

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 574 981**

51 Int. Cl.:

**C12N 15/82** (2006.01)

**A01H 5/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.02.2008 E 08717232 (6)**

97 Fecha y número de publicación de la concesión europea: **13.04.2016 EP 2118286**

54 Título: **Plantas que tienen rasgos potenciados relacionados con el rendimiento y procedimiento de producción de las mismas**

30 Prioridad:

**28.02.2007 EP 07103271**

**15.03.2007 EP 07104197**

**21.03.2007 US 896050 P**

**02.04.2007 US 909510 P**

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

**23.06.2016**

73 Titular/es:

**CROPDESIGN N.V. (50.0%)**

**TECHNOLOPARK 3**

**9052 ZWIJNAARDE, BE y**

**CROP FUNCTIONAL GENOMICS CENTER (50.0%)**

72 Inventor/es:

**CHOI, YANG DO;**

**PARK, YOUN-IL;**

**CHUNG, SUCK WON;**

**HWANG, IN GYU y**

**OH, JONGHEE**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

**ES 2 574 981 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

Plantas que tienen rasgos potenciados relacionados con el rendimiento y procedimiento de producción de las mismas

5 La presente invención se refiere en general al campo de la biología molecular y atañe a un procedimiento para potenciar diversos rasgos económicamente importantes relacionados con el rendimiento en plantas. De forma más específica, la presente invención atañe a un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas, mediante la introducción y expresión en una planta un ácido nucleico que codifica un polipéptido Factor G Asociado a Harpin (en lo sucesivo en este documento denominado "HpaG"). La presente invención también atañe a plantas que tienen aumentada la expresión de un ácido nucleico que codifica un polipéptido HpaG, las cuales tienen el rendimiento de semilla y/o la biomasa aumentada con respecto a las plantas de control. La invención también proporciona construcciones que comprenden ácidos nucleicos que codifican a HpaG, útiles en la realización de los procedimientos de la invención.

15 El continuo aumento de la población mundial y la disminución de suministro de tierra cultivable disponible para la agricultura incentivan la investigación hacia el aumento de la eficiencia de la agricultura. Los medios convencionales para mejoras de cultivo y hortícolas utilizan técnicas de siembra selectiva para identificar plantas que tienen características deseables. Sin embargo, dichas técnicas de siembra selectiva tienen diversos inconvenientes, en concreto, que estas técnicas son típicamente laboriosas y producen plantas que a menudo contienen componentes genéticos heterogéneos que no siempre producen el rasgo deseable que se transmite desde las plantas progenitoras. Los avances en biología molecular han permitido a la humanidad modificar el plasma germinal de animales y plantas. La ingeniería genética de las plantas conlleva el aislamiento y la manipulación de material genético (típicamente en forma de ADN o ARN) y la introducción posterior de ese material genético en una planta. Dicha tecnología tiene la capacidad de producir cultivos o plantas que tienen diversos rasgos económicos, agronómicos u hortícolas mejorados.

25 Un rasgo de interés económico particular es el aumento del rendimiento. El rendimiento se define normalmente como el producto medible de valor económico de un cultivo. Esto puede definirse en términos de cantidad y/o calidad. El rendimiento depende directamente de diversos factores, por ejemplo, del número y tamaño de los órganos, de la arquitectura de la planta (por ejemplo, del número de ramas), de la producción de semillas, de la senescencia de las hojas y de más factores. El desarrollo de raíces, la absorción de nutrientes, la tolerancia al estrés, y el vigor temprano también pueden ser factores importantes en la determinación del rendimiento. Por lo tanto, optimizar los factores mencionados anteriormente puede contribuir a aumentar el rendimiento del cultivo.

30 El rendimiento de las semillas es un rasgo particularmente importante, ya que las semillas de muchas plantas son importantes para la nutrición humana y animal. Cultivos tales como maíz, arroz, trigo, canola y soja representan más de la mitad de la ingesta calórica humana total, bien a través del consumo directo de las propias semillas o a través del consumo de productos cárnicos generados sobre semillas procesadas. También son una fuente de azúcares, aceites y muchos tipos de metabolitos utilizados en los procesos industriales. Las semillas contienen un embrión (la fuente de nuevos brotes y raíces) y un endospermo (la fuente de nutrientes para el crecimiento del embrión durante la germinación y durante el crecimiento temprano de las plántulas). El desarrollo de una semilla implica muchos genes y requiere la transferencia de metabolitos desde las raíces, hojas y tallos al interior de la semilla en crecimiento. El endospermo, en particular, asimila los precursores metabólicos de los carbohidratos, aceites y proteínas y los sintetiza en macromoléculas de almacenamiento para llenar el grano.

45 El índice de cosecha, la proporción entre el rendimiento de semillas y el peso seco de la parte aérea, es relativamente estable en muchas condiciones ambientales y así, a menudo se puede obtener una correlación robusta entre el tamaño de la planta y el rendimiento de grano (por ejemplo, Rebetzke y col., (2002) Crop Science 42:739). Estos procesos están ligados de forma intrínseca, dado que la mayoría de la biomasa del grano depende de la productividad fotosintética actual o almacenada de las hojas y los tallos de la planta (Gardener y col., (1985) Physiology of Crop Plants. Iowa State University Press, pág. 68-73). Por lo tanto, se ha utilizado la selección del tamaño de la planta, incluso en etapas tempranas del desarrollo, como un indicador del futuro rendimiento potencial (por ejemplo, Tiftonell y col. 2005 Agric Ecosys & Environ 105:213). Cuando se prueba el impacto de las diferencias genéticas sobre la tolerancia al estrés, la capacidad de estandarizar las propiedades del suelo, la temperatura, la disponibilidad de agua y de nutrientes, y la intensidad de luz, es una ventaja intrínseca de los entornos de invernaderos y de las cámaras de crecimiento de plantas en comparación con el campo,. Sin embargo, las limitaciones artificiales sobre el rendimiento, debido a una pobre polinización debido a la ausencia de viento o de insectos, o al espacio insuficiente para las raíces maduras o el crecimiento de la cubierta, pueden restringir el uso de estos entornos controlados para probar diferencias en el rendimiento. Por lo tanto, las mediciones del tamaño de la planta en el desarrollo temprano en condiciones estandarizadas en una cámara de crecimiento o un invernadero, son prácticas estándar para proporcionar una indicación de las potenciales ventajas genéticas de rendimiento.

60 Otro rasgo de particular interés económico es el de los rasgos relacionados con el rendimiento de plantas crecidas en condiciones de estrés abiótico potenciados. El estrés abiótico es la causa principal de pérdidas de cultivo en todo el mundo, reduciendo en más de un 50 % los rendimientos promedio para la mayoría de las plantas de cultivo principales (Wang y col. Planta (2003) 218: 1-14). El estrés abiótico puede producirse por sequía, salinidad,

extremos de temperaturas, toxicidad química y estrés oxidativo. La capacidad de potenciar los rasgos relacionados con el rendimiento en plantas crecidas en condiciones de estrés abiótico sería de gran ventaja económica para los granjeros de todo el mundo y permitiría la cosecha de cultivos durante condiciones adversas y en territorios en donde la cosecha de cultivos no pueda ser posible de otra forma.

- 5 La capacidad para aumentar el rendimiento de la planta podría tener muchas aplicaciones en áreas tales como la agricultura, incluyendo en la producción de plantas ornamentales, la arboricultura, la horticultura y la silvicultura. El aumento del rendimiento también puede usarse en la producción de algas para su uso en biorreactores (para la producción biotecnológica de sustancias tales como productos farmacéuticos, anticuerpos o vacunas, o para la bioconversión de residuos orgánicos) y otras de tales áreas.

10 **Antecedentes**

El Sistema de Secreción de Tipo III (SSTT) es una maquinaria de exportación específica de bacterias Gram negativas y se encuentra entre patógenos vegetales y animales, pero también en el organismo endosimbiótico *Rhizobia*. Se postula que el SSTT entrega proteínas a la célula hospedadora a la cual la bacteria está asociada. En bacterias patógenas vegetales, el SSTT es un grupo de genes de respuesta de hipersensibilidad y de patogenicidad que comprende aproximadamente 20 genes, el grupo Hrp. Nueve de estos genes (los conservados harpin o *hrp*) están conservados en los patógenos vegetales y animales, ocho de ellos comparten homología con genes que codifican el aparato de los flagelos (Bogdanove y col., Mol. Microbiol., 20, 681-683, 1996), el noveno, *hrcC*, es homólogo a las secretinas de membrana externa GSP (Deng y Huang., J. Bacteriol 180, 4523-4531, 1999). Los genes *hpa* (*hrp associated*: asociados a *hrp*) contribuyen a la patogenicidad y la inducción de la respuesta de hipersensibilidad (RH) en plantas que no son hospedadoras, pero no son esenciales para las interacciones patógenas de las bacterias con las plantas. Se postula que el aparato de los flagelos y el SSTT evolucionaron a partir de un origen común (Gophna y col., Gene 312, 151-163, 2003); además el SSTT se ha propagado entre especies bacterianas evolutivamente distantes a través de múltiples acontecimientos de transferencia horizontal (Nguyen y col., J. Mol. Microbiol. Biotechnol. 2, 125-144, 2000).

25 Muchas bacterias gram negativas patógenas de plantas poseen dos conjuntos de genes que modulan sus interacciones con las plantas. Los genes de avirulencia determinan la especificidad de hospedador a base de las interacciones "gen para gen", y los genes *hrp* (*hypersensitive reaction and pathogenicity*: reacción de hipersensibilidad y de patogenicidad) están implicados en la patogenicidad y la inducción de las respuestas de hipersensibilidad (RH) en las plantas no hospedadoras. La RH es una muerte de la célula vegetal altamente localizada que se produce cuando se infiltran plantas no hospedadoras o variedades de cultivo resistentes de las plantas hospedadoras, con la planta patógena o las moléculas que desencadenan la RH, tales como las proteínas Avr y las harpin. Se piensa que la RH es una reacción de resistencia de las plantas a los patógenos microbianos.

Las Harpin es un grupo que desencadenantes de la RH que secreta la ruta de secreción de tipo III (SSTT) y desencadenan la RH cuando se infiltran en el apoplasto de las hojas de las plantas no hospedadoras. A diferencia de las proteínas Avr, las que deben entregarse dentro de la célula para ejercer sus funciones, las harpin pueden desencadenar la RH cuando se entregan al espacio intercelular de las células vegetales. Desde que se identificó a HrpN, la primera harpin, a partir de *Erwinia amylovora*, se han notificado muchas harpin de diversas especies, que incluyen *Pseudomonas*, *Ralstonia*, y *Xanthomonas*. Las harpin son proteínas ricas en glicina, estables en calor, que carecen de cisteínas, y se postula que están presentes en todas las bacterias patógenas vegetales que tienen un SSTT (Alfano y Colmer, Annu. Rev. Phytopathol. 42, 385-414, 2004). El mecanismo bioquímico para desencadenar la RH en plantas no hospedadoras mediante las harpin todavía no está claro. La HrpZ de *Pseudomonas syringae* pv. *syringae* se asocia con las paredes celulares en lugar de con las membranas de las células vegetales, y la proteína no desencadena respuesta a partir de protoplastos, que carecen de paredes (Hoyos y col. Mol. Plant-Microbe Interact. 9, 608-616, 1996). Sin embargo, la HrpZ de *P. syringae* pv. *phaseolicola* se une a las bicapas lipídicas y forma un poro de conducción de iones (Lee y col., Proc. Natl. Acad. Sci. USA 98, 289-294, 2001). Los 109 aminoácidos N terminales y los 216 aminoácidos C terminales de HrpZ son capaces de desencadenar la RH a un nivel similar al de HrpZ de longitud completa (Alfano y col., Mol. Microbiol., 19, 715-728, 1996). Kim y col. y Charkowski y col. demostraron que las harpin HrpW de *E. amylovora* y *P. syringae* pv. *tomato* están compuestas de dos dominios -el dominio N terminal de harpin y el dominio C terminal de la Pel (pectato liasa)- y propusieron que HrpW actúa en la pared celular (Charkowski y col., J. Bacteriol. 180, 5211-5217, 1998 ; Kim y Beer, J. Bacteriol 180, 5203-5210, 1998).

Además de las harpin, en las bacterias el grupo SSTT también puede incluir genes que codifican Factores asociados a Harpin. Los polipéptidos HpaG son más pequeños que las harpin, y comparten poca homología de secuencia. Se postula que estas diferencias en la secuencia con las harpin contribuye a la diferencia en la capacidad para desencadenar la RH en plantas entre los polipéptidos HpaG y las harpin (Kim y col., J. Bacteriol. 186, 6239-6247, 2004)

La solicitud de patente de Corea KR20030068302 desvela la proteína HpaG de *Xanthomonas*, la cual, cuando se emplea en plantas o semillas de plantas, confiere resistencia a la enfermedad, en particular resistencia a la infección por *Xanthomonas axonopodis*. Los Factores asociados a Harpin se han utilizado para conferir resistencia a la enfermedad en plantas y como resultado de esta resistencia al estrés biótico, las plantas tenían mejor rendimiento

en comparación con las plantas de control en condiciones de estrés biótico.

Sorprendentemente, ahora se ha encontrado que la introducción y expresión en una planta de un ácido nucleico que codifica un polipéptido Factor G asociado a Harpin (HpaG) proporciona plantas con rendimiento de semilla y/o biomasa aumentados en comparación con las plantas de control. Estos rasgos relacionados con el rendimiento potenciados se obtuvieron en plantas que no estaban expuestas a estrés biótico.

### **Definiciones**

#### Polipéptido(s)/proteína(s)

Los términos “polipéptido” y “proteína” se usan indistintamente en el presente documento y se refieren a aminoácidos en una forma polimérica de cualquier longitud.

#### Polinucleótido(s)/Ácido(s) nucleico(s)/Secuencia(s) de ácidos nucleicos/secuencia(s) de nucleótidos

Los términos “polinucleótido(s)”, “secuencia(s) de ácidos nucleicos”, “secuencia(s) de nucleótidos”, se usan indistintamente y se refieren a nucleótidos, ya sean ribonucleótidos o desoxirribonucleótidos o una combinación de ambos, en una forma polimérica de cualquier longitud.

#### Planta(s) de control

La elección de plantas de control adecuadas es una parte rutinaria de un montaje experimental y puede incluir plantas de tipo silvestre correspondientes o plantas correspondientes sin el gen de interés. La planta de control es típicamente de la misma especie de planta o incluso de la misma variedad de planta que la planta que se va a evaluar. La planta de control también puede ser un nulicigoto de la planta que se va a evaluar. Una “planta de control” como se usa en el presente documento no solo se refiere a plantas completas, sino también a partes de planta, incluyendo semillas y partes de semilla.

#### Homólogo(s)

Los “homólogos” de una proteína incluyen péptidos, oligopéptidos, polipéptidos, proteínas y enzimas que tienen sustituciones, deleciones y/o inserciones de aminoácidos con respecto a la proteína no modificada en cuestión y que tienen actividad funcional y biológica similar como la proteína no modificada de la cual proceden.

Una deleción se refiere a la eliminación de uno o más aminoácidos de una proteína.

Una inserción se refiere a uno o más restos de aminoácido que se introducen en un sitio predeterminado en una proteína. Las inserciones pueden comprender fusiones de N y/o C terminal, así como inserciones intrasecuencia de un solo aminoácido o múltiples aminoácidos. Generalmente, las inserciones en la secuencia de aminoácidos serán más pequeñas que las fusiones de N o C terminal, del orden de aproximadamente 1 a 10 restos. Los ejemplos de proteínas de fusión de N o C terminal o péptidos incluyen el dominio de unión o dominio de activación de un activador transcripcional como se usa en el sistema de doble híbrido de levadura, proteínas de cubierta de fago, etiqueta de (histidina)-6, etiqueta de glutatión S-transferasa, proteína A, proteína de unión a maltosa, dihidrofolato reductasa, epítipo Tag\*100, epítipo c-myc, epítipo FLAG®, lacZ, CMP (péptido de unión a calmodulina), epítipo HA, epítipo de proteína C, epítipo VSV.

Una sustitución se refiere a un reemplazo de aminoácidos de la proteína con otros aminoácidos que tienen propiedades semejantes (tales como hidrofobicidad, hidrofiliidad, antigenicidad, propensidad para formar o romper estructuras  $\alpha$  helicoidales o  $\beta$  laminares semejantes). Las sustituciones de aminoácidos son típicamente de restos únicos, pero pueden agruparse dependiendo de las limitaciones funcionales puestas al polipéptido y variar desde 1 a 10 aminoácidos; las inserciones normalmente serán del orden de aproximadamente 1 a 10 restos de aminoácido. Las sustituciones de aminoácidos son preferentemente sustituciones de aminoácidos conservativas. En la técnica se conocen bien las tablas de sustitución conservativa (véase, por ejemplo, Creighton (1984) Proteins. W. H. Freeman and Company y la Tabla 1 a continuación).

**Tabla 1:** Ejemplos de sustituciones de aminoácidos conservativas

| <b>Resto</b> | <b>Sustituciones conservativas</b> | <b>Resto</b> | <b>Sustituciones conservativas</b> |
|--------------|------------------------------------|--------------|------------------------------------|
| Ala          | Ser                                | Leu          | Ile; Val                           |
| Arg          | Lys                                | Lys          | Arg; Gln                           |
| Asn          | Gln; His                           | Met          | Leu; Ile                           |
| Asp          | Glu                                | Phe          | Met; Leu; Tyr                      |

(continuación)

| Resto | Sustituciones conservativas | Resto | Sustituciones conservativas |
|-------|-----------------------------|-------|-----------------------------|
| Gln   | Asn                         | Ser   | Thr; Gly                    |
| Cys   | Ser                         | Thr   | Ser; Val                    |
| Glu   | Asp                         | Trp   | Tyr                         |
| Gly   | Pro                         | Tyr   | Trp; Phe                    |
| His   | Asn; Gln                    | Val   | Ile; Leu                    |
| Ile   | Leu, Val                    |       |                             |

5 Las sustituciones, deleciones y/o inserciones de aminoácidos pueden realizarse fácilmente usando técnicas sintéticas de péptidos bien conocidas en la técnica, tales como síntesis peptídica de fase sólida y similares, o por manipulación de ADN recombinante. En la técnica se conocen bien procedimientos de manipulación de secuencias de ADN para producir variantes de sustitución, inserción o deleción de una proteína. Por ejemplo, los expertos en la técnica conocen técnicas para realizar mutaciones de sustitución en sitios predeterminados en el ADN e incluyen mutagénesis M13, mutagénesis *in vitro* T7-Gen (USB, Cleveland, OH), mutagénesis dirigida a sitio QuickChange (Stratagene, San Diego, CA), mutagénesis dirigida a sitio mediada por PCR u otros protocolos de mutagénesis dirigida a sitio.

#### Derivados

15 Los “derivados” incluyen péptidos, oligopéptidos, polipéptidos que, en comparación con la secuencia de aminoácidos de forma natural de la proteína tal como la presentada en la SEQ ID NO: 2, pueden comprender sustituciones de aminoácidos con restos de aminoácido de origen no natural, o adiciones de restos de aminoácido de origen no natural. Los “derivados” de una proteína también incluyen péptidos, oligopéptidos, polipéptidos, que comprenden restos de aminoácido modificados de origen natural (glucosilado, acilado, prenilado, fosforilado, miristoilado, sulfatado, etc.) o modificados de origen no natural en comparación con la secuencia de aminoácidos de forma no natural del polipéptido. Un derivado también puede comprender uno o más sustituyentes o adiciones que no son de aminoácido en comparación con la secuencia de aminoácidos de la que procede, por ejemplo una molécula indicadora u otro ligando, unido covalente o no covalentemente a la secuencia de aminoácidos, tal como una molécula indicadora que se une para facilitar su detección, y restos de aminoácido que no son de origen natural con respecto a la secuencia de aminoácidos de una proteína de origen natural.

#### Ortólogo(s)/Parólogo(s)

25 Los ortólogos y parálogos incluyen conceptos evolutivos usados para describir las relaciones antecesoras de los genes. Los parálogos son genes dentro de la misma especie que se han originado a través de la duplicación de un gen antecesor y los ortólogos son genes de diferentes organismos que se han originado a través de la especiación.

#### Dominio

30 El término “dominio” se refiere a un conjunto de aminoácidos conservados en posiciones específicas a lo largo de un alineamiento de secuencias de las proteínas evolutivamente relacionadas. Aunque los aminoácidos en otras posiciones pueden variar entre los homólogos, los aminoácidos que están muy conservados en posiciones específicas indican aminoácidos que son probablemente esenciales en la estructura, estabilidad o actividad de una proteína. Identificado por su alto grado de conservación en las secuencias alineadas de una familia de homólogos de proteína, éstas pueden usarse como identificadores para determinar si cualquier polipéptido en cuestión pertenece a una familia polipeptídica previamente identificada.

#### Motivo/Secuencia consenso/Firma

35 Las expresiones “motivo” o “secuencia consenso” o “firma” se refieren a una región conservada corta en la secuencia de las proteínas evolutivamente relacionadas. Los motivos son partes de dominios frecuentemente muy conservadas, pero también pueden incluir solo parte del dominio, o localizarse fuera del dominio conservado (si todos los aminoácidos del motivo se encuentran fuera de un dominio definido).

#### Hibridación

40 El término “hibridación”, como se define en el presente documento, es un proceso en el que secuencias de nucleótidos complementarias, sustancialmente homólogas, se hibridan entre sí. El proceso de hibridación puede producirse completamente en solución, es decir, ambos ácidos nucleicos complementarios están en solución. El proceso de hibridación también puede producirse con uno de los ácidos nucleicos complementarios inmovilizado en una matriz, tal como perlas magnéticas, perlas de Sefarosa o cualquier otra resina. El proceso de hibridación puede producirse adicionalmente con uno de los ácidos nucleicos complementario inmovilizado en un soporte sólido, tal

como una membrana de nitrocelulosa o de nailon, o inmovilizado, por ejemplo, mediante fotolitografía a, por ejemplo, un soporte vítreo silíceo (el último conocido como matrices o micromatrices de ácido nucleico o como microplacas de ácido nucleico). Para permitir que se produzca la hibridación, las moléculas de ácido nucleico están, generalmente, térmica o químicamente desnaturalizadas para fundir una doble cadena en dos cadenas únicas y/o retirar harpin u otras estructuras secundarias de ácidos nucleicos monocatenarios.

El término "rigurosidad" se refiere a las condiciones en las que se realiza la hibridación. La rigurosidad de la hibridación está influenciada por condiciones tales como temperatura, concentración salina, fuerza iónica y composición del tampón de hibridación. Generalmente, se seleccionan condiciones de baja rigurosidad por ser aproximadamente 30 °C menores que el punto de fusión térmico ( $T_f$ ) para la secuencia específica a una fuerza iónica definida y pH. Las condiciones de rigurosidad media son cuando la temperatura es 20 °C por debajo de la  $T_f$ , y las condiciones de rigurosidad alta son cuando la temperatura es 10 °C por debajo de la  $T_f$ . Las condiciones de hibridación de rigurosidad alta se usan típicamente para aislar secuencias de hibridación que tiene alta similitud de secuencia con la secuencia de ácidos nucleicos diana. Sin embargo, los ácidos nucleicos pueden desviarse en cuanto a la secuencia y aún codificar un polipéptido sustancialmente idéntico, debido a la generación del código genético. Por lo tanto, en ocasiones, pueden necesitarse condiciones de hibridación de rigurosidad media para identificar dichas moléculas de ácidos nucleicos.

La  $T_f$  es la temperatura bajo fuerza iónica definida y pH, en la que el 50 % de la secuencia diana se hibrida con una sonda perfectamente emparejada. La  $T_f$  depende de las condiciones de solución y de la composición base y longitud de la sonda. Por ejemplo, las secuencias más largas hibridan específicamente a temperaturas más altas. La tasa máxima de hibridación se obtiene de aproximadamente 16 °C hasta 32 °C por debajo de la  $T_f$ . La presencia de cationes monovalentes en la solución de hibridación reduce la repulsión electrostática entre las dos cadenas de ácido nucleico por lo cual se promueve la formación del híbrido; este efecto es visible para concentraciones de sodio de hasta 0,4 M (para mayores concentraciones, este efecto puede ignorarse). La formamida reduce la temperatura de fusión de dúplex de ADN-ADN y ADN-ARN con 0,6 a 0,7 °C para cada porcentaje de formamida, y la adición de formamida al 50 % permite que se realice la hibridación de 30 a 40 °C, aunque la tasa de hibridación disminuirá. Los emparejamientos erróneos de los pares de bases reducen la tasa de hibridación y la estabilidad térmica de los dúplex. Por término medio y para sondas grandes, la  $T_f$  disminuye aproximadamente 1 °C por % de emparejamiento erróneo de bases. La  $T_f$  puede calcularse usando las siguientes ecuaciones, dependiendo de los tipos de híbridos:

1) híbridos de ADN-ADN (Meinkoth y Wahl, Anal. Biochem., 138: 267-284, 1984):

$$T_f = 81,5 \text{ °C} + 16,6 \times \log_{10}[(\text{Na}^+)^a + 0,41 \times \%[\text{G/C}^b] - 500 \times [\text{L}^c] - 1 - 0,61 \times \% \text{ formamida}$$

2) híbridos de ADN-ARN o ARN-ARN:

$$T_f = 79,8 + 18,5 (\log_{10}[(\text{Na}^+)^a] + 0,58 (\% \text{G/C}^b) + 11,8 (\% \text{G/C}^b)^2 - 820/\text{L}^c$$

3) híbridos de oligo-ADN u oligo-ARN<sup>d</sup>:

$$\text{Para } <20 \text{ nucleótidos: } T_f = 2 (I_n)$$

$$\text{Para } 20\text{-}35 \text{ nucleótidos: } T_f = 22 + 1,46 (I_n)$$

<sup>a</sup> o para otro catión monovalente, pero solo exacto en el intervalo de 0,01-0,4 M.

<sup>b</sup> solo exacto para % GC en el intervalo de 30 % a 75 %.

<sup>c</sup> L = longitud del dúplex en los pares de bases.

<sup>d</sup> Oligo, oligonucleótido;  $I_n$ , longitud eficaz del cebador =  $2 \times (\text{n.}^\circ \text{ de G/C}) + (\text{n.}^\circ \text{ de A/T})$ .

La unión no específica puede controlarse usando una cualquiera de las numerosas técnicas conocidas, tales como, por ejemplo, bloqueo de la membrana con soluciones que contienen la proteína, adiciones del ARN heterólogo, ADN y SDS al tampón de hibridación y tratamiento con ARNasa. Para sondas no homólogas, puede realizarse una serie de hibridaciones al variar uno de (i) disminuir progresivamente la temperatura de hibridación (por ejemplo de 68 °C a 42 °C) o (ii) disminuir progresivamente la concentración de formamida (por ejemplo del 50 % al 0 %). El experto en la técnica es consciente de que durante la hibridación pueden alterarse diversos parámetros y que se mantendrán o cambiarán las condiciones de rigurosidad.

Además de las condiciones de hibridación, la especificidad de hibridación también depende típicamente de la función de los lavados posthibridación. Para retirar el fondo resultante de la hibridación no específica, las muestras se lavan con soluciones salinas diluidas. Los factores críticos de dichos lavados incluyen la fuerza iónica y la temperatura de la solución de lavado final: a menor concentración salina y mayor temperatura de lavado, mayor rigurosidad de lavado. Las condiciones de lavado se realizan típicamente a, o por debajo de, la rigurosidad de hibridación. Una hibridación positiva da lugar a una señal que es por lo menos dos veces la del fondo. Generalmente, las condiciones de rigurosidad adecuadas para los ensayos de hibridación de ácidos nucleicos o los procedimientos de detección de amplificación de genes son como se expusieron anteriormente. También pueden seleccionarse condiciones más o menos rigurosas. El experto en la técnica es consciente de que, durante el lavado, pueden alterarse diversos parámetros y que mantendrán o cambiarán las condiciones de rigurosidad.

Por ejemplo, las condiciones de hibridación típicas de alta rigurosidad para híbridos de ADN mayores de 50 nucleótidos incluyen una hibridación a 65 °C en 1x SSC o a 42 °C en 1x SSC y formamida al 50 %, seguido de lavado a 65 °C en 0,3x SSC. Los ejemplos de condiciones de hibridación de rigurosidad media para híbridos de ADN más largos de 50 nucleótidos incluyen una hibridación a 50 °C en 4x SSC o a 40 °C en 6x SSC y formamida al 50 %, seguido de lavado a 50 °C en 2x SSC. La longitud del híbrido es la longitud prevista para la hibridación del ácido nucleico. Cuando los ácidos nucleicos de secuencia conocida se hibridan, la longitud del híbrido puede determinarse alineando las secuencias e identificando las regiones conservadas descritas en el presente documento. 1x SSC es NaCl 0,15 M y citrato sódico 15 mM, la solución de hibridación y las soluciones de lavado pueden incluir adicionalmente reactivo de Denhardt 5 x, SDS al 0,5-1,0 %, 100 µg/ml de ADN de esperma de salmón fragmentado, desnaturalizado 100 µg/ml, pirofosfato sódico al 0,5 %.

Con el fin de definir el nivel de rigurosidad, puede hacerse referencia a Sambrook y col. (2001) *Molecular Cloning: a laboratory manual*, 3ª Edición, Cold Spring Harbor Laboratory Press, CSH, Nueva York o a *Current Protocols in Molecular Biology*, John Wiley & Sons, N. Y. (1989 y actualizaciones anuales).

