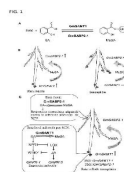


DIRECCIÓN DE NUEVAS CREACIONES

SOLICITUD FASE NACIONAL - PCT



1	Título de la Invención (200 caracteres o espacios máximos)		
DEFENSA SUPERIOR DE SOJA CONTA PLAGAS			
2	Datos del Solicitante / Titular		
Nombre:	UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION	Dirección Electrónica:	sic@castellanosyco.co
Dirección:	600 Henley Street, Suite 211, Knoxville, TN 37996-4122, ESTADOS UNIDOS DE AMERICA	Domicilio/País de constitución:	ESTADOS UNIDOS DE AMÉRICA - TENNESSEE - KNOXVILLE
Identificación:			

☐ CEDULA DE CIUDADANIA
☒ EMPRESA EXTRANJERA
☐ PASAPORTE

☐ CEDULA DE EXTRANJERIA
☐ NIT

Número:

516861772780-

3

Solicitantes

Apellidos - Nombres o Razón Social

Tipo

Identificación

1. UNIVERSITY OF TENNESSEE
RESEARCH FOUNDATION NN 516861772780

4

Datos del Inventor

Nombre:

NEAL C. STEWART

**Dirección
Electrónica:**

sic@castellanosyco.co

Dirección:

1600 Wheaton Place,
Knoxville, TN 37919,
ESTADOS UNIDOS
DE AMERICA

**Domicilio/País de
constitución:**

ESTADOS UNIDOS DE
AMÉRICA -
TENNESSEE -
KNOXVILLE

Identificación:

☐ CEDULA DE
CIUDADANIA
☐ EMPRESA
EXTRANJERA
☐ PASAPORTE

☐ CEDULA DE
EXTRANJERIA
☐ NIT
☒ No Aplica

Número:

-

5

Inventor(es)

Apellidos - Nombres

Domicilio

1. STEWART NEAL C. ESTADOS UNIDOS DE AMÉRICA
2. LIN JINGYU ESTADOS UNIDOS DE AMÉRICA
3. MAZAREI MITRA ESTADOS UNIDOS DE AMÉRICA
4. CHEN FENG ESTADOS UNIDOS DE AMÉRICA

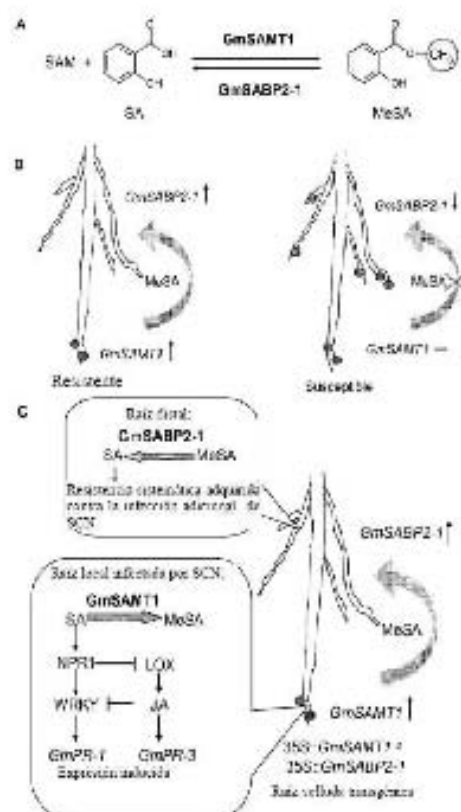
6

Datos Inventor(es)

País de Residencia	Departamento/Estado	Ciudad	Dirección	
1.	ESTADOS UNIDOS DE AMÉRICA	TENNESSEE	KNOXVILLE	1600 Wheaton Place, Knoxville, TN 37919, ESTADOS UNIDOS DE AMERICA
2.	ESTADOS UNIDOS DE AMÉRICA	TENNESSEE	KNOXVILLE	1504 Sutters Mill Lane, Knoxville, TN 37909, ESTADOS UNIDOS DE AMERICA
3.	ESTADOS UNIDOS DE AMÉRICA	TENNESSEE	KNOXVILLE	421 Anteelah Trail, Knoxville, TN 37919, ESTADOS UNIDOS DE AMERICA
4.	ESTADOS UNIDOS DE AMÉRICA	TENNESSEE	KNOXVILLE	1162 Keowee Ave., Knoxville, TN 37919, ESTADOS UNIDOS DE AMERICA
7	Datos del Representante Legal / Apoderado			
Nombre:		MARGARITA CASTELLANOS ABONDANO	Dirección Electrónica:	sic@castellanosyco.co
Dirección:		CALLE 94A 7A 09	Domicilio/País de constitución:	COLOMBIA - BOGOTA D.C. - BOGOTA D.C.
Identificación:				
☒ CEDULA DE CIUDADANIA		☐ CEDULA DE EXTRANJERIA		
☐ EMPRESA EXTRANJERA		☐ NIT		
☐ PASAPORTE				
Número:		39779654-		
Presentación de Poder				
Año de Radicación				
Número de Radicación				

8	Datos Solicitud: PCT / WO			
Número Solicitud:		PCT/US2012/062759	Fecha Solicitud:	31/10/2012
Número Publicacion:		WO 2013/101339	Fecha Publicacion:	04/07/2013
9	Declaraciones de prioridad		☒ SI ☐ NO	
1.	(33) País de origen ESTADOS UNIDOS DE AMÉRICA	Codigo del país US	(31) No. Solicitud 61/581,039	(32) Fecha _ 28/12/2011
10	Reivindicaciones			
Número reivindicaciones:		23	Pago Reivindicaciones:	Si
11	Reducción de tasas.			
<i>Declaro que carezco de medios económicos para presentar la solicitud de patente.</i> ☐ SI ☒ NO Nota: En caso de ser persona natural y carecer de medios económicos, y por lo tanto, aplique la reducción de tasas a que se refiere la resolución vigente en tarifas, debe firmar la presente solicitud bajo la gravedad de juramento.				
☐ Micro, pequeñas y medianas empresas ☐ Universidades públicas o privadas ☐ Entidades sin ánimo de lucro Debe aportar los documentos que se indican en el numeral 17 de anexos				
12	Documentos Anexos			
☒ Reivindicaciones ☒ Descripción ☐ Dibujos y/o Figuras ☒ Resumen ☐ Certificado Depósito Material Biológico ☐ Uso de Conocimiento tradicional ☒ Listado de secuencias ☒ Artes finales 12 x 12 cm ☒ Poderes, si fuere el caso ☐ Copia de la primera solicitud si se reivindica prioridad ☐ Traducción simple de la primera solicitud, si se reivindica prioridad ☒ Otros Anexos				

FIG. 1



DEFENSA SUPERIOR DE SOJA CONTRA PLAGAS

REFERENCIA CRUZADA A LA SOLICITUD RELACIONADA

La presente solicitud reivindica el beneficio de la solicitud de patente provisional de Estados Unidos No. 61/581,039, presentada el 28 de diciembre de 2011, cuyo contenido se incorpora en la presente a modo de referencia en su totalidad.

