectos: hibrida con ácidos nucleicos monocatenarios de manera similar a la de los nucleótidos que se producen de manera natural.

El término "una variante" y variaciones gramaticales del mismo se refiere a una secuencia sustancialmente similar. Para moléculas de ácido nucleico, variantes comprenden delección y/o adición de uno o más nucleótidos en uno o más sitios en la molécula de ácido nucleico nativa, y/o sustitución de uno o más nucleótidos en uno o más sitios en la molécula de ácido nucleico nativa.

El término "proteína" se refiere a un polímero de residuos de aminoácidos. El término se aplica a polímeros de aminoácidos en los que uno o más residuos de aminoácidos son análogos químicos artificiales de correspondientes aminoácidos naturales, y a polímeros de aminoácidos naturales.

Tal como se usa en el presente documento, una proteína o molécula de ácido nucleico "nativa" comprende una secuencia de nucleótidos que se produce de manera natural o una secuencia de aminoácidos, respectivamente.

Tal como se usa en el presente documento, los términos "codificante" o "que codifica" y variaciones gramaticales de los mismos se usan para significar que un ácido nucleico comprende la información deseada, que se especifica por el uso de codones para dirigir la traducción de secuencias de nucleótidos (por ejemplo, leguminosas secuencias) en proteínas específicas. Un ácido nucleico que codifica para una proteína puede comprender una secuencia no traducida (por ejemplo, un intrón) dentro de la región de traducción del ácido nucleico o puede carecer de una secuencia no traducida intermedia de este tipo (por ejemplo, como en el ADNc).

Tal como se utiliza en la presente, el término "alelo" se refiere a uno de dos o más nucleótidos o secuencias de nucleótidos diferentes que aparecen en un locus específico.

Un marcador está "asociado con" un rasgo cuando está enlazado al mismo y cuando la presencia del marcador es un indicador de si y/o en qué medida aparecerá el rasgo o forma de rasgo deseado en una planta/germoplasma que comprende el marcador. De manera similar, un marcador está "asociado con" un alelo cuando está enlazado al mismo y cuando la presencia del marcador es un indicador de si el alelo está presente en una planta/germoplasma que comprende el marcador. Por ejemplo, "un marcador asociado con resistencia a patógenos potenciada" se refiere a un marcador, cuya presencia o ausencia se puede utilizar para predecir si y/o en qué medida una planta exhibirá un fenotipo resistente a patógenos (p. ej., cualquier alelo SNP favorable, tal como se describe en el presente documento, está "asociado con" resistencia a la ASR (roya) en una planta de soja).

Tal como se utiliza en la presente, los términos "retrocruce" y "retrocruzamiento" se refieren al proceso mediante el cual una planta de progenie se cruza repetidamente de vuelta a una de sus plantas parentales. En un esquema de retrocruzamiento, la planta progenitora "donante" se refiere a la planta progenitora con el gen o locus deseado que va a someterse a introgresión. La planta progenitora "receptora" (utilizada una o más veces) o la planta progenitora "recurrente" (utilizada dos o más veces) se refiere a la planta progenitora en la que se está realizando la introgresión del gen o locus. Por ejemplo, véase Ragot, M. et al. *Marker-assisted Backcrossing: A Practical Example*, in *TECHNIQUES ET UTILISATIONS DES MARQUEURS MOLECULAIRES LES COLLOQUES*, Vol. 72, págs. 45-56 (1995); y

Openshaw et al., *Marker-assisted Selection in Backcross Breeding*, in PROCEEDINGS OF THE SYMPOSIUM "ANALYSIS OF MOLECULAR MARKER DATA," págs. 41-53 (1994). El cruce inicial da lugar a la generación F1. El término "BC1" se refiere al segundo uso del progenitor recurrente, "BC2" se refiere al tercer uso del progenitor recurrente, y así sucesivamente.

Un centimorgan ("cM") es una unidad de medida de frecuencia de recombinación. Un cM equivale a una probabilidad de un 1% de que un marcador en un locus genético se separe de un marcador en un segundo locus debido al cruzamiento en una única generación.

Tal como se utiliza en la presente, la expresión "intervalo cromosómico definido por y que incluye", utilizado en referencia a locus y/o alelos particulares, se refiere a un intervalo cromosómico delimitado por y que abarca los locus/alelos establecidos.

Los términos "cruce" o "cruzado", tal como se utilizan en la presente, se refieren a la fusión de gametos mediante la polinización para producir la progenie (p. ej., células, semillas o plantas). El término engloba tanto cruzamientos sexuales (la polinización de una planta por otra) como la autofecundación (autopolinización, por ejemplo, cuando el polen y el óvulo proceden de la misma planta). El término "cruzamiento" se refiere al acto de fusionar gametos mediante la polinización para producir la progenie.

Los términos "cultivar" y "variedad", tal como se utilizan en la presente, se refieren a un grupo de plantas similares que, mediante características estructurales o genéticas y/o de rendimiento, se pueden distinguir de otras variedades dentro de la misma especie.

Tal como se utiliza en la presente, las expresiones "alelo deseado", "alelo favorable" y "alelo de interés" se

utilizan de manera indistinta para referirse a un alelo asociado con un rasgo deseado (p. ej., resistencia a la ASR).

Tal como se usa en el presente documento, los términos "inhibir," "reducir," etc., y variaciones gramaticales de los mismos se refieren a cualquier reducción en la expresión o función de un producto génico diana, incluyendo cualquier reducción relativa en la expresión o función hasta e incluyendo eliminación completa de la expresión o función del producto génico diana.

El término "potenciar" y variaciones gramaticales del mismo se refieren a mejora, aumento, amplificación, reproducción, elevación y/o elevación para reducir uno o más síntomas de la enfermedad.

Tal como se usa en el presente documento, los términos "aumentar", "potenciar" etc., y variaciones gramaticales de los mismos se usan para referirse a cualquier promoción o ganancia o aumento en la expresión, función o actividad de un producto de un gen diana (por ejemplo, un gen de resistencia) en comparación con una planta susceptible, proporcionando de ese modo resistencia aumentada frente a uno o más patógenos (por ejemplo, *Phakopsora*) o enfermedades (por ejemplo, roya). Adicionalmente, tal como se usa en el presente documento, el término "provocar" o "aumentar" y variaciones gramaticales del mismo puede referirse a una expresión mayor de un producto génico diana de modo que el nivel aumento en el 10% o más, el 50% o más, o el 100%, con respecto a una célula o planta que carece de la proteína o gen diana dado a conocer en el presente documento.

El término "inmunidad" o "inmune" se usa en la presente invención para referirse a la ausencia de cualquier síntoma de enfermedad macroscópicamente visible. El término "resistencia parcial" se usa en la presente invención para

referirse a la presencia de lesiones macroscópicamente visibles sin o con formación de esporas limitada y/o una reducción en el alcance o grado de cualquier síntoma de enfermedad y/o un retraso en la progresión de cualquier síntoma de enfermedad pueden, por ejemplo, manifestar una reducción en el número de lesiones o lesiones con formación de esporas reducida. Tal como se usa en el presente documento, el término "susceptibilidad" o la frase "falta de resistencia" en cuanto a roya se refiere a la aparición de una lesión en el caso en que el nivel de formación de esporas es igual o mayor que el nivel de formación de esporas observado en un estándar de referencia, tal como, por ejemplo, la variedad Williams o Peking.

El término "resistencia" se usa en el presente documento se refiere a la ausencia o reducción de uno o más síntomas de la enfermedad provocados por patógenos de plantas en plantas. La resistencia puede significar que los síntomas de la enfermedad, como el número de plantas enfermas, la defoliación y la pérdida de rendimiento asociada, se reducen, minimizan o disminuyen en comparación con las plantas susceptibles a las enfermedades o plantas que no comprenden genes de resistencia efectivos que reducen uno o más síntomas de la enfermedad. Además, la resistencia puede incluir la prevención o el retraso de la proliferación de patógenos. En términos generales, el término "resistencia" incluye inmunidad y resistencia parcial tal como se definió anteriormente.

Tal como se utiliza en la presente, las expresiones "resistencia a patógenos potenciada", "resistencia a patógenos de las plantas potenciada" o "resistencia a enfermedades potenciada" se refieren a una mejora, potenciación o un incremento en la capacidad de una planta de

perdurar y/o progresar a pesar de estar infectada con una enfermedad (p. ej., roya de la soja asiática) en comparación con una o más plantas control (p. ej., uno de los parentales o ambos, o una planta que carece de un marcador asociado con resistencia a patógenos potenciada a una patógeno/enfermedad respectivo). Resistencia a enfermedades potenciada incluye cualquier mecanismo (distinto de la inmunidad o resistencia de la planta entera) que reduce la expresión de síntomas indicativos de infección para una enfermedad respectiva tal como la roya de la soja asiática, el nematodo del quiste de soja, *Pytophthora*, etc.

"Un patógeno vegetal" y variaciones gramaticales del mismo puede usarse en el presente documento para referirse a, por ejemplo, un patógeno fúngico del género *Phakopsora* de la clase *Basidiomycetes* (incluyendo *Phakopsora pachyrhizi* y *Phakopsora meibomiae*). Las enfermedades de las plantas o las enfermedades de cultivos leguminosos pueden ser, por ejemplo, roya.

El término "gen de resistencia a la enfermedad" o "gen de resistencia" se usa en la presente invención para referirse a un gen que codifica para una proteína que puede potenciar o mejorar la defensa o la respuesta del sistema inmune en plantas.

El término "ortólogo" y variaciones gramaticales del mismo se refiere a genes derivados de genes ancestrales comunes y presentes en diferentes especies debido a la especiación.

Una "línea de élite" o una "variedad de élite" es una línea agrónomicamente superior que ha resultado de muchos ciclos de reproducción y selección para un comportamiento agronómico superior. Están disponibles numerosas líneas de élite y son conocidas por los expertos en la técnica del

cultivo de soja. Una "población de élite" es un surtido de individuos o líneas de élite que se pueden utilizar para representar el estado de la técnica en términos de genotipos agronómicamente superiores de una especie de cultivo dada, tal como soja. De manera similar, un "germoplasma de élite" o cepa de germoplasma de élite es un germoplasma agronómicamente superior, derivado típicamente de y/o capaz de dar origen a una planta con comportamiento agronómico superior, tal como una línea de élite de soja existente o recientemente desarrollada.

Una planta de "élite" es cualquier planta de una línea de élite tal que una planta de élite es una planta representativa de una variedad de élite. Ejemplos no limitantes de variedades de soja de élite que se encuentran disponibles comercialmente para granjeros o para obtentores de soja incluyen: AG00802, A0868, AG0902, A1923, AG2403, A2824, A3704, A4324, A5404, AG5903, AG6202 AG0934; AG1435; AG2031; AG2035; AG2433; AG2733; AG2933; AG3334; AG3832; AG4135; AG4632; AG4934; AG5831; AG6534; y AG7231 (Asgrow Seeds, Des Moines, Iowa, USA); BPR0144RR, BPR 4077NRR y BPR 4390NRR (Bio Plant Research, Camp Point, Ill., EE. UU.); DKB17-51 y DKB37-51 (DeKalb Genetics, DeKalb, Ill., EE. UU.); DP 4546 RR, y DP 7870 RR (Delta & Pine Land Company, Lubbock, Tex., EE. UU.); JG 03R501, JG 32R606C ADD y JG 55R503C (JGL Inc., Greencastle, Ind., EE. UU.); NKS 13-K2 (NK Division de Syngenta Seeds, Golden Valley, Minnesota, EE. UU.); 90M01, 91M30, 92M33, 93M11, 94M30, 95M30, 97B52, P008T22R2; P16T17R2; P22T69R; P25T51R; P34T07R2; P35T58R; P39T67R; P47T36R; P46T21R; y P56T03R2 (Pioneer Hi-Bred International, Johnston, Iowa, EE. UU.); SG4771NRR y SG5161NRR/STS (Soygenetics, LLC, Lafayette, Ind., EE. UU.); S00-K5, S11-L2, S28-Y2, S43-B1, S53-A1, S76-L9, S78-G6, S0009-M2; S007-Y4;

S04-D3; S14-A6; S20-T6; S21-M7; S26-P3; S28-N6; S30-V6; S35-C3; S36-Y6; S39-C4; S47-K5; S48-D9; S52-Y2; S58-Z4; S67-R6; S73-S8; y S78-G6 (Syngenta Seeds, Henderson, Ky., EE. UU.); Richer (Northstar Seed Ltd. Alberta, CA); 14RD62 (Stine Seed Co. Ia., EE. UU.); o Armor 4744 (Armor Seed, LLC, Ar., EE. UU.).

Las expresiones "élite agronómica", tal como se utilizan en la presente, significan un genotipo que tiene una culminación de muchos rasgos distinguibles, tales como emergencia, vigor, vigor vegetativo, resistencia a enfermedades, siembra, durabilidad, rendimiento y capacidad de desgrane, que permite a un productor cosechar un producto de importancia comercial.

Tal como se utiliza en la presente, la expresión "rendimiento comercialmente significativo" o "rendimiento agronómicamente aceptable" se refiere a una producción de granos de al menos 100% de una variedad de control comercial tal como AG2703 o DKB23-51.

Tal como se utiliza en la presente, el término "exótico" y las expresiones "línea exótica" y "germoplasma exótico" se refieren a cualquier planta, línea o germoplasma que no es élite. En general, las plantas/germoplasmas exóticos no se han derivado de ninguna planta o germoplasma élite conocidos, sino que en su lugar, se han seleccionado para introducir uno o más elementos genéticos deseados en un programa de reproducción (p. ej., para introducir alelos novedosos en un programa de reproducción).

El término "germoplasma" se usa en la presente invención para referirse a un material genético derivado de un individuo (por ejemplo, una planta), un grupo de individuos (por ejemplo, una línea germinal de la planta, variedad, o familia), o un clon derivado de una cepa, variedad, especie o

cultivo. El germoplasma puede ser parte de un organismo o una célula, o puede aislarse de un organismo o una célula. El germoplasma proporciona material genético que tiene una composición molecular específica que proporciona la base física para algunas o todas las propiedades genéticas de un organismo o cultivo celular.

Un "mapa genético" es una descripción de relaciones de enlace genéticas entre loci en uno o más cromosomas dentro de una especie dada, generalmente representada de una forma diagramática o tabular. Para cada mapa genético, se miden las distancias entre loci mediante las frecuencias de recombinación entre ellos. Se pueden detectar las recombinaciones entre locus utilizando diversos marcadores. Un mapa genético es un producto de la población cartografiada, tipos de marcadores utilizados y el potencial polimórfico de cada marcador entre diferentes poblaciones. El orden y las distancias genéticas entre locus pueden variar entre un mapa genético y otro.

El término "genotipo", tal y como se utiliza en la presente, se refiere a la constitución genética de un individuo (o un grupo de individuos) en uno o más loci genéticos, a diferencia del rasgo observable y/o detectable y/o manifestado (el fenotipo). El genotipo está definido por el alelo o los alelos de uno o más loci conocidos que el individuo ha heredado de sus progenitores. El término genotipo se puede utilizar para referirse a la constitución genética de un individuo en un único locus, en múltiples locus o, más generalmente, el término genotipo se puede utilizar para referirse a los caracteres genéticos de un individuo para todos los genes en su genoma. Los genotipos se pueden caracterizar de forma indirecta, por ejemplo, empleando marcadores y/o se pueden caracterizar de forma

directa mediante la determinación de la secuencia de los ácidos nucleicos.

El término "germoplasma", tal como se utiliza en la presente, se refiere a material genético de un individuo o procedente de este (por ejemplo, una planta), un grupo de individuos (por ejemplo, una línea, variedad o familia vegetal), o un clon derivado de una línea, variedad, especie o cultivo. El germoplasma puede ser parte de un organismo o célula, o puede estar separado del organismo o célula. En general, el germoplasma proporciona material genético con una composición molecular específica que proporciona una base física para algunas o todas las cualidades hereditarias de un organismo o cultivo celular. Tal como se utiliza en la presente, germoplasma se puede referir a semillas, células (incluyendo protoplastos y callos) o tejidos a partir de los cuales pueden crecer nuevas plantas, así como partes de plantas que se pueden cultivar en una planta entera (por ejemplo, tallos, brotes, raíces, hojas, etc.).

Un "haplotipo" es el genotipo de un individuo a una pluralidad de loci genéticos, es decir, una combinación de alelos. Típicamente, los loci genéticos que definen un haplotipo están física y genéticamente enlazados, es decir, en el mismo segmento de cromosoma. El término "haplotipo" se puede referir a polimorfismos en un locus particular tal como un locus de un solo marcador, o polimorfismos en múltiples loci a lo largo de un segmento cromosómico.

El término "heterocigótico", tal como se utiliza en la presente, se refiere a un estado genético en el que se encuentran alelos diferentes en el loci correspondiente en cromosomas homólogos.

El término "homocigótico", tal como se utiliza en la presente, se refiere a un estado genético en el que se

encuentran alelos idénticos en el loci correspondiente en cromosomas homólogos.

Tal como se utiliza en la presente, el término "híbrido" se refiere a una semilla y/o planta producida cuando se cruzan al menos dos parentales genéticamente distintos.

El término "endogámico", tal como se utiliza en la presente, se refiere a una planta o variedad sustancialmente homocigótica. El término se puede referir a una planta o variedad que es sustancialmente homocigótica a lo largo de todo su genoma o que es sustancialmente homocigótica con respecto a una porción del genoma que es de un interés particular.

El término "indel", tal como se utiliza en la presente, se refiere a una inserción o delección en un par de secuencias de nucleótidos, donde de la primera secuencia se dice que tiene una inserción respecto a una segunda secuencia o de la segunda secuencia se dice que tiene una delección respecto a la primera secuencia.

Los términos y expresiones "introgresión", "introgresar" y "sometido a introgresión", tal como se utilizan en la presente, se refieren tanto a la transmisión natural como artificial de un alelo deseado o una combinación de alelos deseados de un locus genético o loci genéticos desde un acervo genético a otro. Por ejemplo, un alelo deseado en un locus específico se puede transmitir a al menos un descendiente mediante un cruce sexual entre dos progenitores de la misma especie, donde al menos uno de los progenitores tiene el alelo deseado en su genoma. Como alternativa, por ejemplo, la transmisión de un alelo se puede producir mediante la recombinación entre dos genomas donadores, por ejemplo, en un protoplasto fusionado, donde al menos uno de los protoplastos donadores posee el alelo deseado en su

genoma. El alelo deseado puede ser un alelo seleccionado de un marcador, un QTL, un transgén o similares. La descendencia que comprende el alelo deseado se puede retrocruzar repetidamente hasta obtener una línea que tiene un acervo genético deseado y seleccionar en función del alelo deseado, para que el resultado consista en fijar el alelo deseado en el acervo genético deseado. Por ejemplo, un marcador asociado con tolerancia potenciada a ASR puede someterse a introgresión desde un donador en un progenitor recurrente que no es resistente a la enfermedad. Los descendientes resultantes se podrían entonces retrocruzar y seleccionar repetidamente hasta que la descendencia posea el o los alelos de tolerancia a ASR en el fondo del progenitor recurrente.

Tal como se utiliza en la presente, el término "enlace" se refiere al grado con el que un locus de un marcador está asociado con otro locus de un marcador o algunos otros locus (por ejemplo, un locus de tolerancia a la ASR). La relación de enlace entre un marcador molecular y un fenotipo se puede dar como "probabilidad" o "probabilidad ajustada". El ligamiento se puede expresar como un intervalo o límite deseado. Por ejemplo, en algunas realizaciones, cualquier marcador está ligado (genética y físicamente) con cualquier otro marcador cuando los marcadores están separados por menos de aproximadamente 50, 40, 30, 25, 20, o 15 unidades de mapa (o cM). Por ejemplo, realizaciones de la invención de este documento, proporcionan locus marcadores ligados estrechamente a intervalos cromosómicos resistentes a ASR que comprenden una secuencia de nucleótidos de una cualquiera de las SEQ ID NO: 2, 11, 12 o 13.

En algunos aspectos de la presente invención, es ventajoso definir un intervalo de ligamiento entre corchetes, por ejemplo de aproximadamente 10 cM y aproximadamente 20 cM,

de aproximadamente 10 cM y aproximadamente 30 cM o de aproximadamente 10 cM y aproximadamente 40 cM. Cuanto más estrecho sea el ligamiento de un marcador a un segundo locus, mejor indicador será el marcador para este segundo locus. Por lo tanto, "loci estrechamente relacionados" tales como un locus de marcador y un segundo locus muestran una frecuencia de recombinación inter-locus de aproximadamente 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, o 2% o menor. En algunas realizaciones, los locus relevantes muestran una frecuencia de recombinación de aproximadamente un 1 % o menos, por ejemplo, aproximadamente un 0.75 %, 0.5 %, 0.25 % o menos. Dos locus que están ubicados en el mismo cromosoma, y una distancia tal que la recombinación entre los dos locus se produce a una frecuencia de menos de aproximadamente un 10 % (por ejemplo, aproximadamente un 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, 0.75%, 0.5%, o 0.25%, o menos) también puede decirse que están "proximales" entre sí. Como un cM es la distancia entre dos marcadores que muestran una frecuencia de recombinación del 1 %, cualquier marcador está estrechamente ligado (genética y físicamente) a cualquier otro marcador que se encuentre muy cercano, por ejemplo, a una distancia de aproximadamente 10 cM o menos. Dos marcadores estrechamente ligados en el mismo cromosoma pueden estar situados aproximadamente 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.75, 0.5 o 0.25 cM o menos uno de otro.

La expresión "desequilibrio de enlace", tal como se utiliza en la presente, se refiere a una segregación no aleatoria de rasgos o loci genéticos (o ambos). En cualquier caso, el desequilibrio de enlace implica que los loci relevantes estén dentro de una proximidad física suficiente a lo largo de una longitud de un cromosoma de modo que se segregan juntas con una frecuencia mayor que la aleatoria (es decir, no aleatoria) (en el caso de rasgos co-segregantes,

los loci en los que se fundamentan los rasgos están en suficiente proximidad uno de otro). Los marcadores que muestran desequilibrio de ligamiento se consideran ligados. Los locus ligados se cosegregan más de un 50 % de las veces, por ejemplo, de aproximadamente un 51 % a aproximadamente 100 % de las veces. En otras palabras, dos marcadores que se cosegregan tienen una frecuencia de recombinación menor de un 50 % (y, por definición, están separados por menos de 50 cM en el mismo cromosoma). Como se utiliza en la presente, el ligamiento puede ser entre dos marcadores, o como alternativa, entre un marcador y un fenotipo. Un locus marcador puede estar "asociado con" (ligado a) un rasgo, por ejemplo, roya de la soja asiática. El grado de relación de un marcador molecular con un rasgo fenotípico se mide, por ejemplo, como una probabilidad estadística de cosegregación de ese marcador molecular con el fenotipo.

El desequilibrio de ligamiento se evalúa muy habitualmente usando la medida r^2 , que se calcula usando la fórmula descrita por Hill y Robertson, *Theor. Appl. Genet.* 38:226 (1968). Cuando $r^2=1$, existe un desequilibrio de ligamiento completo entre los dos locus de marcadores, lo que significa que los marcadores no se han separado por recombinación y tienen la misma frecuencia alélica. Valores para r^2 por encima de 1/3 indican que un desequilibrio de ligamiento lo suficientemente fuerte es útil para el cartografiado. Ardlie et al., *Nature Reviews Genetics* 3:299 (2002). Por lo tanto, alelos se encuentran en desequilibrio de ligamiento cuando valores de r^2 entre locus de marcadores por pares son mayores que o iguales a aproximadamente 0.33, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, o 1.0.

Tal como se utiliza en la presente, la expresión "equilibrio de enlace" describe una situación en la que se

segregan independientemente dos marcadores, es decir, se escogen entre la progenie al azar. Se considera que los marcadores que muestran un equilibrio de ligamiento no están ligados (estén o no en el mismo cromosoma).

Un "locus" es una posición en un cromosoma en el que está ubicado un gen o marcador o alelo. En algunas realizaciones, un locus puede abarcar uno o más nucleótidos.

Como se usa en este documento, la expresiones "marcador" y "marcador genético" se usan indistintamente para hacer referencia a un nucleótido y/o una secuencia de nucleótidos que se ha asociado con un fenotipo, rasgo o tipo de rasgo. En algunas realizaciones, un marcador puede estar asociado con un alelo o alelos de interés y puede ser indicativo de la presencia o ausencia del alelo o alelos de interés en una célula u organismo. Un marcador puede ser, aunque sin limitación, un alelo, un gen, un haplotipo, un polimorfismo de longitud de fragmento de restricción (RFLP), una repetición de secuencia simple (SSR), ADN polimórfico amplificado aleatorio (RAPD), secuencias polimórficas amplificadas escindidas (CAPS) (Rafalski y Tingey, Trends in Genetics 9:275 (1993)), un polimorfismo de longitud de fragmento amplificado (AFLP) (Vos et al., Nucleic Acids Res. 23:4407 (1995)), un polimorfismo mononucleotídico (SNP) (Brookes, Gene 234:177 (1993)), una región amplificada caracterizada por secuencia (SCAR) (Paran y Michelmore, Theor. Appl. Genet. 85:985 (1993)), un sitio marcado por secuencia (STS) (Onozaki et al., Euphytica 138:255 (2004)), un polimorfismo de conformación monocatenaria (SSCP) (Orita et al., Proc. Natl. Acad. Sci. USA 86:2766 (1989)), una repetición de secuencia simple interna (ISSR) (Blair et al., Theor. Appl. Genet. 98:780 (1999)), un polimorfismo amplificado de retrotransposón interno (IRAP), un

polimorfismo amplificado de retrotransposón-microsatélite (REMAP) (Kalendar et al., Theor. Appl. Genet. 98:704 (1999)), un intervalo cromosómico o un producto de escisión de ARN (tal como una marca Lynx). Un marcador puede estar presente en el componente genómico o en ácidos nucleicos expresados (por ejemplo, EST). El término marcador también se puede referir a ácidos nucleicos utilizados como sondas o cebadores (por ejemplo, pares de cebadores) para uso en amplificación, hibridación y/o detección de moléculas de ácido nucleico de acuerdo con métodos bien conocidos en la técnica. En la técnica se conoce un gran número de marcadores moleculares de soja y están publicados o están disponibles de diversas fuentes tal como el recurso de internet SoyBase.

Los marcadores que corresponden a polimorfismos genéticos entre miembros de una población pueden detectarse mediante métodos bien establecidos en la técnica. Estos incluyen, p. ej., secuenciación de ácidos nucleicos, métodos de hibridación, métodos de amplificación (p. ej., métodos de amplificación específica de la secuencia basados en PCR), detección de polimorfismos en la longitud de fragmentos de restricción (RFPL), detección de marcadores de isozima, detección de polimorfismos de polinucleótidos mediante la hibridación específica del alelo (ASH), detección de secuencias variables amplificadas del genoma vegetal, detección de la replicación de la secuencia autosuficiente, detección de repeticiones de secuencia sencilla (SSR), detección de polimorfismos de un único nucleótido (SNP) y/o detección de polimorfismos en la longitud de los fragmentos amplificados (AFLP). También existe constancia de métodos muy establecidos para la detección de etiquetas de secuencia expresadas (EST) y marcadores SSR obtenidos a partir de las

secuencias EST y ADN polimórfico amplificado aleatoriamente (RAPD).

Un "alelo de marcador", también descrito como un "alelo de un locus de marcador" puede aludir a una pluralidad de secuencias de nucleótidos polimórficas que se encuentran en un locus de marcador en una población que es polimórfica para el locus de marcador.

"Selección asistida por marcador" (MAS) es un proceso mediante el cual se seleccionan fenotipos en base a genotipos de marcadores. En algunas realizaciones, los genotipos de marcadores se utilizan para identificar plantas que se seleccionarán para un programa de reproducción o para la plantación. En algunas realizaciones, los genotipos de marcadores se utilizan para identificar plantas que no se seleccionarán para un programa de reproducción o para la plantación, es decir, plantas contraseleccionadas), permitiendo que se separen de la población de reproducción/plantación.

Tal como se utiliza en la presente, las expresiones "locus de marcador" y "loci de marcadores" se refieren a una ubicación o ubicaciones en el cromosoma específicas en el genoma de un organismo en que se pueden encontrar un marcador o marcadores. Un locus marcador se puede utilizar para rastrear la presencia de un segundo locus ligado, por ejemplo, un locus ligado que codifica un rasgo fenotípico o contribuye a la expresión de este. Por ejemplo, se puede utilizar un locus marcador para monitorizar la segregación de alelos en un locus, tales como un QTL o un gen único, que están genética o físicamente ligados al locus marcador.

Tal como se utiliza en la presente, la expresión "sonda de marcador" y el término "sonda" se refieren a una secuencia de nucleótidos o molécula de ácido nucleico que se puede

utilizar para detectar la presencia de uno o más alelos particulares dentro de un locus de marcador (p, ej., una sonda de ácido nucleico que es complementaria a todo o a una porción del marcador o locus de marcador a través de hibridación del ácido nucleico). Pueden usarse sondas marcadoras que comprenden aproximadamente 8, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100 o más nucleótidos contiguos para la hibridación de ácidos nucleicos. Alternativamente, en algunos aspectos, una sonda de marcador se refiere a una sonda de cualquier tipo que es capaz de distinguir (es decir, genotipo) el alelo particular que está presente en un locus de marcador.

Tal como se utiliza en la presente, las expresiones "marcador molecular" o "marcador genético" se puede utilizar para aludir a un marcador genético, tal como se define arriba, o un producto codificado del mismo (p. ej, una proteína) utilizado como un punto de referencia cuando se identifica un locus enlazado. Se puede obtener un marcador molecular a partir de una secuencia de nucleótidos genómicos o a partir de secuencias de nucleótidos expresadas (por ejemplo, a partir de un ARN que ha experimentado corte y empalme, un ADNc, etc.). La expresión también se refiere a secuencias de nucleótidos complementarias a las secuencias de los marcadores, o que las flanquean, tales como secuencias de nucleótidos utilizadas como sondas y/o cebadores capaces de amplificar la secuencia del marcador. Las secuencias de nucleótidos son "complementarias" cuando se hibridan específicamente en solución, p. ej., de acuerdo con las reglas de emparejamiento de bases de Watson-Crick. Algunos de los marcadores descritos en la presente también se denominan marcadores de hibridación cuando se ubican en una región indel. Esto se debe a que la región de inserción es, por

definición, un polimorfismo frente a una planta sin la inserción. Por lo tanto, el marcador únicamente necesita indicar si la región indel está presente o ausente. Se puede utilizar cualquier tecnología de detección de marcadores adecuada, tal como un marcador de hibridación, por ejemplo, se utiliza la tecnología SNP en los ejemplos que se proporcionan en la presente.

