

OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



⑪ Número de publicación: **2 375 933**

⑯ Int. Cl.:
A01H 1/00
(2006.01)

⑫

TRADUCCIÓN DE PATENTE EUROPEA

T3

- ⑯ Número de solicitud europea: **02772254 .5**
⑯ Fecha de presentación: **30.08.2002**
⑯ Número de publicación de la solicitud: **1420631**
⑯ Fecha de publicación de la solicitud: **26.05.2004**

⑭

Título: **GEN DE RESISTENCIA DERIVADO DE PLANTA.**

⑯ Prioridad:
31.08.2001 EP 01120670

⑬ Titular/es:
**MAX-PLANCK-GESELLSCHAFT ZUR
FÖRDERUNG DER WISSENSCHAFTEN E.V.
14195 BERLIN, DE y
KWS SAAT AG**

⑯ Fecha de publicación de la mención BOPI:
07.03.2012

⑬ Inventor/es:
**GEBHARDT, Christiane;
BALLVORA, Agim;
ERCOLANO, Maria, Raffaella;
WEISS, Julia y
SALAMINI, Francesco**

⑯ Fecha de la publicación del folleto de la patente:
07.03.2012

⑬ Agente: **Miltenyi, Peter**

ES 2 375 933 T3

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

DESCRIPCIÓN

Gen de resistencia derivado de planta

5 La presente invención se refiere al gen de resistencia *R1* de la patata. Se refiere además a métodos y materiales que emplean el gen, y a procedimientos para identificar o producir otros genes relacionados. También se refiere generalmente a métodos para identificar agentes protectores de plantas que pueden inducir el gen *R1* o la actividad de su proteína codificada. Además, la presente invención se refiere a plantas transgénicas que se hicieron resistentes a tizón tardío debido a la expresión de un transgén *R1*.

10 Se citan varios documentos a lo largo de todo el texto de esta memoria descriptiva por el nombre. Pueden encontrarse las citas bibliográficas completas al final de la memoria descriptiva, precediendo inmediatamente a la lista de secuencias o las reivindicaciones.

15 El tizón tardío es a nivel mundial la enfermedad más destructora para el cultivo de patata provocando pérdidas de miles de millones de dólares cada año (Kamoun *et al.* 1999). El patógeno causante es *Phytophthora infestans*, un hongo oomiceto que también infecta los tomates (Judelson 1997). La destrucción completa del cultivo de patata por el tizón tardío provocó la "hambruna irlandesa de la patata" a mediados del siglo XIX (Salaman 1985) e inició la búsqueda de plantas resistentes. Se descubrieron genes individuales para la resistencia al tizón tardío (genes *R*) hace casi 100 años en *S. demissum*, una especie de patata silvestre autóctona de México. La introgresión en cultivares de patata de genes *R* que confieren resistencia específica de la variedad proporcionó, sin embargo, sólo una resistencia transitoria al tizón tardío, ya que las nuevas variedades superaron rápidamente la resistencia mediada por genes *R* (Wastie 1991, Fry 20 y Goodwin 1997). También se ha identificado resistencia cuantitativa o de campo al tizón tardío en especies de patata silvestres (Ross 1986). Esta resistencia es más duradera que la medida por genes *R*, pero difícil de mover en variedades cultivadas mediante cruzamiento y selección fenotípica. El tizón tardío se controla todavía en su mayor parte por la aplicación frecuente de fungicidas que pierden eficacia por la selección de aislados resistentes a los fungicidas.

25 25 Varios genes *R* se han mapeado en cromosomas de patata usando marcadores de ADN (Leonards-Schippers *et al.* 1992, El-Kharbotly *et al.* 1994, 1996, Li *et al.* 1998, Ewing *et al.* 2000, Naess *et al.* 2000). *R1* está ubicado en el cromosoma V (Leonards-Schippers *et al.* 1992) en una región en la que también se han mapeado genes para la resistencia al virus X de la patata (Ritter *et al.* 1991, De Jong 1997). La misma región contiene loci de rasgos cuantitativos (LRC) principales para la resistencia al nematodo quístico parasitario de la raíz *Globodera pallida* (Kreike *et al.* 1994, Rouppe van der Voort *et al.* 1997, 2000) y tizón tardío (Leonards-Schippers *et al.* 1994, Oberhagemann *et al.* 1999, Collins 1999). La presencia de un punto caliente de genes de resistencia sugiere su evolución a partir de ancestros comunes mediante duplicación génica local seguida por diversificación funcional (Leonards-Schippers *et al.* 1994, Leister *et al.* 1996, Oberhagemann *et al.* 1999, Gebhardt y Valkonen 2001). Si este es el caso, la clonación molecular del gen *R1* debería abrir la posibilidad de estudiar a nivel molecular varios factores que se mapean en la región y que participan en el control de la resistencia cualitativa y cuantitativa a diversos patógenos.

30 35 Por tanto, el problema técnico que subyace a la presente invención era cumplir con la necesidad de genes de resistencia a patógenos de plantas y sus secuencias reguladoras. Se logra la solución al problema técnico proporcionando las realizaciones caracterizadas en las reivindicaciones y descritas adicionalmente a continuación.

40 40 Según la presente invención, se ha clonado y caracterizado a nivel molecular *R1*, el cebador gen para la resistencia al tizón tardío. Se identificó el gen mediante un enfoque combinado particular de clonación posicional y gen candidato. La estructura molecular del gen permite clasificar *R1* entre los genes de resistencia de plantas que contienen motivos de cremallera de leucina y NBS-LRR conservados (Ellis *et al.* 2000, Dangl y Jones, 2001).

45 45 *R1* se clonó usando una estrategia clonación posicional en combinación con la búsqueda de genes candidatos que tienen una similitud de secuencia de ADN con genes de resistencia de plantas conocidos (Hammond-Kosack y Jones 1997, Ellis *et al.* 2000). Un enfoque similar fue satisfactorio para la clonación de genes de la patata para la resistencia al virus X de la patata (*Rx1*, Bendahmane *et al.* 1999) y el nematodo quístico de la raíz *Globodera pallida* (*Gpa2*, Van der Vossen *et al.* 2000). Se inició un paseo cromosómico hacia *R1* desde dos loci marcadores *SPUD237* y *AFLP1*, que flanquean *R1* a cortas distancias genéticas de 0,1 cM. El paseo desde el marcador *AFLP1* demostró ser improductivo, sin embargo, debido a la escasez de clones BAC y YAC (Leister *et al.* 1997) que tienen el marcador *AFLP1*. La identificación de clones genómicos de patata con insertos solapantes se facilitó mediante la tecnología BAC en combinación con el uso de macromatrizes de clones BAC. La aplicación de este enfoque fue satisfactoria en el arroz (Nakamura *et al.* 1997, Yang *et al.* 1997, Yang *et al.* 1998) y el tomate (Folkertsma *et al.* 1998). El mapa físico (figura 1) cubre al menos 250 kb del genoma de la patata. Aproximadamente 200 kb era una región candidata a contener el gen *R1* basado en ligamiento sin recombinación a marcadores de extremo BAC. La región candidata estaba abierta por los extremos hacia el locus *AFLP1* porque no se incluyó el acontecimiento de recombinación individual que separa *R1* y *AFLP1* en el mapa físico. La información de secuencia parcial de la región candidata identificó un fragmento génico similar al gen de resistencia *RGL* que detectó una familia de genes con miembros presentes en ambos cromosomas que llevan el alelo de susceptibilidad *r1* o el alelo de resistencia *R1*. De hecho, la sonda *RGL* usada para identificar ADNc y clones BAC para *R1* era parte de un alelo de susceptibilidad *r1*. Basándose en un ensayo de PCR específico de

5 alelo derivado de un clon de ADNc que codifica para parte de *R1*, se demostró que estaba presente un miembro funcional de la familia de genes candidatos en plantas que tenían el alelo de resistencia *R1* y que estaba ausente en plantas susceptibles. Este gen candidato se subclonó a partir de BAC BA87d17 y se transformó de manera estable en el cultivar susceptible Desirée. La complementación del fenotipo *R1* en varias plantas transgénicas mostró que el gen candidato era, en efecto, el gen *R1*.

Las realizaciones de la invención se definen por las reivindicaciones.

Por consiguiente, la presente invención se refiere a una molécula de ácido nucleico que codifica para un polipéptido que puede conferir resistencia frente a un patógeno en una planta en la que se expresa dicho polipéptido, comprendiendo dicha molécula de ácido nucleico o consistiendo en una secuencia de nucleótidos seleccionados del grupo que consiste en:

- (a) una secuencia de nucleótidos que codifica para al menos la forma madura de una proteína (R1) que comprende la secuencia de aminoácidos facilitada en SEQ ID NO: 2;
- (b) una secuencia de nucleótidos que comprende las regiones codificantes de la secuencia de ADN facilitada en SEQ ID NO: 1;
- 15 (c) una secuencia de nucleótidos que se hibrida con la cadena complementaria de una secuencia de nucleótidos definida en (a) o (b) en condiciones de hibridación rigurosas, en la que dicha secuencia de nucleótidos es homóloga en al menos el 80% con las secuencias de nucleótidos de (a) o (b);
- (d) una secuencia de nucleótidos que codifica para una proteína derivada de la proteína codificada por una secuencia de nucleótidos de (a) o (b) por medio de sustitución, delección y/o adición de uno o varios aminoácidos de la secuencia de aminoácidos codificada por la secuencia de nucleótidos de (a) o (b), en la que dicha secuencia de nucleótidos es homóloga en al menos el 80% con las secuencias de nucleótidos de (a) o (b);
- 20 (e) una secuencia de nucleótidos que codifica para una proteína que tiene una secuencia de aminoácidos idéntica en al menos el 60% a la secuencia de aminoácidos codificada por la secuencia de nucleótidos de (a) o (b);
- (f) una secuencia de nucleótidos que codifica para al menos un dominio de cremallera de leucina (LZ, *leucine zipper*) correspondiente a la posición de aminoácidos 308-329 de SEQ ID NO: 2, un dominio de sitio de unión a ácido nucleico (NBS, *nucleic binding site*) correspondiente a la posición de aminoácidos 572-682 de SEQ ID NO: 2 y/o un dominio de repetición rica en leucina (LRR, *leucine rich repeat*) correspondiente a la posición de aminoácidos 780-1280 de SEQ ID NO: 2;
- 25 (g) una secuencia de nucleótidos que codifica para una parte que lleva un epítopo de una proteína R1 codificada por una secuencia de nucleótidos de (a) o (b);
- (h) la secuencia de nucleótidos que está degenerada como resultado del código genético con respecto a una secuencia de nucleótidos de una cualquiera de (a) a (g),

30 en la que dicho patógeno se selecciona del grupo que consiste en *Phytophthora sojae*, *Peronospora parasitic*, *Magaporthe grisea*, *Erysiphe spp*, *Pseudomonas syringae*, *Erwinia amylovora*, *Erwinia carotovora*, *Botrytis cinerea*, *Rhizoctonia solani*, *Pythium debaryanum* y *Phytophthora infestans*.

35 Pueden proporcionarse moléculas de ácido nucleico según la presente invención en forma recombinante o libre o sustancialmente libre de ácido nucleico o genes de la especie de interés u origen distinto de la secuencia que codifica para un polipéptido con la función requerida. Las moléculas de ácido nucleico (y sus productos de polipéptido codificados) también pueden (i) aislarse y/o purificarse de su entorno natural (aunque no necesariamente en forma pura *per se*), o (ii) en forma sustancialmente pura u homogénea.

40 El ácido nucleico según la presente invención puede incluir ADNc, ARN, ADN genómico, preferiblemente el gen intacto, y puede ser total o parcialmente sintético (construcciones). Cuando se especifica una secuencia de ADN, por ejemplo con referencia a una figura o SEQ ID NO, a menos que el contexto requiera lo contrario, está englobado el equivalente de ARN, con U sustituido por T donde aparezca. También está englobado el complemento de las diversas secuencias dadas a conocer, que pueden usarse en experimentos con sondas, o en la regulación por disminución de la secuencia.

45 Un aspecto particular de la invención es una molécula de ácido nucleico que tiene la secuencia mostrada en SEQ ID NO: 1. Dentro de SEQ ID NO: 1, hay aparentemente un gran marco de lectura abierto (ORF). La posterior comparación de la secuencia de ADN genómico con las secuencias de ADNc reveló que el gen contiene tres exones y tres intrones; véase el ejemplo 5 y la figura 4. La supuesta secuencia del polipéptido de *R1* se muestra en la figura 4 designada como SEQ ID NO: 2. *R1* parece contener 1293 residuos de aminoácido y tiene un peso molecular de 149,4 kDa. Moléculas de ácido nucleico particulares de este aspecto de la invención incluyen las que codifican para el

5 producto de proteína R1 y ADNc, que se cree que son las bases 2223-6321, excluyendo los intrones marcados tal como se muestra (4878-4970 y 6130-6229 inclusive). Sorprendentemente la estructura primaria de R1 es similar a la de la clase L. Zip/NBS-LRR (Hammond-Kosack y Jones 1997) de proteínas R. Basándose en la secuencia de proteína deducida, R1 pertenece a la clase L.Zip/NBS/LRR de genes de resistencia de plantas (Hammond-Kosack y Jones 1997). Se cree que el motivo de cremallera de leucina (L.Zip) en la región amino-terminal es característico en la dimerización o interacción con otras proteínas. El supuesto dominio de sitio de unión a nucleótidos (NBS) en el sentido 3' puede estar implicado en la ruta de transducción de señales que conduce al inicio de la respuesta de resistencia. El dominio de repetición rica en leucina (LRR) C-terminal coincide con la secuencia consenso para un dominio LRR citoplasmático tal como describen Jones y Jones (1997) y puede funcionar en interacciones proteína-proteína y unión a ligandos. Se ha mostrado que los dominios LRR de alelos del gen de resistencia a la raya del lino L determinan el reconocimiento de variedades específicas del patógeno (Ellis *et al.* 1999). La predicción *in silico* de cuatro sitios de miristilación y 43 de fosforilación en la secuencia de R1 sugiere un posible anclaje de la proteína R1 en la membrana plasmática y etapas de fosforilación/desfosforilación, respectivamente, que participan en la transducción de señales (Dangl y Jones 2001).

10 15 R1 está ubicado en el brazo corto del cromosoma V (Leonards-Schippers *et al.* 1992, Dong *et al.* 2000) y está relacionado en la secuencia con el gen *Prf* del tomate para la resistencia a *Pseudomonas syringae* que está ubicado en el cromosoma 5 del tomate dentro del gen *Pto/Fen* del agrupamiento de genes de resistencia (Salmeron *et al.* 1996). Los cromosomas cinco de la patata y el tomate son colineales entre sí excepto por una inversión paracéntrica del brazo corto (Tanksley *et al.* 1992). El locus *StPto* de la patata correspondiente a *Pto/Fen* se mapea a más de 10 cM de manera proximal a R1 (Leister *et al.* 1996), excluyendo, por tanto, la posibilidad de que R1 y *Prf* se ubiquen en una 20 25 región genómica colineal. Éste es el caso, sin embargo, cuando se considera el gen *Bs4* del tomate que confiere resistencia al patógeno bacteriano *Xanthomonas campestris*. La posición del locus de la patata correspondiente a *Bs4* puede deducirse a partir del ligamiento estrecho (1 cM) entre *Bs4* y el marcador TG432 (Ballvora *et al.* 2001) que se mapea a 3,8 cM de manera distal a *GP21* en el mapa molecular del tomate (Tanksley *et al.* 1992). Esta región del cromosoma 5 del tomate que se extiende de manera distal desde el marcador *GP21* debe ser colineal con el intervalo 30 de la patata *GP21 - GP179* incluyendo R1, cuando se tiene en cuenta la inversión paracéntrica entre los dos genomas.

30 35 Dos genes de la patata para la resistencia al virus X de la patata, Rx2 y Nb, también se mapean en posiciones similares que R1 (Ritter *et al.* 1991, Leonards-Schippers *et al.* 1992, De Jong *et al.* 1997). Se ha clonado el gen Rx2 y es, como R1, un miembro de la clase L.Zip/NBS/LRR de genes de resistencia (Bendahmane *et al.* 2000). Los dos genes de resistencia comparten una identidad de secuencia de sólo el 32% y son, por tanto, miembros bastante diferentes de la misma superfamilia de genes. Nb está ubicado en el intervalo *GP21 - SPUD237* (De Jong *et al.* 1997) que no contiene R1 y está separado genéticamente, por tanto, de R1.

40 45 En un aspecto adicional, se dan a conocer variantes activas, homólogas de las secuencias de R1, que pueden ser por ejemplo mutantes u otros derivados, u homólogos de R1 que se producen de manera natural tales como variantes alélicas, parálogos (de la misma especie, pero en una ubicación diferente, por ejemplo, pseudoalelos en loci ligados), u ortólogos (genes relacionados de diferente especie). Se muestran a continuación ejemplos de éstos. En cada caso, la variante codifica para un producto que es homólogo (similar) a R1, que puede aislarse o producirse basándose en la secuencia, y puede conferir resistencia a patógenos frente a uno o más patógenos.

50 55 Puede someterse a prueba la actividad de genes de resistencia mediante métodos convencionales conocidos en la técnica, según sea apropiado para la naturaleza de la resistencia que esté investigándose. Pueden encontrarse métodos de ejemplo en las siguientes publicaciones: bacterianos (Grant, (1995) *Science* 269, 843-846); fúngicos (Dixon, (1996) *Cell* 84, 451-459; Jones, (1994) *Science* 266, 789-793; Thomas, (1997) *The Plant Cell* 9, 2209-2224; con nematodos y virales (Whitham, (1994) *Cell* 78, 1101-1115). Normalmente, se somete a prueba la actividad mediante la complementación de un rasgo en una planta; véase el ejemplo 4. Esto puede lograrse usando el gen aislado o por ejemplo mediante el acoplamiento de la supuesta variante activa con un promotor y un terminador para la expresión en plantas y transformándola en una planta susceptible que carece de un rasgo de resistencia dado. La actividad de la variante de R1 se confirma entonces mediante exposición con el patógeno apropiado. Alternativamente, puede usarse un ensayo de expresión transitoria para someter a prueba la activación de la variante de R1 análogo al ensayo usado por Mindrinos, (1994) *Cell* 78, 1089-1099. Brevemente, se coexpresa la supuesta variante de R1 activa a partir de un plásmido con un gen derivado de patógeno que es un elicitador de la resistencia especificada por el supuesto homólogo de R1 homólogo y un gen indicador (por ejemplo, GUS). Si la variante se activa mediante la expresión continua del gen derivado de patógeno, entonces resultaría una HR y la actividad del gen indicador se suprimiría. Si no se inició la actividad, entonces sería detectable el gen indicador.

60 65 La similitud u homología entre la variante y R1 puede ser según se define y determina mediante el programa TBLASTN, de Altschul, (1990) *J. Mol. Biol.* 215, 403-10, que es de uso convencional en la técnica, o, y esto puede preferirse, el programa convencional BestFit, que es parte del paquete Wisconsin, versión 10, enero de 1999, (Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, EE.UU., Wisconsin 53711), que se ha usado para calcular homologías de secuencia en la presente solicitud. También puede usarse el software DNASTAR usando el método CLUSTAL con la tabla de pesos de residuos PAM250 (penalización por hueco 10, longitud de hueco 10). La homología (o similitud, o identidad) puede ser a nivel de la secuencia de nucleótidos y/o la secuencia de aminoácidos expresada. La secuencia de ácido nucleico y/o aminoácidos comparte homología con la secuencia codificante o la secuencia codificada por la secuencia de nucleótidos de SEQ ID NO:1 indicada en la reivindicación 1. La homología puede ser en

la longitud completa de la secuencia relevante mostrada en el presente documento, en comparación con la secuencia de aminoácidos o la secuencia de nucleótidos relevante como puede ser el caso.

Se cree que hay más de dos homólogos de *R1* en el genoma de la patata. Es probable que uno o más de estos homólogos sean genes *R* frente a virus, hongos, bacterias o nematodos.

5 Pueden aislar variantes de *R1* que se producen de manera natural, a la luz de la presente descripción, sin problemas a partir de cualquier planta adecuada. Pueden aislar las variantes de *R1* que se producen de manera natural a partir, por ejemplo, de ADN genómico o ADNc. Pueden obtenerse los supuestos genes de resistencia usando materiales (por ejemplo, cebadores o sondas) basados en regiones peculiares de *R1*, por ejemplo designadas en la figura 4. Tal como se trata en los ejemplos adjuntos, se espera que el gen *R1* identificado según la presente invención en patata defina una clase novedosa de genes de resistencia de plantas. Los genes correspondientes que codifican para proteínas que presentan propiedades similares deben estar presentes, por tanto, en otras plantas también. Pueden obtenerse las moléculas de ácido nucleico de la invención, por ejemplo, mediante hibridación de las moléculas de ácido nucleico descritas anteriormente con una (muestra de) molécula(s) de ácido nucleico de cualquier fuente. En general, pueden derivarse moléculas de ácido nucleico que se hibridan con las moléculas de ácido nucleico descritas anteriormente a partir de cualquier planta que tenga tales moléculas, preferiblemente a partir de plantas dicotiledóneas, en particular a partir de cualquier planta de interés en agricultura, horticultura o silvicultura, tales como plantas cultivadas, concretamente las de la familia Solanaceae, tales como patata y tomate pero también de plantas tales como mandioca, plantas leguminosas, plantas oleaginosas, tales como colza, linaza, etc., plantas que usan polipéptido como sustancias de almacenamiento, tales como soja, plantas que usan sacarosa como sustancia de almacenamiento, tales como remolacha azucarera o caña de azúcar, árboles, plantas ornamentales así como plantas que pueden usarse para la producción de biomasa, energía renovable, o materiales de construcción tales como ramio, etc.

10 15 20 25 Por tanto, un aspecto adicional de la presente invención proporciona un método de identificación y/o clonación de genes *R1* homólogos de una planta, empleando el método la totalidad o una parte de una secuencia de nucleótidos tal como se describió anteriormente. Puede usarse la información de secuencia de nucleótidos proporcionada en el presente documento en una búsqueda en base de datos (por ejemplo de EST, o STS, u otra información de secuencia genómica) para hallar secuencias homólogas, cuyos productos de expresión pueden someterse a prueba para determinar la actividad de resistencia a patógenos, por ejemplo, usando métodos basados en los ensayos transitorios de la presente invención, o ensayos de fenotipo convencionales en plantas transgénicas.

30 35 Alternativamente, pueden usarse sondas basadas en la secuencia, por ejemplo, en transferencia de tipo Southern. Por ejemplo, puede extraerse ADN de células tomadas de plantas que presentan el rasgo de resistencia apropiado y digerirse con diferentes enzimas de restricción. Entonces pueden separarse los fragmentos de restricción (por ejemplo, mediante electroforesis en un gel de agarosa) antes de la desnaturación y transferirse a un filtro de nitrocelulosa. Puede hibridarse la sonda marcada con los fragmentos de ADN sobre el filtro y determinarse la unión.

40 45 50 55 Pueden realizarse experimentos preliminares mediante hibridación en condiciones de baja rigurosidad. Para el estudio con sondas, son condiciones preferidas aquéllas que son lo suficientemente rigurosas para que haya un patrón sencillo con un pequeño número de hibridaciones identificadas como positivas que pueden investigarse adicionalmente. Por ejemplo, pueden realizarse hibridaciones usando una solución de hibridación que comprende: 5X SSC (en la que SSC = cloruro de sodio 0,15 M; citrato de sodio 0,15 M; pH 7), 5X reactivo de Denhardt, SDS al 0,5-1,0%, ADN de esperma de salmón fragmentado y desnaturizado 100 µg/ml, pirofosfato de sodio al 0,05% y formamida hasta al 50%. Se lleva a cabo la hibridación a 37-42°C durante al menos seis horas. Tras la hibridación, se lavan los filtros tal como sigue: (1) 5 minutos a temperatura ambiente en 2X SSC y SDS al 1%; (2) 15 minutos a temperatura ambiente en 2X SSC y SDS al 0,1%; (3) 30 minutos-1 hora a 37°C en 1X SSC y SDS al 1%; (4) 2 horas a 42-65°C en 1X SSC y SDS al 1%, cambiando la solución cada 30 minutos.

55 Una fórmula común para calcular las condiciones de rigurosidad requeridas para lograr la hibridación entre moléculas de ácido nucleico de una homología de secuencia especificada es (Sambrook *et al.*, 1989): $T_m = 81,5°C + 16,6\log [Na^+] + 0,41 (\%G+C)-0,63 (\% \text{ de formamida})-600/n.$ de pb en el dúplex. Como ilustración de la fórmula anterior, usando $[Na^+] = [0,368]$ y formamida al 50%, con un contenido en GC del 42% y un tamaño de sonda promedio de 200 bases, la T_m es de 57°C. La T_m de un dúplex de ADN disminuye en 1-1,5°C con cada disminución del 1% de la homología. Por tanto, se observarían dianas con una identidad de secuencia mayor del 75% usando una temperatura de hibridación de 42°C. Se considerará que una secuencia de este tipo es sustancialmente homóloga a la secuencia de ácido nucleico de la presente invención. Se conoce bien en la técnica cómo aumentar la rigurosidad de hibridación gradualmente hasta que sólo queden unos cuantos clones positivos. Otras condiciones adecuadas incluyen, por ejemplo para la detección de secuencias que son idénticas en aproximadamente el 80-90%, hibridación durante la noche a 42°C en Na_2HPO_4 0,25 M, pH 7,2, SDS al 6,5%, sulfato de dextrano al 10% y un lavado final a 55°C en 0,1X SSC, SDS al 0,1%. Para la detección de secuencias que son idénticas en más del 90%, las condiciones adecuadas incluyen hibridación durante la noche a 65°C en Na_2HPO_4 0,25 M, pH 7,2, SDS al 6,5%, sulfato de dextrano al 10% y un lavado final a 60°C en 0,1X SSC, SDS al 0,1%.