CAMPO DE LA DIVULGACIÓN

La presente divulgación se refiere al campo de biología molecular vegetal. Más específicamente, la presente divulgación se refiere a la tolerancia a patógenos conferida por la expresión transgénica de genes en la soja y otras plantas.

ANTECEDENTES

La soja [*Glycine max* (L.) Merr.] es la legumbre más ampliamente cultivada en el mundo, lo que proporciona una fuente importante de proteínas y aceite. La soja puede utilizarse de varias formas tal como ingredientes en la formulación de una gran cantidad de alimentos para humanos, alimentos para animales y productos industriales. Como la semilla oleaginosa domina el comercio mundial, la soja contribuye aproximadamente un 56% de la producción global de semillas oleaginosas. El valor estimado de los productos de soja es de aproximadamente 48.6 mil millones de dólares en el mundo y de aproximadamente 18.7 mil millones de dólares en los Estados Unidos. Por lo tanto, la soja es considerada como uno de los cultivos más importantes a nivel económico. Aunque la producción global de soja ha aumentado de forma constante durante la última década, la futura demanda de soja no puede satisfacerse aún debido a una población mundial en aumento y a fuentes limitadas de tierra.

Durante todo su ciclo de vida, la soja puede ser atacada por varios patógenos, tales como hongos, bacterias, virus y nematodos, y sufrir varias enfermedades en cualquier tejido. Por ejemplo, el patógeno bacteriano *Pseudomonas syringae* pv. *glycinea*, puede ocasionar plaga bacteriana, pasar el invierno en residuos de cultivos y puede transmitirse por semilla. En estudios de laboratorio, *P. syringae* se elige normalmente como el modelo de patógenos bacterianos. Los patógenos fúngicos también pueden ocasionar enfermedades de la soja en los tejidos superiores e inferiores, lo que representa aproximadamente un 50% de todas las pérdidas de enfermedad de soja en los Estados Unidos y alrededor del mundo. El virus del mosaico de la soja (SMV) detectado en semillas y en plantas en el campo, es el virus principal que ocasiona la enfermedad de la soja. Dado que esta enfermedad viral se transmite por la semilla, puede ocasionar problemas en la siguiente temporada de cultivo. Su efecto sobre la calidad de la semilla es la principal preocupación. La soja infectada con el virus puede producir semillas más

pequeñas, en menor cantidad y a veces moteadas. Otra preocupación es que la soja infectada con SMV tiende a estar doblemente infectada por otros virus, lo que aumenta el riesgo de pérdida de rendimiento y reduce aún más la calidad de la semilla. Además, las plagas de insectos de soja y nematodos tales como escarabajos, orugas, áfidos y ácaros también pueden ocasionar pérdida de rendimiento de soja.

El nematodo del quiste de la soja (*Heterodera glycines*, SCN), es la plaga que ocasiona el mayor daño de la soja a nivel económico en los Estados Unidos, lo que causa casi 1 mil millones de pérdidas en rendimiento. El SCN es un ascáride parásito vegetal pequeño y la mayoría de las etapas del SCN no pueden observarse a simple vista. El SCN se alimenta de las raíces de soja y roba nutrientes de la soja. Una característica importante, la heterogeneidad genética de poblaciones del SCN, dota este patógeno para sobrellevar fácilmente la resistencia. Para completar su ciclo vital, los juveniles de segunda etapa infecciosos (J2) ingresan a las raíces hospedadoras y migran intracelularmente dentro del tejido cortical al cilindro vascular. Después, los juveniles inician la formación de sitios de alimentación especializados llamados sincitios, cuya función como metabólico disminuye para alimentar a los nematodos. Los nematodos se alimentan exclusivamente desde sus sincitios ya que se desarrollan en hembras y machos adultos. Una vez fertilizado, el nematodo de soja hembra produce hasta varios cientos de huevos que, por lo general, se retienen dentro del útero del nematodo. Después de la muerte de la hembra, su cuerpo se desarrolla en un quiste protector alrededor de los huevos, lo que le otorga su nombre.

Las infecciones por SCN son difíciles de identificar en el campo ya que los síntomas por arriba del nivel del suelo de las infecciones por SCN se confunden fácilmente con síntomas de deficiencia de nutrientes. Cuando las plantas de soja son gravemente dañadas por nematodos, se vuelven atrofiadas y cloróticas. Estos síntomas pueden confundirse con deficiencia de nutrientes y tensión hídrica. Hasta que la población del SCN alcanza una determinada cantidad, los agricultores no pueden identificar su efecto dañino con claridad. En la actualidad, la forma más precisa para diagnosticar la infección por SCN es analizar una muestra de soja y observar la plaga directamente (Opperman y Bird 1998). Controlar el SCN en producciones comerciales de soja sigue siendo difícil ya que el SCN tiene un ciclo de vida corto y las poblaciones pueden construirse rápidamente. Los cambios frecuentes en la virulencia de población del SCN también contribuyen a la dificultad en el manejo de esta plaga. Además, los quistes del SCN pueden sobrevivir en el suelo por hasta nueve años y luego romperse para liberar los huevos en condiciones adecuadas, lo que aumenta la probabilidad de dispersión de

nematodos a través del suelo infectado.