Una "variedad de soja que se produce de forma no natural" es cualquier variedad de soja que no existe de forma natural en la naturaleza. Una "variedad de soja que se produce de forma no natural" se puede producir por cualquier método conocido en el técnica, incluyendo, pero no limitado a transformar una planta o germoplasma de soja, transfectar una planta o germoplasma de soja y cruzar una variedad de soja que se produce de forma natural con una variedad de soja que no se produce de forma natural. En algunas realizaciones, una "variedad de soja que no se produce de forma natural" puede comprender una o más secuencias de nucleótidos heterólogos. En algunas realizaciones, una "variedad de soja que no se produce de forma natural" puede comprender una o más copias que no se producen de forma natural de una secuencia de nucleótidos que se produce de forma natural (es decir, copias extrañas de un gen que se produce de forma natural en la soja). En algunas realizaciones, una "variedad de soja que no es de origen natural" puede comprender una combinación no natural de dos o más secuencias de nucleótidos que de origen natural (es decir, dos o más genes de origen natural que no existen de forma natural en la misma soja, por ejemplo, genes no encontrados en líneas de *Glycine max*).

El término "transformación" y variaciones gramaticales del mismo se usan en la presente invención para referirse a, por ejemplo, la transferencia de fragmentos de ácido nucleico

en el genoma de un organismo huésped, obteniendo de ese modo herencia genéticamente estable. El organismo huésped que comprende los fragmentos de ácido nucleico transformado se denomina un organismo "transgénico". El término "célula huésped" y variaciones gramaticales del mismo se refieren a una célula en la que la transformación de un constructo de ADN recombinante tiene lugar y puede incluir células de levadura, células bacterianas y/o células vegetales.

Tal como se usa en el presente documento, el término "transgénico" y variaciones gramaticales del mismo se refieren a una planta, incluyendo cualquier parte derivada de la planta, tal como una célula, tejido o órgano, en el que un ácido nucleico exógeno (por ejemplo, un constructo recombinante, vector o casete de expresión que comprende uno o más ácidos nucleicos) se integra en el genoma por un método de modificación por ingeniería genética, tal como transformación por *Agrobacterium*. A través de la tecnología génica, el ácido nucleico exógeno se integra de manera estable en los cromosomas, de modo que la próxima generación también puede ser transgénica. Tal como se usa en el presente documento, "transgénico" y variaciones gramaticales del mismo también abarcan tratamientos biológicos, que incluyen hibridación de plantas y/o recombinación natural.

Tal como se utiliza en la presente, el término "cebador" se refiere a un oligonucleótido que es capaz de hibridar a un objetivo de ácido nucleico y servir de ese modo como punto de iniciación de la síntesis de ADN, cuando se lo coloca en condiciones bajo las cuales se induce la síntesis de un producto de extensión de cebadores (por ejemplo, en presencia de nucleótidos y un agente de polimerización como la polimerasa de ADN y una temperatura y pH apropiados). Un cebador (en algunas realizaciones un cebador de extensión y

en otras realizaciones un cebador de amplificación) en algunas realizaciones es de una sola hebra para obtener la mayor eficiencia de extensión y/o amplificación. En algunas realizaciones, el cebador es un oligodesoxirribonucleótido. Un cebador típicamente es lo suficientemente largo como para cebar la síntesis de productos de extensión y/o amplificación en presencia de un agente para polimerización. La longitud mínima del cebador puede depender de varios factores, inclusive, de manera no exclusiva, de la temperatura y composición (contenido A/T frente a G/C) del cebador. En el contexto de cebadores de amplificación, estos se proporcionan típicamente como un par de cebadores bidireccionales que consisten en un cebador directo y un cebador inverso o se proporcionan como un par de cebadores directos como se utiliza comúnmente en la técnica de amplificación de ADN, por ejemplo, en la amplificación por PCR. Como tal, se entenderá que el término "cebador" tal como se utiliza en la presente, se puede referir a más de un cebador, particularmente en el caso en el que haya algo de ambigüedad en la información que se refiere a la(s) secuencia(s) terminal(es) de la región diana que se quiere amplificar. Por lo tanto, un "cebador" puede incluir un conjunto de oligonucleótidos cebadores que contienen secuencias que representan las posibles variaciones en la secuencia o incluye nucleótidos que permiten un típico emparejamiento de base. Los cebadores pueden prepararse por cualquier método adecuado conocido en la técnica. Los métodos para preparar oligonucleótidos de secuencia específica son conocidos en la técnica e incluyen, por ejemplo, la clonación y restricción de secuencias apropiadas y la síntesis química directa. Los métodos de síntesis química pueden incluir, por ejemplo, el método de fosfo di- o triéster, el método de dietilfosforamidato y el método en soporte sólido divulgado

en la patente de Estados Unidos n.º 4 458 066. Se pueden marcar los cebadores, si se desea, incorporando porciones detectables, por ejemplo, mediante porciones espectroscópicas, fluorescentes, fotoquímicas, bioquímicas, inmunoquímicas o químicas. Pueden crearse cebadores de diagnóstico (es decir, que pueden identificar o seleccionar basándose en la presencia de alelos resistentes a ASR) para resistencia a la ASR para cualquiera SNP favorable como se describe en la Tabla 1. El método de PCR está bien descrito en los manuales y es conocido para los expertos en la materia. Después de la amplificación mediante PCR, los polinucleótidos diana se pueden detectar mediante hibridación con un polinucleótido sonda que forma un híbrido estable con la secuencia diana, en condiciones de hibridación y de lavado rigurosas a moderadamente rigurosas. Si se espera que las sondas sean esencialmente complementarias por completo (es decir, un 99% o más) con respecto a la secuencia objetivo, se pueden utilizar condiciones rigurosas. Si se prevé algún desacierto, por ejemplo si se esperan hebras variantes con el resultado de que la sonda no sea completamente complementaria, se puede reducir la rigurosidad de la hibridación. En algunas realizaciones, las condiciones se eligen para descartar las uniones no específicas/fortuitas. Las condiciones que afectan la hibridación, y que seleccionan para que no se produzcan las uniones no específicas se conocen en la técnica y se describen, por ejemplo, en Sambrook y Russell (2001). Molecular Cloning: A Laboratory Manual, tercera edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, Estados Unidos de América. Generalmente la concentración de sal más baja y la temperatura más elevada aumentan la rigurosidad de las condiciones de hibridación.

Tal como se utilizan en la presente el término "fenotipo" y las expresiones "rasgo fenotípico" o "rasgo" se refieren a uno o más rasgos y/o manifestaciones de un organismo. El fenotipo puede ser una manifestación que es observable a simple vista o por cualesquiera otros medios de evaluación conocidos en la técnica, p. ej., microscopía, análisis bioquímico o un ensayo electromecánico. En algunos casos, un fenotipo o rasgo es directamente controlado por un solo gen o locus genético, es decir, un "rasgo de un solo gen". En otros casos, un fenotipo o rasgo es el resultado de varios genes. Se señala que, tal como se utiliza en la presente, la expresión "fenotipo resistente a enfermedades" tiene en cuenta condiciones medioambientales que pudieran afectar a la enfermedad respectiva de modo que el efecto es real y reproducible.

Como se usa en este documento, el término "planta" y variaciones gramaticales del mismo puede hacer referencia a una planta completa, cualquier parte de la misma, o un cultivo celular o tisular derivado de una planta. Por lo tanto, el término "planta" puede referirse a cualquiera de: plantas enteras, componentes u órganos de plantas (p. ej., raíces, tallos hojas, brotes, flores, vainas, etc.), tejidos vegetales, semillas y/o células vegetales. Las células vegetales son células de plantas, que se obtienen directamente de semillas o plantas, o derivadas de cultivos celulares obtenidos de plantas. Las progenies, variantes y mutantes de plantas regeneradas están dentro del alcance de la presente invención, siempre que estas partes comprendan los genes de resistencia introducidos. Por lo tanto, la expresión "planta de soja" puede referirse a planta de soja entera, a una o más partes de una planta de soja (por ejemplo, raíces, puntas de las raíces, tallos, hojas, brotes,

flores, vainas, semillas, cotiledones, etc.), células de la planta de soja, protoplastos de la planta de soja y/o callos de la planta de soja.

La expresión "parte de planta", tal como se utiliza en la presente, incluye, pero sin limitación, embriones, polen, semillas, hojas, flores (incluyendo pero sin limitarse a anteras, óvulos y similares) frutos, tallos o ramas, raíces, rizomas, células incluidas células que están intactas en plantas y/o partes de plantas, protoplastos, cultivos de células vegetales, callos de plantas, agregados vegetales y similares. Así una parte de planta incluye cultivo de tejido de soja a partir de los que se pueden regenerar plantas de soja. Además, la expresión "célula vegetal", tal como se utiliza en la presente, se refiere a una unidad estructural y fisiológica de la planta que comprende una pared celular y también se puede denominar protoplasto. Una célula vegetal de la presente invención puede estar en forma de una única célula aislada o puede ser una célula cultivada o puede formar parte de una unidad más organizada tal como, por ejemplo, un tejido vegetal o un órgano vegetal.

El término "polimorfismo", tal como se utiliza en la presente, se refiere a una variación en la secuencia de nucleótidos en un locus, donde dicha variación es demasiado común para ser debida únicamente a una mutación espontánea. Un polimorfismo puede ser un polimorfismos de nucleótido único (SNP) o un polimorfismo de inserción/delección, al que también se alude en la presente como un "indel". Además, la variación puede seguir un perfil de transcripción o un patrón de metilación. El sitio o los sitios polimórficos de una secuencia de nucleótidos se pueden determinar comparando las secuencias de nucleótidos en uno o más locus para dos o más entradas de germoplasma.

Como se usa en este documento, la expresión "ligado estrechamente" se refiere a marcadores ligados que presentan una frecuencia de cruce con un marcador dado de aproximadamente un 10 % o menos (por ejemplo, el marcador dado está dentro de aproximadamente 10 cM de un marcador de ASR ligado estrechamente). Dicho de otro modo, los locus ligados estrechamente cosegregan al menos aproximadamente un 90 % de las veces. Con respecto a la posición física en un cromosoma, los marcadores ligados estrechamente pueden estar separados, por ejemplo, en aproximadamente 1 megabase (Mb; 1 millón de nucleótidos), aproximadamente 500 kilobases (Kb; 1000 nucleótidos), aproximadamente 400 Kb, aproximadamente 300 Kb, aproximadamente 200 Kb, aproximadamente 100 Kb, aproximadamente 50 Kb, aproximadamente 25 Kb, aproximadamente 10 Kb, aproximadamente 5 Kb, aproximadamente 4 Kb, aproximadamente 3 Kb, aproximadamente 2 Kb, aproximadamente 1 Kb, aproximadamente 500 nucleótidos, aproximadamente 250 nucleótidos o menos.

El término "población", tal como se utiliza en la presente, se refiere a una colección genéticamente heterogénea de plantas que comparten una derivación genética común.

Tal como se utiliza en la presente, los términos "progenie" y "planta de progenie" se refieren a una planta generada a partir de una reproducción vegetativa o sexual de una o más plantas parentales. Puede obtenerse una planta de progenie mediante la clonación o autofecundación de una única planta parental, o mediante cruzamiento de dos plantas parentales.

La expresión "secuencia de referencia", tal como se utiliza en la presente, se refiere a una secuencia de nucleótidos definida que se utiliza como base para la

comparación de secuencias. Por ejemplo, la secuencia de referencia para un marcador se obtiene genotipando un número de líneas en el locus o los loci de interés, alineando las secuencias de nucleótidos en un programa de alineación de secuencias y, posteriormente, obteniendo la secuencia consenso de la alineación. Así pues, una secuencia de referencia identifica los polimorfismos en alelos en un locus. Una secuencia de referencia puede que no sea una copia de una secuencia de ácido nucleico real procedente de cualquier organismo concreto; sin embargo, es útil para diseñar cebadores y sondas para polimorfismos reales en el locus o los loci.

Tal como se utiliza en la presente, las expresiones "tolerancia a la enfermedad" y "resistente a la enfermedad" se refieren a la capacidad de una planta de perdurar y/o progresar, a pesar de estar infectada con una enfermedad respectiva. Cuando se utiliza en referencia a germoplasma, las expresiones se refieren a la capacidad de una planta que surge del germoplasma de perdurar y/o progresar a pesar de estar infectada con una enfermedad respectiva. En algunas realizaciones, plantas de soja resistentes a enfermedades infectadas pueden proporcionar también (o casi también) plantas de soja no infectadas. En general, una planta o germoplasma se etiqueta como "resistente a enfermedades", si muestra "resistencia a patógenos potenciada".

Un "alelo favorable" de un marcador es un alelo de marcador que segrega con el fenotipo de la planta desfavorable, proporcionando, por lo tanto, el beneficio de identificar plantas que pueden ser separadas de un programa de reproducción o una plantación. Por ejemplo, podrían eliminarse de un programa de cultivo selectivo de plantas

líneas de plantas que portan alelos desfavorables para resistencia a la ASR.

Roya en plantas

La presente invención pertenece a proteínas relacionadas a resistencia a la roya en plantas, secuencia de ácido nucleico que codifica para tales proteínas, y usos de las mismas. La proteína y el gen codificante de la misma y el método dado a conocer en la invención pueden usarse para proteger las plantas de los patógenos de roya. En determinadas realizaciones, la roya es roya de planta leguminosa.

En otras realizaciones, la roya de planta leguminosa es roya de la soja. El patógeno de roya de la soja puede ser *Phakopsora pachyrhizi* o *Phakopsora meibomiae*. Tal como se usa a continuación y en las reivindicaciones cada referencia a la roya de la soja incluye roya de la soja asiática.

En una realización específica de la presente invención, el patógeno de roya de la soja es específicamente la raza fisiológica SS4 de *Phakopsora pachyrhizi*.

En realizaciones preferidas, la presente invención pertenece a proteínas relacionadas con la resistencia a la roya en una planta leguminosa.

La planta leguminosa puede ser plantas *Glycine*, plantas *Cicer*, plantas *Cajanus*, plantas *Lablab*, plantas *Medicago*, plantas *Phaseolus*, plantas *Pisum*, planta *sPueraria*, plantas *Trifolium* o plantas *Vigna*.

Las plantas *Glycine* pueden ser *Glycine arenaria*, *Glycine argyrea*, *Glycine cyrtoloba*, *Glycine canescens*, *Glycine clandestine*, *Glycine curvata*, *Glycine falcata*, *Glycine latifolia*, *Glycine microphylla*, *Glycine pescadrensis*, *Glycine stenophita*, *Glycine syndetica*, *Glycine soja* Seib. et Zucc.,

Glycine max (L.) Merrill., *Glycine tabacina* o *Glycine tomentella*.

Las plantas *Cicer* pueden ser *Cicerarietinum*, *Cicer echinospermum*, *Cicer reticulatum* o *Cicer pinnatifidum*.

Las plantas *Lablab* pueden ser *Lablab purpureus*.

Las plantas *Medicago* pueden ser *Medicago truncatula* o *Medicago sativa*.

Las plantas *Phaseolus* pueden ser *Phaseolus vulgaris*, *Phaseolus lunatus*, *Phaseolus acutifolius*, o *Phaseolus coccineus*.

Las plantas *Pisum* pueden ser *Pisum abyssinicum*, *Pisum sativum*, *Pisum elatius*, *Pisum fulvum*, *Pisum transcaucasicum*, o *Pisum humile*.

Las plantas *Pueraria* pueden ser *Pueraria lobata*.

Las plantas *Trifolium* pueden ser *Trifolium aureum* o *Trifolium occidentale*.

Las plantas *Vigna* pueden ser *Vigna unguiculata*, *Vigna dalzelliana*, *Vigna oblongifolia*, *Vigna parkeri*, *Vignafilicaulis*, *Vigna kirkii*, *Vigna luteola*, *Vigna radiata*, *Vigna trilobata*, *Vigna luteola*, o *Vigna mungo*.

Además, la planta leguminosa puede ser cualquiera de las siguientes: soja, alfalfa, trébol, guisante, frijol común, lenteja, lupino, árbol de árbol de Ghaf, algarrobo, soja, maní o tamarindo.

En una realización específica de la presente invención, la planta es específicamente una variedad de la soja Tianlong No.1 (天隆一号).

La proteína y el gen codificante de la misma y el método dado a conocer en la invención pueden usarse para proteger las plantas de los patógenos de roya. La interacción entre huéspedes y patógenos pueden describirse un continuo de

"inmunidad" a la "resistencia parcial" a la "susceptibilidad".

El método dado a conocer en la invención pueden aumentar, potenciar o mejorar la resistencia de soja a un hongo parásito biotrófico obligatorio *Phakopsora pachyrhizi* (el patógeno principal de la roya) o a *Phakopsora meibomiae*. Por ejemplo, la resistencia aumentada o potenciada frente a los patógenos de la roya puede compararse con el impacto de patógenos en plantas susceptibles. Las manifestaciones de resistencia aumentada o aumentada pueden estar en diferentes niveles, pero están relacionadas con los síntomas de la enfermedad (tal como el color de las manchas de la enfermedad) y la morbilidad observada en plantas o tejidos vegetales (por ejemplo, hojas). Puede proporcionarse el valor de la inmunidad, resistencia y susceptibilidad. Por ejemplo, el valor de la resistencia indica el grado de resistencia de plantas a enfermedades de las plantas (por ejemplo, roya). También pueden usarse los valores para comparar el grado de resistencia entre, por ejemplo, plantas de interés (por ejemplo, plantas leguminosas transgénicas) y plantas susceptibles (por ejemplo, Tianlong No.1 (天隆一号) o Williams) o estándares de referencia.

La proteína y el gen codificante de la misma y los métodos dados a conocer en la presente invención se refiere al aislamiento de un gen de resistencia de especies de leguminosas y posterior transferencia del gen de resistencia a una planta receptora, tal como soja, para proporcionar o potenciar la resistencia a *Phakopsora pachyrhizi*. Una realización de la solicitud incluye transferir el gen de resistencia a especies sexualmente compatibles o incompatibles para producir resistencia a enfermedades. El gen de resistencia de la presente invención puede usarse solo

o en la superposición con otros genes de resistencia o junto con genes de no resistencia para proporcionar o potenciar la resistencia de las especies receptoras frente a la roya.

Por tanto, el método transgénico dado a conocer en la presente invención puede usarse solo o en combinación con otras estrategias para producir o conferir resistencia a la roya en plantas. Otras estrategias disponibles incluyen, pero no se limitan a, bloquear la actividad funcional de los efectores, inhibir la absorción de patógenos o factores patógenos (por ejemplo, hongos) en las células huésped (por ejemplo, células vegetales) y/o fitomejoramiento con resistencia convencional.

El método dado a conocer en la presente invención puede proporcionar o potenciar la resistencia a la roya de plantas, de modo que el patógeno de la roya no puede reproducir o el coeficiente de reproducción del patógeno de la roya se reduce de manera significativa. Por tanto, el método de la presente invención puede aliviar uno o más síntomas (es decir, síntomas de la enfermedad) de roya de plantas leguminosas cuando se compara con plantas susceptibles o tolerantes al género *Phakopsora*. Las plantas a las que se hace referencia en la presente invención también incluyen plantas leguminosas transgénicas (por ejemplo, soja) en las que se introducen genes o proteínas de resistencia a la enfermedad mediante métodos de ingeniería genética para mejorar su resistencia a las enfermedades cuando se exponen a la roya de plantas leguminosas.

Las plantas, partes de plantas o células vegetales de la presente invención se derivan de plantas que incluyen, pero no se limitan a soja, alfalfa, trébol, guisante, frijol común, lenteja, lupino, árbol ghaf, algarroba, maní y tamarindo.

Las plantas de la invención pertenecen a *Leguminosae*. Los ejemplos de *Leguminosae* incluyen, pero no se limitan a, *Phaseolus* (por ejemplo, judía francesa, judía verde, *Phaseolus vulgaris*, *Phaseolus lunatus*, *Phaseolus acutifolius*, y *Phaseolus coccineus*); *Glycine* (por ejemplo, *Glycine soja*, y *Glycine max* (L.)); *Pisum* (por ejemplo, guisante sin vaina (a veces denominado guisante suave o guisante redondo, *Pisum sativum* (菜豌豆)), *Pisum sativum* (皱粒豌豆), *Pisum sativum* (糖荚豌豆) que también se conoce como guisante de nieve, guisante comestible o *Pisum granda*); maní (*Arachis hypogaea*), trébol (*Trifolium* spp.), alfalfa (*Medicago*), kudzu (*Pueraria lobata*), alfalfa común, alfalfa (*Medicago sativa*), garbanzo (*Cicer*), lenteja (*Lens culinaris*), y altramuces (*Lupinus*); veza (*Vicia*), haba (*Vicia faba*), almorta (*Lathyrus*) (por ejemplo, *Lathyrus sativus*, y *Lathyrus tuberosus*); fríjol de oro (por ejemplo, *Vigna aconitifolia*, *Vigna angularis*, *Vigna mungo*, *Vigna radiata*, *Vigna subterranea*, *Vigna umbellata*, *Vigna vexillata*, y *Vigna unguiculata* (también conocida como caupí largo o caupí)); *Cajanus cajan*, *Macrotyloma* (por ejemplo, *Macrotyloma geocarpum*, y *Macrotyloma uniflorum*); *Psophocarpus tetragonolobus*, *Sphenostylis stenocarpa*, frijoles negros egipcios, *Lablab purpureus*, *Pachyrhizus erosus*, y *Cyamopsis tetragonolobus*; y/o *Canavalia* (por ejemplo, *Canavalia ensiformis*, y *Canavalia gladiata*).

Intervalos cromosómicos

Una realización de la invención se refiere a intervalos cromosómicos derivados de la cepa de *Glycine max* SX6907. Estos intervalos cromosómicos cuando se introducen en una planta (por ejemplo, una soja tal como la cepa de *Glycine max* Williams 82) son suficientes para conferir una resistencia aumentada a la roya, tal como por ejemplo resistencia a la

roya de la soja asiática ("ASR"), en comparación con una planta de control que no comprende dicho intervalo cromosómico.

SEQ ID NO 11, 12, o 13 son intervalos cromosómicos derivados de la cepa de *Glycine max* SX6907. Los estudios de mapeo genético indican que la cepa de *Glycine max* SX6907 contiene intervalos cromosómicos asociados con la resistencia a la ASR (por ejemplo, correspondientes a S SEQ ID NO 11, 12, o 13). Estos intervalos cromosómicos o partes de los mismos pueden introducirse (*es decir*, introgresar mediante el uso de fitomejoramiento asistido por marcadores (MAB), o mediante el uso de introducción de GM o GE) en líneas *Glycine max* para crear líneas de *Glycine max* resistentes a diversas enfermedades tales como ASR. Por ejemplo, estos intervalos cromosómicos pueden introducirse en la línea de *Glycine max* Williams 82.

La Tabla 1 indica polimorfismos de nucleótido único (SNP) dentro de las SEQ ID NO: 11, 12 o 13 que están asociados con la resistencia a la ASR.

Tabla 1: Posiciones de SNP dentro de la SEQ ID NO 11, 12, o 13 que están asociadas con resistencia aumentada a la ASR

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56601565	A	AC
Gm18	56601634	T	C
Gm18	56602295	G	T
Gm18	56602866	G	GA
Gm18	56602929	T	A
Gm18	56604482	G	GATATATA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
			T
Gm18	56604846	C	T
Gm18	56605172	GTATATTTATATATATATATA TATA (SEQ ID NO: 14)	G
Gm18	56605243	G	T
Gm18	56605859	A	AATATGCT TT (SEQ ID NO: 15)
Gm18	56605869	ATATG	A
Gm18	56605893	A	G
Gm18	56605895	A	G
Gm18	56605897	A	G
Gm18	56605899	A	G
Gm18	56605901	A	G
Gm18	56606593	TA	T
Gm18	56606764	C	T
Gm18	56607175	T	C
Gm18	56608059	T	C
Gm18	56609308	C	T
Gm18	56609424	A	G
Gm18	56609508	A	T
Gm18	56609511	C	G
Gm18	56609554	T	G
Gm18	56610268	T	C
Gm18	56610625	T	TAAG
Gm18	56611386	T	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56612772	C	CTATATA
Gm18	56613273	C	T
Gm18	56613352	A	G
Gm18	56613400	C	T
Gm18	56613421	C	T
Gm18	56613787	TA	T
Gm18	56614286	T	A
Gm18	56614344	G	GA
Gm18	56614630	A	G
Gm18	56614906	G	GT
Gm18	56615544	CT	C
Gm18	56615626	A	ATG
Gm18	56615677	T	A
Gm18	56616069	C	A
Gm18	56616107	T	A
Gm18	56616637	T	C
Gm18	56616692	A	T
Gm18	56616778	T	A
Gm18	56616824	C	T
Gm18	56616825	T	G
Gm18	56616833	A	G
Gm18	56616839	C	T
Gm18	56616844	T	C
Gm18	56616894	A	G
Gm18	56616903	TGGGTCAACCTGAATGAGTCA GGGTTGACCCAAA (SEQ ID NO: 16)	T
Gm18	56616938	CGA	C

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56616943	AT	A
Gm18	56616965	ATT	A
Gm18	56616973	TAAAAA	T
Gm18	56616996	T	TAA
Gm18	56616997	T	A
Gm18	56617058	A	G
Gm18	56617072	AC	A
Gm18	56617119	C	A
Gm18	56617143	CTTTCTTTAATT (SEQ ID NO: 17)	C
Gm18	56617792	A	C
Gm18	56618125	C	G
Gm18	56618647	TA	T
Gm18	56621101	T	C
Gm18	56621162	C	A
Gm18	56621217	A	C
Gm18	56621385	G	A
Gm18	56621559	C	G
Gm18	56621731	G	A
Gm18	56621822	A	G
Gm18	56621836	T	C
Gm18	56621857	C	T
Gm18	56621899	T	C
Gm18	56621936	T	A
Gm18	56621937	G	GAA
Gm18	56621943	G	A
Gm18	56622058	A	AT
Gm18	56622285	T	C