60 Puede medirse la unión de una sonda a ácido nucleico diana (por ejemplo, ADN) usando cualquiera de una variedad de técnicas a la disposición de los expertos en la técnica. Por ejemplo, las sondas pueden marcarse radiactiva, fluorescente o enzimáticamente. Otros métodos que no emplean el marcaje de sonda incluyen amplificación usando

PCR (incluyendo, cuando sea apropiado, RACE-PCR), protección de ARNasa y estudio con sondas de oligonucleótidos específicos de alelo.

5 Se sigue la identificación de una hibridación satisfactoria mediante el aislamiento del ácido nucleico que se ha hibridado, lo que puede implicar una o más etapas de PCR o amplificación mediante clonación en un vector que se replica en un huésped adecuado.

10 En cada caso, si hubiera necesidad pueden extenderse o complementarse los clones (por ejemplo, lambda, cósmido, plásmido, BAC, biBACS) o fragmentos identificados en la búsqueda. Por ejemplo si se sospecha que están incompletos, puede volver a visitarse la fuente de ADN original (por ejemplo una biblioteca de clones, preparación de ARNm, etc.) para aislar las partes que faltan, por ejemplo, usando secuencias, sondas o cebadores basados en esa parte que ya se ha obtenido para identificar otros clones que contienen una secuencia solapante (véase por ejemplo "Principles of Genome Analysis" de S B Primrose (1995) Pub. Blackwell Science Ltd, Oxford, RU).

15 Las moléculas de ácido nucleico o los genes correspondientes pueden someterse a prueba entonces para determinar su funcionalidad, por ejemplo, tal como se describe en los ejemplos. Un esquema para aislar homólogos de R1 es tal como sigue:

15 I) producir una población en la que se segregá un rasgo de resistencia.

20 II) Amplificar por PCR ADN de miembros individuales de la población con cebadores basados en la secuencia de R1 (pero no de los motivos conservados del gen *R*).

25 III) someter a prueba los productos de PCR (o bien mediante análisis directo de la secuencia o bien mediante digestión con enzimas de restricción) para determinar el polimorfismo de secuencia que se segregá conjuntamente con el rasgo R. Identificar una secuencia de marcador polimórfico apropiado.

IV) Aislar la secuencia codificante completa del gen polimórfico. Esto podría realizarse a partir de una biblioteca clonada apropiada o amplificándola usando cebadores de los extremos 5' y 3' de R1. En cada caso, el producto de PCR polimórfico identificado, o la información de secuencia proporcionada por el mismo, se usa para identificar el gen.

25 Entonces puede someterse a prueba la actividad de codificación del gen de resistencia tal como se describió anteriormente o se describe en los ejemplos.

Un enfoque más específico se basa en la comprensión de los que genes R1 homólogos pueden ligarse en agrupamientos. Ya se ha notificado el agrupamiento de genes *R* en patata (Leister *et al.* 1996; De Jong *et al.* 1997). Uno de los grandes agrupamientos de genes *R* se encuentra en el brazo corto del cromosoma V de la patata.

30 El punto caliente de resistencia en el cromosoma V de la patata que incluye *R1*, también contiene QTL (*Quantitative Trait Loci*, loci de rasgos cuantitativos) principales para la resistencia a *Phytophthora infestans* (Leonards-Schippers *et al.* 1994, Oberhagemann *et al.* 1999, Collins *et al.* 1999) y el nematodo quístico de la raíz *Globodera pallida* (Kreike *et al.* 1994, Rouppe van der Voort *et al.* 1997, 2000). El mapeo del desequilibrio de ligamiento reveló una fuerte asociación entre los marcadores en el intervalo de 0,8 cM *SPUD237 - GP179* que contiene *R1* y la resistencia del follaje y los tubérculos al ligamiento estrecho de soporte del tizón tardío entre *R1* y los factores que controlan la resistencia cuantitativa al tizón tardío. Se ha sugerido, basándose en el ligamiento genético observado, que *R1* y los factores que controlan resistencia cuantitativa al tizón tardío pueden ser alelos del mismo gen o miembros de una familia de genes agrupados (Leonards-Schippers *et al.* 1994, Oberhagemann *et al.* 1999). El primer análisis molecular del locus *R1* locus reveló ahora que la última opción es más favorable ya que *R1* es un miembro de una familia de genes y está presente como copia extra en una inserción de ADN en el cromosoma que lleva *R1*. Se ha notificado un hallazgo similar para el locus *Rpm1* en *Arabidopsis* (Stahl *et al.* 1999). El gen *R1* debe haberse sometido a introgresión en el genoma de *S. tuberosum* de la especie silvestre *S. demissum* a través de sobrecruzamiento cromosómico heterogenético. En cruzamientos entre especies de *Solanum* silvestres y cultivadas, se encuentra frecuentemente el apareamiento cromosómico heterogenético (Singh *et al.* 1989). Un segundo miembro altamente homólogo de la familia del gen *R1*, que tiene dos alelos *r1.1* y *r1.2*, está ubicado próximo físicamente a *R1*. Se requieren estudios adicionales sobre la funcionalidad de este gen. Estando disponible la secuencia de *R1*, pueden identificarse ahora otros miembros de la familia de *R1* que podrían estar presentes en aquellas partes del intervalo *GP21 - GP179* que aún no están cubiertas por el mapa físico y/o en otras partes del genoma de la patata. Pueden aislarse variantes alélicas en *S. tuberosum* y homólogos en otras especies de Solanaceae que están implicadas en la resistencia cuantitativa a *P. infestans*.

50 Por tanto, es una realización preferida de la presente invención que dicho patógeno al que es resistente una planta que expresa una molécula de ácido nucleico de la invención sea *Phytophthora infestans*.

55 La interacción entre *R1* y el patógeno del tizón tardío está de acuerdo con el concepto de gen por gen (Person *et al.* 1962, Flor 1971). La transferencia de un único gen fue suficiente para provocar en una planta huésped susceptible la respuesta de resistencia hipersensible con la infección con una variedad de *P. infestans* que lleva el gen de virulencia *Avr1* (todas las variedades excepto aquéllas con la especificidad de variedad 1). *Avr1* se segregó como un único factor dominante en la descendencia de cepas de *P. infestans* heterocigotas para *Avr1* y se mapeó en el grupo de ligamiento IV del mapa molecular de *P. infestans* (Van der Lee *et al.* 2001). No se ha clonado hasta la fecha ningún

factor de avirulencia de *P. infestans*. La caracterización adicional de *R1* a nivel molecular y la clonación del gen *Avr1* deberían contribuir a clarificar cómo reconoce la proteína de la resistencia la molécula efectora de avirulencia. La clonación de genes de resistencia al tizón tardío que reconocen factores de avirulencia diferentes de *Avr1* podrían permitir la identificación de los motivos moleculares que determinan la especificidad de reconocimiento de efectores y puede ayudar a modificar mediante ingeniería genética proteínas *R* con resistencia al tizón tardío más amplia y más duradera. Pueden aislarse otras variantes de *R1* ligadas (que proporcionan diferentes rasgos *R*) esencialmente tal como se expuso anteriormente, pero en las que el ADN usado para la etapa de amplificación inicial se toma de miembros de la población en la que el rasgo *R* requerido se segregó conjuntamente con el propio *R1* (o una variante de *R1*).

Los presentes inventores han observado que la secuencia de *R1* es similar a la secuencia del gen *Prf* por lo demás no relacionado que confiere resistencia en el tomate frente a un patógeno bacteriano, es decir *P. syringae* (Salmeron *et al.* 1996). A la luz de esta información, parece que la secuencia de *R1* podría modificarse, por ejemplo, mediante mutación dirigida al sitio o al azar, para producir mutantes de *R1* u otros derivados que pueden conferir resistencia frente a (es decir se activan mediante) patógenos que son bastante diferentes de *P. infestans*. Esto puede lograrse tal como se describe a continuación, sometiéndose a prueba los mutantes de *R1* con los métodos de ensayo de expresión transitoria descritos anteriormente.

Preferiblemente, se genera la molécula de ácido nucleico que es el mutante u otro derivado o bien directamente o bien indirectamente (por ejemplo mediante una o etapas de amplificación o replicación) a partir de un ácido nucleico original correspondiente a la totalidad o una parte de la secuencia mostrada en SEQ ID NO:1 u otras secuencias dadas a conocer en el presente documento.

Por tanto, se da a conocer adicionalmente un método de producción de un ácido nucleico que codifica para un derivado de *R1* que comprende la etapa de modificar una molécula de ácido nucleico que codifica para *R1*. El derivado puede incluir cambios a la molécula de ácido nucleico que no influyen en la secuencia de aminoácidos codificada (es decir, equivalente de manera degenerada). Los cambios a una secuencia, para producir un mutante o derivado, pueden ser mediante uno o más de adición, inserción, delección o sustitución de uno o más nucleótidos en el ácido nucleico, lo que conduce a la adición, inserción, delección o sustitución de uno o más aminoácidos en el polipéptido codificado. Además de uno o más cambios dentro de la secuencia de *R1*, un ácido nucleico variante puede codificar para una secuencia de aminoácidos que incluye aminoácidos adicionales en el extremo C-terminal y/o el extremo N-terminal.

Están incluidos específicamente partes o fragmentos (sin embargo producidos) correspondientes a partes de las secuencias proporcionadas, y que codifican para polipéptidos que tienen actividad biológica, por ejemplo resistencia a patógenos o la capacidad para producir o unirse a anticuerpos de unión a *R1*.

En términos generales, pueden ser deseables los cambios por varios motivos, incluyendo introducir o eliminar las siguientes características: secuencias de endonucleasas de restricción; utilización de codones; otros sitios que se requieren para la modificación postraduccional; sitios de escisión en el polipéptido codificado; motivos en el polipéptido codificado para glicosilación, lipoilación etc. Pueden añadirse secuencias líder u otras de direccionamiento a la proteína expresada para determinar su ubicación tras la expresión. Todos estos pueden ayudar eficazmente en la clonación y expresión de un polipéptido activo en forma recombinante (tal como se describe a continuación). Modificaciones preferidas incluyen las que disminuyen la carga negativa neta de la región en o alrededor de los motivos QLPL, CFLY o LHD. El experto en la técnica conoce medios y métodos sobre cómo modificar genes resistentes y se describen, por ejemplo en el documento WO 01/29239 para el gen Rx de *Solanum tuberosum*. Otra mutación deseable puede ser mutagénesis dirigida al sitio o al azar para alterar la actividad (por ejemplo, especificidad) o estabilidad del polipéptido codificado.

Tal como se entiende bien, se determina la homología a nivel de aminoácidos en cuanto a la similitud o identidad de aminoácidos. La similitud permite la variación conservativa, es decir la sustitución de un residuo hidrófobo tal como isoleucina, valina, leucina o metionina por otro, o la sustitución de un residuo polar por otro, tal como arginina por lisina, ácido glutámico por aspártico, o glutamina por asparagina. Tal como conocen bien los expertos en la técnica, la alteración de la estructura primaria de un polipéptido mediante una sustitución conservativa puede no alterar significativamente la actividad de ese péptido porque la cadena lateral del aminoácido que se inserta en la secuencia puede ser capaz de formar enlaces y contactos similares que la cadena lateral del aminoácido que se ha eliminado en la sustitución. Esto es así incluso cuando la sustitución es en una región que es crítica en la determinación de la conformación de los péptidos.

También están incluidos homólogos que tienen sustituciones no conservativas. Como conocen bien los expertos en la técnica, las sustituciones en regiones de un péptido que no son críticas en la determinación de su conformación pueden no afectar mucho a su actividad porque no alteran mucho la estructura tridimensional del péptido. En las regiones que son críticas en la determinación de la actividad o la conformación de los péptidos, tales cambios pueden alterar las propiedades del polipéptido. En efecto, cambios tales como los descritos anteriormente pueden conferir propiedades ligeramente ventajosas en el péptido, por ejemplo, estabilidad o especificidad alterada, en particular especificidad más amplia.

Entonces pueden seleccionarse mutantes que tienen estas propiedades, tal como se describió anteriormente.

5

Otros métodos pueden incluir mezclar o incorporar secuencias de genes de resistencia relacionados en la secuencia de *R1*. Por ejemplo, podrían ligarse fragmentos de enzima de restricción de *R1* junto con fragmentos de un homólogo de *R1* o incluso de un gen no relacionado para generar versiones recombinantes de *R1*. Una estrategia alternativa para modificar *R1* emplearía la PCR tal como se describió anteriormente (Ho *et al.*, 1989 Gene 77, 51-59) o intercambio de ADN (Crameri *et al.*, 1998 Nature 391).

10

Por tanto, los métodos descritos anteriormente, pueden incluir hibridación de uno o más (por ejemplo dos) sondas o cebadores basados en la secuencia de *R1* o bien para seleccionar homólogos de *R1* o bien para producir derivados de *R1*. Tales, oligonucleótidos, sondas o cebadores forman una parte adicional de la presente descripción. Un oligonucleótido para su uso en estudio con sondas o PCR puede ser de aproximadamente 30 o menos nucleótidos de longitud (por ejemplo 18, 21 ó 24). Generalmente cebadores específicos son de más de 14 ó 15 nucleótidos de longitud. Para una especificidad y rentabilidad óptimas, pueden preferirse cebadores de 16-24 nucleótidos de longitud. Los expertos en la técnica están muy versados en el diseño de cebadores para procesos de uso tales como PCR. Si se requiere, puede realizarse el estudio con sondas con fragmentos de restricción completos del gen dado a conocer en el presente documento que puede ser de cientos o incluso miles de nucleótidos de longitud.

15

En un aspecto de la presente invención, la molécula de ácido nucleico descrita anteriormente está en forma de un vector recombinante y preferiblemente replicable.

20

“Vector” se define para incluir, entre otros, cualquier plásmido, cósmido, fago o vector binario de *Agrobacterium* en forma circular o lineal, mono o bicatenaria que puede ser o no autotransmisible o movilizable, y que puede transformar un huésped procariota o eucariota mediante la integración en el genoma celular o existir de manera extracromosómica (por ejemplo plásmido de replicación autónoma con un origen de replicación). Están incluidos específicamente vectores lanzadera, mediante lo que se quiere decir un vehículo de ADN que puede, de manera natural o por diseño, dar la replicación en dos organismos huésped diferentes, que pueden seleccionarse de actinomicetos y especies relacionadas, bacterias y eucariotas (por ejemplo células de plantas superiores, mamíferos, levaduras o fúngicas).

25

No es necesario que un vector que incluye ácido nucleico según la presente invención incluya un promotor u otra secuencia reguladora, particularmente si el vector va a usarse para introducir el ácido nucleico en células para la recombinación en el genoma.

30

Preferiblemente, el ácido nucleico en el vector está bajo el control de, y operativamente unido a, un promotor apropiado u otros elementos reguladores para la transcripción en una célula huésped tal como un microbio, por ejemplo célula bacteriana o vegetal. El vector puede ser un vector de expresión bifuncional que funciona en múltiples huéspedes. En el caso de ADN genómico, éste puede contener su propio promotor u otros elementos reguladores y en el caso de ADNc éste puede estar bajo el control de un promotor apropiado u otros elementos reguladores para la expresión en la célula huésped.

35

Por “promotor” se entiende una secuencia de nucleótidos a partir de la que puede iniciarse la transcripción de ADN operativamente unido aguas abajo (es decir en el sentido de 3' en la cadena homosentido del ADN bicatenario).

“Operativamente unido” significa unido como parte de la misma molécula de ácido nucleico, situado de manera adecuada y orientado para que se inicie la transcripción desde el promotor. ADN operativamente unido a un promotor está “bajo la regulación de inicio de la transcripción” del promotor.

40

Por tanto, este aspecto proporciona un constructo génico, preferiblemente un vector replicable, que comprende un promotor operativamente unido a una secuencia de nucleótidos proporcionada por la presente invención, tal como la región codificante del gen *R1*, o una variante (por ejemplo, mutante, derivado o alelo) de la misma. En términos generales, los expertos en la técnica con muy capaces de construir vectores y protocolos de diseño para la expresión génica recombinante. Pueden elegirse o construirse vectores adecuados, que contienen secuencias reguladoras apropiadas, incluyendo secuencias promotoras, fragmentos de terminación, secuencias de poliadenilación, secuencias potenciadoras, genes marcadores y otras secuencias según sea apropiado. Para detalles adicionales, véase, por ejemplo, Molecular Cloning: a Laboratory Manual: 2^a edición, Sambrook *et al.*, 1989, Cold Spring Harbor Laboratory Press. Se describen muchas técnicas y protocolos para la manipulación de ácido nucleico conocidos, por ejemplo en la preparación de constructos de ácido nucleico, mutagénesis (véase anteriormente), secuenciación, introducción de ADN en células y expresión génica, y análisis de proteínas, en Current Protocols in Molecular Biology, segunda edición, Ausubel *et al.* eds., John Wiley & Sons, 1992. Las descripciones de Sambrook *et al.* y Ausubel *et al.* Se incorporan al presente documento como referencia.

50

En una realización de este aspecto de la presente invención proporciona un constructo génico, preferiblemente un vector replicable, que comprende un promotor inducible operativamente unido a una secuencia de nucleótidos proporcionada por la presente invención.

55

El término “inducible” tal como se aplica a un promotor lo entienden bien los expertos en la técnica. En esencia, la expresión bajo el control de un promotor inducible se “activa” o aumenta en respuesta a un estímulo aplicado. La naturaleza del estímulo varía entre promotores. Algunos promotores inducibles provocan bajos niveles o niveles indetectables de expresión (o la no expresión) en ausencia del estímulo apropiado. Otros promotores inducibles

provocan la expresión constitutiva detectable en ausencia del estímulo. Cualquiera que sea el nivel de expresión en ausencia del estímulo, la expresión de cualquier promotor inducible aumenta en presencia del estímulo correcto. La situación preferible es en la que el nivel de expresión aumenta con la aplicación del estímulo relevante en una cantidad eficaz para alterar una característica fenotípica. Por tanto, puede usarse un promotor inducible (o "activable") que produce un nivel básico de expresión en ausencia del estímulo cuyo nivel es demasiado bajo para provocar un fenotipo deseado (y de hecho puede ser cero). Con la aplicación del estímulo, se aumenta la expresión (o se activa) hasta un nivel que provoca el fenotipo deseado.

Los posibles elementos reguladores que permiten la expresión en células huésped procariotas comprenden, por ejemplo, el promotor P_L , *lac*, *trp* o *tac* en *E. coli*, y ejemplos de elementos reguladores que permiten la expresión en células huésped eucariotas son el promotor *AOX1* o *GAL1* en levadura o el promotor de CMV, SV40, VSR (virus de sarcoma de Rous), potenciador de CMV, potenciador de SV40 o un intrón de globina en células de mamífero y otros animales. En este contexto, se conocen en la técnica vectores de expresión adecuados tales como vector de expresión de ADNC de Okayama-Berg pcDV1 (Pharmacia), pCDM8, pRc/CMV, pDNAc1, pDNAc3 (In-vitrogen), pSPORT1 (GIBCO BRL).

Son particularmente de interés en el presente contexto los vectores de plantas. Se describen procedimientos específicos y vectores usados previamente con gran éxito en plantas por Bevan (Nucl. Acids Res. 12,8711-8721 (1984)) y Guerineau y Mullineaux (1993) (Plant transformation and expression vectors. En: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, págs. 121-148). Los promotores adecuados que operan en plantas incluyen el promotor del gen del virus del mosaico de la coliflor 35S (VMCo 35S) que se expresa en un alto nivel en prácticamente todos los tejidos vegetales (Benfey *et al*, 1990a y 1990b); el promotor meri 5 de la coliflor que se expresa en el meristemo apical vegetativo así como varias posiciones bien localizadas en el cuerpo vegetal, por ejemplo el floema interno, primordio floral, puntos de ramificación en la raíz y los brotes (Medford, 1992; Medford *et al*, 1991) y el promotor LEAFY de *Arabidopsis thaliana* que se expresa de forma muy temprana en el desarrollo floral (Weigel *et al*, 1992). Otros promotores incluyen el promotor de actina del arroz.

El promotor puede incluir uno o más motivos de secuencia o elementos que confieren un control de la expresión regulador en el desarrollo y/o específico de tejido.

Por tanto, los vectores dados a conocer pueden incluir el gen *R1* o una variante del mismo, además de diversas secuencias requeridas para proporcionarles funcionalidad de replicación, integración y/o expresión. Tales vectores pueden usarse, por ejemplo, para preparar plantas en las que se introduce resistencia a *P. infestans* u otros hongos.

Si se desea inducir resistencia de más amplio espectro, están disponibles diversas opciones adicionales a la luz de la presente descripción:

(a) Modificar la secuencia de *R1*, para producir mutantes u otros derivados tal como se trató anteriormente, de manera que su efecto pueden iniciar elicitores o patógenos distintos de *P. infestans* solo o los otros elicidores naturales tratados en el presente documento.

(b) Coexpresar *R1* directamente con un elicitor apropiado (por ejemplo Avr 1 de una cepa avirulenta).

(c) Coexpresar *R1* y un gen elicitor, cuya transcripción o traducción está suprimida por la activación de *R1*.

Esto volvería a acoplar *R1* con su elicitor, e imitaría mejor la respuesta natural a la infección por *P. infestans* que da como resultado un amplio silenciamiento de la especificidad.

(d) Coexpresar *R1* con un gen elicitor, cuya traducción sólo se activa en presencia de patógeno(s).

(e) Coexpresar *R1* con un gen elicitor, mediante lo cual se inactivan uno o ambos, y reactivar el/los gen(es) de manera variada, de modo que se limita la HR sólo a determinados sectores de la planta (por ejemplo, sectores definidos somáticamente) pero mientras que la respuesta defensiva se extiende más allá de estos sectores. Esto podría lograrse, por ejemplo, en analogía a los métodos dados a conocer en el documento WO95/31564, en los que, tras un retrocruzamiento entre una planta que lleva un gen de resistencia marcado con transposición (en ese caso cf-9) más el elicitor intacto (Avr-9) y una planta que lleva una transposasa activadora, la progenie mostró una reactivación somática del cf-9, que condujo a una respuesta necrótica localizada pero resistencia extendida.

Además de los vectores y constructos anteriores, la presente invención también proporciona métodos que comprenden introducir los constructos de *R1* tratados anteriormente (tales como vectores) en una célula huésped y/o inducir la expresión de un constructo dentro de una célula vegetal, mediante la aplicación de un estímulo adecuado, un inductor exógeno eficaz. Los vectores descritos anteriormente pueden introducirse en huéspedes mediante cualquier método apropiado, por ejemplo, conjugación, movilización, transformación, transfección, transducción o electroporación, tal como se describe en mayor detalle a continuación.

En un aspecto adicional de la invención, se da a conocer una célula huésped que contiene ácido nucleico o un vector según la presente invención, especialmente una célula vegetal o microbiana. La célula huésped puede ser

cualquier célula procariota o eucariota, tal como células bacterianas, de insecto, vegetales o animales. Células fúngicas preferidas son, por ejemplo, las del género *Saccharomyces*, en particular las de la especie *S. cerevisiae*.

Para la expresión de las moléculas de ácido nucleico según la invención en orientación homosentido o antisentido en células vegetales, las moléculas se ponen bajo el control de elementos reguladores que garantizan la expresión en células vegetales. Estos elementos reguladores pueden ser heterólogos u homólogos con respecto a la molécula de ácido nucleico que va a expresarse así como con respecto a la especie vegetal que va a transformarse. En general, tales elementos reguladores comprenden un promotor activo en células vegetales. Para obtener la expresión en todos los tejidos de una planta transgénica, se usan preferiblemente promotores constitutivos, tales como el promotor 35 S de VMCo (Odell, *Nature* 313 (1985), 810-812) o promotores de los genes de poliubiquitina del maíz (Christensen, *Plant Mol. Biol.* 18 (1982), 675-689). Para lograr la expresión en tejidos específicos de una planta transgénica, es posible usar promotores específicos de tejido (véase, por ejemplo, Stockhaus, *EMBO J.* 8 (1989), 2245-2251). También se conocen promotores que son específicamente activos en tubérculos de patatas o en semillas de diferentes especies vegetales, tales como maíz, *Vicia*, trigo, cebada, etc. pueden usarse promotores inducibles para poder controlar exactamente la expresión. Un ejemplo de promotores inducibles son los promotores de genes que codifican para proteínas de choque térmico. También se han descrito elementos reguladores específicos de microesporas y sus usos (documento WO96/16182). Además, puede emplearse el sistema Tet químicamente inducible (Gatz, *Mol. Gen. Genet.* 227 (1991); 229-237). El experto en la técnica conoce promotores adecuados adicionales y se describen, por ejemplo, en Ward (*Plant Mol. Biol.* 22 (1993), 361-366). Los elementos reguladores pueden comprender además potenciadores transcripcionales y/o traducionales en células vegetales. Además, los elementos reguladores pueden incluir señales de terminación de la transcripción, tales como una señal de poli-A, que conduce a la adición de una cola de poli-A al transcripto que puede mejorar su estabilidad; para bibliografía véase también lo citado anteriormente.

En el caso de que una molécula de ácido nucleico según la invención se exprese en la orientación homosentido, es en principio posible modificar la secuencia codificante de tal manera que la proteína esté ubicada en cualquier compartimento deseado de la célula vegetal. Éstos incluyen el retículo endoplasmático, la vacuola, la mitocondria, los plástidos, el apoplasto, el citoplasma, etc. El experto en la técnica conoce bien métodos sobre cómo llevar a cabo estas modificaciones y secuencias señal que garantizan la localización en un compartimento deseado.