#### Combinación de genes/Evolución dirigida

La combinación de genes o evolución dirigida consiste en repeticiones de combinación de ADN seguido de exploración y/o selección apropiada para generar variantes de ácidos nucleicos o partes de las mismas que codifican proteínas que tienen una actividad biológica modificada (Castle y col., (2004) *Science* 304(5674): 1151-4; Patentes de Estados Unidos 5.811.238 y 6.395.547).

#### Elemento regulador/Secuencia de control/Promotor

Las expresiones “elemento regulador”, “secuencia de control” y “promotor” se usan todas indistintamente en el presente documento y se toman en un contexto amplio para referirse a una secuencia de ácidos nucleicos reguladora capaz de efectuar la expresión de las secuencias a las que se unen. El término “promotor” se refiere típicamente a una secuencia de control de ácido nucleico localizada cadena arriba del inicio transcripcional de un gen y que está implicado en el reconocimiento y unión de la ARN polimerasa y otras proteínas, dirigiendo así la transcripción de un ácido nucleico unido operativamente. En los términos anteriormente mencionados se incluyen las secuencias reguladoras transcripcionales procedentes de un gen genómico eucariota clásico (incluyendo la caja TATA que es necesaria para el inicio de la transcripción adecuado, con o sin secuencia de caja CCAAT) y elementos reguladores adicionales (es decir, secuencias activadoras cadena arriba, potenciadores y silenciadores) que alteran la expresión génica en respuesta al desarrollo y/o estímulo externo, o de una manera específica de tejido. También se incluye en el término una secuencia reguladora transcripcional de un gen procarionta clásico, en cuyo caso puede incluir una secuencia de caja -35 y/o secuencias reguladores transcripcionales de caja -10. La expresión “elemento regulador” también incluye una molécula de fusión sintética o un derivado que confiere, activa o potencia la expresión de una molécula de ácido nucleico en una célula, tejido u órgano.

Un “promotor de planta” comprende elementos reguladores, que median la expresión de un segmento de secuencia codificante en células de plantas. Por consiguiente, un promotor de planta no tiene que ser de origen vegetal, sino que puede originarse de virus o microorganismos, por ejemplo de virus que atacan a las células de planta. El “promotor de planta” también puede originarse de una célula de planta, por ejemplo, de la planta que se transforma con la secuencia de ácidos nucleicos a expresar en el proceso de la invención y descrito en el presente documento. Esto también se aplica a las otras señales reguladoras de “planta”, tales como terminadores de “planta”. Los promotores cadena arriba de la secuencias de nucleótidos útiles en los procedimientos de la presente invención pueden modificarse mediante una o más sustituciones, inserciones y/o deleciones de nucleótido sin interferir con la funcionalidad o actividad de cualquiera de los promotores, la fase de lectura abierta (ORF, *Open Reading Frame*) o la región reguladora 3' tal como terminadores u otras regiones reguladoras 3' que se localizan fuera de la ORF. Además es posible que la actividad de los promotores aumente por la modificación de su secuencia o que se reemplacen completamente mediante más promotores activos, incluso promotores de organismos heterólogos. Para la expresión en plantas, la molécula de ácido nucleico debe, como se ha descrito anteriormente, estar unida operativamente a, o comprender, un promotor adecuado que exprese el gen de manera adecuada, en tiempo y con el patrón de expresión espacial requerido.

#### Unido operativamente

La expresión “unido operativamente” como se usa en el presente documento se refiere a un enlace funcional entre la secuencia promotora y el gen de interés, de tal manera que la secuencia promotora puede iniciar la transcripción del gen de interés.

#### Promotor constitutivo

Un “promotor constitutivo” se refiere a un promotor que es transcripcionalmente activo durante la mayor parte, aunque no necesariamente en todas, las fases de crecimiento y desarrollo y en la mayor parte de las condiciones ambientales, en por lo menos una célula, tejido u órgano. La siguiente Tabla 2a proporciona ejemplos de promotores constitutivos.

**Tabla 2a:** Ejemplos de promotores constitutivos

| Fuente del gen            | Referencia   |
|---------------------------|--|
| Actina                    | McElroy y col, Plant Cell, 2: 163-171, 1990                    |
| HMGP                      | WO 2004/070039   |
| CAMV 35S                  | Odell y col, Nature, 313: 810-812, 1985                        |
| CaMV 19S                  | Nilsson y col., Physiol. Plant. 100:456-462, 1997              |
| GOS2                      | de Pater y col, Plant J Nov; 2(6):837-44, 1992, WO 2004/065596 |
| Ubiquitina                | Christensen y col, Plant Mol. Biol. 18: 675-689, 1992          |
| Ciclofilina de arroz      | Buchholz y col, Plant Mol Biol. 25(5): 837-43, 1994            |
| Histona H3 de maíz        | Lepetit y col, Mol. Gen. Genet. 231:276-285, 1992              |
| Histona H3 de alfalfa     | Wu y col. Plant Mol. Biol. 11:641-649, 1988                    |
| Actina n 2                | An y col, Plant J. 10(1); 107-121, 1996                        |
| 34S FMV                   | Sanger y col., Plant. Mol. Biol., 14, 1990: 433-443            |
| Subunidad pequeña Rubisco | US 4,962,028   |
| OCS                       | Leisner (1988) Proc Natl Acad Sci USA 85(5): 2553              |
| SAD1                      | Jain y col., Crop Science, 39 (6), 1999: 1696                  |
| SAD2                      | Jain y col., Crop Science, 39 (6), 1999: 1696                  |
| Nos                       | Shaw y col. (1984) Nucleic Acids Res. 12(20): 7831-7846        |
| V-ATPasa                  | WO 01/14572  |
| Súper promotor            | WO 95/14098  |
| Proteínas de la caja G    | WO 94/12015  |

Promotor ubicuo

Un promotor ubicuo es activo en sustancialmente todos los tejidos o células de un organismo.

5 Promotor regulado evolutivamente

Un promotor regulado evolutivamente es activo durante determinadas etapas del desarrollo o en partes de la planta que experimentan cambios evolutivos.

Promotor inducible

10 Un promotor inducible ha inducido o aumentado el inicio de la transcripción en respuesta a un estímulo químico (para una revisión véase Gatz 1997, Annu. Rev. Plant Physiol. Plant Mol. Biol., 48: 89-108), ambiental o físico, o puede ser “inducible por estrés”, es decir, activado cuando una planta se expone a diversas condiciones de estrés, o un “patógeno inducible,” es decir, activado cuando una planta está expuesta a la exposición a diversos patógenos.

Promotor específico de órgano/específico de tejido

15 Un promotor específico de órgano o específico de tejido es uno que es capaz de iniciar preferentemente la transcripción en determinados órganos o tejidos, tales como hojas, raíces tejido de semillas, etc. Por ejemplo, un “promotor específico de raíz” es un promotor que es transcripcionalmente activo predominantemente en las raíces las de plantas, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta. En el presente documento a los promotores capaces de iniciar la transcripción solo en determinadas células se les denomina promotores “específicos de célula”.

20 En la siguiente Tabla 2b se indican ejemplos de promotores específicos de raíz:

**Tabla 2b:** Ejemplos de promotores específicos de raíz

| Fuente del gen                       | Referencia   |
|--------------------------------------|--|
| RCc3                                 | Plant Mol Biol. 1995 Jan;27(2):237-48                      |
| Arabidopsis PHT1                     | Kovama y col., 2005; Mudge y col. (2002, Plant J. 31: 341) |
| Transportador de fosfato en Medicago | Xiao y col., 2006  |



(continuación)

| Fuente del gen                          | Referencia   |
|---|--|
| Arabidopsis Pyk10                       | Nitz y col. (2001) Plant Sci 161(2): 337-346                               |
| genes expresables en raíz               | Tingey y col., EMBO J. 6: 1, 1987.   |
| gen inducible por auxina en tabaco      | Van der Zaal y col., Plant Mol. Biol. 16, 983, 1991.                       |
| $\beta$ -tubulina                       | Oppenheimer, y col., Gene 63: 87, 1988.                                    |
| genes específicos de la raíz del tabaco | Conkling, y col., Plant Physiol. 93: 1203, 1990.                           |
| gen G1-3b de B. napus                   | Patente de Estados Unidos No. 5.401.836                                    |
| SbPRP1                                  | Suzuki y col., Plant Mol. Biol. 21: 109-119, 1993.                         |
| LRX1                                    | Baumberger y col. 2001, Genes & Dev. 15:1128                               |
| BTG-26 de Brassica napus                | US 20050044585   |
| LeAMT1 (tomate)                         | Lauter y col. (1996, PNAS 3:8139)  |
| The LeNRT1-1 (tomate)                   | Lauter y col. (1996, PNAS 3:8139)  |
| gen de la patatina de clase I (patata)  | Liu y col., Plant Mol. Biol. 153:386-395, 1991.                            |
| KDC1 (Daucus carota)                    | Downey y col. (2000, J. Biol. Chem. 275:39420)                             |
| gen de TobRB7                           | W Song (1997) PhD Thesis, North Carolina State University, Raleigh, NC USA |
| OsRAB5a (raíz)                          | Wang y col. 2002, Plant Sci. 163:273                                       |
| ALF5 (Arabidopsis)                      | Diener y col. (2001, Plant Cell 13:1625)                                   |
| NRT2;1Np (N. plumbaginifolia)           | Quesada y col. (1997, Plant Mol. Biol. 34:265)                             |

- 5 Un promotor específico de semilla es transcripcionalmente activo predominantemente en tejidos de semilla, pero no necesariamente de forma exclusiva en tejidos de semilla (en casos de expresión parcial). El promotor específico de semilla puede estar activo durante el desarrollo y/o durante la germinación de la semilla. El promotor específico de semilla puede ser específico de endospermo y/o aleurona y/o embrión. A continuación en las Tablas 2c, d, e, f, se muestran ejemplos de promotores específicos de semilla (específicos de endospermo/aleurona/embrión). Otros ejemplos de promotores específicos de semilla se proporcionan en Qing Qu y Takaiwa (Plant Biotechnol. J. 2, 113-125, 2004), cuya divulgación se incorpora por referencia en el presente documento como si se expusiera completamente.
- 10

**Tabla 2c:** Ejemplos de promotores específicos de semilla

| Fuente del gen                   | Referencia  |
|----------------------------------|---|
| genes específicos de semilla     | Simon y col., Plant Mol. Biol. 5: 191, 1985;        |
|                                  | Scofield y col., J. Biol. Chem. 262: 12202, 1987;   |
|                                  | Baszczynski y col., Plant Mol. Biol. 14: 633, 1990. |
| albúmina de Nuez del Brasil      | Pearson y col., Plant Mol. Biol. 18: 235-245, 1992. |
| Legumina                         | Ellis y col., Plant Mol. Biol. 10: 203-214, 1988.   |
| Glutelina (arroz)                | Takaiwa y col., Mol. Gen. Genet. 208: 15-22, 1986;  |
|                                  | Takaiwa y col., FEBS Letts. 221: 43-47, 1987.       |
| Zeína                            | Matzke y col Plant Mol Biol, 14(3):323-32 1990      |
| napA                             | Stalberg y col, Planta 199: 515-519, 1996.          |
| gluteína-1 de BPM y APM de trigo | Mol Gen Genet 216:81-90, 1989; NAR 17:461-2, 1989   |

## ES 2 574 981 T3

(continuación)

| <b>Fuente del gen</b>                             | <b>Referencia</b>  |
|---|--|
| SPA de trigo                                      | Albani y col, Plant Cell, 9: 171-184, 1997   |
| $\alpha$ , $\beta$ y $\gamma$ gliadinas de trigo  | EMBO J. 3:1409-15, 1984  |
| promotor Itr1 de cebada                           | Diaz y col. (1995) Mol Gen Genet 248(5):592-8  |
| hordeína B1, C, D de cebada                       | Theor Appl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750-60, 1996              |
| DOF de cebada                                     | Mena y col, The Plant Journal, 116(1): 53-62, 1998   |
| blz2  | EP99106056.7   |
| Promotor sintético                                | Vicente-Carbajosa y col., Plant J. 13: 629-640, 1998.  |
| prolamina NRP33 de arroz                          | Wu y col, Plant Cell Physiology 39(8) 885-889, 1998  |
| $\alpha$ globulina Glb-1 de arroz                 | Wu y col, Plant Cell Physiology 39(8) 885-889, 1998  |
| OSH1 de arroz                                     | Sato y col, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996  |
| $\alpha$ globulina REB/OHP-1 de arroz             | Nakase y col. Plant Mol. Biol. 33: 513-522, 1997   |
| ADP-glucosa pirofosforilasa de arroz              | Trans Res 6:157-68, 1997   |
| familia del gen ESR de maíz                       | Plant J 12:235-46, 1997  |
| $\alpha$ -kafirina de sorgo                       | DeRose y col., Plant Mol. Biol 32:1029-35, 1996  |
| KNOX  | Postma-Haarsma y col, Plant Mol. Biol. 39:257-71, 1999   |
| oleosina de arroz                                 | Wu y col, J. Biochem. 123:386, 1998  |
| oleosina de girasol                               | Cummins y col., Plant Mol. Biol. 19: 873-876, 1992   |
| PRO0117, supuesta proteína ribosomal 40S de arroz | WO 2004/070039   |
| PRO0136, alanina aminotransferasa de arroz        | No publicado   |
| PRO0147, inhibidor de tripsina ITR1 (cebada)      | No publicado   |
| PRO0151, WSI18 de arroz                           | WO 2004/070039   |
| PRO0175, RAB21 de arroz                           | WO 2004/070039   |
| PRO005  | WO 2004/070039   |
| PRO0095   | WO 2004/070039   |
| $\alpha$ -amilasa (Amy32b)                        | Lanahan y col, Plant Cell 4: 203-211, 1992; Skriver y col, Proc Natl Acad Sci USA 88:7266-7270, 1991 |
| gen similar a la $\beta$ catepsina                | Cejudo y col, Plant Mol Biol 20: 849-856, 1992   |
| Ltp2 de cebada                                    | Kalla y col., Plant J. 6: 849-60, 1994   |
| Chi26   | Leah y col., Plant J. 4: 579-89, 1994  |
| B-Perú de maíz                                    | Selinger y col., Genetics 149; 1125-38, 1998   |

**Tabla 2d:** ejemplos de promotores específicos de endospermo

| <b>Fuente del gen</b> | <b>Referencia</b>  |
|-----------------------|--|
| glutelina (arroz)     | Takaiwa y col. (1986) Mol Gen Genet 208:15-22; Takaiwa y col. (1987) FEBS Letts. 221:43-47 |

(continuación)

| Fuente del gen                       | Referencia  |
|--------------------------------------|---|
| zeína                                | Matzke y col., (1990) Plant Mol Biol 14(3): 323-32  |
| gluteína-1 de BPM y APM de trigo     | Colot y col. (1989) Mol Gen Genet 216:81-90, Anderson y col. (1989) NAR 17:461-2  |
| SPA de trigo                         | Albani y col. (1997) Plant Cell 9:171-184   |
| gliadinas de trigo                   | Rafalski y col. (1984) EMBO 3:1409-15   |
| promotor ltr1 de cebada              | Diaz y col. (1995) Mol Gen Genet 248(5):592-8   |
| hordeína B1, C, D de cebada          | Cho y col. (1999) Theor Appl Genet 98:1253-62; Muller y col. (1993) Plant J 4:343-55; Sorenson y col. (1996) Mol Gen Genet 250:750-60 |
| DOF de cebada                        | Mena y col, (1998) Plant J 116(1): 53-62  |
| blz2                                 | Oate y col. (1999) J Biol Chem 274(14):9175-82  |
| Promotor sintético                   | Vicente-Carbajosa y col. (1998) Plant J 13:629-640  |
| prolamina NRP33 de arroz             | Wu y col, (1998) Plant Cell Physiol 39(8) 885-889   |
| globulina Glb-1 de arroz             | Wu y col. (1998) Plant Cell Physiol 39(8) 885-889   |
| globulina REB/OHP-1 de arroz         | Nakase y col. (1997) Plant Molec Biol 33: 513-522   |
| ADP-glucosa pirofosforilasa de arroz | Russell y col. (1997) Trans Res 6:157-68  |
| familia del gen ESR de maíz          | Opsahl-Ferstad y col. (1997) Plant J 12: 235-46   |
| Kafirina de sorgo                    | DeRose y col. (1996) Plant Mol Biol 32: 1029-35   |

**Tabla 2e:** Ejemplos de promotores específicos de embrión:

| Fuente del gen | Referencia  |
|----------------|---|
| OSH1 de arroz  | Sato y col, Proc. Natl. Acad. Sci. USA, 93: 8117-8122, 1996 |
| KNOX           | Postma-Haarsma y col, Plant Mol. Biol. 39: 257-71, 1999     |
| PRO0151        | WO 2004/070039  |
| PRO0175        | WO 2004/070039  |
| PRO005         | WO 2004/070039  |
| PRO0095        | WO 2004/070039  |

5

**Tabla 2f:** Ejemplos de promotores específicos de aleurona:

| Fuente del gen                  | Referencia  |
|---------------------------------|---|
| $\alpha$ -amilasa (Amy32b)      | Lanahan y col, Plant Cell 4:203-211, 1992; Skriver y col, Proc Natl Acad Sci USA 88:7266-7270, 1991 |
| gen similar a $\beta$ catepsina | Cejudo y col, Plant Mol Biol 20:849-856, 1992   |
| Ltp2 de cebada                  | Kalla y col., Plant J. 6:849-60, 1994   |
| Chi26                           | Leah y col., Plant J. 4:579-89, 1994  |
| B-Perú de maíz                  | Selinger y col., Genetics 149:1125-38, 1998   |

Un promotor específico de tejido verde como se define en el presente documento es un promotor que es transcripcionalmente activo predominantemente en tejido verde, sustancialmente con la exclusión de cualquier otra

parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta.

Los ejemplos de promotores específicos de tejido verde que pueden usarse para realizar los procedimientos de la invención se muestran a continuación en la Tabla 2g.

**Tabla 2g:** Ejemplo de promotores específicos de tejido verde

| Gen  | Expresión           | Referencia             |
|--|---------------------|------------------------|
| Ortofosfato diquinasa de maíz                  | Específica de hoja  | Fukavama y col., 2001  |
| Fosfoenolpiruvato carboxilasa de maíz          | Específica de hoja  | Kausch y col., 2001    |
| Fosfoenolpiruvato carboxilasa de arroz         | Específica de hoja  | Liu y col., 2003       |
| Pequeña subunidad de Rubisco de arroz          | Específica de hoja  | Nomura y col., 2000    |
| Beta expansina EXBP9 de arroz                  | Específica de brote | WO 2004/070039         |
| Pequeña subunidad de Rubisco de frijol de palo | Específica de hoja  | Panguluri y col., 2005 |
| RBCS3A de guisante                             | Específica de hoja  |                        |

5

Otro ejemplo de un promotor específico de tejido es un promotor específico de meristemo, que es transcripcionalmente activo predominantemente en tejido meristemático, sustancialmente con la exclusión de cualquier otra parte de una planta, mientras que aún permite cualquier expresión parcial en estas otras partes de la planta. Los ejemplos de promotores específicos de meristemo verde que pueden usarse para realizar los procedimientos de la invención se muestran en la siguiente Tabla 2h.

10

**Tabla 2h:** Ejemplos de promotores específicos de meristemo

| Fuente del gen          | Patrón de expresión   | Referencia   |
|-------------------------|---|--|
| OSH1 de arroz           | Meristemo apical de brote, desde la fase globular embrionaria a la fase de plántula | Sato y col. (1996) Proc. Natl. Acad. Sci. USA, 93: 8117-8122 |
| Metalotioneína de arroz | Específico de meristemo   | BAD87835.1   |
| WAK1 y WAK 2            | Meristemos apicales de brote y raíz y en hojas y sépalos en expansión               | Wagner y Kohorn (2001) Plant Cell 13(2): 303-318             |

### Terminador

15

El término "terminador" incluye una secuencia de control que es una secuencia de ADN en el extremo de una unidad transcripcional que señala el procesamiento y la poliadenilación en dirección 3' de un transcripto primario y la terminación de la transcripción. El terminador puede proceder del gen natural, de una variedad de otros genes de planta, o de ADN T. El terminador a añadir puede proceder, por ejemplo, de genes de nopalina sintasa u octopina sintasa, o como alternativa, de otro gen de planta, o menos preferentemente, de cualquier otro gen eucariota.

### Marcador de selección (gen)/Gen indicador

20

Un "marcador de selección", "gen marcador de selección" o "gen indicador" incluye cualquier gen que confiere un fenotipo a una célula en la que se expresa para facilitar la identificación y/o selección de células que se transfectan o transforman con una construcción de ácido nucleico de la invención. Estos genes marcadores permiten la identificación de una transferencia satisfactoria de las moléculas de ácido nucleico mediante una serie de principios diferentes. Los marcadores adecuados pueden seleccionarse de marcadores que confieren resistencia a antibióticos o a herbicidas, que introducen un nuevo rasgo metabólico o que permiten selección visual. Los ejemplos de genes marcadores de selección incluyen genes que confieren resistencia a antibióticos (tales como nptII que fosforila neomicina y kanamicina, o hpt, que fosforila higromicina, o genes que confieren resistencia a, por ejemplo, bleomicina, estreptomycin, tetraciclina, cloranfenicol, ampicilina, gentamicina, geneticina (G418), espectinomycin o blastidina), a herbicidas (por ejemplo bar que proporciona resistencia a Basta®; aroA o gox que proporciona resistencia contra glifosato, o los genes que confieren resistencia a, por ejemplo, imidazolinona, fosfinotricina o sulfonilurea), o genes que proporcionan un rasgo metabólico (tal como manA que permite a las plantas usar manosa como única fuente de carbono o xilosa isomerasa para la utilización de xilosa, o marcadores antinutritivos tales como

25

30

la resistencia a 2-desoxiglucosa). La expresión de genes marcadores visuales da como resultado la formación de color (por ejemplo  $\beta$ -glucuronidasa, GUS o  $\beta$ -galactosidasa con sus sustratos de color, por ejemplo X Gal), luminiscencia (tal como el sistema luciferina/luciferasa) o fluorescencia (Proteína Fluorescente Verde, GFP y derivados de los mismos). Esta lista solo representa una pequeña cantidad de marcadores posibles. El experto está familiarizado con dichos marcadores. Se prefieren diferentes marcadores dependiendo del organismo y del procedimiento de selección.

#### Transgénico/Transgén/Recombinante

Para los fines de la invención, "transgénico", "transgén" o "recombinante" significa con respecto a, por ejemplo, una secuencia de ácidos nucleicos, un casete de expresión, una construcción génica o un vector que comprende la secuencia de ácidos nucleicos o un organismo transformado con las secuencias de ácidos nucleicos, casetes o vectores de expresión de acuerdo con la invención, todas estas construcciones realizadas por procedimientos recombinantes en los que

- (a) las secuencias de ácidos nucleicos que codifican las proteínas útiles en los procedimientos de la invención o
- (b) la secuencia (o secuencias) de control genético que se une operativamente con la secuencia de ácidos nucleicos de acuerdo con la invención, por ejemplo un promotor, o
- (c) a) y b)

no se localizan en su ambiente genético natural o se han modificado por procedimientos recombinantes, siendo posible que se produzca la modificación de, por ejemplo, una sustitución, adición, delección, inversión o inserción de uno o más restos de nucleótido. Se entiende que, ambiente genético natural, significa el locus genómico o cromosómico natural en la planta original o la presencia en una genoteca. En el caso de una genoteca, el ambiente genético natural de la secuencia de ácidos nucleicos se conserva preferentemente, por lo menos en parte. El ambiente flanquea la secuencia de ácidos nucleicos por lo menos en un lado y tiene una longitud de secuencia de por lo menos 50 pb, preferentemente por lo menos 500 pb, en especial preferentemente por lo menos 1000 pb, más preferentemente por lo menos 5000 pb. Un casete de expresión de origen natural - por ejemplo la combinación de origen natural del promotor natural de las secuencias de ácidos nucleicos con la secuencia de ácidos nucleicos correspondiente que codifica un polipéptido útil en los procedimientos de la presente invención, como se ha definido anteriormente - llega a ser un casete de expresión transgénica cuando esta casete de expresión se modifica mediante procedimientos sintéticos no naturales ("artificiales") tal como, por ejemplo, tratamiento mutagénico. Se describen procedimientos adecuados, por ejemplo, en los documentos US 5.565.350 o WO 00/15815.

Por lo tanto, una planta transgénica, para los fines de la invención, se entiende que significa, como se ha indicado anteriormente, que los ácidos nucleicos usados en el procedimiento de la invención no están en su locus natural en el genoma de dicha planta, siendo posible que los ácidos nucleicos se expresen de manera homóloga o heteróloga. Sin embargo, como se ha mencionado, transgénico también significa que, aunque los ácidos nucleicos de acuerdo con la invención o usados en el procedimiento de la invención están en su posición natural en el genoma de una planta, la secuencia se ha modificado con respecto a la secuencia natural y/o que las secuencias reguladoras de las secuencias naturales se han modificado. Por transgénico se entiende preferentemente que significa la expresión de los ácidos nucleicos de acuerdo con la invención en un locus no natural en el genoma, es decir, se produce la expresión homóloga o, preferentemente, heteróloga de los ácidos nucleicos. En el presente documento se mencionan las plantas transgénicas preferidas.

#### Transformación

El término "introducción" o "transformación" como se denomina en el presente documento, incluye la transferencia de un polinucleótido exógeno en una célula huésped, independientemente del procedimiento usado para la transferencia. El tejido de planta capaz de propagación clonal posterior, mediante organogénesis o embriogénesis, puede transformarse con una construcción genética de la presente invención y a partir del mismo regenerarse una planta completa. El tejido particular seleccionado variará dependiendo de los sistemas de propagación clonales disponibles para, y mejor adaptados a, la especie particular que se va a transformar. Las dianas tisulares ejemplares incluyen discos foliares, polen, embriones, cotiledones, hipocotiledones, megagametofitos, tejido de callo, tejido meristemático existente (por ejemplo, meristemo apical, brotes axilares y meristemos radiculares) e inducen tejido meristemático (por ejemplo, meristemo de cotiledón y meristemo de hipocotiledón). El polinucleótido puede introducirse de manera estable o transitoria en una célula huésped y puede mantenerse no integrado, por ejemplo, como un plásmido. Como alternativa, éste puede integrarse en el genoma huésped. La célula de la planta transformada resultante puede después usarse para regenerar una planta transformada de una manera conocida por los expertos en la técnica.

La transferencia de genes externos dentro del genoma de una planta se denomina transformación. La transformación de especies plantas es actualmente una técnica muy habitual. Ventajosamente, cualquiera de los diversos procedimientos de transformación puede usarse para introducir el gen de interés en una célula antecesora adecuada. Los procedimientos descritos para la transformación y regeneración de plantas de tejidos de plantas o células de plantas pueden utilizarse para la transformación transitoria o estable. Los procedimientos de transformación incluyen el uso de liposomas, electroporación, productos químicos que aumentan la captación de

ADN libre, inyección del ADN directamente en la planta, pistola de bombardeo de partículas, transformación usando virus o polen y microproyección. Pueden seleccionarse procedimientos entre, el procedimiento de calcio/poli-etilenglicol para protoplastos (Krens, F. A. y col., (1982) *Nature* 296, 72-74; Negrutiu I y col. (1987) *Plant Mol Biol* 8: 363-373); electroporación de protoplastos (Shillito R. D. y col. (1985) *Bio/Technol* 3, 1099-1102); microinyección en el material de planta (Crossway A y col., (1986) *Mol. Gen Genet* 202: 179-185); bombardeo de partículas revestido con ADN o ARN (Klein TM y col., (1987) *Nature* 327: 70) infección con virus (no integrantes) y similares. Las plantas transgénicas, incluyendo plantas de cultivo transgénicas, se producen preferentemente mediante transformación mediada por *Agrobacterium*. Un procedimiento de transformación ventajoso es la transformación en la planta. Para esta finalidad, es posible, por ejemplo, permitir que la agrobacterias actúen en las semillas de la planta o inocular las agrobacterias en el meristemo de la planta. Se ha demostrado de un modo particularmente conveniente, de acuerdo con la invención, permitir que una suspensión de agrobacterias transformadas actúe en la planta intacta o por lo menos en los primordios florales. Posteriormente la planta crece hasta que se obtienen las semillas de la planta tratada (Clough y Bent, *Plant J.* (1998) 16, 735-743). Los procedimientos para la transformación de arroz mediada por *Agrobacterium* incluyen procedimientos bien conocidos para la transformación de arroz, tales como los descritos en cualquiera de las siguientes referencias bibliográficas: solicitud de patente Europea EP 1198985 A1, Aldemita y Hodges (*Planta* 199: 612-617, 1996); Chan y col. (*Plant Mol Biol* 22 (3): 491-506, 1993), Hiei y col. (*Plant J* 6 (2): 271-282, 1994), cuyas divulgaciones se incorporan por referencia en el presente documento como si se expusieran completamente. En el caso de la transformación de maíz, el procedimiento preferido es como se describe en Ishida y col. (*Nat. Biotechnol* 14(6): 745-50, 1996) o en Frame y col. (*Plant Physiol* 129(1): 13-22, 2002), cuyas divulgaciones se incorporan por referencia en el presente documento como si se expusieran completamente. Dichos procedimientos se describen adicionalmente, a modo de ejemplo, en B. Jené y col., *Techniques for Gene Transfer*, en: *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press (1993) 128-143 y en Potrykus *Annu. Rev. Plant Physiol. Plant Molec. Biol.* 42 (1991) 205-225). Los ácidos nucleicos o la construcción a expresar se clona preferentemente en un vector, que es adecuado para transformar *Agrobacterium tumefaciens*, por ejemplo, pBin19 (Bevan y col., *Nucl. Acids Res.* 12 (1984) 8711). Las agrobacterias transformadas por dicho vector pueden después usarse de una manera conocida para la transformación de plantas, tales como plantas usadas como un modelo, como *Arabidopsis* (*Arabidopsis thaliana* que se encuentra dentro del ámbito de la presente invención no se considera como una planta de cultivo) o plantas de cultivo tales como, a modo de ejemplo, plantas de tabaco, por ejemplo, sumergiendo hojas laceradas o picadas en una solución agrobacteriana y después cultivarlas en un medio adecuado. La transformación de plantas mediante *Agrobacterium tumefaciens* la describen, por ejemplo, Höfgen y Willmitzer en *Nucl. Acid Res.* (1988) 16, 9877 o se conoce, entre otros, de F. F. White, *Vectors for Gene Transfer in Higher Plants*; in *Transgenic Plants*, Vol. 1, Engineering and Utilization, eds. S. D. Kung y R. Wu, Academic Press, 1993, págs. 15-38.