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56622585	A	ATGG
Gm18	56622602	A	T
Gm18	56622637	A	G
Gm18	56622642	T	G
Gm18	56622853	T	A
Gm18	56622862	A	G
Gm18	56622906	G	T
Gm18	56622925	C	T
Gm18	56623192	TC	T
Gm18	56623667	A	G
Gm18	56623903	C	T
Gm18	56624060	C	T
Gm18	56624065	T	G
Gm18	56624185	AT	A
Gm18	56624351	AC	A
Gm18	56624404	A	AAAGT
Gm18	56624427	T	C
Gm18	56624464	A	G
Gm18	56624551	G	T
Gm18	56624606	C	A
Gm18	56624609	TTAA	T
Gm18	56624664	G	A
Gm18	56624955	C	T
Gm18	56625057	C	A
Gm18	56625513	A	G
Gm18	56625553	A	AATAT
Gm18	56625571	A	C
Gm18	56625591	C	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56625648	C	T
Gm18	56625668	AGAT	A
Gm18	56625720	C	T
Gm18	56625722	T	A
Gm18	56625799	A	G
Gm18	56625966	A	G
Gm18	56626008	G	A
Gm18	56626041	G	T
Gm18	56626052	C	T
Gm18	56626063	T	G
Gm18	56626080	T	G
Gm18	56626449	T	G
Gm18	56626566	G	A
Gm18	56626575	A	G
Gm18	56626903	T	C
Gm18	56626915	C	G
Gm18	56626974	ACATAC	A
Gm18	56627161	A	AATAT
Gm18	56627238	C	T
Gm18	56627264	A	C
Gm18	56627287	GT	G
Gm18	56627317	A	G
Gm18	56627338	G	A
Gm18	56627360	T	TA
Gm18	56627382	C	A
Gm18	56627386	C	T
Gm18	56627412	C	G
Gm18	56627575	A	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56627577	T	C
Gm18	56627857	T	G
Gm18	56627967	A	G
Gm18	56628395	A	G
Gm18	56628450	C	G
Gm18	56629061	CTG	C
Gm18	56629276	G	A
Gm18	56629622	T	C
Gm18	56629688	C	T
Gm18	56629764	T	A
Gm18	56630232	T	A
Gm18	56630293	AT	A
Gm18	56630337	C	T
Gm18	56630348	C	T
Gm18	56630383	C	G
Gm18	56630411	T	C
Gm18	56630490	G	A
Gm18	56630497	A	ATTCAAAA ATATTTTT TTAATAAT T (SEQ ID NO: 18)
Gm18	56630555	C	T
Gm18	56630572	A	G
Gm18	56630585	A	C
Gm18	56630748	T	C
Gm18	56630769	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56630804	T	A
Gm18	56630811	A	G
Gm18	56630892	T	C
Gm18	56630893	G	A
Gm18	56630923	C	T
Gm18	56630934	A	T
Gm18	56630961	T	C
Gm18	56630985	C	T
Gm18	56631041	T	C
Gm18	56631056	T	C
Gm18	56631116	T	C
Gm18	56631151	A	T
Gm18	56631166	C	T
Gm18	56631200	A	T
Gm18	56631266	C	CAT
Gm18	56631376	A	G
Gm18	56631609	T	G
Gm18	56631641	G	T
Gm18	56631665	A	T
Gm18	56632093	C	T
Gm18	56632157	G	A
Gm18	56632282	C	A
Gm18	56632296	G	C
Gm18	56633543	C	T
Gm18	56634232	G	A
Gm18	56635842	A	T
Gm18	56635932	A	G
Gm18	56636250	C	CA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56636303	T	C
Gm18	56637522	G	C
Gm18	56637889	A	AT
Gm18	56639048	T	C
Gm18	56639540	G	C
Gm18	56640035	A	C
Gm18	56640169	C	T
Gm18	56640187	A	T
Gm18	56640293	C	G
Gm18	56640394	AGGGG	A
Gm18	56640465	T	TA
Gm18	56640696	CAAA	C
Gm18	56640962	A	G
Gm18	56641591	C	A
Gm18	56641739	G	A
Gm18	56641791	T	C
Gm18	56641865	G	T
Gm18	56644131	CAA	C
Gm18	56644396	G	A
Gm18	56645035	A	C
Gm18	56645098	A	ATAC
Gm18	56645223	ATTAAATTTAAATTGATTGTT AAT (SEQ ID NO: 19)	A
Gm18	56645355	T	C
Gm18	56645442	G	T
Gm18	56645465	A	C
Gm18	56645519	T	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56645639	G	A
Gm18	56645668	C	G
Gm18	56645752	T	G
Gm18	56645780	T	A
Gm18	56645880	G	T
Gm18	56645901	TTA	T
Gm18	56645929	G	T
Gm18	56645937	C	CAA
Gm18	56645953	A	T
Gm18	56645958	TAAA	T
Gm18	56645984	T	C
Gm18	56646008	G	A
Gm18	56646045	C	A
Gm18	56646087	A	AT
Gm18	56646107	T	C
Gm18	56646143	G	A
Gm18	56646148	A	AT
Gm18	56646159	C	G
Gm18	56646193	G	A
Gm18	56646220	T	TTA
Gm18	56646239	C	CTATTATA CACCGA (SEQ ID NO: 20)
Gm18	56646257	G	T
Gm18	56646309	T	C
Gm18	56646347	C	T
Gm18	56646350	T	TGA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56646354	AT	A
Gm18	56646357	CA	C
Gm18	56646360	T	G
Gm18	56646361	A	G
Gm18	56646362	T	TG
Gm18	56646364	T	A
Gm18	56646369	T	C
Gm18	56646371	T	C
Gm18	56646372	AAGTG	A
Gm18	56646382	G	GT
Gm18	56646383	AGG	A
Gm18	56646387	G	T
Gm18	56646389	C	T
Gm18	56646390	A	T
Gm18	56646392	A	T
Gm18	56646394	ACAC	A
Gm18	56646398	C	G
Gm18	56646400	A	G
Gm18	56646401	C	A
Gm18	56646415	C	T
Gm18	56646471	A	G
Gm18	56646489	G	A
Gm18	56646513	C	T
Gm18	56646524	C	T
Gm18	56646536	T	C
Gm18	56646578	T	C
Gm18	56646638	G	A
Gm18	56646658	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56646659	G	T
Gm18	56646720	TCTAAA	T
Gm18	56646727	CATTAAGGCCT (SEQ ID NO: 21)	C
Gm18	56646741	AGATG	A
Gm18	56646799	T	A
Gm18	56646810	G	A
Gm18	56646825	ACG	A
Gm18	56646842	G	T
Gm18	56646855	G	T
Gm18	56646871	TAAGTTC	T
Gm18	56646931	C	T
Gm18	56646932	G	A
Gm18	56647024	T	C
Gm18	56647026	A	G
Gm18	56647039	C	T
Gm18	56647040	C	T
Gm18	56647041	T	C
Gm18	56647053	C	T
Gm18	56647057	T	G
Gm18	56647059	G	A
Gm18	56647065	T	G
Gm18	56647068	C	T
Gm18	56647069	T	C
Gm18	56647104	C	T
Gm18	56647112	C	T
Gm18	56647127	A	G
Gm18	56647145	T	TATTA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56647174	A	AG
Gm18	56647186	G	A
Gm18	56647187	A	G
Gm18	56647196	C	A
Gm18	56647200	A	G
Gm18	56647209	C	G
Gm18	56647211	A	G
Gm18	56647214	A	G
Gm18	56647221	G	A
Gm18	56647239	C	T
Gm18	56647248	A	G
Gm18	56647294	G	A
Gm18	56647477	T	C
Gm18	56647482	C	T
Gm18	56647487	T	G
Gm18	56647498	G	A
Gm18	56647512	T	A
Gm18	56647543	TTAAGTATATG (SEQ ID NO: 22)	T
Gm18	56647567	G	A
Gm18	56647605	A	G
Gm18	56647659	G	T
Gm18	56647792	C	T
Gm18	56647797	A	T
Gm18	56647842	G	A
Gm18	56647893	T	C
Gm18	56647916	A	G
Gm18	56647955	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56647957	T	G
Gm18	56648064	T	G
Gm18	56648082	T	A
Gm18	56648111	A	G
Gm18	56648129	T	C
Gm18	56648240	C	A
Gm18	56648316	C	T
Gm18	56648404	T	C
Gm18	56648436	T	C
Gm18	56648528	G	GT
Gm18	56648611	A	AG
Gm18	56648613	T	A
Gm18	56648624	C	T
Gm18	56648718	G	GC
Gm18	56648732	TA	T
Gm18	56648734	A	T
Gm18	56648765	A	ATACATAC
Gm18	56648914	A	T
Gm18	56649033	C	T
Gm18	56649889	G	C
Gm18	56650186	G	GT
Gm18	56650603	C	G
Gm18	56650733	C	A
Gm18	56651552	C	T
Gm18	56651585	T	C
Gm18	56651696	C	A
Gm18	56651711	T	C
Gm18	56651853	A	AT

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56651998	G	GA
Gm18	56652522	T	C
Gm18	56652542	C	CCTAA
Gm18	56652838	AT	A
Gm18	56653055	C	T
Gm18	56653492	G	A
Gm18	56653844	G	T
Gm18	56653850	TGGGGCG	T
Gm18	56653858	G	T
Gm18	56653860	G	T
Gm18	56653874	TGGG	T
Gm18	56654008	A	C
Gm18	56654425	A	ATATGATA G
Gm18	56654689	G	C
Gm18	56656609	ATTTCTTCTTTAATTGTTTT TTTTTTTTTTTGCTCAGCAAA ATTAAATATATTATAGATGAG TACCAGAGGTACTAAAATATA CAGATTTAGAGCCATATTACA AGTAGTTTTGGACAGACAATG ATACAGTAGCTGAAATATCCC AAAACTACTCAAATAAGACT TGGAGCTATACTCTAAACCTA TGATGCTGTCCTAAGAAAAGC ATCTTTGAGATTTGAAGACCA TTGATTGAAATGTAGAGAGAA ATCTTTTTCAAACATCTGAG	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		CCATGTCCATAAAATAAACAC TGCATCATCCATTAATTTATG AGTGTCAA AATTAGCATTGGA GAAAATGATGCTGTTTCCTATG CTTCCAAATCGAGTAAGTGAG AGCGAACCACCACCACTGCCA TCTTTTTTCCCTGCAACCCTA CGGAAGCTCCAAAGATATGCT GACTAAAATGCTGATCCGGCT GATAAGGGAAAACACCCATCA TATTTACCCAAGACTGTGACT CCCACCATAGAGGACTAACTT TACTGCAGTGGAAAAATAAGT GACTTGCAGTCTCTTCCACAG ATCTGCATAAGGGGCACAAGT ACTCTTGTA ACTCAACCCGCT TCTTTCTCAAATTGGCTTTAG TTGGCAGTTTGTCTTGAATCA ACCTCCACGCGAAGATAGCCA CTTTCAATGGGACCCTAAGCT TCCATAATTCCTTGA ACTTTC CATCCTGCTCTTCTCCCACTG TGACATGATGTATTGCCTTAT ATGCACTCTTAGTCGAATAAC AGCCACTAGGCTCTGCTGCCC ACTTCCATTGGTCACTTAATT CCGGCCTGATTGTGAACCCTT CCAGCTGTTGGAGGAAAGCTA	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		CCGCCATATCTATCTCGCTAT CAAACAAAGGTCTCCTCCACT TAAGACTCCATTCCCACCCAC CTTCTTTCGCAGCACCTATTT GATGAATGAAATGATGTTTTT GAGCTGATATGGTATAACAATC TAGGATATTTATCAGCTAAGC AGTTGTCTCCGCCTATCCACC TATCCTCCCCAAACTTAAACT TGTCCCCACATCCGACCCTCC ACAATATCAACCTATTCAGCT GTTGACCTTGATTCATGCTTT GATTTACTATCTTTAGATCCC TCCACCATGATGATTCAGTAC TAGCCCTCGAAGCTCCATCAA GACTCCTCCATCCTCCATACT TTGATTCAAGGACTCTAGCCC AAGTCTCCTCCTGATTCTGAA ACATCTCCCATTTCCATTTTC CTAATAATGATGTGTTGAACG AGATGATATCCTTGACTCCCA ACCCACCTTGTTCCCTTTGGCA AACACACCTTTTCCCACCTGA TCCAAGGGATCTTATTTTGAT CAGCTGCACCACCCATAAGA AGTTCCTTTGTAATCTGGTTA TCTTGCTAATCACCTGCCTAG GGACCCTGAAAAAAGATAGAA	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		AATAAATTGGAATAGATGTTA GCACTGAATTTATGAGAATCA CTCTCCCCC AAAAGACAAGT GTCTCTGTTTCCATCCTGCCA GTTTTCTCTCACACTTATTGA TTAAAGGTTCCCACGTCTGAC ATCTTCTAGGATTAGCACCAA TAGGTATGCCAAGATATGTGA AAGGAAAAGACATCAATCCAC AGTTGAGATAACTGGATGCAC CATAAGTCCACTGCTCAGGCA CCCCAAAGGCCCCACA ACTAC TCTTGCGGAAATTTATCTTTA ACCCCGATGACATCTCAAAG CACGAAGTATTGCATTGATTG TTCTTACATTGCTAATGTTG CCTCTCCCCAAAATAGTGTCA TCTGCATACTGAAGCAGGCTA ATCTCCACCTTCTTTGAGCCC ACTAAAAATCCCTTATAGAAT CCTCCCTCAATTGCTTTGGTC ATTAGACCGCTCAAACCTTCT GCAACAATATTAAACAAAAGG GGTGATAATGGATCCCCCTGT CTAAGACCTTTCTGTGGATAG AACTCCTTCGTAGGGCTTCCA TTTACCAAGACTGAGATGGAA GCTGATTTTAAACATCTCTGT	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		ATCCACTGAATCCACTTCGGT CCAAAACCCAATCTCCTCATC ATGTATATCAAAAAATCCCAA GACACTAAGTCATACGCCTTC TCATAGTCCACCTTGAAAACC ATGCATGGTTTTTGACATCTT TTGGCCTCTTCAACTACCTCA TTTGCAATCACCACACTATGA AGCATGTGCCTACCTTGTATG AATGCTGACTGACGATCACTA ATTATGAGGGGCATAACTCTC TTCAGTCTATTAGGCAACACT TTAGCGAGGATTTTATAGGTA CATCCGATGAGTGAAATAGGT CTATATTCATTCAACCCTTGA GGGTCCGCCACCTTCGGGATT AAAGCTATGAATGATGCGTTC AAGCCCCTCGGGAAGACTCCG TTGACATAGAACTCGTCAAAG AATCTAAGAAAATCTGGTTTG ATGACCTCCCAAATTGCTTG ATGAATTTGAAATTGAACCCG TCTGGGCCCCGGACTTTTATCA CTTCCACAAC TCAAACAGCC CTCCTTATCTCCTCTTCCTGA AAGTGCTCGACTAACATGGCA TTCTGATGGGAATCAATGGTA TTGAAGCTGATCCCATTCAGA	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		GTTGGCCTATCAAAATCTGTT TCCTGGAACCTCTGTGAAAAA AATCTCCTAACTTCCTCTTTG ACTTCAGCCGGCTCCTCCTTC CACATCCCATCAACCATCAGT CCATTTAGGCTATTGGATCGC CTCCTGGAATTTATCAGCAAA TGTTAATACCTCGAATTACAA TCACCCTCCTTAATCCACCTG GATCTTGCCTTCTGACGTAAAC AGCGACTCATGAGTTTGGGCT GCCTCCCAAACCTTCTGCTGC AACTTCTTTTTTGAGTATCCGC TCATTATCATCCAGTGGTCTG TCTGCTGATTCCCTCTTAGT TTGTTCAACTCTTCCTCTATC CTCTTGAAACGGCTTAAAGTG TCTCCAAAGTGATCTTTATTC CACGCCTTTATCTTCTGTTTG AGGGCCTTGAATTTGTTTTTT AAAACATGACCCCCCATCCA CTCTGAGTATTGGATGACCAG GTTTCGGACACCAACTTCTTA AAAGATACATCAGATAACCAA CAATCCAATAGTCTGAAAGGT TTTGGGCCCCAATCAATGGTT TTAGATCTAAGCAATATGGGA CAATGGTCCGAAAAATTCCTG	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		GCTAGCGGTGTTTGGACCGAT CCGGGCCATTTGGAAAGCCAC TCTGGGGAAACGAAGGCTCTA TCCAACTTGCTTTTAGCTGTA CCGTTTCGGTCTAAACCATGTA AATCTCTTTCCCACCCAAGGC GCTTCTTCTAATTCCAACCTCC TCAATCCAGCTATTAAAATCC CGTATGCTTCCATCCACCATC CCCCTCTGCGAAGACCCCACT CTTTCTCCACTTATCCTGATG TTGTTAAAGTCCCCTATAATA CACCAATAACCATTTGGATTC TGGATTTTCAGCTGAGTGAAG TTATCCCATAACACCCTTTTG TCCTGCAGGTTACATGGGGAG TATATATTTACAACATGCACT GCTATGGATTCTTGACCCAC TTTCTGACAGTAGTATGAAG CCCGATCCAGCTGATTGCTC TCCACCTTCAGAGACTTTTGA TTCCAAATACAGAGAATACCA CCAGCTGTATTTACGGCTGGA ACCTCCTCCCAGCAGAAATCA GAATTCCCCCATAGAGCCCTG CACATACTCTTATTGACCGTT TCCTTCTTTGTTTCTTGTAAG CATAAAAGGTCAATATGATAC	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		TCTTTGACCATACGTCTAATT GCTGGCCATTTAACCCCCCTC CCCAACCCCCTTACATTATAA GAGATGATATTCATGAGTCCC TCCCCCTTCCTCCCAAGCTAT CTGCCTCTTTTCCATCTCTAT CTTCCATTTCTGTCAACTGTT GAACATAATCTCTTTGAGTTT TTCCTGTTGTCAGTCCCAACT CCTTGATCATTCTCCATAGAC TTTCTTCGTGCGAGTCAAAAG TTTCCTCCTGGAGCTGATTTT CCTGCCTATTGTGAGTTGATG ATGTTTCTGTTTGTATTTGGG GGTTGGCTGATCCACGTCCCT CGTTATCTCGCTGGAATAGTC CATTTTGAACATGCTCTGGGC TTGAAAGCACTTGGTTCGAAA CAACCTCACCTATTTGGGCTT TAACTCTCTTCATTTTCCCTT TACGGGAATACGTTAACAAAG GGGTTAGGAGAGGTGTTGTTT TTAAAACTGGGCCATTTTGG GAGAAGGACATAATTTAGGTG GGGTGTTAATAGCTTGGCCCA AATCCAGCAGCCCCATCATCT CCTTGTCACGTGACTTGGGTG AGGGGGATTCAAAATTATGTT	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		TGTCTTCCTCGTACACCATGC CTCCAACGTTTCTTGTCTGCC TCCTCATACTATTTTCATACA AATTGCCCCAAACCGTCAGGGC CGACATCCCCCCCCGAGGTTAT CTTCTTCCCCTAAATCGGTCT TCCCTTCTTCTGCGACACTG GTAGCTCACGGTCCCTGGTCT GCACGCATTTATCACCTGTCA GTGCATTATCTGCTTGCTTCA CGACCCATCGGACAATAGCTT TGTCCCTTTGGGTCATGTCCGC TCTTCGTCCCGATCGCCTTTC CGTTTTGGTCCGCACCGGCCG CGACTGTCGCCGGTCCTTCTG ACCGAGAGACCTGCCTGAGTC GCCCTTCTTCCACACCACGCG GCGACCTCTCCTCGTCGCCAC CCCTATCCTCTGTGTCATCCA CCGTCTCCGACTGGGCTGGAG CTCCCCCTTTGGGTTTTAAATG GCACGATTTCACTTCTCCCTC TGCGGTTATCTTCCTCGCCCC CCACCACTCCCTCCGTGCCA TATCCTCCATGGCTTCTGCCT CCAAGGTCCCTCCCTTCGCTA CTGTTTCGTATTTCTCCGCCA CGCTCTCTATTGCGCTTTGT	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		ACCGCCACCCCTGCAGCTCAC TACCAGACCACGCGTCATCGT TACATTCGTCAGAAAGGATTT CCTCTGATGAGCTGAAGTCAC GTTCCCTTTTTGGCTGCCACC TAGCTCCAAACAACCCACTCT CCTCCGCTATGTGCACCGGAA AAATCTCCCCCTGTATATGAA TATTTACCGTGTGCTGAATCA GCGGTTTCCATGGCGTTTTGA CCAGTATCCTCGCGACGTCGA ACCTCCTTTGATCTTCTAAGC TCTCCTCAGCCTCCACCATCT CACCTATTCCCGCGACTATCT GCTTGATACAGTTCATGTCCC ATGCAACCAAAGGTATTCCCC AACATTGGATCCAAGTTAGTC TAAAACCAGGGCGTAGACTTG GATTCCACTTCTCTATGGAGT AGAACAGGTCGCCCCATCCTT CATCTTCTTCATTACCATTT TCTCTGCGTTCTCTTCTGTGA GCCCCAGTAACAAAACCATGT CATCACCGATGTACTTTGGTG ATATGTTTTGTCCGCTATCCC ACCATATGGCTTCTTCTATGC TATCAAAGCTAGCCAAGTTTT TCAGTCTACCCACCCAGGCCT	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		CCTTGAGCCATTGTTTACCCG TCATCGAAATGTCCAGGTTCA CTTCCGATGTGGCATTGGAAC TTATGTATTGGAGTTTGGCTG TTGTTCTCCTCTGTGCTGGTT TCGGGATATTCGTGTTGACCA CCTCCGCATACGACCGTCCTT GGATTCCCATATTCGCTTCTG GTAACCTGCTCAGGTACACTT CTGTTCTTCCCTTTTGTGTTT GTCCTGTCTCTGCTGTCCCCC TCTCGACTGTTTGTGATTCTC CCTTCTTTGGTCTCCCATATT TAGGGATATTAACATTCATCT TCATTCCCTCCGAAGAACCTAT TGTCCAAATGCTGCTGAAGTT GATGCACATCTCTCACCCCCT TATACCTCACAAAACCATATC TCCTTCCTTGGTTGTTTCTTT TCCTGGGTATGAAAACCTCTC TCACGTCTCCTTGCTGTTTAA AATGGAACCACAAATCCTTTG CTGTTGCGTCTTCCGGAACC TCGTAAAGTAAAAGGATATAA CGTCCTTATGGTCTCTCCAGT TCGCTCGCGTGTAGGATTGAC GCTGTCCTGTGGCTAACCCTT CGATGGTCAGATGGGAACCCT	

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
		CTGACTCTCGAAATCTGACTC TAGCTCTCCTCTCTTTCCTAG TCCTAACTCTCTCCCACCCAC TGTTTCTCTCTCTACTCTCTC TCTCTCTCCATGATTGGATTG TTCT (SEQ ID NO: 23)	
Gm18	56662924	C	T
Gm18	56663558	A	T
Gm18	56663630	T	A
Gm18	56663642	T	C
Gm18	56663911	T	G
Gm18	56664225	A	G
Gm18	56664232	C	T
Gm18	56664489	T	C
Gm18	56664670	C	G
Gm18	56665263	G	T
Gm18	56665270	TTG	T
Gm18	56665533	T	TA
Gm18	56665602	G	A
Gm18	56666140	CA	C
Gm18	56666204	T	A
Gm18	56666205	T	A
Gm18	56666616	G	C
Gm18	56666694	G	A
Gm18	56666897	G	A
Gm18	56666919	AT	A
Gm18	56667054	GT	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56667061	C	G
Gm18	56667091	C	T
Gm18	56667228	A	G
Gm18	56667372	A	T
Gm18	56667399	C	G
Gm18	56667482	CAG	C
Gm18	56667528	T	C
Gm18	56667541	A	ACAGCTCG AGTTAATA T (SEQ ID NO: 24)
Gm18	56667755	G	T
Gm18	56667840	T	A
Gm18	56667860	T	A
Gm18	56667861	A	AGG
Gm18	56668194	A	AGAAATGG AGAAAGTG (SEQ ID NO: 25)
Gm18	56668230	CT	C
Gm18	56668264	T	TA
Gm18	56668288	TA	T
Gm18	56668290	A	T
Gm18	56668400	A	AAAT
Gm18	56668492	A	T
Gm18	56668671	A	C
Gm18	56668721	A	ATAT

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56668741	C	T
Gm18	56668770	A	T
Gm18	56668923	CCG	C
Gm18	56668926	TCACC	T
Gm18	56669216	T	G
Gm18	56676220	C	A
Gm18	56676566	T	G
Gm18	56676638	C	T
Gm18	56680021	C	G
Gm18	56680179	C	T
Gm18	56682143	A	C
Gm18	56682234	C	T
Gm18	56682368	C	CCTTCTTC AGTT (SEQ ID NO: 26)
Gm18	56682411	A	G
Gm18	56682511	C	T
Gm18	56682528	C	A
Gm18	56682538	G	C
Gm18	56682658	T	A
Gm18	56682675	C	A
Gm18	56682700	C	T
Gm18	56682703	C	T
Gm18	56682726	T	TTTATCA TGAAA (SEQ ID NO: 27)

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56682755	T	C
Gm18	56682839	T	G
Gm18	56682841	A	C
Gm18	56682852	G	A
Gm18	56682858	T	A
Gm18	56682870	T	A
Gm18	56682968	C	A
Gm18	56683053	T	C
Gm18	56683054	G	A
Gm18	56683110	TTC	T
Gm18	56683296	A	C
Gm18	56683332	A	ATATATAT ATATATAT (SEQ ID NO: 28)
Gm18	56683334	C	A
Gm18	56683428	A	ATCC
Gm18	56683643	A	G
Gm18	56683761	A	G
Gm18	56683817	T	C
Gm18	56683881	ATGTGTGTGTG (SEQ ID NO: 82)	A
Gm18	56683894	TGTA	T
Gm18	56684003	A	G
Gm18	56684101	A	T
Gm18	56684693	T	G
Gm18	56684788	G	A
Gm18	56684790	C	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56684820	T	A
Gm18	56684873	T	G
Gm18	56684957	A	G
Gm18	56684968	C	T
Gm18	56685074	T	G
Gm18	56685278	TAA	T
Gm18	56685521	C	G
Gm18	56685535	A	AAGGGGGG AATGG (SEQ ID NO: 29)
Gm18	56685566	C	A
Gm18	56685685	T	C
Gm18	56685835	T	C
Gm18	56685874	C	A
Gm18	56685895	C	A
Gm18	56686085	C	T
Gm18	56686209	T	C
Gm18	56686219	G	A
Gm18	56686290	TA	T
Gm18	56686387	T	C
Gm18	56686538	C	T
Gm18	56686683	CA	C
Gm18	56686768	C	A
Gm18	56686773	G	A
Gm18	56686786	A	ATAAAAT
Gm18	56686895	C	T
Gm18	56686900	T	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56686921	G	A
Gm18	56686940	AG	A
Gm18	56687096	T	C
Gm18	56687120	C	T
Gm18	56687358	A	C
Gm18	56687362	T	C
Gm18	56687459	A	C
Gm18	56687461	G	T
Gm18	56687462	A	C
Gm18	56687515	GAAAGGTGGA (SEQ ID NO: 83)	G
Gm18	56687733	A	C
Gm18	56687735	C	T
Gm18	56687743	CCTATGCG	C
Gm18	56687842	T	C
Gm18	56688141	T	C
Gm18	56688170	A	G
Gm18	56688353	C	T
Gm18	56688392	AT	A
Gm18	56688800	T	A
Gm18	56689257	T	G
Gm18	56689445	T	G
Gm18	56689598	G	GT
Gm18	56689744	T	G
Gm18	56689862	CA	C
Gm18	56689878	TAA	T
Gm18	56689972	GAA	G
Gm18	56690034	A	AT

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56690170	A	G
Gm18	56690474	CA	C
Gm18	56690559	A	ATT
Gm18	56690747	A	T
Gm18	56690889	T	C
Gm18	56690976	T	G
Gm18	56691178	C	A
Gm18	56691193	A	C
Gm18	56691290	G	T
Gm18	56691305	T	A
Gm18	56691433	G	T
Gm18	56691440	C	T
Gm18	56692305	T	TAATATAA TTTATTTA TTAA (SEQ ID NO: 84)
Gm18	56692901	G	GA
Gm18	56692988	G	C
Gm18	56693008	T	C
Gm18	56693110	T	A
Gm18	56693247	AAATATATATATATATATATA TATATATATATATATATATAT (SEQ ID NO: 30)	A
Gm18	56693306	TA	T
Gm18	56693344	TA	T
Gm18	56693516	T	C
Gm18	56695064	C	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56697201	T	TCA
Gm18	56697257	G	C
Gm18	56697691	G	T
Gm18	56697723	G	A
Gm18	56698131	G	T
Gm18	56698364	G	A
Gm18	56698379	A	G
Gm18	56698390	T	A
Gm18	56698424	G	A
Gm18	56698426	G	A
Gm18	56698434	G	A
Gm18	56698465	C	A
Gm18	56698489	C	A
Gm18	56698566	GT	G
Gm18	56698734	A	G
Gm18	56698802	T	C
Gm18	56698808	G	GA
Gm18	56698810	G	A
Gm18	56698818	T	G
Gm18	56698825	C	T
Gm18	56699167	CA	C
Gm18	56699334	G	T
Gm18	56699339	T	C
Gm18	56699344	A	C
Gm18	56699389	AT	A
Gm18	56699399	C	A
Gm18	56699401	A	ATT
Gm18	56699504	G	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56699527	C	CAAAAATA A
Gm18	56699584	G	A
Gm18	56699588	TA	T
Gm18	56699607	TTCA	T
Gm18	56699622	C	CT
Gm18	56699653	T	A
Gm18	56699721	G	A
Gm18	56699759	A	ATAAATAA ATAATAAT AGT (SEQ ID NO: 31)
Gm18	56699770	C	G
Gm18	56699840	A	G
Gm18	56699965	T	A
Gm18	56700056	G	T
Gm18	56700068	G	A
Gm18	56700082	G	A
Gm18	56700086	T	A
Gm18	56700348	T	C
Gm18	56700360	C	G
Gm18	56700362	T	C
Gm18	56700612	A	C
Gm18	56700818	T	TG
Gm18	56700967	G	C
Gm18	56701142	A	G
Gm18	56701149	T	C

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56701244	A	C
Gm18	56701322	A	G
Gm18	56701339	A	G
Gm18	56701342	T	C
Gm18	56701385	C	T
Gm18	56701428	C	T
Gm18	56701439	G	T
Gm18	56701455	T	C
Gm18	56701492	G	A
Gm18	56701673	C	T
Gm18	56701736	T	C
Gm18	56701824	A	AG
Gm18	56701832	A	ATT
Gm18	56702001	C	T
Gm18	56702064	T	C
Gm18	56702101	T	G
Gm18	56702360	C	T
Gm18	56702698	AT	A
Gm18	56702760	A	C
Gm18	56702767	G	GC
Gm18	56702770	GATTT	G
Gm18	56702777	T	G
Gm18	56702779	A	G
Gm18	56702780	C	T
Gm18	56702788	G	A
Gm18	56702824	A	T
Gm18	56702882	A	G
Gm18	56702888	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56702894	T	C
Gm18	56702927	A	G
Gm18	56702944	A	G
Gm18	56703048	T	A
Gm18	56703067	A	G
Gm18	56703127	T	C
Gm18	56703131	G	A
Gm18	56703135	T	A
Gm18	56703188	C	T
Gm18	56703195	A	G
Gm18	56703209	T	C
Gm18	56703284	A	AT
Gm18	56703290	T	A
Gm18	56703334	T	C
Gm18	56703375	A	T
Gm18	56703399	T	A
Gm18	56703404	T	G
Gm18	56703506	A	G
Gm18	56703564	G	A
Gm18	56703591	C	CTTCTTAC TTGTAACT AAGTCTTT GA (SEQ ID NO: 32)
Gm18	56703608	G	A
Gm18	56703653	T	TA
Gm18	56703658	G	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56703672	GAA	G
Gm18	56703682	TA	T
Gm18	56703785	T	A
Gm18	56703798	T	G
Gm18	56703824	T	G
Gm18	56703833	A	AAATT
Gm18	56703842	C	A
Gm18	56703861	AG	A
Gm18	56703866	G	A
Gm18	56704030	CT	C
Gm18	56704201	A	G
Gm18	56704273	T	A
Gm18	56704278	C	A
Gm18	56704284	G	C
Gm18	56704352	T	TG
Gm18	56704382	G	C
Gm18	56704398	ATACT	A
Gm18	56704417	T	C
Gm18	56704420	TA	T
Gm18	56704443	T	A
Gm18	56704456	A	G
Gm18	56704501	T	C
Gm18	56704508	T	A
Gm18	56704531	C	T
Gm18	56704539	T	C
Gm18	56704548	A	C
Gm18	56704611	C	G
Gm18	56704650	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56704669	A	T
Gm18	56704693	T	C
Gm18	56704697	G	A
Gm18	56704748	A	T
Gm18	56704760	A	ACCATG
Gm18	56704766	A	T
Gm18	56704771	G	A
Gm18	56704777	T	C
Gm18	56704845	C	T
Gm18	56704940	TAA	T
Gm18	56704963	G	A
Gm18	56705061	G	A
Gm18	56705077	T	C
Gm18	56705088	A	C
Gm18	56705094	G	C
Gm18	56705113	C	T
Gm18	56705146	A	T
Gm18	56705200	A	G
Gm18	56705219	A	G
Gm18	56705251	T	TATATTTT TTGTGTTT GGTCTTTG ATAAATTT TTCTTTTCG AATTTGGA TTCTAATG TTTTAAAT TTTTTATC

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
			TTAAAGTC TCTATCAT TAATATGT AATCCTAC ATGACTCT TAATTAGA CACGTAGA AGTGTGAT GTGTCATG TCAAATAT AACCTTTG TGATTTCT ATTTATAC ATATTGGA TTATTAAA AAAAAATA TTTTCTTT GTTAAAAA ATAATAAA ATCCAAC TATATAAA CAATAATC CGACATAA TACATCAC ATTTTAC GTGTGTCC AGTTCATA C (SEQ