Los métodos utilizados para controlar y manejar el SCN en la producción de soja incluyen la rotación de cultivos, la utilización de cultivares resistentes al SCN y la aplicación de nematicidas, que a menudo se utilizan de forma integrada. Sin embargo, estos enfoques siempre enfrentan restricciones económicas, consumen tiempo y el uso de nematicidas puede ocasionar problemas ambientales. Además, en algunas áreas, los factores económicos pueden limitar el uso de rotación de cultivos. Para considerar las demandas a largo plazo de soja, es importante desarrollar cultivares de soja resistentes mediante cultivo para manejar enfermedades de la soja. Aun para la rotación de cultivos, también se requieren cultivares de soja resistentes. De este modo, existe una necesidad de creación de plantas, tales como soja, que sean resistentes a plagas, tal como resistentes a infecciones por SCN. Esta divulgación cumple con esa necesidad.

SUMARIO DE LA DIVULGACIÓN

Se divulgan genes con resistencia a plagas no reconocidos previamente aislados de la soja. En una realización, el gen resistente a las plagas codifica una carboxil metiltransferasa dependiente de S-adenosil-L-metionina (SAM) y se conoce como *GmSAMT1*. En otra realización, el gen resistente a las plagas codifica una esterasa de salicilato de metilo y se conoce como *GmSABP2-1*.

En varias realizaciones, el gen resistente a las plagas aislado de la soja incluye una secuencia de ácidos nucleicos al menos un 80% idéntica a la SEQ ID NO: 1 o SEQ ID NO: 2 o una variante degenerada de esta, o un fragmento de esta que codifica una carboxil metiltransferasa dependiente de S-adenosil-L-metionina (SAM) activa conocida como *GmSAMT1* o esterasa de salicilato de metilo conocida como *GmSABP2-1*. Los ácidos nucleicos aislados que codifican dichos genes se divulgan como proteínas codificadas por dichos genes. Los ácidos nucleicos aislados pueden unirse de forma operativa al promotor. También se divulgan vectores de expresión y construcciones que incluyen dichos ácidos nucleicos. En algunas realizaciones, el vector de expresión es pJL-OFP-35S::*GmSAMT1* o pJL-OFP-35S::*GmSABP2-1*.

Las construcciones divulgadas proporcionan una resistencia mejorada a las plagas, tal como infección por nematodo del quiste de la soja. De este modo, se divulgan los métodos para producir una planta transgénica que tiene resistencia mejorada a las plagas, tal como infección por nematodo del quiste de la soja. También se divulgan plantas transgénicas, células o tejidos vegetales (tales como plantas dicotiledóneas o monocotiledóneas, células o tejidos vegetales) transformados con las construcciones divulgadas. También se proporcionan semillas vegetales, frutas, hojas, raíces, brotes,

flores, esquejes y otro material de reproducción útil para la propagación sexual o asexual, plantas de progenie incluidos híbridos de F1, plantas estériles macho y otras plantas y productos vegetales derivados de la planta transgénica divulgada. También se proporcionan en la presente métodos para producir las plantas transgénicas, células o tejidos vegetales divulgados.

Las características y ventajas anteriores y otras resultarán más evidentes a partir de la siguiente descripción detallada de varias realizaciones, lo que continúa con referencia a las figuras adjuntas.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

Las FIG. 1A-1C son un conjunto de esquemas que muestra un modelo propuesto de resistencia a la soja contra el SCN y las funciones de GmSAMT1 y GmSABP2-1. 1A muestra las reacciones entre SA y MeSA que son catalizadas por GmSAMT1 y GmSABP2-1. SAM representa S-adenosil-L-metionina que funciona como el donante de un grupo metilo. 1B muestra que en la raíz de soja resistente, la expresión génica de *GmSAMT1* y *GmSABP2-1* se regula positivamente cuando la soja es infectada por el SCN. El transporte eficaz de la señal móvil MeSA aumenta la resistencia de la soja contra el SCN. En la raíz de la soja susceptible, la expresión génica de GmSAMT1 no cambia y la expresión génica de GmSABP2-1 se regula negativamente cuando la soja es infectada por el SCN. El transporte de la señal móvil MeSA está bloqueado, lo que disminuye la resistencia de la soja contra el SCN. Hay más SCN que completan su ciclo de vida en la raíz de la soja susceptible en comparación con la resistente. 1C muestra que en la raíz de soja local infectada con el SCN, SA se biosintetiza y el nivel de SA aumenta. El nivel de SA puede manipularse por GmSAMT1 mediante conversión a la señal móvil MeSA. El nivel residual de SA es suficiente para inducir la expresión génica de *GmPR-1*. El nivel disminuido de SA beneficia la senda de defensa que depende de ácido jasmónico (JA), lo que induce la expresión génica de *GmPR-3*. En la raíz de soja no dañada, MeSA puede catalizarse por GmSABP2-1 para formar SA, lo que juega un papel de resistencia adquirida sistémica contra la infección de SCN adicional. Ningún expresor de genes relacionados con patogénesis 1 (NPR1), factor de transcripción de la familia WRKY y lipoxigenasa (LOX) puede funcionar como punto de regulación. Los puntos representan los SCN hembra adultos dentro de las raíces de soja.

La FIG. 2 muestra la alineación de la secuencia de la sonda de Affymetrix, Gma.12911.1.A1_S_AT (SEQ ID NO: 17) con cinco genes candidatos de la soja (SEQ ID NOS: 18-22). La secuencia de la sonda de Affymetrix, Gma.12911.1.A1_S_AT se alinea con el extremo de la secuencia de marco de lectura abierto de los genes de soja y cubre

la parte de la secuencia UTR 3' (TAA en el recuadro es el codón de detención). Con base en la alineación, únicamente *Glyma02g06070* tiene una secuencia idéntica a la sonda (indicado en el recuadro).