Además de la transformación de células somáticas, que después deben regenerarse en plantas intactas, también es posible transformar las células de meristemos de planta y en particular aquellas células que se desarrollan en gametos. En este caso, los gametos transformados siguen el desarrollo natural de la planta, dando lugar a plantas transgénicas. Por tanto, por ejemplo, semillas de *Arabidopsis* se tratan con agrobacterias y se obtienen semillas del desarrollo de plantas de las que se transforma una determinada proporción y por tanto transgénicas [Feldman, KA y Marks MD (1987). *Mol Gen Genet* 208: 274-289; Feldmann K (1992). En: C Koncz, N-H Chua y J Shell, eds, *Methods in Arabidopsis Research*. World Scientific, Singapore, pág. 274-289]. Los procedimientos alternativos se basan en la retirada repetida de las inflorescencias e incubación del sitio de escisión en el centro de la roseta con agrobacterias transformadas, por lo que las semillas transformadas pueden obtenerse de modo similar en un punto de tiempo final (Chang (1994). *Plant J.* 5: 551-558; Katavic (1994). *Mol Gen Genet*, 245: 363-370). Sin embargo, un procedimiento especialmente eficaz es el procedimiento de infiltración por vacío con sus modificaciones tal como el procedimiento de "inmersión floral". En el caso de la infiltración por vacío de *Arabidopsis*, las plantas intactas bajo presión reducida se tratan con una suspensión agrobacteriana [Bechthold, N (1993). *C R Acad Sci Paris Life Sci*, 316: 1194-1199], mientras que en el caso del procedimiento de "inmersión floral" el tejido floral en desarrollo se incuba brevemente con una suspensión agrobacteriana tratada con tensioactivo [Clough, SJ y Bent, AF (1998). *The Plant J.* 16, 735-743]. En ambos casos se recoge una determinada proporción de semillas transgénicas y estas semillas pueden diferenciarse de semillas no transgénicas al cultivar bajo las condiciones selectivas descritas anteriormente. Además, la transformación estable de los plástidos es ventajosa ya que estos se heredan por vía materna en la mayor parte de los cultivos reduciendo o eliminando el riesgo de flujo transgénico a través del polen. La transformación del genoma del cloroplasto generalmente se realiza mediante un proceso que han presentado esquemáticamente Klaus y col., 2004 [*Nature Biotechnology* 22 (2), 225-229]. En resumen, las secuencias a transforman se clonan junto con un gen marcador de selección entre las secuencias flanqueantes homólogas al genoma del cloroplasto. Estas secuencias flanqueantes homólogas dirigen la integración específica de sitio en el plastoma. La transformación plastidial se ha descrito para muchas especies de plantas diferentes y se ofrece una revisión en Bock (2001) *Transgenic plastids in basic research and plant biotechnology*. *J Mol Biol.* 2001 Sep 21; 312 (3):425-38 o Maliga, P (2003) *Progress towards commercialization of plastid transformation technology*. *Trends Biotechnol.* 21, 20-28. Recientemente se han descrito progresos biotecnológicos adicionales, en forma de transformantes plastidiales sin marcador, que pueden producirse mediante un gen marcador cointegrado transitorio (Klaus y col., 2004, *Nature Biotechnology* 22(2), 225-229).

TILLING

TILLING (Targeted Induced Local Lesions In Genomes, Inducción Dirigida de Lesiones Locales en el Genoma) es una tecnología de mutagénesis útil para generar y/o identificar ácidos nucleicos que codifican proteínas con expresión y/o actividad modificada. El TILLING también permite la selección de plantas que llevan dichas variantes mutantes. Estas variantes mutantes pueden presentar expresión modificada, en fuerza o en localización o en tiempo (si las mutaciones afectan, por ejemplo, al promotor). Estas variantes mutantes pueden presentar mayor actividad que la mostrada por el gen en su forma natural. El TILLING combina mutagénesis de alta densidad con procedimientos de detección de alto rendimiento. Las etapas seguidas típicamente en el TILLING son: (a) mutagénesis EMS (Redei GP y Koncz C (1992) In *Methods in Arabidopsis Research*, Koncz C, Chua NH, Schell J, eds. Singapore, World Scientific Publishing Co, páginas 16-82; Feldmann y col., (1994) In Meyerowitz EM, Somerville CR, eds, *Arabidopsis*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 137-172; Lightner J y Caspar T (1998) In J Martinez-Zapater, J Salinas, eds, *Methods on Molecular Biology*, Vol. 82. Humana Press, Totowa, NJ, páginas 91-104); (b) preparación y agrupamiento de ADN en individuos; (c) amplificación por PCR de una región de interés; (d) desnaturalización e hibridación para permitir la formación de heterodúplex; (e) DHPLC, donde la presencia de un heterodúplex en un grupo se detecta como un pico extra en el cromatograma; (f) identificación del mutante individual; y (g) secuenciación del producto PCR mutante. Los procedimientos de TILLING son bien conocidos en la técnica (McCallum y col., (2000) *Nat Biotechnol* 18: 455-457; revisado por Stemple (2004) *Nat Rev Genet* 5(2): 145-50).

Rendimiento

El término “rendimiento” en general significa un producto medible de valor económico, típicamente relacionado con un cultivo, un área y un periodo de tiempo específicos. Las partes de planta individuales contribuyen directamente al rendimiento en base a su número, tamaño y/o peso, o el rendimiento real es el rendimiento por metro cuadrado de un cultivo y año, que se determina dividiendo la producción total (incluyendo la producción tanto cosechada como valorada) por metro cuadrado plantado.

Aumento/Mejora/Potenciación

Los términos “aumento”, “mejora” o “potenciación” son intercambiables y significan en el sentido de la solicitud por lo menos el 5 %, 6 %, 7 %, 8 %, 9 % o 10 %, preferentemente por lo menos el 15 % o el 20 %, más preferentemente el 25 %, 30 %, 35 % o 40 % más rendimiento y/o crecimiento en comparación con plantas de control como se define en el presente documento.

Rendimiento de semilla

El rendimiento de semilla aumentado puede manifestarse por sí mismo como uno o más de los siguientes: a) un aumento en biomasa de semilla (peso total de semilla) que puede establecerse sobre una semilla individual y/o por planta y/o hectárea o metro cuadrado; b) número de flores por planta aumentado; c) número de semillas (llenas) aumentado; d) tasa de llenado de semilla aumentado (que se expresa como la proporción entre el número de semillas llenas dividido entre el número total de semillas); e) índice de cosecha aumentado, que se expresa como una proporción del rendimiento de partes cosechables, tales como semillas, dividido entre la biomasa total, y f) peso de mil granos (PMG), que se extrapola a partir del número de semillas llenas contadas y su peso total. Un PMG aumentado puede ser el resultado de un tamaño y/o peso de semilla aumentado y también puede ser el resultado de un aumento en el tamaño del embrión y/o endospermo.

También puede manifestarse un aumento en el rendimiento de la semilla como un aumento del tamaño de la semilla y/o volumen de la semilla. Además, también puede manifestarse un aumento en el rendimiento de la semilla en sí mismo como un aumento en el área de semilla y/o longitud de semilla y/o anchura de semilla y/o un perímetro de semilla. El rendimiento aumentado puede ser también el resultado de arquitectura modificada, o puede ocurrir debido a la arquitectura modificada.

Planta

El término “planta” como se usa en el presente documento incluye plantas completas, antecesores y progenie de las plantas y partes de las plantas, incluyendo semillas, brotes, tallos, hojas, raíces (incluyendo tubérculos), flores y tejidos y órganos, donde cada uno de los elementos anteriormente mencionados comprende el gen/ácido nucleico de interés. El término “planta” también incluye células de planta, cultivos en suspensión, tejido de callo, embriones, regiones meristemáticas, gametofitos, esporofitos, polen y microesporas, de nuevo donde cada uno de los elementos mencionados anteriormente comprende el gen/ácido nucleico de interés.

Las plantas que son particularmente útiles en los procedimientos de la invención incluyen todas las plantas que pertenecen a la superfamilia *Viridiplantae*, en particular plantas monocotiledóneas y dicotiledóneas incluyendo forraje o leguminosas forrajeras, plantas ornamentales, cultivos alimenticios, árboles o arbustos seleccionados de la lista que comprende *Acer spp.*, *Actinidia spp.*, *Abelmoschus spp.*, *Agave sisalana*, *Agropyron spp.*, *Agrostis stolonifera*, *Allium spp.*, *Amaranthus spp.*, *Ammophila arenaria*, *Ananas comosus*, *Annona spp.*, *Apium graveolens*, *Arachis spp.*, *Artocarpus spp.*, *Asparagus officinalis*, *Avena spp.* (por ejemplo, *Avena sativa*, *Avena fatua*, *Avena byzantina*, *Avena*

*fatua var. sativa, Avena hybrida), Averrhoa carambola, Bambusa sp., Benincasa hispida, Bertholletia excelsea, Beta vulgaris, Brassica spp. (por ejemplo, Brassica napus, Brassica rapa ssp. [canola, colza, nabo]), Cadaba farinosa, Camellia sinensis, Canna indica, Cannabis sativa, Capsicum spp., Carex elata, Carica papaya, Carissa macrocarpa, Carya spp., Carthamus tinctorius, Castanea spp., Ceiba pentandra, Cichorium endivia, Cinnamomum spp., Citrullus lanatus, Citrus spp., Cocos spp., Coffea spp., Colocasia esculenta, Cola spp., Corchorus sp., Coriandrum sativum, Corylus spp., Crataegus spp., Crocus sativus, Cucurbita spp., Cucumis spp., Cynara spp., Daucus carota, Desmodium spp., Dimocarpus longan, Dioscorea spp., Diospyros spp., Echinochloa spp., Elaeis (por ejemplo, Elaeis guineensis, Elaeis oleifera), Eleusine coracana, Eragrostis tef, Erianthus sp., Eriobotrya japonica, Eucalyptus sp., Eugenia uniflora, Fagopyrum spp., Fagus spp., Festuca arundinacea, Ficus carica, Fortunella spp., Fragaria spp., Ginkgo biloba, Glycine spp. (por ejemplo, Glycine max, Soja hispida o Soja max), Gossypium hirsutum, Helianthus spp. (por ejemplo, Helianthus annuus), Hemerocallis fulva, Hibiscus spp., Hordeum spp. (por ejemplo, Hordeum vulgare), Ipomoea batatas, Juglans spp., Lactuca sativa, Lathyrus spp., Lens culinaris, Linum usitatissimum, Litchi chinensis, Lotus spp., Luffa acutangula, Lupinus spp., Luzula sylvatica, Lycopersicon spp. (por ejemplo, Lycopersicon esculentum, Lycopersicon lycopersicum, Lycopersicon pyriforme), Macrotyloma spp., Malus spp., Malpighia emarginata, Mamea americana, Mangifera indica, Manihot spp., Manilkara zapota, Medicago sativa, Melilotus spp., Mentha spp., Miscanthus sinensis, Momordica spp., Morus nigra, Musa spp., Nicotiana spp., Olea spp., Opuntia spp., Ornithopus spp., Oryza spp. (por ejemplo Oryza sativa, Oryza latifolia), Panicum miliaceum, Panicum virgatum, Passiflora edulis, Pastinaca sativa, Pennisetum sp., Persea spp., Petroselinum crispum, Phalaris arundinacea, Phaseolus spp., Phleum pratense, Phoenix spp., Phragmites australis, Physalis spp., Pinus spp., Pistacia vera, Pisum spp., Poa spp., Populus spp., Prosopis spp., Prunus spp., Psidium spp., Punica granatum, Pyrus communis, Quercus spp., Raphanus sativus, Rheum rhabarbarum, Ribes spp., Ricinus communis, Rubus spp., Saccharum spp., Salix sp., Sambucus spp., Secale cereale, Sesamum spp., Sinapis sp., Solanum spp. (por ejemplo, Solanum tuberosum, Solanum integrifolium o Solanum lycopersicum), Sorghum bicolor, Spinacia spp., Syzygium spp., Tagetes spp., Tamarindus indica, Theobroma cacao, Trifolium spp., Triticale sp., Triticosecale rimpai, Triticum spp. (por ejemplo, Triticum aestivum, Triticum durum, Triticum turgidum, Triticum hybernum, Triticum macha, Triticum sativum o Triticum vulgare), Tropaeolum minus, Tropaeolum majus, Vaccinium spp., Vicia spp., Vigna spp., Viola odorata, Vitis spp., Zea mays, Zizania palustris, Ziziphus spp., entre otras.*

#### **Descripción detallada de la invención**

De acuerdo con una primera realización, la presente invención proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas, que comprende introducir y expresar en una planta un ácido nucleico que codifica un polipéptido Factor G asociado a Harpin (en lo sucesivo en este documento denominado "HpaG") según se definió en el presente documento, en el que dicho aumento del rendimiento de semilla y/o de la biomasa se obtiene en condiciones sin estrés o de estrés abiótico o en el que dichos rendimiento de semillas y/o biomasa aumentados no son el resultado de la resistencia aumentada al estrés biótico.

Un procedimiento preferente para aumentarla expresión de un ácido nucleico que codifica un polipéptido HpaG es mediante la introducción y expresión en una planta de un ácido nucleico que codifica un polipéptido HpaG.

En lo sucesivo, cualquier referencia en el presente documento a "proteína útil en los procedimientos de la invención" se considera que significa un polipéptido HpaG como se define en el presente documento... En lo sucesivo, cualquier referencia en el presente documento a un "ácido nucleico útil en los procedimientos de la invención" se considera que significa un ácido nucleico capaz de codificar tal polipéptido HpaG. El ácido nucleico que se introducirá en una planta (y por lo tanto útil en la realización de los procedimientos de la invención) es cualquier ácido nucleico que codifique el tipo de proteína que será descrita ahora, en lo sucesivo también llamado "ácido nucleico de HpaG" o "gen de HpaG".

Un polipéptido HpaG como se define en el presente documento comprende cualquier polipéptido que tenga las siguientes características:

- (i) en orden creciente de preferencia, por lo menos el 80 % o más de identidad de secuencia con la secuencia del polipéptido HpaG representada en la SEQ ID NO: 2; y
- (ii) una composición de aminoácidos en la que el contenido de glicina varía desde entre aproximadamente el 13 % y aproximadamente el 25 %, el contenido de glutamina varía entre aproximadamente el 13 % y aproximadamente el 20 %, el contenido de cisteína varía entre aproximadamente el 0 % y aproximadamente el 1 %, el contenido de histidina varía desde aproximadamente el 0 % y aproximadamente el 1 %, y en la que el triptófano está ausente.

Preferentemente, la longitud del polipéptido HpaG varía entre aproximadamente 121 y aproximadamente 143 aminoácidos.

Preferentemente, la proteína HpaG también comprende el motivo conservado 1 (SEQ ID NO: 3)  
 G (G/E/D) (N/E)X(Q/R/P)Q(A/S)GX(N/D)G  
 en el que X en la posición 4 puede ser cualquier aminoácido, preferentemente uno de S, N, P, R o Q, y en el que X en la posición 9 puede ser cualquier aminoácido, preferentemente uno de Q, E, S o P; y/o el motivo conservado 2 (SEQ ID NO: 4)



(P/A/V)S(P/Q/A)(F/L/Y)TQ(M/A)LM(H/N/Q)IV(G/M)(E/D/Q)

De forma opcional, la proteína HpaG también tiene el motivo conservado 3:

QGISEKQLDQLL

Y/o el motivo conservado 4:

5 ILQAQN

Además, los polipéptidos HpaG (por lo menos en su forma nativa) desencadenan una respuesta de hipersensibilidad en *Arabidopsis thaliana* ecotipo Cvi-0 (Kim y col., J. Bacteriol. 185, 3155-3166, 2003).

10 En el presente documento se describe un homólogo de una proteína HpaG que tiene en orden creciente de preferencia por lo menos el 25 %, 26 %, 27 %, 28 %, 29 %, 30 %, 31 %, 32 %, 33 %, 34 %, 35 %, 36 %, 37 %, 38 %, 39 %, 40 %, 41 %, 42 %, 43 %, 44 %, 45 %, 46 %, 47 %, 48 %, 49 %, 50 %, 51 %, 52 %, 53 %, 54 %, 55 %, 56 %, 57 %, 58 %, 59 %, 60 %, 61 %, 62 %, 63 %, 64 %, 65 %, 66 %, 67 %, 68 %, 69 %, 70 %, 71 %, 72 %, 73 %, 74 %, 75 %, 76 %, 77 %, 78 %, 79 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o 99 % de identidad de secuencia global con los aminoácidos representados en la SEQ ID NO: 2, a condición de que la proteína homóloga comprenda los motivos conservados como se describió anteriormente. La identidad de secuencia global se determina utilizando un algoritmo de alineamiento global, tal como el algoritmo Needleman Wunsch en el programa GAP (GCG Wisconsin Package, Accelrys), preferentemente con los parámetros predeterminados. En comparación con la identidad de secuencia global, en general la identidad de secuencia será mayor cuando se consideren solamente los dominios o motivos conservados.

20 Los términos “dominio” y “motivo” son como se definen en la sección de “Definiciones” del presente documento. Existen bases de datos especializadas para la identificación de dominios, por ejemplo, SMART (Schultz y col. (1998) Proc. Natl. Acad. Sci. USA 95, 5857-5864; Letunic y col. (2002) Nucleic Acids Res 30, 242-244, InterPro (Mulder y col., (2003) Nucl. Acids. Res. 31, 315-318, Prosite (Bucher y Bairoch (1994), A generalized profile syntax for biomolecular SECuencias motifs and its function in automatic SECuencia interpretation. (In) ISMB-94; Proceedings 2nd International Conference on Intelligent Systems for Molecular Biology. Altman R., Brutlag D., Karp P., Lathrop R., Searls D., Eds., páginas 53-61, AAAIPress, Menlo Park; Hulo y col., Nucl. Acids. Res. 32:D134-D137, (2004), o Pfam (Bateman y col., Nucleic Acids Research 30(1): 276-280 (2002). En el servidor de proteomics ExpASy (alojado en el Swiss Institute of Bioinformatics (Gasteiger y col., ExpASY: the proteomics server for in-depth protein knowledge and analysis, Nucleic Acids Res. 31: 3784-3788(2003)) se encuentra disponible un conjunto de herramientas para realizar análisis informatizados de secuencias de proteínas. Los dominios también pueden identificarse usando técnicas rutinarias, tales como mediante alineamiento de secuencias.

30 Los procedimientos para el alineamiento de las secuencias para comparación son bien conocidos en la técnica, dichos procedimientos incluyen GAP, BESTFIT, BLAST, FASTA y TFASTA. GAP utiliza el algoritmo de Needleman Wunsch ((1970) J Mol Biol 48: 443-453) para encontrar el alineamiento global (es decir el tramo de la secuencia completa) de dos secuencias que maximizan el número de emparejamientos y minimiza el número de espacios. El algoritmo BLAST (Altschul y col. (1990) J Mol Biol 215: 403-10) calcula el porcentaje de identidad de secuencia y realiza un análisis estadístico de la similitud entre las dos secuencias. El programa informático para realizar el análisis BLAST está públicamente disponible a través del National Centre for Biotechnology Information (NCBI). Los homólogos pueden identificarse fácilmente usando, por ejemplo, el algoritmo de alineamiento de secuencia múltiple ClustalW (versión 1,83), con los parámetros de alineamiento por pares predeterminados y un procedimiento de clasificación en porcentaje. Los porcentajes globales de similitud e identidad también pueden determinarse utilizando uno de los procedimientos disponibles en el paquete informático MatGAT (Campanella y col., BMC Bioinformatics. Jul 2003 10; 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de ADN o proteína). Como sería evidente para un experto en la materia, puede realizarse una edición manual menor para optimizar el alineamiento entre los motivos conservados. Adicionalmente, también pueden usarse dominios específicos en lugar de usar secuencias de longitud completa para la identificación de homólogos. Los valores de identidad de secuencia pueden determinarse sobre toda la secuencia de ácidos nucleicos o de aminoácidos o sobre los dominios seleccionados o uno o más motivos conservados, usando los programas mencionados anteriormente usando los parámetros predeterminados.

40 La presente invención se ilustra mediante la transformación de plantas con la secuencia de ácidos nucleicos representada por la SEQ ID NO: 1, que codifica la secuencia de polipéptidos de la SEQ ID NO: 2. Sin embargo, la realización de la invención no está limitada a estas secuencias; pudiendo realizarse los procedimientos de la invención de forma ventajosa usando cualquier ácido nucleico que codifique a HpaG o un polipéptido HpaG, como se define en el presente documento.

55 En la Tabla A del Ejemplo 1 del presente documento, se proporcionan ejemplos de ácidos nucleicos que codifican los polipéptidos HpaG. Dichos ácidos nucleicos son útiles para realizar los procedimientos de la invención. Las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1 son secuencias ejemplo de ortólogos y parálogos del polipéptido HpaG representado en la SEQ ID NO: 2, siendo los términos “ortólogos” y “parálogos” como se define en el presente documento. Adicionalmente, los ortólogos y parálogos pueden identificarse fácilmente

realizando una búsqueda BLAST denominada recíproca. Típicamente, esto implica un primer BLAST que implica una secuencia de consulta BLASTing (por ejemplo usando cualquiera de las secuencias indicadas en la Tabla A del Ejemplo 1) contra cualquier base de datos de secuencia, tal como la base de datos NCBI públicamente disponible. BLASTN o TBLASTX (utilizando valores predeterminados convencionales) pueden usarse cuando se parte de una secuencia de nucleótidos y BLASTP o TBLASTN (usando valores predeterminados convencionales) pueden usarse cuando se parte de una secuencia de proteína. Los resultados BLAST pueden filtrarse opcionalmente. Las secuencias de longitud completa de los resultados filtrados o no filtrados vuelven después a buscarse con BLAST (segundo BLAST) contra las secuencias del organismo del cual procede la secuencia de consulta (donde la secuencia de consulta es la SEQ ID NO: 1 o SEQ ID NO: 2, por lo que el segundo BLAST se realizaría contra las secuencias de *Xanthomonas*). Después se comparan los resultados del primer y segundo BLAST. Se identifican un parálogo si un acierto de alta clasificación del primer BLAST es de la misma especie de la que procede la secuencia de consulta, después, un BLAST posterior da como resultado idealmente la secuencia de consulta entre los mayores aciertos; un ortólogo se identifica si un acierto de alta clasificación en el primer BLAST no es de la misma especie de la que procede la secuencia de consulta, y preferentemente los resultados basados en un BLAST posterior en la secuencia de consulta se encuentran entre los mayores aciertos.

Los aciertos de alta clasificación son aquellos que tienen un valor E bajo. Cuanto menor es el valor E, más significativa la clasificación (o en otras palabras menor es la probabilidad de que se encuentre el acierto casualmente). El cálculo del valor E es bien conocido en la técnica. Además de los valores E, las comparaciones también se clasifican mediante el porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácidos nucleicos (o polipéptido) sobre una longitud particular. En el caso de grandes familias, puede usarse ClustalW, seguido de la generación de un árbol de unión del vecino más próximo para ayudar a visualizar el agrupamiento de los genes relacionados y para identificar los ortólogos y parálogos.

Las variantes de ácidos nucleicos también pueden ser útiles en la realización práctica de los procedimientos de la invención. Los ejemplos de dichas variantes incluyen ácidos nucleicos que codifican homólogos y derivados de una cualquiera de las secuencias de aminoácidos proporcionada en la Tabla A del Ejemplo 1, siendo los términos "homólogo" y "derivado" como se definen en el presente documento. También son útiles en los procedimientos de la invención los ácidos nucleicos que codifican homólogos y derivados de ortólogos o parálogos de una cualquiera de las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1. Los homólogos y derivados útiles en los procedimientos de la presente invención tienen esencialmente la misma actividad biológica y funcional que la proteína no modificada de la cual proceden.

Las variantes de ácidos nucleicos adicionales útiles en la realización práctica de los procedimientos de la invención incluyen porciones de ácidos nucleicos que codifican polipéptidos HpaG, ácidos nucleicos que se hibridan con ácidos nucleicos que codifican polipéptidos HpaG y variantes de ácidos nucleicos que codifican polipéptidos HpaG obtenidos por combinación de genes. Las expresiones secuencia de hibridación y combinación de genes son como se describen en el presente documento.

Los ácidos nucleicos que codifican polipéptidos HpaG no tienen por qué ser ácidos nucleicos de longitud completa, dado que la realización de los procedimientos de la invención no se basa en el uso de secuencias de ácidos nucleicos completas. De acuerdo con la presente invención, se proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas, que comprende introducir y expresar en una planta una porción de una cualquiera de las secuencias de ácidos nucleicos proporcionadas en la Tabla A del Ejemplo 1, o una porción de un ácido nucleico que codifica un ortólogo, parálogo y homólogo de cualquiera de las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1.

Una porción de un ácido nucleico puede prepararse, por ejemplo, realizando una o más deleciones en el ácido nucleico. Las porciones pueden usarse en forma aislada o pueden fusionarse a otras secuencias codificantes (o no codificantes) para producir por ejemplo, una proteína que combine diversas actividades. Cuando se fusiona con otras secuencias codificantes, el polipéptido resultante producido después de la traducción puede ser más grande que el esperado para la porción proteica.

Las porciones útiles en los procedimientos de la invención, codifican un polipéptido HpaG como se define en el presente documento, y tienen sustancialmente la misma actividad biológica que la secuencia de aminoácidos proporcionada en la Tabla A del Ejemplo 1. Preferentemente, la porción es una porción de uno cualquiera de los ácidos nucleicos proporcionados en la Tabla A del Ejemplo 1, o es una porción de un ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1. Preferentemente la porción es, en orden creciente de preferencia, de por lo menos 70, 90, 110, 130 nucleótidos consecutivos de longitud, siendo los nucleótidos consecutivos de una cualquiera de las secuencias de ácido nucleico proporcionadas en la Tabla A del Ejemplo 1, o de un ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1. Muy preferentemente la porción es una porción del ácido nucleico de la SEQ ID NO: 1. Preferentemente la porción codifica una secuencia de aminoácidos que cuando se utiliza en la construcción de un árbol filogenético, tal como el representado en la Figura 2, tiende a agruparse con el grupo de polipéptidos HpaG que comprende la secuencia de aminoácidos representada en la SEQ ID NO: 2 en lugar de con cualquier otro grupo.

Otra variante de ácido nucleico útil en los procedimientos de la invención es un ácido nucleico capaz de hibridar, en condiciones de rigurosidad reducida, preferentemente en condiciones rigurosas, con un ácido nucleico que codifica un polipéptido HpaG según se define en el presente documento, o con una porción según se define en el presente documento.

5 De acuerdo con la presente invención, se proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas, que comprende introducir y expresar en una planta un ácido nucleico capaz de hibridar con uno cualquiera de los ácidos nucleicos dados en la Tabla A del Ejemplo 1, o que comprende introducir y expresar en una planta un ácido nucleico capaz de hibridar con un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A del Ejemplo 1.

10 Las secuencias de hibridación útiles en los procedimientos de la invención codifican un polipéptido HpaG como se define en el presente documento, y tienen sustancialmente la misma actividad biológica que las secuencias de aminoácidos proporcionadas en la Tabla A del Ejemplo 1. Preferentemente, la secuencia de hibridación es capaz de hibridar con uno cualquiera de los ácidos nucleicos proporcionados en la Tabla A del Ejemplo 1, o con una porción de cualquiera de estas secuencias, siendo una porción como se define anteriormente, o en la que la secuencia de hibridación es capaz de hibridar a un ácido nucleico que codifica un ortólogo o parálogo de una cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A del Ejemplo 1. Más preferentemente, la secuencia de hibridación es capaz de hibridar con el complemento de un ácido nucleico como se representa en la SEQ ID NO: 1 o con una porción del mismo.

20 Preferentemente, la secuencia de hibridación codifica una secuencia de aminoácidos que, cuando se usa en la construcción de un árbol filogenético, tal como el representado en la Figura 2, tiende a agruparse con el grupo de polipéptidos HpaG que comprende la secuencia de aminoácidos representada en la SEQ ID NO: 2, en lugar de con cualquier otro grupo.

25 La combinación de genes o evolución dirigida también puede usarse para generar variantes de ácidos nucleicos que codifican polipéptidos HpaG como se define anteriormente; siendo la expresión "combinación de genes" como se define en el presente documento.

30 De acuerdo con la presente invención, se proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas, que comprende introducir y expresar en una planta una variante de una cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A del Ejemplo 1, o que comprende introducir y expresar en una planta una variante de un ácido nucleico que codifica un ortólogo, parálogo u homólogo de cualquiera de las secuencias de ácidos nucleicos dadas en la Tabla A del Ejemplo 1, cuyo ácido nucleico variante se obtiene mediante combinación de genes.