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
			ID NO: 33)
Gm18	56705253	G	A
Gm18	56705254	G	A
Gm18	56711483	TA	T
Gm18	56711576	G	GA
Gm18	56711711	A	G
Gm18	56711752	T	TA
Gm18	56712550	T	C
Gm18	56712561	T	C
Gm18	56712563	A	G
Gm18	56712847	T	A
Gm18	56712894	CTATATA	C
Gm18	56712934	A	G
Gm18	56712958	C	CCA
Gm18	56712959	T	TGC
Gm18	56712988	C	T
Gm18	56713017	C	A
Gm18	56713045	G	T
Gm18	56713064	C	T
Gm18	56713220	T	C
Gm18	56713259	G	A
Gm18	56713332	T	C
Gm18	56713478	G	C
Gm18	56713624	C	A
Gm18	56713748	A	G
Gm18	56713764	C	A
Gm18	56713852	A	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56713995	C	A
Gm18	56714029	G	A
Gm18	56714182	G	C
Gm18	56714236	A	G
Gm18	56714245	G	T
Gm18	56714295	T	G
Gm18	56714368	G	T
Gm18	56714611	T	C
Gm18	56714635	T	C
Gm18	56714850	A	C
Gm18	56714865	A	G
Gm18	56715062	T	A
Gm18	56715180	C	A
Gm18	56715529	C	CTT
Gm18	56715568	C	T
Gm18	56715595	C	T
Gm18	56715695	A	G
Gm18	56715700	T	TGGA
Gm18	56715702	T	G
Gm18	56715704	AC	A
Gm18	56715707	T	TGGA
Gm18	56715709	T	TGG
Gm18	56715710	A	G
Gm18	56715712	C	T
Gm18	56715714	G	T
Gm18	56715716	G	GAGA
Gm18	56715717	G	GAA
Gm18	56715720	T	TTTTGACA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
			AAAAC (SEQ ID NO: 34)
Gm18	56715724	C	CTTAAA
Gm18	56715725	C	G
Gm18	56715727	T	TTTTAAAA G
Gm18	56715732	A	T
Gm18	56715747	T	C
Gm18	56715775	C	CAACACTA ACAAATTT TA (SEQ ID NO: 35)
Gm18	56715781	A	G
Gm18	56715841	T	C
Gm18	56715854	T	A
Gm18	56715872	G	A
Gm18	56715929	C	T
Gm18	56715931	A	G
Gm18	56715950	G	C
Gm18	56715958	G	A
Gm18	56715971	C	T
Gm18	56715981	T	C
Gm18	56715982	A	G
Gm18	56715993	G	A
Gm18	56716009	A	T
Gm18	56716021	ATT	A

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56716031	A	T
Gm18	56716262	C	CG
Gm18	56716273	G	A
Gm18	56716285	G	A
Gm18	56716310	G	A
Gm18	56716354	TTAA	T
Gm18	56716388	A	T
Gm18	56716452	T	C
Gm18	56716547	TA	T
Gm18	56716576	T	C
Gm18	56716720	T	TA
Gm18	56716984	G	A
Gm18	56717056	T	TGC
Gm18	56717086	A	G
Gm18	56717118	T	C
Gm18	56717139	T	C
Gm18	56717216	G	A
Gm18	56717267	T	A
Gm18	56717293	T	C
Gm18	56717338	C	T
Gm18	56717407	G	C
Gm18	56717408	G	C
Gm18	56717438	C	T
Gm18	56717695	G	A
Gm18	56718157	A	T
Gm18	56718193	G	T
Gm18	56718199	A	AT
Gm18	56718208	C	T

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56718334	A	G
Gm18	56718585	A	AT
Gm18	56718665	A	T
Gm18	56718667	G	A
Gm18	56718831	A	G
Gm18	56718997	G	A
Gm18	56719142	C	CAAGTATG CAGGCTTT TGTACCCA TAGTACCA CTGGTACT ATTTCAAT CTATAATA TATATATT TTTGCTGA GCAAAAAA AAAA (SEQ ID NO: 36)
Gm18	56719289	T	C
Gm18	56719463	C	A
Gm18	56719488	C	T
Gm18	56719491	T	C
Gm18	56719540	T	C
Gm18	56719572	C	T
Gm18	56719814	A	G
Gm18	56719828	G	A
Gm18	56719975	A	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56720147	AAACCCATGATAACTGGTTTT AATTGTGGGCTGTCTCCATAC TCTACACAAGCT (SEQ ID NO: 37)	A
Gm18	56720232	G	T
Gm18	56720354	G	T
Gm18	56720393	G	A
Gm18	56720436	A	T
Gm18	56720464	T	C
Gm18	56720492	T	G
Gm18	56720496	C	CTGTCCAG CCAAGATC TTGACTGT TGTA GTTG AACTTAGT AGCTGAAG AGGAAAGA GAATGTGA TGGGTGGT TGAGATTT GGGAAGGA GAGAAACC TTGCTGGG AGAGGCTG CAGAGGAT CCAGATTC CTGCTATA TATTGTCA

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
			TGATAACT GTCAAGTG TGAAATTG AGAGCTTG CTAATCTT GTAGAATA TATAAACC ATTTTTGA CTTTTTTT TTTTAAAA AAATGATT TGATCATA TGGCATTG ATGTTTGT TTGAGTTG TAGCAGTT TCTTTCTG TTCCCATA AATTGTTA TCATTCTT TTTTGAGA GTTGGATT TACTGGTT TGAG (SEQ ID NO: 38)
Gm18	56720499	A	G
Gm18	56720501	T	G

Tabla 1: Posiciones de SNP			
Cromosoma	Posición	Alelo de referencia	Alelo Rpp6907
Gm18	56720502	T	C
Gm18	56720534	T	C
Gm18	56720672	G	C
Gm18	56720684	T	C
Gm18	56720710	A	C
Gm18	56720828	C	T
Gm18	56720902	C	T
Gm18	56720938	T	C
Gm18	56720942	C	T
Gm18	56721082	C	T
Gm18	56721119	T	C
Gm18	56721214	G	A
Gm18	56721542	G	A
Gm18	56721704	C	T
Gm18	56721865	G	A
Gm18	56722043	G	A
Gm18	56722203	G	T
Gm18	56722409	G	T
Gm18	56722545	A	C
Gm18	56722832	C	T
Gm18	56722859	A	C
Gm18	56723297	G	A

Cebadores de oligonucleótidos (en la presente, 'cebadores') se pueden desarrollar y utilizar para identificar plantas que portan uno cualquiera de los intervalos cromosómicos representados en SEQ ID NO: 11, 12, o 13 que se encuentran altamente asociados con la resistencia a la ASR. Los cebadores también pueden desarrollarse y usarse

para identificar plantas que portan la SEQ ID NO: 2. Específicamente, un experto en la materia puede desarrollar cebadores para detectar cualquier polimorfismo mononucleotídico (en este documento "SNP") como se define en la Tablas 1 respecto a la identificación o producción de líneas de soja que tienen uno cualquiera o una parte de los intervalos cromosómicos representados en las SEQ ID NO: 11, 12, o 13 que están asociados con resistencia a la ASR. Puede emplearse un ensayo TAQMAN® (por ejemplo, generalmente un ensayo de discriminación alélica de dos etapas o similar), un ensayo KASP™ (generalmente un ensayo de discriminación alélica de una etapa definido a continuación o similar) o ambos para identificar los SNP que están asociados con resistencia a la ASR aumentada como se divulga en este documento (por ejemplo, alelos favorables como se representa en la Tabla 1 anteriormente). En un ensayo de dos etapas ilustrativo se emplean un cebador directo, un cebador inverso y dos sondas de ensayo (u oligos de hibridación). Los cebadores directo e inverso se emplean para amplificar locus genéticos que comprenden SNP que están asociados con locus de resistencia a la ASR (por ejemplo, cualquiera de los alelos favorables, tal como se muestra en la Tabla 1). Los nucleótidos particulares que se presentan en las posiciones de SNP se someten entonces a ensayo utilizando los cebadores de ensayo (que en algunas realizaciones están marcados diferencialmente con, por ejemplo, fluoróforos para permitir distinguir entre las dos sondas de ensayo en una sola reacción), que en cada uno de los pares difieren entre sí con respecto a los nucleótidos que están presentes en la posición de SNP (aunque se señala que en cualquier par dado, las sondas pueden diferir en sus extremos 5' o 3' sin impactar en sus capacidades de diferenciarse entre nucleótidos presentes

en las posiciones de SNP correspondientes). En algunas realizaciones, los cebadores de ensayo y las condiciones de reacción se diseñan de modo que un cebador de ensayo hibridará únicamente con el complemento inverso de una secuencia perfectamente coincidente del 100 %, permitiendo de ese modo la identificación del alelo o alelos que están presentes basándose en la detección de las hibridaciones.

En una realización de la invención, el siguiente ensayo puede emplearse para identificar los SNP que se asocian con la resistencia aumentada a la ASR tal como se da a conocer en el presente documento:

Tabla 2: Ensayos para la detección de los SNP a través del intervalo					
Ensayo	Tipo de variante (SNP o InDel)	Posición en la Secuencia o BAC	Posición física en el cromosoma	Alelo favorable	Alelo desfavorable
1	SNP	282	56,723,297	T	C
2	SNP	173,304	56,607,175	G	A
3	Inserción:	42631-42641	56,682,368	aactgaagaag-	
4	SNP	42,775	56,682,234	A	G
5	SNP	42,866	56,682,143	G	T
6	SNP	44,833	56,680,179	A	G
7	SNP	48,513	56,676,638	A	G
8	SNP	48,585	56,676,566	C	A
9	SNP	87,329	56,676,220	T	G
10	SNP	106,768	56,680,021	C	G

Tabla 2: Ensayos para la detección de los SNP a través del intervalo (continuación)	
	Ensayo de discriminación alélica de TAQ

Ensayo	secuencia de sonda 1	secuencia de sonda 2	secuencia de cebador 1	secuencia de cebador 2
1	AATAGCCCCACCCTT AATT (SEQ ID NO: 39)	CAATAGCCCCATCCT TAA (SEQ ID NO: 49)	GGATGGAATGGGAAA GTCTTGATAA (SEQ ID NO: 59)	GCAAGCAAACCTCGA ACCTA (SEQ ID NO: 69)
2	TCATCATGCTAAAAC TGAGTA (SEQ ID NO: 40)	CTCATCATGCTAAAA TTGAGTA (SEQ ID NO: 50)	AAAACCTGCGAAGCT CCTTTACTA (SEQ ID NO: 60)	CCCACCCGTTTATCA TTTGCACTT (SEQ ID NO: 70)
3	TTGAGGAAGAACTGA AGAAGGTAGC (SEQ ID NO: 41)	ACGGATTGAGGAAGG TAGCAGAA (SEQ ID NO: 51)	GTTTGTGTCCACGC CTGAA (SEQ ID NO: 61)	TGTCATCAAGCATCA CCAGGTA (SEQ ID NO: 71)
4	ATTGAAAATTGTTCC TGTAC (SEQ ID NO: 42)	CCTATTGAAAATTGT TCTTGAC (SEQ ID NO: 52)	TTACGTCAATGCTCT TAAGGTCTAA (SEQ ID NO: 62)	ACTAATAACAAGGAG GTGGCATAC (SEQ ID NO: 72)
5	TGGCTTAGGAAGAAA GCTC (SEQ ID NO: 43)	TGGCTTAGGACGAAA GCT (SEQ ID NO: 53)	GCCTTGACCTCATCC ATTGATAGAG (SEQ ID NO: 63)	ACTAATAACAAGGAG GTGGCATAC (SEQ ID NO: 73)
6	ATCTCTGATCTGGCT CAC (SEQ ID NO: 44)	TCTCTGATTTGGCTC AC (SEQ ID NO: 54)	GGTGTCAACCACATC TTCAGCTTT (SEQ ID NO: 64)	TTCTGGCGTGAGGA CAAA (SEQ ID NO: 74)
7	CAACAAATACATATA CCAAAGTC (SEQ ID NO: 45)	AGCAACAAATACATA TATCAAAGTC (SEQ ID NO: 55)	ACTTGCTCTACTTCA TAAATGGGTA (SEQ ID NO: 65)	GTCAGTGTCTAGTT TTCCTTGA (SEQ ID NO: 75)
8	ACAAAATCATCACCA AGTTGA (SEQ ID NO: 46)	CAAAAATCATCACCCA GTTG (SEQ ID NO: 56)	ACCCATTTATGAAGT AGAGCAAGTA (SEQ ID NO: 66)	AGCGAACGGGAACAG TCCAT (SEQ ID NO: 76)
9	ATCTGCATGCTAATC CTG (SEQ ID NO: 47)	CTGCATGCTCATCCT G (SEQ ID NO: 57)	TCCGAGAAAGGAATA GTTCTGTG (SEQ ID NO: 67)	GGACGACACTTTTAC ATCAAACAC (SEQ ID NO: 77)
10	TTCGATCTGACTCCT GAT (SEQ ID NO: 48)	CGATCTGAGTCCTGA TC (SEQ ID NO: 50)	CACCTTCTCCAATGC CATATCTGT (SEQ ID NO: 68)	GCTTCACCAAGTCAA CTCTGA (SEQ ID NO: 78)

Mapeo Genético

Locus genéticos con fenotipos particulares tales como resistencia a la roya se pueden cartografiar en el genoma de un organismo. Al identificar un marcador o grupo de marcadores que co-segregan con un rasgo de interés, el obtentor es capaz de seleccionar rápidamente un fenotipo deseado, seleccionando el marcador apropiado (un proceso denominado selección asistida por marcador, o "MAS"). Los reproductores también pueden usar dichos marcadores para diseñar fenotipos *in silico* y para llevar a la práctica la selección del genoma completo.

En determinadas realizaciones, la presente invención proporciona marcadores asociados con la resistencia potenciada a la roya (por ejemplo, roya de la soja asiática). La detección de estos marcadores y/u otros marcadores ligados puede usarse para identificar, seleccionar y/o producir plantas resistentes a la roya, más específicamente resistentes a la roya de la soja asiática (en el presente documento, "ASR") y/o para eliminar plantas que no son resistentes a enfermedades de los programas de cultivo selectivo o plantación.

Locus genéticos de asociados con resistencia potenciada a enfermedades

En este documento se identifican marcadores asociados con resistencia potenciada a enfermedades (véase la Tabla 1, que indican marcadores favorables asociados con resistencia potenciada a ASR). Un marcador de la presente invención puede comprender un único alelo o una combinación de alelos en uno o más locus genéticos (por ejemplo, cualquier combinación de marcadores favorables de la Tabla 1. Por ejemplo, el marcador puede comprender uno o más alelos de marcadores ubicados dentro un primer intervalo cromosómico (por ejemplo, SEQ ID

NO: 11) y uno o más alelos de marcadores ubicados dentro un segundo intervalo cromosómico (por ejemplo, SEQ ID NO: 12, o SEQ ID NO: 13).

Selección asistida por marcadores

Se pueden utilizar los marcadores en una gran variedad de aplicaciones del cultivo selectivo de las plantas. Véase, por ejemplo, Staub *et al.*, *Hortscience* 31: 729 (1996); Tanksley, *Plant Molecular Biology Reporter* 1: 3 (1983). Una de las principales áreas de interés es el aumento de la eficacia del retrocruzamiento e introgresión de genes utilizando la selección asistida por marcador (MAS). En general, MAS se aprovecha de marcadores genéticos que se han identificado por tener una probabilidad significativa de cosegregarse con un rasgo deseado. Se supone que tales marcadores están en el gen o genes que dan lugar al fenotipo deseado, o cerca de estos, y su presencia indica que la planta poseerá el rasgo deseado. Se espera que las plantas que poseen el marcador transfieran el fenotipo deseado a su descendencia.

Un marcador del que se haya demostrado que está ligado con un locus que afecta a un rasgo fenotípico deseado proporciona una herramienta útil para la selección del rasgo en una población de plantas. Esto es especialmente cierto cuando es difícil someter al fenotipo a un ensayo o este ocurre en una etapa tardía del desarrollo vegetal. Ya que los ensayos con marcadores del ADN son menos laboriosos y requieren menos espacio físico que el fenotipado en el campo, se pueden someter a ensayo poblaciones mucho más grandes lo que aumenta las probabilidades de encontrar un recombinante en el que el segmento diana procedente de la línea donadora se haya transferido a la línea receptora. Cuanto más cercano sea el ligamiento, más útil será el marcador, ya que es menos

probable que la recombinación ocurra entre el marcador y el gen que provoca o confiere el rasgo. La presencia de marcadores flanqueantes disminuye las probabilidades de que ocurra una selección con un falso positivo. La situación ideal es tener un marcador dentro del propio gen causante, de modo que la recombinación no puede producirse entre el marcador y el gen. Un marcador de este tipo se denomina un "marcador perfecto".

Cuando un gen se somete a introgresión por MAS, no solamente se introduce el gen sino también las regiones flanqueantes. Gepts, *Crop Sci* 42:1780 (2002). Esto se denomina "reducción del ligamiento". En el caso en el que la planta donadora sea sumamente ajena a la planta receptora, estas regiones flanqueantes portan genes adicionales que pueden codificar rasgos agronómicamente indeseables. Esta "reducción del ligamiento" puede resultar también en un rendimiento reducido o en otras características agronómicas negativas, incluso después de múltiples ciclos de retrocruzamiento en la línea de soja de élite. Esto también se denomina a veces "reducción del rendimiento". El tamaño de la región flanqueante puede disminuirse por retrocruzamiento adicional, aunque esto no siempre es satisfactorio, ya que los obtentores no tienen control sobre el tamaño de la región o los puntos de interrupción de recombinación. Young *et al.*, *Genetics* 120:579 (1998). En el cultivo selectivo clásico, habitualmente únicamente la probabilidad hace que se seleccionen las recombinaciones que contribuyen a una reducción en el tamaño del segmento donador. Tanksley *et al.*, *Biotechnology* 7: 257 (1989). Incluso después de 20 retrocruzamientos, se podría encontrar un trozo dimensionable del cromosoma donador aún ligado al gen que se está seleccionando. Sin embargo, con los

marcadores es posible seleccionar esos inusuales individuos que han experimentado recombinación cerca del gen de interés. En 150 plantas sometidas a retrocruzamiento, existe una probabilidad de un 95 % de que al menos una planta haya experimentado recombinación por meiosis en 1 cM del gen, en función de una distancia de mapa de meiosis única. Los marcadores permiten la identificación inequívoca de esos individuos. Con un retrocruzamiento adicional de 300 plantas, existiría una probabilidad de un 95 % de recombinación por meiosis en una distancia de mapa de meiosis única de 1 cM del otro lado del gen, lo que genera un segmento alrededor del gen diana de menos de 2 cM en función de una distancia de mapa de meiosis única. Esto se puede lograr en dos generaciones con marcadores, mientras que requeriría un promedio de 100 generaciones sin marcadores. Véase Tanksley et al., citado anteriormente. Cuando se conoce la ubicación exacta de un gen, se pueden utilizar los marcadores flanqueantes que rodean al gen para seleccionar recombinaciones en diferentes tamaños poblacionales. Por ejemplo, en tamaños poblacionales menores, se podrían esperar recombinaciones más alejadas del gen, de manera que se requerirían marcadores flanqueantes más distales para detectar la recombinación.

La disponibilidad de mapas de enlace integrados del genoma de soja que contiene densidades crecientes de marcadores de soja públicos ha facilitado el mapeo genético de soja y MAS.

De todos los tipos de marcadores moleculares, los SNP son los más abundantes y poseen el potencial de proporcionar la resolución más elevada del mapa genético. Bhatramakki et al., *Plant Molec. Biol.* 48:539 (2002). Los SNP se pueden someter a ensayo de la manera denominada de "resolución

ultraelevada" debido a que no requieren cantidades grandes de ácido nucleico y la automatización del ensayo es sencilla. Los SNP también presentan el beneficio de ser sistemas relativamente baratos. En conjunto, estos tres factores convierten a los SNP en sumamente atractivos para su uso en MAS. Se disponen de varios métodos para el genotipado de SNP que incluyen, sin carácter limitante, la hibridación, extensión del cebador, ligamiento de oligonucleótidos, escisión con nucleasas, minisequenciación y esferas codificadas. Este tipo de métodos han sido objetos de revisiones en varias publicaciones: Gut, *Hum. Mutat.* 17:475 (2001); Shi, *Clin. Chem.* 47:164 (2001); Kwok, *Pharmacogenomics* 1:95 (2000); Bhatramakki y Rafalski, *Discovery and application of single nucleotide polymorphism markers in plants*, en *PLANT GENOTYPING: THE DNA FINGERPRINTING OF PLANTS*, CABI Publishing, Wallingford (2001). Una amplia gama de tecnologías comercializadas utilizan estos y otros métodos para estudiar SNP, incluidos Masscode™ (Qiagen, Germantown, MD), Invader® (Hologic, Madison, WI), SnapShot® (Applied Biosystems, Foster City, CA), Taqman® (Applied Biosystems, Foster City, CA) y Beadarrays™ (Illumina, San Diego, CA).

Se puede utilizar un cierto número de alelos de SNP juntos dentro de una secuencia o a través de secuencias enlazadas para describir un haplotipo para cualquier genotipo particular. Ching et al., *BMC Genet.* 3:19 (2002); Gupta et al., (2001), Rafalski, *Plant Sci.* 162:329 (2002b). Los haplotipos pueden ser más informativos que los SNP únicos y pueden ser más descriptivos de un genotipo particular. Por ejemplo, un SNP único puede ser el alelo "T" para una línea o variedad resistente a enfermedades específica, pero el alelo "T" también puede aparecer en la población de reproducción de soja que esté siendo utilizada para parentales recurrentes.

En este caso, también puede ser más informativa una combinación de alelos en SNP ligados. Una vez que se ha asignado un haplotipo único a una región cromosómica donadora, se puede utilizar ese haplotipo en esa población o en cualquiera de sus subconjuntos para determinar si un individuo tiene un gen particular. El uso de plataformas de detección de marcadores de resolución elevada automáticas conocidas por los expertos en la materia hace este proceso sumamente eficiente y eficaz.

Los marcadores de la presente invención se pueden utilizar en protocolos de selección asistidos por marcadores para identificar y/o seleccionar descendencia con resistencia a la roya de la soja asiática potenciada. Tales métodos pueden comprender, consistir esencialmente en o consistir en cruzar una primera planta o germoplasma de soja con una segunda planta o germoplasma de soja, donde la primera planta o germoplasma de soja comprende un intervalo cromosómico derivado de SX6907 donde dicho intervalo cromosómico comprende las SEQ ID NO 11, 12, o 13, o una parte de las mismas que codifica la resistencia a la ASR, o donde el intervalo cromosómico comprende la SEQ ID NO: 2 o un ácido nucleico que codifica para la SEQ ID NO: 1, y seleccionar una planta de la descendencia que posee el marcador. Una cualquiera de la primera y la segunda plantas de soja, o ambas, puede ser de una variedad de soja que se produce de forma no natural. En algunas realizaciones, la segunda planta o germoplasma de soja es de una variedad de élite de soja. En algunas realizaciones, el genoma de la segunda planta o germoplasma de soja es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, o 100% idéntico al de una variedad de élite de soja. En otra realización, la primera planta de soja comprende un intervalo

cromosómico derivado de SX6907 donde dicho intervalo cromosómico comprende SEQ ID NO 11, 12, o 13 y donde el intervalo cromosómico comprende además al menos un alelo tal como se representa en la Tabla 1. En otra realización, la primera soja comprende un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR, o la secuencia de ácido nucleico que codifica para la proteína de la SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

Produciendo plantas resistentes a enfermedades

Los métodos para identificar y/o seleccionar una planta o germoplasma de soja resistente a enfermedades puede comprender, consistir esencialmente en o consistir en detectar la presencia de un marcador asociado con tolerancia potenciada a ASR. El marcador se puede detectar en cualquier muestra tomada de la planta o del germoplasma, incluyendo, pero no limitado a la planta o el germoplasma entero, una porción de dicha planta o germoplasma (por ejemplo, un chip de semilla, un disco perforador de hojas o una célula de dicha planta o germoplasma) o una secuencia de nucleótidos de dicha planta o germoplasma. Una muestra de este tipo puede tomarse de la planta o el germoplasma utilizando cualquier método actual o futuro conocido en la técnica, incluyendo, pero no limitado a métodos automatizados de separar una porción de endosperma con una cuchilla afilada, practicar un pequeño agujero en la semilla y recoger el polvo resultante, cortar la semilla con un láser y perforar un disco de hoja. La planta de soja puede ser de una variedad de soja que se produce de forma no natural. En algunas realizaciones, el genoma de la planta o germoplasma de soja es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%,

90%, 95%, 97%, 99% o 100% idéntico al de una variedad de elite de soja. En algunas realizaciones, el marcador detectado en la muestra puede comprender, consistir esencialmente en o consistir en uno o más alelos marcadores ubicados dentro de un intervalo cromosómico seleccionado de:

1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR; o

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

Alternativamente, el uno o más alelos marcadores pueden ubicarse dentro de la SEQ ID NO: 2.

Métodos para producir una planta de soja resistente a enfermedades pueden comprender, consistir esencialmente en o consistir en detectar, en un germoplasma, un marcador asociado con resistencia potenciada a enfermedades (por ejemplo, ASR), donde dicho marcador se selecciona de la Tabla 1, o donde el marcador son locus estrechamente ligados de cualquier marcador descrito en la Tabla 1 y producir una planta de soja a partir de dicho germoplasma. Alternativamente, los métodos pueden comprender, consistir esencialmente en o consistir en detectar, en un germoplasma, un marcador asociado con SEQ ID NO: 2. El marcador se puede

detectar en cualquier muestra tomada del germoplasma, incluyendo, pero no limitado a una porción de dicho germoplasma (por ejemplo, una viruta de semilla o una célula de dicho germoplasma) o una secuencia de nucleótidos de dicho germoplasma. Una muestra de este tipo puede tomarse del germoplasma utilizando cualquier método actual o futuro conocido en la técnica, incluyendo, pero no limitado a métodos automatizados de separar una porción de endosperma con una cuchilla afilada, practicar un pequeño agujero en la semilla y recoger el polvo resultante, cortar la semilla con un láser y perforar un disco de hoja. El germoplasma puede ser de una variedad de soja que se produce de forma no natural. En algunas realizaciones, el genoma del germoplasma es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99%, o 100% idéntico al de una variedad de elite de soja. Una planta de soja resistente a enfermedades se produce entonces a partir del germoplasma identificado por tener el marcador asociado con resistencia potenciada a enfermedades (por ejemplo, ASR) de acuerdo con métodos bien conocidos en la técnica para reproducir y producir plantas a partir de germoplasma.

En algunas realizaciones, el marcador detectado en el germoplasma puede comprender, consistir esencialmente en o consistir en uno o más alelos marcadores ubicados dentro de un intervalo cromosómico seleccionado de:

- 1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.
- 2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR; o

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

En otra realización, el marcador se ubica con el ácido nucleico de la SEQ ID NO: 2. En algunas realizaciones, el marcador detectado en el germoplasma puede comprender, consistir esencialmente en o consistir en uno o más alelos marcadores seleccionados de la Tabla 1.

Métodos para producir y/o seleccionar una planta o germoplasma de soja resistente/tolerante a la roya de la soja asiática pueden comprender cruzar una primera planta o germoplasma de soja con una segunda planta o germoplasma de soja, en donde dicha primera planta o germoplasma de soja comprende un intervalo cromosómico seleccionado de:

1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR;

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde

el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1; o

5) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP mostrado en la Tabla 1; y cruzar con una segunda planta de soja que no comprende el intervalo cromosómico, produciendo entonces una planta de la descendencia con resistencia aumentada a la ASR. Una cualquiera de la primera y la segunda plantas o germoplasmas de soja, o ambos, puede ser de una variedad de soja que se produce de forma no natural. En algunas realizaciones, la segunda planta o germoplasma de soja es de una variedad de élite de soja. En algunas realizaciones, el genoma de la segunda planta o germoplasma de soja es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% o 100% idéntico al de una variedad de élite de soja.

En una realización, la planta de soja puede usarse para introducir resistencia/tolerancia a la roya de la soja asiática en la cepa *Glycine max* Williams 82.

Alternativamente, los métodos para producir y/o seleccionar una planta o germoplasma de soja resistente/tolerante a la roya de la soja asiática pueden comprender cruzar una primera planta o germoplasma de soja con una segunda planta o germoplasma de soja, en donde dicha primera planta o germoplasma de soja comprende la secuencia de ácido nucleico de la SEQ ID NO: 2 o una proteína que codifica para la secuencia de aminoácidos de la SEQ ID NO: 1.

También se proporciona en la presente un método de introgresar un alelo asociado con resistencia/tolerancia

aumentada a enfermedades (por ejemplo, ASR, SCN, SDS, RKN, *Phytophthora*, etc.) en una planta de soja. Dichos métodos para someter a introgresión un alelo asociado con resistencia/tolerancia potenciada a enfermedades (por ejemplo, ASR, SCN, SDS, RKN, *Phytophthora*, etc.) en una planta o germoplasma de soja puede comprender, consistir esencialmente en o consistir en cruzar una primera planta o germoplasma de soja que comprende dicho alelo (el donador), donde dicho alelo se selecciona de cualquier alelo enumerado en la Tabla 1 o un marcador en "cercana proximidad" a un marcador enumerado en la Tabla 1 con una segunda planta o germoplasma de soja que carece de dicho alelo (el progenitor recurrente) y retrocruzar repetidamente la descendencia que comprende dicho alelo con el progenitor recurrente. La descendencia que comprende dicho alelo se puede identificar detectando, en sus genomas, la presencia de un marcador asociado con resistencia/tolerancia aumentada a enfermedades (por ejemplo, ASR, SCN, SDS, RKN, *Phytophthora*, etc.). El marcador se puede detectar en cualquier muestra tomada de la descendencia, incluyendo, pero no limitado a una porción de dicha descendencia (por ejemplo, una viruta de semilla, un disco perforado de hojas o una célula de dicha planta o germoplasma) o una secuencia de nucleótidos de dicha descendencia. Una muestra de este tipo puede tomarse de la descendencia utilizando cualquier método actual o futuro conocido en la técnica, incluyendo, pero no limitado a métodos automatizados de separar una porción de endosperma con una cuchilla afilada, practicar un pequeño agujero en la semilla y recoger el polvo resultante, cortar la semilla con un láser y perforar un disco de hoja. Uno cualquiera del donador o el progenitor recurrente, o ambos, puede ser de una variedad de soja que se produce de forma no natural. En

algunas realizaciones, el progenitor recurrente es de una variedad de elite de soja. En algunas realizaciones, el genoma del progenitor recurrente es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% o 100% idéntico al de una variedad de elite de soja.

En algunas realizaciones, el marcador utilizado para identificar la descendencia comprende un alelo asociado con resistencia/tolerancia potenciada a la roya puede comprender, consistir esencialmente en o consistir en uno o más alelos marcadores ubicados dentro de un intervalo cromosómico seleccionado de:

- 1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

- 2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

- 3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR;

- 4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1; o

- 5) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM desde un SNP marcador que está asociado con resistencia aumentada a la ASR en la soja, donde el SNP marcador se selecciona del grupo que consiste en cualquier SNP marcador presentado en la Tabla 1 o cualquier marcador ligado estrechamente en cercana proximidad a dichos intervalos (SEQ ID NO 11, 12, o 13).