La FIG. 3 muestra un árbol filogenético de Glyma02g06070 y las proteínas de la familia SABATH funcionalmente caracterizadas de otras plantas. Un árbol de unión circundante con base en el grado de similitud de secuencia entre Glyma02g06070 y SABATH funcionalmente caracterizado que se selecciona de otra especie vegetal que utiliza el programa ClustalX. Glyma02g06070 se encuentra en la cepa SAMT mostrada en el óvalo. CbSAMT, *Clarkia breweri* SAMT (AF133053); AmSAMT, *Antirrhinum majus* (snapdragon) SAMT (AF515284); SfSAMT, *Stephanotis floribunda* SAMT (AJ308570); HcSAMT, *Hoya carnosa* SAMT (AJ863118); DwSAMT, *Datura wrightii* SAMT (EF472972); AmBAMT, *Antirrhinum majus* BAMT (AF198492); NsBSMT, *Nicotiana suaveolens* BSMT (AJ628349); AtBSMT, *Arabidopsis thaliana* BSMT (BT022049); AIBSMT, *A. lyrata* BSMT (AY224596); AbSAMT, *Atropa belladonna* SAMT (AB049752); OsBSMT1, *Oryza sativa* BSMT1 (XM467504); PhBSMT, *Petunia hybrida* BSMT (AY233465); OslAMT1, *O. sativa* IAMT1 (EU375746); PtiAMT1, *Populus trichocarpa* IAMT1 (XP_002298843); AtIAMT, *A. thaliana* IAMT (AK175586); AtGAMT1, *A. thaliana* GAMT1 (At4g26420); AtGAMT2, *A. thaliana* GAMT2 (At5g56300); CaCaS1, *Coffea arabica* sintasa de cafeína 1 (AB086414); CaXMT1, *C. arabica* XMT1 (AB048793); CaDXMT1, *C. arabica* DXMT1 (AB084125). Las cadenas se extrajeron para modificar la escala de la barra que indica 0.1 sustituciones por sitio.

La FIG. 4 es una imagen digital de un gel SDS-PAGE de proteína GmSAMT1 recombinante purificada. GmSAMT1 marcado con His expresado en *E. coli* se purificó tal como se describió en el ejemplo 1. La vía M contenía marcadores de peso molecular de proteínas. La vía 1 contenía extracto bruto y la vía 2 contenía aproximadamente 2 µg de proteína GmSAMT1 purificada. El gel se tiñó con azul Coomassie.

Las FIG. 5A-5D son un conjunto de gráficas que muestran las propiedades bioquímicas de GmSAMT1. La FIG. 5A muestra el efecto del pH sobre la actividad de GmSAMT1. El nivel de actividad de GmSAMT1 en tampón Tris-HCl de 50 mM, pH 7.5, se estableció arbitrariamente en 1.0. La FIG. 5B muestra la actividad de GmSAMT1 en respuesta a la temperatura que varía de 0 °C a 60 °C. El nivel de actividad de GmSAMT1 a 25 °C se estableció arbitrariamente en 1.0. La FIG. 5C muestra los efectos de los iones metálicos sobre la actividad de GmSAMT1. Se agregaron iones metálicos a las reacciones en la forma de sales de cloruro en concentraciones finales de 5 mM. El nivel de actividad de GmSAMT1 sin ningún ion metálico agregado como control (Ctr) se

estableció arbitrariamente en 1.0. La FIG. 5D muestra que GmSAMT1 presentó una cinética parcial de Michaelis-Menten, con un valor aparente de K_m de $46.2 \pm 4.2 \mu M$ para el ácido salicílico.

La FIG. 6 es un esquema que muestra la construcción de pJL-OFP-35S::GmSAMT1, un plásmido para la sobreexpresión del gen *GmSAMT1* en raíces de soja. Se modificó pJL-OFP de pCAMBIA 1305.2 mediante inserción de terminador *ppor RFP-NOS* de promotor 35S en el sitio *Nocl*. *ppor RFP* codifica una proteína fluorescente anaranjada. pJL-OFP es un plásmido binario compatible con *Agrobacteria* para la transformación vegetal y se utiliza para un plásmido de control de vector. pJL-OFP-35S::GUS es un vector intermediario y se utiliza para analizar si *OFP* y *GUS* dentro de dos casetes se coexpresan en raíces de soja. pJL-OFP-35S::GmSAMT1 se obtiene mediante el reemplazo del gen *GUS* con *GmSAMT1* a través de digestión y ligadura entre los sitios *BamH I* y *Sac I*.

Las FIG. 7A y 7B son imágenes digitales que muestran una prueba de expresión génica reportera en hojas de tabaco y raíces velludas de soja. La FIG. 7A muestra la expresión transitoria de pJL-OFP y del vector vacío en hojas de tabaco. La FIG. 7B muestra la coexpresión de dos genes reporteros en raíces velludas transgénicas. Todas las raíces velludas transgénicas analizadas con señal de OFP mostraron tinción de GUS. Se muestra al comienzo una representación esquemática de la construcción utilizada para la coexpresión de un gen reportero de proteína fluorescente anaranjada (OFP) y un gen reportero GUS. "35S-Pro" y "NOS-ter" representan el promotor CaMV 35S y el terminador NOS, respectivamente. Análisis de raíces velludas transgénicas transformadas con la construcción (inicio). Las raíces velludas transgénicas positivas de OFP se separaron de las raíces velludas negativas de OFP en el filtro OFP. Las raíces velludas positivas de OFP demostraron además ser positivas para la tinción de GUS.

Las FIG. 8A y 8B muestran la producción de raíces velludas transgénicas de soja que sobreexpresan *GmSAMT1*. La FIG. 8A. es una representación esquemática de la construcción utilizada para la coexpresión de una proteína fluorescente anaranjada (OFP) y *GmSAMT1*. "35S-Pro" y "NOS-ter" representan el promotor CaMV 35S y el terminador NOS, respectivamente. La FIG. 8B. muestra la expresión de *GmSAMT1* en cinco tipos de raíces velludas transgénicas a partir de tres antecedentes genéticos TN02-226, TN02-275 y Williams 82. "GmSAMT1" representa raíces velludas transgénicas que contienen el transgen *GmSAMT1*. "VC" representa raíces velludas transgénicas producidas que utilizan un vector de control que contiene marcador OFP pero no el gen *GmSAMT1*. qRT-PCR se llevó a cabo con cebadores específicos de *GmSAMT1*. Los valores de expresión

se normalizaron hasta los niveles de expresión del gen de ubiquitina 3 de soja (*GmUBI3*) en muestras respectivas. El nivel de expresión de *GmSAMT1* en las raíces velludas transgénicas de “VC” de TN02-275 se estableció arbitrariamente en 1.0. Cada barra representa el nivel de expresión relativo promedio de tres experimentos independientes con errores estándar, donde cada uno contiene un grupo de cinco plantas. Las barras con asteriscos son significativamente diferentes (LSD $p < 0.01$).