35 Preferentemente, la secuencia de aminoácidos que codifica el ácido nucleico variante obtenido por combinación de genes, cuando se utiliza en la construcción de un árbol filogenético tal como el representado en la Figura 2, tiende a agruparse con el grupo de polipéptidos HpaG que comprende la secuencia de aminoácidos representada en la SEQ ID NO: 2, en lugar de con cualquier otro grupo.

40 Además, las variantes de ácidos nucleicos se pueden también obtener mediante mutagénesis dirigida. Hay disponibles varios procedimientos para conseguir la mutagénesis dirigida, siendo los más comunes los procedimientos basados en PCR (Current Protocols in Molecular Biology. Wiley Eds.).

45 Los ácidos nucleicos que codifican los polipéptidos HpaG se pueden obtener a partir de cualquier fuente natural o artificial. El ácido nucleico se puede modificar en composición y/o entorno genómico a partir de su forma nativa a través de la manipulación deliberada del ser humano. Preferentemente el ácido nucleico que codifica el polipéptido HpaG es de origen procariota, preferentemente de una bacteria Gram negativa que posee un SSTT, además preferentemente de una bacteria patógena vegetal que posee un SSTT, más preferentemente de la familia *Pseudomonaceae*, además preferentemente del género *Xanthomonas*, muy preferentemente el ácido nucleico es de *Xanthomonas axonopodis*.

50 La realización de los procedimientos de la invención proporciona plantas que tienen biomasa aumentada y/o rendimiento de semilla aumentado con respecto a las plantas de control. Las expresiones "rendimiento" y "rendimiento de semilla" se describen con más detalle en la sección "definiciones" en el presente documento.

55 En el presente documento la referencia a rasgos potenciados relacionados con el rendimiento se considera que significa un aumento en la biomasa (peso) de una o más partes de una planta, las que pueden incluir partes aéreas (cosechables) y/o partes por debajo del suelo (cosechables). En particular, tales partes cosechables son semillas, y la realización de los procedimientos de la invención da como resultado plantas que tienen rendimiento de semilla aumentado en comparación con el rendimiento de semilla de las plantas de control adecuadas.

Tomando al maíz como ejemplo, un aumento en el rendimiento puede manifestarse como uno o más de los siguientes: el aumento del número de plantas establecidas por hectárea o metro cuadrado, un aumento en el número de mazorcas por planta, un aumento en el número de hileras, el número de granos por hilera, el peso del

grano, el peso de mil granos, el longitud/diámetro de la mazorca, el aumento de la tasa de llenado de semilla (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), entre otros. Tomando el arroz como ejemplo, un incremento en el rendimiento puede manifestarse en sí mismo como un aumento en uno o más de los siguientes: el número de plantas por hectárea metro cuadrado, el número de panículas por planta, el número de espigas por panícula, el número de flores (floretes) por panícula (que se expresa como una proporción del número de semillas llenas sobre el número de panículas principales), el aumento en la tasa de llenado de semilla (que es el número de semillas llenas dividido entre el número total de semillas y multiplicado por 100), el aumento en el peso de mil granos, entre otros.

La presente invención proporciona un procedimiento para aumentar la biomasa y/o el rendimiento de semilla de las plantas, con respecto a las plantas de control, cuyo procedimiento comprende introducir y expresar, preferentemente aumentar la expresión, en una planta, de un ácido nucleico que codifica un polipéptido HpaG según se define en el presente documento. Debe indicarse que el aumento del rendimiento observado no es el resultado de la resistencia aumentada al estrés biótico.

Dado que las plantas transgénicas de acuerdo con la presente invención tienen rendimiento aumentado, es probable que estas plantas muestren una velocidad de crecimiento aumentada (durante por lo menos parte de su ciclo de vida), con respecto a la velocidad de crecimiento de las plantas de control en una etapa correspondiente de su ciclo de vida. Además de la capacidad de rendimiento aumentada, una eficiencia de captación de nutrientes aumentada puede también contribuir al aumento del rendimiento. Se observa que las plantas de acuerdo con la presente invención muestran una mayor eficacia en la captación de nutrientes. La eficacia aumentada de la captación de nutrientes permite el mejor crecimiento de la planta.

La tasa de crecimiento aumentada puede ser específica a una o más partes de una planta (incluyendo semillas), o puede ser sustancialmente en toda la planta. Las plantas que tienen una tasa de crecimiento aumentada pueden tener un ciclo de vida más corto. El ciclo de vida de una planta puede significar el tiempo necesario para crecer a partir de una semilla madura hasta la etapa en la que la planta ha producido semillas maduras, similar al material de partida. Este ciclo de vida puede estar influenciado por factores tales como el vigor temprano, tasa de crecimiento, índice de verdor, tiempo de floración y velocidad de la maduración de la semilla. El aumento en la tasa de crecimiento puede tener lugar en una o más etapas en el ciclo de vida de una planta o durante sustancialmente el ciclo de vida completo de una planta. Una tasa de crecimiento aumentada durante las etapas tempranas en el ciclo de vida de una planta puede reflejar vigor potenciado. El aumento en la tasa de crecimiento puede alterar el ciclo de la cosecha de una planta lo que permite que las plantas se siembren más tarde y/o se cosechen más pronto de lo que de otra manera sería posible (puede obtenerse un efecto similar con un tiempo de floración más temprano). Si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de semillas de la misma especie de la planta (por ejemplo siembra y cosecha de plantas de arroz seguido de siembra y cosecha de plantas de arroz adicionales todas dentro de un periodo de crecimiento convencional). De manera similar, si la tasa de crecimiento se aumenta suficientemente, puede permitir la siembra adicional de semillas de diferentes especies de plantas (por ejemplo, la siembra y cosecha de plantas de maíz seguido de, por ejemplo, la siembra y cosecha opcional de plantas de soja, patata o cualquier otra planta adecuada). Los tiempos de cosecha adicionales del mismo rizoma en el caso de algunas plantas de cultivo también pueden ser posibles. Alterar el ciclo de cosecha de una planta puede conducir a un aumento en producción de biomasa anual por metro cuadrado (debido a un aumento en el número de tiempos (dicho en un año) que se puede cultivar y cosechar cualquier planta particular). Un aumento en la tasa de crecimiento también puede permitir el cultivo de plantas transgénicas en un área geográfica más amplia que sus homólogos de tipo silvestre, debido a que las limitaciones territoriales para cosechar un cultivo a menudo se determinan por condiciones ambientales adversas bien en el momento de plantar (estación temprana) o bien en el momento de cosechar (estación tardía). Dichas condiciones adversas pueden evitarse si se acorta el ciclo de la cosecha. La tasa de crecimiento puede determinarse derivando diversos parámetros de curvas de crecimiento, dichos parámetros pueden ser: T-Medio (el tiempo que tardan las planta en alcanzar el 50 % de su tamaño máximo) y T-90 (el tiempo que tardan las plantas en alcanzar el 90 % de su tamaño máximo) entre otros.

De acuerdo con una característica preferida de la presente invención, la realización de los procedimientos de la invención proporciona plantas que tienen una tasa de crecimiento aumentada en comparación con plantas de control. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para aumentar la tasa de crecimiento en plantas, cuyo procedimiento comprende modular la expresión, preferentemente aumentar la expresión, en una planta de un ácido nucleico que codifica un polipéptido HpaG como se define en el presente documento. Debe indicarse que el aumento de la tasa de crecimiento observado no es el resultado de la resistencia al estrés biótico.

Un aumento en el rendimiento y/o tasa de crecimiento se produce si la planta está en condiciones sin estrés o si la planta está expuesta a diversos estreses abióticos en comparación con las plantas de control. Típicamente, las plantas responden a la exposición al estrés abiótico creciendo más lentamente. En condiciones de estrés intenso la planta puede incluso detener por completo el crecimiento. Por otro lado, el estrés leve se define en el presente documento como que es cualquier estrés al que se expone la planta que no da como resultado el cese total del crecimiento de la planta sin la capacidad de reanudar el crecimiento. El estrés leve, en el sentido de la invención, conduce a una reducción en el crecimiento de la planta estresada de menos del 40 %, 35 % o 30 %, preferentemente menos del 25 %, 20 % o 15 %, más preferentemente menos del 14 %, 13 %, 12 %, 11 % o 10 % o

menos en comparación con la planta de control en condiciones sin estrés. Debido a los avances en las prácticas agrícolas (riego, fertilización, tratamientos con pesticidas) las formas de estrés intenso no se encuentran frecuentemente en las plantas cultivadas. Como consecuencia, el crecimiento comprometido inducido por estrés leve es a menudo una característica indeseable para la agricultura. La expresión "estrés leve" son los estreses abióticos (ambientales) diarios a los cuales está expuesta la planta. El estrés abiótico puede deberse a sequía o a exceso de agua, a estrés anaeróbico, estrés salino, toxicidad química, estrés oxidativo y a temperaturas altas, frías o heladas. El estrés abiótico puede ser un estrés osmótico causado por estrés hídrico (particularmente debido a sequía), estrés salino, estrés oxidativo y estrés iónico.

La expresión "estrés abiótico" como se define en el presente documento, se entiende que se refiere a uno cualquiera o más de: estrés hídrico (debido a sequía o exceso de agua), estrés aneróbico, estrés salino, estrés térmico (debido a temperaturas elevadas, frías o de congelación), estrés por toxicidad química y estrés oxidativo. De acuerdo con un aspecto de la invención, el estrés abiótico es un estrés osmótico, seleccionado de estrés hídrico, estrés salino, estrés oxidativo y estrés iónico. Preferentemente, el estrés hídrico es estrés por sequía. La expresión estrés salino no se restringe a la sal común (NaCl), sino que puede ser una cualquiera o más de: NaCl, KCl, LiCl, MgCl<sub>2</sub>, CaCl<sub>2</sub>, entre otras.

Otro ejemplo de estrés ambiental abiótico es la disponibilidad reducida de uno o más de los nutrientes que las plantas necesitan asimilar para el crecimiento y el desarrollo. Debido a la fuerte influencia de la eficacia en la utilización de la nutrición sobre el rendimiento de la planta y la calidad del producto, se vierte una enorme cantidad de fertilizantes en los campos para optimizar el crecimiento de la planta y la calidad. Normalmente limitan la productividad de las plantas los tres nutrientes principales, fósforo, potasio y nitrógeno, el que entre estos tres es usualmente el elemento limitante de la tasa de crecimiento de la planta. Por lo tanto, el principal elemento nutricional necesario para el crecimiento de la planta es el nitrógeno (N). Es el constituyente de numerosos compuestos importantes encontrados en las células vivas, incluyendo aminoácidos, proteínas, (enzimas), ácidos nucleicos, y clorofila. Del 1,5 % y el 2 % de la materia seca vegetal es nitrógeno, y es aproximadamente el 16 % de la proteína total de la planta. Por lo tanto, la disponibilidad de nitrógeno es un factor limitante principal del crecimiento y la producción de la plantas de cultivo (Frink y col. (1999) Proc Natl Acad Sci USA 96(4): 1175-1180), y tiene también un impacto principal en la acumulación de proteína y en la composición de aminoácidos. Por lo tanto, son de gran interés las plantas de cultivo con un rendimiento aumentado cuando crecen en condiciones limitantes de nitrógeno.

Los estreses bióticos son normalmente los estreses que provocan patógenos, tales como bacterias, virus, hongos, nematodos e insectos.

En particular, los procedimientos de la presente invención pueden realizarse en condiciones sin estrés o en condiciones de sequía para proporcionar plantas que tienen rendimiento aumentado con respecto a las plantas de control. Como se informa en Wang y col. (Planta (2003) 218: 1-14), el estrés abiótico conduce a una serie de cambios morfológicos, fisiológicos, bioquímicos y moleculares que afectan de forma adversa el crecimiento y la productividad de la planta. Se sabe que la sequía, la salinidad, las temperaturas extremas y el estrés oxidativo están interconectados y pueden inducir daño en el crecimiento y daño celular a través de mecanismos similares. Rabbani y col. (Plant Physiol (2003) 133: 1755-1767) describen un particularmente alto grado de "interferencia" entre estrés por sequía y estrés por alta salinidad. Por ejemplo, la sequía y/o salinización se manifiestan principalmente como estrés osmótico, dando como resultado la alteración de la homeostasis y de la distribución iónica en la célula. El estrés oxidativo, que de forma frecuente acompaña a alta o baja temperatura, salinidad o estrés por sequía, puede provocar la desnaturalización de proteínas funcionales y estructurales. En consecuencia, estos diversos estreses medioambientales a menudo activan rutas de señalización celulares y respuestas celulares semejantes, tales como la producción de proteínas de estrés, la regulación positiva de antioxidantes, la acumulación de solutos compatibles y la detención del crecimiento.

La expresión condiciones "sin estrés", como se usa en el presente documento, son aquellas condiciones medioambientales que permiten el crecimiento óptimo de las plantas. Los expertos en la técnica saben cuáles son las condiciones normales en el suelo y las condiciones climáticas para cualquier localización determinada.

La realización de los procedimientos de la invención da lugar a plantas, crecidas en condiciones sin estrés o en condiciones de sequía, de rendimiento de semilla y/o biomasa aumentado con respecto a plantas de control adecuadas que crecen en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas que crecen en condiciones sin estrés o en condiciones de sequía, cuyo procedimiento comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido HpaG.

Además, la realización de los procedimientos de la invención da lugar a plantas que crecen en condiciones de déficit de nutrientes, particularmente en condiciones de déficit de nitrógeno, con rendimiento aumentado con respecto a plantas de control que se cultivan en condiciones comparables. Por lo tanto, de acuerdo con la presente invención, también se proporciona un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas crecidas en condiciones de déficit de nutrientes, cuyo procedimiento comprende aumentar la expresión en una planta de un ácido nucleico que codifica un polipéptido HpaG.

La realización de los procedimientos de la invención también proporciona plantas que tienen el vigor de la planta aumentado con respecto a las plantas de control, en particular durante las etapas tempranas del desarrollo de la planta (normalmente tres, cuatro semanas postgerminación en el caso del arroz y el maíz, pero esto variará de especie a especie) conduciendo al vigor temprano. Por lo tanto, de acuerdo con la presente invención, se proporciona un procedimiento para aumentar el vigor temprano de la planta, cuyo procedimiento comprende modular, preferentemente aumentar, la expresión en una planta de un ácido nucleico que codifica un polipéptido HpaG. Preferentemente el aumento en el vigor de la plántula se consigue mediante la expresión del ácido nucleico que codifica el polipéptido HpaG bajo el control de un promotor específico de brote. También se proporciona un procedimiento para la producción de plantas que tienen vigor temprano con respecto a las plantas de control, cuyo procedimiento comprende modular, preferentemente aumentar, la expresión en una planta de un ácido nucleico que codifica un polipéptido HpaG.

El vigor temprano también puede resultar de la eficacia biológica aumentada de la planta debida a que, por ejemplo, las plantas estén mejor adaptadas a su entorno (es decir que optimizan la utilización de los recursos de energía y la distribución entre brote y raíz). Las plantas que tienen vigor temprano también muestran supervivencia de las plántulas aumentada y un mejor establecimiento del cultivo, lo que a menudo da como resultado campos altamente uniformes (con el cultivo creciendo de una manera uniforme, es decir, con la mayoría de las plantas alcanzando las diversas etapas del desarrollo esencialmente al mismo tiempo), y con frecuencia mejor y mayor rendimiento. Por lo tanto, el vigor temprano puede determinarse midiendo diversos factores, tales como el peso de 1000 granos, la germinación porcentual, la emergencia porcentual, el crecimiento de la plántula, la altura de la plántula, la longitud radicular, la biomasa radicular y del brote y muchos más.

La presente invención abarca a plantas o partes de las mismas (incluyendo semillas) obtenibles por los procedimientos de acuerdo con la presente invención. Las plantas o partes de las mismas comprenden un transgén de ácido nucleico que codifica un polipéptido HpaG como se define anteriormente.

La invención también proporciona construcciones génicas y vectores para facilitar la introducción y/o expresión en plantas de ácidos nucleicos que codifican polipéptidos HpaG. Las construcciones génicas pueden insertarse en vectores, que pueden estar disponibles en el comercio, adecuados para la transformación en plantas y adecuados para la expresión del gen de interés en las células transformadas. La invención también proporciona el uso de una construcción génica como se define en el presente documento en los procedimientos de la invención.

Más específicamente, la presente invención proporciona una construcción que comprende:

- (a) un ácido nucleico que codifica un polipéptido HpaG como se ha definido anteriormente;
- (b) una o más secuencias de control capaces de conducir la expresión de la secuencia de ácidos nucleicos de (a); y opcionalmente
- (c) una secuencia de terminación de la transcripción.

Preferentemente, el ácido nucleico que codifica a HpaG es

- (i) un ácido nucleico como se representa en la SEQ ID NO: 1 o el complemento de la misma,
- (ii) un ácido nucleico que codifica un polipéptido HpaG como se define anteriormente.

La expresión "secuencia de control" y "secuencia de terminación" son como se define en el presente documento.

Las plantas se transforman con un vector que comprende cualquiera de los ácidos nucleicos descritos anteriormente. El experto en la técnica conocerá los elementos genéticos que deben estar presentes en un vector para transformar, seleccionar y propagar satisfactoriamente células hospedadoras que contengan la secuencia de interés. La secuencia de interés está unida operativamente a una o más secuencias de control (por lo menos a un promotor).

Ventajosamente, puede usarse para dirigir la expresión de la secuencia de ácidos nucleicos cualquier tipo de promotor, ya sea natural o sintético. Un promotor constitutivo, o un promotor específico de tejido verde, es particularmente útil en los procedimientos. Véase la sección "Definiciones" del presente documento para las definiciones de los diversos tipos de promotores.

Preferentemente, el ácido nucleico de HpaG o variante del mismo, está unido operativamente a un promotor constitutivo. Un promotor constitutivo preferente es uno que se expresa también de forma esencialmente ubicua. Además preferentemente el promotor se obtiene a partir de una planta, más preferentemente de una planta monocotiledónea. Es muy preferente el uso de un promotor GOS2 (de arroz) (SEQ ID NO: 5). Debería estar claro que la aplicabilidad de la presente invención no se limita al ácido nucleico de *HpaG* representado en la SEQ ID NO: 1, ni está la aplicabilidad de la invención limitada a la expresión de un ácido nucleico de *HpaG* cuando la dirige un promotor GOS2. Los ejemplos de otros promotores constitutivos que también pueden utilizarse para dirigir la expresión de un ácido nucleico de *HpaG* se muestran en la Tabla 2a en la sección de Definiciones en el presente documento.

Preferentemente, el promotor constitutivo es de fuerza media y tiene actividad más débil que el promotor CaMV 35S.

Como alternativa, el ácido nucleico de *HpaG*, o la variante del mismo, está unido operativamente a un promotor específico de tejido verde. Un promotor específico de tejido verde como se define en el presente documento es un promotor que está transcripcionalmente activo de forma predominante en tejido verde, esencialmente con la exclusión de cualquier otra parte de una planta, mientras que todavía permite cualquier expresión parcial en estas otras partes de la planta. El promotor específico de tejido verde es preferentemente un promotor de la protoclorofilida reductasa, más preferentemente el promotor de la protoclorofilida reductasa representado por una secuencia de ácido nucleico esencialmente similar a la SEQ ID NO: 6, muy preferentemente el promotor es como se representa en la SEQ ID NO: 6. Debería estar claro que la aplicabilidad de la presente invención no se limita al ácido nucleico que codifica *HpaG* representado en la SEQ ID NO: 1, ni está la aplicabilidad de la invención limitada a la expresión de tal ácido nucleico que codifica *HpaG* cuando la dirige un promotor de la protoclorofilida reductasa. Los ejemplos de otros promotores específicos de tejido verde que también pueden utilizarse para realizar los procedimientos de la invención se muestran en la sección de definiciones en el presente documento.

Para la identificación de promotores funcionalmente equivalentes, la fuerza del promotor y/o el patrón de expresión de un promotor candidato pueden analizarse, por ejemplo, uniendo operativamente el promotor a un gen indicador y evaluando el nivel de expresión y el patrón del gen indicador en diversos tejidos de la planta. Los genes indicadores bien conocidos adecuados incluyen, por ejemplo, la beta glucuronidasa o la beta galactosidasa. La actividad del promotor se evalúa midiendo la actividad enzimática de la beta glucuronidasa o beta galactosidasa. La fuerza del promotor y/o el patrón de expresión pueden después compararse con la de un promotor de referencia (tal como el que se usa en los procedimientos de la presente invención). Como alternativa, la fuerza del promotor puede evaluarse cuantificando los niveles de ARNm o comparando los niveles de ARNm del ácido nucleico utilizado en los procedimientos de la presente invención, con los niveles de ARNm de genes constitutivos tales como ARNr 18S, utilizando procedimientos conocidos en la técnica, tales como transferencia de Northern con análisis densitométrico de autorradiogramas, PCR cuantitativa en tiempo real o RT-PCR (Heid y col., 1996 Genome Methods 6: 986-994). En general por "promotor débil" se refiere a un promotor que dirige la expresión de una secuencia codificante a un nivel bajo. Por "nivel bajo" se entiende a niveles de aproximadamente 1/10.000 transcritos a aproximadamente 1/100.000 transcritos, a aproximadamente 1/500.000 transcritos por célula. Por el contrario, un "promotor fuerte" dirige la expresión de una secuencia codificante a alto nivel, o a aproximadamente 1/10 transcritos a aproximadamente 1/100 transcritos a aproximadamente 1/1.000 transcritos por célula.

De forma opcional, en la construcción introducida en una planta se pueden utilizar una o más secuencias terminadoras. Los elementos reguladores adicionales pueden incluir potenciadores de la transcripción así como de la traducción. Los expertos en la materia conocerán las secuencias terminadoras o potenciadoras que pueden ser adecuadas para su uso en la realización de la invención. Un experto en la materia conocería tales secuencias, o podría obtenerlas fácilmente.

También se puede añadir una secuencia de intrón a la región 5' no traducida (UTR) o en la secuencia codificante para aumentar la cantidad de mensajero maduro que se acumula en el citosol. La inclusión de un intrón que se puede cortar y empalmar en la unidad de transcripción en las construcciones de expresión tanto vegetal como animal, ha demostrado aumentar la expresión génica a nivel de ARNm y de proteína hasta 1000 veces (Buchman y Berg, Mol Cell Biol 8: 4395-4405 (1988); Callis y col, genes Dev. 1: 1183-1200 (1987)). Tal potenciación por intrón de la expresión génica es normalmente mayor cuando se pone cerca del extremo 5' de la unidad de transcripción. El uso de los intrones de maíz *Adh1-S* intrones 1, 2, y 6, el intrón *Bronze-1*, son conocidos en la técnica. Para información general, véase *The Maize Handbook*, Capítulo 116, Freeling and Walbot, Eds., Springer, N. Y. (1994).

Otras secuencias de control (además del promotor, potenciador, silenciador, secuencias de intrón, regiones 3' UTR y/o 5' UTR) pueden ser elementos que estabilizan proteínas y/o ARN. Un experto en la materia conocería tales secuencias o podría obtenerlas fácilmente. Además, el uso de codones de la secuencia codificante a insertar en la construcción puede optimizarse con referencia a la célula hospedadora en la cual se introducirá la construcción. Aunque el código genético es degenerado, los organismos tienden a utilizar un codón en particular para un aminoácido más que otros codones para ese mismo aminoácido. Se conocen en la técnica las tablas con el uso preferente de codones para diversos organismos.

Las construcciones genéticas de la invención pueden incluir adicionalmente un origen de secuencia de replicación que es necesario para la conservación y/o replicación en un tipo de célula específico. Un ejemplo es cuando en una célula bacteriana se requiere conservar una construcción genética como un elemento genético episomal (por ejemplo, una molécula de plásmido o de cósmido). Como orígenes de replicación preferidos se incluyen, pero sin limitación *f1-ori* y *colE1*.

Para la detección de la transferencia satisfactoria de las secuencias de ácidos nucleicos como se usa en los procedimientos de la invención y/o selección de plantas transgénicas que comprenden estos ácidos nucleicos, es ventajoso usar genes marcadores (o genes indicadores). Por lo tanto, la construcción genética puede comprender opcionalmente un gen marcador de selección. Los marcadores de selección se describen con más detalle en la sección "definiciones" del presente documento.

Se sabe que, después de la integración estable o transitoria de los ácidos nucleicos en células vegetales, solo una minoría de las células capta el ADN exógeno y, si se desea, lo integra en su genoma, dependiendo del vector de

expresión usado y de la técnica de transfección usada. Para identificar y seleccionar estos integrantes, normalmente se introduce un gen que codifica un marcador de selección (tal como los descritos anteriormente) en las células huésped junto con el gen de interés. Estos marcadores pueden usarse, por ejemplo, en mutantes en los que estos genes no son funcionales mediante, por ejemplo, delección por procedimientos convencionales. Además, las moléculas de ácido nucleico que codifican un marcador de selección pueden introducirse en una célula hospedadora en el mismo vector que comprende la secuencia que codifica los polipéptidos de la invención o usarse en los procedimientos de la invención, o incluso en un vector distinto. Las células que se han transfectedo de manera estable con el ácido nucleico introducido pueden identificarse, por ejemplo, por selección (por ejemplo, células que tienen integrado el marcador de selección sobreviven mientras que las otras células mueren).

Dado que los genes marcadores, en particular los genes de resistencia a antibióticos y herbicidas, ya no se necesitan o no se desean en la célula hospedadora transgénica una vez que se han introducido satisfactoriamente los ácidos nucleicos, el procedimiento de acuerdo con la invención para introducir los ácidos nucleicos emplea ventajosamente técnicas que permiten la retirada o escisión de estos genes marcadores. Un procedimiento de este tipo es lo que se conoce como cotransformación. El procedimiento de cotransformación emplea para la transformación dos vectores de forma simultánea, un vector que lleva el ácido nucleico de acuerdo con la invención y un segundo que lleva el gen (o genes) marcador. Una proporción grande de transformantes recibe o, en el caso de plantas, comprende (hasta el 40 % o más de los transformantes), ambos vectores. En el caso de transformación con *Agrobacterium*, los transformantes normalmente solo reciben una parte del vector, es decir, la secuencia flanqueada por el ADN-T, que normalmente representa el casete de expresión. Posteriormente los genes marcadores pueden retirarse de la planta transformada realizando cruzamientos. En otro procedimiento, para la transformación junto con el ácido nucleico deseado se usan genes marcadores integrados en un transposón (conocido como tecnología Ac/Ds). Los transformantes pueden cruzarse con una fuente transposasa o los transformantes se transforman con una construcción de ácido nucleico que confiere expresión de una transposasa, de manera transitoria o estable. En algunos casos (aprox. 10 %), el transposón salta del genoma de la célula hospedadora una vez que se ha producido la transformación satisfactoriamente y se pierde. En una serie de casos adicionales, el transposón salta a una localización diferente. En estos casos, el gen marcador debe eliminarse realizando cruzamientos. En microbiología, se desarrollaron técnicas que hacen posible, o facilitan, la detección de tales acontecimientos. Un procedimiento ventajoso adicional se basa en lo que se conoce como sistemas de recombinación; cuya ventaja es que la eliminación puede dispensarse por cruzamiento. El sistema de este tipo mejor conocido es lo que se conoce como sistema Cre/lox. Cre1 es una recombinasa que retira las secuencias localizadas entre las secuencias loxP. Si el gen marcador está integrado entre las secuencias loxP, éste se retira una vez se haya producido la transformación satisfactoriamente, por expresión de la recombinasa. Otros sistemas de recombinación son el sistema HIN/HIX, FLP/FRT y REP/STB (Tribble y col., *J. Biol. Chem.*, 275, 2000: 22255-22267; Velmurugan y col., *J. Cell Biol.*, 149, 2000: 553-566). De acuerdo con la invención, es posible una integración específica de sitio de las secuencias de ácidos nucleicos en el genoma de la planta. Naturalmente, estos procedimientos también pueden aplicarse a microorganismos tales como levaduras, hongos o bacterias.

La invención también proporciona un procedimiento para la producción de plantas transgénicas que tienen rendimiento de semilla y/o biomasa aumentados con respecto a las plantas de control, que comprende la introducción y expresión en una planta de cualquier ácido nucleico que codifica un polipéptido HpaG según se define anteriormente en el presente documento.

De forma más específica, la presente invención proporciona un procedimiento para la producción de plantas transgénicas que tienen la biomasa y/o el rendimiento de semilla aumentados, comprendiendo el procedimiento:

- (i) introducir y expresar en una planta o célula vegetal un ácido nucleico que codifica el polipéptido HpaG; y
- (ii) cultivar la célula vegetal en condiciones sin estrés o en condiciones de estrés abiótico, en el que dicho rendimiento aumentado no es el resultado de la resistencia aumentada al estrés biótico.

El ácido nucleico de (i) puede ser cualquiera de los ácidos nucleicos capaces de codificar un polipéptido HpaG como se define en el presente documento.

El ácido nucleico puede introducirse directamente en una célula de planta o en la propia planta (incluyendo la introducción en un tejido, órgano o en cualquier otra parte de una planta). De acuerdo con una característica preferida de la presente invención, el ácido nucleico se introduce preferentemente en una planta por transformación. El término "transformación" se describe con más detalle en la sección "definiciones" del presente documento.

Las células de planta modificadas genéticamente pueden regenerarse mediante todos los procedimientos con los que está familiarizado el experto en la técnica. Pueden encontrarse procedimientos adecuados en las publicaciones anteriormente mencionadas de S. D. Kung y R. Wu, Potrykus o Höfgen y Willmitzer.