También se conocen bien en la técnica métodos para la introducción de ADN foráneo en plantas. Estos incluyen, por ejemplo, la transformación de células o tejidos vegetales con ADN-T usando *Agrobacterium tumefaciens* o *Agrobacterium rhizogenes*, la fusión de protoplastos, transferencia de genes directas (véase, por ejemplo, el documento EP-A 164 575), inyección, electroporación, métodos biolíticos tales como bombardeo con partículas y otros métodos conocidos en la técnica. Los vectores usados en el método de la invención pueden contener elementos funcionales adicionales, por ejemplo secuencias de "borde izquierdo" y de "borde derecho" del ADN-T de *Agrobacterium* que permiten la integración estable en el genoma de la planta. Además, el experto en la técnica conoce métodos y vectores que permiten la generación de plantas transgénicas libres de marcador, es decir el gen marcador seleccionable o puntuable se pierde en una determinada fase del desarrollo de la planta o el cultivo de la planta. Esto puede lograrse, por ejemplo, mediante cotransformación (Lyznik, *Plant Mol. Biol.* 13 (1989), 151-161; Peng, *Plant Mol. Biol.* 27 (1995), 91-104) y/o usando sistemas que utilizan enzimas que pueden fomentar la recombinación homóloga en plantas (véase, por ejemplo, el documento WO97/08331; Bayley, *Plant Mol. Biol.* 18 (1992), 353-361; Lloyd, *Mol. Gen. Genet.* 242 (1994), 653-657; Maeser, *Mol. Gen. Genet.* 230 (1991), 170-176; Onouchi, *Nucl. Acids Res.* 19 (1991), 6373-6378). Se describen métodos para la preparación de vectores apropiados, por ejemplo, por Sambrook (*Molecular Cloning: A Laboratory Manual*, 2^a edición (1989), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

Los expertos en la técnica conocen bien cepas adecuadas de *Agrobacterium tumefaciens* y vectores así como la transformación de *Agrobacterium* y medios de crecimiento y selección apropiados y se describen en la técnica anterior (GV3101 (pMK90RK), Koncz, *Mol. Gen. Genet.* 204 (1986), 383-396; C58C1 (pGV 3850kan), Deblaere, *Nucl. Acid Res.* 13 (1985), 4777; Bevan, *Nucleic Acid Res.* 12(1984), 8711; Koncz, *Proc. Natl. Acad. Sci. USA* 86 (1989), 8467-8471; Koncz, *Plant Mol. Biol.* 20 (1992), 963-976; Koncz, Specialized vectors for gene tagging and expression studies. En: *Plant Molecular Biology Manual* vol. 2, Gelvin y Schilperoort (Eds.), Dordrecht, Países Bajos: Kluwer Academic Publ. (1994), 1-22; documento EP-A-120 516; Hoekema: *The Binary Plant Vector System*, Offsetdrukkerij Kanters B.V., Albllasserdam (1985), capítulo V, Fraley, *Crit. Rev. Plant. Sci.*, 4, 1-46; An, *EMBO J.* 4 (1985), 277-287). Aunque se prefiere el uso de *Agrobacterium tumefaciens* en el método de la invención, pueden usarse otras cepas de *Agrobacterium*, tales como *Agrobacterium rhizogenes*, por ejemplo si se desea un fenotipo conferido por dicha cepa.

El experto en la técnica conoce bien métodos para la transformación usando métodos biolíticos; véase, por ejemplo, Wan, *Plant Physiol.* 104 (1994), 37-48; Vasil, *Bio/Technology* 11 (1993), 1553-1558 y Christou (1996) *Trends in Plant Science* 1, 423-431. Puede realizarse la microinyección tal como se describe en Potrykus y Spangenberg (eds.), *Gene Transfer To Plants*. Springer Verlag, Berlín, NY (1995).

La transformación de la mayoría de las plantas dicotiledóneas es posible con los métodos descritos anteriormente. Pero también se han desarrollado varias técnicas de transformación satisfactorias para la transformación de plantas monocotiledóneas. Estas incluyen la transformación usando métodos biolíticos tales como, por ejemplo, los descritos anteriormente así como la transformación de protoplastos, electroporación de células parcialmente permeabilizadas, introducción de ADN usando fibras de vidrio, etc. Entonces puede usarse la célula vegetal transformada resultante para regenerar una planta transformada de una manera conocida por un experto en la técnica.

Esto puede encontrarse, por ejemplo, en Hood, Molecular Breeding 3 (1997), 291-306; Coleman, Proc. Natl. Acad. Sci. USA 94 (1997), 7094-7097; Shilto, Biotechnology 7 (1989), 581-587.

En general, las plantas que pueden modificarse según la invención y que o bien muestran sobreexpresión de una proteína según la invención o bien una reducción de la síntesis de una proteína de este tipo pueden derivarse de cualquier especie vegetal deseada. Pueden ser plantas monocotiledóneas o plantas dicotiledóneas, preferiblemente pertenecen a una especie vegetal de interés en agricultura, silvicultura u horticultura, tal como plantas cultivadas (por ejemplo maíz, arroz, cebada, trigo, centeno, avenas etc.), patatas, plantas oleaginosas (por ejemplo colza, girasol, cacahuate, semilla de soja, etc.), algodón, remolacha azucarera, caña de azúcar, plantas leguminosas (por ejemplo judías, guisantes etc.), plantas productoras de madera, preferiblemente árboles, etc.

La elección particular de una tecnología de transformación se determinará mediante su eficacia para transformar una determinada especie vegetal así como la experiencia y preferencia de la persona que pone en práctica la invención con una metodología de elección particular. Resultará evidente para el experto en la técnica que la elección particular de un sistema de transformación para introducir ácido nucleico en células vegetales no es esencial para, o una limitación de la invención, ni lo es la elección de la técnica para la regeneración de la planta. Si se desea, pueden usarse marcadores genéticos seleccionables que consisten en genes químicos que confieren fenotipos seleccionables tales como resistencia a antibióticos tales como kanamicina, higromicina, fosfinotricina, clorsulfuron, metotrexato, gentamicina, espectinomicina, imidazolinonas y glifosato.

Por tanto un aspecto adicional proporciona un método de transformación de una célula vegetal que supone la introducción de un vector que comprende un ácido nucleico de la presente invención (por ejemplo *R1* o variante de *R1*) en una célula vegetal y provocar o permitir la recombinación entre el vector y el genoma de la célula vegetal para introducir la secuencia de nucleótidos en el genoma.

La invención abarca además una célula huésped transformada con una molécula de ácido nucleico o un vector según la presente invención, especialmente una célula vegetal o microbiana. En la célula vegetal transgénica (es decir transgénica para el ácido nucleico en cuestión) el transgén puede estar en un vector extragenómico o incorporarse, preferiblemente de manera estable, en el genoma. Puede haber más de una secuencia de nucleótidos heteróloga por genoma haploide.

El término "heterólogo" se usa ampliamente en este aspecto para indicar que el gen/la secuencia de nucleótidos en cuestión se han introducido en dichas células de la planta o un ancestro de las mismas, usando ingeniería genética, es decir mediante intervención humana. Un gen heterólogo puede ser adicional a un gen endógeno correspondiente. Ácido nucleico heterólogo, o exógeno o foráneo, a una célula vegetal puede no producirse de manera natural en células de ese tipo, variedad o especie. Por tanto, el ácido nucleico heterólogo puede comprender una secuencia codificante de, o derivada de, un tipo particular de célula vegetal o especie o variedad vegetal, colocada dentro del contexto de una célula vegetal de un tipo o especie o variedad vegetal diferentes.

Tras la transformación, puede regenerarse una planta, por ejemplo a partir de células individuales, tejido de callo o discos de hojas, tal como es convencional en la técnica. Casi cualquier planta puede regenerarse completamente a partir de células, tejidos y órganos de la planta. Se revisan técnicas disponibles en Vasil *et al.*, Cell Culture and Somatic Cell Genetics of Plants, Vol I, II y III, Laboratory Procedures and Their Applications, Academic Press, 1984, y Weissbach y Weissbach, Methods for Plant Molecular Biology, Academic Press, 1989.

La generación de plantas transgénicas fértiles se ha logrado en los cereales arroz, maíz, trigo, avena y cebada (revisado en Shimamoto, K. (1994) Current Opinion in Biotechnology 5,158-162.; Vasil, *et al.* (1992) Bio/Technology 10,667-674; Vain *et al.*, 1995, Biotechnology Advances 13 (4): 653-671; Vasil, 1996, Nature Biotechnology 14 página 702).

También se proporcionan plantas que incluyen una célula vegetal según la invención, junto con cualquier parte o propágulo de las mismas, semilla, progenie y descendientes propios o híbridos. Una planta según la presente invención puede ser una cuyos descendientes no son idénticos en una o más propiedades. Pueden excluirse variedades vegetales, particularmente variedades vegetales registrables según Plant Breeders' Rights. Se observa que una planta no necesita considerarse una "variedad vegetal" simplemente porque contenga estable dentro de su genoma un transgén, introducido en una célula de la planta o un ancestro de la misma.

En una realización preferida de la invención, la planta transgénica de la invención con la presencia del gen *R1* de la invención obtuvo una resistencia o resistencia mejorada frente a un patógeno al que era susceptible la planta silvestre correspondiente.

El término "resistencia" cubre la gama de protección desde un retardo hasta la inhibición completa del desarrollo de la enfermedad. Los ejemplos de patógenos de importancia comprenden *Phytophthora infestans*, el agente causante de la enfermedad de tizón tardío de la patata, *Phytophthora sojae*, patógeno de podredumbre de la raíz de la soja, *Peronospora parasitica* (mildiú), *Magnaporthe grisea*, agente causante de la enfermedad de quemado del arroz, *Erysiphe* spp (óidio), *Pseudomonas syringae* (agente de tizón bacteriano), *Erwinia amylovora* (enfermedad de fuego bacteriano), *Erwinia carotovora* (podredumbre húmeda), *Botrytis cinerea* (mildiú de la vid), *Rhizoctonia solani* y *Pythium*

debaryanum (agentes de tizón de semillas o enfermedad de mal de los semilleros). Preferiblemente, la planta transgénica de la invención obtiene resistencia frente a *P. infestans*.

Además de la planta regenerada, la presente invención abarca todo de lo siguiente: un clon de tal planta, semilla, descendientes y progenie propios o híbridos (por ejemplo descendientes F1 y F2) y cualquier parte de cualquiera de los mismos, tal como esquejes, semillas. La invención también proporciona un propágulo vegetal a partir de tal planta, que es cualquier parte que puede usarse en la reproducción o propagación, sexual o asexual, incluyendo esquejes, semilla etcétera.

Como alternativa a los métodos basados en biología molecular de introducción de *R1* (o variantes del mismo) en plantas, las secuencias dadas a conocer en el presente documento pueden usarse para facilitar la selección de plantas en las que se desea introducir el rasgo de resistencia usando métodos de reproducción de plantas convencionales. Puede identificarse fácilmente la progenie de cruces que llevan el gen examinando basándose en la secuencia de *R1*, particularmente la secuencia distintiva de *R1*.

Los métodos dados a conocer en el presente documento para identificar marcadores proximales al locus de *R1* pueden ser generalmente aplicables a otros genes encontrados en agrupaciones (por ejemplo genes de resistencia derivados de plantas). Tales métodos se caracterizan porque emplean una etapa que usa PCR de baja rigurosidad con cebadores no degenerados que evitan motivos de secuencia conservados. El enfoque general puede resumirse de la siguiente manera: (a) preparar una población en la que se segregá el gen de interés, (b) identificar homólogo(s) del/de los gen(es) de resistencia unido(s) al locus de interés basándose en motivos (de gen de resistencia) altamente conservados y cebadores altamente degenerados (Leister *et al.*, 1996) *Nature Genet.* 14,421-428, (c) identificar marcadores adicionales correspondientes a genes homólogos, que están dentro del locus (de resistencia) y que están más cerca del gen, usando PCR de baja rigurosidad con cebadores no degenerados que evitan motivos de secuencia conservados, (d) usar dichos marcadores adicionales para identificar un clon que lleva la biblioteca genómica del gen (de resistencia) de interés de una planta resistente, opcionalmente junto con ensayos transitorios para determinar la actividad (Mindrinos *et al* (1994) o tal como se describe en el presente documento), (e) opcionalmente, confirmar la identidad del gen clonado basándose en el fenotipo en plantas transgénicas.

La presente descripción también abarca el producto de expresión de cualquiera de las secuencias de ácido nucleico de *R1* o variante dadas a conocer anteriormente, y métodos de preparación del producto de expresión mediante expresión a partir de moléculas de ácido nucleico codificantes por tanto en condiciones adecuadas, que puede ser en células huésped adecuadas *in vitro*, o sintetizado químicamente, en particular si se desean antígenos para preparar anticuerpos.

Pueden prepararse anticuerpos frente a un péptido o polipéptido *R1*/variante purificado mediante cualquier método conocido en la técnica (para un resumen, véase por ejemplo "Immunology- 5^a edición" de Roitt, Brostoff, Male: Pub 1998-Mosby Press, Londres). Tales anticuerpos, o fragmentos o derivados de los mismos, pueden usarse para unirse a *R1* o en la identificación y/o el aislamiento de proteínas homólogas a *R1* (es decir que comparten epítopos con la misma), que a su vez pueden proporcionar la base de un método alternativo a los descritos anteriormente para aislar sus genes codificantes.

Asimismo, pueden emplearse aptómeros que se unen al polipéptido *R1* de la invención. La preparación de aptómeros la conoce el experto en la técnica; véase, por ejemplo, Thomas, y Dinshaw (2000) *Adaptive recognition by nucleic aptamers.* *Science* 287:820-825.

La invención proporciona además un método para influir o afectar un rasgo de resistencia en una planta, mediante el cual el método incluye la etapa de provocar o permitir la expresión de una secuencia de ácido nucleico heteróloga tal como se comentó anteriormente (por ejemplo *R1* o variante de *R1*, en cada caso, más un elicitor opcional) dentro de células de la planta.

Como alternativa, puede ser deseable regular por disminución la actividad de *R1*. Esto puede lograrse por ejemplo usando tecnología antisentido (que se revisa en Bourque, (1995), *Plant Science* 105,125-149, y Flavell, (1994) *PNAS USA* 91, 3490-3496). Una alternativa a la tecnología antisentido es usar una copia de la totalidad o una parte del gen diana insertado en sentido, que es la misma orientación que el gen diana, para lograr la reducción en la expresión del gen diana mediante cosupresión; véase, por ejemplo, van der Krol *et al.*, (1990) *The Plant Cell* 2, 291-299; Napoli *et al.*, (1990) *The Plant Cell* 2, 279-289; Zhang *et al.*, (1992) *The Plant Cell* 4, 1575-1588, y el documento US-A-5,231,020.

Por tanto, un aspecto adicional también se refiere a una célula vegetal transgénica (y a plantas transgénicas que comprenden tales células vegetales) que contiene, preferiblemente integrada de manera estable en el genoma, una molécula de ácido nucleico según la invención o parte de la misma, en la que la transcripción y/o expresión de la molécula de ácido nucleico o parte de la misma conduce a la reducción de la síntesis de una proteína *R1*. En una realización preferida, la reducción se logra mediante una cadena antisentido, cadena sentido, ribozima, cosupresión, efecto mutante dominante, o mutante desactivado en el gen *R1*.

Sin embargo, preferiblemente la descripción proporciona un método que incluye expresar SEQ ID NO: 1 o una variante de la misma dentro de las células de una planta (produciendo así el polipéptido codificado), tras una etapa anterior de introducción del ácido nucleico en una célula de la planta o un ancestro de la misma. Generalmente tal

método puede usarse para introducir resistencia fúngica en la planta mediante lo cual se activa una resistencia mediada por *R1* mediante contacto con un elicitor fúngico apropiado y otro incitador o inductor. En general, el elicitor u otro activador puede codificarse directamente por los hongos invasores (tal como la proteína de virulencia de *P. infestans* u otros determinados hongos).

5 Alternativamente puede expresarse mediante un transgén o constructo separado que a su vez se activa o se regula por incremento mediante la infección fúngica. Adicionalmente, en ambos de esos casos, la modificación de la secuencia de *R1* (variante) puede permitir activarse mediante un elicitor no natural, si se prefiere.

10 Los formatos descritos anteriormente para evaluar la función de *R1* o derivado de *R1* con respecto a un elicitor supuesto o conocido, forman por sí mismos un aspecto adicional de la presente invención. En particular los métodos, para establecer compatibilidad gen a gen entre el elicitor y el gen de resistencia, se caracterizan porque incluyen las etapas de: (a) provocar o permitir la co-expresión en la célula de *R1* o un derivado de *R1* con el elicitor, (b) observar dicha célula para detectar una HR, (c) correlacionar el resultado de la observación realizada en (b) con la especificidad del *R1* o el derivado de *R1* para el elicitor.

15 Según lo anterior, la presente descripción también se refiere a tales plantas transgénicas que son más sensibles a la infección por tizón tardío en comparación con una planta silvestre correspondiente. Asimismo, la presente descripción se refiere a partes que pueden cosecharse y material de propagación de tales plantas.

20 Tal como se describe en los ejemplos, se ha aislado un gen de *R1* que tras la transformación en un cultivar de planta susceptible *Desireé* confirió resistencia frente a *P. infestans*. Dado que el clon genómico cuya secuencia de ADN correspondiente se representa en SEQ ID NO:1 pudo dar lugar a este efecto, resulta evidente que las secuencias reguladoras del gen *R1* necesarias y suficientes para mediar la expresión del polipéptido *R1* tras la infección por el patógeno están contenidas en la secuencia de ADN aislada. Resulta inmediatamente evidente para el experto en la técnica que tales secuencias reguladoras tienen importantes aplicaciones en sí mismas, por ejemplo para la expresión de secuencias de ADN heterólogas específicamente tras la infección por el patógeno, por ejemplo, para la inducción de una respuesta hipersensible a un patógeno dado.

25 Por consiguiente, la presente descripción también se refiere a una secuencia reguladora de un promotor que regula de manera natural la expresión de una molécula de ácido nucleico de la invención descrita anteriormente o de una molécula de ácido nucleico homóloga a una molécula de ácido nucleico de la invención, pudiendo dicha secuencia reguladora conferir o modular la expresión de una secuencia de ADN heteróloga tras la infección por el patógeno.

30 En contexto con la presente invención, la expresión "secuencia reguladora" se refiere a secuencias que influyen sobre la especificidad y/o el nivel de expresión, por ejemplo en el sentido de que confieren especificidad de celular y/o tisular. Tales regiones pueden ubicarse en sentido 5' del sitio de inicio de la transcripción, pero también pueden ubicarse en sentido 3' del mismo, por ejemplo, en secuencias líder transcritas pero no traducidas, o en intrones.

35 El término "promotor", dentro del significado de la presente invención se refiere a una secuencia de nucleótidos necesaria para el inicio de la transcripción, es decir unión a ARN polimerasa e inicio satisfactorio de la formación de transcripto de procesado, y también puede incluir, por ejemplo, la caja TATA.

40 La expresión "molécula de ácido nucleico homóloga a una molécula de ácido nucleico de la invención", tal como se usa en el presente documento incluye regiones promotoras y secuencias reguladoras de otros genes de *R1*, tales como genes de otra especie, por ejemplo, tomate que son homólogos a genes de *R1* de la patata y que presentan sustancialmente el mismo patrón de expresión. Tales promotores se caracterizan por su capacidad para conferir de manera preferiblemente exclusiva la expresión de una secuencia de ADN heteróloga en una planta tras la infección por el patógeno.

45 La expresión "que puede conferir o modular la expresión de una secuencia de ADN heteróloga tras la infección por el patógeno" tal como se usa en el presente documento significa que dicho promotor puede controlar la expresión de una secuencia de ADN heteróloga en plantas en sitios de infección, análoga o estrechamente relacionada con la expresión controlada de genes relacionados con el patógeno que participan en la resistencia natural en la mayoría de las interacciones huésped/patógeno incompatibles, tales como la muerte celular por hipersensibilidad en sitios de infección de una parte de una planta. Por tanto, la secuencia reguladora de la invención se caracteriza por su capacidad de mediar la activación transcripcional localizada selectivamente en respuesta a un ataque por patógeno o en respuesta a estímulos que imitan un ataque por patógeno tales como elicidores preparados, por ejemplo, a partir patógenos tales como hongos o bacterias o derivados de los mismos. La activación transcripcional por la secuencia reguladora de la invención también puede producirse en células que rodean al sitio de infección real debido a interacciones célula-célula. La secuencia reguladora de la invención y promotores químéricos que comprenden tales secuencias pueden no ser inducibles, o serlo sólo en un pequeño grado, con otros estímulos tales como estrés abiótico. Preferiblemente, la inducción a partir del promotor químérico tras un ataque por patógeno o tratamiento con elicitor es al menos aproximadamente 10 veces superior, preferiblemente 20 veces superior y particularmente 30 veces superior a su activación, si la hay, mediante estrés abiótico.

50 Sin embargo, la especificidad de expresión conferida por las secuencias reguladoras de la invención puede no limitarse a la expresión génica local debida a patógenos, por ejemplo, puede combinarse con secuencias reguladoras

adicionales que proporcionan expresión génica específica de tejidos. El patrón de expresión particular también puede depender del sistema de planta/vector empleado. Sin embargo, la expresión de secuencias de ADN heterólogas impulsada por las secuencias reguladoras de la invención se produce predominantemente tras la infección por patógeno o el tratamiento con un elicitor correspondiente a menos que el experto en la técnica tome y diseñe determinados elementos de la invención para controlar la expresión de una secuencia de ADN heteróloga en determinados tipos de células.

Por tanto, según la presente invención, pueden usarse secuencias reguladoras de otras especies que son funcionalmente homólogas a las secuencias reguladoras del promotor de las moléculas de ácido nucleico específicas de *R1* definidas anteriormente, o promotores de genes que presentan un patrón de expresión idéntico o similar. El patrón de expresión particular también puede depender del sistema de planta/vector empleado. Sin embargo, la expresión de secuencias de ADN heterólogas activada por las secuencias reguladoras de la invención se produce predominantemente en cualquier célula infectada por un patógeno particular a menos que el experto en la técnica tome y diseñe determinados elementos de las secuencias reguladoras de la invención para controlar la expresión de una secuencia de ADN heteróloga en un tejido particular o controlarla de otro modo. Según la presente invención, pueden aislarse secuencias reguladoras novedosas de genes *R1*, y se han mostrado a modo de ejemplo para la secuencia reguladora del gen *R1* de la patata. Por ejemplo, puede digerirse ADN genómico con enzimas de restricción apropiadas, desnaturalizarse y permitirse que se hibride con un cebador inverso derivado de la secuencia de ADNc de la invención. Tras la extensión con cebador, puede ligarse un adaptador de extremos romos y puede realizarse una PCR usando un cebador inverso anidado derivado del ADNc mencionado, y un cebador directo derivado de la secuencia adaptadora. En otra estrategia para la clonación de las secuencias reguladoras de la invención puede construirse un mapa físico de las secuencias genómicas en sentido 5' de la región codificante por medio de análisis de tipo Southern genómico. Con esta información, puede digerirse ADN genómico con enzimas de restricción seleccionadas, pueden purificarse en gel fragmentos genómicos que contienen una parte de las secuencias en sentido 5' y la secuencia codificante y auto-ligarse en un gran volumen para favorecer la formación de moléculas circulares, que posteriormente pueden amplificarse mediante PCR con cebadores directos e inversos, derivados de la secuencia codificante del gen. Dentro de la secuencia genómica clonada, el sitio de inicio de la transcripción puede determinarse mediante procedimientos convencionales bien conocidos por todos los expertos en la técnica, tales como 5'-RACE, extensión con cebadores o mapeo S1. Para definir elementos reguladores en *cis* en sentido 5' del sitio de inicio de la transcripción (es decir dentro de la supuesta región promotora), se fusiona la región respectiva con genes de marcador tales como genes que codifican para GUS o GFP, y se generan derivados con delección en 5' de estos constructos. Se transforman en material vegetal adecuado, y se determina la expresión del gen de marcador dependiente de la secuencia en sentido 5' restante (supuesto promotor). Un experto en la técnica conoce bien estas técnicas.

En una realización la secuencia reguladora de la invención comprende una secuencia de ADN seleccionada del grupo que consiste en

35 (a) secuencias de ADN que comprenden la secuencia de nucleótidos tal como se representa en SEQ ID NO. 1 desde los nucleótidos 1 hasta 2222 o de 1 a 2164.

Las secuencias reguladoras homólogas difieren en una o más posiciones de la secuencia reguladora de (a) o (b) pero todavía tienen la misma especificidad, concretamente comprenden motivos de secuencia iguales o similares, preferiblemente de 6 a 10 nucleótidos de longitud, responsables del patrón de expresión descrito anteriormente. Preferiblemente, tales secuencias reguladoras se hibridan con una de las secuencias reguladoras mencionadas anteriormente, lo más preferiblemente en condiciones rigurosas. Se dan a conocer adicionalmente secuencias reguladoras que comparten al menos el 85%, más preferiblemente el 90-95% y lo más preferiblemente el 96-99% de identidad de secuencia con una de las secuencias reguladoras mencionadas anteriormente y tienen la misma o sustancialmente la misma especificidad. Tales secuencias reguladoras también comprenden las que están alteradas, por ejemplo mediante una o más delecciones, inserciones, sustituciones, adiciones y/o recombinaciones de nucleótidos y/o cualquier otra modificación conocida en la técnica o bien sola o bien en combinación en comparación con la secuencia de nucleótidos descrita anteriormente. Los expertos en la técnica conocen bien métodos para introducir tales modificaciones en la secuencia de nucleótidos de las secuencias reguladoras de la invención. También es inmediatamente evidente para el experto en la técnica que puede añadirse cualquier elemento regulador adicional a las secuencias reguladoras de la invención. Por ejemplo, pueden emplearse secuencias y/o potenciadores de la transcripción que permiten la expresión inducida de las secuencias reguladoras de la invención. Un sistema inducible adecuado es por ejemplo la expresión génica regulada por tetraciclina tal como se describe, por ejemplo, por Gatz, citado anteriormente.