Las FIG. 9A y 9B son un conjunto de gráficas de barra que muestran análisis de qRT-PCR en raíces velludas transgénicas para *GmPR-1* y *GmPR-3*. “*GmSAMT1*” representa raíces velludas transgénicas que contienen el transgen *GmSAMT1*. “VC” representa raíces velludas transgénicas producidas que utilizan un vector de control que contiene marcador OFP pero no el gen *GmSAMT1*. Cada barra representa el nivel de expresión relativo promedio de tres experimentos independientes con errores estándar. En las gráficas mostradas en la FIG. 9A y FIG. 9B, las barras con letras diferentes son significativamente diferentes (LSD $p < 0.05$).

La FIG. 10 es un conjunto de imágenes digitales que muestra las etapas de J2, J3 y J4 de nematodos del quiste de la soja mostrados en raíces velludas transgénicas mediante despeje y tinción con fucsina ácida.

La FIG. 11 es una gráfica de barras y un conjunto de imágenes digitales que muestra un ensayo del SCN entre raíces velludas transgénicas con sobreexpresión de *GmSAMT1* y raíces peludas de control. Cada barra representa el promedio del índice de resistencia con error estándar de 10 raíces velludas de soja. El índice de resistencia es la relación de la suma de los nematodos J3 y J4 con respecto al total de nematodos hembra infecciosos mostrados en raíces velludas transformadas en *A. rhizogenes* después de dos semanas de inoculación del SCN. VS y VR representan las líneas de raíces velludas que albergan el plásmido pJL-OFP en el ámbito susceptible (TN02-275 y Williams 82) y resistente (TN02-226), lo que es útil como líneas de control susceptibles y resistentes, respectivamente. SAMT representa la línea de la raíz velluda con sobreexpresión de *GmSAMT1* y gen reportero OFP en el ámbito susceptible. Las barras con asteriscos son significativamente diferentes (LSD $p < 0.01$). La última foto muestra las etapas representativas del SCN en las líneas de raíces velludas transgénicas con respuesta.

Las FIG. 12A y 12B son gráficas de barras que muestran abundancia de transcripción relativa de *GmICS1* (**12A**) y *GmICS2* (**12B**) en raíces velludas transgénicas con o sin infección por SCN. Se produjeron y analizaron dos tipos de raíces velludas transgénicas del ámbito Williams 82. “*GmSAMT1*” representa raíces velludas transgénicas que contienen el transgen *GmSAMT1*. “VC” representa raíces velludas transgénicas

producidas que utilizan un vector de control que contiene el gen marcador *OFP* pero no el *GmSAMT1*. "SCN" representa la administración de infección por SCN. "Ctr" representa el control no infectado. qRT-PCR se llevó a cabo con cebadores específicos de *GmSAMT1*. Los valores de expresión se normalizaron hasta los niveles de expresión del gen de ubiquitina 3 de soja (*GmUBI3*) en muestras respectivas. El nivel de expresión de *GmSAMT1* en las raíces velludas transgénicas de "VC" sin infección por SCN ("Ctr") se estableció arbitrariamente en 1.0. Cada barra representa el nivel de expresión relativo promedio de tres experimentos independientes con errores estándar, donde cada uno contiene un grupo de cinco plantas. Las barras con letras diferentes son significativamente diferentes (LSD $p < 0.05$).

Las FIG. 13A y 13B son gráficas de barra que muestran abundancia de transcripción relativa de *GmNPR1-1* (**13A**) y *GmNPR1-2* (**13B**) en raíces velludas transgénicas con o sin infección por SCN. Se produjeron y analizaron dos tipos de raíces velludas transgénicas del ámbito Williams 82. "*GmSAMT1*" representa raíces velludas transgénicas que contienen el transgen *GmSAMT1*. "VC" representa raíces velludas transgénicas producidas que utilizan un vector de control que contiene el gen marcador *OFP* pero no el *GmSAMT1*. "SCN" representa la administración de infección por SCN. "Ctr" representa el control no infectado. qRT-PCR se llevó a cabo con cebadores específicos de *GmSAMT1*. Los valores de expresión se normalizaron hasta los niveles de expresión del gen de ubiquitina 3 de soja (*GmUBI3*) en muestras respectivas. El nivel de expresión de *GmSAMT1* en las raíces velludas transgénicas de "VC" sin infección por SCN ("Ctr") se estableció arbitrariamente en 1.0. Cada barra representa el nivel de expresión relativo promedio de tres experimentos independientes con errores estándar, donde cada uno contiene un grupo de cinco plantas. Las barras con letras diferentes son significativamente diferentes (LSD $p < 0.05$).

La FIG. 14 es una gráfica de barras que muestra el análisis de qRT-PCR en hojas de soja infectadas con virus del mosaico para *GmSAMT1* y *GmPR-1*. Las barras blancas y grises representan la expresión relativa de genes de plantas de soja con las respuestas resistentes y susceptibles, respectivamente. Cada barra representa el nivel de expresión relativo promedio de tres experimentos independientes con errores estándar. Las barras con estrella o estrellas significan que la expresión génica son significativamente diferentes a las barras que no tienen una o más estrellas (LSD $p < 0.05$).

La FIG. 15 muestra una alineación de secuencia múltiple de soja Glyma16g26060 (SEQ ID NO: 23), tabaco SABP2 (NtSABP2) (SEQ ID NO: 24) y un ortólogo de Arabidopsis de SABP2 (At4g37150) (SEQ ID NO: 25). Los residuos idénticos se

encuentran coloreados en color negro y los residuos similares en gris. Los residuos de la tríada catalítica se indican con flechas y los residuos que están en contacto con ácido salicílico se indican con triángulos.

La FIG. 16 es una imagen digital de un gel SDS-PAGE de proteína GmSABP2-1 recombinante purificada. El GmSABP2-1 marcado con His expresado en *E. coli* se purificó tal como se describió en el ejemplo 2. La vía M contenía marcadores de peso molecular de proteínas. La vía 1 contenía extracto bruto y la vía 2 contenía aproximadamente 2 µg de proteína GmSABP2-1 purificada. El gel se tiñó con azul Coomassie.