En algunas realizaciones, el marcador puede comprender, consistir esencialmente en o consistir en alelos de marcadores ubicados en al menos dos intervalos cromosómicos diferentes. Por ejemplo, el marcador puede comprender uno o más alelos ubicados en el intervalo cromosómico definido por y que incluye cualesquiera dos marcadores en la SEQ ID NO 11, 12, o 13.

Plantas de soja resistentes a enfermedades y germoplasmas

En otra realización, la presente invención proporciona plantas de soja y germoplasmas que son resistentes a la roya. Tal como se ha comentado anteriormente, los métodos de la presente invención se pueden utilizar para identificar, producir y/o seleccionar una planta o un germoplasma de soja resistente a enfermedades (por ejemplo, una planta de soja resistente o que tiene tolerancia aumentada a la roya de la soja asiática). Además de los métodos descritos anteriormente, una planta de soja o germoplasma resistente a ASR puede producirse por cualquier método en el que un marcador asociado con tolerancia potenciada a enfermedades se introduce en la planta o el germoplasma de soja, incluyendo, pero no limitado a transformación, transformación o fusión del protoplasto, una técnica haploide doble, rescate de embriones, edición de genes y/o por cualquier otro sistema de transferencia de ácidos nucleicos.

En algunas realizaciones, la planta o el germoplasma de soja comprende una variedad de soja que se produce de forma no natural. En algunas realizaciones, la planta o germoplasma de soja es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% o 100% idéntico al de una variedad de elite de soja.

La planta o el germoplasma de soja resistente a enfermedades puede ser la descendencia de un cruce entre una variedad de elite de soja y una variedad de soja que comprende un alelo asociado con resistencia a la roya potenciada (por ejemplo, ASR), donde el alelo está dentro de un intervalo cromosómico seleccionado de:

1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR;

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

En otras realizaciones, la planta o el germoplasma de soja resistente a enfermedades puede ser la descendencia de un cruce entre una variedad de elite de soja y una variedad de soja que comprende un alelo asociado con resistencia a la roya potenciada (por ejemplo, ASR), donde el alelo comprende la SEQ ID NO: 2 o codifica para una proteína de SEQ ID NO: 1.

Una realización de la invención es una planta de *Glycine max* que tiene resistencia/tolerancia a la roya de la soja asiática y que comprende la secuencia de ácido nucleico de SEQ ID NO: 2 o un intervalo cromosómico seleccionado de:

1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR; o

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

En determinadas realizaciones, el intervalo cromosómico confiere resistencia aumentada a la roya de la soja asiática (ASR) en comparación con una planta de control que no comprende dicho intervalo cromosómico. La planta de *Glycine max* puede derivarse de la cepa Williams 82

La planta o el germoplasma de soja resistente a enfermedades puede ser la descendencia de una introgresión, donde el progenitor recurrente es una variedad de elite de soja y el donador comprende un alelo asociado con tolerancia y/o resistencia potenciada a enfermedades, donde el donador porta un intervalo cromosómico o una porción del mismo que comprende una cualquiera de las SEQ ID NO: 11, 12, o 13, y donde el intervalo cromosómico comprende al menos un alelo seleccionado respectivamente de la Tabla 1.

La planta o el germoplasma de soja resistente a enfermedades puede ser la descendencia de un cruce entre una primera variedad de elite de soja (por ejemplo, una línea de ensayo) y la descendencia de un cruce entre una segunda

variedad de elite de soja (por ejemplo, un progenitor recurrente) y una variedad de soja que comprende un alelo asociado con tolerancia a la ASR potenciada (por ejemplo, un donador).

La planta o el germoplasma de soja resistente a enfermedades puede ser la descendencia de un cruce entre una primera variedad de elite de soja y la descendencia de una introgresión, donde el progenitor recurrente es una segunda variedad de elite de soja y el donador comprende un alelo asociado con tolerancia potenciada a ASR.

Una planta y un germoplasma de soja resistente a enfermedades de la presente invención puede comprender uno o más marcadores de la presente invención (por ejemplo, cualquier marcador descrito en la Tabla 1; o cualquier marcador en estrecha proximidad con cualquier marcador tal como se describe en la Tabla 1).

En algunas realizaciones, la planta y el germoplasma de soja resistente a enfermedades puede comprender dentro de su genoma un marcador asociado con tolerancia potenciada a ASR, donde dicho marcador está ubicado dentro de un intervalo cromosómico seleccionado de:

- 1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

- 2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

- 3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR; o

- 4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está

asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

En otras realizaciones, el marcador se ubica con la secuencia de ácido nucleico de la SEQ ID NO: 2.

En algunas realizaciones, la planta o el germoplasma de soja resistente a enfermedades puede comprender dentro de su genoma un marcador que comprende, consiste esencialmente en o consiste en alelos de marcadores ubicados en al menos dos intervalos cromosómicos diferentes. Por ejemplo, el marcador puede comprender uno o más alelos ubicados en el intervalo cromosómico definido por y que incluye cualquier combinación de dos marcadores en la Tabla 1 y uno o más alelos ubicados en el intervalo cromosómico definido por y que incluye cualquier combinación de dos marcadores ubicados en la Tabla 1.

En determinadas realizaciones, la planta de soja resistente a enfermedades se deriva de la cepa de *Glycine max* Williams 82.

Semillas de soja resistentes a enfermedades

La presente invención también proporciona semillas de soja resistentes a enfermedades. Tal como se ha comentado anteriormente, los métodos de la presente invención se pueden utilizar para identificar, producir y/o seleccionar una semilla de soja resistente a enfermedades. Además de los métodos descritos anteriormente, una semilla de soja resistente a enfermedades se puede producir por cualquier método en el que un marcador asociado con tolerancia potenciada a ASR se introduce en la semilla de soja, incluyendo, pero no limitado a transformación, transformación o fusión del protoplasto, una técnica haploide doble, rescate de embriones, edición de genes (por ejemplo, CRISPR o TALEN o

MegaNucleasas) y/o por cualquier otro sistema de transferencia de ácidos nucleicos.

Una realización de la invención es una semilla de la cepa de *Glycine max* Williams 82 que se ha modificado para que tenga resistencia/tolerancia a la roya de la soja asiática i

En algunas realizaciones, la semilla de soja resistente a enfermedades comprende una variedad de soja que se produce de forma no natural. En algunas realizaciones, la semilla de soja es al menos aproximadamente un 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 99% o 100% idéntico al de una variedad de elite de soja.

La semilla de soja resistente a enfermedades se puede producir por una planta de soja resistente a enfermedades identificada, producida o seleccionada por los métodos de la presente invención. En algunas realizaciones, la semilla de soja resistente a enfermedades se produce por una planta de soja resistente a enfermedades o una planta de glicina silvestre (por ejemplo, *Glycine tomentella*) que comprende uno cualquiera de los intervalos cromosómicos correspondientes a las SEQ ID NO: 11-13 o un intervalo cromosómico que comprende SEQ ID NO: 2, o un intervalo cromosómico que codifica para una proteína de la SEQ ID NO: 1, o cualquier parte de estos intervalos que codifican la resistencia a la ASR.

Una semilla de soja resistente a enfermedades de la presente invención puede comprender, puede seleccionarse por o producirse mediante el uso de uno o más marcadores de la Tabla 1 de la presente invención.

En algunas realizaciones, la semilla de soja resistente a enfermedades puede comprender dentro de su genoma un marcador asociado con tolerancia potenciada a ASR, donde dicho marcador está ubicado dentro de un intervalo cromosómico seleccionado del grupo que consiste en:

1) un intervalo cromosómico que comprende SEQ ID NO: 2, o una parte de la misma que codifica la resistencia a la ASR.

2) un intervalo cromosómico que codifica para la secuencia de proteínas de SEQ ID NO: 1, o una parte de la misma que codifica la resistencia a la ASR.

3) un intervalo cromosómico que comprende SEQ ID NO 11, 12, o 13, o una parte de la misma que codifica para resistencia a la ASR; o

4) un intervalo cromosómico que abarca 20 cM, 15 cM, 10 cM, 5 cM, 1 cM, 0.5 cM de un marcador de SNP que está asociado con resistencia aumentada a la ASR en la soja, donde el marcador de SNP se selecciona del grupo que consiste en cualquier marcador de SNP favorable mostrado en la Tabla 1.

En otras realizaciones, el marcador se ubica con la secuencia de ácido nucleico de la SEQ ID NO: 2.

Proteínas que confieren resistencia a la roya

Además de proporcionar plantas *Glycine max* que tienen una resistencia aumentada a la roya de la soja asiática, en determinadas realizaciones, la invención proporciona proteínas que están relacionadas con la resistencia a la roya, en particular la resistencia a la roya de la soja asiática (en el presente documento, "ASR"). En realizaciones particulares, estas proteínas confieren resistencia aumentada a la soja asiática. La proteína y el gen codificante de la misma pueden usarse para proteger las plantas de los patógenos de roya.

En determinadas realizaciones, las proteínas se codifican por la secuencia de ácido nucleico de la SEQ ID NO: 2. En otras realizaciones, las proteínas tienen al menos el 75%, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% de identidad

con una proteína codificada por la secuencia de ácido nucleico de la SEQ ID NO: 1.

En otras realizaciones, las proteínas se codifican por un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ NO: 11, 12, o 13. En otras realizaciones, las proteínas tienen al menos el 75%, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% de identidad con una proteína codificada por un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ NO: 11, 12, o 13.

En determinada realización, la proteína de las presentes descripciones se deriva de la soja y se denomina RppRC1. En una realización de la invención, la proteína tiene la secuencia de aminoácidos de la SEQ ID NO: 1. En otra realización de la invención, la proteína tiene una secuencia de aminoácidos de al menos el 75%, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% idéntica a la secuencia de aminoácidos de la SEQ ID NO: 2. En una realización, la proteína de la presente descripción puede ser una cualquiera de las siguientes proteínas:

(A1) una proteína que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A2) una proteína que tiene sustitución y/o delección y/o adición de uno o varios residuos de aminoácidos de y que tiene la misma función que la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A3) una proteína que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con y que tiene la misma función que la secuencia de aminoácidos definida en o bien (A1) o bien (A2); y

(A4) una proteína de fusión obtenida etiquetando en el N terminal y/o C terminal de la proteína definida en uno cualquiera de (A1) a (A3).

En otra realización, la proteína de la presente descripción puede ser una cualquiera de las siguientes proteínas:

(A5) una proteína que tiene la secuencia de aminoácidos codificada por un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ NO: 11, 12, o 13;

(A6) una proteína que tiene sustitución y/o delección y/o adición de uno o varios residuos de aminoácidos de y que tiene la misma función que la secuencia de aminoácidos codificada por un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ NO: 11, 12, o 13;

(A7) una proteína que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con y que tiene la misma función que la secuencia de aminoácidos definida en o bien (A5) o bien (A6); y

(A8) una proteína de fusión obtenida etiquetando en el N terminal y/o C terminal de la proteína definida en uno cualquiera de (A5) a (A7).

En estas proteínas, la etiqueta se refiere a un polipéptido o proteína que se fusiona y se expresa junto con la proteína de interés usando tecnología de recombinación *in vitro* de ADN, para facilitar la expresión, detección, rastreo y/o purificación de la proteína de interés. La etiqueta puede ser una etiqueta FLAG, una etiqueta His, una etiqueta MBP, una etiqueta HA, una etiqueta myc, una etiqueta GST y/o una etiqueta SUMO, etc.

En estas proteínas, identificar se refiere a la identidad entre secuencias de aminoácidos. Los sitios web de recuperación de homología en Internet pueden usarse para

determinar la identidad entre secuencias de aminoácidos, tal como la página web BLAST en el sitio web de la página de inicio de NCBI. Por ejemplo, el valor de identidad (%) puede obtenerse en BLAST2.1 avanzado usando blastp como el programa, configurando el valor de Expect a 10, configurando todos los filtros en Apagado, usando BLOSUM62 como Matrix, configurando el coste de existencia de espacio, por residuo coste de espacio y relación de Lambda a 11, 1 y 0.85 (valores por defecto), respectivamente, y recuperando la identidad de un par de secuencias de aminoácidos para el cálculo.

Las proteínas de la presente invención pueden producirse a partir de las moléculas de ácido nucleico dadas a conocer en el presente documento o mediante el uso de técnicas de biología molecular estándar.

La presente invención abarca una proteína aislada o sustancialmente purificada. La proteína "aislada" o "purificada" o una porción biológicamente activa de la misma está sustancialmente o en gran medida libre de componentes concomitantes con o que interactúan con la proteína que normalmente están presentes en el entorno natural de la proteína. La proteína que está sustancialmente libre de materiales celulares incluye formulaciones de proteínas que tienen menos de aproximadamente el 30%, aproximadamente el 20%, aproximadamente el 10%, aproximadamente el 5%, o aproximadamente el 1% (en peso seco) de proteínas contaminantes. Cuando la proteína de la porción biológicamente activa de la misma en las realizaciones se producen por métodos recombinantes, lo más preferiblemente, el medio tiene menos de aproximadamente el 30%, aproximadamente el 20%, aproximadamente el 10%, aproximadamente el 5% o aproximadamente el 1% (en peso seco) de precursores químicos o productos químicos que no son

proteínas de interés. Los fragmentos y variantes relacionados con las proteínas están dentro del alcance de la presente descripción.

Las proteínas variantes abarcadas por la presente invención son bioactivas, es decir, continúan teniendo la bioactividad requerida (es decir, la capacidad de mejorar la resistencia de la planta (es decir, la resistencia de la planta frente a los patógenos fúngicos) tal como se describe en la presente invención) de proteínas nativas. Tales variantes pueden obtenerse, por ejemplo, por un polimorfismo genético o por manipulación humana. Las variantes bioactivas de la proteína nativa de la presente invención pueden tener al menos aproximadamente el 40%, aproximadamente el 45%, aproximadamente el 50%, aproximadamente el 55%, aproximadamente el 60%, aproximadamente el 65%, aproximadamente el 70%, aproximadamente el 75%, aproximadamente el 80%, aproximadamente el 85%, aproximadamente el 90%, aproximadamente el 91%, aproximadamente el 92%, aproximadamente el 93%, aproximadamente el 94%, aproximadamente el 95%, aproximadamente el 96%, aproximadamente el 97%, aproximadamente el 98%, aproximadamente el 99% o más de identidad de secuencia con la secuencia de aminoácidos de la proteína nativa o a SEQ ID NO: 1, tal como se determina por los programas de alineación de secuencia conocidos en la técnica. Las variantes biológicamente activas de la proteína dada a conocer en la presente invención pueden diferir de la proteína por tan solo aproximadamente de 1 a 15 residuos de aminoácidos, tan solo aproximadamente de 1 a 10 (por ejemplo, aproximadamente de 6 a 10), tan solo aproximadamente 5, tan solo 4, 3, 2 o incluso 1 residuo de aminoácido.

Las proteínas dadas a conocer en la presente solicitud puede modificarse, por ejemplo, incluyendo la sustitución, delección, truncamiento e inserción de aminoácidos. Los métodos de tal manipulación se conocen en la técnica. Por ejemplo, pueden prepararse variantes de secuencia de aminoácidos y fragmentos de proteínas resistentes mediante mutación en el ADN. Se conocen en la técnica métodos de mutagénesis y modificación de polinucleótidos.

En una realización, la proteína de la invención es un fragmento biológicamente activo de la SEQ ID NO: 1, que puede proteger plantas de patógenos de la roya.

Las proteínas dadas a conocer en la presente invención también abarcan proteínas que se producen de manera natural y variantes, fragmentos y formas modificadas de las mismas. Tales variantes y fragmentos aún tendrán la capacidad requerida para conferir o potenciar resistencia de la planta frente a los patógenos fúngicos.

Moléculas de ácido nucleico

Otra realización de la invención se refiere a moléculas de ácido nucleico relacionadas con la resistencia a la roya. Estas moléculas de ácido nucleico codifican para una proteína de la invención, es decir, una proteína que confiere una resistencia aumentada a la roya, tal como por ejemplo, la roya de la soja asiática. La molécula de ácido nucleico puede ser ARN, tal como ADNc, ADN genómico o ADN recombinante. La molécula de ácido nucleico también puede ser ARN, tal como ARNm.

En determinadas realizaciones, la molécula de ácido nucleico es el gen *RppRC1* (denominado *RppRC1*). En una realización de la invención, la molécula de ácido nucleico tiene la secuencia de ácido nucleico de la SEQ ID NO: 2. En otra realización, la molécula de ácido nucleico tiene una

secuencia de ácido nucleico de al menos el 75%, al menos el 85%, al menos el 90%, al menos el, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% idéntica a la secuencia de ácido nucleico de la SEQ ID NO: 2. Las realizaciones de las moléculas de ácido nucleico de la presente descripción incluyen:

(B1) una molécula de ADN que tiene la secuencia de ácido nucleico de la SEQ ID NO: 2;

(B2) una molécula de ADN que hibrida con la secuencia de ácido nucleico de la SEQ ID NO: 2 en una condición estricta y que codifica para la proteína descrita anteriormente; y

(B3) una molécula de ADN que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con las secuencias de ADN definidas en (B1) y (B2) y que codifica para la proteína descrita anteriormente.

En otras realizaciones, la molécula de ácido nucleico comprende un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ ID NO: 11, 12, o 13, o una parte de la misma que codifica la resistencia a la ASR. En otra realización, la molécula de ácido nucleico tiene una secuencia de ácido nucleico de al menos el 75%, al menos el 85%, al menos el 90%, al menos el, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% idéntica a un intervalo cromosómico que comprende la secuencia de ácido nucleico de la SEQ ID NO: 11, 12, o 13. En determinadas realizaciones, las moléculas de ácido nucleico codifican la resistencia a la ASR. Las realizaciones de tales moléculas de ácido nucleico incluyen:

(B4) una molécula de ADN que tiene la secuencia de ácido nucleico de la SEQ ID NO: 11, 12, o 13, o una parte de la misma que codifica la resistencia a la ASR;

(B5) una molécula de ADN que hibrida con la secuencia de ácido nucleico de la SEQ ID NO: 11, 12 o 13 en una condición estricta y que codifica para la proteína descrita anteriormente; y

(B6) una molécula de ADN que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con las secuencias de ADN definidas en (B4) y (B4) y que codifica para una proteína que confiere resistencia aumentada a la roya frente a una planta.

En cuanto a los genes anteriores, la condición estricta puede ser tal como sigue: hibridación a 50°C en una disolución mezclada de dodecil sulfato de sodio al 7% (SDS), Na₃PO₄ 0.5 M y EDTA 1 mM, y aclaramiento a 50°C en 2 × SSC, SDS al 0.1%; la condición estricta también puede ser: hibridación a 50°C en una disolución mezclada de SDS al 7%, Na₃PO₄ 0.5 M y EDTA 1 mM, y aclaramiento a 50°C en 1 × SSC, SDS al 0.1%; la condición estricta también puede ser: hibridación a 50°C en una disolución mezclada de SDS al 7%, Na₃PO₄ 0.5 M y EDTA 1 mM, y aclaramiento a 50°C en 0.5 × SSC, SDS 0.1%; la condición estricta también puede ser: hibridación a 50°C en una disolución mezclada de SDS al 7%, Na₃PO₄ 0.5 M y EDTA 1 mM, y aclaramiento a 50°C en 0.1 × SSC, SDS al 0.1%; la condición estricta también puede ser: hibridación a 50°C en una disolución mezclada de SDS al 7%, Na₃PO₄ 0.5 M y EDTA 1 mM, y aclaramiento a 65°C en 0.1 × SSC, SDS al 0.1%; la condición estricta también puede ser: hibridación a 65°C en una disolución de 6 × SSC y SDS al 0.5%, y luego lavado de la membrana una vez con 2 × SSC y SDS al 0.1%, y una vez con 1 × SSC y SDS al 0.1%, respectivamente.

En otra realización, la molécula de ácido nucleico codifica para el aminoácido de la SEQ ID NO: 1 o una proteína que tiene una secuencia de aminoácidos de al menos el 75%, al menos el 85%, al menos el 90%, al menos el 95%, al menos el 97%, al menos el 98%, o al menos el 99% idéntica a la secuencia de aminoácidos de la SEQ ID NO: 1.

La presente invención abarca una molécula de ácido nucleico aislada o sustancialmente purificada. La molécula de ácido nucleico "aislada" o "purificada" o una porción biológicamente activa de la misma está sustancialmente o en gran medida libre de componentes concomitantes con o que interactúan con la molécula de ácido nucleico que normalmente están presentes en el entorno natural de la molécula de ácido nucleico. Por tanto, la molécula de ácido nucleico aislada o purificada o proteína está sustancialmente libre de otros materiales o medios celulares cuando se produce mediante técnicas recombinantes (tal como la amplificación por PCR), o precursores químicos u otros productos químicos cuando se sintetiza por métodos químicos. Lo más preferiblemente, la molécula de ácido nucleico "aislada" no comprende secuencias (por ejemplo, secuencias de codificación de proteínas) que se ubican de manera natural flanqueando la molécula de ácido nucleico (es decir, secuencias ubicadas en los extremos 5' y 3' de la molécula de ácido nucleico) en el ADN genómico del organismo del cual se deriva la molécula de ácido nucleico. Por ejemplo, en algunas realizaciones de la presente invención, la molécula de ácido nucleico aislada puede comprender menos de aproximadamente 5 kb, aproximadamente 4 kb, aproximadamente 3 kb, aproximadamente 2 kb, aproximadamente 1 kb, aproximadamente 0.5 kb, o aproximadamente 0.1 kb de secuencias de nucleótidos que se ubican de manera natural flanqueando la molécula de ácido

nucleico en el ADN genómico de la célula de la que se deriva la molécula de ácido nucleico. Los fragmentos y variantes relacionados con las secuencias de nucleótidos codificadas están dentro del alcance de la presente descripción. "Un fragmento" y variaciones gramaticales del mismo se refieren a una porción de una secuencia de nucleótidos o una porción de una secuencia de aminoácidos y una proteína codificada de ese modo. Los fragmentos de la secuencia de nucleótidos pueden codificar para fragmentos de proteína que retienen la actividad biológica de proteínas naturales y tienen la capacidad de conferir resistencia (es decir, antifúngica) en plantas. Alternativamente, los fragmentos de la secuencia de nucleótidos que pueden usarse como sondas de hibridación no codifican necesariamente para fragmentos de proteína que mantienen la actividad biológica. Por tanto, el fragmento de la secuencia de nucleótidos puede estar en el intervalo de al menos aproximadamente 15 nucleótidos, aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos y como máximo la secuencia de nucleótidos de longitud completa que codifica para la proteína descrita en el presente documento.

El fragmento de la secuencia de nucleótidos que codifica para la porción biológicamente activa de la proteína dada a conocer puede codificar para al menos aproximadamente 15, aproximadamente 25, aproximadamente 30, aproximadamente 40, o 45, aproximadamente 50 aminoácidos consecutivos o como máximo el número total de aminoácidos presentes en la proteína de longitud completa de esta realización (por ejemplo, 857 aminoácidos para SEQ ID NO: 1). Los fragmentos de secuencias de nucleótidos que pueden usarse como sondas de hibridación o cebadores de PCR normalmente no tienen que codificar para porciones biológicamente activas de proteínas.

Cuando se hace referencia a una molécula de ácido nucleico especificada, el término "secuencia de longitud completa" se refiere a la secuencia completa de ácido nucleico de una secuencia nativa. "Una secuencia nativa" y variaciones gramaticales de la misma se usan en la presente invención para referirse a una secuencia endógena, es decir, una secuencia no modificada por ingeniería presente en el genoma de un organismo.

Por tanto, el fragmento de la secuencia de nucleótidos dada a conocer en la presente invención puede codificar para una porción biológicamente activa de una proteína, o puede ser un fragmento usado como una sonda de hibridación o cebador de PCR. En determinadas realizaciones, la molécula de ácido nucleico de la presente invención comprende al menos aproximadamente 15, aproximadamente 20, aproximadamente 50, aproximadamente 75, aproximadamente 100, o aproximadamente 150 nucleótidos o como máximo el número de nucleótidos presentes en la secuencia de nucleótidos de longitud completa dada a conocer en el presente documento (por ejemplo, 2574 nucleótidos para SEQ ID NO: 2).

Los expertos en la técnica reconocerán que las variantes de ácido nucleico de la presente invención se configurarán de modo que se mantiene el marco de lectura abierta. Para las moléculas de ácido nucleico, las variantes conservadas comprenden aquellas secuencias que codifican para secuencias de aminoácidos en las proteínas de la presente invención debido a la degeneración del código genético. Las variantes alélicas nativas pueden identificarse mediante técnicas biológicas moleculares bien conocidas, tales como las técnicas de reacción en cadena de la polimerasa (PCR) y técnicas de hibridación. Las moléculas de ácido nucleico variantes también comprenden moléculas de ácido nucleico

sintéticas, tales como las generadas mediante el uso de mutagénesis dirigida al sitio pero que todavía codifican para las proteínas de la presente invención. En general, las variantes de una molécula de ácido nucleico particular dada a conocer en el presente documento pueden tener al menos aproximadamente el 40%, aproximadamente el 45%, aproximadamente el 50%, aproximadamente el 55%, aproximadamente el 60%, aproximadamente el 65%, aproximadamente el 70%, aproximadamente el 75%, aproximadamente el 80%, aproximadamente el 85%, aproximadamente el 90%, aproximadamente el 91%, aproximadamente el 92%, aproximadamente el 93%, aproximadamente el 94%, aproximadamente el 95%, aproximadamente el 96%, aproximadamente el 97%, aproximadamente el 98%, aproximadamente el 99% o más de identidad de secuencia con la molécula de ácido nucleico particular, tal como se determina por los programas de alineación de secuencia bien conocidos en la técnica.

Las variantes de una molécula de ácido nucleico particular (es decir, una molécula de ácido nucleico de referencia) de la presente invención también pueden evaluarse comparando los porcentajes de identidad de secuencia entre la proteína codificada por la molécula de ácido nucleico variante y la proteína codificada por la molécula de ácido nucleico de referencia. El porcentaje de identidad de secuencia entre dos proteínas cualquiera puede calcularse usando programas de alineación de secuencia conocidos en la técnica. En el caso en el que se evalúa cualquier par de moléculas de ácido nucleico dado de la presente invención comparando el porcentaje de identidad de secuencia compartida por las dos proteínas codificadas por el par de moléculas de ácido nucleico dado, el porcentaje de identidad de secuencia

entre las dos proteínas codificadas es de al menos aproximadamente el 40%, aproximadamente el 45%, aproximadamente el 50%, aproximadamente el 55%, aproximadamente el 60%, aproximadamente el 65%, aproximadamente el 70%, aproximadamente el 75%, aproximadamente el 80%, aproximadamente el 85%, aproximadamente el 90%, aproximadamente el 91%, aproximadamente el 92%, aproximadamente el 93%, aproximadamente el 94%, aproximadamente el 95%, aproximadamente el 96%, aproximadamente el 97%, aproximadamente el 98%, aproximadamente el 99%, o más de identidad de secuencia.

Las proteínas y moléculas de ácido nucleico variantes también abarcan secuencias y proteínas obtenidas mediante mutagénesis y procedimientos de recombinación, incluyendo pero sin limitarse a procedimientos tales como intercambio de ADN. Una biblioteca de polinucleótidos recombinantes puede generarse de un grupo de polinucleótidos con secuencias relacionadas que comprenden regiones en las secuencias que tienen identidad de secuencia sustancial y son capaces de recombinación homóloga *in vitro* o *in vivo*. Por ejemplo, usando este método, un motivo de secuencia que codifica para un dominio de interés pueden intercambiarse entre el gen de la proteína dado a conocer en la presente invención y otros genes de la proteínas conocidos para obtener un nuevo gen que codifica para una proteína que tiene propiedades mejoradas de interés, tal como una capacidad mejorada para conferir o potenciar resistencia de plantas a patógenos fúngicos. Tales estrategias de intercambio de ADN se conocen en la técnica.

La presente descripción abarca secuencias que se aíslan basándose en su identidad de secuencia con la secuencia entera mostrada en el presente documento o las variantes y

fragmentos de la misma. Tales secuencias incluyen secuencias que son ortólogas de las secuencias dadas a conocer. Se consideran que los genes presentes en diferentes especies son ortólogos cuando sus secuencias de nucleótidos y/o secuencias de proteínas codificadas comparten por lo tanto al menos aproximadamente el 60%, aproximadamente el 70%, aproximadamente el 75%, aproximadamente el 80%, aproximadamente el 85%, aproximadamente el 90%, aproximadamente el 91%, aproximadamente el 92%, aproximadamente el 93%, aproximadamente el 94%, aproximadamente el 95%, aproximadamente el 96%, aproximadamente el 97%, aproximadamente el 98%, aproximadamente el 99% o más de identidad de secuencia. La función de los ortólogos a menudo está altamente conservada en diversas especies. Por tanto, la presente descripción abarca moléculas de ácido nucleico aisladas que codifican para proteínas que confieren o potencian resistencia a hongos fitopatógenos e se hibridan con las secuencias dadas a conocer en la presente invención o variantes o fragmentos de las mismas.

En los métodos de PCR, pueden diseñarse cebadores de oligonucleótidos para reacciones de PCR para amplificar secuencias de ADN correspondientes a partir de ADNc o ADN genómico extraído de cualquier organismo de interés. Se conocen en la técnica métodos para diseñar cebadores de PCR y para clonar mediante PCR y se dan a conocer en los siguientes documentos: Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual* (2^a ed., Cold Spring Harbor Laboratory Press, Plainview, N.Y.). Los métodos de PCR conocidos incluyen, pero no se limitan a, métodos que usan cebadores emparejados, cebadores anidados, cebadores específicos individuales, cebadores degenerados, cebadores específicos de

genes, cebadores específicos de vectores, cebadores con apareamiento parcialmente erróneo, etc.

En las técnicas de hibridación, se usa la totalidad o una parte de una molécula de ácido nucleico conocida como sonda que hibrida selectivamente con otros polinucleótidos correspondientes presentes en un conjunto de fragmentos de ADN genómicos o fragmentos de ADNc clonados (es decir, biblioteca genómica o de ADNc) de un organismo seleccionado. La sonda de hibridación puede ser un fragmento de ADN genómico, un fragmento de ADNc, un fragmento de ARN u otro oligonucleótido y puede etiquetarse con grupos detectables tal como ^{32}P u otros marcadores detectables. Por tanto, por ejemplo, la sonda de hibridación puede prepararse etiquetando el oligonucleótido sintético basándose en el polinucleótido de esta realización. Se conocen en la técnica métodos para preparar las sondas de hibridación y construir bibliotecas de ADNc y bibliotecas genómicas.