55 Existe la posibilidad de modificar las secuencias reguladoras descritas anteriormente o motivos de secuencias de las mismas mediante, por ejemplo, sustituciones de nucleótidos que no afectan a la estructura global o el motivo de unión de la secuencia reguladora de modo que sigue pudiendo conferir expresión génica tras la infección con patógenos. La secuencia reguladora de la invención puede derivarse de los genes *R1* de patata (véanse los ejemplos) aunque otras plantas pueden ser fuentes adecuadas para tales secuencias reguladoras también. Además, las secuencias de nucleótidos de la invención pueden compararse ya que pueden usarse programas informáticos apropiados conocidos en la técnica tales como BLAST, que significa Herramienta de Búsqueda de Alineación Local Básica "Basic Local Alignment Search Tool" (Altschul, 1997; Altschul, J. Mol. Evol. 36 (1993), 290-390; Altschul, J. Mol. Biol. 215 (1990); 403-410) para buscar alineaciones de secuencias locales. BLAST produce alineaciones de secuencias

de nucleótidos para determinar la similitud de secuencia. Debido a la naturaleza local de las alineaciones, BLAST es especialmente útil en la determinación de coincidencias exactas o en la identificación de homólogos. Con tales medios, es posible identificar secuencias de nucleótidos conservadas que pueden desempeñar un papel en la expresión específica del patógeno.

5 Habitualmente, dicha secuencia reguladora es parte de una molécula de ADN recombinante. En una realización preferida de la presente invención, la secuencia reguladora en la molécula de ADN recombinante está operativamente unida a una secuencia de ADN heteróloga. El término heterólogo con respecto a la secuencia de ADN que está operativamente unida a la secuencia reguladora de la invención significa que dicha secuencia de ADN no está unida de manera natural a la secuencia reguladora de la invención. La expresión de dicha secuencia de ADN heteróloga comprende la transcripción de la secuencia de ADN, preferiblemente dando un ARNm traducible. Los expertos en la técnica conocen bien elementos reguladores que garantizan la expresión en células eucariotas, preferiblemente células vegetales. Comprenden habitualmente señales de poli-A que garantizan la terminación de la transcripción y la estabilización del transcripto, véase también anteriormente. Los elementos reguladores adicionales pueden incluir potenciadores de la transcripción así como de la traducción; véase anteriormente.

10 15 En una realización preferida, la secuencia de ADN heteróloga de las moléculas de ADN recombinante descritas anteriormente codifica para un péptido, proteína, ARN antisentido, ARN homosentido y/o ribozima. La molécula de ADN recombinante de la invención puede usarse sola o como parte de un vector para expresar secuencias de ADN heterólogas, que, por ejemplo, codifican para proteínas para, por ejemplo, proteínas de almacenamiento de semillas, 20 toxinas, anticuerpos ("planticuerpos") o diagnósticos de la expresión de genes relacionados con R1. La molécula de ADN recombinante o el vector que contiene la secuencia de ADN que codifica para una proteína de interés se introduce en las células que a su vez producen la proteína de interés. Por ejemplo, las secuencias reguladoras de la invención pueden estar operativamente unidas a secuencias que codifican para barstar y barnasa, respectivamente, para su uso en la producción de una respuesta HR en plantas. Aplicaciones de las secuencias reguladoras de la invención son evidentes para el experto en la técnica y pueden deducirse de la bibliografía, por ejemplo, Strittmatter y Wegener, 25 Zeitschrift fr Naturforschung 48c (1993), 673-688; Kahl, J. Microbiol. Biotechnol. 11 (1995), 449-460 y referencias citadas en el mismo.

30 35 Por otro lado, dicha proteína puede ser un marcador puntuable, por ejemplo, luciferasa, proteína fluorescente verde o β -galactosidasa. Esta realización es particularmente útil para métodos de selección rápidos y sencillos para compuestos y sustancias descritos en el presente documento a continuación que pueden modular la expresión de genes R1. Por ejemplo, puede cultivarse una planta transgénica en presencia y ausencia de un compuesto candidato con el fin de determinar si el compuesto afecta a la expresión de genes que están bajo el control de secuencias reguladoras de la invención, que puede medirse, por ejemplo, monitorizando la expresión del marcador mencionado anteriormente. También es inmediatamente evidente para los expertos en la técnica que pueden emplearse otros genes marcadores también, que codifican para, por ejemplo, un marcador seleccionable que proporciona la selección directa de compuestos que inducen o inhiben la expresión de dicho marcador.

40 45 Las secuencias reguladoras de la invención pueden usarse también en métodos de enfoques antisentido. El ARN antisentido puede ser una secuencia de nucleótidos corta (generalmente de al menos 10, preferiblemente de al menos 14 nucleótidos, y opcionalmente de hasta 100 o más nucleótidos) formulada para que sea complementaria a una parte de una secuencia de ADN y/o una secuencia de ARNm específica del gen de interés. Se han descrito métodos convencionales referentes a la tecnología antisentido; véase, por ejemplo, Klann, Plant Physiol. 112 (1996), 1321-1330 y citado anteriormente. Tras la transcripción de la secuencia de ADN dando ARN antisentido, el ARN antisentido se une a su secuencia diana dentro de una célula, inhibiendo de ese modo la traducción del ARNm y regulando por disminución la expresión de la proteína codificada por el ARNm.

50 55 60 En un aspecto adicional, la descripción se refiere a moléculas de ácido nucleico de al menos 15 nucleótidos de longitud que se hibridan específicamente con una secuencia reguladora descrita anteriormente o con una cadena complementaria de la misma. La hibridación específica se produce preferiblemente en condiciones rigurosas e implica ninguna o muy poca hibridación cruzada con secuencias de nucleótidos que no tienen propiedades reguladoras o tienen propiedades reguladoras sustancialmente diferentes. Tales moléculas de ácido nucleico pueden usarse como sondas y/o para el control de la expresión génica. Los expertos en la técnica conocen bien la tecnología de sondas de ácido nucleico, los cuales apreciarán fácilmente que tales sondas pueden variar en longitud. Se prefieren sondas de ácido nucleico de 17 a 35 nucleótidos de longitud. Por supuesto, también puede ser apropiado usar ácidos nucleicos de hasta 100 y más nucleótidos de longitud. Las sondas de ácido nucleico son útiles para diversas aplicaciones. Por un lado, pueden usarse como cebadores de PCR para la amplificación de secuencias reguladoras según la invención. Otra aplicación es el uso como sonda de hibridación para identificar secuencias reguladoras que se hibridan con las secuencias reguladoras de la invención mediante examen de homología de bibliotecas de ADN genómico. Las moléculas de ácido nucleico que son complementarias a una secuencia reguladora descrita anteriormente pueden usarse también para la represión de la expresión de un gen que comprende tales secuencias reguladoras, por ejemplo debido a un efecto antisentido, de cosupresión o de triple hélice o para la construcción de ribozimas apropiadas (véanse, por ejemplo, los documento EP-B1 0 291 533, EP-A1 0 321 201, EP-A2 0 360 257) que escinden específicamente el (pre)-ARNm de un gen que comprende una secuencia reguladora de la invención. La selección de sitios diana apropiados y ribozimas correspondientes puede realizarse tal como se describe por ejemplo en Steinecke, Ribozymes, Methods in Cell Biology 50, Galbraith *et al.* eds. Academic Press, Inc. (1995), 449-460. Además, el experto

en la técnica es muy consciente de que también es posible marcar una sonda de ácido nucleico de este tipo con un marcador apropiado para aplicaciones específicas, tales como para la detección de la presencia de una molécula de ácido nucleico de la invención en una muestra derivada de un organismo.

5 Las moléculas de ácido nucleico descritas anteriormente pueden ser o bien ADN o bien ARN o bien un híbrido de los mismos. Además, dicha molécula de ácido nucleico puede contener, por ejemplo, enlaces tioéster y/o análogos de nucleótidos, comúnmente usados en enfoques antisentido de oligonucleótidos; véase anteriormente.

10 La presente invención también se refiere a vectores, particularmente plásmidos, cósmidos, virus y bacteriófagos usados convencionalmente en ingeniería genética que comprenden una secuencia reguladora o molécula de ARN recombinante correspondiente de la invención. Preferiblemente, dicho vector es un vector de expresión y/o un vector que comprende adicionalmente un marcador de selección para plantas. Para un ejemplo de marcadores selectores adecuados, véase anteriormente. Pueden usarse métodos que conocen bien los expertos en la técnica para construir vectores recombinantes; véanse, por ejemplo, las técnicas descritas en Sambrook, Molecular Cloning A Laboratory Manual, Cold Spring Harbor Laboratory (1989) N.Y. y Ausubel, Current Protocols in Molecular Biology, Green Publishing Associates y Wiley Interscience, N.Y. (1989). Alternativamente, las moléculas de ADN recombinantes y los vectores de la invención pueden reconstituirse en liposomas para su administración a células diana.

15 La presente invención se refiere además a células huésped transformadas con una secuencia reguladora, una molécula de ADN o vector de la invención. Dicha célula huésped puede ser una célula procariota o eucariota. La secuencia reguladora, vector o molécula de ADN recombinante de la invención que está presente en la célula huésped puede estar o bien integrada en el genoma de la célula huésped o bien puede mantenerse de manera extracromosómica. La célula huésped puede ser cualquier célula procariota o eucariota, tal como una célula bacteriana, de insecto, fúngica, vegetal, animal o humana. Células preferidas son células vegetales.

20 En una realización preferida adicional, la presente invención proporciona un método para la producción de plantas transgénicas, células vegetales o tejido vegetal que comprende la introducción de una molécula de ácido nucleico, molécula de ADN recombinante o vector de la invención en el genoma de dicha planta, célula vegetal o tejido vegetal. Para la expresión de la secuencia de ADN heteróloga bajo el control de la secuencia reguladora según la invención en células vegetales, pueden fusionarse secuencias reguladoras adicionales tales como cola de poli A, preferiblemente en 3' con respecto a la secuencia de ADN heteróloga, véase también anteriormente. Posibilidades adicionales podrían ser añadir potenciadores de la transcripción o la traducción que aumentan la expresión génica, o secuencias que aumentan la estabilidad del ARNm. Se describen anteriormente métodos para la introducción de ADN foráneo en plantas, células vegetales y tejido vegetal.

25 Por tanto, la presente invención se refiere también a células vegetales transgénicas que contienen, preferiblemente integrada de manera estable en el genoma, una secuencia reguladora, una molécula de ADN recombinante o un vector según la invención. Además, la presente invención también se refiere a plantas transgénicas y tejido vegetal que comprende las células vegetales transgénicas descritas anteriormente.

30 35 Además, la presente invención se refiere a un método para la identificación de un agente protector de plantas que comprende las etapas de:

- 40 (a) cultivar un tejido o célula vegetal o mantener una planta que comprende una molécula de ADN recombinante que comprende un sistema de lectura operativamente unida a una secuencia reguladora de la presente invención en presencia de un compuesto o una muestra que comprende una pluralidad de compuestos en condiciones que permiten la expresión de dicho sistema de lectura;
- (b) identificar o verificar una muestra y compuesto, respectivamente, que conduce a la supresión o activación y/o potenciación de la expresión de dicho sistema de lectura en dicha planta, célula vegetal o tejido vegetal.

45 La expresión "sistema de lectura" en el contexto de la presente invención significa una secuencia de ADN que tras su transcripción y/o expresión en una célula, tejido u organismo proporciona un fenotipo puntuable y/o seleccionable. Tales sistemas de lectura los conocen bien los expertos en la técnica y comprenden, por ejemplo, moléculas de ADN recombinante y genes marcadores tal como se describió anteriormente.

50 La expresión "pluralidad de compuestos" en un método de la invención debe entenderse como una pluralidad de sustancias que pueden ser o no idénticas.

55 Dicho compuesto o pluralidad de compuestos pueden ser inorgánicos u orgánicos, compuestos que se producen de manera natural o artificiales y pueden estar comprendido en, por ejemplo, muestras, por ejemplo, extractos celulares de, por ejemplo, plantas, animales o microorganismos. Además, dicho(s) compuesto(s) pueden conocerse en la técnica pero hasta la fecha no se sabía que pudieran suprimir o activar y/o potenciar la transcripción de un gen R1. La pluralidad de compuestos puede añadirse, por ejemplo, al medio de cultivo o inyectarse en la planta, células vegetales o tejido o pulverizarse sobre la planta o suministrarse en el suelo.

55 Si se identifica una muestra que contiene un compuesto o una pluralidad de compuestos en el método de la invención, entonces es posible o bien aislar el compuesto de la muestra original que se identificó que contenía el

- 5 compuesto que podía suprimir o activar y/o potenciar la transcripción de un gen R1, o bien puede subdividirse la muestra original, por ejemplo, si consiste en una pluralidad de diferentes compuestos, de modo que se reduce el número de diferentes sustancias por muestra y se repite el método con las subdivisiones de la muestra original. Dependiendo de la complejidad de las muestras, las etapas descritas anteriormente pueden realizarse varias veces, preferiblemente hasta que la muestra identificada según el método de la invención comprende sólo un número limitado de sustancias o sólo una sustancia. Preferiblemente, dicha muestra comprende sustancias de propiedades químicas y/o físicas similares, y lo más preferiblemente dichas sustancias son idénticas. El compuesto identificado según el método descrito anteriormente puede formularse además en una forma adecuada para la aplicación en la reproducción de plantas o el cultivo de tejidos y células vegetales.
- 10 Los compuestos que pueden someterse a prueba e identificarse según un método de la invención pueden ser bibliotecas de expresión, por ejemplo, bibliotecas de expresión de ADNc, péptidos, proteínas, ácidos nucleicos, anticuerpos, compuestos orgánicos pequeños, hormonas, peptidomiméticos, PNA o similares (Milner, *Nature Medicine* 1 (1995), 879-880; Hupp, *Cell* 83 (1995), 237-245; Gibbs, *Cell* 79 (1994), 193-198 y referencias citadas anteriormente). Además, pueden identificarse genes que codifican para un supuesto regulador de un gen R1 usando, por ejemplo, mutagénesis por inserción usando, por ejemplo, vectores de direccionamiento génico conocidos en la técnica (véase, por ejemplo, Hayashi, *Science* 258 (1992), 1350-1353; Fritze y Walden, *Gene activación by T-DNA tagging*. En *Methods in Molecular biology* 44 (Gartland, K.M.A. y Davey, M.R., eds). Totowa: Human Press (1995), 281-294) o etiquetado de transposones (Chandlee, *Physiología Plantarum* 78 (1990), 105-115). Dichos compuestos pueden ser también análogos o derivados funcionales de activadores o inhibidores conocidos. Los expertos en la técnica conocen bien métodos para la preparación de análogos y derivados químicos y se describen en, por ejemplo, Beilstein, *Handbook of Organic Chemistry*, Springer edition New York Inc., 175 Fifth Avenue, Nueva York, N.Y. 10010 EE.UU. y *Organic Synthesis*, Wiley, Nueva York, EE.UU. Además, dichos análogos y derivados pueden someterse a prueba para determinar sus efectos según métodos conocidos en la técnica. Además, pueden usarse peptidomiméticos y/o diseño asistido por ordenador de derivados apropiados, por ejemplo, según los métodos descritos anteriormente.
- 15 25 La determinación de si un compuesto puede suprimir o activar y/o potenciar la transcripción de un gen R1 puede hacerse, por ejemplo, en plantas monitorizando el gen indicador. Puede hacerse adicionalmente monitorizando las características fenotípicas de la planta transgénica de la invención que se pone en contacto con los compuestos y comparándolas con las de plantas de tipo natural. En una realización adicional, dichas características pueden compararse con las de una planta transgénica que se pone en contacto con un compuesto que se sabe que puede o no puede suprimir o activar y/o potenciar la expresión del gen R1 o la actividad de la proteína. Se espera que los compuestos identificados según el método de la invención sean muy beneficiosos puesto que los promotores que se conocen hasta la fecha son sólo de uso limitado debido a la especificidad de patógeno no regulada o no estrechamente regulada de sus secuencias reguladoras.
- 30 35 40 45 50 55 60 El inhibidor o activador identificado mediante el método descrito anteriormente puede ser útil como herbicida, pesticida y/o como regulador del crecimiento de plantas. Tales compuestos útiles pueden ser por ejemplo factores de actuación en *trans* que se unen a la secuencia reguladora de la invención. La identificación de factores que actúan en *trans* puede llevarse a cabo usando métodos convencionales en la técnica (véase, por ejemplo, Sambrook, citado anteriormente, y Ausubel, citado anteriormente). Para determinar si una proteína se une a las secuencias reguladoras de la invención, pueden llevarse a cabo análisis de desplazamiento en gel nativo y/o de obtención de huellas de ADN convencionales. Con el fin de identificar un factor de actuación en *trans* que se une a las secuencias reguladoras de la invención, las secuencias reguladoras pueden usarse como un reactivo de afinidad en métodos de purificación de proteínas convencionales, o como una sonda para el examen de una biblioteca de expresión. Una vez identificado el factor de actuación en *trans*, puede proseguirse con la modulación de esta unión a las secuencias reguladoras de la invención, comenzando con, por ejemplo, la selección de inhibidores frente a la unión del factor de actuación en *trans* a las secuencias reguladoras de la presente invención. La activación o represión de genes R1 podría lograrse entonces en plantas mediante la aplicación del factor de actuación en *trans* (o su inhibidor) o los genes que lo codifican, por ejemplo en un vector para plantas transgénicas. Además, si la forma activa del factor de actuación en *trans* es un dímero, podrían prepararse mutantes dominantes negativos del factor de actuación en *trans* con el fin de inhibir su actividad. Además, tras la identificación del factor de actuación en *trans*, pueden identificarse entonces componentes adicionales en la ruta que conduce a la activación (por ejemplo transducción de señales) o represión de un gen bajo el control de las secuencias reguladoras de la presente invención. Puede proseguirse con la modulación de las actividades de estos compuestos, con el fin de desarrollar fármacos y métodos adicionales para modular la expresión de un gen bajo el control de las secuencias reguladoras de la presente invención.
- Preferiblemente, el compuesto identificado según el método descrito anteriormente o su análogo o derivado se formula adicionalmente en una forma adecuada para la aplicación en la reproducción de plantas o el cultivo de tejidos y células vegetales. Por ejemplo, puede combinarse con un portador aceptable en agricultura conocido en la técnica. La composición para la protección de plantas puede prepararse empleando el método de la invención descrito anteriormente y sintetizando el compuesto identificado como inhibidor o activador en una cantidad suficiente para su uso en agricultura. Por tanto, la presente descripción también se refiere a un método para la preparación de una composición para la protección de plantas agrícolas que comprende las etapas descritas anteriormente del método de la invención y sintetizar el compuesto así identificado o un análogo o derivado del mismo.

5

En la composición para la protección de plantas, el compuesto identificado mediante el método descrito anteriormente puede formularse preferentemente por medios convencionales comúnmente usados para la aplicación de, por ejemplo, herbicidas y pesticidas o agentes que pueden inducir resistencia adquirida sistémica (SAR). Por ejemplo, pueden usarse determinados aditivos conocidos por los expertos en la técnica que comprenden estabilizantes o sustancias que facilitan la captación por la célula vegetal, tejido vegetal o planta, por ejemplo, carborundo, o solución de SDS al 0,01% (dodecilsulfato de sodio).

En una realización todavía adicional, la presente invención se refiere a un método para identificar y obtener un factor de avirulencia o de virulencia de un patógeno que comprende las etapas de:

- 10 (a) examinar la proteína R1 de la presente invención o un fragmento de la misma frente a una biblioteca de expresión de proteínas o péptidos derivada de un patógeno en un sistema de lectura en condiciones adecuadas que permiten la interacción de la proteína y el péptido en dicho sistema de lectura;
- (b) identificar o verificar un ADNc que conduce a la supresión o activación del sistema de lectura.

15 Además de las posibilidades descritas anteriormente de usar las moléculas de ácido nucleico según la invención para la ingeniería genética de plantas con características modificadas y su uso para identificar moléculas homólogas, las moléculas de ácido nucleico descritas pueden usarse también para otras varias aplicaciones, por ejemplo, para la identificación de moléculas de ácido nucleico que codifican para proteínas que interaccionan con las proteínas R1 descritas anteriormente. Esto puede lograrse mediante ensayos bien conocidos en la técnica, por ejemplo, tal como se describe en Scofield (Science 274 (1996), 2063-2065) mediante el uso del denominado "sistema de dos híbridos" de levadura. En este sistema, la proteína codificada por las moléculas de ácido nucleico según la invención o una parte más pequeña de las mismas se une al dominio de unión a ADN del factor de transcripción GAL4. Una cepa de levadura que expresa esta proteína de fusión y que comprende un gen indicador lacZ dirigido por un promotor apropiado, que reconoce el factor de transcripción GAL4, se transforma con una biblioteca de ADNc que expresará péptidos o proteínas vegetales de los mismos fusionados a un dominio de activación. Por tanto, si un péptido codificado por uno de los ADNc puede interaccionar con el péptido de fusión que comprende un péptido de una proteína de la invención, el complejo puede dirigir la expresión del gen indicador. De este modo, las moléculas de ácido nucleico según la invención y el péptido codificado pueden usarse para identificar péptidos y proteínas que interaccionan con proteína R1. Este método también puede emplearse para identificar inhibidores y activadores tal como se describió anteriormente.

20 30 Otros métodos para identificar compuestos que interaccionan con las proteínas según la invención o moléculas de ácido nucleico que codifican para tales moléculas son, por ejemplo, la selección *in vitro* con el sistema de presentación en fago así como ensayos de unión en filtros o medición "en tiempo real" de la interacción usando, por ejemplo, el aparato BIACore (Pharmacia); véanse las referencias citadas anteriormente.

25 Puede seguirse una estrategia similar con el denominado sistema de tres híbridos.

30 35 El sistema de dos híbridos de levadura se ha descrito originalmente por Fields y Song (Nature 340 (1989), 245-246; véase también para revisión Vidal, M, en Bartel, P.L. y Fields, S. (eds.), The yeast two-hybrid system. Oxford University Press, Nueva York, NY, (1997), 109-147). Una versión modificada del sistema de dos híbridos de levadura se ha descrito por (Gyuris, Cell 75 (1993), 223-232; Zervos, Cell 72 (1993), 223-232). En resumen, se usa un dominio del polipéptido como cebo para la unión de compuestos. Se seleccionan entonces los positivos por su capacidad para crecer sobre placas que carecen de leucina, y entonces se someten a prueba adicionalmente para determinar su capacidad para volverse azules sobre placas con X-gal, tal como se describió anteriormente en gran detalle; véase también el documento WO 95/31544. Una versión modificada es el "sistema de dos híbridos de levadura inverso" que permite seleccionar para la interacción de alelos defectuosos usando una estrategia de selección negativa tal como se describe, por ejemplo, en (Vidal, Proc. Natl Acad Sci. USA 93 (1996), 10321-10326; Vidal, Proc. Natl Acad Sci. USA 93 (1996), 10315-10320). Este sistema usa el gen indicador seleccionable contrario URA3. Las células de levadura que expresan Ura3p convierten el compuesto ácido 5-fluoroorótico (FOA) en el derivado tóxico 5-flourouracilo. Por tanto, una interacción de dos híbridos que conduce a la activación del gen indicador URA3 puede seleccionarse de manera contraria en presencia de FOA y pueden seleccionarse específicamente mutantes de pérdida de función seleccionados fuera de un gran conjunto de alelos de tipo natural.

40 45 50 55 Otro método conveniente, por ejemplo, podría ser el sistema de tres híbridos de levadura tal como se describe (SenGupta, Proc. Natl. Acad. Sci. USA 93 (1996), 8496-8501). El sistema de selección de tres híbridos de levadura se desarrolló para aislar los genes de las proteínas que interaccionan con ARN, y para estudiar las interacciones ARN-proteína. Este sistema, basado en el sistema de dos híbridos de levadura, consiste en un dominio de unión a ADN fusionado a una proteína de unión a ARN conocida, un dominio de activación fusionado a una proteína de unión a ARN prospectiva y un ARN híbrido. La transcripción de genes indicadores sólo se produce cuando ambas proteínas híbridas interaccionan con el ARN híbrido. En el sistema de tres híbridos inverso, la interacción de las proteínas con el ARN híbrido da como resultado la expresión de un gen indicador cuyo producto es tóxico para células de levadura. Todos estos métodos pueden emplearse según el método de la invención descrito anteriormente con la proteína R1 o fragmentos peptídicos de la misma como cebo para identificar y obtener un factor de avirulencia o de virulencia y sus ADNc codificantes o partes de los mismos. El experto en la técnica conoce métodos para obtener la secuencia de ADN

de esos clones sometidos a prueba positivos en el ensayo de selección y se describen en las publicaciones a las que se hizo referencia anteriormente.

La presente descripción también se refiere al ADNc y su producto codificado obtenido o identificado mediante el método descrito anteriormente.

5 La descripción también se refiere a composiciones que comprenden al menos una de las moléculas de ácido nucleico mencionadas anteriormente y/o que comprenden una molécula de ácido nucleico que es complementaria para una molécula de ácido nucleico de este tipo, un vector, una proteína R1 o un fragmento inmunológica o biológicamente activo de la misma o un anticuerpo o aptámero que reconoce específicamente un fragmento o proteína de este tipo; una secuencia regulatoria o ADN recombinante, o un vector correspondiente, un compuesto diseñado orientado según la proteína de la descripción y/o identificado según el método descrito anteriormente y/o un anticuerpo que reconoce específicamente un compuesto de este tipo o una secuencia reguladora, y opcionalmente medios adecuados para la detección o medios imitables para el cultivo de tejidos y células vegetales.

10 15 Pueden usarse composiciones de diagnóstico para métodos para detectar la expresión del gen R1 detectando la presencia de ARNm correspondiente que comprende el aislamiento de ARNm de una célula y poner en contacto el ARNm así obtenido con una sonda que comprende una sonda de ácido nucleico tal como se describió anteriormente en condiciones de hibridación, y detectar la presencia de ARNm hibridado y de ese modo la expresión del gen por la célula. Métodos adicionales de detección de la presencia de una proteína según la presente invención comprenden inmunotécnicas bien conocidas en la técnica, por ejemplo ensayo inmunoabsorbente ligado a enzimas.