Generalmente, después de la transformación, las células o grupos de células de planta se seleccionan para detectar la presencia de uno o más marcadores que están codificados por genes que pueden expresarse en plantas cotransferidos con el gen de interés, después de lo cual el material transformado se regenera en una planta completa. Para seleccionar plantas transformadas, el material de planta obtenido en la transformación, como norma, se somete a condiciones selectivas de tal manera que las plantas transformadas pueden diferenciarse de las plantas



no transformadas. Por ejemplo, las señales obtenidas de la manera descrita anteriormente pueden plantarse y, después de un periodo de crecimiento inicial, someterse a una selección adecuada por pulverización. Una posibilidad adicional consiste en cultivar las semillas, si fuera apropiado, después de esterilización, en placas de agar usando un agente de selección adecuado de tal manera que solamente las semillas transformadas puedan desarrollarse en plantas. Como alternativa, las plantas transformadas se exploran para determinar la presencia de un marcador de selección tal como el descrito anteriormente.

Después de la transferencia y regeneración de ADN, las plantas supuestamente transformadas pueden evaluarse utilizando, por ejemplo, análisis de Southern, para determinar la presencia del gen de interés, número de copias y/u organización genómica. Como alternativa o adicionalmente, los niveles de expresión del ADN recién introducido pueden supervisarse usando análisis de Northern y/o Western, ambas técnicas bien conocidas por los expertos habituales en la materia.

Las plantas transformadas generadas pueden propagarse mediante diversos medios, tales como mediante técnicas clásicas de propagación clonal o de reproducción. Por ejemplo, una primera generación (o T1) de planta transformada puede reproducirse asexualmente y puede seleccionarse una segunda generación homocigota (o T2) y después las plantas T2 pueden propagarse adicionalmente a través de técnicas de reproducción clásicas.

Los organismos transformados generados pueden adoptar diversas formas. Por ejemplo, pueden ser quimeras de células transformadas y de células no transformadas; transformantes clonales (por ejemplo, todas las células transformadas que contengan el casete de expresión); injertos de tejidos transformados y no transformados (por ejemplo, en plantas, una planta madre transformada injertada a un vástago no transformado).

La presente invención se amplía claramente a cualquier célula de planta o planta producida mediante cualquiera de los procedimientos de la divulgación descritos en el presente documento, y a todas las partes de la planta y propágulos de la misma. La presente invención se amplía adicionalmente para incluir la progenie de una célula, tejido u órgano primario transformado o transfectado o la planta completa que se ha producido mediante cualquiera de los procedimientos anteriormente mencionados, siendo el único requisito que la progenie presente las mismas características, o características, genotípicas y/o fenotípicas que las producidas por el precursor en los procedimientos de acuerdo con la invención.

La invención también incluye células huésped que contienen un ácido nucleico aislado que codifica un polipéptido HpaG como se ha definido anteriormente en el presente documento. De acuerdo con la invención, las células huésped preferentes son células de plantas. Las plantas huésped para los ácidos nucleicos o el vector usados en el procedimiento de acuerdo con la invención, el casete de expresión o construcción o vector son, en principio, ventajosamente todas las plantas, que pueden sintetizar los polipéptidos usados en el procedimiento inventivo.

Los procedimientos de la invención son ventajosamente aplicables a cualquier planta.

Las plantas que son particularmente útiles en los procedimientos de la invención incluyen todas las plantas que pertenecen a la superfamilia *Viridiplantae*, en particular plantas monocotiledóneas y dicotiledóneas incluyendo forraje o leguminosas forrajeras, plantas ornamentales, cultivos alimenticios, árboles o arbustos. De acuerdo con una realización preferida de la presente invención, la planta es una planta de cultivo. Los ejemplos de plantas de cultivo incluyen soja, girasol, canola, alfalfa, colza, algodón, tomate, patata y tabaco. Adicionalmente de modo preferente, la planta es una planta monocotiledónea. Ejemplos de plantas monocotiledóneas incluyen caña de azúcar. Más preferentemente la planta es un cereal. Ejemplos de cereales incluyen arroz, maíz, trigo, cebada, mijo, triticale, centeno, sorgo y avena.

La invención también se extiende a las partes cosechables de una planta tales como, pero sin limitación, semillas, hojas, frutos, flores, tallos, rizomas, tubérculos y bulbos. También se describen en el presente documento los productos obtenidos, preferentemente obtenidos de forma directa, de una parte cosechable de tal planta, tal como gránulos o polvos secos, aceite, grasa, y ácidos grasos, almidón o proteínas.

De acuerdo con la invención, la expresión es expresión aumentada. Los procedimientos para el aumento de la expresión de ácidos nucleicos o genes, o de productos génicos, están bien documentados en la técnica e incluyen, por ejemplo, la sobreexpresión dirigida mediante promotores apropiados, el uso de potenciadores de la transcripción o de potenciadores de la traducción. Se pueden introducir en una posición apropiada (normalmente cadena arriba) ácidos nucleicos aislados que sirven como elementos promotores o potenciadores de una forma no heteróloga de un polinucleótido, de forma tal que se regule de forma positiva la expresión. Por ejemplo, los promotores endógenos pueden alterarse *in vivo* mediante mutación, delección, y/o sustitución (véase, Kmiec, Patente de Estados Unidos n.º 5.565.350; Zarling y col, documento PCT/US93/03868), o se pueden introducir promotores aislados en una célula vegetal en la orientación y distancia apropiadas con respecto a un gen de la presente invención de forma que controlen la expresión del gen.

Si se desea la expresión del polipéptido, en general es conveniente incluir una región de poliadenilación en el extremo 3' de la región codificante de un polinucleótido. La región de poliadenilación se puede obtener del gen natural, a partir de una diversidad de otros genes vegetales, o a partir de ADN-T. La secuencia del extremo 3' a añadir se puede obtener de, por ejemplo, los genes de la nopalina sintasa o de la octopina sintasa, o como

alternativa a partir de otros genes vegetales, o menos preferentemente a partir de cualquier otro gen eucariota.

La presente invención también abarca el uso de ácidos nucleicos que codifican polipéptidos HpaG como se describe en el presente documento y el uso de estos polipéptidos HpaG en el aumento del rendimiento de semilla y/o de la biomasa en plantas.

- 5 Los procedimientos de acuerdo con la presente invención dan como resultado plantas que tienen el rendimiento de semilla y/o la biomasa aumentados, como se ha descrito anteriormente en el presente documento. Estos rasgos también pueden combinarse con otros rasgos económicamente ventajosos, tales como rasgos que potencian adicionalmente el rendimiento, la tolerancia a diversos estreses abióticos y bióticos, rasgos que modifican diversas características arquitectónicas y/o características bioquímicas y/o fisiológicas.

## 10 Descripción de las figuras

La presente invención se describirá ahora con referencia a las siguientes figuras, en las que:

La **Figura 1** muestra un alineamiento de los polipéptidos HpaG con los motivos 1 y 2 indicados en negrita y subrayados para la SEQ ID NO: 2.

- 15 La **Figura 2** muestra un árbol filogenético con el grupo de polipéptidos HpaG distinguidos de otras proteínas bacterianas y de proteínas vegetales (las diversas secuencias se indican mediante sus números de referencia GenBank y/o números gi).

La **Figura 3** muestra el vector binario para la expresión aumentada en *Oryza sativa* de un ácido nucleico que codifica la proteína HpaG de *Xanthomonas* bajo el control de un promotor GOS2 (pGOS2) de arroz.

- 20 La **Figura 4** detalla los ejemplos de secuencias Harpin útiles en la realización de los procedimientos de acuerdo con la presente invención.

## Ejemplos

La presente invención se describirá ahora con referencia a los siguientes ejemplos, que son solo a modo ilustrativo. Los siguientes ejemplos no pretenden definir completamente o de otra manera limitar el alcance de la invención.

### **Ejemplo 1: Identificación de secuencias HpaG**

- 25 Las secuencias (ADNc de longitud completa, EST o genómicas) relacionadas con la SEQ ID NO: 1 y/o las secuencias de proteínas relacionados con la SEQ ID NO: 2 se identificaron entre las conservadas en la base de datos de Nucleótidos Entrez en el National Center for Biotechnology Information (NCBI) usando herramientas de búsqueda de secuencias de bases de datos, tales como la Herramienta Básica de Alineamiento Local (BLAST, *Basic Local Alignment Tool*) (Altschul y col. (1990) J. Mol. Biol. 215:403-410; and Altschul y col. (1997) Nucleic Acids Res. 25:3389-3402). La búsqueda de secuencias también se realizó en otras bases de datos públicas y de patente. Se usaron herramientas de búsqueda de secuencia tales como la Herramienta de Alineamiento Local Básica (BLAST, *Basic Local Alignment Tool*) (Altschul y col. (1990) J. Mol. Biol. 215: 403-410; y Altschul y col. (1997) Nucleic Acids Res. 25: 3389-3402). El programa se usó para encontrar regiones de similitud local entre las secuencias comparando las secuencias de ácidos nucleicos o polipéptidos con las bases de datos de secuencia y calculando el significado estadístico de los emparejamientos. El polipéptido que codifica la SEQ ID NO: 1 se usó para el algoritmo TBLASTN, con configuraciones predeterminadas y el filtro para ignorar la activación de secuencias de baja complejidad. El resultado de los análisis se observó por comparación por pares, y se clasificó de acuerdo con la puntuación de probabilidad (valor E), en el que la puntuación refleja la probabilidad de que se produzca un alineamiento particular por casualidad (cuanto menor sea el valor E más significativo es el acierto). Además de los valores E, las comparaciones también se puntuaron por porcentaje de identidad. El porcentaje de identidad se refiere al número de nucleótidos (o aminoácidos) idénticos entre las dos secuencias de ácidos nucleicos (o de polipéptidos) comparadas sobre una longitud particular. En algunos casos, los parámetros predeterminados se pueden ajustar para modificar la rigurosidad de la búsqueda. La Tabla A proporciona un listado de secuencias de ácidos nucleicos y proteínas relacionados con la secuencia de ácidos nucleicos como se representa en la SEQ ID NO: 1 y la secuencia de proteína representada en la SEQ ID NO: 2.

**Tabla A:** secuencias de ácidos nucleicos que codifican HpaG y polipéptidos HpaG útiles en los procedimientos de la presente invención.

| Nombre    | Organismo fuente              | Ácido nucleico SEQ ID NO: | Polipéptido SEQ ID NO: | Estado            |
|-----------|-------------------------------|---------------------------|------------------------|-------------------|
| HpaG      | <i>Xanthomonas axonopodis</i> | 1                         | 2                      | Longitud completa |
| HpaG_T44C | Construcción sintética        | 7                         | 8                      | Longitud completa |

(continuación)

| Nombre  | Organismo fuente  | Ácido nucleico SEQ ID NO: | Polipéptido SEQ ID NO: | Estado            |
|---|---|---------------------------|------------------------|-------------------|
| HpaG-T  | Construcción sintética  | 9                         | 10                     | Longitud completa |
| Hpa1  | <i>Xanthomonas axonopodis</i> pv. citri cep. 306                    | 11                        | 12                     | Longitud completa |
| HpaG-N  | Construcción sintética  | 13                        | 14                     | Longitud completa |
| HpaG_G  | <i>Xanthomonas axonopodis</i>                                       | 15                        | 16                     | Longitud completa |
| Hrp   | <i>Xanthomonas smithii</i> subsp. <i>smithii</i>                    | 17                        | 18                     | Longitud completa |
| factor A de funcionamiento de respuesta hipersensible | <i>Xanthomonas oryzae</i> pv. <i>oryzae</i> cepa JXOIII             | 19                        | 20                     | Longitud completa |
| Hpa1  | <i>Xanthomonas oryzae</i> pv. <i>oryzae</i>                         | 21                        | 22                     | Longitud completa |
| Hpa1  | <i>Xanthomonas oryzae</i> pv. <i>oryzae</i>                         | 23                        | 24                     | Longitud completa |
| hpaGXooc  | <i>Xanthomonas oryzae</i> pv. <i>oryzicola</i>                      | 25                        | 26                     | Longitud completa |
| Hpa1  | <i>Xanthomonas campestris</i> pv. <i>campestris</i> cep. ATCC 33913 | 27                        | 28                     | Longitud completa |

**Ejemplo 2: Alineamiento de las secuencias del polipéptido HpaG**

5 El alineamiento de las secuencias de polipéptidos (Figura 1) se realizó utilizando el programa ClustalW que es a base del popular algoritmo de alineamiento progresivo Clustal (Thompson y col. (1997) Nucleic Acids Res 25:4876-4882; Chenna y col. (2003). Nucleic Acids Res 31:3497-3500). Los valores predeterminados son: para la penalización de apertura de hueco de 10, para la penalización de la extensión de hueco de 0,1 y la matriz de peso seleccionada es Blosum 62 (si se alinean polipéptidos). Para optimizar adicionalmente el alineamiento se realizó edición manual menor.

10 Utilizando un algoritmo de agrupamiento con el vecino más cercano, como se proporciona en el programa AlignX del Vector NTI (Invitrogen), se construyó un árbol filogenético de polipéptidos HpaG (Figura 2).

**Ejemplo 3: Cálculo del porcentaje de identidad global entre las secuencias de polipéptidos útiles en la realización de los procedimientos de la invención**

15 Los porcentajes de similitud e identidad globales entre las secuencias de polipéptidos de longitud completa útiles en la realización de los procedimientos de la invención se determinaron usando uno de los procedimientos disponibles en la técnica, el programa informático MatGAT (Matrix Global Alignment Tool) (Campanella et al., BMC Bioinformatics. 2003 4: 29. MatGAT: una aplicación que genera matrices de similitud/identidad utilizando secuencias de proteína o de ADN.). El programa informático MatGAT genera matrices de similitud/identidad para secuencias de ADN o proteína sin la necesidad de alinear previamente los datos. El programa realiza una serie de alineamientos por pares usando el algoritmo de alineamiento global de Myers y Miller (con una penalización por apertura de hueco de 12, y una penalización por extensión de hueco de 2), calcula la similitud e identidad usando, por ejemplo, Blosum 62 (para polipéptidos), y después se pusieron los resultados en una matriz de distancia. La similitud de secuencia se muestra en la mitad inferior de la línea divisoria y la identidad de secuencia se muestra en la mitad superior de la línea divisoria diagonal.

25 Los parámetros utilizados en la comparación fueron:

**Matriz de puntuación: Blosum62**

Hueco inicial: 12

Extensión de hueco: 2

Los resultados del análisis del programa informático para la similitud e identidad globales a lo largo de la longitud completa de las secuencias de los polipéptidos (que excluyen las secuencias de polipéptidos parciales) se muestran en la Tabla B. La identidad porcentual se proporciona en negrita por encima de la diagonal y la similitud porcentual se proporciona por debajo de la diagonal (caracteres normales).

- 5 La identidad porcentual entre las secuencias de polipéptidos HpaG útiles en la realización de los procedimientos de la invención, puede ser tan baja como el 37 % de identidad de aminoácidos en comparación con la SEQ ID NO: 9.

**Tabla B:** resultados de MatGAT para la similitud e identidad globales a lo largo de la longitud completa de las secuencias de los polipéptidos.

|                        | 1    | 2           | 3           | 4           | 5           | 6           | 7           | 8           | 9            | 10           | 11          | 12          |
|------------------------|------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|-------------|-------------|
| <b>1. SEQ ID NO: 2</b> |      | <b>99,2</b> | <b>94,0</b> | <b>91,2</b> | <b>91,0</b> | <b>90,2</b> | <b>85,4</b> | <b>66,7</b> | <b>66,7</b>  | <b>66,7</b>  | <b>59,6</b> | <b>37,7</b> |
| <b>2. ABK51589</b>     | 99,2 |             | <b>93,2</b> | <b>90,5</b> | <b>90,2</b> | <b>89,5</b> | <b>84,7</b> | <b>67,4</b> | <b>67,4</b>  | <b>67,4</b>  | <b>60,3</b> | <b>37,7</b> |
| <b>3. ABK51587</b>     | 94,0 | 93,2        |             | <b>85,4</b> | <b>85,0</b> | <b>92,0</b> | <b>79,6</b> | <b>60,3</b> | <b>60,3</b>  | <b>60,3</b>  | <b>56,4</b> | <b>33,3</b> |
| <b>4. AAM35307</b>     | 92,0 | 91,2        | 86,1        |             | <b>82,5</b> | <b>81,8</b> | <b>89,8</b> | <b>70,9</b> | <b>70,9</b>  | <b>70,9</b>  | <b>61,4</b> | <b>36,6</b> |
| <b>5. ABK51590</b>     | 91,0 | 90,2        | 90,4        | 83,2        |             | <b>81,2</b> | <b>76,6</b> | <b>57,4</b> | <b>57,4</b>  | <b>57,4</b>  | <b>50,7</b> | <b>32,8</b> |
| <b>6. ABK51588</b>     | 90,2 | 89,5        | 92,0        | 82,5        | 89,3        |             | 75,2        | <b>58,2</b> | <b>58,2</b>  | <b>58,2</b>  | <b>56,4</b> | <b>33,8</b> |
| <b>7. ABG36696</b>     | 89,5 | 88,7        | 83,5        | 92,7        | 80,5        | 79,7        |             | <b>70,7</b> | <b>70,7</b>  | <b>70,7</b>  | <b>58,8</b> | <b>37,0</b> |
| <b>8. ABJ97680</b>     | 77,0 | 77,7        | 70,5        | 80,6        | 67,6        | 68,3        | 81,3        |             | <b>100,0</b> | <b>100,0</b> | <b>64,5</b> | <b>35,0</b> |
| <b>9. AAC95121</b>     | 77,0 | 77,7        | 70,5        | 80,6        | 67,6        | 68,3        | 81,3        | 100,0       |              | <b>100,0</b> | <b>64,5</b> | <b>35,0</b> |
| <b>10. BAD29979</b>    | 77,0 | 77,7        | 70,5        | 80,6        | 67,6        | 68,3        | 81,3        | 100,0       | 100,0        |              | <b>64,5</b> | <b>35,0</b> |
| <b>11. ABB72197</b>    | 72,9 | 73,7        | 72,8        | 73,7        | 68,0        | 72,8        | 72,9        | 72,7        | 72,7         | 72,7         |             | <b>34,6</b> |
| <b>12. AAM40538</b>    | 51,9 | 51,9        | 48,0        | 49,6        | 46,3        | 50,4        | 50,4        | 45,3        | 45,3         | 45,3         | 53,6        |             |

10 **Ejemplo 4: Clonación y construcción del vector**

A menos que se manifieste de otro modo, las técnicas de ADN recombinante se realizaron de acuerdo a protocolos estándar descritos en (Sambrook (2001) Molecular Cloning: a laboratory manual, 3<sup>a</sup> Edición Cold Spring Harbor Laboratory Press, CSH, Nueva York) o en los Volúmenes 1 y 2 de Ausubel y col. (1994), Current Protocols in Molecular Biology, Current Protocols. Los materiales y procedimientos estándares para el trabajo molecular de plantas se describen Plant Molecular Biology Labfax (1993) por R.D.D. Croy, publicado por BIOS Scientific Publications Ltd (RU) y Blackwell Scientific Publications (RU).

15 La secuencia codificante de HpaG de *Xanthomonas* se amplificó mediante PCR a partir de una biblioteca de ADN de *Xanthomonas axonopodis*. Se purificó el fragmento de PCR de la longitud esperada y posteriormente se clonó en un vector Gateway® utilizando tecnología convencional. Después se utilizó en una reacción LR el clon de entrada que comprendía la SEQ ID NO: 1 con un vector de destino utilizado para la transformación de *Oryza sativa*. Este vector contenía como elementos funcionales dentro de los límites del ADN-T: un marcador de selección de plantas; un casete de expresión del marcador de selección; y un casete Gateway destinado a la recombinación LR *in vivo* con la secuencia de ácido nucleico de interés ya clonada en el vector de entrada. Se colocó cadena arriba de este casete Gateway un promotor GOS2 (SEQ ID NO: 5) de arroz para la expresión constitutiva. Como alternativa, mostró ser igualmente útil un promotor específico de tejido verde, tal como el promotor de la protoclороfilida reductasa (SEQ ID NO: 6).

Después de la etapa de recombinación LR, el vector de expresión resultante pGOS2::HpaG se transformó en la cepa LBA4044 de *Agrobacterium* de acuerdo con los procedimientos bien conocidos en la técnica.

**Ejemplo 5: Transformación de plantas**

30 **Transformación de arroz**

El *Agrobacterium* que contenía el vector de expresión se usó para transformar plantas de *Oryza sativa*. Se descascarillaron semillas secas maduras de la variedad de cultivo japónica de arroz Nipponbare. La esterilización se realizó incubando durante un minuto en etanol al 70 %, seguido de 30 minutos en HgCl<sub>2</sub> al 0,2 %, seguido de un

lavado de 6 veces durante 15 minutos con agua destilada estéril. Las semillas estériles germinaron después en un medio que contenía 2,4-D (medio de inducción de callo). Después de incubación en la oscuridad durante cuatro semanas, los callos embrionarios, derivados de escutelo, se cortaron y se propagaron en el mismo medio. Después de 2 semanas, los callos se multiplicaron o se propagaron mediante subcultivo en el mismo medio durante otras 2 semanas. Trozos de callos embriogénicos se subcultivaron en medio reciente 3 días antes del cocultivo (para reforzar la actividad de la división celular).

Para el cocultivo se usó la cepa LBA4404 de *Agrobacterium* que contenía el vector de expresión. Se inoculó *Agrobacterium* en medio AB con los anticuerpos apropiados y se cultivó durante 3 días a 28 °C. Después, las bacterias se recogieron y se suspendieron en medio de cocultivo líquido a una densidad (DO<sub>600</sub>) de aproximadamente 1. Después, la suspensión se transfirió a una placa de Petri y los callos se sumergieron en la suspensión durante 15 minutos. Después, los tejidos de callo se secaron por transferencia en un papel de filtro y se transfirieron a medio de cocultivo solidificado y se incubaron durante 3 días en la oscuridad a 25 °C. Los callos cocultivados crecieron sobre un medio que contenía 2,4-D durante 4 semanas en la oscuridad a 28 °C en presencia de un agente de selección. Durante este periodo, rápidamente se desarrollaron islas de callos resistentes a crecimiento. Después de la transferencia de este material a un medio de regeneración e incubación a la luz, el potencial embriogénico se liberó y se desarrollaron brotes las siguientes cuatro a cinco semanas. Los brotes se extirparon de los callos y se incubaron durante 2 a 3 semanas en un medio que contenía auxina desde el cual se transfirieron al suelo. Los brotes endurecidos se cultivaron a alta humedad y en días cortos en un invernadero.

Se generaron aproximadamente 35 transformantes de arroz T0 independientes para una construcción. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero. Después de un análisis de PCR cuantitativa para verificar el número de copias del inserto de ADN-T, solo una única copia de plantas transgénicas que presentaban tolerancia al agente de selección, se guardaron para la cosecha de la semilla T1. Después, las semillas se cosecharon de tres a cinco meses después del trasplante. El procedimiento produjo transformantes de locus único a una tasa de aproximadamente 50 % (Aldemita y Hodges 1996, Chan y col. 1993, Hiei y col. 1994).

#### *Transformación de maíz*

La transformación de maíz (*Zea mays*) se realizó con una modificación del procedimiento descrito por Ishida y col. (1996) Nature Biotech 14(6): 745-50. La transformación es dependiente del genotipo en maíz y solo genotipos específicos son susceptibles a transformación y regeneración. Como progenitores, la línea endogámica A188 (University of Minnesota) o híbridos con A188, son buenas fuentes de material donante para la transformación aunque también pueden usarse otros genotipos satisfactoriamente. Se cosecharon espigas de plantas de maíz aproximadamente 11 días después de la polinización (DDP) cuando la longitud del embrión inmaduro era de aproximadamente 1 a 1,2 mm. Los embriones inmaduros se co-cultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través de organogénesis. Los embriones cortados se cultivaron en medio de inducción de callo y después en medio de la regeneración de maíz, que contenía el agente de selección (por ejemplo imidazolinona, aunque pueden usarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas, o hasta que se desarrollasen los brotes. Los brotes verdes se transfirieron desde cada embrión al medio de enraizamiento de maíz y se incubaron a 25 °C durante 2-3 semanas, hasta el desarrollo de las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

#### *Transformación de trigo*

La transformación de trigo se realizó con el procedimiento descrito por Ishida y col. (1996) Nature Biotech 14(6): 745-50. Normalmente se utiliza la variedad de cultivo Bobwhite (disponible en CIMMYT, México) para la transformación. Embriones inmaduros se cocultivaron con *Agrobacterium tumefaciens* que contenía el vector de expresión y las plantas transgénicas se recuperaron a través de organogénesis. Después de la incubación con *Agrobacterium*, los embriones se cultivaron *in vitro* en medio de inducción de callo, después en medio de regeneración que contiene el agente de selección (por ejemplo imidazolinona, aunque pueden utilizarse diversos marcadores de selección). Las placas de Petri se incubaron a la luz a 25 °C durante 2-3 semanas o hasta que se desarrollasen los brotes. Los brotes verdes se transfirieron desde cada embrión al medio de enraizamiento y se incubaron a 25 °C durante 2-3 semanas, hasta el desarrollo de las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

#### *Transformación de soja*

La soja se transformó de acuerdo con una modificación del procedimiento descrito en el Texas A y M patente de Estados Unidos 5.164.310. Diversas variedades comerciales de soja son susceptibles a transformación mediante este procedimiento. La variedad de cultivo Jack (disponible en la fundación Illinois Seed) se utiliza normalmente para la transformación. Se esterilizaron semillas de soja para la siembra *in vitro*. El hipocotiledóneo, el radículo y un cotiledón se extirparon de plántulas jóvenes de siete días de vida. El epicótilo y el cotiledón restantes se cultivaron

adicionalmente para desarrollar nodos axilares. Estos nodos axilares se extirparon y se incubaron con *Agrobacterium tumefaciens* que contenía el vector de expresión. Después del tratamiento de cocultivo, los explantes se lavaron y se transfirieron a medios de selección. Los brotes regenerados se extirparon y se colocaron en un medio de elongación de brote. Los brotes no mayores de 1 cm se colocaron en medio de enraizamiento hasta que se desarrollaron las raíces. Los brotes enraizados se trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

#### Transformación de colza/canola

Peciolos cotiledonarios e hipocótilos de plántulas jóvenes de 5-6 días de vida se utilizaron como explantes para el cultivo tisular y se transformaron de acuerdo con Babic y col. (1998, Plant Cell Rep 17: 183-188). La variedad de cultivo comercial Westar (Agriculture Canadá) es la variedad convencional usada para transformación, pero también pueden usarse otras variedades. Se esterilizaron semillas de canola en la superficie para la siembra *in vitro*. Explantes de peciolos cotiledonarios con el cotiledón unido se extirparon de las plántulas *in vitro* y se inocularon con *Agrobacterium* (que contenía el vector de expresión) sumergiendo el extremo cortado del explante del peciolo en la suspensión bacteriana. Después, los explantes se cultivaron durante 2 días en medio MSBAP-3 que contenía BAP 3 mg/l, sacarosa al 3 %, Fitagar al 0,7 % a 23 °C, 16 h de luz. Después de dos días de cocultivo con *Agrobacterium*, los explantes de peciolos se transfirieron a medio MSBAP-3 que contenía BAP 3 mg/l, cefotaxima, carbenicilina o timentina (300 mg/l) durante 7 días y después se cultivaron en medio MSBAP-3 con cefotaxima, carbenicilina o timentina y agente de selección hasta la regeneración del brote. Cuando los brotes tenía una longitud de 5 - 10 mm, se cortaron y se transfirieron al medio de elongación de brote (MSBAP-0,5 que contenía BAP 0,5 mg/l). Los brotes de aproximadamente 2 cm de longitud se transfirieron al medio de enraizamiento (MS0) para la inducción de raíces. Los brotes enraizados trasplantaron al suelo en el invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

#### Transformación de alfalfa

Utilizando el procedimiento de (McKersie y col., 1999 Plant Physiol 119: 839-847) se transformó un clon de regeneración de alfalfa (*Medicago sativa*). La regeneración y transformación de la alfalfa es dependiente del genotipo y por lo tanto se requiere una planta regenerada. Se han descrito procedimientos para obtener plantas regeneradas. Por ejemplo, estas pueden seleccionarse de variedades de cultivo Rangelander (Agriculture Canadá) o de cualquier otra variedad comercial de alfalfa como describen Brown DCW y A Atanassov (1985. Plant Cell Tissue Organ Culture 4: 111-112). Como alternativa, se ha seleccionado la variedad RA3 (University of Wisconsin) para su uso en cultivo tisular (Walker y col., 1978 Am J Bot 65: 654-659). Explantes de peciolos se co-cultivaron durante una noche con un cultivo de *Agrobacterium tumefaciens* C58C1 pMP90 (McKersie y col., 1999 Plant Physiol 119: 839-847) o LBA4404 que contenía el vector de expresión. Los explantes se co-cultivaron durante 3 días en la oscuridad en un medio de inducción SH que contenía Pro 288 mg/l, tioprolina 53 mg/l, K<sub>2</sub>SO<sub>4</sub> 4,35 g/l y acetosiringinona 100 µm. Los explantes se lavaron en medio Murashige-Skoog de fuerza media (Murashige y Skoog, 1962) y se sembraron en placas en el mismo medio de inducción SH con acetosiringinona pero con un agente de selección adecuado y un antibiótico adecuado para inhibir el crecimiento de *Agrobacterium*. Después de varias semanas, los embriones somáticos se transfirieron al medio de desarrollo BOi2Y que no contenía reguladores de crecimiento, sin antibióticos, y sacarosa 50 g/l. Posteriormente, los embriones somáticos germinaron en medio Murashige-Skoog de fuerza media. Las plántulas enraizadas se trasplantaron en macetas y se cultivaron en un invernadero. Se produjeron semillas T1 de plantas que presentaban tolerancia al agente de selección y que contenían una sola copia del inserto de ADN-T.