Pueden usarse diversos procedimientos para determinar la presencia o ausencia de secuencias específicas de ADN, ARN o proteína. Estos incluyen, por ejemplo, inmunotransferencia de tipo Southern, inmunotransferencia de tipo Northern, inmunotransferencia de tipo Western y análisis ELISA. Se conocen bien estas técnicas en la técnica.

La proteína y el gen codificante de la misma y métodos dados a conocer en la presente invención pueden usarse para regular el contenido de una o más proteínas en plantas. El término "regular" y variaciones gramaticales del mismo se usan en la presente invención para referirse a un aumento o disminución en el contenido en proteína en plantas genéticamente modificadas (es decir, transformadas) con respecto al contenido en proteína de las correspondientes plantas no transformadas (es decir, plantas que no se han

modificado genéticamente según los métodos de la presente descripción).

Tal como se usa en el presente documento, el término "expresión" y variaciones gramaticales del mismo se refieren a la biosíntesis o procedimiento por los cuales se genera un polinucleótido, incluyendo la transcripción y/o traducción de un producto génico. Por ejemplo, las moléculas de ácido nucleico dadas a conocer en la presente invención pueden transcribirse a partir de moldes de ADN (tales como en ARNm u otros en transcriptos de ARN) y/o el ARNm transcrito se traduce posteriormente en las proteínas. El término "producto génico" y variaciones gramaticales del mismo pueden referirse a, por ejemplo, transcriptos y proteínas codificadas. La inhibición (o aumento) de la expresión o función del producto génico (es decir, el producto génico de interés) puede realizarse en un entorno en el que se hacen comparaciones entre dos plantas, por ejemplo, la expresión o función de un producto génico en Las plantas modificadas genéticamente son relativas a la expresión o función del producto génico en plantas de tipo natural u otras plantas susceptibles correspondientes pero susceptibles. El nivel de expresión de un producto génico en una planta de tipo natural puede no existir. Por ejemplo, una planta de "tipo natural" puede ser una planta, una célula vegetal o una parte de la planta que no expresa un gen de resistencia exógena.

Alternativamente, la inhibición (o aumento) de la expresión o función del producto génico diana están en un entorno en el que se hacen comparaciones entre células vegetales, orgánulos, órganos, tejidos de partes de plantas dentro de la misma planta o plantas diferentes e incluyen comparaciones entre las etapas de desarrollo o temporales de la misma planta o plantas diferentes. Cualquier método o

composición que regula por disminución la expresión del producto génico diana o regula por disminución la actividad funcional del producto génico diana en el nivel de transcripción o traducción puede usarse para lograr la inhibición de la expresión o función del producto génico diana. De manera similar, cualquier método o composición que induce o regula por incremento la expresión del producto génico diana en el nivel de transcripción o traducción, o aumenta o activa o regula por incremento la actividad funcional del producto génico diana, puede usarse para lograr la expresión o función aumentada del gen o proteína diana. Se conocen bien en la técnica métodos para inhibir o potenciar la expresión génica.

Los genes y moléculas de ácido nucleico dados a conocer en la presente invención incluyen mutantes y secuencias que se producen de manera natural o formas modificadas de los mismos. Las proteínas dadas a conocer en la presente invención también abarcan proteínas que se producen de manera natural y variantes, fragmentos y formas modificadas de las mismas. Tales variantes y fragmentos aún tendrán la capacidad requerida para conferir o potenciar resistencia de la planta frente a los patógenos fúngicos. En una realización, las mutaciones que van a realizarse en el ADN que codifica para las variantes o fragmentos en general no colocan la secuencia fuera del marco de lectura y, preferiblemente, no producirán regiones complementarias, que pueden producir estructuras de ARNm secundarias.

Casetes de expresión, vectores recombinantes, bacterias recombinantes, o líneas celulares transgénicas

La presente solicitud también proporciona un casete de expresión, un vector recombinante, una bacteria recombinante

o una línea celular transgénica que contiene la molécula de ácido nucleico descrita anteriormente.

El casete de expresión se refiere a ADN que puede expresar la proteína en una célula huésped, y el ADN no solo incluye un promotor que inicia la transcripción del gen que codifica para la proteína, sino también incluye un terminador que termina la transcripción. Además, el casete de expresión puede incluir adicionalmente una secuencia potenciadora. Los promotores que pueden usarse en la presente invención incluyen, pero no se limitan a: promotores constitutivos, promotores específicos de tejido, órgano y desarrollo y promotores inducibles. Los ejemplos de promotores incluyen, pero no se limitan a: promotor constitutivo 35S de virus del mosaico de la coliflor; un promotor inducible por trauma de tomate, leucina aminopeptidasa ("LAP", Chao *et al.*, (1999) *Plant Physiol* 120: 979-992); un promotor químicamente inducible del tabaco, patogénesis relacionada 1 (PR1) (inducido por ácido salicílico y BTH (éster S-metílico de benzotiadiazol-7-tiohidroxiácido)); promotor del inhibidor de proteasa de tomate II (PIN2) o promotor LAP (ambos inducidos por jasmonato de metilo); promotor de choque térmico; promotor inducible por tetraciclina; promotores específicos de semillas, tales como el promotor específico de semillas de mijo pF128, promotores específicos de proteínas de almacenamiento de semillas (por ejemplo, promotores de beta conglucina de faseolina, napina, oleosina y soja (Beachy *et al.*, (1985) *EMBO J.* 4: 3047-3053)). Se pueden usar solos o en combinación con otros promotores de plantas. Los terminadores de la transcripción adecuados incluyen, pero no se limitan a: Terminador de nopalina sintasa de *Agrobacterium* (terminador NOS), terminador del virus del mosaico de la coliflor CaMV 35S, terminador tml, terminador pea rbcS E9 y terminador de

nopalina y octopina sintasa (véase, por ejemplo: Odell et al., (1985) *Nature*, 313: 810; Rosenberg et al., (1987) *Gene*, 56: 125; Guerineau et al., (1991) *Mol. Gen. Genet*, 262: 141; Proudfoot (1991) *Cell*, 64: 671; Sanfacon et al., *Genes Dev.*, 5: 141; Mogen et al., (1990) *Plant Cell*, 2: 1261; Munroe et al., (1990) *Gene*, 91: 151; Ballad et al., (1989) *Nucleic Acids Res.*, 17: 7891; Joshi et al., (1987) *Nucleic Acid Res.*, 15: 9627).

Construcción de un vector de expresión recombinante que contiene la molécula de ácido nucleico. Los vectores de expresión de plantas usados pueden ser vectores del sistema Gateway o vectores binarios de *Agrobacterium*, tales como pGWB411, pGWB412, pGWB405, pBin438, pCAMBIA1302, pCAMBIA2301, pCAMBIA1301, pCAMBIA1300, pBI121, pCAMBIA1391-Xa o pCAMBIA1391-Xb.

Cuando se usa *RppRC1* para construir un vector de expresión recombinante, se puede añadir cualquier promotor inducible o específico de tejido, constitutivo, mejorado antes del inicio de la transcripción de nucleótido, tal como el promotor 35S del virus del mosaico de la coliflor (CAMV), el promotor ubiquitina del gen de ubiquitina (pUbi), etc. ., que puede usarse solo o en combinación con otros promotores de plantas. Además, cuando se usa el gen y/o intervalos cromosómicos de la presente invención para construir un vector de expresión vegetal, también puede usarse potenciadores, incluyendo potenciadores de la traducción o potenciadores de la transcripción, y estas regiones potenciadoras pueden ser codones de inicio de ATG o regiones adyacentes a los codones de inicio, etc. , pero debe estar en el mismo marco de lectura que la secuencia de codificación para garantizar la traducción correcta de toda la secuencia. Las fuentes de las señales de control de traducción y los

codones de inicio son extensas y pueden ser naturales o sintéticas. La región de iniciación traduccional puede ser a partir de una región de iniciación transcripcional o un gen estructural.

Con el fin de facilitar la identificación y selección de células vegetales transgénicas o plantas, el vector de expresión vegetal puede modificarse por ingeniería, por ejemplo, añadiendo un gen (gen GUS, gen luciferasa, etc.) que puede expresarse en una planta y codifica para una enzima que puede producir un cambio de color o un compuesto luminiscente, un marcador antibiótico que tiene resistencia (marcador de gentamicina, marcador de kanamicina, etc.), o un gen marcador que tiene resistencia frente a un reactivo químico (tal como un gen resistente a herbicidas), etc.

La línea celular transgénica puede ser o bien un material propagador o bien un material no propagador.

En una realización específica, el promotor para la iniciación de la transcripción del gen codificante de la proteína en el casete de expresión es específicamente el promotor endógeno original del gen codificante, y la secuencia de nucleótidos del promotor endógeno original del gen codificante se muestra en SEQ ID NO: 7.

En otra realización de la invención, el vector recombinante es específicamente un plásmido recombinante obtenido clonando la molécula de ácido nucleico (SEQ ID NO: 2) descrita anteriormente entre los sitios attR1 y attR2 del vector pB2GW7, y reemplazando el promotor 35S entre los sitios de digestión de enzimas SacI y SpeI con el promotor endógeno del gen *RppRC1* mostrado en la SEQ ID NO: 7.

El gen de resistencia dado a conocer en la presente invención puede expresarse como un transgén para producir una planta resistente a la roya. El uso de diferentes promotores

descritos en la presente invención o conocidos por los expertos en la técnica provocarán expresión génica que va a regularse en diferentes condiciones (es decir, pueden seleccionarse promotores basándose en resultados deseados). Por ejemplo, un mayor nivel de expresión en un sistema u órgano tisular particular (por ejemplo, hojas) puede requerirse para potenciar la resistencia. El gen completo (por ejemplo, tanto el promotor natural como la secuencia codificante) puede insertarse como un transgén, lo que permite una rápida combinación con otros rasgos como la resistencia a los insectos o la resistencia a los herbicidas.

En algunas realizaciones, las secuencias de ácido nucleico pueden superponerse con cualquier combinación de secuencias moleculares de ácido nucleico de interés para formar plantas que tengan los fenotipos deseados. Esta superposición puede lograrse por una combinación de genes en un constructo de ADN, o por hibridación de una o más plantas con transgenes con otra cepa vegetal que comprende una combinación deseada. Por ejemplo, las moléculas de ácido nucleico o fragmentos de las mismas dadas a conocer en la presente invención pueden superponerse con cualquier otra molécula de ácido nucleico u otros genes. La combinación resultante también puede incluir múltiples copias de una cualquiera de las moléculas de ácido nucleico de interés. Las moléculas de ácido nucleico dadas a conocer en la presente invención también pueden superponerse con cualquier otro gen o combinación de genes para producir una planta que tiene una combinación deseada de una pluralidad de rasgos. Los rasgos incluyen, pero no se limitan a, rasgos deseados como piensos para animales, tales como genes con alto contenido en aceite, aminoácidos balanceados, digestibilidad incrementada, resistencia a insectos, resistencia a enfermedades o

resistencia a herbicidas, genes no tóxicos y resistentes a enfermedades, rasgos agronómicos (por ejemplo, esterilidad masculina, tiempo de floración) y/o rasgos tecnológicos de transformación (por ejemplo, regulación del ciclo celular o selección de genes).

Cualquier método, incluyendo pero sin limitarse al cruzamiento de plantas mediante cualquier método convencional o conocido o transformación genética, puede usarse para reunir diferentes genes. Si los rasgos se apilan por plantas genéticamente transformadas, pueden combinarse las secuencias de polinucleótido de interés en cualquier orden en cualquier momento. Por ejemplo, las plantas transgénicas que comprenden uno o más rasgos deseados pueden usarse como dianas para introducir más rasgos a través de transformación posterior. En el esquema de co-transformación, pueden introducirse los rasgos simultáneamente con el polinucleótido de interés, que se proporciona por cualquier combinación de casetes de transformación. Por ejemplo, si se van a introducir dos secuencias, las dos secuencias pueden estar comprendidas en casetes de transformación independientes (trans) o en el mismo casete de transformación (cis). La expresión de las secuencias puede estar controlada por el mismo promotor o promotores diferentes. En algunos casos, es deseable introducir un casete de transformación que puede inhibir la expresión de una molécula de ácido nucleico de interés. Este puede combinarse con cualquier combinación de otros casetes de inhibición o casetes de sobreexpresión para generar una combinación de rasgos deseada en plantas.

El vector construido o casete de expresión no existe en el genoma de la planta inicial o el genoma de la planta transgénica, y no está ubicado en el locus nativo en el genoma de la planta inicial.

Las composiciones dadas a conocer en la presente invención pueden producirse o mantenerse por un método para la introgresión génica. La introgresión génica a veces se denomina "retrocruzamiento" cuando ese método se repite dos o más veces. En la introgresión génica o retrocruzamiento, los originales "donantes" se refieren a plantas originales con genes o loci requeridos que van a introgresarse. Los originales "receptores" (usados una vez o más) o los originales "recurrentes" (usados dos veces o más) se refieren a plantas progenitoras en las que se introgresan los genes o loci. La hibridación inicial produce generación F1. El término "BC1" se refiere al segundo uso de los progenitores recurrentes, y "BC2" se refiere al tercer uso de los progenitores recurrente, y así sucesivamente.

La presente invención también puede incluir las secuencias descritas que pueden proporcionarse a partir de un casete de expresión o constructo de ADN expresado en plantas de interés. El casete de expresión puede incluir secuencias reguladoras heterólogas 5' y 3' unidas operativamente a las secuencias dadas a conocer en la presente invención. El término "operativamente unido" se usa en la presente invención para significar que un ácido nucleico que va a expresarse se une a una secuencia reguladora, incluyendo un promotor, un terminador, un potenciador y/u otros elementos de control de expresión (por ejemplo, señales de poliadenilación) de manera que permite la expresión del ácido nucleico (es decir, cuando un vector se introduce en una célula vegetal huésped, él ácido nucleico se expresa en la célula vegetal huésped). Tales secuencias reguladoras se conocen bien en la técnica e incluyen aquellas secuencias de nucleótidos que pueden expresarse directamente constitutivamente en una variedad de células huésped y

expresarse directamente en células huésped específicas o en condiciones específicas. El diseño del vector puede depender de, por ejemplo, el tipo de células huésped que van a transformarse, o el nivel de expresión deseada de ácidos nucleicos. El casete de expresión puede comprender uno o más genes adicionales que van a co-transformarse en las plantas. Además, cualquier gen adicional puede proporcionarse en una pluralidad de casetes de expresión.

El casete de expresión de la presente invención puede comprender una pluralidad de sitios de digestión enzimática restrictivos para la inserción de la secuencia de nucleótidos para estar bajo la regulación transcripcional de una región reguladora. El casete de expresión también puede comprender genes marcadores selectivos.

El casete de expresión también puede comprender, en la dirección de transcripción 5'-3', una región de iniciación de la transcripción y traducción, una secuencia de ADN de la presente descripción, y una región de terminación de la transcripción y traducción que funciona en plantas. La región de iniciación de la transcripción, un promotor, puede ser nativa o similar o foránea o heteróloga en relación con los huéspedes de la planta. Además, el promotor puede ser una secuencia nativa o alternativamente una secuencia sintética. El término "heterólogo" significa que la región de transcripción inicial no existe en la planta nativa en la que se introduce la región de transcripción inicial región. Tal como se usa en el presente documento, un gen quimérico comprende una secuencia codificante unida operativamente a una región de iniciación de la transcripción que es heteróloga a la secuencia codificante. Los ejemplos de promotores incluyen, pero no se limitan a, virus del mosaico de la coliflor 35S y ubiquitina de soja 6.

Aunque los promotores heterólogos pueden usarse preferiblemente para expresar secuencias, pueden usarse promotores homólogos o secuencias promotoras nativas. Tales constructos alterarán el nivel de expresión en células huésped (es decir, plantas o células vegetales). Por tanto, se cambian los fenotipos de las células huésped (es decir, la planta o célula vegetal) .

La región de terminación puede tener naturalmente una región de iniciación de la transcripción, tener naturalmente una secuencia de ADN unida operativamente de interés, o originarse de otra fuente. Se puede obtener una región de terminación fácilmente disponible (tal como regiones de terminación de octopina sintasa y nopalina sintasa) del plásmido Ti de *Agrobacterium tumefaciens*.

El ortólogo endógeno o resistente al gen fuente puede alterarse mediante un método de recombinación homólogo o no homólogo, tal como, por ejemplo, mediante la edición del genoma. Cuando se compara con una secuencia no modificada, tal alteración significa que la secuencia de nucleótidos tiene al menos una modificación e incluye, por ejemplo: (i) reemplazo de al menos un nucleótido, (ii) delección de al menos un nucleótido, (iii) inserción de al menos un nucleótido, o (iv) cualquier combinación de (i) - (iii).

En algunas realizaciones, las técnicas de edición de genomas pueden usarse para introducir los genes de resistencia dados a conocer en la presente invención en el genoma de una planta, o técnicas de edición de genomas pueden usarse para editar genes de resistencia introducidos previamente en el genoma de una planta.

La edición del genoma puede implementarse usando cualquier método de edición de genes disponible. Por ejemplo, la edición de genes puede lograrse introduciendo un molde de

modificación de polinucleótidos (a veces denominado oligonucleótido de reparación de genes) en una célula huésped, en el que el molde de modificación de polinucleótidos comprende modificaciones dirigidas de genes dentro del genoma de la célula huésped. El molde de modificación de polinucleótido puede ser monocatenaria o bicatenaria.

Uno o más genes pueden optimizarse según se desee para aumentar la expresión en plantas transformadas. Por ejemplo, los codones preferidos por las plantas se usan para sintetizar genes para mejorar la expresión. Se conocen en la técnica los métodos para sintetizar un gen preferido por las plantas.

Se sabe que las modificaciones de secuencia adicional potencian la expresión génica en un huésped celular. Estas modificaciones de secuencia incluyen la eliminación de las siguientes secuencias: señales de pseudo-poliadenilación codificadas, señales de sitio de corte y empalme de exón-intrón, secuencias de repetición de tipo transposón y otras secuencias completamente caracterizadas que pueden ser dañinas para la expresión génica. El contenido en G-C en una secuencia puede ajustarse al nivel promedio de un huésped celular dado, cuyo nivel puede calcularse a partir de genes conocidos expresados en la célula huésped. La secuencia se puede modificar si es necesario para evitar una posible estructura secundaria de ARNm de horquilla.

Un casete de expresión puede comprender adicionalmente una secuencia líder 5' en un constructo del casete de expresión. Una secuencia líder de este tipo puede potenciar la traducción. Se conocen en la técnica secuencias líderes de traducción e incluyen: secuencia líder del virus de ácido ribonucleico pequeño, tal como secuencia líder EMCV (región

no codificante de encefalomiocarditis 5`); secuencia líder del grupo del virus de la patata, tal como secuencia líder TEV ((virus de ataque químico del tabaco), y proteína de unión a la cadena pesada de la inmunoglobulina humana (BiP); secuencia líder no traducida del ARNm de proteína de cubierta (AMVRNA 4) del virus del mosaico de alfalfa; secuencia líder del virus del mosaico del tabaco (TMV); así como secuencia líder del virus del moteado clorótico del maíz (MCMV) (Lommel et al., (1991) Virology 81: 382-385). También pueden utilizarse otros métodos conocidos para potenciar la traducción, tales como intrones.

Pueden manipularse diversos fragmentos de ADN en un casete de expresión en marcos de lectura apropiados según las necesidades para garantizar que las secuencias de ADN estén en la dirección correcta. Para este fin, pueden usarse adaptadores o moléculas unión para unir fragmentos de ADN. Además, también pueden usarse otras manipulaciones para proporcionar sitios de restricción convenientes, eliminar el exceso de ADN o eliminar sitios de restricción. Para este fin, la mutagénesis *in vitro*, la reparación del cebador, la restricción, el apareamiento y el re-reemplazo (por ejemplo, conversión y transversión) pueden estar implicados.

En general, un casete de expresión puede comprender genes marcadores selectivos para seleccionar células transformadas. Los genes marcadores selectivos se usan para seleccionar tejidos o células transformadas. Los genes marcadores incluyen genes que codifican para resistencia a antibióticos, tales como genes que codifican para neomicina fosfotransferasa II (NEO) e higromicina fosfotransferasa (HPT), así como genes que confieren resistencia frente a compuestos herbicidas como glufosinato, fosfinotricina, bromoxinil, imidazolinona y ácido 2,4-diclorofenoxiacético

(2,4-D). La lista anterior de genes marcadores selectivos no pretende ser limitativa. Puede usarse cualquier gen marcador selectivo en la presente descripción.

Con el fin de expresar la proteína y/o gen diana de la presente invención en una planta o célula vegetal, el método de la presente invención comprende la transformación de la planta o célula vegetal con una molécula de ácido nucleico que codifica para la proteína diana. La molécula de ácido nucleico de la presente invención puede unirse operativamente a un promotor que acciona la expresión en una célula vegetal. Cualquier promotor conocido en la técnica puede usarse en el método de la presente invención, incluyendo, pero sin limitarse a, promotores constitutivos, promotores inducibles por patógenos, promotores inducibles por heridas, promotores preferidos de tejido y promotores regulados químicamente. La selección del promotor puede depender de la ubicación y el tiempo de expresión deseado en una planta transformada, así como otros factores conocidos por los expertos en la técnica. Las células o plantas transformadas pueden plantarse o cultivarse para formar una planta que comprende uno o más polinucleótidos introducidos, por ejemplo, en células o plantas que codifican para proteínas R.

Puede usarse una variedad de promotores para colocar en práctica la presente invención. Los promotores pueden seleccionarse según el resultado deseado. Es decir, puede combinarse un ácido nucleico con un promotor constitutivo, un promotor preferido de tejido, u otros promotores y expresarse en una célula huésped de interés. Tales promotores constitutivos incluyen, por ejemplo, el promotor central del promotor Rsyn7 y otros promotores constitutivos dados a conocer en el documento WO 99/43838 y la patente estadounidense No. 6,072,050, promotor CaMV 35S, actina de

arroz, ubiquitina, pEMU, MAS, ALS, etc. Tales promotores constitutivos se conocen en la técnica y se contemplan para su uso en la presente descripción.

En general, el uso de promotores inducibles, especialmente los de patógenos, para expresar el gen de la presente invención es beneficioso para la aplicación del gen de la presente invención. Tales promotores incluyen los promotores de proteínas relacionadas con la patogénesis (proteínas PR), cuyas proteínas se inducen para formar por la infección por patógenos y son, por ejemplo, proteínas PR, proteínas SAR, β -1,3-glucanasa, quitosanasas, etc.

Los promotores expresados localmente en o cerca del sitio de infección por patógenos merecen atención. Además, debido a que los patógenos pueden entrar en las plantas a través de heridas o lesiones provocadas por insectos, los promotores inducibles por heridas también pueden usarse para la construcción del vector de la presente invención. Tales promotores inducibles por heridas incluyen el gen inhibidor de proteasa de patata (pinII), wun1 y wun2, win1 y win2, sistemina, WIP1, gen MPI, etc.

Los promotores regulados químicamente pueden regular la expresión génica en plantas mediante la aplicación de agentes reguladores químicos exógenos. De acuerdo con el fin de la utilización, los promotores pueden ser promotores químicamente inducibles, tal como inducción de la expresión génica mediante la aplicación de productos químicos, o promotores químicamente reprimibles, tal como inhibición de la expresión génica mediante la aplicación de productos químicos. Los promotores inducibles con un agente químico son conocidos en la técnica e incluyen, sin carácter limitativo, el promotor In2-2 del maíz, (el cual se activa por acción de protectores contra herbicidas de tipo bencenosulfonamida), el

promotor GST del maíz, (el cual se activa por acción de compuestos electrófilos hidrófobos usados como herbicidas de preemergencia), y el promotor PR-1a del tabaco (el cual se activa por acción del ácido salicílico). Otros promotores regulados con un agente químico de interés incluyen promotores sensibles a los esteroides (por ejemplo, promotores inducibles por glucocorticoides, promotores inducibles por tetraciclina y promotores represibles por tetraciclina).

Los promotores preferidos de tejido pueden usarse para expresión dirigida mejorada de genes o proteínas diana (por ejemplo, secuencias polinucleótidas que codifican para polipéptidos NB-LRR derivados de plantas leguminosas) en tejidos vegetales específicos. Los promotores preferidos para tales tejidos incluyen, pero no se limitan a, promotores específicos de hoja, promotores específicos de raíz, promotores específicos de semilla, y promotores específicos de tallo. Los promotores específicos del tejido incluyen Yamamoto et al., (1997) *Plant J.* 12 (2): 255-265; Kawamata et al., (1997) *Plant Cell Physiol.* 38 (7): 792-803; Hansen et al., (1997) *Mol. Gen Genet.* 254 (3): 337-343; Russell et al., (1997) *Transgenic Res.* 6 (2): 157-168; Rinehart et al., (1996) *Plant Physiol.* 112 (3): 1331-1341; Van Camp et al., (1996) *Plant Physiol.* 112 (2): 525-535; Canevascini et al., (1996) *Plant Physiol.* 112 (2): 513-524; Yamamoto et al., (1994) *Plant Cell Physiol.* 35 (5): 773-778; Lam, (1994) *Results Probl. Cell Differ.* 20: 181-196; Orozco et al., (1993) *Plant Mol Biol.* 23 (6): 1129-1138; Matsuoka et al., (1993) *Proc Natl. Acad. Sci. USA*, 90 (20): 9586-9590; y Guevara-Garcia et al., (1993) *Plant J.* 4 (3): 495-505. Tales promotores pueden usarse para modificar las secuencias de nucleótidos de la presente descripción.

En la técnica se conocen promotores específicos para las hojas. Véase, por ejemplo, Yamamoto *et al.*, (1997) *Plant J.* 12 (2): 255-265; Kwon *et al.*, (1994) *Plant Physiol.* 105: 357-67; Yamamoto *et al.*, (1994) *Plant Cell Physiol.* 35 (5): 773-778; Gotor *et al.*, (1993) *Plant J.* 3: 509-18; Orozco *et al.*, (1993) *Plant Mol. Biol.* 23 (6): 1129-1138; y Matsuoka *et al.*, (1993) *Proc. Natl. Acad. Sci. USA*, 90 (20): 9586-9590.

Los promotores "preferidos por semillas" incluyen promotores "específicos de semillas" (aquellos que están activos durante el desarrollo de semillas, tales como los promotores de la proteína de almacenamiento de semillas) y promotores de "germinación de semillas" (aquellos que están activos durante la germinación de semillas). Tales promotores preferidos de semillas incluyen, pero no se limitan a, Cim1 (citoquinina inducción información), milps (inositol-1-fosfato sintasa) y celA (celulosa sintasa). Globin-1 (Glob-1) es el promotor específico de embrión preferido. Para el dicotiledóneo, los promotores específicos de semillas incluyen, pero no se limitan a, promotor común del gen de la faseolina del frijol, promotor del gen de la napina, promotor del gen de la β -conglucina, promotor del gen de la lectina de soja, promotor del gen de la proteína crucífera, etc.

La expresión de las moléculas de ácido nucleico de la presente invención puede implicar el uso de genes de resistencia nativos completos, en la que la expresión se acciona por una secuencia promotora hacia 5'homóloga u otros promotores heterólogos. Los expertos en la técnica podrán identificar los genes de resistencia para evaluar el nivel de expresión de los mismos y seleccionar una secuencia promotora preferida que pueda usarse para expresar los genes de resistencia de interés. El uso de secuencias promotoras de genes de resistencia homólogas o heterólogas proporciona

opciones para regular la expresión de proteínas para evitar o minimizar cualquier resultado potencialmente inapropiado o indeseable relacionado con la activación de defensa de la planta.

Los promotores de soja específicos incluyen, pero no se limitan a, promotores de ubiquitina de soja (subi-1), factor de elongación 1A, S-adenosilmetionina sintasa para expresión constitutiva, Rpp4, y RPG1-B, y promotores comprendidos en modelos génicos, tales como promotores Glyma conocidos por los expertos en la técnica para regular más estrechamente la expresión proporcionada por los promotores del gen NB-LRR.

En el contexto de la presente invención, el germoplasma incluye células, semillas o tejidos a partir de los cuales pueden generarse nuevas plantas, o partes de plantas tales como hojas, tallos, polen, o células que pueden cultivarse en plantas completas.

Plantas transformadoras

Dado que el gen de resistencia descrito en la presente invención puede expresarse como un transgén para producir una planta resistente a la roya, determinada realización de la invención se refiere a métodos para introducir una molécula de ácido nucleico en una planta. Además de las plantas de soja transformantes descritas anteriormente, los ácidos nucleicos de la invención pueden usarse también en otras plantas.

Tal como se usa en el presente documento, el término "introducción" y variaciones gramaticales del mismo se refieren a proporcionar una planta con una molécula de ácido nucleico. En algunas realizaciones, la molécula de ácido nucleico puede existir de tal manera que la secuencia entra en el interior de una célula vegetal, incluyendo su posible inserción en el genoma de la planta. El método dado a conocer

en la presente invención no depende de métodos específicos para introducir una secuencia en una planta, siempre que una molécula de ácido nucleico entra en el interior de al menos una célula de la planta. Se conocen en la técnica métodos para introducir una molécula de ácido nucleico en una planta e incluyen, pero no se limitan a métodos de transformación estables, métodos de transformación transitoria y métodos mediados por virus.

Los ejemplos de métodos para la transformación de plantas incluyen la transformación mediada por *Agrobacterium* y el bombardeo de partículas, luego la planta transformada puede regenerarse mediante métodos conocidos por los expertos en la técnica.

Una molécula de ácido nucleico puede introducirse de manera transitoria o estable en una célula huésped y puede permanecer sin integrar, por ejemplo, en forma de plásmido. "Transformación estable" o la "transformado de manera estable" y las variaciones gramaticales de los mismos significan que un constructo de nucleótidos introducido en una planta se integra en el genoma de la planta y puede ser heredada por las progenies de la planta. Tal como se usa en el presente documento, "transformación transitoria" y variaciones gramaticales de la misma significan que una molécula de ácido nucleico se introduce en una planta pero no se integra en el genoma de la planta, o una proteína se introduce en una planta.

Los métodos de transformación y los métodos para introducir una secuencia molecular de ácido nucleico en una planta pueden depender del tipo de planta o célula vegetal que va a transformarse. Los métodos adecuados para introducir una proteína o una molécula de ácido nucleico en una célula vegetal incluyen, pero no se limitan a, microinyección,

electroporación, transferencia de genes directa, transformación de *Lec1* y aceleración de partículas balísticas. Como los métodos actualizados están disponibles, tales métodos también pueden usarse en la presente invención debido a que los métodos de transformación o transfección no son críticos.