20 25 30 Además, la presente descripción se refiere a un kit que comprende al menos uno de las moléculas de ácido nucleico, vectores, proteína, compuestos, anticuerpos o aptámeros mencionados anteriormente de la invención. El kit puede contener componentes adicionales tales como marcadores de selección y componentes para medios selectivos adecuados para la generación de plantas, tejido vegetal o células vegetales transgénicas. Además, el kit puede incluir tampones y sustratos para genes indicadores que pueden estar presentes en el gen recombinante o vector de la invención. El kit puede usarse ventajosamente para llevar a cabo el método de la invención y podría emplearse, entre otras, en una variedad de aplicaciones a las que se hace referencia en el presente documento, por ejemplo, en el campo del diagnóstico o como herramienta de investigación. Las partes del kit pueden envasarse individualmente en viales o en combinación en recipientes o unidades de múltiples recipientes. La fabricación del kit sigue preferiblemente procedimientos convencionales que conoce el experto en la técnica. El kit o sus componentes según la invención pueden usarse en cultivos de tejidos vegetales y células vegetales, por ejemplo, para cualquiera de los métodos descritos anteriormente para detectar inhibidores y activadores de genes R1. Se espera que el kit y sus componentes sean muy útiles en la reproducción de nuevas variedades de, por ejemplo, plantas que presentan propiedades mejoradas tales como valor nutritivo o resistencia a enfermedades.

35 También es inmediatamente evidente para el experto en la técnica que las secuencias reguladoras, las moléculas de ADN recombinante, los vectores y los compuestos de la presente invención pueden emplearse para producir plantas transgénicas con un rasgo deseado; véase por revisión TIPTEC Plant Product & Crop Biotechnology 13 (1995), 312-397.

40 Además, es posible usar las moléculas de ácido nucleico según la invención como marcadores moleculares en la reproducción de plantas. Además, la sobreexpresión de moléculas de ácido nucleico según la invención puede ser útil para la alteración o modificación de la interacción planta/patógeno. El término "patógeno" incluye, por ejemplo, bacterias, virus y hongos así como protozoos. Preferiblemente, dicho patógeno es *P. infestans*.

45 Además, la presente invención se refiere al uso de una molécula de ácido nucleico, vector, célula huésped, proteína, una secuencia reguladora, un aptámero, molécula de ADN recombinante, un vector, un compuesto, un aptámero y/o el anticuerpo de la invención para su uso en un método de selección para la identificación de genes de virulencia y avirulencia de patógenos, para la selección de compuestos protectores para plantas, para inducir resistencia a patógenos en plantas, como marcador en la reproducción de plantas asistida por marcador. La secuencia reguladora o una molécula de ADN recombinante de la presente invención se usa preferiblemente para la expresión de una secuencia de ADN heteróloga.

50 55 Estas y otras realizaciones se dan a conocer y se abarcan mediante la descripción y los ejemplos de la presente invención. Puede obtenerse bibliografía adicional referente a uno cualquiera de los métodos, usos y compuestos que van a emplearse según la presente invención de bibliotecas públicas, usando por ejemplo dispositivos electrónicos. Por ejemplo, puede utilizarse la base de datos pública "Medline" que está disponible en Internet, por ejemplo en <http://www.ncbi.nlm.nih.gov/PubMed/medline.html>. Los expertos en la técnica conocen bases de datos y direcciones adicionales, tales como <http://www.ncbi.nlm.nih.gov/>, <http://www.infobiogen.fr/>, http://www.fmi.ch/biology/research_tools.html, <http://www.tigr.org/>, y pueden obtenerse también usando, por ejemplo, <http://www.lycos.com>. Se facilita una visión general de información de patentes en biotecnología y un estudio de fuentes relevantes de información de patentes útiles para investigación retrospectiva y para reconocimiento actual en Berks, TIBTECH 12 (1994), 352-364.

La presente invención se describe adicionalmente mediante referencia a las siguientes figuras y ejemplos no limitativos.

Las figuras muestran:

Figura 1. Mapa genético y físico de la región *R1*. *GP21* y *GP179* son los marcadores usados para construir el mapa de alta resolución de la región *R1*. *SPUD237* y *AFLP1* son marcadores de AFLP convertidos (Meksem *et al.* 1995, De Jong *et al.* 1997) que flanquean a *R1*. Las distancias genéticas se dan en cM. *CosS* es un clon de cósmido seleccionado con *SPUD237*. Los clones restantes en el mapa físico son BAC con longitudes entre 70 y 90 kb. Barras negras continuas: BAC del cromosoma que porta *R1*. Barras grises: BAC del cromosoma que porta *r1*. Barra blanca: BAC de origen no determinado. Los extremos de BAC mapeados se indican mediante el número de recombinantes que separan el extremo de BAC de *R1*. Los extremos de cósmidos y BAC usados para el paseo cromosómico se indican mediante las flechas verticales. *RGL*: fragmento similar a gen de resistencia.

Figura 2. Amplificación por PCR de un fragmento de 1,4 kb del gen *R1* usando los cebadores específicos de alelo 76-2sf2 y 76-2SR y ADN molde de (A) Desirée; (B) original resistente P41; (C) original susceptible P40; (D) planta Desirée transgénica 10-2_5;

(E) clon de BAC BA87d17 que porta el alelo *R1*; (F), (G), (H), (I), (J) clones de BAC BA122p13, BA12101, BA76011, BA47F2 y BA27c1, respectivamente, (K) control negativo.

Figura 3. Prueba de complementación de *R1*. Se muestran síntomas de enfermedad 9 días tras la inoculación sobre folíolos de (A) Desirée susceptible; (B) línea de Desirée transgénica no 10-2_5 transformada con el clon g10-2 y (C) el original resistente P41 (*R1r1*).

Figura 4. El gen *R1*. (A) Organización estructural. Se muestran los exones como recuadros y los intrones como líneas en ángulo. (B) La secuencia de aminoácidos deducida. El motivo de cremallera de leucina está subrayado dos veces. La región LRR está indicada en cursiva. Los motivos quinasa pronosticados están indicados dentro de la región recuadrada y los sitios de N-glicosilación están indicados en negrita. Los motivos conservados QLPL, CFLY y LHD específicos para proteínas de resistencia de plantas están subrayados. Los códigos de una única letra para los aminoácidos son convencionales.

Figura 5. Representación esquemática de la región del cromosoma 5 alrededor del locus *R1*. Los recuadros con líneas verticales representan regiones homólogas entre los cromosomas que llevan los alelos *R1* y *r1*. El alelo funcional *R1.1* y el locus *R1.2/r1.2* están marcados con recuadros abiertos. La línea en ángulo indica la delección presente en el cromosoma *r1* en comparación con el cromosoma *R1*.

EJEMPLOS

Material y Métodos:

Material vegetal:

Se usó la descendencia F1 de un cruce entre los clones diploides heterocigotos H79.1506/1 (*R1 r1*) y H80.696/4 (*r1 r1*), denominados P41 y P40, respectivamente (Gebhardt *et al.* 1989, Leonards-Schippers *et al.* 1992) para el mapeo genético de alta resolución de *R1*. Se seleccionaron recombinantes en el intervalo de marcador *GP21* - *GP179* que se originan a partir del original P41 (*R1r1*) tal como se describe (Meksem *et al.* 1995). Se usó el clon híbrido P6/210 derivado del cruce P41 x P40 (Leister *et al.* 1997) que porta *R1* en estado heterocigoto para construir un cósmido genómico y bibliotecas de BAC. Se usó el original P41 (*R1r1*) para la construcción de bibliotecas de ADNc.

Prueba para la resistencia a *Phytophthora infestans*:

Se determinó la resistencia a un aislado de *P. infestans* que tenía el factor de avirulencia correspondiente *Avr1* (variedad 4) tal como se describe (Leonards-Schippers *et al.* 1992), excepto porque se usaron folíolos completos en lugar de discos de hojas para la inoculación. Se puntuó la presencia o ausencia de respuesta hipersensible (HR) 8-10 días tras la inoculación.

Bibliotecas genómicas de patata:

Se suministró la biblioteca de BAC por LION Bioscience AG (Heidelberg, Alemania). La biblioteca se ha construido a partir de ADN genómico de alto peso molecular parcialmente digerido con *HindIII* del clon P6/210 en el vector binario pCLD04541 (Jones *et al.* 1992) tal como se describe (Meksem *et al.* 2000). La biblioteca de BAC consiste en 101.376 clones con un tamaño de inserto promedio de 70 kb. Se almacenaron las colonias en 264 384 placas de microtitulación (Genetix, Oxford, RU) en medio 2YT (Sambrook *et al.* 1989) con tampón de congelación (glicina al 5,5% p/v, $(\text{NH}_4)\text{SO}_4$ 7 mM, Na-citrato 1,5 mM, MgSO_4 0,3 mM, KH_2PO_4 13 mM, K_2HPO_4 27 mM).

Se construyó una biblioteca de cósmidos de aproximadamente 150.000 clones usando procedimientos convencionales (Sambrook *et al.* 1989) a partir de ADN genómico parcialmente digerido con *Sau3AI* (fragmentos de 17 - 23 kb) de P6/210 y en el mismo vector (sitio de clonación *BamH1*) que la biblioteca de BAC. Se empaquetaron los

5

cósmidos usando el extracto Gigapack II Gold Packaging (Stratagene, CA, EE.UU.) y se transfecaron en la cepa de *E. coli* SURE TM (Stratagene, CA, EE.UU.). Se extrajo el ADN de plásmido a partir de conjuntos de aproximadamente 1.500 colonias bacterianas (Sambrook *et al.* 1989). Se generaron ciento tres conjuntos de cósmidos y se examinaron mediante PCR usando cebadores específicos de SPUD237 (De Jong *et al.* 1997). Se sembraron en placa los conjuntos positivos y se examinaron mediante hibridación de colonias usando protocolos convencionales (Sambrook *et al.* 1989).

Examen de la biblioteca de BAC y construcción de cíntigos:

Se prepararon filtros de colonias de alta densidad para el examen basado en hibridación de la biblioteca de BAC usando un robot BioGRID (Oxford, RU). Se dispusieron en cuadrículas los clones en puntos dobles usando un alineamiento de 5 x 5 con 6 x 384 alineamientos por membrana de nailon de 22,5 x 22,5 cm (PALL, Biodyne, Portsmouth, RU). Cada alineamiento de 5 x 5 contenía 2 x 12 colonias con la posición de control del alineamiento ocupado por el clon pSW1 (PE Biosystems, Foster City, CA EE.UU.). Este patrón de cuadrícula permitía que 27.648 colonias estuviesen representadas dos veces en cada filtro. Se realizó el examen de la biblioteca usando un conjunto de cuatro filtros que portaban 101.376 clones. Se incubaron los filtros de colonias sobre medio LB durante 15 h a 37°C y se procesaron para la hibridación de colonias usando técnicas convencionales (Sambrook *et al.* 1989). Se realizó la hibridación de filtros tal como se describe (Gebhardt *et al.* 1989), excepto porque se marcaron 300 pg de inserto de control pSW1 y se hibridaron junto con la sonda para facilitar la determinación de las direcciones de clones positivos. Se purificó el ADN de plásmido a partir de clones positivos y se secuenciaron las inserciones a partir de ambos extremos empleando oligonucleótidos de T3 y T7 como cebadores de secuenciación. Se usó la información de secuencia de ADN de los extremos de inserción de BAC para diseñar pares de cebadores de PCR específicos. Se usaron los productos de PCR amplificados con estos cebadores y los BAC respectivos como molde como sondas para una nueva hibridación de filtros para identificar clones de BAC solapantes, para determinar la orientación de los clones de BAC solapantes unos en relación con otros y para mapear en las plantas recombinantes. Se confirmaron los solapamientos secuenciando los productos de PCR. Se verificó la dirección de extensión de cíntigos mediante mapeo genético usando las plantas recombinantes y análisis de marcadores basado en RFLP o PCR. Para determinar el tamaño de las inserciones de BAC, se digirió el ADN de BAC con *NotI* y se separaron los fragmentos mediante electroforesis en gel de campo pulsado sobre un aparato CHEF DRIII (BioRad, Hercules, CA, EE.UU.) durante 12 h a 11°C con un tiempo de pulso inicial de 5 s y un tiempo de pulso final de 10 s, en un ángulo de 120° y 6 V/cm.

Aislamiento de ADN de BAC:

Se extrajo ADN de BAC usando QIAfilter Plasmid Purificación Kit 100 (Qiagen, Hilden, Alemania) según las instrucciones de los fabricantes con modificaciones menores. Se cultivó previamente una única colonia en medio LB líquido durante 8 h a 37°C. Se añadieron 75 µl de cultivo previo a 75 ml de medio LB y se incubó adicionalmente durante 15 h a 37°C. Se introdujo una etapa de centrifugación antes de hacer pasar el sobrenadante a través del QIAfilter para eliminar los desechos celulares bacterianos.

Preparación de sondas a partir de inserciones de BAC:

Se digirieron 1,5 µg de ADN de BAC hasta el final con *HindIII* más *EcoRI* y se separaron del vector sobre agarosa de baja temperatura de fusión al 0,8% (Sea Plaque GTG Agarose, Bioproducts, Rockland, Maine, EE.UU.). Se disolvió el ADN insertado a partir del gel usando el sistema GELase (Epicentre Technologies o Biozym) siguiendo las instrucciones del proveedor. Se precipitó el ADN con etanol, se disolvió en agua y se marcó con ³²P-dCTP mediante marcado por cebado aleatorio (Feinberg y Vogelstein 1984).

Subclonación de BAC BA87d17:

Se digirieron parcialmente 10 µg de ADN de BAC con 1U de *Tsp509I* durante 15 min. a 65°C y se separó por tamaño en un gel de agarosa de baja temperatura de fusión al 0,8% (Sea Plaque GTG Agarose, Bioproducts, Rockland, Maine, EE.UU.). Se eluyeron fragmentos de aproximadamente 10 kb de tamaño usando el sistema GELase (Epicentre Technologies, Madison, EE.UU.), siguiendo las instrucciones del proveedor. Se clonaron los fragmentos purificados en el vector binario pCLD04541 linealizado con *EcoRI*, se desfosforilaron usando la fosfatasa SHRIMP (Roche, Alemania) y se transformaron en la cepa de *E. coli* DH10B (Life Technologies, EE.UU.). Se recogieron doscientas colonias recombinantes en placas de microtitulación.

Construcción y examen de la biblioteca de ADNc

Se infectaron brotes cortados de plantas de aproximadamente 8 semanas de edad del original P41 (*R1r1*) y del cv Desirée susceptible con la variedad 4 de *P. infestans* y se mantuvieron en un cilindro de vidrio (para aumentar la humedad) en agua en una cámara de crecimiento a 17°C con 16 h de luz. En estas condiciones, se produjo un sobrecrecimiento de micelio de *P. infestans* sobre las hojas del control susceptible tras 8 días. Se recogieron cantidades iguales de hojas no infectadas del original P41 y hojas infectadas 2 h, 19 h, 3 d, 7 d y 9 d tras la inoculación. Se aisló el ARN de Poli A⁺ usando el RNeasy Plant Mini Kit o el Oligotex mRNA Mini Kit (Qiagen, Hilden, Alemania) según las instrucciones del proveedor. Se construyó una biblioteca de ADNc Lambda ZAP II (Stratagene, CA, EE.UU.) a partir del ARN de poli-A⁺, siguiendo las instrucciones del fabricante. Se agruparon las diferentes preparaciones de ADNc antes

del ligamiento en el vector Lambda ZAP. Se sembraron en placa 5×10^5 ufp y se examinaron mediante hibridación de placas (Sambrook *et al.* 1989) usando como sonda las inserciones de BAC BA121o1 y BA76o11.

Análisis de amplificación rápida en 5' de extremos de ADNc (RACE):

5 Se aisló el ARN total a partir de tejido de hojas no infectadas de P41 (*R1r1*) usando el RNeasy Plant Mini Kit (Qiagen, Hilden, Alemania) según las instrucciones del proveedor. Se realizó el análisis de RACE con 1 μ g de ARN total usando el kit de amplificación de ADNc SMART TM Race (Clontech, CA, EE.UU.) siguiendo las instrucciones del fabricante. Los cebadores específicos de gen anidados usados para la amplificación por PCR fueron en primer lugar RT1-1: 5'-AAACCCGGTGTCCAAATCTAACACT-3' (SEQ ID NO: 3) y en segundo lugar RT2-1: 5'-CATGTAGTGAGGATATGTCACGAGTG-3'. (SEQ ID NO: 4). Se clonaron los productos de PCR finales de la reacción de RACE en el vector pGEM-T (Promega, CA, EE.UU.). Se secuenciaron dos clones independientes.

10 Análisis de la secuencia de ADN:

15 La secuenciación del ADN por encargo la realizó la unidad ADIS en el Instituto Max Planck para la Investigación de Cultivos. Se empleó el método de secuenciación por terminación de cadena didesoxi usando un ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit y un secuenciador de ADN automatizado ABI377 (PE Biosystems, Foster City, CA EE.UU.).

20 Se realizó el análisis de la secuencia de ADN usando el paquete Wisconsin versión 10.0, Genetics Computer Group (GCG), Madison, Wisc, EE.UU. Se realizaron búsquedas en bases de datos de secuencias con BlastX y otros algoritmos disponibles a través del Centro Nacional de Información Biotecnológica, Bethesda, MD, EE.UU. y el servidor web ExPASY (Appel *et al.* 1994).

25 Transformación de *Agrobacterium tumefaciens*:

30 Se sometió a electroporación el subclón g10 de BAC BA87d17 en la cepa de *A. tumefaciens* LBA4404 según Wen-jun y Forde (1989). Se usaron tres cepas de *Agrobacterium*, LBAg10-2, LBAg10-5 y LBAg10-23 para la transformación de patatas.

35 Transformación de patatas mediada por *Agrobacterium tumefaciens* y análisis de plantas transgénicas:

40 Se usó el cultivar de patata susceptible Desirée en todos los experimentos de transformación. Se realizó la transformación mediada por *Agrobacterium tumefaciens* tal como se describe por Rocha-Sosa *et al.* (1989), excepto porque el medio MS contenía 250 mg/l de Cloran. Se sometieron a prueba plantas transgénicas resistentes a kanamicina mediante la reacción en cadena de la polimerasa (PCR) para determinar la presencia del inserto de g10 usando los cebadores específicos de inserto 87e (5'-ATTACAATGGGTTGAACTCAG-3' (SEQ ID NO: 5)) y 87s (5'-ACCTCTTCATTGTTCTGGT-3' (SEQ ID NO: 6)). Las condiciones de la PCR eran: Ta a 55°C durante 45 s y polimerización a 72°C durante 60 segundos. Se examinaron plantas transgénicas con los cebadores específicos de R1 76-2sf2 (5'-CACTCGTGACATATCCTCACTA-3' (SEQ ID NO: 7)) y 76-2SR (5'-CAACCCCTGGCATGCCACG-3' (SEQ ID NO: 8)) derivados del ADNc c76-2. Las condiciones de la PCR eran: Ta a 55°C durante 45 s y polimerización a 72°C durante 90 s. Se realizaron pruebas para determinar la resistencia a la variedad 4 de *P. infestans* usando tres foliolos por planta en cada prueba.

45 **Ejemplo 1: Mapeo genético de alta resolución del locus *R1***

50 Para facilitar el mapeo físico del locus *R1*, se seleccionaron 16 recombinantes entre los marcadores GP21 y GP179 que flanqueaban a *R1* (Leonards-Schippers *et al.* 1992) a partir de 588 plantas y se sometieron a prueba para determinar la resistencia a un *P. infestans* que tenía el factor de virulencia correspondiente Avr1. Junto con 15 recombinantes previamente seleccionados en el mismo intervalo (Meksem *et al.* 1995), estaban disponibles 31 recombinantes en total en el intervalo GP21 - GP179 a partir de 1049 plantas, correspondientes a una frecuencia de recombinación del 3,0% (3 cM). Las frecuencias de recombinación entre GP21 y *R1* y entre *R1* y GP179 eran del 2,2% y el 0,8%, respectivamente (tabla 1).

Tabla 1. Número de individuos recombinantes en los intervalos *GP21-R1*, *GP179-R1* y *GP21-GP179*, seleccionados entre 1049 plantas de una población F1 de segregación.

	<i>GP21-R1</i>	<i>GP179-R1</i>	<i>GP21-GP179</i>
Número de recombinantes	23	8	31
Recombinantes con genotipo <i>R1r1</i>	12	4	16
Recombinantes con genotipo <i>r1r1</i>	11	4	15
Frecuencia de recombinación (%)	2,2	0,8	3,0

Los marcadores *SPUD237* y *AFLP1*, que se mapean ambos en el intervalo *GP21 - GP179* (De Jong *et al.* 1997, Meksem *et al.* 1995) flanquean el locus *R1*. Ambos marcadores estaban separados de *R1* por un acontecimiento de recombinación en 1049 plantas (0,1 cM, figura 1).

Ejemplo 2: Paseo cromosómico hacia el locus *R1* e identificación de un gen *R1* candidato

Se usó el marcador *SPUD237* como sonda para el examen de la biblioteca de cósmidos. Se identificó un clon CosS positivo (figura 1). La secuenciación de extremos del inserto de CosS generó un nuevo marcador separado por un acontecimiento de recombinación (0,1 cM) del locus *R1*. El examen de la biblioteca de BAC con este marcador identificó el clon de BAC BA100e13. Tres acontecimientos de recombinación separaban el extremo distal de BA100e13 de *R1*. El extremo de BA100e13 proximal con respecto a *R1* se identificó como BA47f2. El extremo de BA47f2 distal con respecto a *R1* se solapaba con BA100e13 y estaba separado de *R1* por un acontecimiento de recombinación. El extremo proximal se segregaba conjuntamente con *R1*, como todos los extremos de BAC posteriores analizados (parte derecha de la figura 1). El extremo de BA47f2 que se segregaba conjuntamente con *R1* se identificó como el clon BA27c1. El extremo de BA27c1 que no se solapaba con BA47f2 se identificó como los clones BA122p13 y BA121o1. El extremo de BA121o1 que no se solapaba con BA27c1 mostró una similitud de secuencia altamente significativa (37% de identidad, 56% de similitud de la secuencia de aminoácidos traducida) con el gen *Prf* de tomate para la resistencia a *Pseudomonas syringae* (Salmeron *et al.* 1996). Se usó este fragmento similar a gen de resistencia (RGL) como sonda para volver a examinar la biblioteca de BAC. La sonda RGL identificó, además de BA122p13, varios nuevos clones positivos de los que dos, BA87d17 y BA76o11, se analizaron adicionalmente. Contenían copias de longitud completa del gen RGL que se concebió como un posible candidato a *R1*. Los extremos no solapantes de BA76o11 y BA87d17 se segregaban conjuntamente con *R1*.

Se usaron también marcadores de extremos de BAC decisivos para la construcción del mapa físico para asignar clones de BAC al cromosoma P6/210 (*R1r1*) que portaba o bien un alelo *r1* o bien el *R1*. Los clones BA100e13, BA47f2 y BA87d17 (figura 1) estaban en *cis* con el alelo *R1*, mientras que los clones BA121o1, BA122p13 y BA76o11 se derivaban del homólogo que tenía *r1* (figura 1). El clon BA27c1 no pudo asignarse a un cromosoma *r1* o *R1*, basándose en los marcadores usados.

Ejemplo 3: Clones de ADNc candidatos a *R1*

Usando las inserciones completas de BAC BA121o1 y BA76o11 como sondas, se aislaron seis y ocho clones de ADNc, respectivamente, a partir de una biblioteca de ADNc preparada a partir de hojas infectadas del genotipo P41 (*R1r1*). Ocho de los 14 clones de ADNc eran similares a genes de resistencia de plantas conocidos. Se obtuvo la mayor similitud con el gen *Prf* de tomate para la resistencia a *Pseudomonas syringae* (Salmeron *et al.* 1996). Las secuencias de los ocho ADNc candidatos compartían aproximadamente un 80-90% de identidad entre sí. El clon de ADNc c76-2, de 2292 nucleótidos de largo, era idéntico, con la excepción de los intrones, a la secuencia genómica del clon g10, un subclón que representa parte de BA87d17 (véase a continuación). La comparación de secuencias con genes de resistencia conocidos en la base de datos indicó que c76-2 no era de longitud completa. Usando análisis de RACE, se extendió el ADNc hasta el extremo 5' mediante 1943 nucleótidos, dando como resultado una secuencia de ADNc de longitud completa de 4235 nucleótidos incluyendo una región no traducida en 5' de 59 nucleótidos y una secuencia no traducida en 3' de 297 nucleótidos entre el codón de terminación la cola de poli-A. El ADNc incluía un codón de iniciación en la posición 2223 de la secuencia genómica correspondiente a la primera metionina en la secuencia de aminoácidos deducida a partir del clon g10 (figura 4B). Se identificaron dos adeninas en las posiciones -3 y +4 (en las que la a del ATG es +1) denominadas secuencia de reconocimiento del ribosoma en plantas, insectos, levadura y mamíferos (Kozak 1991).

5 Se diseñaron cebadores de PCR específicos para el ADNc de c76-2 basándose en una alineación de secuencias con los otros siete ADNc candidatos. Los cebadores 76-2sf2 y 76-2SR (véase material y métodos) generaron un producto de PCR de 1,4 kb sólo en el original P41 (*R1r1*) pero no en el original P40 (*r1r1*) (figura 2). Este polimorfismo sugería la posibilidad de que BAC BA87d17 (derivado del cromosoma que alberga R1) contenía el gen *R1*, incluso si los datos de mapeo todavía indicaban la ausencia de recombinación entre el extremo distal de este clon de BAC y R1.

Ejemplo 4: Complementación del fenotipo R1

10 Se construyó una sub-biblioteca genómica en el vector binario pCLD04541 (véase Material y métodos) con, en promedio, inserciones de 10 kb a partir de BA87d17 (76 kb). Se examinó la biblioteca mediante hibridación de colonias con la sonda RGL de BA121o1. Se evaluaron clones positivos para determinar la presencia de la copia completa del gen RGL candidato mediante el tamaño de los productos de amplificación mediante PCR con cebadores directos a partir de los bordes del vector (T3 y T7) y cebadores inversos a partir del RGL. También se sometieron a prueba los clones usando los cebadores específicos de ADNc de c76-2 76-2sf2 y 76-2SR. Se seleccionó el subclón g10-2 y se transformó en *A. tumefaciens*. Se usaron tres colonias bacterianas diferentes para transformar el cultivar susceptible Desirée. A partir de tres experimentos de transformación, se regeneraron quince líneas transgénicas independientes y se sometieron a prueba en cuatro experimentos independientes para determinar la expresión de resistencia a la variedad 4 de *P. infestans* (tabla 2).