La FIG. 17 es una gráfica que muestra el cálculo del valor de Km de GmSABP2-1 que utiliza MeSA como sustrato en concentraciones de 5 µM a 100 µM. El valor de Km aparente se obtuvo mediante el diagrama de Lineweaver-Burk.

La FIG. 18 es una gráfica que muestra el efecto del pH sobre la actividad de GmSABP2-1. La actividad del GmSABP2-1 purificado se analizó a partir de pH 6.0-9.0. El nivel de actividad de GmSABP2-1 en tampón Tris-HCl de 50 mM, pH 7.0, se estableció arbitrariamente en 1.0.

La FIG. 19 es una gráfica de barras que muestra un ensayo demográfico de nematodos para la resistencia al SCN entre raíces velludas transgénicas con sobreexpresión de *GmSABP2-1* y líneas de soja de control. La relación de la suma de los nematodos J3 y J4 con respecto al total de nematodos infecciosos mostrados en raíces velludas se utiliza como el índice de resistencia. Se describió el índice de resistencia, WS, WR, VS y VR. WS y WR son líneas de raíces velludas de tipo salvaje en el ámbito susceptible y resistente, respectivamente. VS y VR son líneas de raíces velludas que albergan el gen reportero OFP en el ámbito susceptible y resistente, respectivamente, como las líneas de control del vector. SABP2 representa la línea de la raíz velluda con sobreexpresión de *GmSABP2-1* y el gen reportero OFP en el ámbito TN02-275. Las barras con letras diferentes son significativamente diferentes (LSD $p < 0.05$).

La FIG. 20 muestra un mapa de plásmidos del vector pJL-OFP-35S::GmSAMT1.

La FIG. 21 muestra un mapa de plásmidos del vector pJL-OFP-35S::GmSABP2-1.

30 BREVE DESCRIPCIÓN DEL LISTADO DE SECUENCIAS

Las secuencias de aminoácidos y ácidos nucleicos mostradas en la presente se muestran mediante el uso de abreviaturas de letras estándar para bases de nucleótidos y aminoácidos, tal como se define en el capítulo 37 del C.F.R. 1.822. Se muestra únicamente una cadena de cada secuencia de ácidos nucleicos, pero se entiende que la cadena complementaria está incluida mediante cualquier referencia a la cadena mostrada.

El listado de secuencias se presenta como un archivo de texto ASCII en la forma del archivo llamado UTK_0141WP_ST25.txt, que fue creado el 30 de octubre de 2012, tiene 56 kilobytes y se incorpora en la presente a modo de referencia.

SEQ ID NO: 1 es un ejemplo de secuencia de ácidos nucleicos de *GmSAMT1*.

5 SEQ ID NO: 2 es un ejemplo de secuencia de ácidos nucleicos de *GmSABP2-1*.

SEQ ID NO: 3 es un ejemplo de secuencia de ácidos nucleicos de pJL-OFP-35S::GmSAMT1.

SEQ ID NO: 4 es un ejemplo de secuencia de ácidos nucleicos de pJL-OFP-35S::GmSABP2-1.

10 SEQ ID NOS: 5-16 y 26-33 son las secuencias de ácidos nucleicos de cebadores.

SEQ ID NO: 17 es la secuencia de ácidos nucleicos de la sonda de Affymetrix, Gma.12911.1.A1_S_AT.

SEQ ID NOS: 18-22 son las secuencias de ácido nucleico de cinco genes candidatos de soja.

15 SEQ ID NO: 23 es una secuencia de aminoácidos de Glyma16g26060.

SEQ ID NO: 24 es una secuencia de aminoácidos de SABP2 de tabaco.

SEQ ID NO: 25 es una secuencia de aminoácidos del ortólogo de Arabidopsis de SABP2.

DESCRIPCIÓN DETALLADA

20 I. Introducción

El nematodo del quiste de la soja (*Heterodera glycines*, SCN) es la plaga más seria de la soja [*Glycine max* (L.) Merr.]. Tal como se describe en la presente, a través de análisis de microensayo en la expresión génica de dos líneas de soja relacionadas genéticamente (líneas endogámicas recombinantes, RIL) con diferentes respuestas al
25 SCN, se identificó que los genes de soja *Glyma02g06070* y *Glyma16g26060* son inducidos por el SCN en la línea resistente pero no en la línea susceptible.

A partir de los datos de la secuencia, se predijo que *Glyma02g06070* codifica una carboxil metiltransferasa dependiente de S-adenosil-L-metionina (SAM). Con base en la actividad bioquímica de la proteína expresada por *Escherichia coli* codificada mediante
30 este gen, *Glyma02g06070* se identificó como un gen metiltransferasa de ácido salicílico y conocido como *GmSAMT1*. *GmSAMT1* cataliza la conversión de ácido salicílico a salicilato de metilo (MeSA). Para determinar el papel biológico de *GmSAMT1* en la resistencia de la soja contra el SCN, los inventores generaron raíces velludas transgénicas en el ámbito de la soja susceptible con un vector binario que alberga casetes
35 para la sobreexpresión de *GmSAMT1* y del gen de proteína fluorescente naranjada

(OFP). Después de dos semanas de la inoculación de los SCN, se comparó la demografía del SCN en raíces velludas transgénicas y raíces velludas de control con sobreexpresión de *GmSAMT1*. Los resultados mostraron que las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* en la línea de soja susceptible presentaron más resistencia, similar a la de la línea de soja resistente. Para entender mejor su rol de defensa en la soja, también se comparó el nivel de expresión de *GmSAMT1* con las reacciones compatibles e incompatibles entre la soja y los virus del mosaico de la soja. Se descubrió que la expresión de *GmSAMT1* regula positivamente la reacción incompatible y no la reacción compatible. Tomados en conjunto, los resultados descritos en la presente indican que el *GmSAMT1* juega un papel importante en la defensa de la soja contra el SCN y el virus del mosaico de la soja.