#### Transformación de algodón

El algodón se transformó utilizando *Agrobacterium tumefaciens* de acuerdo con el procedimiento descrito en el documento US 5.159.135. Se esterilizó la superficie de semillas de algodón en una solución de hipoclorito sódico al 3 % durante 20 minutos y se lavaron en agua destilada con cefotaxima 500 µg/ml. Las semillas después se transfirieron a medio SH con benomil 50 µg/ml para la germinación. Se retiraron hipocótilos de plántulas de de 4 a 6 días, se cortaron en trozos de 0,5 cm y se pusieron en agar al 0,8 %. Para la inoculación de los explantes de hipocótilo se utilizó una suspensión de *Agrobacterium* (aprox. 108 células por ml, diluidas a partir de un cultivo de cultivado durante toda la noche transformado con el gen de interés y con marcadores de selección adecuados). Después de 3 días a temperatura ambiente e iluminación, se transfirieron los tejidos a medio sólido (Gelrite 1,6 g/l), con sales de Murashige y Skoog con vitaminas B5 (Gamborg y col., Exp. Cell Res. 50: 151-158 (1968)), 2,4-D 0,1 mg/l, 6-furfurilaminopurina 0,1 mg/ml y MgCl<sub>2</sub> 750 µg/2, y con de cefotaxima 50 a 100 µg/ml y carbenicilina 400-500 µg/ml para eliminar bacterias residuales. Después de dos a tres meses se aislaron líneas celulares individuales (con subcultivos cada cuatro a seis semanas) y se cultivaron adicionalmente en medio selectivo para amplificación tisular (30 °C, fotoperíodo de 16 h). Posteriormente, los tejidos transformados se cultivaron adicionalmente en medio no selectivo durante dos a tres meses para obtener los embriones somáticos. Se transfirieron embriones de apariencia saludable de por lo menos 4 mm de longitud a tubos con medio SH en vermiculita fina, complementado con ácido indolacético 0,1 mg/l, 6 furfurilaminopurina y ácido giberélico. Los embriones se cultivaron a 30 °C con una fotoperíodo de 16 h, y los plantones en la etapa de 2 a 3 hojas se transfirieron a recipientes con vermiculita y nutrientes. Las plantas se endurecieron y posteriormente se trasladaron al invernadero para cultivo adicional.

**Ejemplo 6: Procedimiento de evaluación fenotípica**6.1 Configuración de la evaluación

Se generaron aproximadamente 35 transformantes de arroz T0 independientes. Los transformantes primarios se transfirieron desde una cámara de cultivo tisular a un invernadero para siembra y cosecha de la semilla T1. Se conservaron seis acontecimientos, de los cuales la progenie T1 segregó 3:1 para la presencia/ausencia del transgén. Para cada uno de estos acontecimientos, se seleccionan aproximadamente 10 plántulas T1 que contenían el transgén (hetero y homocigotos) y aproximadamente 10 plántulas T1 que carecían del transgén (nulicigotos) supervisando la expresión con marcador visual. Las plantas transgénicas y los correspondientes nulicigotos se cultivaron a cada lado en posiciones aleatorias. Las condiciones en el invernadero eran días cortos (12 horas de luz), 28 °C con luz y 22 °C en oscuridad, y una humedad relativa del 70 %.

Se evaluaron adicionalmente 4 acontecimientos T1 en la generación T2, siguiendo el mismo procedimiento de evaluación que para la generación T1, pero con más individuos por acontecimiento. Desde la etapa de siembra hasta la etapa de madurez las plantas pasaron varias veces a través de una cabina de formación de imágenes digital. En cada punto de tiempo se tomaron imágenes digitales (2048x1536 píxeles, 16 millones de colores) de cada planta desde por lo menos 6 ángulos diferentes.

*Detección de sequía*

Se crecieron plantas de seis acontecimientos (semillas T2) en suelo de maceta en condiciones normales hasta que alcanzaron la etapa de encabezamiento. Después se transfirieron a una sección "seca" donde se mantenía el riego. Se insertaron sondas de humedad en macetas seleccionadas al azar para supervisar el contenido hídrico del suelo (CHS). Cuando el CHS estaba por debajo de determinados umbrales, las plantas volvieron a regarse automáticamente de manera continua hasta que de nuevo alcanzaron un nivel normal. Después las plantas se volvieron a transferir a condiciones normales. El resto del cultivo (maduración de planta, cosecha de semilla) fue igual que para las plantas que no se cultivaron en condiciones de estrés abiótico. Los parámetros de cultivo y rendimiento se registraron como se detalla para el cultivo en condiciones normales.

*Detección de la eficiencia del uso de nitrógeno*

Se cultivaron plantas de arroz de semillas T2 en suelo de maceta en condiciones normales excepto para la solución de nutrientes. Las macetas se riegan desde el trasplante hasta la maduración con una solución de nutrientes específica que contenía un contenido de nitrógeno N (N) reducido, normalmente entre 7 a 8 veces menor. El resto del cultivo (maduración de planta, cosecha de semilla) es igual que para las plantas que no se cultivaron bajo estrés abiótico. Los parámetros de cultivo y rendimiento se registraron como se detalla para el cultivo en condiciones normales.

*Detección de estrés salino*

Las plantas se crecen en un sustrato hecho de fibras de coco y argex (proporción 3 a 1). Se utiliza una solución de nutrientes normal durante las dos primeras semanas después del trasplante de los plantones en el invernadero. Después de las dos primeras semanas, se añade a la solución de nutrientes sal 25 mM (NaCl), hasta que las plantas se cosechan. Después se miden los parámetros relacionados con la semilla.

6.2 Análisis estadístico: prueba F

Se utilizó un ANOVA (análisis de varianza) de dos factores como un modelo estadístico para la evaluación global de las características fenotípicas de la planta. Se realizó un ensayo F en todos los parámetros medidos de todas las plantas de todos los acontecimientos transformados con el gen de la presente invención. El ensayo F se realizó para verificar un efecto del gen sobre todos los acontecimientos de transformación y para verificar un efecto global del gen, también conocido efecto génico global. El umbral de significado de un efecto génico global verdadero se establece a un nivel de probabilidad de 5 % para el ensayo F. Un valor significativo del ensayo F para un efecto génico, significa que no solo la mera presencia o la posición del gen ocasionan las diferencias en el fenotipo.

Debido a que se llevaron a cabo dos experimentos con acontecimientos solapantes, se realizó un análisis combinado. Esto es útil para comprobar la consistencia de los efectos sobre los dos experimentos, y si este es el caso, para acumular evidencia de ambos experimentos para aumentar la aceptación de la conclusión. El procedimiento utilizado fue una estrategia de modelo mixto que toma en cuenta la estructura de multinivel de los datos (es decir experimento - acontecimiento - segregantes). Los valores P se obtuvieron comparando la prueba de proporción de probabilidades con las distribuciones de chi cuadrado.

6.3 Parámetros medidos**Medición de parámetros relacionados con biomasa**

Desde la etapa de siembra hasta la etapa de madurez las plantas pasaron varias veces a través de una cabina de formación de imágenes digital. En cada punto de tiempo se tomaron imágenes digitales (2048x1536 píxeles, 16

millones de colores) de cada planta desde por lo menos 6 ángulos diferentes.

El área aérea de la planta (o biomasa foliar) se determinó contando el número total de píxeles en las imágenes digitales de las partes de la planta por encima de la superficie discriminadas del fondo. Este valor se promedió para las imágenes tomadas en el mismo momento de los diferentes ángulos y se convirtió a un valor de superficie física expresado en mm cuadrados por calibración. Los experimentos mostraron que el área de la planta por encima de la superficie medida de este modo se correlacionaba con la biomasa de las partes de la planta por encima de la superficie. El área por encima de la superficie es el área medida en el momento en el que la planta ha alcanzado su máxima biomasa foliar. El vigor temprano es el área por encima de la superficie de la planta (plántula) tres semanas después de la germinación. El aumento de la biomasa radicular se expresa como un incremento en la biomasa radicular total (medido como biomasa radicular máxima observada durante la vida útil de una planta); o como un aumento en el índice de raíz/brote (medido como la relación entre la masa radicular y la masa de brote en el período de crecimiento activo de raíces y brotes).

El vigor temprano se determinó contando el número total de píxeles de las partes de la planta por encima de la superficie discriminadas del fondo. Este valor se promedió para las imágenes tomadas en el mismo momento desde diferentes ángulos y se convirtió a un valor de superficie física expresado en mm cuadrados por calibración. Los resultados descritos a continuación son para plantas tres semanas después de la germinación.

**Mediciones de parámetros relacionados con la semilla**

Las panículas primarias maduras se cosecharon, se contaron, se metieron en bolsas, se etiquetaron con código de barras y después se secaron durante tres días en un horno a 37 ° C. Después, las panículas se trillaron y todas las semillas se recogieron y se contaron. Las vainas llenas se separaron de las vacías utilizando un dispositivo de soplado de aire. Las vainas vacías se desecharon y la fracción restante volvió a contarse. Las vainas llenas se pesaron en una balanza analítica. El número de semillas llenas se determinó contando el número de vainas llenas que quedaron después de la etapa de separación. El rendimiento total de la semilla se midió pesando todas las vainas llenas cosechadas de una planta. El número total de semillas por planta se midió contando el número de vainas cosechadas de una planta. El peso de mil granos (PMG) se extrapoló del número de semillas llenas contadas y de su peso total. El índice de cosecha (IC) en la presente invención se define como la relación entre el rendimiento de semilla total y el área por encima de la superficie (mm<sup>2</sup>), multiplicado por un factor de 10<sup>6</sup>. El número total de flores por panícula como se define en la presente invención es la relación entre el número total de semillas y el número de panículas primarias maduras. La tasa de llenado de semilla, como se define en la presente invención, es la proporción (expresada en %) del número de semillas llenas sobre el número total de semillas (o floretes).

**Ejemplo 7: Resultados de la evaluación fenotípica de las plantas transgénicas**

Se presentan a continuación los resultados de la evaluación de las plantas de arroz transgénicas que expresan un ácido nucleico de HpaG en condiciones sin estrés. Se observó un aumento de la biomasa de la parte aérea (AreaMax), del vigor de emergencia (vigor temprano), del rendimiento total de semilla, del número de semillas llenas, de la tasa de llenado, del número de flores por panoja, del índice de cosecha, y del peso de 1000 granos (véase la Tabla C).

**Tabla C:** Resultados de las mediciones del aumento del rendimiento en condiciones sin estrés

| Parámetro              | Aumento global (en %) | valor p de la prueba F |
|------------------------|-----------------------|------------------------|
| AreaMax                | 13                    | 0,0000                 |
| Vigor temprano         | 25                    | 0,0041                 |
| Peso total de semillas | 30                    | 0,0000                 |
| N.º de semillas llenas | 26                    | 0,0000                 |
| Tasa de llenado        | 9                     | 0,0000                 |
| Flores por panícula    | 12                    | 0,0371                 |
| Índice de Cosecha      | 18                    | 0,0000                 |
| Peso de 1000 Granos    | 4                     | 0,0000                 |

A continuación se presentan abajo los resultados de la evaluación de las plantas de arroz transgénicas que expresan un ácido nucleico de HpaG en condiciones de estrés por sequía. Se observó un aumento del peso de total semilla, del número de semillas llenas, de la tasa de llenado, del índice de cosecha y del peso de 1000 granos (Tabla D).



**Tabla D:** resultados de las mediciones del aumento del rendimiento en condiciones de estrés por sequía

| Parámetro              | Aumento global (en %) | valor p de la prueba F |
|------------------------|-----------------------|------------------------|
| Peso total de semillas | 40                    | 0,0000                 |
| N.º de semillas llenas | 37                    | 0,0000                 |
| Tasa de llenado        | 30                    | 0,0000                 |
| Índice de Cosecha      | 37                    | 0,0000                 |
| Peso de 1000 Granos    | 3                     | 0,0001                 |

LISTADO DE SECUENCIAS

- 5 <110> Cropdesign N.V. Crop Functional Genomics Center
- <120> Plantas que tienen rasgos potenciados relacionados con el rendimiento y un procedimiento de producción de las mismas
- 10 <130> PF58891
- <160> 114
- 15 <170> PatentIn versión 3.3
- <210> 1
- <211> 402
- <212> ADN
- 20 <213> *Xanthomonas axonopodis*
- <400> 1
- ```

atgaattctt tgaacacaca gctcggcgcc aactcgtcct tctttcaggt tgaccccggc      60
cagaacacgc aatctagtcc gaaccagggc aaccagggca tctcggaaaa gcaactggac      120
cagctgctga cccagctcat catggccctg cttcagcaga gcaacaatgc cgagcagggt      180
cagggtcaag gccagggtgg tgactctggc ggtcagggcg gcaatccgcg gcaggccggg      240
cagtcaaag gctccccctc gcaatacacc caggecctga tgaatatcgt cggagacatt      300
ctccaggegc agaatggtgg cggettcggc ggcggettgg gtggtggett cgggtggcacc      360
ctcgtcacca gccttgcgag cgacaccgga togatgcagt aa                          402

```
- 25 <210> 2
- <211> 133
- <212> PRT
- <213> *Xanthomonas axonopodis*
- 30 <400> 2

ES 2 574 981 T3

```

Met Asn Ser Leu Asn Thr Gln Leu Gly Ala Asn Ser Ser Phe Phe Gln
1      5      10      15
Val Asp Pro Gly Gln Asn Thr Gln Ser Ser Pro Asn Gln Gly Asn Gln
      20      25      30
Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Thr Gln Leu Ile Met
      35      40      45
Ala Leu Leu Gln Gln Ser Asn Asn Ala Glu Gln Gly Gln Gly Gln Gly
      50      55      60
Gln Gly Gly Asp Ser Gly Gly Gln Gly Gly Asn Pro Arg Gln Ala Gly
65      70      75      80
Gln Ser Asn Gly Ser Pro Ser Gln Tyr Thr Gln Ala Leu Met Asn Ile
      85      90      95
Val Gly Asp Ile Leu Gln Ala Gln Asn Gly Gly Gly Phe Gly Gly Gly
      100      105      110
Phe Gly Gly Gly Phe Gly Gly Ile Leu Val Thr Ser Leu Ala Ser Asp
      115      120      125
Thr Gly Ser Met Gln
      130

```

- 5 <210> 3
- <211> 11
- <212> PRT
- <213> Secuencia artificial
  
- 10 <220>
- <223> motivo conservado 1
  
- 15 <220>
- <221> VARIANTE
- <222> (2)..(2)
- <223> /reemplaza = "Glu" /reemplaza = "Asp"
  
- 20 <220>
- <221> VARIANTE
- <222> (3)..(3)
- <223> /reemplaza = "Glu"
  
- 25 <220>
- <221> INCIERTO
- <222> (4)..(9)
- <223> Xaa puede ser cualquier aminoácido de origen natural
  
- 30 <220>
- <221> VARIANTE
- <222> (5)..(5)
- <223> /reemplaza = "Arg" /reemplaza = "Pro"
  
- 35 <220>
- <221> VARIANTE
- <222> (7)..(7)
- <223> /reemplaza = "Ser"
  
- 40 <220>
- <221> INCIERTO
- <222> (9)..(9)
- <223> Xaa puede ser cualquier aminoácido de origen natural
  
- 45 <220>
- <221> VARIANTE
- <222> (10)..(10)
- <223> /reemplaza = "Asp"
  
- <400> 3

ES 2 574 981 T3

Gly Gly Asn Xaa Gln Gln Ala Gly Xaa Asn Gly  
 1 5 10

5 <210> 4  
 <211> 14  
 <212> PRT  
 <213> Secuencia artificial

10 <220>  
 <223> motivo conservado 2

15 <220>  
 <221> VARIANTE  
 <222> (1)..(1)  
 <223> /reemplaza = "Ala" /reemplaza = "Val"

20 <220>  
 <221> VARIANTE  
 <222> (3)..(3)  
 <223> /reemplaza = "Gln" /reemplaza = "Ala"

25 <220>  
 <221> VARIANTE  
 <222> (4)..(4)  
 <223> /reemplaza = "Leu" /reemplaza = "Tyr"

30 <220>  
 <221> VARIANTE  
 <222> (7)..(7)  
 <223> /reemplaza = "Ala"

35 <220>  
 <221> VARIANTE  
 <222> (10)..(10)  
 <223> /reemplaza = "Asn" /reemplaza = "Gln"

40 <220>  
 <221> VARIANTE  
 <222> (13)..(13)  
 <223> /reemplaza = "Met"

45 <220>  
 <221> VARIANTE  
 <222> (14)..(14)  
 <223> /reemplaza = "Asp" /reemplaza = "Gln"

<400> 4

Pro Ser Pro Phe Thr Gln Met Leu Met His Ile Val Gly Glu  
 1 5 10

50 <210> 5  
 <211> 2194  
 <212> ADN  
 <213> *Oryza sativa*

<400> 5

ES 2 574 981 T3

aatccgaaaa gttttctgcac cgttttcacc ccctaactaa caatataggg aacgtgtgct 60  
 aaatataaaa tgagacctta tatatgtagc gctgataact agaactatgc aagaaaaact 120  
 catocaccta ctttagtggc aatcgggcta aataaaaaag agtcgctaca ctagtctcgt 180  
 tttccttagt aattaagtgg gaaaaatgaaa tcattattgc ttagaatata cgttcacatc 240  
 tctgtcatga agttaaatta ttogaggtag ccataattgt catcaaactc ttcttgaata 300  
 aaaaaatctt tctagctgaa ctcaatgggt aaagagagag atttttttta aaaaaataga 360  
 atgaagatat tetgaacgta ttggcaaaga tttaaacata taattatata attttatagt 420  
 ttgtgcattc gtcatatcgc acatcattaa ggacatgtct tactccatcc caatttttat 480  
 ttagtaatta aagacaattg acttattttt attattttat ttttttcgat tagatgcaag 540

gtacttaacg acacactttg tgetcatgtg catgtgtgag tgcacctcct caatacacgt 600  
 tcaactagca acacatctct aatatacact gcctatttaa tacatttagg tagcaatata 660  
 tgaattcaag cactccacca tcaccagacc acttttaata atatctaaaa tacaanaaat 720  
 aattttacag aatagcatga aaagtatgaa acgaaactatt taggtttttc acatacaaaa 780  
 aaaaaaagaa ttttgetcgt gcgcgagcgc caatctccca tattgggcac acaggcaaca 840  
 acagagtggc tgcccacaga acaacccaca aaaaacgatg atctaacgga ggacagcaag 900  
 tccgcaacaa ccttttaaca gcaggctttg cggccaggag agaggaggag aggcaaaaga 960  
 aaccaagcat cctccttctc ccatactataa attcctcccc cctttcccc tctctatata 1020  
 ggaggcatcc aagccaagaa gagggagagc accaaggaca cgcgactagc agaagccgag 1080  
 cgaccgcctt ctgcgatccat atcttccggt cgagttcttg gtcgatctct tccctctctc 1140  
 acctctctct cacaggggat gtgcctcctt tgggttgttc ttggatttat tgttctaggt 1200  
 tgtgtagtag gggcgttgat gttaggaag gggatctgta tctgtgatga ttctgtttct 1260  
 tggatttggg atagaggggt tcttgatggt gcattgttata ggttcggttt gattagtagt 1320  
 atggttttca atcgtctgga gagctctatg gaaatgaaat ggtttaggga tcggaatctt 1380  
 gcgattttgt gactaccttt tgtttgaggt aaaatcagag caccgggtgat tttgcttggg 1440  
 gtaataaagt acggttgttt ggtcctcgat tctggtagtg atgettctcg atttgacgaa 1500  
 gctatecttt gtttattecc tattgaacaa aaataatcca actttgaaga cggteccggt 1560  
 gatgagattg aatgattgat tcttaagcct gtccaaaatt tgcagctgg cttgtttaga 1620  
 tacagtagtc cccatcacga aattcatgga aacagttata atcctcagga acaggggatt 1680  
 cctgtttctt ccgatttget ttagtcccag aattttttt cccaaatata ttaaaaagtc 1740  
 actttctggg tcagttcaat gaattgattg ctacaaataa tgettttata gcgttatcct 1800  
 agctgtagtt cagttaatag gtaatacccc tatagtttag tcaggagaag aacttatccg 1860  
 atttctgatc tccattttta attatatgaa atgaactgta gcataagcag tattcatttg 1920  
 gattattttt tttattagct ctaccccctt cattattctg agctgaaagt ctggcatgaa 1980  
 ctgtcctcaa ttttgttttc aaattccatc cgattatcta tgcattatcc tcttgtatct 2040  
 acctgtagaa gtttcttttt ggttattcct tgactgcttg attacagaaa gaaatttatg 2100  
 aagctgtaat cgggatagtt atactgcttg ttcttatgat tcatttcctt tgtgcagttc 2160  
 ttggtgtagc ttgccacttt caccagcaaa gttc 2194

- <210> 6
- <211> 1179
- 5 <212> ADN
- <213> *Oryza sativa*
- <400> 6

ES 2 574 981 T3

```

ttgcagttgt gaccaagtaa gctgagcatg ccttaactt cacctagaaa aaagtatact 60
tggcttaact gctagtaaga catttcagaa ctgagactgg tgtacgcatt tcatgcaagc 120
cattaccact ttacctgaca ttttggacag agattagaaa tagtttcgta ctacctgcaa 180
gttgcaactt gaaaagtgaa atttgttctt tgctaataata ttggcgtgta attcttttat 240
gcgtagcgt aaaaagttga aatttgggtc agttactgg tcagattaac cagtaactgg 300
ttaaagttga aagatggtct tttagtaatg gagggagtac tacactatcc tcagctgatt 360
taaateetat tccgtcgggtg gtgatttcgt caatctccca acttagtttt tcaatatatt 420
cataggatag agtgtgcata tgtgtgttta tagggatgag tctacgcgcc ttatgaacac 480
ctacttttgt actgtatttg tcaatgaaaa gaaaatctta ccaatgctgc gatgctgaca 540
ccaagaagag gcgatgaaaa gtgcaacgga tctcgtgcc cgtcgggttc caagtcagca 600
cagacccaat gggcctttcc taegtgtctc ggccacagcc agtcggttac cgcacgttca 660
catgggcacg aactcgcgtc atcttcccac gcaaaaacgac agatctgccc tatctggtec 720
caccatcag tggcccacac ctcccattgt gcattatttg cgaactccat cccgtcctcc 780
accgccaac accgcacacg ggtcgcgata gccacgaccc aatcacacaa cgcacgtca 840
ccatatgtta cgggcagcca tgcgcagaag atcccgcgac gtcgctgtcc cccgtgtcgg 900
ttacgaaaaa atatcccacc acgtgtcgtc ttcacaggac aatatctcga aggaaaaaaa 960
tcgtagcgga aaatccgagg cacgagctgc gattggctgg gagcgtcca gcgtgggtgg 1020
gggcccacc ccttactctt agcccgtggc gctcctcgtc cctcgggtcc gtgtataaat 1080
accctcggga actcactctt gctggtcacc aacacgaage aaaaggacac cagaacata 1140

```

```

gtacacttga gctcactcca aactcaaaca ctcacacca 1179

```

5 <210> 7  
 <211> 402  
 <212> ADN  
 <213> Secuencia artificial

10 <220>  
 <223> gen de HpaG\_T44C desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética

400> 7

```

atgaattctt tgaacacaca gctcggcgcc aactcgtctc tctttcaggt tgaccccggc 60
cagaacacgc aatctagtcc gaaccagggc aaccagggca tctcggaaaa gcaactggac 120
cagctgctgt gccagctcat catggccttg cttcagcaga gcaacaatgc cgagcagggt 180
cagggtcaag gccagggtgg tgaactctgc ggtcagggcg gcaatccgag gcaggccggg 240
cagtcacaacg gctccccctc gcaatacacc caggcgtgta tgaatatcgt cggagacatt 300
ctcagggcg agaatggtgg cggcttcggc ggccgctttg gtgggtgctt cggtggtcgc 360
ctcgtcacca gccttgcgag cgacaccgga tcatgcagt aa 402

```

15 <210> 8  
 <211> 133  
 <212> PRT  
 <213> Secuencia artificial

20 <220>  
 <223> HpaG\_T44C desencadenante de la respuesta de hipersensibilidad mutante

25 <400> 8

ES 2 574 981 T3

```

Met Asn Ser Leu Asn Thr Gln Leu Gly Ala Asn Ser Ser Phe Phe Gln
1          5          10          15
Val Asp Pro Gly Gln Asn Thr Gln Ser Ser Pro Asn Gln Gly Asn Gln
20          25          30
Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Cys Gln Leu Ile Met
35          40          45
Ala Leu Leu Gln Gln Ser Asn Ala Glu Gln Gly Gln Gly Gln Gly
50          55          60
Gln Gly Gly Asp Ser Gly Gly Gln Gly Gly Asn Pro Arg Gln Ala Gly
65          70          75          80
Gln Ser Asn Gly Ser Pro Ser Gln Tyr Thr Gln Ala Leu Met Asn Ile
85          90          95
Val Gly Asp Ile Leu Gln Ala Gln Asn Gly Gly Gly Phe Gly Gly Gly
100         105         110
Phe Gly Gly Gly Phe Gly Gly Ile Leu Val Thr Ser Leu Ala Ser Asp
115         120         125
Thr Gly Ser Met Gln
130

```

- 5 <210> 9
- <211> 378
- <212> ADN
- <213> Secuencia artificial
- 10 <220>
- <223> gen de HpaG-T desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética
- <400> 9

```

atgaattcctt tgaacacaca gctcggcgcc aactcgtcct tctttcaggt tgaccccggc      60
cagaacacgc aatctagtc gaaccagggc aaccagggca tctcggaaaa gcaactggac      120
cagctgctga cccagctcat catggccctg cttcagcaga gcaacaatgc cgagcagggt      180
cagggtaag gccaggggtg tgactctggc ggtcagggcg gcaatccgcg gcaggccggg      240
cagtccaacg gctccccctc gcaatacaac caggcctga tgaatatcgt cgggagcggc      300
ttcggcggcg gctttggtgg tggtctcggg ggcatacctc tcaccagcct tgcgagcgac      360
accggatcga tgcagtaa
378

```

- 15 <210> 10
- <211> 125
- <212> PRT
- <213> Secuencia artificial
- 20 <220>
- <223> HpaG-T desencadenante de la respuesta de hipersensibilidad mutante
- <400> 10

ES 2 574 981 T3

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Asn | Thr | Gln | Leu | Gly | Ala | Asn | Ser | Ser | Phe | Phe | Gln |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Val | Asp | Pro | Gly | Gln | Asn | Thr | Gln | Ser | Ser | Pro | Asn | Gln | Gly | Asn | Gln |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Gly | Ile | Ser | Glu | Lys | Gln | Leu | Asp | Gln | Leu | Leu | Thr | Gln | Leu | Ile | Met |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ala | Leu | Leu | Gln | Gln | Ser | Asn | Asn | Ala | Glu | Gln | Gly | Gln | Gly | Gln | Gly |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gln | Gly | Gly | Asp | Ser | Gly | Gly | Gln | Gly | Gly | Asn | Pro | Arg | Gln | Ala | Gly |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Gln | Ser | Asn | Gly | Ser | Pro | Ser | Gln | Tyr | Thr | Gln | Ala | Leu | Met | Asn | Ile |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Val | Gly | Asp | Gly | Phe | Gly | Gly | Gly | Phe | Gly | Gly | Gly | Phe | Gly | Gly | Ile |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Leu | Val | Thr | Ser | Leu | Ala | Ser | Asp | Thr | Gly | Ser | Met | Gln |     |     |     |
|     |     |     | 115 |     |     |     | 120 |     |     |     |     | 125 |     |     |     |

<210> 11  
 <211> 414  
 <212> ADN  
 <213> *Xanthomonas axonopodis* pv. *citri*

5

<400> 11

|             |             |            |             |             |             |     |
|-------------|-------------|------------|-------------|-------------|-------------|-----|
| ttactgcatac | gatccgggtgt | cgctcgcaag | gctgggtgccg | aggetgggtgc | cgaggccgcc  | 60  |
| gccgaagcca  | ccaccaaacg  | cgccgccgaa | gccaccacca  | ttctgcccct  | ggagaatgtc  | 120 |
| tccgacgata  | ttcatcagca  | tctgggtgta | ttgctgagggg | gagccgttgg  | actgaccggc  | 180 |
| ctgctgcca   | ttgcccccct  | gaccaccaga | gtcaccacc   | tggccttgac  | cctgaccctg  | 240 |
| ctcggcattg  | ttgctctgct  | gaagcagggc | catgatgagc  | tgggtcagca  | gctgggtccag | 300 |
| ttgcttttcc  | gagatgccct  | ggttgccctg | gttcgaacca  | gattgctgtg  | tctggctggg  | 360 |
| gtcaacctga  | aagaaggagc  | agttggcgcc | gagctgtgtg  | ttcaaagaat  | tcat        | 414 |