Las células transformadas pueden cultivarse en plantas según los métodos convencionales. Luego estas plantas pueden hacerse crecer y polinizarse con la misma línea de transformación o diferentes líneas, y luego se pueden identificar las progenies con expresión constitutiva con las características fenotípicas requeridas. Se pueden cultivar dos o más generaciones de plantas para asegurar que la expresión de las características fenotípicas requeridas se mantenga de manera estable y heredada. Se recogen las semillas para asegurar que se ha logrado la expresión de las características fenotípicas requeridas. En algunas realizaciones, las semillas transformadas o semillas transgénicas tienen constructos de nucleótidos o casetes de expresión incorporados de manera estable en sus genomas.

La presente invención abarca semillas que comprenden las secuencias moleculares de ácido nucleico dadas a conocer en la presente invención, las semillas pueden desarrollarse en o usarse para desarrollar una o más plantas que tienen resistencia potenciada frente a patógeno (por ejemplo, hongos) o infecciones formadas por patógenos en comparación con, por ejemplo, las semillas de plantas de variedades de tipo natural. La presente invención se caracteriza por semillas de plantas leguminosas transgénicas, en las que las semillas comprenden las moléculas de ácido nucleico dadas a conocer en la presente invención.

Las plantas de interés incluyen especies de cultivos leguminosos, incluyendo pero sin limitarse a alfalfa (*Medicago sativa*); trébol o trifolio (*Trifolium* spp.); guisante, incluyendo *Pisum sativum*, *Gajanus cajan*, *Vigna unguiculata* y *Lathyrus* spp.; frijol común (*Fabaceae* o *Leguminosae*); lenteja (*Lens culinaris*); lupinus (*Lupinus* spp.); árbol de Ghaf (*Prosopis* spp.); frijol largo (*Ceratonia siliqua*), soja (*Glycine max*), maní (*Arachis hypogaea*) o tamarindo (*Tamarindus indica*). Los términos "especies leguminosas" y "especies de cultivos leguminosos" se usan en el presente documento para referirse a plantas y pueden, por ejemplo, ser plantas de interés. Las especies leguminosas o especies de cultivos leguminosos pueden ser plantas, partes de plantas o células vegetales.

Regulación y mejora de la resistencia a la roya

La presente invención también proporciona para el uso de las proteínas, vectores recombinantes de ácidos nucleicos, bacteria recombinante o línea celular transgénica descritos en el presente documento para regular la resistencia de las plantas frente a la roya. En el uso, el nivel de expresión y/o la actividad de la proteína o el gen codificante de la misma en las plantas aumenta, y la resistencia de las plantas frente a la roya se potencia.

Además, la invención se refiere a un método para potenciar la resistencia de plantas frente a enfermedades de plantas tales como la roya. El método puede incluir conferir resistencia contra patógenos (por ejemplo, roya) al introgesar genes de resistencia de plantas leguminosas en germoplasma en un procedimiento de fitomejoramiento (es decir, procedimiento de fitomejoramiento para resistencia a la roya).

Los métodos de introgresión tal como se da a conocer en el presente documento también pueden usarse para producir plantas de soja que tengan una resistencia aumentada a una cualquiera de las siguientes: nematodo del quiste de soja, pústula bacteriana, nematodo del nudo de la raíz, mancha de la hoja del ojo de rana, fitoptora, pudrición del tallo marrón, nematodo, tizón, *Golovinomyces cichoracearum*, *Erysiphe cichoracearum*, *Blumeria graminis*, *Podosphaera xanthii*, *Sphaerotheca fuliginea*, *Pythium ultimum*, *Uncinula necator*, *Mycosphaerella pinodes*, *Magnapotha grisea*, *Bipolaris oryzae*, *Magnapotha grisea*, *Rhizoctonia solani*, *Phytophthora sojae*, *Schizaphis graminum*, *Bemisia tabaci*, *Rhopalosiphum maidis*, *Deroceras reticulatum*, *Diatraea saccharalis*, *Schizaphis graminum*, o *Myzus persicae*.

En una realización, el método mejora la resistencia frente a la roya. El método para mejorar la resistencia de las plantas frente a la roya puede incluir el aumento del nivel de expresión y/o la actividad de las proteínas mencionadas anteriormente en las plantas, donde la mejora de la resistencia de plantas frente a la roya no solo puede conferir resistencia a la roya frente a las plantas que no están originalmente resistente a la roya, sino también puede potenciar adicionalmente la resistencia a la roya de las plantas que originalmente son resistentes a la roya. El aumento del nivel de expresión y/o la actividad de las proteínas en las plantas no solo puede hacer que las plantas que originalmente no expresan las proteínas expresen las proteínas, sino que también puede aumentar adicionalmente el nivel de expresión y/o la actividad de las proteínas en las plantas que originalmente expresa las proteínas. Además, en determinadas realizaciones, el aumento del nivel de expresión y/o la actividad de las proteínas en las plantas puede

realizarse o bien por medios transgénicos o por hibridación sexual.

Método para fitomejorar una variedad de plantas

En otra realización, la presente invención proporciona un método para fitomejorar una variedad de plantas con resistencia mejorada frente a la roya.

El método para fitomejorar una variedad de plantas con resistencia aumentada frente a la roya puede incluir la etapa de aumentar el nivel de expresión y/o la actividad de las proteínas mencionadas anteriormente en una planta receptora, donde la mejora de la resistencia frente a la roya no solo puede conferir resistencia a la roya frente a las plantas que no están originalmente resistente a la roya, sino también puede potenciar adicionalmente la resistencia a la roya de las plantas que originalmente son resistentes a la roya. El aumento del nivel de expresión y/o la actividad de las proteínas mencionadas anteriormente en la planta receptora no solo puede hacer que las plantas que originalmente no expresan las proteínas expresen las proteínas, sino que también puede aumentar adicionalmente el nivel de expresión y/o la actividad de las proteínas en las plantas que originalmente expresa las proteínas.

Además, el aumento del nivel de expresión y/o la actividad de las proteínas en la planta puede realizarse o bien por medios transgénicos o por hibridación sexual.

Métodos para el fitomejoramiento

En otra realización, la invención proporciona un método para fitomejorar una planta transgénica con resistencia mejorada frente a la roya.

El método para fitomejorar una planta transgénica con resistencia mejorada frente a la roya puede incluir la introducción de una molécula de ácido nucleico tal como se

describe en el presente documento a una planta receptora para obtener una planta transgénica que tiene resistencia mejorada frente a la roya en comparación con la planta receptora, donde la mejora de la resistencia frente a la roya no solo puede conferir resistencia a la roya frente a las plantas que no están originalmente resistente a la roya, sino también puede potenciar adicionalmente la resistencia a la roya de las plantas que originalmente son resistentes a la roya.

Además, la introducción de la molécula de ácido nucleico a la planta receptora puede realizarse introduciendo el casete de expresión o el vector recombinante descrito anteriormente en la planta receptora.

En el método anterior, la introducción del casete de expresión o el vector recombinante en la planta receptora puede ser particularmente: transformando tejidos o células vegetales usando métodos biológicos convencionales tales como plásmido Ti, plásmido Ri, vector de virus vegetal, transformación directa de ADN, microinyección, electroporación, transformación mediada por *Agrobacterium*, etc., y cultivando los tejidos vegetales transformados en plantas.

La invención también se refiere a plantas transgénicas obtenidas por los métodos de fitomejoramiento. Por consiguiente, en una realización, la invención proporciona una planta transgénica con resistencia mejorada frente a la roya obtenida mediante fitomejoramiento usando el método descrito en el presente documento, o es soja (*Glycine max*) SX6907, o una planta de progenie que comprende la molécula de ácido nucleico descrita en el segundo aspecto anterior obtenida después de la hibridación sexual usando la soja (*Glycine max*) SX6907 como progenitor. En una realización, puede usarse un derivado de la soja (*Glycine max*) SX6907 que

tiene el número de registro CGMCC No. 17575 en el Centro de Colección de Cultivo Microbiológico General de China.

La planta en el presente documento puede ser una planta completa, o puede ser una célula vegetal, semilla o tejido, o una parte de la planta tal como una hoja, tallo, polen o célula que se puede cultivar en una planta completa.

Ensayos y kits

En determinadas realizaciones, las moléculas de aminoácidos o ácidos nucleicos de la invención pueden usarse para evaluar las plantas para la resistencia a la roya. Por consiguiente, la invención también comprende un kit para el ensayo descrito en el presente documento. Las proteínas o moléculas de ácido nucleico o casetes de expresión, vectores recombinantes o bacterias recombinantes o líneas celulares transgénicas que comprenden las moléculas de ácido nucleico también pueden empaquetarse junto con las instrucciones como componentes del kit para completar el ensayo dado a conocer en la presente invención. El kit de la presente invención puede incluir cualquier combinación de las proteínas o moléculas de ácido nucleico de la presente invención o casetes de expresión, vectores recombinantes o bacterias recombinantes o líneas celulares transgénicas que comprenden las moléculas de ácido nucleico y instrucciones adecuadas (escritas y/o proporcionadas como materiales de audio, visual o audiovisual). Por ejemplo, el kit también puede comprender una sonda específica que tiene una secuencia correspondiente a o complementaria a una secuencia que tiene del 80% al 100% de identidad de secuencia con una región específica del acontecimiento transgénico. El kit puede comprender cualquier reactivo y material requerido para realizar el ensayo o método de detección.

La realización de la invención también incluye cualquiera de los siguientes materiales biológicos o aplicaciones:

(D1) un par de cebadores para amplificar la molécula de ácido nucleico descrita anterior; En determinadas realizaciones, el par de cebadores es un par de cebadores 1 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 3 y SEQ ID NO: 4, respectivamente, o un par de cebadores 2 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 5 y SEQ ID NO: 6, respectivamente.

(D2) una sonda para amplificar la molécula de ácido nucleico descrita anterior;

(D3) un kit que comprende el par de cebadores y/o la sonda.

(D4) una planta que comprende la molécula de ácido nucleico descrita anteriormente.

(D5) uso del par de cebadores o la sonda o el kit en la identificación de si la planta que va a someterse a prueba comprende la molécula de ácido nucleico descrita anteriormente.

(D6) uso del par de cebadores o la sonda o el kit en la identificación de si la planta que va a someterse a prueba tiene resistencia frente a la roya conferida por la molécula de ácido nucleico descrita anteriormente.

(D7) una molécula de ADN, que es una molécula de ADN mostrada en SEQ ID NO: 7, o una molécula de ADN con función promotora obtenida por la adición, delección y/o sustitución de uno o más nucleótidos de SEQ ID NO:7.

(D8) uso de la molécula de ADN en el inicio de la expresión de un gen de interés. En determinadas realizaciones, el gen de interés es la molécula de ácido nucleico descrita anteriormente.

Información de registro para el derivado de soja
(*Glycine max*) SX6907 útil en la invención reivindicada:

Nombre de la categoría:	Soja
Nombre científico:	<i>Glycine max</i>
Depositario:	Centro de Colección de Cultivo Microbiológico General de China
Abreviatura de depósito:	CGMCC
Dirección:	No. 3, Yard 1, Beichen West Road, Chaoyang District, Beijing
Fecha de registro:	Viernes, 28 de junio de 2019
Número de registro del centro de depósito:	CGMCC No. 17575

Ejemplos

Los siguientes ejemplos facilitan un mejor entendimiento de la invención, pero no limitan la invención. Los métodos experimentales en los siguientes ejemplos son métodos convencionales. Los materiales de prueba usados en los siguientes ejemplos, a menos que se especifique lo contrario, se adquirieron de almacenes de reactivos bioquímicos generales. Para las pruebas cuantitativas en los siguientes ejemplos, se establecieron tres pruebas repetidas y se promediaron los resultados.

Ejemplo 1: Clonación del gen RppRC1

La soja es una antigua planta leguminosa tetraploide con autopolinización y tiene un tamaño del genoma de aproximadamente 1.1 Gbp. La respuesta del germoplasma de soja con resistencia diferente a *Phakopsora pachyrhizi* es obviamente diferente. La soja (*Glycine max*) SX6907 es un

recurso de resistencia a la roya seleccionado del germoplasma de soja china por el Instituto de Investigación de Cultivos de Aceite, Academia China de Ciencias Agrícolas. La variedad se conserva actualmente en el Centro de Colección de Cultivo Microbiológico General de China bajo el número de registro CGMCC No. 17575, y la respuesta de la variedad a *Phakopsora pachyrhizi* es la inmunidad.

Se usó el genoma de soja SX6907 (la fuente del gen *RppRC1*) como molde para la amplificación por PCR usando el cebador F y el cebador R. Las secuencias de cebadores son tal como sigue:

5'-ATGGCAGATAGTGTGTTGCTTTTCTGC-3' (SEQ ID NO: 3);

y

5'-TCACAGTTCATTAGAGATTTTGAGCTTACAGC-3' (SEQ ID NO:

4).

El producto de amplificación obtenido es la secuencia mostrada en SEQ ID NO: 2. SEQ ID NO: 2 es la secuencia de nucleótidos del gen *RppRC1*, que codifica para la proteína mostrada en SEQ ID NO 1, cuya proteína se denomina *RppRC1*.

Ejemplo 2: Transformación de soja con el gen *RppRC1* de la soja

1. Construcción del Vector de Expresión Recombinante

Llevando las hojas de la variedad de soja SX6907 cultivadas en condiciones normales a la fase de crecimiento de la primera hoja compuesta ternada , extrayendo ADN, usando el ADN como molde y amplificando por un método de PCR convencional bajo la guía del cebador F y cebador R (véase el ejemplo 1 para secuencias de cebadores), realizando una detección de electroforesis en gel de agarosa al 1% en el producto amplificado por PCR después de completar la reacción, recuperando y purificando aproximadamente 2.5 kb de fragmento de ADN. Al mismo tiempo, usando el ADN como molde y

amplificando por un método de PCR convencional en la guía de PF y PR para obtener el fragmento pRppRC1 del promotor de 2525 pb, y uniendo el fragmento al vector pB2GW7 (el vector proviene de biología de sistemas de plantas VIB, sitio web: <https://gateway.psb.ugent.be/search>) mediante una digestión enzimática y un método de enlace para obtener el vector recombinante pB2GW7-pRppRC1.

PF: 5'-GAGCTCAAAGGCTTTTTTGTTAAGGGAAGGT-3' (la parte subrayada es la secuencia de reconocimiento de SacI);

PR: 5'- ACTAGTTTCTGTGAAACAGGAAATCTTGGGT-3' (la parte subrayada es la secuencia de reconocimiento de SpeI).

Luego, el gen *RppRC1* (SEQ ID NO: 2) se clonó en el vector recombinante pB2GW7-pRppRC1 obtenido mediante el método de Gateway para obtener el vector recombinante pB2GW7-*RppRC1*. La secuencia confirmó que el vector recombinante pB2GW7-*RppRC1* es un plásmido recombinante obtenido insertando el fragmento de ADN de 2574 pb que se muestra en la SEQ ID NO: 2 entre los sitios attR1 y attR2 del vector pB2GW7, y reemplazando el promotor 35S entre los sitios de digestión de enzimas SacI y SpeI con el promotor endógeno del gen *RppRC1* mostrado en la SEQ ID NO: 7.

En el vector de expresión recombinante final pB2GW7-*RppRC1*, el promotor que inicia la transcripción del gen *RppRC1* es el promotor original pRppRC1. Además de los elementos anteriores, el vector también comprende un gen de resistencia a espectinomicina para la selección bacteriana y un gen Bar de soja resistente a los herbicidas como marcador selectivo de plantas (figura 1).

2. Transformación de la soja

Se transfirió el vector de expresión recombinante pB2GW7-*RppRC1* que comprende el gen *RppRC1* construido en la etapa 1 en *Agrobacterium tumefaciens* EHA105 mediante el

método de congelar-descongelar. Luego se seleccionaron transformantes positivos para la transformación genética de soja.

Se usó transformación de nodo de cotiledón mediada por *Agrobacterium* (Paz, M.M., Martinez, J.C., Kalvig, A.B., et al., (2006) *Plant Cell Report*, 25, 206-213) para la transformación de la soja, y el receptor de transformación fue la variedad de soja Tianlong No.1 (se fitomejóro la variedad por el Instituto de Investigación de Cultivos de Aceite, Academia China de Ciencias Agrícolas. (examen nacional número 2008023) y está disponible a partir de ahí). Los medios usados para la transformación y regeneración de la planta de la soja eran tal como sigue:

Medio de cultivo de *Agrobacterium*: levadura en polvo 10 g/L, polvo de proteína pancreática 20 g/L, agar 20 g/L, y rifampicina 50 mg/L y espectinomicina 100 mg/L, como antibióticos. Tras la esterilización, verter la mezcla de los reactivos en una placa de Petri con 9 cm de diámetro para su uso.

Medio líquido de *Agrobacterium*: levadura en polvo 10 g/L, polvo de proteína pancreática 20 g/L, y rifampicina 50 mg/L y espectinomicina 100 mg/L, como antibióticos. Esterilización para uso posterior.

Medio de co-cultivo: Se añadieron medio B5 0.32 g/L, sacarosa 30 g/L, ácido etanosulfónico, (2-(N-Morfolino) ácido etanosulfónico (MES)) 6.0 g/L, pH ajustado hasta 5.4; tras la esterilización, 6-bencilaminopurina (6-BA) 1.67 mg/L, L-cisteína 400 mg/L, DL-Ditiotreitol (DTT) 150 mg/L y acetosiringona 200 µg/L.

Medio de inducción de brote de regeneración: Se añadieron medio MS 4.4 g/L, sacarosa 30 g/L, MES 0.6 g/L, agar 8 g/L, pH ajustado hasta 5.8; tras la esterilización,

6-BA 1.67 mg/L, cefotoxina (Cef) 200 mg/L, vancomicina (Van) 50 mg/L, timentim (Tim) 100 mg/L y glufosinato 8 mg/L. Verter la mezcla de los reactivos en una placa de Petri con 9 cm de diámetro para su uso.

Medio de elongación de brote de regeneración: Se añadieron medio MS 4.4 g/L, sacarosa 30 g/L, MES 0.6 g/L, agar 8 g/L, pH ajustado hasta 5.8; tras la esterilización, ácido giberelina (GA₃) 0.5 mg/L, Cef 200 mg/L, Van 50mg/L, Tim 100 mg/L y glufosinato 8 mg/L. Verter la mezcla de los reactivos en una placa de Petri con 9 cm de diámetro para su uso.

Disolución de enraizamiento: Se disolvieron 30 mg de ácido indolebutírico (IBA), en 10 ml de agua clara; se almacenaron a 4°C y se diluyeron 1000 veces cuando se usa.

El procedimiento de transformación fue el siguiente:

Desinfección de semillas: tomar semillas limpias, esparcir las semillas 1-2 capas en una placa de Petri de 9 cm, colocar la placa de Petri en una secadora de 300 ml, colocar un vaso de precipitados en la secadora, añadir 10 ml de disolución de hipoclorito de sodio al 3% y luego 10 ml de disolución de ácido clorhídrico al 15% en el vaso de precipitados, cubrir la cubierta de la secadora para sellar durante 16-20 horas, luego sacar la placa de Petri cargada con semillas, colocar la placa de Petri en un banco súper limpio durante aproximadamente 30 minutos y se retira el cloro mediante soplado, luego añadir aproximadamente 40 ml de agua estéril en la placa de Petri, con transformación tras 8-12 horas. Se realizaron esta etapa y las siguientes operaciones en condiciones asépticas a menos que se enfatice lo contrario.

Preparación de Agrobacterium: Se tomó Agrobacterium monoclonal (transferido a pB2GW7-RppRC1), se rayó en una

placa de Petri con medio de cultivo de Agrobacterium que contenía los antibióticos correspondientes, se añadieron 3 ml de medio de cultivo líquido de Agrobacterium, se giró ligeramente la placa de Petri para permitir que el medio de cultivo líquido de Agrobacterium cubriera la placa de Petri, y se incubó durante la noche a 28°C. Al día siguiente, se lavó el césped con el medio de co-cultivo para preparar una suspensión bacteriana con un valor de DO de aproximadamente 0.8-1.2 para su uso posterior.

Preparación de explantes: Se cortó verticalmente el hipocótilo de una semilla, y se separaron uniformemente dos cotiledones a lo largo de la línea media del hipocótilo. Se rascó la articulación del cotiledón y el hipocótilo, y se retiró la hoja verdadera del nódulo cotiledón. Cada semilla puede hacerse en dos explantes.

Infección y co-cultivo: Se colocaron los explantes en la suspensión bacteriana para garantizar que todos los explantes estaban sumergidos en el medio de co-cultivo. Tras 20-40 min, se retiró el líquido bacteriano con una pipeta. Se colocaron dos piezas de papel de filtro redondo estéril en la placa de co-cultivo (15 cm de diámetro), el diámetro del papel de filtro (aproximadamente 13-14 cm) era ligeramente menor que el diámetro de la placa, y se añadieron 10 ml del medio de co-cultivo a cada placa. Se esparcieron los explantes infectados en el papel de filtro con la incisión hacia arriba. Se selló la placa y se incubó a 22°C en un fotoperiodo de 18 h durante 5 días.

Inducción de brote de regeneración: Tras el co-cultivo, se cortaron los explantes de los hipocótilos elongados y se colocaron en el medio de inducción de brotes regenerados. Se sumergió la región de hipocótilo en el medio, con 6-7 explantes por placa (9 cm de diámetro). Se incubó la placa a

24°C a un fotoperiodo de 18 h. Dos semanas después, se cortaron los callos cultivados a partir de los hipocótilos y se transfirieron a un medio de inducción de brotes regenerados frescos para un cultivo adicional durante dos semanas.

Elongación de brote de regeneración: Se retiraron los cotiledones de los explantes, y se cortaron los nuevos callos crecidos de los hipocótilos. Se transfirieron los explantes resultantes a un medio de elongación de brote de regeneración y se subcultivaron cada dos meses. Se eliminaron los nuevos callos crecidos a partir de los hipocótilos en cada transferencia. Cuando los brotes se elongaron a más de 3 cm, se cortaron los brotes elongados (> 3 cm) de los explantes, y los explantes restantes continuaron siendo cultivados en el medio de elongación del brote de regeneración.

Enraizamiento del brote de regeneración (esta etapa puede operarse en condiciones abiertas): tomar una placa vacía (15 cm de diámetro), colocar una pieza de papel filtro que tiene un diámetro ligeramente menor que el diámetro de la placa en la placa, añadir agua a remoje bien el papel, sumergir el extremo del brote elongado en 3 mg/L de disolución de IBA durante 10-20 segundos, sacar el brote y luego envolver el extremo del brote con una pieza de papel absorbente, esparcir el brote en el papel de filtro remojado, cubrir la placa, y cultivar el brote a 24°C en un fotoperiodo de 18 h. Se abrió la tapa todos los días para ventilación, y se complementó el agua adecuadamente para mantener el papel de filtro húmedo. Cuando las nuevas raíces crecieron hasta 2-3 cm, se transfirió el cultivo al suelo y se cultivó en el invernadero hasta la fructificación.

Al mismo tiempo, se estableció un control (control de vector vacío) en este experimento al introducir el vector vacío pB2GW7 en la variedad de soja Tianlong No.1.

3. Identificación de soja transgénica

A. Identificación de PCR

Se seleccionaron al azar plantas T₁ parciales de 2 T₀ plantas transgénicas de soja Tpp *RppRC1* (denominadas acontecimiento de transformación L1 y acontecimiento de transformación L2) (6 plantas, respectivamente registradas como L1-1, L1-2, L1-3, L2-1, L2-2 y L2-3; L1-1, L1-2 y L1-3 eran plantas individuales T₁ del acontecimiento de transformación L1, y L2-1, L2-2 y L2-3 eran plantas individuales T₁ del acontecimiento de transformación L2), y se seleccionaron una planta no transgénica Tianlong No. 1 (negativa) y una planta SX6907 (positiva) ; se extrajo el ADN genómico de estas plantas, respectivamente, y se amplificó el gen diana *RppRC1* mediante PCR con el cebador F y cebador R del ejemplo 1 y el tamaño del producto diana era 2574 pb; se sometió el producto amplificado a electroforesis en gel de agarosa al 1%; y se registraron las plantas de las cuales se obtuvo una banda de 2574 pb como positivas.

Resultados: Las plantas L1-1, L1-2, L1-3, L2-1, L2-2 y L2-3 eran todas positivas, las plantas no transgénicas eran negativas, y las plantas SX6907 eran positivas, tal como se muestra en la figura 2.

B. Identificación de RT-PCR

Se tomaron la soja transgénica T₁ *RppRC1* obtenida en la etapa 2 (es decir L1-2 y L2-1 en la etapa 1), la línea de soja transformada con vector vacío (CK) y la planta no transgénica Tianlong No.1 y se sometieron a una extracción de ARN total, respectivamente; el ARN resultante se transcribió inversamente para obtener ADNc, se usó el ADNc resultante

como molde para realizar amplificación por PCR cuantitativa de fluorescencia en tiempo real en el ADNc del gen *RppRC1* con cebadores específicos *F1* y *R1*, en el que se usó β -actina de soja como una referencia interna que se amplificó con los cebadores *FC* y *RC*. Se ejecutó la PCR cuantitativa de fluorescencia en tiempo real en un instrumento de PCR cuantitativa de fluorescencia en tiempo real CFX Connect™ con 3 réplicas en una prueba paralela. El nivel de expresión relativa se calculó utilizando el método informado por Livak KJ y Schmittgen TD (2001), es decir, $2^{-\Delta\Delta C_T}$.

$$\Delta\Delta C_T = (C_{T,Diana} - C_{T,Actina})_{Tiempo\ x} - (C_{T,Diana} - C_{T,Actina})_{Tiempo\ 0}$$

Tiempo x representa cualquier punto en el tiempo, y Tiempo₀ representa una cantidad doble de la expresión del gen diana tras la corrección de β -actina.

Las secuencias de los cebadores anteriores son tal como sigue:

5'-TCGGCAAAGTTGGTTTTTCATCT-3' (SEQ ID
1: NO: 5);
5'-CCATTCCTGGGCTCCACATT-3' (SEQ ID
1: NO: 6);
5'-ATTGGACTCTGGTGATGGTG-3' (SEQ ID
C NO: 8); y
5'-TCAGCAGAGGTGGTGAACATT-3' (SEQ ID
C: NO: 9).

Los resultados se muestran en la figura 3. El gen diana *RppRC1* prácticamente no se expresó en plantas Tianlong No.1 y no transgénicas con vector vacío; sin embargo, el gen diana *RppRC1* se expresó altamente en *RppRC1* transgénico Tianlong No.1 L1-2 y L2-1.

C. Detección de Southern

Las plantas de soja transgénicas T1 *RppRC1* (L1-1, L1-2, L1-3, L1-4, L1-5, L2-1, y L2-1) obtenidas en el paso 2, la

planta no transgénica Tianlong No. 1 (negativo), y las plantas T1 L3-1, L3-2, L3-3, L3-4 y L3-5 transformadas con vector vacío se tomaron y sometidas a extracción de ADN genómico, respectivamente; el ADN genómico resultante se digirió con endonucleasa HindIII, y luego los productos digeridos con enzimas se sometieron a detección Southern usando el kit de detección de hibridación digoxina II (método de quimioluminiscencia), en el que se usó el gen BAR como sonda, y los cebadores de sonda fueron tal como sigue:

5'-AGAAACCCACGTCATGCCAGTT-3' (SEQ ID NO: 10); y

2:

5'-ATCGTCAACCACTACATCG-3' (SEQ ID NO: 80) (421 bp).

2:

Los resultados se muestran en las figuras 4A y 4B. Las plantas L1-1, L1-2, L1-3, L1-4 y L1-5 eran copias dobles, L2-1 y L2-1 eran copias individuales, y las plantas L3-1, L3-2, L3-3, L3-4 y L3-5 con vector vacío eran copias individuales. En cuanto a la soja transgénica *RppRC1*, estos pueden representar el número de copias del gen *RppRC1*.

Ejemplo 3: Detección de la resistencia a la roya de la soja transgénica

En este ejemplo, se inocularon las plantas transgénicas obtenidas en el Ejemplo 2 con la raza fisiológica SS4 de *Phakopsora pachyrhizi* (descrita en "Shan Zhihui et al., Pathogenic responses of *Phakospora pachyrhizi* en diferentes huéspedes de de legumbres. Chinese Journal of Oil Crop Sciences, 2008, 30(4): 497-500", que está disponible para el público por parte del solicitante y solo puede usarse para repetir los experimentos de la presente invención, pero no para otros fines), y los síntomas de la enfermedad de las plantas se calificaron para determinar el efecto del gen *RppRC1* frente a la roya

En primer lugar, se sometieron a prueba preliminarmente los acontecimientos de transformación T0 L1 y L2 para evaluar el efecto del transgén *RppRC1* sobre la infección con roya. Las operaciones específicas fueron tal como sigue: Se tomaron hojas nuevas completamente desplegadas de plantas T0, se pulverizaron e se inocularon con la suspensión de la raza fisiológica SS4 de *Phakopsora pachyrhizi* (1×10^5 esporas/ml) a un tamaño de inóculo de 10 μ l por centímetro cuadrado. El genotipo receptor no transformado Tianlong No.1 (control negativo) y la planta transformada con el vector vacío del mismo acontecimiento (control de vector vacío) se usaron como controles susceptibles, y la variedad de resistencia a la enfermedad no transformada SX6907 se usó como control de resistencia a la enfermedad (control positivo). Tras la inoculación, se cultivaron las plantas en un invernadero a 25°C con un fotoperiodo de 16 horas de luz/8 horas de oscuridad y una humedad relativa del 65%-85%. Se calificaron los síntomas de la enfermedad 12-15 días tras la inoculación. Se determinó la resistencia a enfermedades de las plantas según la naturaleza de las manchas de la enfermedad y la ruptura de los soros. Se calificaron cualitativamente las plantas como inmunidad (IM: sin lesiones), alta resistencia (R: manchas de enfermedad de color negro rojizo, una pequeña cantidad de formación de esporas) y susceptibilidad (S: manchas de enfermedad de color rojizo, una gran cantidad de formación de esporas). Se hace referencia en "Bromfield, KR., Melching, JS., Kingsolver, CH. 1980, Virulence y Aggressiveness of *Phakopsora pachyrhizi* Isolates Causing Soybean Rust. *Phytopathology*. 70: 17-21". Los resultados mostraron que tanto el control positivo y las plantas transformadas (acontecimientos de transformación T0 L1 y L2) mostraron inmunidad, mientras que las plantas con el vector

vacío y las plantas de control negativo mostraron susceptibilidad, como se muestra en la figura 5 y Tabla -1.