A partir de los datos de la secuencia, se predijo que *Glyma16g26060* codifica una hidrolasa de pliegue α/β . Después de clonar este gen y de realizar un ensayo de actividad bioquímica, se identificó que este gen codifica una esterasa de salicilato de metilo. Dado que este gen era el gen ortólogo de esterasa de salicilato de metilo a partir de tabaco, denominado *NtSABP2*, se denominó el gen de soja *GmSABP2-1*. La enzima recombinante codificada por *GmSABP2-1* puede catalizar la conversión de MeSA en ácido salicílico (SA). Para explorar el papel biológico de *GmSABP2-1* en la resistencia de la soja contra el SCN, los inventores generaron raíces velludas transgénicas con sobreexpresión de *GmSABP2-1* y de gen de proteína fluorescente anaranjada (OFP) en el ámbito de la soja susceptible. Después de dos semanas de la inoculación de los huevos de SCN, se comparó la demografía de SCN en las raíces velludas transgénicas y en las raíces velludas de control con sobreexpresión de *GmSABP2-1*. Los resultados mostraron que las raíces velludas transgénicas con sobreexpresión de *GmSABP2-1* en la línea de soja susceptible tuvieron más resistencia, lo que fue similar a la de la línea de soja resistente. Por lo tanto, se determinó que el *GmSABP2-1* juega un papel en la resistencia de la soja contra el SCN mediante la vía de señalización dependiente de SA.

Tomados en conjunto, los dos componentes, *GmSAMT1* y *GmSABP2-1*, trabajan para el transporte de la señal móvil MeSA y juegan un papel importante en la defensa de la soja contra el SCN. Sin restringirse por ninguna consideración teórica, se proponen funciones de *GmSAMT1* y *GmSABP2-1* y un modelo de resistencia de la soja contra el SCN (véase FIG. 1). En la raíz local infectada con SCN, el *GmSAMT1* juega un papel importante para regular los niveles de SA y MeSA basales. En la búsqueda de resistencia al nematodo de nudo de raíces, se descubrió que el nivel residual de SA libre es suficiente para conferir la resistencia basal. Esto indica que el nivel de SA requerido para

la defensa basal contra el SCN puede ser bajo. El SA biosintetizado puede cambiarse a otras formas, tal como MeSA, para inducir SAR en el tejido de raíz distal. Parece que el alto nivel de SA basal inhibe la actividad de GmSABP2-1 y, por lo tanto, inhibe la conversión de MeSA en SA en el sitio local. En el tejido de raíz distal, el alto nivel relativo de MeSA y el bajo nivel de SA favorecen la actividad de GmSABP2-1, lo que conlleva a la conversión de MeSA en SA. De este modo, tal como se describió en la presente, mediante promotores inducibles y específicos adecuados que conducen *GmSAMT1* y *GmSABP2-1* en plantas transgénicas se puede lograr la resistencia de la soja contra el SCN a través de la manipulación de los niveles de SA y MeSA.

II. Resumen de los términos

A menos que se exprese lo contrario, los términos técnicos se utilizan de conformidad con el uso convencional. Las definiciones de términos comunes en la biología molecular pueden encontrarse en Benjamin Lewin, *Genes IX*, publicado por Jones y Bartlet, 2008 (ISBN 0763752223); Kendrew *et al.* (eds.), *The Encyclopedia of Molecular Biology*, publicada por Blackwell Science Ltd., 1994 (ISBN 0632021829); y Robert A. Meyers (ed.), *Molecular Biology and Biotechnology: a Comprehensive Desk Reference*, publicado por VCH Publishers, Inc., 1995 (ISBN 9780471185710).

Los términos en singular "un", "una" y "el/la" incluyen las referencias plurales a menos que el contexto claramente establezca lo contrario. De manera similar, la palabra "o" pretende incluir "y" a menos que el contexto claramente establezca lo contrario. El término "comprende" significa "incluye". En caso de conflicto, la presente memoria descriptiva, incluidas las explicaciones de términos, prevalecerá.

Para facilitar la revisión de las diversas realizaciones de la presente divulgación, se proporcionan las siguientes explicaciones de términos:

5' y/o 3': Se dice que las moléculas de ácido nucleico (tales como, ADN y ARN) tienen "**extremos 5'**" y "**extremos 3'**" ya que se hacen reaccionar mononucleótidos para realizar polinucleótidos de forma tal que el fosfato 5' de un anillo de pentosa mononucleótido esté ligado al oxígeno 3' de su vecino en una dirección a través de una unión de fosfodiéster. Por lo tanto, un extremo de un polinucleótido es conocido como el "**extremo 5'**" cuando su fosfato 5' no está ligado al oxígeno 3' de un anillo de pentosa mononucleótido. El otro extremo de un polinucleótido se conoce como el "**extremo 3'**" cuando su oxígeno 3' no está ligado al fosfato 5' de otro anillo de pentosa mononucleótido. Sin perjuicio de que un fosfato 5' de un anillo de pentosa mononucleótido se encuentre ligado al oxígeno 3' de su vecino, también se puede decir que una secuencia de ácidos nucleicos internos tiene extremos 5' y 3'.

Tanto en una molécula de ácido nucleico circular como lineal, los elementos internos separados son conocidos como que están “**corriente arriba**” o 5' de “**corriente abajo**” o elementos 3'. Con respecto al ADN, esta terminología refleja que la transcripción procede en una dirección 5' a 3' junto con una hebra de ADN. Los elementos potenciadores y promotores, que dirigen la transcripción de un gen ligado, se ubican generalmente en 5' o corriente arriba de la región codificante. Sin embargo, los elementos potenciadores pueden ejercer su efecto aun cuando se encuentren ubicados en 3' del elemento promotor y la región codificante. Las señales de poliadenilación y terminación de transcripción están ubicadas en 3' o corriente abajo de la región codificante.