10

<210> 12  
 <211> 137  
 <212> PRT  
 <213> *Xanthomonas axonopodis* pv. *citri*

15

<400> 12

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Asn | Thr | Gln | Leu | Gly | Ala | Asn | Ser | Ser | Phe | Phe | Gln |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Val | Asp | Pro | Ser | Gln | Asn | Thr | Gln | Ser | Gly | Ser | Asn | Gln | Gly | Asn | Gln |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Gly | Ile | Ser | Glu | Lys | Gln | Leu | Asp | Gln | Leu | Leu | Thr | Gln | Leu | Ile | Met |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ala | Leu | Leu | Gln | Gln | Ser | Asn | Asn | Ala | Glu | Gln | Gly | Gln | Gly | Gln | Gly |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gln | Gly | Gly | Asp | Ser | Gly | Gly | Gln | Gly | Gly | Asn | Arg | Gln | Gln | Ala | Gly |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Gln | Ser | Asn | Gly | Ser | Pro | Ser | Gln | Tyr | Thr | Gln | Met | Leu | Met | Asn | Ile |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Val | Gly | Asp | Ile | Leu | Gln | Ala | Gln | Asn | Gly | Gly | Gly | Phe | Gly | Gly | Gly |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Phe | Gly | Gly | Gly | Phe | Gly | Gly | Gly | Leu | Gly | Thr | Ser | Leu | Gly | Thr | Ser |
|     |     |     | 115 |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Leu | Ala | Ser | Asp | Thr | Gly | Ser | Met | Gln |     |     |     |     |     |     |     |
|     |     |     | 130 |     |     |     | 135 |     |     |     |     |     |     |     |     |

20

<210> 13  
 <211> 366  
 <212> ADN  
 <213> Secuencia artificial

ES 2 574 981 T3

<220>

<223> gen de HpaG-N desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética

<400> 13

5

```

atgaattctt tgaacacaca gctcggcgcc aactcgtcct tctttcaggt tgaccccggc      60
cagaacacgc aatctagtc gaaccagggc aacaccagc tcatcatggc cctgcttcag      120
cagagcaaca atgcecgagca gggtcagggt caaggccagg gtggtgactc tggcggtcag      180
ggcggcaatc cgcggcaggc cgggcagtc aacggctccc cctcgcaata caccagggc      240
ctgatgaata tcgtcggaga cattctccag gcgcagaatg gtggcggctt cggcggcggc      300
tttggtggtg gcttcgggtg cctcctcgtc accagccttg cgagcgacac cggatcgatg      360
cagtaa
    
```

<210> 14

<211> 121

10

<212> PRT

<213> Secuencia artificial

<220>

<223> HpaG-N desencadenante de la respuesta de hipersensibilidad mutante

15

<400> 14

```

Met Asn Ser Leu Asn Thr Gln Leu Gly Ala Asn Ser Ser Phe Phe Gln
1      5      10
Val Asp Pro Gly Gln Asn Thr Gln Ser Ser Pro Asn Gln Gly Asn Thr
      20      25      30

Gln Leu Ile Met Ala Leu Leu Gln Gln Ser Asn Asn Ala Glu Gln Gly
      35      40      45
Gln Gly Gln Gly Gln Gly Gly Asp Ser Gly Gly Gln Gly Gly Asn Pro
      50      55      60
Arg Gln Ala Gly Gln Ser Asn Gly Ser Pro Ser Gln Tyr Thr Gln Ala
      65      70      75      80
Leu Met Asn Ile Val Gly Asp Ile Leu Gln Ala Gln Asn Gly Gly Gly
      85      90      95
Phe Gly Gly Gly Phe Gly Gly Gly Phe Gly Gly Ile Leu Val Thr Ser
      100      105      110
Leu Ala Ser Asp Thr Gly Ser Met Gln
      115      120
    
```

20

<210> 15

<211> 366

<212> ADN

<213> *Xanthomonas axonopodis*

25

<400> 15

```

atgaattctt tgaacacaca gctcggcgcc aactcgtcct tctttcaggt tgaccccggc      60
cagaacacgc aatctagtc gaaccagggc aaccagggca tctcggaaaa gcaactggac      120
cagctgctga cccagctcat catggccttg cttcagcaga gcaacaatgc cgagcagggt      180
cagggtaacg gccaggggtg tgactctggc ggtcagggcg gcaatccgcg gcaggccggg      240
cagtccaacg gctccccctc gcaatacacc caggecgtga tgaatatcgt cggagacatt      300
ctccagggcg agaatggctt tctcctcgtc accagccttg cgagcgacac cggatcgatg      360
cagtaa
    
```

30

<210> 16

<211> 121

<212> PRT

<213> *Xanthomonas axonopodis*



ES 2 574 981 T3

<400> 16

```

Met Asn Ser Leu Asn Thr Gln Leu Gly Ala Asn Ser Ser Phe Phe Gln
1      5      10      15
Val Asp Pro Gly Gln Asn Thr Gln Ser Ser Pro Asn Gln Gly Asn Gln
      20      25      30
Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Thr Gln Leu Ile Met
      35      40      45
Ala Leu Leu Gln Gln Ser Asn Asn Ala Glu Gln Gly Gln Gly Gln Gly
      50      55      60
Gln Gly Gly Asp Ser Gly Gly Gln Gly Gly Asn Pro Arg Gln Ala Gly
65      70      75      80
Gln Ser Asn Gly Ser Pro Ser Gln Tyr Thr Gln Ala Leu Met Asn Ile
      85      90      95
Val Gly Asp Ile Leu Gln Ala Gln Asn Gly Phe Ile Leu Val Thr Ser
      100      105      110
Leu Ala Ser Asp Thr Gly Ser Met Gln
      115      120

```

5 <210> 17  
 <211> 402  
 <212> ADN  
 <213> *Xanthomonas smithii* subsp. *smithii*

10 <400> 17

```

atgaattctt tgaacacaca gatcggcgcc aactcgtcct tcttgcaggt cgacccgagc      60
cagaacacgc aattcgggtcc gaaccagggc aatcaaggca tctcggaaaa gcagctggac      120
cagctgctga cccagctcat catggccctg cttcagcaga gcaacaatgc cgaccagggc      180
cagggtggtg actctggttg tcaaggcggc aattcggcggc aggccgggca gcccaatggt      240
tccccctcgg catacaccca gatgctgatg aatatacgtcg gagacattct ccagggcgcag      300
aatgggtggtg gcttcggcggc cgggttcggc ggtggctttg gtggcgggct cggcaccagc      360
ctcggcagca gccttgcgag cgacaccgga tcgatgcagt aa      402

```

15 <210> 18  
 <211> 133  
 <212> PRT  
 <213> *Xanthomonas smithii* subsp. *smithii*

20 <400> 18

```

Met Asn Ser Leu Asn Thr Gln Ile Gly Ala Asn Ser Ser Phe Leu Gln
1      5      10      15
Val Asp Pro Ser Gln Asn Thr Gln Phe Gly Pro Asn Gln Gly Asn Gln
      20      25      30
Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Thr Gln Leu Ile Met
      35      40      45
Ala Leu Leu Gln Gln Ser Asn Asn Ala Asp Gln Gly Gln Gly Gly Asp
      50      55      60
Ser Gly Gly Gln Gly Gly Asn Ser Arg Gln Ala Gly Gln Pro Asn Gly
65      70      75      80
Ser Pro Ser Ala Tyr Thr Gln Met Leu Met Asn Ile Val Gly Asp Ile
      85      90      95
Leu Gln Ala Gln Asn Gly Gly Gly Phe Gly Gly Gly Phe Gly Gly Gly
      100      105      110
Phe Gly Gly Gly Leu Gly Thr Ser Leu Gly Ser Ser Leu Ala Ser Asp
      115      120      125
Thr Gly Ser Met Gln
      130

```

<210> 19  
 <211> 420

ES 2 574 981 T3

<212> ADN  
 <213> *Xanthomonas oryzae* pv. *oryzae*

<400> 19

5

```

atgaactctt tgaacacaca attcggcggc agcacgtcca accttcaggt tggcccaagc      60
caggacacaa cgttcgggtc gaaccagggc ggcaaccagg gcatctcgga aaagcaactg      120
gaccagttgc tgtgccagct catctcggcc ctgcttcagt cgagcaaaaa tgctgaggag      180
ggtaagggtc aggggtggcga taatggcggg gccacagggc gcaattcgca gcaggccggg      240
cagcagaatg gcccctcgcc attcaccagc atgctgatgc atatcgtcgg agagattctc      300
cagggcgaga atggtgggtg tgctgggtgg ggcgggttcg gcggcggggt cggcggcgac      360
tttagtgggc acctcggcct cggcaaccaac ctctcgagcg acagcgcate aatgcagtaa      420
    
```

<210> 20  
 <211> 139  
 <212> PRT  
 <213> *Xanthomonas oryzae* pv. *oryzae*

10

<400> 20

```

Met Asn Ser Leu Asn Thr Gln Phe Gly Gly Ser Thr Ser Asn Leu Gln
1          5          10          15
Val Gly Pro Ser Gln Asp Thr Thr Phe Gly Ser Asn Gln Gly Gly Asn
20         25         30
Gln Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Cys Gln Leu Ile
35         40         45
Ser Ala Leu Leu Gln Ser Ser Lys Asn Ala Glu Glu Gly Lys Gly Gln
50         55         60
Gly Gly Asp Asn Gly Gly Gly Gln Gly Gly Asn Ser Gln Gln Ala Gly
65         70         75         80
Gln Gln Asn Gly Pro Ser Pro Phe Thr Gln Met Leu Met His Ile Val
85         90         95
Gly Glu Ile Leu Gln Ala Gln Asn Gly Gly Gly Ala Gly Gly Gly Gly
100        105        110
Phe Gly Gly Gly Phe Gly Gly Asp Phe Ser Gly Asp Leu Gly Leu Gly
115        120        125
Thr Asn Leu Ser Ser Asp Ser Ala Ser Met Gln
130        135
    
```

15

<210> 21  
 <211> 420  
 <212> ADN  
 <213> *Xanthomonas oryzae* pv. *oryzae*

20

<400> 21

```

atgaattctt tgaacacaca attcggcggc agcacgtcca accttcaggt tggcccaagc      60
caggacacaa cgttcgggtc gaaccagggc ggcaaccagg gcatctcgga aaagcaactg      120
gaccagttgc tgtgccagct catctcggcc ctgcttcagt cgagcaaaaa tgctgaggag      180
ggtaagggtc aggggtggcga taatggcggg gccacagggc gcaattcgca gcaggctggg      240
cagcagaatg gcccctcgcc attcaccagc atgctgatgc atatcgtcgg agagattctc      300
cagggcgaga atggtgggtg tgctgggtgg ggcgggttcg gcggcggggt cggcgggtgac      360
tttagtgggc acctcggcct cggcaaccaac ctctcgagcg acagcgcate gatgcagtaa      420
    
```

25

<210> 22  
 <211> 139  
 <212> PRT  
 <213> *Xanthomonas oryzae* pv. *oryzae*

30

<400> 22

ES 2 574 981 T3

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Asn | Thr | Gln | Phe | Gly | Gly | Ser | Thr | Ser | Asn | Leu | Gln |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Val | Gly | Pro | Ser | Gln | Asp | Thr | Thr | Phe | Gly | Ser | Asn | Gln | Gly | Gly | Asn |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Gln | Gly | Ile | Ser | Glu | Lys | Gln | Leu | Asp | Gln | Leu | Leu | Cys | Gln | Leu | Ile |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ser | Ala | Leu | Leu | Gln | Ser | Ser | Lys | Asn | Ala | Glu | Glu | Gly | Lys | Gly | Gln |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gly | Gly | Asp | Asn | Gly | Gly | Gly | Gln | Gly | Gly | Asn | Ser | Gln | Gln | Ala | Gly |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Gln | Gln | Asn | Gly | Pro | Ser | Pro | Phe | Thr | Gln | Met | Leu | Met | His | Ile | Val |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Gly | Glu | Ile | Leu | Gln | Ala | Gln | Asn | Gly | Gly | Gly | Ala | Gly | Gly | Gly | Gly |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| Phe | Gly | Gly | Gly | Phe | Gly | Gly | Asp | Phe | Ser | Gly | Asp | Leu | Gly | Leu | Gly |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Thr | Asn | Leu | Ser | Ser | Asp | Ser | Ala | Ser | Met | Gln |     |     |     |     |     |
|     |     | 130 |     |     |     |     | 135 |     |     |     |     |     |     |     |     |

<210> 23  
 <211> 420  
 <212> ADN  
 <213> *Xanthomonas oryzae* pv. *oryzae*

5

<400> 23

|            |            |            |            |             |             |     |
|------------|------------|------------|------------|-------------|-------------|-----|
| atgaattctt | tgaacacaca | attcggcggc | agcaagtc   | accttcaggt  | tggccaagc   | 60  |
| caggacacaa | cgttcgggtc | gaaccagggc | ggcaaccagg | gcattctcgga | aaagcaactg  | 120 |
| gaccagttgc | tgtgccagct | catctcggcc | ctgcttcagt | cgagcaaaaa  | tgctgaggag  | 180 |
| ggtaagggtc | aggggtggca | taatggcggg | ggccagggcg | gcaattcgca  | gcagggccggg | 240 |
| cagcagaatg | gcccctcgcc | attcaccag  | atgctgatgc | atatcgctcg  | agagattctc  | 300 |
| cagggcgaga | atgggtggtg | tgetggtggc | ggcgggttcg | gcggcgggtt  | cggcgggtgac | 360 |
| tttagtgggc | acctcggcct | cggcaccaac | ctctcgagcg | acagcgcctc  | gatgcagtaa  | 420 |

10

<210> 24  
 <211> 139  
 <212> PRT  
 <213> *Xanthomonas oryzae* pv. *oryzae*

15

<400> 24

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Asn | Ser | Leu | Asn | Thr | Gln | Phe | Gly | Gly | Ser | Thr | Ser | Asn | Leu | Gln |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |     |
| Val | Gly | Pro | Ser | Gln | Asp | Thr | Thr | Phe | Gly | Ser | Asn | Gln | Gly | Gly | Asn |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |     |     |
| Gln | Gly | Ile | Ser | Glu | Lys | Gln | Leu | Asp | Gln | Leu | Leu | Cys | Gln | Leu | Ile |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |
| Ser | Ala | Leu | Leu | Gln | Ser | Ser | Lys | Asn | Ala | Glu | Glu | Gly | Lys | Gly | Gln |
|     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gly | Gly | Asp | Asn | Gly | Gly | Gly | Gln | Gly | Gly | Asn | Ser | Gln | Gln | Ala | Gly |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Gln | Gln | Asn | Gly | Pro | Ser | Pro | Phe | Thr | Gln | Met | Leu | Met | His | Ile | Val |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Gly | Glu | Ile | Leu | Gln | Ala | Gln | Asn | Gly | Gly | Gly | Ala | Gly | Gly | Gly | Gly |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Phe | Gly | Gly | Gly | Phe | Gly | Gly | Asp | Phe | Ser | Gly | Asp | Leu | Gly | Leu | Gly |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Thr | Asn | Leu | Ser | Ser | Asp | Ser | Ala | Ser | Met | Gln |     |     |     |     |     |
|     |     | 130 |     |     |     |     | 135 |     |     |     |     |     |     |     |     |

ES 2 574 981 T3

<210> 25  
 <211> 378  
 <212> ADN  
 <213> *Xanthomonas oryzae* pv. *oryzicola*

5

<400> 25

```

atgaattctt tgaacacaca attcggcggc agcgcgtcca acttccaggt tgaccaaaagc      60
cagaacgcgc aatccgattc gagccagggc agcaatggca gccagggtat ctccgaaaag      120
caactggacc agttgctgtg ccagctcacc caggccctgc ttcagccgaa caaaaatgct      180
gaggaagta  agggtcagca ggggtggcag aataatcagc aggccgggaa ggagaatggc      240

gcctcggcac tcaccagat gctgatgaat atcgtcggag agattctcca ggcgcagaat      300
gcggcgcca gcagcggcgg cgactttggt ggcagtttcg ccagcagctt ctccaacgac      360
agcggatcga tgcagtaa                                     378
    
```

<210> 26  
 <211> 125  
 <212> PRT  
 <213> *Xanthomonas oryzae* pv. *oryzicola*

15 <400> 26

```

Met Asn Ser Leu Asn Thr Gln Phe Gly Gly Ser Ala Ser Asn Phe Gln
 1                    5                    10                    15
Val Asp Gln Ser Gln Asn Ala Gln Ser Asp Ser Ser Gln Gly Ser Asn
                20                    25                    30
Gly Ser Gln Gly Ile Ser Glu Lys Gln Leu Asp Gln Leu Leu Cys Gln
                35                    40                    45
Leu Ile Gln Ala Leu Leu Gln Pro Asn Lys Asn Ala Glu Glu Gly Lys
 50                    55                    60
Gly Gln Gln Gly Gly Glu Asn Asn Gln Gln Ala Gly Lys Glu Asn Gly
65                    70                    75                    80
Ala Ser Pro Leu Thr Gln Met Leu Met Asn Ile Val Gly Glu Ile Leu
                85                    90                    95
Gln Ala Gln Asn Ala Gly Gly Ser Ser Gly Gly Asp Phe Gly Gly Ser
                100                    105                    110
Phe Ala Ser Ser Phe Ser Asn Asp Ser Gly Ser Met Gln
                115                    120                    125
    
```

<210> 27  
 <211> 366  
 <212> ADN  
 <213> *Xanthomonas campestris* pv. *campestris*

25

<400> 27

```

tcaggcttgg ccggtgatgc tcgacaggtt ggcattgaag ccgccacca agctgggtgcc      60
gcccatgcgc gcgccgcctt ggttctgcat cagctgcacc acgatctgca tcagcatctg      120
cgtaaacgga ctcacaccgt cctggtgacc gctctggcgt tgttcgtctc cgcactcctg      180
atcggcatcg ctgccctggc tctggtggag catcatcatg atgaacatgg cgagcagctg      240
atccagctgc tgctcggagt cagccgaagg cgagcctgga ctggagttct gggtttgcctg      300
gggcccgatg cccatcgtct gcaggttgat gaagttggaa aatttgtttc cgatagatga      360
gtccat                                     366
    
```

<210> 28  
 <211> 121  
 <212> PRT  
 <213> *Xanthomonas campestris* pv. *campestris*

30

<400> 28

ES 2 574 981 T3

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

50 55 60  
 Cys Gly Asp Glu Gln Pro Gln Ser Gly Gln Gln Asp Gly Val Ser Pro  
 65 70 75 80  
 Leu Thr Gln Met Leu Met Gln Ile Val Met Gln Leu Met Gln Asn Gln  
 85 90 95  
 Gly Gly Ala Gly Met Gly Gly Thr Ser Leu Gly Gly Gly Phe Asn Ala  
 100 105 110  
 Asn Leu Ser Ser Ile Thr Gly Gln Ala  
 115 120

## REIVINDICACIONES

1. Un procedimiento para aumentar el rendimiento de semilla y/o la biomasa en plantas con respecto a las plantas de control, que comprende introducir y expresar en una planta un ácido nucleico que codifica un polipéptido HpaG que comprende:
- 5 (i) por lo menos el 80 % o más de identidad de secuencia con la secuencia del polipéptido HpaG representado por la SEQ ID NO: 2; y  
(ii) una composición de aminoácidos en la que el contenido de glicina varía entre el 13 % y el 25 %, el contenido de glutamina varía entre el 13 % y el 20 %, el contenido de cisteína varía entre el 0 % y el 1 %, el contenido de histidina varía entre el 0 % y el 1 %, y en la que el triptófano está ausente,
- 10 en el que dicho aumento del rendimiento de semilla y/o de la biomasa se obtiene en condiciones sin estrés o de estrés abiótico o en el que dicho rendimiento de semilla y/o biomasa aumentado no es el resultado de la resistencia aumentada al estrés biótico.
2. Procedimiento de acuerdo con la reivindicación 1, en el que dicho polipéptido HpaG comprende además uno o más de los siguientes motivos:
- 15 (i) (motivo 1): G(G/E/D)(N/E)X(O/R/P)O(A/S)GX(N/D)G (SEQ ID NO: 3), en el que X en la posición 4 puede ser cualquier aminoácido, preferentemente uno de S, N, P, R o Q, y en el que X en la posición 9 puede ser cualquier aminoácido, preferentemente uno de Q, E, S, o P; y  
(ii) (motivo 2): (P/A/V)S(P/Q/A)(F/L/Y)TQ(M/A)LM(H/N/Q)IV(G/M)(E/D/Q) (SEQ ID NO: 4),
- 20 3. Procedimiento de acuerdo con cualquier reivindicación precedente, en el que uno cualquiera de los ácidos nucleicos enumerados en la Tabla A o una porción de los mismos representa dicho ácido nucleico que codifica un polipéptido Hpag, o una secuencia capaz de hibridar con uno cualquiera de los ácidos nucleicos proporcionados en la Tabla A.
4. Procedimiento de acuerdo con cualquier reivindicación precedente, en el que dicha secuencia de ácido nucleico codifica un ortólogo o parálogo de cualquiera de las proteínas proporcionadas en la Tabla A.
- 25 5. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicho ácido nucleico está unido operativamente a un promotor constitutivo, preferentemente a un promotor GOS2, muy preferentemente a un promotor GOS2 de arroz.
6. Procedimiento de acuerdo con una cualquiera de las reivindicaciones 1 a 4, en el que dicho ácido nucleico está unido operativamente a un promotor específico de tejido verde, preferentemente a un promotor de la protoclorofilida reductasa, muy preferentemente a un promotor de la protoclorofilida reductasa del arroz.
- 30 7. Procedimiento de acuerdo con cualquier reivindicación precedente, en el que dicho ácido nucleico que codifica un polipéptido HpaG es de origen procarionte, preferentemente de una bacteria patógena vegetal que posee un Sistema de Secreción de Tipo Tres (SSTT), además preferentemente de la familia *Pseudomonaceae*, más preferentemente del género *Xanthomonas*, muy preferentemente de *Xanthomonas axonopodis*.
- 35 8. Planta o parte de la misma, incluyendo semillas, obtenible mediante un procedimiento de acuerdo con cualquier reivindicación precedente, en la que dicha planta o parte de la misma comprende un ácido nucleico recombinante que codifica un polipéptido HpaG unido operativamente a un promotor GOS2 o a un promotor específico de tejido verde.
9. Construcción que comprende:
- 40 (a) ácido nucleico que codifica un polipéptido HpaG según se define en las reivindicaciones 1 a 4 o 7;  
(b) una o más secuencias de control capaces de dirigir la expresión de la secuencia de ácido nucleico de (a), comprendiendo dichas secuencias de control un promotor GOS2 o un promotor específico de tejido verde; y opcionalmente  
(c) una secuencia de terminación de la transcripción.
- 45 10. Construcción de acuerdo con la reivindicación 9, en la que dicho promotor GOS2 es de arroz o en la que dicho promotor específico de tejido verde es un promotor de la protoclorofilida reductasa o un promotor de la protoclorofilida reductasa de arroz.
11. Uso de una construcción de acuerdo con la reivindicación 9 o 10 en un procedimiento para preparar plantas que tienen biomasa aumentada y/o rendimiento de semilla aumentado con respecto a las plantas de control, en el que dichas plantas se crecen en condiciones sin estrés o condiciones de estrés abiótico.
- 50 12. Planta, parte de planta o célula vegetal transformada con una construcción de acuerdo con la reivindicación 9 o 10.

13. Procedimiento de producción de una planta transgénica que tiene biomasa aumentada y/o rendimiento de semilla aumentado con respecto a las plantas de control, que comprende:

- 5 (i) introducir y expresar en una planta un ácido nucleico que codifica un polipéptido HpaG según se define en la reivindicación 1 a 4 o 7; y  
(ii) cultivar la célula vegetal en condiciones sin estrés o condiciones de estrés abiótico,

en el que dicho rendimiento aumentado no es resultado de la resistencia aumentada al estrés biótico.

10 14. Planta transgénica que tiene biomasa aumentada y/o rendimiento de semilla aumentado, con respecto a las plantas de control, que es el resultado de la expresión aumentada de un ácido nucleico que codifica un polipéptido HpaG según se define en la reivindicación 1 a 4 o 7 unido operativamente a un promotor GOS2 o a un promotor específico de tejido verde, o una célula vegetal transgénica obtenida de dicha planta transgénica

15. Planta transgénica de acuerdo con la reivindicación 8, 12 o 14, o una célula vegetal transgénica obtenida de la misma, en la que dicha planta es una planta de cultivo o una monocotiledónea o un cereal, tal como arroz, maíz, trigo, cebada, mijo, centeno, sorgo o avena.

15 16. Partes cosechables de una planta de acuerdo con la reivindicación 15, en las que dichas partes cosechables son preferentemente semillas y en las que dichas partes cosechables comprenden un ácido nucleico que codifica un polipéptido HpaG según se define en la reivindicación 1 a 4 o 7 unido operativamente a un promotor GOS2 o a un promotor específico de tejido verde.

20 17. Uso de un ácido nucleico que codifica un polipéptido HpaG según se define en la reivindicación 1 a 4 o 7 en el aumento del rendimiento de semilla y/o de la biomasa, en plantas con respecto a las plantas de control, en el que dicho rendimiento de semilla y/o biomasa aumentado no es el resultado de la resistencia aumentada al estrés biótico o en el que dicho rendimiento de semilla y/o biomasa aumentado se obtiene en condiciones sin estrés o de estrés abiótico.

Alineamiento de múltiples secuencias por CLUSTAL W (1.83)

```

ABJ97680      MNSLNTQFGGTSNLQVGPSQ--DTTFGSNQG--GNQGISEKQLDQLLQCLISALLQSSK
AAC95121      MNSLNTQFGGTSNLQVGPSQ--DTTFGSNQG--GNQGISEKQLDQLLQCLISALLQSSK
BAD29979      MNSLNTQFGGTSNLQVGPSQ--DTTFGSNQG--GNQGISEKQLDQLLQCLISALLQSSK
ABB72197      MNSLNTQFGGSASNFQVDQSQ--NAQSDSSQGSNGSQGISEKQLDQLLQCLIQALLQPNK
SEQID2_      MNSLNTQLGANSSEFFQVDPGQ--NTQSS--PNQGNQGISEKQLDQLLQCLIMALLQSSN
ABK51590      MNSLNTQLGANSSEFFQVDPGQ--NTQSS--PNQGN--PNQGN-----TQLIMALLQSSN
ABK51589      MNSLNTQLGANSSEFFQVDPGQ--NTQSS--PNQGNQGISEKQLDQLLQCLIMALLQSSN
ABK51587      MNSLNTQLGANSSEFFQVDPGQ--NTQSS--PNQGNQGISEKQLDQLLQCLIMALLQSSN
ABK51588      MNSLNTQLGANSSEFFQVDPGQ--NTQSS--PNQGNQGISEKQLDQLLQCLIMALLQSSN
AAM35307      MNSLNTQLGANSSEFFQVDPGQ--NTQSG--SNQGNQGISEKQLDQLLQCLIMALLQSSN
ABG36696      MNSLNTQLGANSSEFLQVDPGQ--NTQFG--PNQGNQGISEKQLDQLLQCLIMALLQSSN
AAM40538      --MDSSIGNKFSNFINLQTMGIGPQQQTQNSSQSPSADSEQQLDQLLAMFTMMMLQSSQ
                :*: * :
                . .
                .