15

20 **Tabla 2.** Prueba para la resistencia a la variedad 4 de *P. infestans* de líneas de patata transgénicas transformadas con el clon g10. Se sometieron a prueba las líneas transgénicas en cuatro experimentos independientes con tres folíolos a partir de cada línea para la expresión de resistencia hipersensible a la variedad 4 de *P. infestans*.

Línea transgénica n. ^º	Resistencia ^c
10-2 ^a 1 ^b	S
10-2 2	R
10-2 3	R
10-2 4	R
10-5 1	R
10-5 2	S
10-5 3	n.d.
10-5 4	n.d.
10-5 5	R
10-23 1	n.d.
10-23 2	R
10-23 3	R
10-23 4	R
10-23 5	R
10-23 6	S

25 Nueve líneas transgénicas mostraban de manera constante una respuesta HR típica, similar a la línea resistente P41 que alberga *R1* (figura 3); tres dieron resultados inconstantes y las tres restantes eran susceptibles, como el control Desirée no transformado. Basándose en estos resultados, se concluyó que el subclón g10 contenía un gen *R1* funcional.

Todas las líneas transgénicas con el fenotipo *R1* contenían el gen correspondiente al ADNc de 76-2, tal como se demostró por la presencia del producto de PCR de 1,4 kb amplificado por los cebadores 76-2sf2 y 76-2SR. Este

producto estaba ausente en Desirée no transformado (control) y en todos los clones de BAC notificados en la figura 1 como miembros del cónigo alrededor de R1, excepto por BA87d17 (figura 2).

Ejemplo 5: Estructura del gen R1

Se secuenció el subclón g10 que contenía el gen R1; véase SEQ ID NO: 1. La secuencia tenía 10.388 nucleótidos de largo y contenía un gen con similitud de secuencia con otros genes de resistencia de plantas. No se identificó ningún otro marco de lectura abierto o secuencia de homología en la base de datos GenEMBL. La alineación de secuencia con el ADNc de c76-2 y el producto de 5' RACE reveló la presencia de tres exones y tres intrones. Dos intrones de 92 pb (posición 4878 a 4970) y 126 pb (posición 6103 a 6229) interrumpen la región codificante. El tercer intrón de 81 pb (posición 6323 a 6404) está ubicado en la región no traducida en 3' inmediatamente en el sentido de 3' del codón de terminación (figura 4A). La secuencia de aminoácidos deducida pronostica un polipéptido de 1293 aminoácidos con una masa molecular de 149,4 kDa (figura 4B). La secuencia de aminoácidos deducida del gen R1 es lo más similar (40% de identidad) al gen *Prf* para la resistencia a *P. syringae* de tomate (Salmeron *et al.* 1996). La proteína R1 pronosticada tiene un supuesto dominio de sitio de unión a nucleótidos (NBS) que consiste en motivos de bucle P (aminoácidos 572-578), quinasa 2 (aminoácidos 649-653) y quinasa 3a (aminoácidos 677-682) (figura 4B). En el sentido de 3' de los motivos de quinasa estaban otras secuencias con similitud con dominios de función desconocida conservada entre genes de resistencia: GPL (QLPL (SEQ ID NO: 10) en R1), CKLY (CFLY (SEQ ID NO: 12) en R1) y MHD (LHD en R1). Al buscar motivos conservados usando el algoritmo ExPASY, se encontraron 4 supuestos sitios de miristilación, 9 de glicosilación, 43 de fosforilación y 1 de amidación en la secuencia de aminoácidos de R1 deducida. El supuesto dominio de repetición rica en leucina (LRR) de R1 tiene 15-16 repeticiones imperfectas ubicadas en la parte carboxi-terminal del gen. Como algunas proteínas R de plantas con LRR citoplasmáticas, la proteína R1 contiene una cremallera de leucina desde la posición 308 hasta 329 (Hammond-Kosack y Jones, 1997).

Ejemplo 6: Organización genómica del locus R1

El análisis de transferencia en gel de tipo Southern mostró que R1 es un miembro de una familia génica.

Los cebadores específicos de R1 76-2sf2 y 76-2SR amplificaron el fragmento de 1,4 kb en BA87d17 (R1) pero no en los clones solapantes BA121o1, BA122p13 y BA76o11 (r1) (figura 2). El análisis de la secuencia de ADN de BAC BA87d17 (R1) y BA122p13 (r1) reveló que BA87d17 contenía dos miembros altamente homólogos de la familia del gen R1, correspondiendo R1 al gen R1 funcional y siendo r1.1 ortólogo con el alelo r1.2 en BA122p13. El gen R1 funcional era parte de una inserción de 15 kb presente en el cromosoma que lleva R1 en la región representada por BA87d17, pero estaba ausente en el cromosoma que alberga r1 (figura 5).

BIBLIOGRAFÍA

- Appel, R-D., Bairoch, A. and Hochstrasser, D.F. (1994). A new generation of information retrieval tools for biologists: the example of the ExPASY www server. *Trends Biochem. Sci.* 19:258-260
- Balvora, A., Schornack, S., Baker, B.J., Ganal, M., Bonas, U. and Lahaye, T. (2001) Chromosome landing at the tomato Bs4 locus. *Mol Gen Genet*, in press.
- Bendahmane, A., Kanyuka, K. and Baulcombe, D.C. (1999) The Rx gene from potato controls separate virus resistance and cell death responses. *The Plant Cell* 11:781-791
- Bendahmane, A., Querci, M., Kanyuka, K. and Baulcombe, D.C. (2000) Agrobacterium transient expression system as a tool for the isolation of disease resistance genes: application to the Rx2 locus in potato. *The Plant Journal* 21: 73-81
- Dangl, J.L. and Jones, J.D.G. (2001). Plant pathogens and integrated defence responses to infection. *Nature* 411:826-833
- De Jong, W., Forsyth, A., Leister, D., Gebhardt, C. and Baulcombe, D.C. (1997). A potato hypersensitive resistance gene against potato virus X maps to a resistance gene cluster on chromosome V. *Theor Appl Genet* 95:153-62.
- Dong F, Song J, Naess SK, Helgeson JP, Gebhardt C, Jiang J (2000) Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato. *Theor Appl Genet* 101: 1001-07.
- El-Kharbotly, A., Leonards-Schippers, C., Huigen, D.J., Jacobsen, E., Pereira, A., Stiekema, W.J., Salamini, F. and Gebhardt, C. (1994). Segregation analysis and RFLP mapping of the R1 and R3 alleles conferring race specific resistance to *Phytophthora infestans* in progenies of dihaploid potato parents. *Mol. Gen. Genet.* 242: 749-754.
- El-Kharbotly, A., Palomino-Sánchez, C., Salamini, F., Jacobsen, E. and Gebhardt, C. (1996). R6 and R7 alleles of potato conferring race-specific resistance to *Phytophthora infestans* (Mont.) de Bary identified genetic loci clustering with the R3 locus on chromosome XI. *Theor Appl Genet* 92: 880-884.
- Ellis, J.G., Lawrence, G.J., Luck, J.E. and Dodds, P.N. (1999). Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *Plant Cell* 11:495-506

- Ellis, J.G., Dodds, P.N. and Pryor, T. (2000). Structure, function and evolution of plant disease resistance genes. *Curr Opin Plant Biol* 3:278-284
- 5 Ewing, E.E., Simko, I., Smart, C.D., Bonierbale, M.W., Mizubuti, E.S.G., May, G.D., and Fry, W.E. (2000). Genetic mapping from field tests of quantitative and qualitative resistance to *Phytophthora infestans* in a population derived from *Solanum tuberosum* and *Solanum berthaultii*. *Mol Breeding* 6: 25-36.
- Feinberg, A.P. and Vogelstein, B. (1984). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity (addendum). *Anal Biochem* 137: 266-267.
- 10 Flor, H.H (1971) Current status of the gene-for-gene concept. *Annu Rev Phytopathol* 9: 275-296
- Folkertsma, R.T., Spassova, M.I., Prins, M., Stevens, M.R., Hille, J. and Goldbach R.W. (1999). Construction of a bacterial artificial chromosome (BAC) library of *Lycopersicon esculentum* cv. Stevens and its application to physically map the Sw-5 locus. *Molecular Breeding* 5:197-207
- 15 Fry, W.E. and Goodwin, S.B. (1997). Resurgence of the Irish potato famine fungus. *Bioscience* 47: 363-371.
- Gebhardt, C., Ritter, E., Debener, T., Schachtschabel, U., Walkemeier, B., Uhrig, U. and Salamini, F. (1989). RFLP analysis and linkage mapping in *Solanum tuberosum*. *Theor Appl Genet* 78:65-75
- 20 15 Gebhardt, C. and Valkonen, J.P.T. (2001). Organization of genes controlling disease resistance in the potato genome. *Annu Rev Phytopathol.* 39: 79-102.
- Hammond_Kosack, K. and Jones, J.D. (1997). Plant disease resistance genes *Annu Rev Plant Physiol Plant Mol Biol* 48: 575-607
- 25 Jones, D.A., Thomas, C.M., Hammond-Kosack. K.E., Balint-Kurti, P.J. and Jones, D.J.G.(1992). Effective vectors for transformation, expression of heterologous genes, and assaying transposon excision in transgenic plants. *Transgenic Res.* 1:285-297
- Jones, D.A. and Jones, D.J.G.(1997). The roles of leucine rich repeats in plant defences. *Adv Bot Res Adv Plant Pathol* 24:90-167
- 25 25 Judelson, H.S. (1997). The genetics and biology of *Phytophthora infestans*: Modern approaches to a historical challenge. *Fungal Genet and Biol* 22: 65-76.
- Kamoun, S. (2001). Nonhost resistance to *Phytophthora*: Novel prospects for a classical problem. *Curr Opin Plant Biol* 4:295-300
- Kreike, C.M., De Koning, J.R.A., Vinke, J.H., Van Ooijen, J.W., and Stiekema, W.J. (1994). Quantitatively-inherited resistance to *Globodera pallida* is dominated by one major locus in *Solanum spegazzinii*. *Theor Appl Genet* 88: 764-69.
- 30 30 Kozak, M. (1991) Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J Biol Chem* 266:19867-19870
- Leister, D., Ballvora, A., Salamini., F, and Gebhardt, C. (1996). A PCR based approach for isolating pathogen resistance genes from potato with potential for wide application in plants. *Nature Genetics* 14: 421-429.
- 35 35 Leister, D., Berger, A., Thelen, H., Lehmann, W., Salamini, F. and Gebhardt, C. (1997). Construction of a potato YAC library and identification of clones linked to the disease resistance loci R1 and Gro1. *Theor Appl Genet* 95: 954-960.
- Leonards-Schippers, C., Gieffers, W., Gebhardt, C. and Salamini, F. (1992). The R1 gene conferring race-specific resistance to *Phytophthora infestans* in potato is located on potato chromosome V. *Mol Gen Genet* 233:278-283
- 40 40 Leonards-Schippers, C., Gieffers, W., Schäfer-Pregl, R., Ritter, E., Knapp, S.J., Salamini, F. and Gebhardt, C. (1994). Quantitative resistance to *Phytophthora infestans* in potato: a case study for QTL mapping in an allogamous plant species. *Genetics* 137: 67-77.
- Li,X., van Eck, H.J., Rouppe van der Voort, J., Huigem, D-J., Stam, P., and Jacobsen, E. (1998). Autotetraploids and genetic mapping using common AFLP markers: The R2 allele conferring resistance to *Phytophthora infestans* mapped on potato chromosome 4. *Theor Appl Genet* 96: 1121-28.
- 45 45 Meksem, K. Leister, D. Peleman, J., Zabeau, M., Salamini, F. and Gebhardt, C. (1995). A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers *Mol Gen Genet* 249:74-81
- Meksem, K., Zobrist, K., Ruben, E. Hyten, D., Quanzhou, T., Zhang, H-B. and Lightfoot, D.A. (2000). Two large-insert soybean genomic libraries constructed in a binary vector: applications in chromosome walking and genome wide physical mapping. *Theor Appl Genet* 101:747-755

- Naess, S.K., Bradeen, J.M., Wielgus, S.M., Haberlach, G.T., McGrath, J.M., and Helgeson, J.P. (2000). Resistance to late blight in *Solanum bulbocastanum* is mapped to chromosome 8. *Theor Appl Genet* 101: 697-704.
- Nakamura, S., Asakawa, S., Ohmido, N., Fukui, K., Shimizu, N. and Kawasaki, S. (1997). Construction of an 800-kb contig near-centromeric region of the rice blast resistance gene Pi-ta<2> using a highly representative rice BAC library. *Mol Gen Genet* 254:611-620
- Oberhagemann, P., Chatot-Balandras, C., Bonnel, E., Schäfer-Pregl, R., Wegener, D., Palomino, C., Salamini, F. and Gebhardt, C. (1999). A genetic analysis of quantitative resistance to late blight in potato: Towards marker assisted selection. *Mol Breed* 5: 399-415.
- Person, C., Samborski, D.J. and Rohringer, R. (1962) The gene-for-gene concept. *Nature* 194: 561-562
- 10 Rocha-Sosa, M., Sonnewald, U., Frommer, W., Stratmann, M., Schell, J. and Willmitzer, L. (1989). Both developmental and metabolic signals activate the promoter of a class I patatin gene. *The EMBO Journal* 8 (1):23-29.
- Ross, H. (1986). Potato breeding. Problems and perspectives. *Adv Plant Breed, Supplement* 13.
- 15 Reiser, L., Modrusan, Z., Margossian, L., Samach, A., Ohad, N., Haughn, G.W. and Fischer, R. (1995). The BELL1 Gene Encodes a Homeodomain Protein Involved in Pattern Formation in the *Arabidopsis* Ovule Primordium. *Cell* 83:735-742
- Ritter, E., Debener, T., Barone, A., Salamini, F. and Gebhardt, C. (1991). RFLP mapping on potato chromosomes of two genes controlling extreme resistance to potato virus X (PVX). *Mol Gen Genet* 227: 81-85.
- 20 Rouppe van der Voort, J., Wolters, P., Folkertsma, R., Hutton, R., van Zandvoort, P., Vinke, H., Kanyuka, K., Bendahmane, A., Jacobsen, E., Janssen, R. and Bakker, J. (1997). Mapping of the cyst nematode resistance locus Gpa2 in potato using a strategy based on comigrating AFLP markers. *Theor Appl Genet* 95: 874-80.
- Rouppe van der Voort, J., van der Vossen, E., Bakker, E., Overmars, H., van Zandvoort, P., Hutton, R., Klein-Lankhorst, R. and Bakker, J. (2000). Two additive QTLs conferring broad-spectrum resistance in potato to *Globodera pallida* are localized on resistance gene clusters. *Theor Appl Genet* 101: 1122-30.
- 25 Salaman, R.N. (1985). The potato famine: its causes and consequences. In *The History and Social Influence of the Potato*, revised impression, ed. JG Hawkes, pp 289-316. Cambridge/ New York/New Rochelle/Melbourne/Sydney: Cambridge University Press.
- Salmeron, J.M., Oldroyd, G.E., Rommens C.M., Scofield, S.R., Kim, H-S., Lavelle, D.T., Dahlbeck, D. and Staskawicz, B.J. (1996). Tomato Prf is a member of the leucine-rich repeat class of plant disease resistance genes and lies embedded within the Pto kinase gene cluster. *Cell*, Vol. 86:123-133
- 30 Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). *Molecular Cloning: A laboratory Manual*. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
- Singh, A.K., Salamini, F. and Uhrig, H. (1989). Chromosome pairing in 14 F1 hybrids among 11 diploid potato species. *J Genet Breed* 43:1-5
- 35 Stahl, E.A., Dwyer, G., Mauricio, R., Kreitman, M. and Bergelson, J. (1999) Dynamics of disease resistance polymorphism at the Rpm1 locus in *Arabidopsis*. *Nature* 400:667-671
- Tanksley SD, Ganal MW, Prince JP, de Vicente MC, Bonierbale MW, Broun P, Fulton TM, Giovannoni JJ, Grandillo S, Martin GB, Messeguer R, Miller JC, Miller L, Paterson AH, Pineda O, Röder MS, Wing RA, Wu W, Young ND. 1992. High density molecular linkage maps of the tomato and potato genomes. *Genetics* 132: 1141-60.
- 40 Van der Lee, T., Robold, A., Testa, A., van't Klooster, J.W. and Govers, F. (2001). Mapping of Avirulence Genes in *Phytophthora infestans* With Amplified Fragment Length Polymorphism Markers Selected by Bulked Segregant Analysis. *Genetics* 157: 949-956
- Van der Vossen, E.A.G., Rouppe van der Voort, J.N.A.M., Kanyuka, K., Bendahmane, A., Sandbrink, H., Baulcombe, D.C., Bakker, J., Stiekema, W.J. and Klein-Lankhorst, R.M. (2000) Homologues of a single resistance-gene cluster in potato confer resistance to distinct pathogens: a virus and a nematode. *The Plant Journal* 23: 567-576
- 45 Wastie, R.L. (1991). Breeding for resistance. *Adv Plant Pathol* 7: 193-223.
- Wen-jun, S. and Forde, B.G. (1989). Efficient transformation of *Agrobacterium* spp. By high voltage electroporation. *Nuc Acids Res* 17:8385
- Yang, D., Parco, A., Nandi, S., Subudhi, P., Zhu, Y., Wang, G. and Huang, N. (1997). Construction of a bacterial artificial chromosome (BAC) library and identification of overlapping BAC clones with chromosome 4-specific RFLP markers

ES 2 375 933 T3

Yang, D., Sanzhes, A., Khush, G.S., Zhu, Y., and Huang, N. (1998). Construction of a BAC contig containing the xa5 locus in rice. *Theor Appl Genet* 97:1120-1124

REIVINDICACIONES

- 1 Molécula de ácido nucleico que codifica para un polipéptido que puede conferir resistencia frente a un patógeno en una planta en la que dicho polipéptido se expresa, comprendiendo o consistiendo dicha molécula de ácido nucleico en una secuencia de nucleótidos seleccionada del grupo que consiste en:
- 5 (a) una secuencia de nucleótidos que codifica para al menos la forma madura de una proteína (R1) que comprende la secuencia de aminoácidos facilitada en SEQ ID NO: 2;
- (b) una secuencia de nucleótidos que comprende las regiones codificantes de la secuencia de ADN facilitada en SEQ ID NO: 1;
- 10 (c) una secuencia de nucleótidos que se hibrida con la cadena complementaria de una secuencia de nucleótidos definida en (a) o (b) en condiciones de hibridación rigurosas, en la que dicha secuencia de nucleótidos es homóloga en al menos el 80% con las secuencias de nucleótidos de (a) o (b);
- (d) una secuencia de nucleótidos que codifica para una proteína derivada de la proteína codificada por una secuencia de nucleótidos de (a) o (b) por medio de sustitución, delección y/o adición de uno o varios aminoácidos de la secuencia de aminoácidos codificada por la secuencia de nucleótidos de (a) o (b), en la que dicha secuencia de nucleótidos es homóloga en al menos el 80% con las secuencias de nucleótidos de (a) o (b);
- 15 (e) una secuencia de nucleótidos que codifica para una proteína que tiene una secuencia de aminoácidos idéntica en al menos el 60% a la secuencia de aminoácidos codificada por la secuencia de nucleótidos de (a) o (b);
- 20 (f) una secuencia de nucleótidos que codifica para al menos un dominio de cremallera de leucina (LZ) correspondiente a la posición de aminoácidos 308-329 de SEQ ID NO: 2, un dominio de sitio de unión a ácido nucleico (NBS) correspondiente a la posición de aminoácidos 572-682 de SEQ ID NO: 2 y/o un dominio de repetición rica en leucina (LRR) correspondiente a la posición de aminoácidos 780-1280 de SEQ ID NO: 2;
- 25 (g) una secuencia de nucleótidos que codifica para una parte que lleva un epítopo de una proteína (R1) codificada por una secuencia de nucleótidos de (a) o (b), en la que dicha secuencia de nucleótidos es homóloga en al menos el 80% con las secuencias de nucleótidos de (a) o (b);
- (h) cuya secuencia de nucleótidos está degenerada como resultado del código genético con respecto a una secuencia de nucleótidos de una cualquiera de (a) a (g),
- 30 en la que dicho patógeno se selecciona del grupo que consiste en *Phytophthora sojae*, *Peronospora parasitic*, *Magnaporthe grisea*, *Erysiphe spp*, *Pseudomonas syringae*, *Erwinia amylovora*, *Erwinia carotovora*, *Botrytis cinerea*, *Rhizoctonia solani*, *Pythium debaryanum* y *Phytophthora infestans*.
2. Molécula de ácido nucleico según la reivindicación 1, en la que dicho patógeno es *Phytophthora infestans*.
3. Vector que comprende una molécula de ácido nucleico según la reivindicación 1 ó 2.
- 35 4. Vector según la reivindicación 3, que es un vector de expresión en el que la molécula de ácido nucleico está operativamente unida a una o más secuencias de control que permiten la transcripción y opcionalmente la expresión en células huésped procariotas y/o eucariotas.
5. Célula huésped que contiene un vector según la reivindicación 3 ó 4 o molécula de ácido nucleico según la reivindicación 1 ó 2.
- 40 6. Proteína o fragmento funcional de la misma que puede conferir resistencia frente a un patógeno en una planta en la que se expresa, codificada por una molécula de ácido nucleico según la reivindicación 1 ó 2.
7. Anticuerpo o aptámero que reconoce específicamente la proteína según la reivindicación 6.
8. Célula vegetal transgénica que comprende una molécula de ácido nucleico según la reivindicación 1 ó 2 que está operativamente unida a elementos reguladores que permiten la transcripción y/o expresión de la secuencia de ADN en células vegetales.
- 45 9. Planta transgénica o tejido vegetal que comprende células vegetales según la reivindicación 8.
10. Secuencia reguladora de un promotor que regula la expresión de un gen que comprende la molécula de ácido nucleico según la reivindicación 1 ó 2, pudiendo conferir o modular dicha secuencia reguladora la expresión de una secuencia de ADN heteróloga tras la infección con patógenos y comprendiendo la secuencia de ADN de SEQ ID NO.1 desde los nucleótidos 1 hasta 2222 o desde los nucleótidos 1 hasta 2164.

11. Molécula de ADN recombinante que comprende la secuencia reguladora según la reivindicación 10.
12. Molécula de ADN recombinante según la reivindicación 11, en la que dicha secuencia reguladora está operativamente unida a una secuencia de ADN heteróloga.
- 5 13. Célula huésped transformada con una secuencia reguladora según la reivindicación 10 o con una molécula de ADN recombinante según la reivindicación 11 ó 12.
14. Planta transgénica, tejido vegetal o célula vegetal que comprende la secuencia reguladora según la reivindicación 10 o la molécula de ADN recombinante según la reivindicación 11 ó 12.
15. Método para la identificación de un agente protector de plantas que comprende las etapas de:
- 10 (a) cultivar un tejido o célula vegetal o mantener una planta que comprende una molécula de ADN recombinante que comprende un sistema de lectura operativamente unida a una secuencia reguladora según la reivindicación 10 en presencia de un compuesto o una muestra que comprende una pluralidad de compuestos en condiciones que permiten la expresión de dicho sistema de lectura;
- 15 (b) identificar o verificar una muestra y compuesto, respectivamente, que conduce a la supresión o activación y/o potenciación de la expresión de dicho sistema de lectura en dicha planta, célula vegetal o tejido vegetal.
16. Método para identificar y obtener un factor de avirulencia o de virulencia de un patógeno que comprende las etapas de:
- 20 (a) examinar la proteína R1 según la reivindicación 6 o un fragmento de la misma frente a una biblioteca de expresión de péptidos o proteínas derivada de un patógeno en un sistema de lectura en condiciones adecuadas que permiten la interacción de la proteína y el péptido en dicho sistema de lectura;
- (b) identificar o verificar un ADNc que conduce a la supresión o activación del sistema de lectura.

LISTA DE SECUENCIAS

<110> Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.