Luego, se sometió a prueba la eficacia del gen *RppRC1* frente a *Phakopsora pachyrhizi* para las plantas transgénicas T1 L1-1, L1-2, L1-3, L2-1 y L2-1, y la planta L3-1 transformada con el vector vacío. Se plantaron las semillas T1 en condiciones de cámara de crecimiento, y la inoculación e identificación se llevaron a cabo cuando la planta creció hasta tener dos hojas verdaderas completamente desplegadas. Se usó la suspensión de esporas de la raza fisiológica SS4 de *Phakopsora pachyrhizi* para la inoculación. El método de inoculación era el mismo que lo anterior. La planta de la variedad no transformada Tianlong No.1 era el control susceptible y la planta de la variedad no transformada SX6907 con resistencia a enfermedades era el control de resistencia a enfermedades. Los síntomas se observaron 12 días después. Los resultados mostraron que no se observó formación de manchas de enfermedad y formación de soro en las hojas de las plantas transgénicas L1-1, L1-2, L1-3, L2-1 y L2-1 con el grado de resistencia de inmunidad. La detección molecular mostró que las plantas transgénicas que comprenden la longitud total de y tenían un alto nivel de expresión del gen *RppRC1*. Las hojas de la planta no transformada y las hojas de la planta L3-1 transformada con el vector vacío produjeron manchas de enfermedad rojizas y produjeron una gran cantidad de esporas, mostrando una respuesta susceptible (figura 6 y Tabla 3-2).

Los resultados de estas pruebas de infección por roya mostraron que el gen *RppRC1* puede proporcionar resistencia frente a *Phakopsora pachyrhizi* a través de la transgénesis.

Tabla 3-1: Resistencia de plantas transgénicas T0 que portan

el gen <i>RppRC1</i> frente a la roya	
Acontecimientos de de transformación	Resistencia
L1	Inmunidad
L2	Inmunidad
Vector vacío L3	Susceptibilidad
Tianlong No.1 de control negativo	Susceptibilidad
SX6907 de control positivo	Inmunidad

Tabla 3-2: Resistencia de plantas transgénicas T1 que portan el gen <i>RppRC1</i> frente a la roya	
Plantas transformadas	Resistencia
L1-1	Inmunidad
L1-2	Inmunidad
L1-3	Inmunidad
L2-1	Inmunidad
L2-2	Inmunidad
Vector vacío L3-1	Susceptibilidad
Tianlong No.1 de control negativo	Susceptibilidad
SX6907 de control positivo	Inmunidad

Los ejemplos 1'3 establecen que la resistencia frente a la roya de la soja transgénica obtenida mediante la transformación del gen *RppRC1* en la variedad de soja susceptible Tianlong No.1 es significativamente mayor que la del Tianlong No.1 progenitor receptor, lo que indica que *RppRC1* y el gen codificante del mismo puede regular y controlar la resistencia de las plantas leguminosas frente a la roya, y mejorar la resistencia a la roya de las plantas después de la sobreexpresión. *RppRC1* y el gen codificante del mismo pueden usarse para mejorar la resistencia a enfermedades de cultivos leguminosos y son de gran

importancia para el fitomejoramiento de nuevas variedades con resistencia a enfermedades.

Ejemplo 4: Uso de la edición de genes para el reemplazo de alelos de genes de resistencia a la roya.

En este ejemplo, se usa la edición de genes para reemplazar un gen de tipo natural con un intervalo o gen que confiere una resistencia aumentada a la roya frente a *Phakopsora pachyrhizi*.

En primer lugar, se diseñan dos ARNg para seleccionar como diana la región de inserción. En este ejemplo, se diseñaron ARNg para seleccionar como diana la región con 736 pb y la región con 1642 pb de Glycine max Williams 82. Se diseñó una secuencia de ADN del donante que incluye una parte de 6057 pb del intervalo de la SEQ ID NO: 13, que se modificó adicionalmente para incluir brazos homólogos de 500 pb en cada lado. Otros ejemplos incluyen otras porciones del intervalo, pero los ejemplos típicos incluirán partes que contienen una secuencia de ácido nucleico que codifica para la proteína de la reivindicación 1.

En segundo lugar, la maquinaria de edición de Cas12a, los ARNg y el ADN del donante se administran al menos a una célula vegetal usando transformación mediada biolística.

Entonces, los acontecimientos de inserción dirigidos pueden detectarse con PCR y secuenciación, por ejemplo. También puede usarse la evaluación fenotípica.

Realizaciones numeradas de la invención

Lo siguiente es todas las realizaciones contempladas y abarcadas dentro de la invención: En determinados aspectos, la realización numerada de la invención incluye:

1. Una proteína seleccionada del grupo constituido por:

(A1) una proteína que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A2) una proteína que tiene sustitución y/o delección y/o adición de uno o varios residuos de aminoácidos de y que tiene la misma función que la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A3) una proteína que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con y que tiene la misma función que la secuencia de aminoácidos definida en o bien (A1) o bien (A2); y

(A4) una proteína de fusión obtenida etiquetando en el N terminal y/o C terminal de la proteína definida en uno cualquiera de (A1) a (A3).

2. Una molécula de ácido nucleico que codifica para la proteína de la realización 1.

3. La molécula de ácido nucleico de la realización 2, en la que la molécula de ácido nucleico es un gen y en la que el gen es una molécula de ADN de cualquiera de:

(B1) una molécula de ADN mostrada en SEQ ID No: 2;

(B2) una molécula de ADN que hibrida con la molécula de ADN definida en (B1) en una condición estricta y que codifica para la proteína;

(B3) una molécula de ADN que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con las secuencias de ADN definidas en (B1) y (B2) y que codifica para la proteína.

4. Un casete de expresión, un vector recombinante, una bacteria recombinante o una línea celular transgénica que comprende la molécula de ácido nucleico de la realización 2 o 3.

5. El casete de expresión de la realización 4, caracterizado porque el promotor para iniciar la transcripción de la molécula de ácido nucleico en el casete de expresión es un promotor endógeno original, y la secuencia

de nucleótidos del promotor endógeno original se muestra en la SEQ ID NO: 7.

6.El vector recombinante de la realización 4, caracterizado porque el vector recombinante es un plásmido recombinante obtenido clonando la molécula de ácido nucleico entre los sitios attR1 y attR2 del vector pB2GW7, y reemplazando el promotor 35S entre los sitios de digestión enzimática SacI y SpeI con el promotor endógeno del gen RppRC1 mostrado en SEQ ID NO: 7.

7.Uso de la proteína de la realización 1 o la molécula de ácido nucleico de la realización 2 o 3, o el vector recombinante, bacteria recombinante o línea celular transgénica de una cualquiera de las realizaciones 4 a 6 en la regulación de la resistencia de una planta frente a la roya.

8.El uso de la realización 7, caracterizado porque en el uso, el nivel de expresión y/o la actividad de la proteína o el gen codificante de la misma en las plantas aumenta, y la resistencia de las plantas frente a la roya se potencia.

9.Un método seleccionado entre

(C1) un método para mejorar la resistencia de una planta frente a la roya, que comprende las siguientes etapas: aumentar el nivel de expresión y/o actividad de la proteína de la realización 1 en la planta; o

(C2) un método para fitomejorar una variedad de plantas con resistencia mejorada frente a la roya, que comprende la etapa de aumentar el nivel de expresión y/o actividad de la proteína de la realización 1 en una planta receptora.

10. El método de la realización 9, caracterizado porque el aumento del nivel de expresión y/o la actividad de la proteína en la planta puede realizarse por medios transgénicos o por hibridación sexual.

11. Un método para fitomejorar una planta transgénica con resistencia mejorada frente a la roya, que comprende la siguiente etapa: introducir la molécula de ácido nucleico de la realización 2 o 3 a una planta receptora para obtener una planta transgénica; la planta transgénica tiene resistencia mejorada frente a la roya en comparación con la planta receptora.

12. El método de la realización 11, caracterizado porque la introducción de la molécula de ácido nucleico a la planta receptora se realiza introduciendo el casete de expresión o el vector recombinante de una cualquiera de las realizaciones 4-6 en la planta receptora.

13. Un par de cebadores para amplificar la molécula de ácido nucleico de la realización 2 o 3.

14. El par de cebadores de la realización 13, caracterizado porque el par de cebadores es un par de cebadores 1 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 3 y SEQ ID NO: 4, respectivamente, o un par de cebadores 2 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 5 y SEQ ID NO: 6, respectivamente.

15. Una sonda para amplificar la molécula de ácido nucleico de la realización 2 o 3.

16. Un kit que comprende el par de cebadores de la realización 13 o 14 y/o la sonda de la realización 15.

17. Una planta que comprende la molécula de ácido nucleico de la realización 2 o 3.

18. La planta de la realización 17, caracterizada porque la planta es una planta transgénica con resistencia mejorada frente a la roya obtenida mediante fitomejoramiento usando el método de la realización 11 o 12, o es soja SX6907, o una planta de progenie que comprende la molécula de ácido nucleico de la realización 2 o 3 obtenida después de la

hibridación sexual usando la soja SX6907 como progenitor; la soja SX6907 tiene el número de registro CGMCC No. 17575 en el Centro de Colección de Cultivo Microbiológico General de China.

19. Uso del par de cebadores de la realización 13 o 14 o la sonda de la realización 15 o el kit de la realización 16 en la identificación de si una planta que va a someterse a prueba comprende la molécula de ácido nucleico de la realización 2 o 3.

20. Uso del par de cebadores de la realización 13 o 14 o la sonda de la realización 15 o el kit de la realización 16 en la identificación de si una planta que va a someterse a prueba tiene resistencia frente a la roya conferida por la molécula de ácido nucleico de la realización 2 o 3.

21. El uso o método o par de cebadores o sonda o kit o planta de una cualquiera de las realizaciones 7-20, caracterizado porque la roya es roya de planta leguminosa.

22. El uso o método o par de cebadores o sonda o kit o planta de la realización 21, caracterizado porque la roya de planta leguminosa es roya de soja.

23. El uso o método o par de cebadores o sonda o kit o planta de la realización 22, caracterizado porque el patógeno de la roya de la soja es *Phakopsora pachyrhizi* o *Phakopsora meibomiae*.

24. El uso o método o par de cebadores o sonda o kit o planta de la realización 23, caracterizado porque *Phakopsora pachyrhizi* es la raza fisiológica SS4 de *Phakopsora pachyrhizi*.

25. El uso o método o par de cebadores o sonda o kit o planta de una cualquiera de las realizaciones 7-24, caracterizado porque la planta es una planta leguminosa.

26. El uso o método o par de cebadores o sonda o kit o planta de la realización 25, caracterizado porque la planta leguminosa es cualquiera de: soja, alfalfa, trébol, guisante, frijol común, lenteja, lupino, árbol de ghaf, algarroba, soja, maní o tamarindo.

[0100] En otro aspecto, las realizaciones numeradas de la invención incluyen:

1. Una planta elite *Glycine max* que tiene en su genoma un intervalo cromosómico de una segunda planta de glicina, en el que dicho intervalo cromosómico confiere una resistencia a la roya de la soja asiática (ASR) aumentada en comparación con una planta de control que no comprende dicho intervalo cromosómico.

2. La planta de la realización 1, en la que el intervalo cromosómico se deriva de la cepa de *Glycine max* SX6907.

3. La planta de las realizaciones 12, donde el intervalo cromosómico comprende la SEQ ID NO: 2 o cualquier parte de la misma, donde la parte confiere resistencia aumentada a la ASR en la planta;

4. La planta de la realización 4, en la que el intervalo cromosómico comprende una secuencia de ácido nucleico que es al menos el 85%, al menos el 90%, o al menos el 95% idéntica a SEQ ID NO: 2.

5. La planta de una cualquiera de las realizaciones 1-4, en la que el intervalo cromosómico codifica para una proteína de SEQ ID NO: 1 y en la que la proteína confiere resistencia aumentada a la ASR en la planta.

6. La planta de una cualquiera de las realizaciones 1-5, en la que el intervalo cromosómico de la segunda planta de glicina se inserta en el cromosoma 18 de la planta.

7. La planta de una cualquiera de las realizaciones 1-6, en la que el intervalo cromosómico de la segunda planta de glicina se inserta en la región que comienza en aproximadamente el par de bases 56,680,416 y que termina en aproximadamente el par de bases 56,677,361 en el cromosoma 18 de la cepa de *Glicina max* Williams 82 o equivalente del mismo en otras cepas de *Glycine max*.

8. La planta de una cualquiera de las realizaciones 1-7, donde el intervalo cromosómico comprende la SEQ ID NO: 11-13 o una parte de cualquiera de las mismas en la que dicha parte confiere en dicha planta una resistencia a la ASR aumentada.

9. La planta de una cualquiera de las realizaciones 1-7, en la que el intervalo cromosómico comprende un marcador SNP asociado con resistencia a la ASR aumentada en donde dicho marcador SNP corresponde con uno cualquiera de los marcadores SNP favorables como se enumeran en la Tabla 1.

10. La planta de una cualquiera de las realizaciones 1-9, en la que el intervalo se deriva del cromosoma 18 de la segunda planta.

11. La planta de una cualquiera de las realizaciones 1-10, donde al menos una línea progenitora de dicha planta se seleccionó o identificó mediante selección de marcadores moleculares, donde dicha línea progenitora se selecciona o identifica basándose en la presencia de un marcador molecular ubicado dentro o estrechamente unido con dicho intervalo cromosómico correspondiente a una cualquiera de las SEQ ID NO: 11-13 donde dicho marcador molecular está asociado con la resistencia a la ASR aumentada.

12. La planta de la realización 11, donde el marcador molecular es un polimorfismo mononucleotídico (SNP), un locus de rasgo cuantitativo (QTL), un polimorfismo de longitud de

fragmento amplificado (AFLP), ADN polimórfico amplificado aleatoriamente (RAPD), un polimorfismo de longitud de fragmento de restricción (RFLP) o un microsatélite.

13. La planta de las realizaciones 11o 12, donde el marcador molecular es un SNP marcador y el marcador molecular es cualquier marcador favorable mostrado en las tablas 1.

14. La planta de una cualquiera de las realizaciones 1-13, donde la planta es una planta agronómicamente de elite de *Glycine max* que tiene una producción comercialmente significativa y/o vigor comercialmente susceptible, formación de semillas, capacidad de resistir erguido o facilidad de desgranarse.

15. La planta de una cualquiera de las realizaciones 1-13, donde dicho intervalo se introduce en dicho genoma de la planta mediante edición genómica de secuencias correspondientes a y que comprenden una cualquiera de las SEQ ID NO: 2 o 11.

16. La planta de la realización 15, donde el intervalo se introduce mediante edición genómica de una región genómica de *Glycine max* homóloga a, o un ortólogo de cualquiera de los intervalos correspondientes a las SEQ ID NO: 11-13 y haciendo además al menos una edición genómica a dicha región genómica de *Glycine max* para que incluya al menos 1 cambio de alelo correspondiente a cualquier alelo favorable como se describe en cualquiera de las tablas 1, donde dicha región genómica de *Glycine max* no comprendía dicho cambio de alelo antes de la edición genómica y además, donde dicha edición genómica confiere en una planta una resistencia a la ASR aumentada.

17. La planta de la realización 16, donde la edición genómica se consigue mediante CRISPR, TALEN, meganucleasas o mediante modificación de ácidos nucleicos genómicos.

18. La planta de una cualquiera de las realizaciones 1-17, donde dicho intervalo se introduce en dicho genoma de la planta mediante expresión transgénica de secuencias correspondientes a y que comprenden una cualquiera de las SEQ ID NO: 11-13.

19. La planta de una cualquiera de las realizaciones 1-18, donde el intervalo cromosómico comprende la SEQ ID NO: 11, o una parte de la misma que confiere resistencia a la ASR.

20. La planta de una cualquiera de las realizaciones 1-18, donde el intervalo cromosómico comprende la SEQ ID NO: 12, o una parte de la misma que confiere resistencia a la ASR.

21. La planta de una cualquiera de las realizaciones 1-18, donde el intervalo cromosómico comprende la SEQ ID NO: 13, o una parte de la misma que confiere resistencia a la ASR.

22. Una planta de *Glycine max* agronómicamente de élite que tiene una producción comercialmente significativa, que comprende un intervalo cromosómico derivado de *Glycine max* SX6907, un intervalo cromosómico que comprende SEQ ID NO: 2, un intervalo cromosómico que comprende SEQ ID NO: 11-13, un intervalo cromosómico que codifica para la proteína de la SEQ ID NO: 1 o una parte de las mismas, donde dicho intervalo cromosómico o parte del mismo confiere resistencia aumentada a la ASR en dicha planta, en comparación con una planta de control que no comprende dicho intervalo cromosómico.

23. Una célula vegetal, semilla, o parte de la planta derivada de la planta de una cualquiera de las realizaciones 1-22.

24. Una planta de progenie de las plantas de una cualquiera de las realizaciones 1-23.

25. Un método de producción de una planta de *Glycine max* que tiene resistencia aumentada a la roya de la soja asiática (ASR), comprendiendo el método las etapas de:

a) proporcionar una primera planta de *Glycine max* que comprende en su genoma un intervalo cromosómico correspondiente a SEQ ID NO: 11, 12 o 13, o un intervalo cromosómico que codifica para la proteína de SEQ ID NO: 1 o un intervalo cromosómico que comprende SEQ ID NO: 2, donde dicha primera planta de *Glycine max* tiene resistencia aumentada a la ASR;

b) cruzar la planta de *Glycine max* de a) con una segunda planta de *Glycine max* que no comprende dicho intervalo cromosómico; y

c) seleccionar una planta de la progenie del cruce de b) aislando un ácido nucleico de dicha planta de la progenie y detectar dentro de dicho ácido nucleico un alelo que se asocia con resistencia aumentada a la ASR y en el que, además, dicho alelo está estrechamente unido con o ubicado dentro de los intervalos cromosómicos correspondientes a SEQ ID NO: 11, 12 o 13, o un intervalo cromosómico que codifica para la proteína de SEQ ID NO: 1 o cromosoma que comprende SEQ ID NO: 2, produciendo con ello una planta de *Glycine max* que tiene resistencia aumentada a la ASR.

26. El método de la realización 25, donde el alelo corresponde a cualquiera de los alelos favorables tal como se representa en la Tabla 1.

27. El método de las realizaciones 25 o 26, en el que o bien la primera o bien la segunda planta de *Glycine max* es una planta de *Glycine max* de élite.

28. Un método de producción de una planta de *Glycine max* con resistencia aumentada a la roya de la soja asiática (ASR), comprendiendo el método las etapas de:

a) aislar un ácido nucleico de una planta de *Glycine max*;

b) detectar en el ácido nucleico de a) al menos un marcador molecular asociado con resistencia aumentada a la ASR, donde dicho marcador molecular está ubicado dentro de 20 cM, 10 cM, 5 cM, 1 cM 0.5 cM, o está ligado estrechamente con un intervalo cromosómico correspondiente a una región genómica del cromosoma 18 de *Glycine max* que comprende SEQ ID NO 2, o SEQ ID NO: 11 o una parte de las mismas en la que dicha parte confiere a una planta una resistencia aumentada a la ASR.

c) seleccionar una planta en base a la presencia del marcador molecular detectado en b); y

d) producir una planta de la progenie de *Glycine max* de la planta de c) identificada por tener dicho alelo asociado con resistencia aumentada a la ASR.

29. El método de la realización 28, en el que el marcador molecular está estrechamente unido con o consiste en uno cualquiera de los alelos favorables tal como se representa en la Tabla 1.

30. Un método de identificar o seleccionar una planta de *Glycine max* que tiene una resistencia aumentada a la ASR, comprendiendo el método las etapas de:

a) aislar un ácido nucleico de una planta de *Glycine max*;

b) detectar en el ácido nucleico la presencia de un marcador molecular que se asocia con resistencia aumentada a la ASR, en donde el marcador molecular está ubicado dentro de 20cM, 10cM, 5cM, 1cM, 0.5cM de un marcador tal como se describe en la Tabla 1; y

c) identificar o seleccionar una planta de *Glycine max* que tiene resistencia aumentada a la ASR en función del marcador molecular detectado en b).

31. El método de la realización 30, en el que el alelo detectado en b) consiste en cualquiera de los marcadores favorables tal como se describen en la Tabla 1.

32. El método de una cualquiera de las realizaciones 28-31, donde el marcador molecular es un polimorfismo mononucleotídico (SNP), un locus de rasgo cuantitativo (QTL), un polimorfismo de longitud de fragmento amplificado (AFLP), ADN polimórfico amplificado aleatoriamente (RAPD), un polimorfismo de longitud de fragmento de restricción (RFLP) o un microsatélite.

33. El método de una cualquiera de las realizaciones 28-32, en el que la detección comprende amplificar el locus de un marcador o una porción del locus del marcador y detectar el amplicón del marcador amplificado resultante.

34. El método de la realización 33, en el que la amplificación comprende: a) mezclar un cebador de amplificación o un par de cebadores de amplificación con un ácido nucleico aislado de la primera planta de *Glycine max* o germoplasma de soja, en donde el cebador o el par de cebadores es complementario o parcialmente complementario a al menos una porción del locus del marcador, y es capaz de iniciar la polimerización del ADN mediante una ADN polimerasa utilizando el ácido nucleico de la *Glycine max* como un molde; y b) extender el cebador o el par de cebadores en una reacción de polimerización de ADN que comprende una ADN polimerasa y un ácido nucleico molde para generar al menos un amplicón.

35. El método de las realizaciones 33 o 34, en el que la amplificación comprende emplear una reacción en cadena de

la polimerasa (PCR) o una reacción en cadena de la ligasa (LCR) utilizando un ácido nucleico aislado de una planta o germoplasma de soja como un molde en la PCR o LCR.

36. El método de una cualquiera de las realizaciones 28-35, en el que el ácido nucleico se selecciona de ADN o ARN.

37. Un cebador diagnóstico para la resistencia a la ASR, en donde dicho cebador se puede utilizar en una reacción PCR para indicar la presencia de un alelo asociado con la resistencia a la ASR, en donde dicho alelo es cualquier alelo favorable según se describe en la Tabla 1.

38. Un método de conferir resistencia a la ASR frente a plantas de *Glycine max* que comprenden:

a) proporcionar una molécula de ácido nucleico del cromosoma 18 de una planta de *Glycine max* que tiene resistencia a la ASR, donde dicho ácido nucleico codifica la resistencia a la ASR; y

b) insertar la molécula de ácido nucleico en el cromosoma 18 de una cepa de *Glycine max* que carece de resistencia a la ASR para producir de este modo una planta que tenga resistencia aumentada a la ASR en comparación con una planta de control que no comprende el ácido nucleico,

donde la molécula de ácido nucleico se selecciona del grupo que consiste en:

un ácido nucleico que codifica para la proteína de SEQ ID NO: 1;

un ácido nucleico que comprende la secuencia de ácido nucleico de SEQ ID NO: 2 o cualquier parte de la misma, donde la parte confiere resistencia aumentada a la ASR en la planta; y

un intervalo cromosómico que comprende SEQ ID NO: 11.

39. El método de la realización 38, donde la molécula de ácido nucleico se deriva de la cepa de *Glycine max* SX6907.

40. El método de la realización 38 o la realización 39, en el que la molécula de ácido nucleico se inserta en el cromosoma 18 de una planta de *Glycine max*.

41. El método de una cualquiera de las realizaciones 38-40, en el que la molécula de ácido nucleico se inserta en una región que empieza en aproximadamente el par de bases 56,680,416 y que termina en aproximadamente el par de bases 56,677,361.

42. El método de una cualquiera de las realizaciones 39-41, en el que el método comprende el reemplazo del gen mediado por Cas12a.

43. El método de una cualquiera de las realizaciones 38-42, en el que el método comprende 2 ARNg.

44. El método de la realización 43, en el que el método comprende un ARNg de SEQ ID NO: 79 y/o SEQ ID NO: 81.

45. El método de una cualquiera de las realizaciones 38-44, en el que el método comprende la detección de la inserción localizada con PCR y/o secuenciación.

46. Una proteína seleccionada de:

(A1) una proteína que tiene la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A2) una proteína que tiene sustitución y/o delección y/o adición de uno o varios residuos de aminoácidos de y que tiene la misma función que la secuencia de aminoácidos mostrada en SEQ ID NO: 1;

(A3) una proteína que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con y que tiene la misma función que la secuencia de aminoácidos definida en o bien (A1) o bien (A2); o

(A4) una proteína de fusión obtenida etiquetando en el N terminal y/o C terminal de la proteína definida en uno cualquiera de (A1) a (A3).

47. Una molécula de ácido nucleico que codifica para la proteína de la realización 46.

48. La molécula de ácido nucleico de la realización 47, en la que la molécula de ácido nucleico es cualquiera de:

(B1) una molécula de ADN mostrada en SEQ ID No: 2;

(B2) una molécula de ADN que hibrida con la molécula de ADN definida en (B1) en una condición estricta y que codifica para la proteína; o

(B3) una molécula de ADN que tiene más del 99%, más del 95%, más del 90%, más del 85%, o más del 80% de homología con las secuencias de ADN definidas en (B1) y (B2) y que codifica para la proteína.

49. Un casete de expresión, un vector recombinante, una bacteria recombinante o una línea celular transgénica que comprende la molécula de ácido nucleico de la realización 47 o 48.

50. El casete de expresión de la realización 49, caracterizado porque el promotor para iniciar la transcripción de la molécula de ácido nucleico en el casete de expresión es un promotor endógeno original, y la secuencia de nucleótidos del promotor endógeno original se muestra en la SEQ ID NO: 7.

51. El vector recombinante de la realización 50, caracterizado porque el vector recombinante es un plásmido recombinante obtenido clonando la molécula de ácido nucleico entre los sitios attR1 y attR2 del vector pB2GW7, y reemplazando el promotor 35S entre los sitios de digestión enzimática SacI y SpeI con el promotor endógeno del gen *RppRC1* mostrado en SEQ ID NO: 7.

52. Uso de la proteína de la realización 46 o la molécula de ácido nucleico de la realización 47 o 48, o el vector recombinante, bacteria recombinante o línea celular transgénica de una cualquiera de las realizaciones 49 a 51 en la regulación de la resistencia de una planta frente a la roya.

53. El uso de la realización 47, en el que se aumenta el nivel de expresión y/o actividad de la proteína del gen codificador de la misma en la planta, y se aumenta la resistencia de la planta frente a la roya.

54. Un método para mejorar la resistencia de una planta frente a la roya, que comprende aumentar el nivel de expresión y/o actividad de la proteína de la realización 46 en la planta;

55. Un método para fitomejorar una variedad de plantas con resistencia mejorada frente a la roya, que comprende aumentar el nivel de expresión y/o actividad de la proteína de la realización 46 en una planta receptora.

56. El método de las realizaciones 54 o 55, en el que el aumento del nivel de expresión y/o la actividad de la proteína en la planta puede realizarse por medios transgénicos o por hibridación sexual.

57. Un método para fitomejorar una planta transgénica con resistencia mejorada frente a la roya, que comprende la siguiente etapa: introducir la molécula de ácido nucleico de la realización 47 o 48 a una planta receptora para obtener una planta transgénica; la planta transgénica tiene resistencia mejorada frente a la roya en comparación con la planta receptora.

58. El método de la realización 56, en el que la introducción de la molécula de ácido nucleico a la planta receptora se realiza introduciendo el casete de expresión o

el vector recombinante de una cualquiera de las realizaciones 49-51 en la planta receptora.

59. Un par de cebadores para amplificar la molécula de ácido nucleico de la realización 47 o 48.

60. El par de cebadores de la realización 59, en el que el par de cebadores es un par de cebadores 1 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 3 y SEQ ID NO: 4, respectivamente, o un par de cebadores 2 compuesto por dos ADN monocatenarios mostrados en SEQ ID NO: 5 y SEQ ID NO: 6, respectivamente.

61. Una sonda para amplificar la molécula de ácido nucleico de la realización 47 o 48.

62. Un kit que comprende el par de cebadores de la realización 59 o 60 y/o la sonda de la realización 61.

63. Una planta que comprende la molécula de ácido nucleico de la realización 47 o 48.

64. La planta de la realización 63, en la que la planta es una planta transgénica con resistencia mejorada frente a la roya obtenida mediante fitomejoramiento usando el método de la realización 56 o 57, o es soja SX6907, o una planta de progenie que comprende la molécula de ácido nucleico de la realización 47 o 48 obtenida después de la hibridación sexual usando la soja SX6907 como progenitor; la soja SX6907 tiene el número de registro CGMCC No. 17575 en el Centro de Colección de Cultivo Microbiológico General de China.

65. Uso del par de cebadores de la realización 59 o 60 o la sonda de la realización 61 o el kit de la realización 62 en la identificación de si una planta que va a someterse a prueba comprende la molécula de ácido nucleico de la realización 45 o 46.

66. Uso del par de cebadores de la realización 59 o 60 o la sonda de la realización 61 o el kit de la realización 62 en la identificación de si una planta que va a someterse a prueba tiene resistencia frente a la roya conferida por la molécula de ácido nucleico de la realización 47 o 48.

67. El uso o método o par de cebadores o sonda o kit o planta de una cualquiera de las realizaciones 52-66, donde la roya es roya de planta leguminosa.

68. El uso o método o par de cebadores o sonda o kit o planta de la realización 67, donde la roya de planta leguminosa es roya de la soja.

69. El uso o método o par de cebadores o sonda o kit o planta de la realización 68, donde el patógeno de la roya de la soja es *Phakopsora pachyrhizi* o *Phakopsora meibomiaae*.

70. El uso o método o par de cebadores o sonda o kit o planta de la realización 69, en el que *Phakopsora pachyrhizi* es la raza fisiológica SS4 de *Phakopsora pachyrhizi*.

71. El uso o método o par de cebadores o sonda o kit o planta de una cualquiera de las realizaciones 52-70, en el que la planta es una planta leguminosa.

72. El uso o método o par de cebadores o sonda o kit o planta de la realización 73, en el que la planta leguminosa es cualquiera de: soja, alfalfa, trébol, guisante, frijol común, lenteja, lupino, árbol de ghaf, algarroba, soja, maní o tamarindo.

Se hace constar que con relación a esta fecha, el mejor método conocido por la solicitante para llevar a la práctica la citada invención, es el que resulta claro de la presente descripción de la invención.