```

FIGURA 1





|          |                                 |
|----------|---------------------------------|
| ABJ97680 | GGFGGDFS-----GDLGLGTLNLSSDSASMQ |
| AAC95121 | GGFGGDFS-----GDLGLGTLNLSSDSASMQ |
| BAD29979 | GGFGGDFS-----GDLGLGTLNLSSDSASMQ |
| ABB72197 | GSEASSFS-----NDSGSMQ-----       |
| SEQID2   | GFGGILVT-----SLASDTGSMQ-----    |
| ABK51590 | GFGGILVT-----SLASDTGSMQ-----    |
| ABK51589 | GFGGILVT-----SLASDTGSMQ-----    |
| ABK51587 | GFGGILVT-----SLASDTGSMQ-----    |
| ABK51588 | --GFILVT-----SLASDTGSMQ-----    |
| AAM35307 | GFGGGLGTSLGFTSLASDTGSMQ-----    |
| ABG36696 | GFGGGLGTSLGSSLASDTGSMQ-----     |
| AAM40538 | GGFNANLS-----SITGQA-----        |

: . \*

**FIGURA 1 (continuación)**

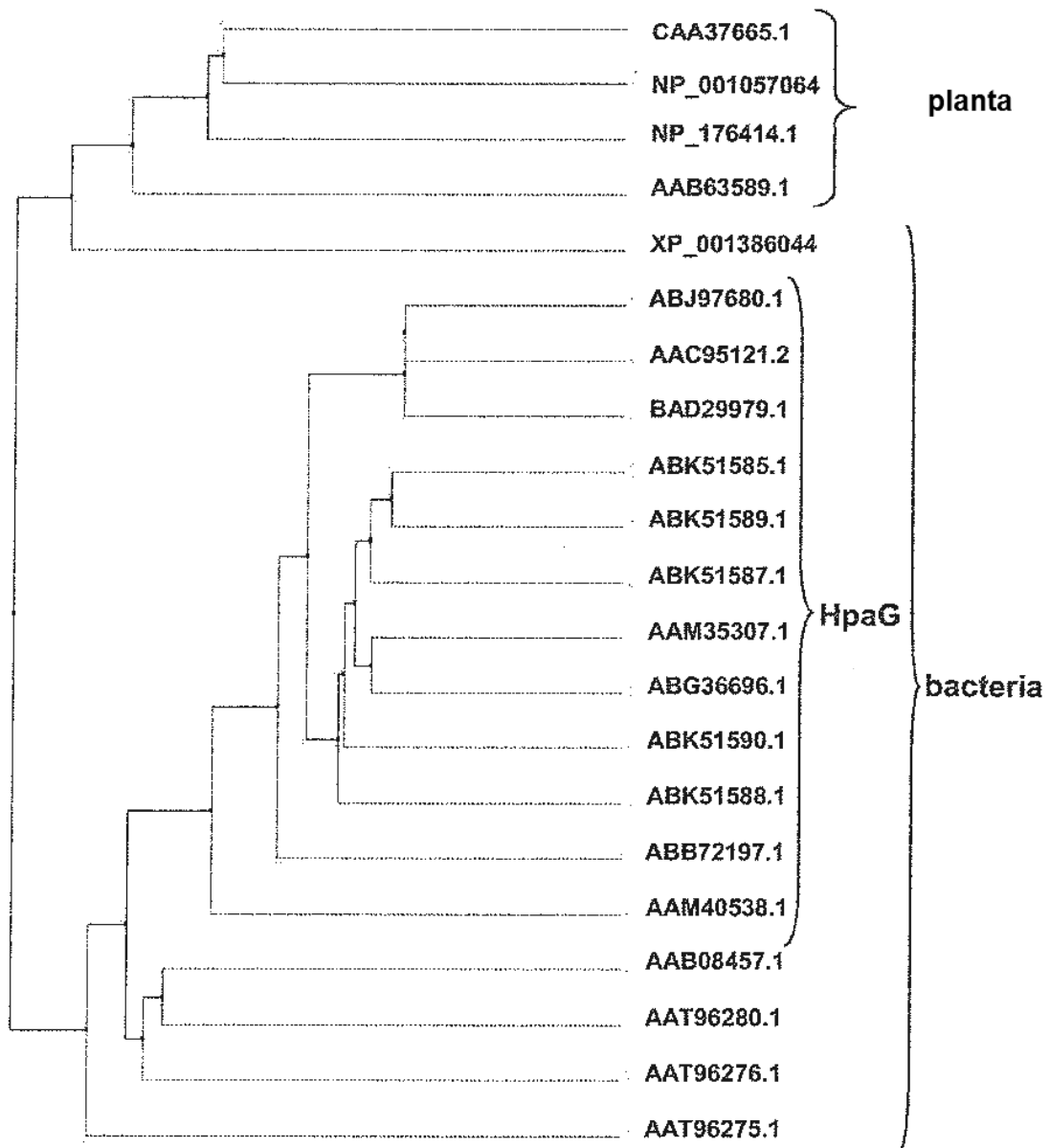


FIGURA 2

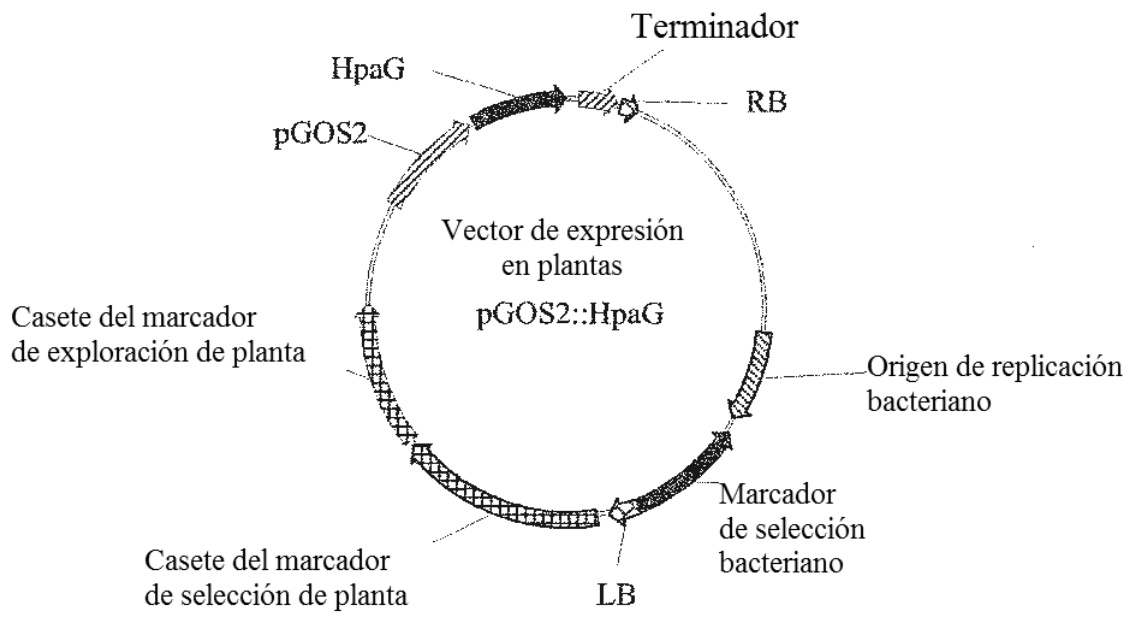


FIGURA 3

**SEQ ID NO: 1, EF050509.1, gen de HpaG (hpaG) desencadenante de la respuesta de hipersensibilidad en *Xanthomonas axonopodis*, cds completo**

ATGAATTCTTTGAACACACAGCTCGGGCGCAACTCGTCCTTCTTTCAGGTTGACCCCGGCCAGAAC  
 ACGCAATCTAGTCCGAACCAGGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGCTGCTGACC  
 CAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAATGCCGAGCAGGGTCAGGGTCAAGGCCAGGGT  
 GGTGACTCTGGCGGTGAGGGCGGCAATCCGCGGCAGGCCGGGCAGTCCAACGGCTCCCCCTCGCAA  
 TACACCCAGGCGCTGATGAATATCGTCGGAGACATTCTCCAGGCGCAGAATGGTGGCGGCTTCGGC  
 GCGGCTTTGGTGGTGGCTTCGGTGGCATCCTCGTCAACCAGCCTTGGCAGCGACACCGGATCGATG  
 CAGTAA

**SEQ ID NO: 2, ABK51582.1, HpaG desencadenante de la respuesta de hipersensibilidad**

[*Xanthomonas axonopodis*]

MNSLNTQLGANSSFFQVDFGQNTQSSPNQGNQGI SEKQLDQLLTQLIMALLQQSNNAEQGQGGQGG  
 GDSGGQGGNPRQAGQSNQSPSQTQALMNIVGDILQAQNGGGFGGGFGGFGGILVTSLASDTGSM  
 Q

**SEQ ID NO: 3, motivo conservado 1**

G (G/E/D) (N/E) X (Q/R/P) Q (A/S) GX (N/D) G

**SEQ ID NO: 4, motivo conservado 2**

(P/A/V) S (P/Q/A) (F/L/Y) TQ (M/A) LM (H/N/Q) IV (G/M) (E/D/Q)

**SEQ ID NO: 5, promotor constitutivo GOS2**

AATCCGAAAAGTTTTCTGCACCGTTTTCCACCCCTAACTAACAATATAGGGAAACGTGTGCTAAATAT  
 AAAATGAGACCTTATATATGTAGCGCTGATAACTAGAACTATGCAAGAAAACTCATCCACCTACT  
 TTAGTGGCAATCGGGCTAAATAAAAAAGAGTCGCTACACTAGTTTCGTTTTCTTAGTAATTAAGT  
 GGGAAAATGAAATCATTATTGCTTAGAATATACGTTACATCTCTGTCATGAAGTTAAATTATTCCG  
 AGGTAGCCATAATTGTCATCAAACCTTCTTGAATAAAAAAATCTTCTAGCTGAACTCAATGGGT  
 AAAGAGAGAGATTTTTTTTAAAAAATAGAATGAAGATATCTGAAAGTATTGGCAAAGATTTAAA  
 CATATAATTATATAATTTTATAGTTTGTGCATTCGTCATATCGCACATCATTAAAGGACATGTCTTA  
 CTCCATCCCAATTTTTATTTAGTAATTAAGACAATTGACTTATTTTTATTATTTATCTTTTTTCG  
 ATTAGATGCAAGGTACTTACGCACACACTTTGTGCTCATGTGCATGTGTGAGTGCACCTCCTCAAT  
 ACACGTTCAACTAGCAACACATCTCTAATATCACTCGCCTATTTAATAACATTTAGGTAGCAATATC  
 TGAATTCAGCACTCCACCATCACCAGACCACTTTTAATAATATCTAAAATACAAAAAATAATTTT  
 ACAGAATAGCATGAAAAGTATGAAAACGAACTATTTAGGTTTTTACATACAAAAAAAGAAATT  
 TTGCTCGTGCGGAGCGCAATCTCCCATATTGGGCACACAGGCAACAACAGAGTGGCTGCCACA  
 GAACAACCCACAAAAACGATGATCTAACGGAGGACAGCAAGTCCGCAACAACCTTTTAACAGCAG  
 GCTTTGCGGCCAGGAGAGAGGAGGAGGCAAGAAAACCAAGCATCCTCCTTCTCCCATCTATAA  
 ATTCTCCCCCTTTTCCCCTCTCTATATAGGAGGCATCCAAGCCAAGAAGAGGGAGAGCACC  
 GACACGCGACTAGCAGAAGCCGAGCGACCGCCTTCTCGATCCATATCTTCCGGTCGAGTTCTTGGT  
 CGATCTCTTCCCTCCTCCACCTCCTCCTCACAGGGTATGTGCCTCCTTCGGTTGTTCTTGGATTT  
 ATTTGTTCTAGGTTGTGTAGTACGGGCGTTGATGTTAGGAAAGGGGATCTGTATCTGTGATGATCC  
 TGTCTTTGGATTTGGGATAGAGGGGTTCTTGATGTTGCATGTTATCGGTTCCGGTTGATTAGTAGT  
 ATGGTTTTCAATCGTCTGGAGAGCTCTATGGAAATGAAATGGTTTTAGGGATCGGAATCTTGCATT  
 TTGTGAGTACCTTTTGTGTTGAGGTAATAATCAGAGCACCAGGTTGTTTTGCTTGGTGTAAATAAGTAC  
 GGTGTTTTGGTCCCTCGATTCTGGTAGTGATGCTTCTCGATTTGACGAAGCTATCCTTTGTTTATTC  
 CCTATTGAACAAAAATAATCCAACCTTTGAAGACGGTCCCGTTGATGAGATTGAATGATTGATTCCT  
 AAGCCTGTCCAAAATTTCCGAGCTGGCTTGTGTTAGATACAGTAGTCCCATCACGAAATTCATGGA

**FIGURA 4**

AACAGTTATAATCCTCAGGAACAGGGGATTCCTGTTCTCCGATTTGCTTTAGTCCCAGAATTTT  
 TTTTCCCAAATATCTTAAAAAGTCACTTTCTGGTTCAGTTCAATGAATTGATTGCTACAAATAATG  
 CTTTTATAGCGTTATCCTAGCTGTAGTTCAGTTAATAGGTAATACCCCTATAGTTTAGTCAGGAGA  
 AGAACTTATCCGATTTCTGATCTCCATTTTAATTATATGAAATGAACTGTAGCATAAGCAGTATT  
 CATTGGATTATTTTTTTTATTAGCTCTCACCCCTTCATTATTCTGAGCTGAAAGTCTGGCATGAA  
 CTGTCCCAATTTGTTTTCAAATTCACATCGATTATCTATGCATTATCCTCTTGTATCTACCTGT  
 AGAAGTTTCTTTTGGTTATTCCTTGACTGCTTGATTACAGAAAGAAATTTATGAAGCTGTAATCG  
 GGATAGTTATACTGCTTGTCTTATGATTCATTTCCCTTGTGCAGTTCTTGGTGTAGCTTGCCACT  
 TTCACCAGCAAAGTTC

**SEQ ID NO: 6, PCR del promotor específico de tejido verde**

TTGCAGTTGTGACCAAGTAAGCTGAGCATGCCCTTAACTTCACCTAGAAAAAGTATACTTGGCTT  
 AACTGCTAGTAAGACATTTCAGAAGTGAAGTGGTGTACGCATTTTCATGCAAGCCATTACCACTTT  
 ACCTGACATTTTGGACAGAGATTAGAAATAGTTTTCGTACTACCTGCAAGTTGCAACTTGAAAAGTG  
 AAATTTGTTCTTGGCTAATATATTGGCGTGTAATTCCTTTTATGCGTTAGCGTAAAAAGTTGAAATT  
 TGGGTCAAGTTACTGGTCAGATTAACCAGTAAGTGGTTAAAGTTGAAAGATGGTCTTTTAGTAATG  
 GAGGGAGTACTACACTATCCTCAGCTGATTTAAATCTTATTCCGTCGGTGGTGATTTTCGTCAATCT  
 CCAACTTAGTTTTTCAATATATTCATAGGATAGAGTGTGCATATGTGTGTTTATAGGGATGAGTC  
 TACGCGCCTTATGAACACCTACTTTTGTACTGTATTTGTCAATGAAAAGAAAATCTTACCAATGCT  
 GCGATGCTGACACCAAGAAGAGGCGATGAAAAGTGAACGGATATCGTGCCACGTCGGTTGCCAAG  
 TCAGCACAGACCCAATGGGCCTTTCCTACGTGTCTCGGCCACAGCCAGTCGTTACCGCACGTTCA  
 CATGGGCACGAACTCGCGTCATCTTCCCACGCAAAACGACAGATCTGCCCTATCTGGTCCCACCCA  
 TCAGTGGCCACACCTCCCATGCTGCATTATTTGCGACTCCCATCCCGTCTCCACGCCCAAACAC  
 CGCACACGGGTGCGGATAGCCACGACCCAATCACACAACGCCACGTCACCATATGTTACGGGCAGC  
 CATGCGCAGAAGATCCCGCGACGTCGCTGTCCCCCGTGTGCGGTTACGAAAAAATATCCCACCACGT  
 GTCGCTTTCACAGGACAATATCTCGAAGGAAAAAATCGTAGCGGAAAATCCGAGGCACGAGCTGC  
 GATTGGCTGGGAGGCGTCCAGCGTGGTGGGGGGCCACCCCTTATCCTTAGCCCGTGGCGCTCCT  
 CGCTCCTCGGGTCCGTGTATAAATACCCTCCGGAAGTCACTCTTGTGGTCAACAACAGAAAGCAA  
 AAGGACACCAGAAACATAGTACACTTGAGCTCACTCCAAACTCAAACACTCACACCA

**SEQ ID NO: 7, EF042294, gen de HpaG\_T44C desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética, cds completo**

ATGAATTCTTTGAACACACAGCTCGGCGCCAACTCGTCTTCTTTTTCAGGTTGACCCCGGCCAGAAC  
 ACGCAATCTAGTCCGAACCAGGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGCTGCTGTGC  
 CAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAATGCCGAGCAGGGTCAGGGTCAAGGCCAGGGT  
 GGTGACTCTGGCGGTGAGGGCGGCAATCCGCGGCAGGCCGGGCAGTCCAACGGCTCCCCCTCGCAA  
 TACACCAGGGCGTGTGAATATCGTCCGAGACATTTCCAGGCGCAGAATGGTGGCGGCTTCGGC  
 GCGGCTTTGGTGGTGGCTTCGGTGGCATCCTCGTCAACCAGCCTTCCGAGCGACACCGGATCGATG  
 CAGTAA

**SEQ ID NO: 8, ABK51589, HpaG\_T44C desencadenante de la respuesta de hipersensibilidad mutante [construcción sintética]**

MNSLNTQLGANSSFFQVDPGQNTQSSPNQGNQGI SEKQLDQLLQQLIMALLQOSNNAEQGGQGGQ  
 GDSGGQGGNPRQAGQSNQSPSYTQALMNI VGDILQAQNGGGFGGGFGGFGLVTSLASDTGSM  
 Q

**FIGURA 4 (continuación)**

**SEQ ID NO: 9, EF042292, gen de HpaG-T desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética, cds completo**

ATGAATTCTTTGAACACACAGCTCGGCGCCAACCTCGTCCTTCTTTCAGGTTGACCCCGGCCAGAAC  
ACGCAATCTAGTCCGAACCAGGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGCTGCTGACC  
CAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAATGCCGAGCAGGGTCAGGGTCAAGGCCAGGGT  
GGTGACTCTGGCGGTGTCAGGGCGGCAATCCGCGGCAGGGCCGGGCAGTCCAACGGCTCCCCCTCGCAA  
TACACCCAGGCGCTGATGAATATCGTCGGAGACGGCTTCGGCGGCGGCTTTGGTGGTGGCTTCGGT  
GGCATCCTCGTCACCAGCCTTGCGAGCGACACCGGATCGATGCAGTAA

**SEQ ID NO: 10, ABK51587, HpaG-T desencadenante de la respuesta de hipersensibilidad mutante [construcción sintética]**

MNSLNTQLGANSSFFQVDPGQNTQSSPNQGNQGI SEKQLDQLLTQLIMALLQOSNNAEQGQGQGG  
GDSGGQGGNPRQAGQSNQSPSYTQALMNI VGDGFGGGFGGGFGGILVTSLASDTGSMQ

**SEQ ID NO: 11, 21106495:2613-3026 *Xanthomonas axonopodis* pv. *citri* cep. 306, sección 45 de 469 del genoma completo**

TTACTGCATCGATCCGGTGTGCTCGCAAGGCTGGTGCCGAGGCTGGTGCCGAGGCCGCCGCCGAA  
GCCACCACCAAAGCCGCCGCCGAAGCCACCACCATTCTGCGCCTGGAGAATGTCTCCGACGATATT  
CATCAGCATCTGGGTGTATTGCGAGGGGGAGCCGTTGGACTGACCGGCTGCTGCCGATTGCCGCC  
CTGACCACCAGAGTCAACCACCCTGGCCTTGACCCTGACCCTGCTCGGCATTGTTGCTCTGCTGAAG  
CAGGGCCATGATGAGCTGGGTGAGCAGCTGGTCCAGTTGCTTTTCCGAGATGCCCTGGTTGCCCTG  
GTTCCGAACCAGATTGCGTGTCTGGCTGGGGTCAACCTGAAAGAAGGACGAGTTGGCGCCGAGCTG  
TGTGTTCAAAGAATTCAT

**SEQ ID NO: 12, AAM35307, proteína Hpa1 [*Xanthomonas axonopodis* pv. *citri* cep. 306]**

MNSLNTQLGANSSFFQVDPSQNTQSGSNQGNQGI SEKQLDQLLTQLIMALLQOSNNAEQGQGQGG  
GDSGGQGGNRQAGQSNQSPSYTQMLMNI VGDILQAQNGGGFGGGFGGGLTSLGTSLASD  
TGSMQ

**SEQ ID NO: 13, EF042295, gen de HpaG-N desencadenante de la respuesta de hipersensibilidad mutante de construcción sintética, cds completo**

ATGAATTCTTTGAACACACAGCTCGGCGCCAACCTCGTCCTTCTTTTCAGGTTGACCCCGGCCAGAAC  
ACGCAATCTAGTCCGAACCAGGGCAACACCAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAAT  
GCCGAGCAGGGTCAGGGTCAAGGCCAGGGTGGTACTCTGGCGGTGAGGGCGGCAATCCGCGGCAG  
GCCGGGCAGTCCAACGGCTCCCCCTCGCAATACACCAGGCGCTGATGAATATCGTCGGAGACATT  
CTCCAGGCGCAGAATGGTGGCGGCTTCGGCGGCGGCTTTGGTGGTGGCTTCGGTGGCATCCTCGTC  
ACCAGCCTTGCGAGCGACACCGGATCGATGCAGTAA

**SEQ ID NO: 14, ABK51590, HpaG-N desencadenante de la respuesta de hipersensibilidad mutante [construcción sintética]**

MNSLNTQLGANSSFFQVDPGQNTQSSPNQGNQGI SEKQLDQLLTQLIMALLQOSNNAEQGQGQGG  
GDSGGQGGNPRQAGQSNQSPSYTQALMNI VGDILQAQNGGGFGGGFGGILVTSLASDTGSMQ

**SEQ ID NO: 15, EF042293, gen HpaG\_G de *Xanthomonas axonopodis*, cds completo**

ATGAATTCTTTGAACACACAGCTCGGCGCCAACCTCGTCCTTCTTTCAGGTTGACCCCGGCCAGAAC  
ACGCAATCTAGTCCGAACCAGGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGCTGCTGACC

**FIGURA 4 (continuación)**

CAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAATGCCGAGCAGGGTCAGGGTCAAGGCCAGGGT  
GGTGACTCTGGCGGTTCAGGGCGGCAATCCGCGGCAGGCCGGGCAGTCCAACGGCTCCCCCTCGCAA  
TACACCCAGGCGCTGATGAATATCGTCGGAGACATTCTCCAGGCGCAGAATGGCTTTATCCTCGTC  
ACCAGCCTTGCGAGCGACACCGGATCGATGCAGTAA

**SEQ ID NO: 16, ABK51588, HpaG\_G [*Xanthomonas axonopodis*]**

MNSLNTQLGANSSFFQVDPGQNTQSSPNQGNQGI SEKQLDQLLTQLIMALLQQSNNAEQGQGQGG  
GDSGGQGGNPRQAGQSNGPSQYTOALMNIIVGDI LQAQNGFILLVTSLASDTGSMQ

**SEQ ID NO: 17, DQ643828, gen Hrp de *Xanthomonas smithii* subsp. *smithii*, cds completo**

ATGAATTCTTTGAACACACAGATCGGCGCCAACCTCGTCTTCTTGCAAGTTCGACCCGAGCCAGAAC  
ACGCAATTCGGTCCGAACCAGGGCAATCAAGGCATCTCGGAAAAGCAGCTGGACCAGCTGCTGACC  
CAGCTCATCATGGCCCTGCTTCAGCAGAGCAACAATGCCGACCAGGGTCAGGGTGGTACTCTGGT  
GGTCAAGGCGGCAATTCGCGGCAGGCCGGGCAGCCCAATGGTTCCTCCCTCGGCATACACCCAGATG  
CTGATGAATATCGTCGGAGACATTCTCCAGGCGCAGAATGGTGGTGGCTTCGGCGGCGGGTTCGGC  
GGTGGCTTTGGTGGCGGGCTCGGCACCAGCCTCGGCAGCAGCCTTGCGAGCGACACCGGATCGATG  
CAGTAA

**SEQ ID NO: 18, ABG36696, Hrp [*Xanthomonas smithii* subsp. *smithii*]**

MNSLNTQIGANSSFLQVDPSONTQFGPNQGNQGI SEKQLDQLLTQLIMALLQQSNNADQGGGDSG  
GQGGNSRQAGQPNGSPSAYTQMLMNIIVGDI LQAQNGGGFGGGFGGGFGGGLGTSLGSSLASDTGSM  
Q

**SEQ ID NO: 19, gi|116292746:1016-1435 grupo de genes hrp de *Xanthomonas oryzae* pv. *oryzae* cepa JXOIII, secuencia parcial**

ATGAACTCTTTGAACACACAATTCGGCGGCAGCACGTCCAACCTTCAGGTTGGCCCAAGCCAGGAC  
ACAACGTTTCGGTTCGAACCAGGGCGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGTTGCTG  
TGCCAGCTCATCTCGGCCCTGCTTCAGTCGAGCAAAAATGCTGAGGAGGGTAAGGGTCAGGGTGGC  
GATAATGGCGGTGGCCAGGGCGGCAATTCGCAGCAGGCGGGCAGCAGAATGGCCCCTCGCCATTC  
ACCCAGATGCTGATGCATATCGTCGGAGAGATTCTCCAGGCGCAGAATGGTGGTGGTGGTGGTGGC  
GGCGGTTTCGGCGGCGGGTTCGGCGGCGACTTTAGTGGCGACCTCGGCCTCGGCACCAACCTCTCG  
AGCGACAGCGCATCAATGCAGTAA

**SEQ ID NO: 20, ABJ97680, factor A de funcionamiento de la respuesta de hipersensibilidad [*Xanthomonas oryzae* pv. *oryzae*]**

MNSLNTQFGGSTSNLQVGPQDFTFGSNQGGNQGI SEKQLDQLLCQLISALLQSSKNAEEGKGGGG  
DNGGGQGGNSQQAGQONGPSPFTQMLMHIVGEILQAQNGGGAGGGGFGGGFGGDFSGDLGLGTNLS  
SDSASMQ

**SEQ ID NO: 21, gi|42717988:1136-1555 grupo de genes hrp de *Xanthomonas oryzae* pv. *oryzae*, secuencia parcial**

ATGAATTCTTTGAACACACAATTCGGCGGCAGCACGTCCAACCTTCAGGTTGGCCCAAGCCAGGAC  
ACAACGTTTCGGTTCGAACCAGGGCGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGTTGCTG  
TGCCAGCTCATCTCGGCCCTGCTTCAGTCGAGCAAAAATGCTGAGGAGGGTAAGGGTCAGGGTGGC  
GATAATGGCGGTGGCCAGGGCGGCAATTCGCAGCAGGCTGGGCAGCAGAATGGCCCCTCGCCATTC  
ACCCAGATGCTGATGCATATCGTCGGAGAGATTCTCCAGGCGCAGAATGGTGGTGGTGGTGGTGGC  
GGCGGTTTCGGCGGCGGGTTCGGCGGTGACTTTAGTGGCGACCTCGGCCTCGGCACCAACCTCTCG  
AGCGACAGCGCATCGATGCAGTAA

**FIGURA 4 (continuación)**



**SEQ ID NO: 22, AAC95121.2 | Hpa1 [*Xanthomonas oryzae* pv. *oryzae*]**

MNSLNTQFGGSTSNLQVGPSQDTTFGSNQGNGQGISEKQLDQLLCQLISALLQSSKNAEEGKGQGG  
DNNGGGQGGNSQQAGQQNGPSPFTQMLMHIVGEILQAQNGGGAGGGGFGGGFGGDFSGDLGLGTNLS  
SDSASMQ

**SEQ ID NO: 23, gi|50428340:1138-1557 grupo de genes hrp de *Xanthomonas oryzae* pv. *oryzae*, cds completo**

ATGAATTCTTTGAACACACAATTCGGCGGCAGCACGTCCAACCTTCAGGTTGGCCCAAGCCAGGAC  
ACAACGTTCCGGTTCGAACCAGGGCGGCAACCAGGGCATCTCGGAAAAGCAACTGGACCAGTTGCTG  
TGCCAGCTCATCTCGGCCCTGCTTCAGTCGAGCAAAAATGCTGAGGAGGGTAAGGGTCAGGGTGGC  
GATAATGGCGGTGGCCAGGGCGGCAATTCGCAGCAGGCCGGGCAGCAGAATGGCCCCCTCGCCATTC  
ACCCAGATGCTGATGCATATCGTCGGAGAGATTCTCCAGGCGCAGAATGGTGGTGGTGGTGGTGGC  
GGCGGGTTCGGCGGCGGGTTCGGCGGTGACTTTAGTGGCGACCTCGGCCTCGGCACCAACCTCTCG  
AGCGACAGCGCATCGATGCAGTAA

**SEQ ID NO: 24, BAD29979, Hpa1 [*Xanthomonas oryzae* pv. *oryzae*]**

MNSLNTQFGGSTSNLQVGPSQDTTFGSNQGNGQGISEKQLDQLLCQLISALLQSSKNAEEGKGQGG  
DNNGGGQGGNSQQAGQQNGPSPFTQMLMHIVGEILQAQNGGGAGGGGFGGGFGGDFSGDLGLGTNLS  
SDSASMQ

**SEQ ID NO: 25, gi|82393799:1-378 gen hpaGXooc de *Xanthomonas oryzae* pv. *oryzicola*, cds completo**

ATGAATTCTTTGAACACACAATTCGGCGGCAGCGCGTCCAACCTTCAGGTTGACCAAAGCCAGAAC  
GCGCAATCCGATTTCGAGCCAGGGCAGCAATGGCAGCCAGGGTATCTCGGAAAAGCAACTGGACCAG  
TTGCTGTGCCAGCTCATCCAGGCCCTGCTTCAGCCGAACAAAAATGCTGAGGAAGGTAAGGGTTCAG  
CAGGGTGGCGAGAATAATCAGCAGGCCGGGAAGGAGAATGGCGCCTCGCCACTCACCCAGATGCTG  
ATGAATATCGTCGGAGAGATTCTCCAGGCGCAGAATGCCGGCGGCAGCAGCGGCGGCGACTTTGGT  
GGCAGTTTCGCCAGCAGCTTCTCGAACGACAGCGGATCGATGCAGTAA

**SEQ ID NO: 26, ABB72197, hpaGXooc [*Xanthomonas oryzae* pv. *oryzicola*]**

MNSLNTQFGGSASNFQVDQSQNAQSDSSQGSNGSQGISEKQLDQLLCQLIQALLQPNKNAEEGKGQ  
QGGENNOQAGKENGASPLTQMLMNIIVGEILQAQNAAGSSGGDFGGSFASSFSNDSGSMQ

**SEQ ID NO: 27, gi|21112286:70-435 *Xanthomonas campestris* pv. *campestris* cep. ATCC 33913, sección 131 de 460 del genoma completo**

TCAGGCTTGGCCGGTGATGCTCGACAGGTTGGCATTGAAGCCGCCACCCAAGCTGGTGGCGCCCAT  
GCCGGCGCCGCCTTGGTTCTGCATCAGCTGCATCACGATCTGCATCAGCATCTGCGTCAACGGACT  
CACACCGTCCTGTTGACCGCTCTGCGGTTGTTTCGTCTCCGCACTCCTGATCGGCATCGCTGCCCTG  
GCTCTGTTGGAGCATCATCATGATGAACATGGCGAGCAGCTGATCCAGCTGCTGCTCGGAGTCAGC  
CGAAGGCGAGCGCTGACTGGAGTTCTGGGTTTGGTGGGGCCCGATGCCCATCGTCTGCAGTTGAT  
GAAGTTGGAAAATTTGTTTCCGATAGATGAGTCCAT

**SEQ ID NO: 28, AAM40538, proteína Hpa1 [*Xanthomonas campestris* pv. *campestris* cep. ATCC 33913]**

MDSSIGNKFSNFINLQTMGIGPQQTQNSSQRSPSADSEQQLDQLLAFIMMMLQQSQGSADADQECG  
DEQPQSGQQDGVSPLTQMLMQIVMQLMQNGGAGMGGTSLGGGFNANLSSITGQA

**FIGURA 4 (continuación)**