KWS Saat AG

5 <120> Gen de resistencia derivado de planta

<130> F 2298 PCT

<150> EP 01120670.3

10 <151> 31-08-2001

<160> 12

<170> PatentIn versión 3.1

15

<210> 1

<211> 10388

<212> ADN

<213> *Solanum tuberosum*

20

<220>

<221> Intrón

<222> (4878)..(4970)

<223>

25

<220>

<221> Intrón

<222> (6103)..(6229)

<223>

30

<220>

<221> Intrón

<222> (6323)..(6404)

<223>

35

<220>

<221> CDS

<222> (2223)..(4877)

<223>

<220>

5 <221> *CDS

<222> (4971)..(6102)

<223>

<220>

10 <221> CDS

<222> (6230)..(6321)

<223>

<220>

15 <221> 5'-UTR

<222> (2164)..(2222)

<223>

<220>

20 <221> promotor

<222> (1)..(2164)

<223>

<220>

25 <221> 3'-UTR

<222> (6405)..(6702)

<223>

<400> 1

ES 2 375 933 T3

ttaatatata gatggaatcg gtgtttaaa aggcagggcg cgaggcgaga cgtttactt	60
agtatagagc gaggcgtaag cctaatgatt cattttctta agaacatata taattcatta	120
aattacttaa ttataataaa ttatataact tcaaatacac ttggatataa ataagtaatt	180
atccttcacg agattcaaat gaaaaataag tggagttat tagagtaag tagagtaatt	240
taaacacttt agtctgaatc ttatatacatt atacaaaaaa agtaattata ttccaccaaa	300
ttcaaataatgga aattaaaaat atatgaagat aattacaaca caagtgttat gtgtcagttg	360
gaaagctcaa gcgtgggtcc taccatactc catgacattt cacttttagg gtatgattcg	420
taatttaatg aaaatgatga cttttttttt tggagttatg aatgaggctt aaataactaa	480
acatagagga caaccctttt aagcaagcaa atcttgcattt aacatttcaa agaccatgt	540
atccctcaat ttttttattaa tgactaaaaa ctaacatgtt aaactctcct gtgtattatt	600
cattgtataa ttttttttgg ttaaatcattt cattgtaaaa taaattcattt atacataatg	660
ttaattttt cttataatc aaatattattt catcgatata ttactaaaaa tattcaatgt	720
atgatgatga gagaataaac tatataagaa atataagaaa ttatgaaa ccatattcaa	780
aatggcttc tcaatgtgtc aaaaaatcaa caatgacaga tcaaatacat tatcttattt	840
ctttaatttg tgtagatata ttgttactt ttagaggattt attaattttt aatatcaatg	900
aagccatata ttatataag agtctctga aaatataattt tcttattttt ttatcaaaat	960
ggatgatttt ttccattgtat ccaagtcggg accaaaaaaag aatattatct caaagagatt	1020
attaatttttcaaaatattaa tttagtgagg ttactgtat atttgggtgt ggtccatgac	1080
catatatattt ttggaaaaaa actgcattgtt aaattccaaat ttggaaacgac atttctacag	1140
ccaatgttga aatactattt ttgtcctgtat taggagactt attatttcat ttcatatata	1200
gataggtccc ttgagaacta gaaagattaa attaaagatt gagatccaaat aatgcataatg	1260
aacacagaac atttgcatttt ttccaaagg ggaccatata tatataagat gtcattgtgc	1320
tttatgtatg gaagagagaa taacgatctc atatataatat ctcataatata tatataatcat	1380

ES 2 375 933 T3

ctaaaataag atgttttaaa ccatctggta ttccgtataac aatttacact aaaaagacca	1440
aacaggtggg aaggcacaaa acattagatc aaaaattaag gttaaagtatc tcagatata	1500
agaggaacaa tatactaatt ggaacaaatt aaagtatcct cacttacaat ggtcatat	1560
agaagctact taggtataac tctcactatc cctaattatt tgcacttt taaaatgc	1620
cacctattaa taaaacaatt attggcatag tgagtttacc attttacctt tttaattat	1680
aagcgaatga attaaaaact taagatatta aaaaaattct gccttaaca aagtaattat	1740
ttgagggtat aataggtaaa aagaaattgt cctttttta tttgtcaaaa tgaacaagta	1800
gttagggaca actaaaaaaag gaaaaatgga tgagtaatta ggaacggagg gatataaaa	1860
cactgtcatc actcaaaaaa tatgagtatc ttgacttgca caacataggt acttaatcaa	1920
agactcaata tacaaatctc taaagtaat ttgtatgtt atatacagtc tcttgaaag	1980
cccaatttgt ataaaatatt taaatgcagc tagatataca aacggaaatt agcatagcaa	2040
ctgaaaactat agatatagaa cataattagg caatgacttt gtttttgtt tgctgcctc	2100
acactttatt tgactgcctt cttgaatac tttgaatatt ctaagtaacgc cagctataag	2160
gtgaagaaag aattaaacta taatactctg tattgctctt cttccataat agtgaacaa	2220
gg atg aat ttc aac aat gaa ttg tct gat ctg aaa aat cgc ttc cta Met Asn Phe Asn Asn Glu Leu Ser Asp Leu Lys Asn Arg Phe Leu 1 5 10 15	2267
ttt agg acg ctg aga gcc cag aaa tgc tcg gat gtt gca aga gat cga Phe Arg Thr Leu Arg Ala Gln Lys Cys Ser Asp Val Ala Arg Asp Arg 20 25 30	2315
ata gat ttc ttt ata tgg gag tta aaa ttc ctt aat tgt ttt ctc cat Ile Asp Phe Phe Ile Trp Glu Leu Lys Phe Leu Asn Cys Phe Leu His 35 40 45	2363
ttg cag agc ttc gct ttt gca agt gaa tgt ggt atg cta gat atc tca Leu Gln Ser Phe Ala Phe Ala Ser Glu Cys Gly Met Leu Asp Ile Ser 50 55 60	2411
cag aaa atg ata gaa att tgc aag agg ttt aat aca cca cct cca cat Gln Lys Met Ile Glu Ile Cys Lys Arg Phe Asn Thr Pro Pro Pro His 65 70 75	2459
aat tca ttt gca tac tgg aag gag gta att tgc aag agg ctg tgc gct Asn Ser Phe Ala Tyr Trp Lys Glu Val Ile Cys Lys Arg Leu Cys Ala 80 85 90 95	2507
att agc atc cag ccg gat gct agt tca gat gat gga ttt gca tgc tgg Ile Ser Ile Gln Pro Asp Ala Ser Ser Asp Asp Gly Phe Ala Cys Trp 100 105 110	2555
aag aaa gta att tgg aag act aag caa gaa ttc aga gct aaa tac tcc Lys Lys Val Ile Trp Lys Thr Lys Gln Glu Phe Arg Ala Lys Tyr Ser 115 120 125	2603

ES 2 375 933 T3

ttt cca aaa aca cta ctt gca gac aac aag gta tat gat gat gat Phe Pro Lys Thr Leu Leu Ala Asp Asn Lys Val Tyr Asp Asp Asp 130 135 140	2651
act aat ccc aaa ttt gtg atg gaa ttc atc gat gct gtt gtg ggg aat Thr Asn Pro Lys Phe Val Met Glu Phe Ile Asp Ala Val Val Gly Asn 145 150 155	2699
ctc aat gtt cta gtc aag atc aat gat cca tct tca ttg ctt ttt gtt Leu Asn Val Leu Val Lys Ile Asn Asp Pro Ser Ser Leu Leu Phe Val 160 165 170 175	2747
cca gga ccc aag gaa caa ata gaa caa gtg tta aag gag ttg aag tta Pro Gly Pro Lys Glu Gln Ile Glu Gln Val Leu Lys Glu Leu Lys Leu 180 185 190	2795
ttg aga ttt ttt gtc tgc ttt gtt tca aac aaa tgt ata gag cct caa Leu Arg Phe Phe Val Cys Phe Val Ser Asn Lys Cys Ile Glu Pro Gln 195 200 205	2843
tac caa cat act act ttt tat act cac gct tta att gag gct agc cac Tyr Gln His Thr Thr Phe Tyr Thr His Ala Leu Ile Glu Ala Ser His 210 215 220	2891
atc gca atg gtt gtg tgg ttg aat ttg cca atc tat gga aac aga aat Ile Ala Met Val Val Trp Leu Asn Leu Pro Ile Tyr Gly Asn Arg Asn 225 230 235	2939
caa gac ttg gct tca agt gaa gtt agt tgt ttg ctt tct gat ttc atg Gln Asp Leu Ala Ser Ser Glu Val Ser Cys Leu Leu Ser Asp Phe Met 240 245 250 255	2987
gaa atg aag att aag tcc att cag cca gac atc agc cgc aac aat att Glu Met Lys Ile Lys Ser Ile Gln Pro Asp Ile Ser Arg Asn Asn Ile 260 265 270	3035
tat att gat gtc ttg agg gcg ttg aag tca acc ata cca caa gct caa Tyr Ile Asp Val Leu Arg Ala Leu Lys Ser Thr Ile Pro Gln Ala Gln 275 280 285	3083
gat aag cat gct gag agt ggc att gtg gag act cca aca cac aat Asp Lys His Ala Ala Glu Ser Gly Ile Val Glu Thr Pro Thr His Asn 290 295 300	3131
ctg atg gtt ggt ttg agt gat caa atg gcc aac ctt cag gag atg ctc Leu Met Val Gly Leu Ser Asp Gln Met Ala Asn Leu Gln Glu Met Leu 305 310 315	3179
tgc ctt cta aga gac aat ctc att cat ctg cca ata cta gat ctg gaa Cys Leu Leu Arg Asp Asn Leu Ile His Leu Pro Ile Leu Asp Leu Glu 320 325 330 335	3227
ttt cat ctt caa gat atg gat tct gtt att gtt gat gcc gga ctt ctt Phe His Leu Gln Asp Met Asp Ser Val Ile Val Asp Ala Gly Leu Leu 340 345 350	3275
att tac tca tta tat gat atc aag ggg cag aag gaa gac aca aca ttg Ile Tyr Ser Leu Tyr Asp Ile Lys Gly Gln Lys Glu Asp Thr Thr Leu 355 360 365	3323

ES 2 375 933 T3

gag gat atc aac cag gca ctt ggt ttt gat ctt ccc aga aac att gag Glu Asp Ile Asn Gln Ala Leu Gly Phe Asp Leu Pro Arg Asn Ile Glu 370 375 380	3371
cct atc aag gca atg atc aac ctt gtc atg caa aag gca ttt caa tgt Pro Ile Lys Ala Met Ile Asn Leu Val Met Gln Lys Ala Phe Gln Cys 385 390 395	3419
aac ttg cca agg att cat gga cta ggt tat gtc gat ttt cta ttg aaa Asn Leu Pro Arg Ile His Gly Leu Gly Tyr Val Asp Phe Leu Leu Lys 400 405 410 415	3467
aac ctg aag gat ttc caa ggc cgt tat tca gat tca ctc gat ttc ctc Asn Leu Lys Asp Phe Gln Gly Arg Tyr Ser Asp Ser Leu Asp Phe Leu 420 425 430	3515
aag aat caa ctt caa gtt att caa act gaa ttt gag agc ttg caa cct Lys Asn Gln Leu Gln Val Ile Gln Thr Glu Phe Glu Ser Leu Gln Pro 435 440 445	3563
ttc ttg aag gtt gtc gta gaa gag cca cac aat aag ctc aag aca ctg Phe Leu Lys Val Val Glu Glu Pro His Asn Lys Leu Lys Thr Leu 450 455 460	3611
aat gaa gat tgt gct aca cag ata att agg aaa gca tat gag gtg gaa Asn Glu Asp Cys Ala Thr Gln Ile Ile Arg Lys Ala Tyr Glu Val Glu 465 470 475	3659
tat gta gtt gat gct tgt ata aac aaa gag gtt cct cag tgg tgc atc Tyr Val Val Asp Ala Cys Ile Asn Lys Glu Val Pro Gln Trp Cys Ile 480 485 490 495	3707
gag cgt tgg ctc ctg gat atc ata gag gag att act tgt atc aaa gca Glu Arg Trp Leu Leu Asp Ile Ile Glu Glu Ile Thr Cys Ile Lys Ala 500 505 510	3755
aag att cag gaa aag aac acg gtt gag gat aca atg aag act gtc att Lys Ile Gln Glu Lys Asn Thr Val Glu Asp Thr Met Lys Thr Val Ile 515 520 525	3803
gct cgt aca tca tca aaa ctg gca agg act cca agg atg aat gaa gag Ala Arg Thr Ser Ser Lys Leu Ala Arg Thr Pro Arg Met Asn Glu Glu 530 535 540	3851
att gtt ggg ttt gag gat gtc ata gaa aat tta aga aaa aaa cta ctg Ile Val Gly Phe Glu Asp Val Ile Glu Asn Leu Arg Lys Lys Leu Leu 545 550 555	3899
aat gga acc aaa ggg caa gat gtc att tca att cac ggc atg cca ggt Asn Gly Thr Lys Gly Gln Asp Val Ile Ser Ile His Gly Met Pro Gly 560 565 570 575	3947
tta ggt aag acg act tta gcc aac agt ctc tat tct gac agg tca gtt Leu Gly Lys Thr Thr Leu Ala Asn Ser Leu Tyr Ser Asp Arg Ser Val 580 585 590	3995
ttt tct caa ttt gat att tgt gca caa tgt tgt gtg tct caa gta tat Phe Ser Gln Phe Asp Ile Cys Ala Gln Cys Cys Val Ser Gln Val Tyr 595 600 605	4043

ES 2 375 933 T3

tct tat aag gac tta ata ttg gcc ttg cta cgt gat gct att ggt gag Ser Tyr Lys Asp Leu Ile Leu Ala Leu Leu Arg Asp Ala Ile Gly Glu 610 615 620	4091
ggt tct gtg cgt aga gaa ctt cat gcc aat gaa tta gct gat atg ctt Gly Ser Val Arg Arg Glu Leu His Ala Asn Glu Leu Ala Asp Met Leu 625 630 635	4139
cgc aaa act cta ttg ccc cga agg tac ctt atc ctt gtt gat gac gtg Arg Lys Thr Leu Leu Pro Arg Arg Tyr Leu Ile Leu Val Asp Asp Val 640 645 650 655	4187
tgg gaa aat agt gtt tgg gat gat tta aga ggt ttt cca gat gtc Trp Glu Asn Ser Val Trp Asp Asp Leu Arg Gly Cys Phe Pro Asp Val 660 665 670	4235
aat aac aga agc aga atc att cta aca aca aga cat cat gaa gtt gcc Asn Asn Arg Ser Arg Ile Ile Leu Thr Thr Arg His His Glu Val Ala 675 680 685	4283
aaa tat gct agt gtt cat agt gat ccc ctt cat ctt cgt atg ttt gac Lys Tyr Ala Ser Val His Ser Asp Pro Leu His Leu Arg Met Phe Asp 690 695 700	4331
gaa gtt gaa agt tgg aag ttg ctt gaa aag aaa gtg ttt ggt gaa gaa Glu Val Glu Ser Trp Lys Leu Leu Glu Lys Lys Val Phe Gly Glu Glu 705 710 715	4379
agc tgt tcc cct ctc cta aaa aat gtt ggg cta aga ata gca aaa atg Ser Cys Ser Pro Leu Leu Lys Asn Val Gly Leu Arg Ile Ala Lys Met 720 725 730 735	4427
tgt gga caa cta cct ctt tca att gtt ctg gtg gct ggt att ctg tca Cys Gly Gln Leu Pro Leu Ser Ile Val Leu Val Ala Gly Ile Leu Ser 740 745 750	4475
gag atg gaa aag gaa gta gaa tgt tgg gaa caa gtg gcc aac aat ttg Glu Met Glu Lys Val Glu Cys Trp Glu Gln Val Ala Asn Asn Leu 755 760 765	4523
ggt tcc tac att cac aat gac tca aga gcc att gta gac aaa agt tat Gly Ser Tyr Ile His Asn Asp Ser Arg Ala Ile Val Asp Lys Ser Tyr 770 775 780	4571
cat gtt tta cct tgt cat ctt aag tct tgc ttc ctt tat ttt gga gca His Val Leu Pro Cys His Leu Lys Ser Cys Phe Leu Tyr Phe Gly Ala 785 790 795	4619
ttt tta gaa gat aga gtg att gac att tca agg tta ata agg cta tgg Phe Leu Glu Asp Arg Val Ile Asp Ile Ser Arg Leu Ile Arg Leu Trp 800 805 810 815	4667
ata tca gaa gca ttt ata aaa agt agt gaa ggc agg agg ttg gag gat Ile Ser Glu Ala Phe Ile Lys Ser Ser Glu Gly Arg Arg Leu Glu Asp 820 825 830	4715
ata gca gaa ggt tac ttg gag aat ctt att gga aga aat cta gta atg Ile Ala Glu Gly Tyr Leu Glu Asn Leu Ile Gly Arg Asn Leu Val Met 835 840 845	4763

ES 2 375 933 T3

gtt act cag agg tcc att tca gat ggt aag gcg aaa gaa tgt cgc ctt	4811
Val Thr Gln Arg Ser Ile Ser Asp Gly Lys Ala Lys Glu Cys Arg Leu	
850 855 860	
cat gat gta tta ctc gac ttc tgc aag gaa aga gca gct gag gag aat	4859
His Asp Val Leu Leu Asp Phe Cys Lys Glu Arg Ala Ala Glu Glu Asn	
865 870 875	
ttt cta cta tgg ata aat aggtaatatg ataagtaact gtactttcaa	4907
Phe Leu Leu Trp Ile Asn	
880 885	
tcaatcaagt atttcaagtt atatctgaaa attaatgata tgatttgct aattgatata	4967
ttc agg gat cag att acc aaa cct tct tcc tgt gtt tac tct cac aag	5015
Arg Asp Gln Ile Thr Lys Pro Ser Ser Cys Val Tyr Ser His Lys	
890 895 900	
cag cat gct cac ttg gcc ttc act gaa atg cat aat ctt gta gaa tgg	5063
Gln His Ala His Leu Ala Phe Thr Glu Met His Asn Leu Val Glu Trp	
905 910 915	
agt gcg tct tgc tca ttt gtt ggc tgc gta gta ctt tcc aat aaa tat	5111
Ser Ala Ser Cys Ser Phe Val Gly Ser Val Val Leu Ser Asn Lys Tyr	
920 925 930	
gac tca tac ttt tcc act cgt gac ata tcc tca cta cat gat ttt tca	5159
Asp Ser Tyr Phe Ser Thr Arg Asp Ile Ser Ser Leu His Asp Phe Ser	
935 940 945	
att tca cgc att tta cca aat ttc aag ttt cta aaa gtg tta gat ttg	5207
Ile Ser Arg Ile Leu Pro Asn Phe Lys Phe Leu Lys Val Leu Asp Leu	
950 955 960	
gaa cac cgg gtt ttt att gat ttt att cca act gag ctt gtt tac ttg	5255
Glu His Arg Val Phe Ile Asp Phe Ile Pro Thr Glu Leu Val Tyr Leu	
965 970 975 980	
aag tat ttt tct gca cac att gaa cag aat tca att cct tca agc ata	5303
Lys Tyr Phe Ser Ala His Ile Glu Gln Asn Ser Ile Pro Ser Ser Ile	
985 990 995	
tcc aat ctt tgg aac ctt gaa act ctt ata tta aaa agt cca ata	5348
Ser Asn Leu Trp Asn Leu Glu Thr Leu Ile Leu Lys Ser Pro Ile	
1000 1005 1010	
tat gcg tta cgt tgc acg cta cta cta cct agt aca gtt tgg gat	5393
Tyr Ala Leu Arg Cys Thr Leu Leu Leu Pro Ser Thr Val Trp Asp	
1015 1020 1025	
atg gtt aaa ttg aga cat ctg tat att cct gac ttc agc aca agg	5438
Met Val Lys Leu Arg His Leu Tyr Ile Pro Asp Phe Ser Thr Arg	
1030 1035 1040	
att gaa gca gca tta ctt gag aac tct gca aaa ctt tat aat ttg	5483
Ile Glu Ala Ala Leu Leu Glu Asn Ser Ala Lys Leu Tyr Asn Leu	
1045 1050 1055	
gaa acc ctt tcc act cta tat ttc tct cgt gtt gag gat gca gaa	5528
Glu Thr Leu Ser Thr Leu Tyr Phe Ser Arg Val Glu Asp Ala Glu	

ES 2 375 933 T3

1060	1065	1070	
ttg atg ctg aga	aaa aca cct aat ctt	cga aaa ctg ata tgt	5573
Leu Met Leu Arg	Lys Thr Pro Asn Leu	Arg Lys Leu Ile Cys	
1075	1080	1085	
gtt gaa tgt tta	gaa tac ccc cct cag	tac cat gtg ttg aat	5618
Val Glu Cys Leu	Glu Tyr Pro Pro Gln	Tyr His Val Leu Asn	
1090	1095	1100	
cca ata cgg ctt	gaa ata cta aag ctt	tat cga tca aaa ttt	5663
Pro Ile Arg Leu	Glu Ile Leu Lys Leu	Tyr Arg Ser Lys Phe	
1105	1110	1115	
acc atc ccc ttt	tgc atc tct gca cca	aat ctc aaa tac ttg	5708
Thr Ile Pro Phe	Cys Ile Ser Ala Pro	Asn Leu Lys Tyr Leu	
1120	1125	1130	
ctc tgt ggc ttt	tcc ctg gat tct cag	tac tta tca gaa act	5753
Leu Cys Gly Phe	Ser Leu Asp Ser Gln	Tyr Leu Ser Glu Thr	
1135	1140	1145	
gat cat ctc aag	cac ctt gag gta ctc	ata ctg tac aag gtt	5798
Asp His Leu Lys	His Leu Glu Val Leu	Ile Leu Tyr Lys Val	
1150	1155	1160	
ttt ggt gat cat	agg gaa tgg aaa gtg	agc aat ggc aag ttc	5843
Phe Gly Asp His	Arg Glu Trp Lys Val	Ser Asn Gly Lys Phe	
1165	1170	1175	
caa ctc aaa atc	ttg aaa cta gaa tat	ttg tcc ttg gtg aaa	5888
Gln Leu Lys Ile	Leu Lys Leu Glu Tyr	Leu Ser Leu Val Lys	
1180	1185	1190	
att gta gct gat	gat gcc ttt cct aac	ctt gaa caa ttg gtt	5933
Ile Val Ala Asp	Asp Ala Phe Pro Asn	Leu Glu Gln Leu Val	
1195	1200	1205	
cgt gga tgt caa	gat ctt atg gag atc	cct tct tgt ttc atg	5978
Arg Gly Cys Gln	Asp Leu Met Glu Ile	Pro Ser Cys Phe Met	
1210	1215	1220	
atc ctt tct ctc	aag tac atc ggg gta	gaa tac tgc aat gag	6023
Ile Leu Ser Leu	Lys Tyr Ile Gly Val	Glu Tyr Cys Asn Glu	
1225	1230	1235	
gtt gtc aag tca	gcc ttg aat ata caa	gaa aca caa gtc gaa	6068
Val Val Lys Ser	Ala Leu Asn Ile Gln	Glu Thr Gln Val Glu	
1240	1245	1250	
tat caa aat act	aat ttc aag ctc gtt	ctc atc g aggtacacta	6112
Tyr Gln Asn Thr	Asn Phe Lys Leu Val	Leu Ile	
1255	1260		
ctgaaaaaaag	ctttatctg	atgattttg	6172
ctgttttctc	agttatctt	acctcggtgc	6229
ag ttt tct ttg	cag aaa aag gcg	tgg aaa tta aat tta	6273
Glu Phe Ser	Leu Gln Lys Lys Ala	Trp Lys Leu Asn Leu	
		Thr Asp	

ES 2 375 933 T3

1265	1270	1275	
gac gaa gat . atg cac aat gca gta	aaa aat att ctt gca	gaa ata	6318
Ala Glu Asp Met His Asn Ala Val	Lys Asn Ile Leu Ala	Glu Ile	
1280	1285	1290	
aga tagtactac tttttttttt ttctttcctt tttttaaata caccaaatag			6371
Arg			
atagattcat cttttttgtc ttttcgatat gaaaggata gaatcagttt catctgatga			6431
gaaagagaag aaacttactg tgaccggaga tggatgtct gatgaagttc aatttagttgt			6491
ggagaaaactg agaaagcgtg gcatgccagg gttgttagtcc caacttgtca acacaaaatgt			6551
gctatactca ttttgcttac tgtaataccca tttcatgaca cacacacacaa aacattaact			6611
gtagtaaaagt tttgatggat cagtaaatct gagttcaacc cattgtaatc cgttcaaatt			6671
caactcaaaa aattcccatt gagttattct ttaacagggt atccagagtt ttagctgga			6731
gcaatttggaa atatcacatg taatttcttt atgagttat tcgaaaaata aaagattctg			6791
taaaacgtqc aacggctgtt gcattcattg taaactaaat atatctcagt atgtaactat			6851
tgaacaaatt tttcattttt gtcctgagg tttgatgtaa gtcatttagat tttacggatc			6911
ctgaagtgaa tggttttagc cttttctatt ttcttatgag ttcaccaaaa tggatgtatc			6971
ccactctgct acatgttaga gaaatgagaa tggtagcacc cgagagtatg gcctagcggt			7031
caatcaatga agcaggtgaa aacaacaaaa gcaaaaaata ctaagagatt tcttcacatc			7091
tatctaagta ccgctaagca aagatactgt aatgaccctc ctggcattt atgtgtcttg			7151
ccttcgtgt gtgtttttaga gtgttccat agcgacccca agtcatttat gacttgcgtt			7211
gactaacggt tcggtcacat ggtcggtcgat ttggttttgg tgcgagttt tggatgtatc			7271
agcttatgaa tcttgaacga tgatttcga tcaaaaaattc aagaagatga catcgaaatc			7331
catttctaac gattccatca gctccggaaag ggtcattttt ggtttttttt tggatgtatc			7391
tgtactcccg tgcgatttagg ctttttaact ttaagttaa gcctaagttt gactttggc			7451
aacattctga gttaaacgcgc tcggatgaga attccgtcag tgcggatgtc tccggaaatgt			7511
caagtttgggt ttagattgac cttttttttt ggtttttttt tttttttttt tttttttttt			7571
tcttttgcgtt gttttgactt aaaatggcat ttgggtgtgg aatccacttt ttgtcaagat			7631
gacccgtat agaaattttg gctgtccat tgagtccgaa atatcgaatt tgatatgatt			7691
gcataatctcg ttgtgtgcg aaaaaatccg aacgagttcg gagaaccccg tgcagttt			7751
taaattttgg ggtaagtgcg aaaaaatctg cactttggaa aacccaaaaa aacccatcc			7811
ctctctcatc cctctctcat ctctcaatca tttaggcgtat tgcgatgtg gaaactttgt			7871

atttttgatg	gcttcgctgt	agagtattca	gaagctgttgcgt	ggtgccgtat	gttcgaggta	7931
aatttcgtaa	aatacctgtct	gccacatcc	ctttttgtgg	cttgattttg	gaatttttga	7991
gatttgtttt	cttagccattt	tttggtccga	tttcagtgtat	tcttgaggct	atcttgagtg	8051
gtttttcgag	gagagcatcg	tgggtttgtc	ggaattttact	gtagaccacc	catttttggt	8111
aaatatgctg	aattaacttc	tgtcccattt	ctttagtttt	tgagaaaaat	ttgggttttg	8171
gttgcattgtt	ggttatgtgt	tgttttgc	ccccgaatgg	tgtcccatca	tggaaacacaa	8231
tttggggagc	tgttaagaac	ctatttttgg	ggtaattcc	ggagtttcca	gcgcgggtcc	8291
cacttctccc	gttttgaccc	cgaaatttgc	atgtctccgt	ttcttgcgat	tttagtgtct	8351
aaacgaccgt	aataacatttgc	tgactctatt	tttgatagcg	gggcagcg	tcgaggccgt	8411
tcgaaaaggg	aaagctccgg	agaagtgtt	tttggagcgt	gcgtgatctg	cccacaagta	8471
gggtatggtt	tccctctt	agatttgagct	tgagagtgtg	aatgcattgtt	gattagttgg	8531
gatttgggtt	ggttagttt	gaatcatgca	taggtgttta	gaaatcatgt	tttggccttt	8591
tcgggaattt	tcgggtaact	gtgagcatgc	tatgtgttac	taattgaccc	tccttgctat	8651
gtggagtgtct	tgaatgcttgc	attactat	atctgaagca	tgttggcct	tagtttaggt	8711
ttgactaggg	cttgccttag	agatgcatt	ttcgggat	ataggcctt	gtttttgccc	8771
cgacgtcgct	cggtcgactt	agatccatgt	agactgggt	agcaacttga	gtctgatagt	8831
ttggccttta	gctaggcgt	acgcttgctc	cgatgataat	tatcttcttc	ttctttttt	8891
cgttggttacg	gcttcacgag	ttacgttgc	gacattgatt	ctgcttcg	atttgaagtt	8951
gattttttat	tcgggttccaa	ggacttacat	tgattggcta	agtgtggacg	gcgttccacg	9011
gaaatttata	agcatggatc	gattgagacc	ctttcagcag	ctacattggc	acttatata	9071
agcatccgat	ttaagggtccg	gcctctatcg	cccaaataact	tatata	agac aaccgg	9131
aggccggcc	tctatcgccc	agatacttgc	atagagcatc	cggtagagg	tctggcctca	9191
gttacttgat	acttggatt	ggttacttgg	gtacttttgc	tgagcatccg	gttcgagg	9251
cggcctccgt	actgtcagat	tctactaatt	ggggttgaga	ttctgggtcg	atgtttccg	9311
ttcttgggtt	cttatatgca	attttcttta	gttata	gttac	tcatcggt	9371
tatggggatc	cgttttaggtt	tttatttaac	ttgcgcacga	gtgtac	tttgc	9431
ggggccca	gttttaggtt	taggtgtat	tagcttata	attacttgc	ttacacttgc	9491
gtgccttcca	tggtttactt	agattgtcat	tcttgac	ttttgtgtt	attccttctt	9551
ttcatattgc	ctttactttt	caagttcagt	cggcctataa	tgcata	acttgg	9611
ttttggta	ctatgtacgc	tctgcattgt	tttcgtat	caggtctgag	caccagtggc	9671

ES 2 375 933 T3

cagcgttgat ccagtttggaa gtagtctgat ccggagacgg gggtgagcac atggcgttt	9731
gtactatttc agtctccatc tgtgtatata gacttgcctt ttaccttgc agacagtcca	9791
atctctgtgg tccacttttg ggacttgcac tcgtttgtt agtagctcg tactgggtac	9851
ttcctggttc taggagggat ctttatgtt atatatgtt tggttcgctt ccgcctgttt	9911
atattgttat cataaaattt tgccctactct tgtagtttc taccctcaga cccattactt	9971
gttattccgg gttacgggtt ggcttaccta ctgggtgggtt atagtatgtg ccaccatgac	10031
tcgagaaaatc gggtcgtgac agatacatga tgccctttt gttgaaagaa gcgggtactc	10091
agtcaaatgg tcgaggtgag ctcgacacca tcaataacat cccaaaaaag gaacaatgag	10151
aaagttacaa atcacaatac atgtccatat gctttggAAC taaagaattc aaagcacaca	10211
atgtatTTCA ataatctttt atctctgcct gcagttgaat ataccagata tcagatctga	10271
gacgatgttt aaaaaggaaa ctattattcg accctattcc tttctcaaac ctcgaaacca	10331
acaccagtta tacaacaata tatgcagaac cctttaacta tatactatat acaaatt	10388

<210> 2

<211> 1293

<212> PRT

5 <213> *Solanum tuberosum*

<400> 2

Met Asn Phe Asn Asn Glu Leu Ser Asp Leu Lys Asn Arg Phe Leu Phe			
1	5	10	15
10	15		

Arg Thr Leu Arg Ala Gln Lys Cys Ser Asp Val Ala Arg Asp Arg Ile			
20	25	30	
30			

Asp Phe Phe Ile Trp Glu Leu Lys Phe Leu Asn Cys Phe Leu His Leu			
35	40	45	
45			

Gln Ser Phe Ala Phe Ala Ser Glu Cys Gly Met Leu Asp Ile Ser Gln			
50	55	60	
60			

Lys Met Ile Glu Ile Cys Lys Arg Phe Asn Thr Pro Pro Pro His Asn			
65	70	75	80
75	80		

Ser Phe Ala Tyr Trp Lys Glu Val Ile Cys Lys Arg Leu Cys Ala Ile			
85	90	95	
95			

ES 2 375 933 T3

Ser Ile Gln Pro Asp Ala Ser Ser Asp Asp Gly Phe Ala Cys Trp Lys
100 105 110

Lys Val Ile Trp Lys Thr Lys Gln Glu Phe Arg Ala Lys Tyr Ser Phe
115 120 125

Pro Lys Thr Leu Leu Ala Asp Asn Lys Val Tyr Asp Asp Asp Asp Thr
130 135 140

Asn Pro Lys Phe Val Met Glu Phe Ile Asp Ala Val Val Gly Asn Leu
145 150 155 160

Asn Val Leu Val Lys Ile Asn Asp Pro Ser Ser Leu Leu Phe Val Pro
165 170 175

Gly Pro Lys Glu Gln Ile Glu Gln Val Leu Lys Glu Leu Lys Leu Leu
180 185 190

Arg Phe Phe Val Cys Phe Val Ser Asn Lys Cys Ile Glu Pro Gln Tyr
195 200 205

Gln His Thr Thr Phe Tyr Thr His Ala Leu Ile Glu Ala Ser His Ile
210 215 220

Ala Met Val Val Trp Leu Asn Leu Pro Ile Tyr Gly Asn Arg Asn Gln
225 230 235 240

Asp Leu Ala Ser Ser Glu Val Ser Cys Leu Leu Ser Asp Phe Met Glu
245 250 255

Met Lys Ile Lys Ser Ile Gln Pro Asp Ile Ser Arg Asn Asn Ile Tyr
260 265 270

Ile Asp Val Leu Arg Ala Leu Lys Ser Thr Ile Pro Gln Ala Gln Asp
275 280 285

Lys His Ala Ala Glu Ser Gly Ile Val Glu Thr Pro Thr His Asn Leu
290 295 300

Met Val Gly Leu Ser Asp Gln Met Ala Asn Leu Gln Glu Met Leu Cys
305 310 315 320

Leu Leu Arg Asp Asn Leu Ile His Leu Pro Ile Leu Asp Leu Glu Phe
325 330 335

ES 2 375 933 T3

His Leu Gln Asp Met Asp Ser Val Ile Val Asp Ala Gly Leu Leu Ile
340 345 350

Tyr Ser Leu Tyr Asp Ile Lys Gly Gln Lys Glu Asp Thr Thr Leu Glu
355 360 365

Asp Ile Asn Gln Ala Leu Gly Phe Asp Leu Pro Arg Asn Ile Glu Pro
370 375 380

Ile Lys Ala Met Ile Asn Leu Val Met Gln Lys Ala Phe Gln Cys Asn
385 390 395 400

Leu Pro Arg Ile His Gly Leu Gly Tyr Val Asp Phe Leu Leu Lys Asn
405 410 415

Leu Lys Asp Phe Gln Gly Arg Tyr Ser Asp Ser Leu Asp Phe Leu Lys
420 425 430

Asn Gln Leu Gln Val Ile Gln Thr Glu Phe Glu Ser Leu Gln Pro Phe
435 440 445

Leu Lys Val Val Val Glu Glu Pro His Asn Lys Leu Lys Thr Leu Asn
450 455 460

Glu Asp Cys Ala Thr Gln Ile Ile Arg Lys Ala Tyr Glu Val Glu Tyr
465 470 475 480

Val Val Asp Ala Cys Ile Asn Lys Glu Val Pro Gln Trp Cys Ile Glu
485 490 495

Arg Trp Leu Leu Asp Ile Ile Glu Glu Ile Thr Cys Ile Lys Ala Lys
500 505 510

Ile Gln Glu Lys Asn Thr Val Glu Asp Thr Met Lys Thr Val Ile Ala
515 520 525

Arg Thr Ser Ser Lys Leu Ala Arg Thr Pro Arg Met Asn Glu Glu Ile
530 535 540

Val Gly Phe Glu Asp Val Ile Glu Asn Leu Arg Lys Lys Leu Leu Asn
545 550 555 560

Gly Thr Lys Gly Gln Asp Val Ile Ser Ile His Gly Met Pro Gly Leu
565 570 575

ES 2 375 933 T3

Gly Lys Thr Thr Leu Ala Asn Ser Leu Tyr Ser Asp Arg Ser Val Phe
580 585 590

Ser Gln Phe Asp Ile Cys Ala Gln Cys Cys Val Ser Gln Val Tyr Ser
595 600 605

Tyr Lys Asp Leu Ile Leu Ala Leu Leu Arg Asp Ala Ile Gly Glu Gly
610 615 620

Ser Val Arg Arg Glu Leu His Ala Asn Glu Leu Ala Asp Met Leu Arg
625 630 635 640

Lys Thr Leu Leu Pro Arg Arg Tyr Leu Ile Leu Val Asp Asp Val Trp
645 650 655

Glu Asn Ser Val Trp Asp Asp Leu Arg Gly Cys Phe Pro Asp Val Asn
660 665 670

Asn Arg Ser Arg Ile Ile Leu Thr Thr Arg His His Glu Val Ala Lys
675 680 685

Tyr Ala Ser Val His Ser Asp Pro Leu His Leu Arg Met Phe Asp Glu
690 695 700

Val Glu Ser Trp Lys Leu Leu Glu Lys Lys Val Phe Gly Glu Glu Ser
705 710 715 720

Cys Ser Pro Leu Leu Lys Asn Val Gly Leu Arg Ile Ala Lys Met Cys
725 730 735

Gly Gln Leu Pro Leu Ser Ile Val Leu Val Ala Gly Ile Leu Ser Glu
740 745 750

Met Glu Lys Glu Val Glu Cys Trp Glu Gln Val Ala Asn Asn Leu Gly
755 760 765

Ser Tyr Ile His Asn Asp Ser Arg Ala Ile Val Asp Lys Ser Tyr His
770 775 780

Val Leu Pro Cys His Leu Lys Ser Cys Phe Leu Tyr Phe Gly Ala Phe
785 790 795 800

Leu Glu Asp Arg Val Ile Asp Ile Ser Arg Leu Ile Arg Leu Trp Ile
805 810 815

ES 2 375 933 T3

Ser Glu Ala Phe Ile Lys Ser Ser Glu Gly Arg Arg Leu Glu Asp Ile
820 825 830

Ala Glu Gly Tyr Leu Glu Asn Leu Ile Gly Arg Asn Leu Val Met Val
835 840 845

Thr Gln Arg Ser Ile Ser Asp Gly Lys Ala Lys Glu Cys Arg Leu His
850 855 860

Asp Val Leu Leu Asp Phe Cys Lys Glu Arg Ala Ala Glu Glu Asn Phe
865 870 875 880

Leu Leu Trp Ile Asn Arg Asp Gln Ile Thr Lys Pro Ser Ser Cys Val
885 890 895

Tyr Ser His Lys Gln His Ala His Leu Ala Phe Thr Glu Met His Asn
900 905 910

Leu Val Glu Trp Ser Ala Ser Cys Ser Phe Val Gly Ser Val Val Leu
915 920 925

Ser Asn Lys Tyr Asp Ser Tyr Phe Ser Thr Arg Asp Ile Ser Ser Leu
930 935 940

His Asp Phe Ser Ile Ser Arg Ile Leu Pro Asn Phe Lys Phe Leu Lys
945 950 955 960

Val Leu Asp Leu Glu His Arg Val Phe Ile Asp Phe Ile Pro Thr Glu
965 970 975

Leu Val Tyr Leu Lys Tyr Phe Ser Ala His Ile Glu Gln Asn Ser Ile
980 985 990

Pro Ser Ser Ile Ser Asn Leu Trp Asn Leu Glu Thr Leu Ile Leu Lys
995 1000 1005

Ser Pro Ile Tyr Ala Leu Arg Cys Thr Leu Leu Leu Pro Ser Thr
1010 1015 1020

Val Trp Asp Met Val Lys Leu Arg His Leu Tyr Ile Pro Asp Phe
1025 1030 1035

Ser Thr Arg Ile Glu Ala Ala Leu Leu Glu Asn Ser Ala Lys Leu
1040 1045 1050

ES 2 375 933 T3

Tyr Asn Leu Glu Thr Leu Ser Thr Leu Tyr Phe Ser Arg Val Glu
1055 1060 1065

Asp Ala Glu Leu Met Leu Arg Lys Thr Pro Asn Leu Arg Lys Leu
1070 1075 1080

Ile Cys Glu Val Glu Cys Leu Glu Tyr Pro Pro Gln Tyr His Val
1085 1090 1095

Leu Asn Phe Pro Ile Arg Leu Glu Ile Leu Lys Leu Tyr Arg Ser
1100 1105 1110

Lys Phe Lys Thr Ile Pro Phe Cys Ile Ser Ala Pro Asn Leu Lys
1115 1120 1125

Tyr Leu Lys Leu Cys Gly Phe Ser Leu Asp Ser Gln Tyr Leu Ser
1130 1135 1140

Glu Thr Ala Asp His Leu Lys His Leu Glu Val Leu Ile Leu Tyr
1145 1150 1155

Lys Val Glu Phe Gly Asp His Arg Glu Trp Lys Val Ser Asn Gly
1160 1165 1170

Lys Phe Pro Gln Leu Lys Ile Leu Lys Leu Glu Tyr Leu Ser Leu
1175 1180 1185

Val Lys Trp Ile Val Ala Asp Asp Ala Phe Pro Asn Leu Glu Gln
1190 1195 1200

Leu Val Leu Arg Gly Cys Gln Asp Leu Met Glu Ile Pro Ser Cys
1205 1210 1215

Phe Met Asp Ile Leu Ser Leu Lys Tyr Ile Gly Val Glu Tyr Cys
1220 1225 1230

Asn Glu Ser Val Val Lys Ser Ala Leu Asn Ile Gln Glu Thr Gln
1235 1240 1245

Val Glu Asp Tyr Gln Asn Thr Asn Phe Lys Leu Val Leu Ile Glu
1250 1255 1260

Phe Ser Leu Gln Lys Lys Ala Trp Lys Leu Asn Leu Thr Asp Ala
1265 1270 1275

	Glu Asp Met His Asn Ala Val Lys Asn Ile Leu Ala Glu Ile Arg	
	1280 1285 1290	
	<210> 3	
	<211> 26	
	<212> ADN	
5	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido	
10	<400> 3	
	aaacccggtg ttccaaatct aacact	26
	<210> 4	
	<211> 26	
	<212> ADN	
15	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido	
20	<400> 4	
	catgtagtga ggatatgtca cgagtg	26
	<210> 5	
	<211> 21	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido	
30	<400> 5	
	attacaatgg gttgaactca g	21
	<210> 6	
	<211> 22	
	<212> ADN	
35	<213> Secuencia artificial	

	<220>	
	<223> Oligonucleótido	
5	<400> 6	
	ac ct ttt ca at gtt ctgg tg	22
	<210> 7	
	<211> 22	
	<212> ADN	
10	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido	
15	<400> 7	
	ca tc qt gac at atc c tc ac ta	22
	<210> 8	
	<211> 18	
	<212> ADN	
20	<213> Secuencia artificial	
	<220>	
	<223> Oligonucleótido	
25	<400> 8	
	ca ac c ct g gc at gc c ac g	18
	<210> 9	
	<211> 4	
	<212> PRT	
30	<213> Secuencia artificial	
	<220>	
	<223> Dominio conservado	
35	<400> 9	

Gly Leu Pro Leu

1

<210> 10

<211> 4

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Dominio conservado

10 <400> 10

Gln Leu Pro Leu

1

<210> 11

<211> 4

<212> PRT

15 <213> Secuencia artificial

<220>

<223> Dominio conservado

20 <400> 11

Cys Lys Leu Tyr

1

<210> 12

<211> 4

<212> PRT

25 <213> Secuencia artificial

<220>

<223> Dominio conservado

30 <400> 12

Cys Phe L^{eu} Tyr
1

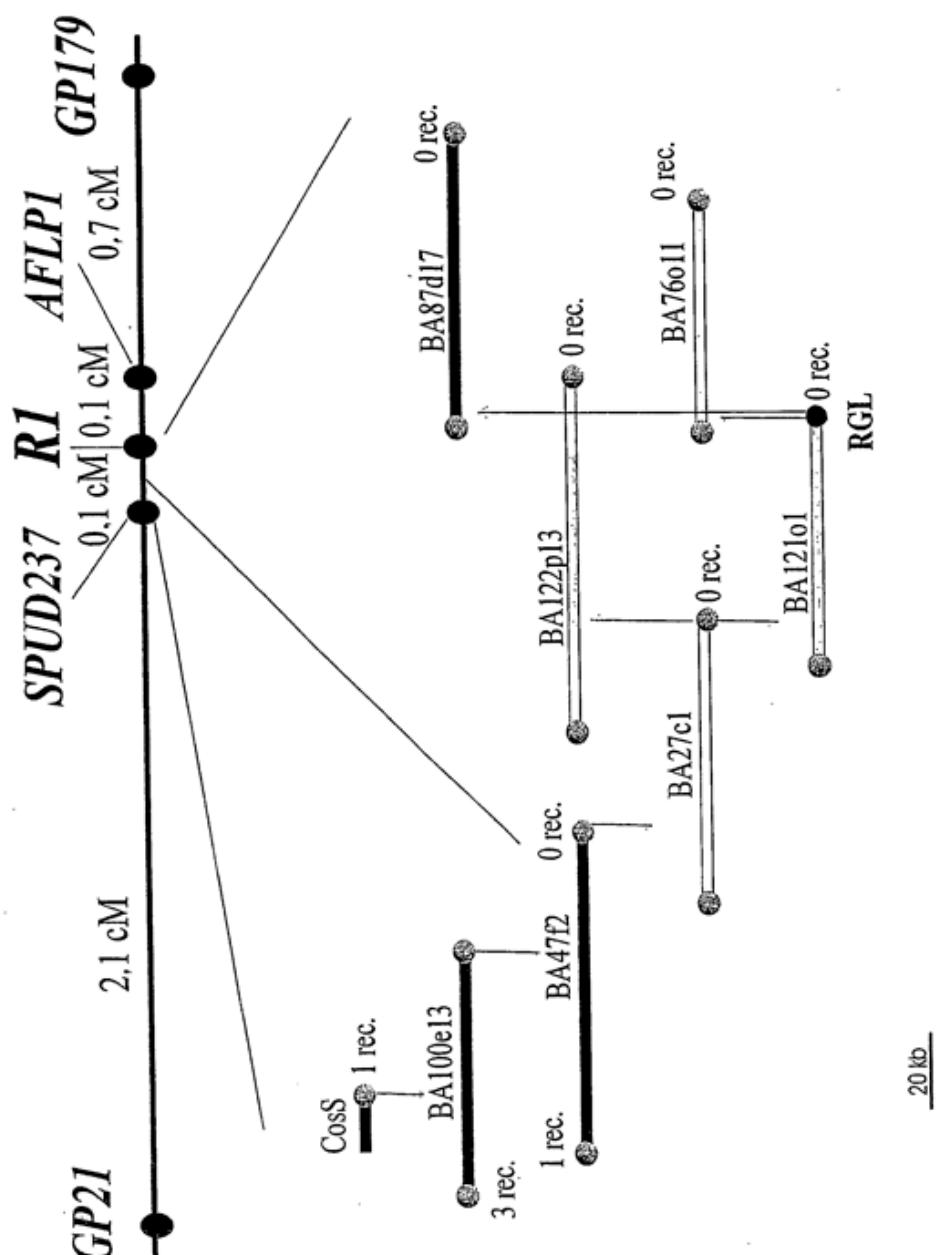


Figura 1

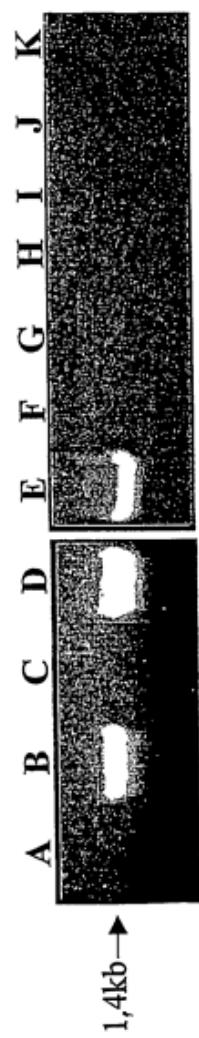


Figura 2

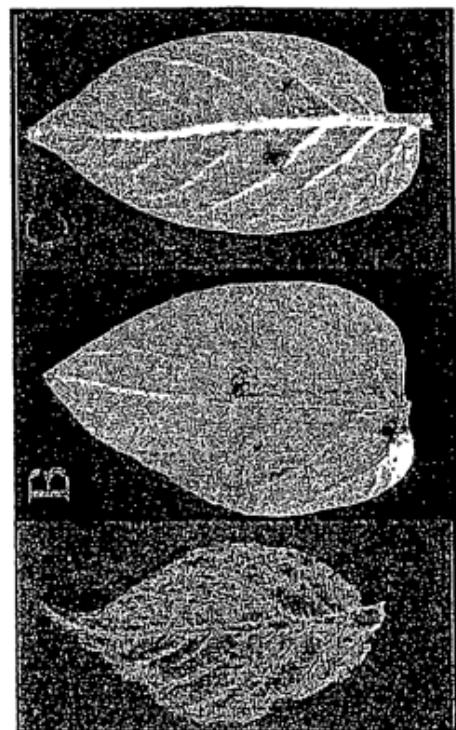


Figura 3



B

1 MNFINNELSDL KNRFLFRTLR AQKCSDVYARD RDUFFIWELK FLNCFLHLQS
 51 FAFASECGLM DISQKMCIEIC KRFNTPPPHN SFAYWKEVIC KRLCAISIQP
 101 DASSDDGFAC WKKVWTWKTKQ EFRAKYSPPK TLLADNKVYD DDDTNPKFVM
 151 EFIDAVVGNL NVLVKINDS SLLFVPGPKK QIEQVLKELK LRRFFVCFS
 201 NKCIIEPQYQH TTFYTHALIE ASHIAMVVWL NLPIYGNRNQ DLASSEVSCL
 251 LSDFMEMKIK SIOPDISRNN IYDVLRAKL STIPOAQDKH AAESGIVETP
 301 THNLMVGLSD OMANLOEMLC LLRDNLHLP LDLEFHLQD MDSVIVDAGL
 351 LIYSLYDIKG QKEDTILEDINQALGFJLPR NEPIKAMIN LYMQKAFCQCN
 401 LPRHGLGYV DFLLKNLQDF QGRYSDSLSDF LKNQLQVQT EFESLQPLK
 451 VVVEEPHENKL KTLNEDCATQ IRKA YEVY VVACINKEV P QWCCTERWLL
 501 DIBBETCK AKIQEKNTVB DTMKTKVART SSCLARTPRM NEELVGFEDV
 551 IENLRKLLN GIIGQDVISI H[GMPGLGK]T LANSLSYDRS VFSQFDICAQ
 601 CCVSQVSYK DJIALLRDA IGEGSVIRREL HANELADMRL KTLPPRLI
 651 LVDDYWWNSV WDDLRGCFPD VNVAKS[ILT]TRHEVAKYA SVHSDPLHLR
 701 MFDEYESWKL LEKKVFGES CSPLKLNVGL RIAKMCGLPLSIVLVAGL
 751 SEMKEVECW EQVANNLGSY IHNSRAIVD KSYHVLPCHL KSCPLYFGAF
 801 LEDRVIDSR LIRLWSEAF KISSEGRLLE DIAEGYLENL IGNLUVMTQ
 851 RSIISDGKAKE CRHDVLLDF CKERAAENF LLWINDQIT KFSSCVYSHK
 901 OHAAHAFTEM HNLVEWSASC SFYGSVYLSN KIDSYFSTRD ISSLHDFSI
 951 RILPNFKFLK VLDLEHRYFI DFIPTELVIL KTFSAHIEQN SIPSSSNLW
 1001 NLETLLKSP NYALRCILL PSTYWDWYKL RHLYTPDFST REAALLENS
 1051 AKLYNLETLS TLTTSRVEDA ELMRKTPW RLICEYECI EPPQHVLN
 1101 FPIRLELLKL YRSFKTIPF C3APNLKYL KLCGFSLSQ YLSETADHLK
 1151 HLEVLLYKV EFGDHREWVKV SNGKFPOLKI LKLETSLVK WIVADDAPN
 1201 LEQLVLRGCQ DLMEIPSFCM DILSLKTYGV EYCNESVTKS ALNQETQVE
 1251 DYQNTNEFKLV LIEFSLOKKA WKLNYDIAED MHNAVKNILA EIR*

Figura 4

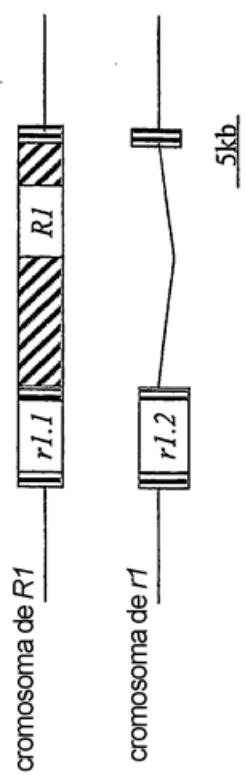


Figura 5