Rasgo agronómico: La característica de una planta, cuyas características incluyen, a modo no taxativo, morfología, fisiología, crecimiento y desarrollo vegetal, rendimiento, mejora nutritiva, enfermedad o resistencia a las plagas, o tolerancia química o ambiental. Un "rasgo agronómico mejorado" se refiere a una mejora mensurable en un rasgo agronómico que incluye, a modo no taxativo, mejora del rendimiento, incluido el rendimiento mejorado en condiciones sin estrés y rendimiento mejorado en condiciones de estrés ambiental. Las condiciones de estrés pueden incluir, por ejemplo, sequía, sombra, enfermedad fúngica, enfermedad viral, enfermedad bacteriana, infección de insectos, infección de nematodos, exposición a bajas temperaturas, exposición al calor, estrés osmótico, disponibilidad de nutrientes de nitrógeno reducido, disponibilidad de nutrientes de fósforo reducido y alta densidad vegetal. El "rendimiento" puede verse afectado por varias propiedades incluidas, a modo no taxativo, altura de la planta, número de vaina, posición de vaina en la planta, cantidad de internodos, incidencia de destrucción de vaina, tamaño del grano, eficacia de nodulación y fijación del nitrógeno, eficacia de asimilación de nutrientes, resistencia al estrés biótico y abiótico, asimilación del carbono, arquitectura vegetal, resistencia a alojamiento, germinación porcentual de semillas, vigor de semilla y rasgos juveniles. En algunas realizaciones, un rasgo agronómico es la resistencia a una plaga, tal como resistencia de la soja al SCN.

Nivel de alteración de producción o expresión: Cambiar, ya sea mediante aumento o disminución, el nivel de producción o expresión de una molécula de ácido nucleico o una molécula de aminoácido (por ejemplo, un gen, un polipéptido, un péptido), en comparación con un nivel de control de producción o expresión.

Aumento: Cuando se utiliza en referencia a un ácido nucleico, esto se refiere a técnicas que aumentan la cantidad de copias de una molécula de ácido nucleico en una muestra o espécimen. Un ejemplo de aumento es la reacción en cadena de la polimerasa, donde una muestra biológica recogida de un sujeto se pone en contacto con un par de

cebadores oligonucleótidos en condiciones que permiten la hibridación de los cebadores a la plantilla de ácido nucleico en la muestra. Los cebadores se extienden en condiciones adecuadas, separados de la plantilla y luego se vuelven a templar, extender y separar para aumentar la cantidad de copias del ácido nucleico. El producto de aumento *in vitro* se puede caracterizar mediante electroforesis, patrones de escisión de endonucleasa de restricción, ligación o hibridación de oligonucleótidos y/o secuenciamiento de ácido nucleico, mediante el uso de técnicas estándar. Otros ejemplos de técnicas de aumento *in vitro* incluyen aumento de desplazamiento de hebras (véase Patente de Estados Unidos No. 5,744,311); aumento isotérmico libre de transcripción (véase Patente de Estados Unidos No. 6,033,881); aumento de la reacción de cadena de reparación (véase WO 90/01069); aumento de la reacción de la cadena de ligasa (véase EP-A-320 308); aumento de la reacción de ligasa con relleno de brecha (véase Patente de Estados Unidos No. 5,427,930); detección de ligasa acoplada y PCR (véase Patente de Estados Unidos No. 6,027,889); y aumento libre de transcripción de ARN NASBA™ (véase Patente de Estados Unidos No. 6,025,134).

Casete: Un fragmento manipulable de ADN que porta (y es capaz de expresar) uno o más genes de interés entre uno o más conjuntos de sitios de restricción. Un casete puede transferirse desde una secuencia de ADN (normalmente en un vector) a otra mediante "corte" del fragmento con el uso de enzimas de restricción y mediante "pegado" nuevamente en el nuevo contexto. En realizaciones divulgadas, el casete incluye uno o más de los promotores inducibles divulgados.

ADNc (ADN complementario): Una pieza de ADN que carece de segmentos no codificantes internos (intrones) y secuencias regulatorias de transcripción. El ADNc también puede contener regiones no traducidas (UTR) que son responsables del control de la traducción en la molécula de ARN correspondiente. El ADNc normalmente se sintetiza en el laboratorio mediante transcripción inversa del ARN mensajero extraído de células u otras muestras. En algunos ejemplos, el ADNc se utiliza como una fuente de un gen de interés, tal como un gen que codifica GmSAMT1 o GmSABP2-1.

Construcción: Toda molécula de polinucleótido recombinante tal como un plásmido, cósmido, virus, molécula de polinucleótido autónomamente replicadora, fago, o molécula de polinucleótido de ARN o ADN de cadena simple o cadena doble lineal o circular, derivada de cualquier fuente, capaz de la integración genómica o replicación autónoma, que comprende una molécula de polinucleótido donde se han ligado una o más moléculas de polinucleótido que pueden transcribirse.

Planta de control: Una planta que no contiene ADN recombinante que confiere (por ejemplo) un rasgo agronómico alterado o mejorado en una planta transgénica, se utiliza como base de referencia para la comparación, por ejemplo, con el fin de identificar un rasgo agronómico modificado o mejorado en la planta transgénica. Una planta de control adecuada puede ser una planta no transgénica de la línea parental utilizada para generar una planta transgénica, o una planta que al menos no es transgénica para el rasgo particular sujeto a examen (esto es, la planta de control puede haber sido modificada para contener otras secuencias heterólogas o moléculas de ADN recombinante). De este modo, una planta de control puede, en algunos casos, ser una línea de planta transgénica que comprende un vector vacío o gen marcador, pero que no contiene el ADN recombinante, o no contiene todos los ADN recombinantes, en la planta de análisis.

Variante degenerativa y variante conservativa: Un polinucleótido que codifica un polipéptido o un anticuerpo que incluye una secuencia que se degenera como un resultado del código genético. Existen 20 aminoácidos naturales, cuya mayoría se especifica por más de un codón. Por lo tanto, se incluyen todas las secuencias de nucleótidos degenerativas siempre y cuando no cambie la secuencia de aminoácidos del polipéptido codificado por la secuencia de nucleótidos. Debido a la degeneración del código genético, una gran cantidad de ácidos nucleicos funcionalmente idénticos codifican cualquier polipéptido dado. Por ejemplo, los codones CGU, CGC, CGA, CGG, AGA y AGG codifican la arginina del aminoácido. Así, en cada posición donde se especifica una arginina dentro de una secuencia que codifica una proteína, el codón puede alterarse con cualquier codón correspondiente descrito sin alterar la proteína codificada. Dichas variaciones de ácido nucleico son "variaciones silenciosas", que son una especie de variaciones conservadoras. Cada secuencia de ácidos nucleicos en la presente que codifica un polipéptido también describe cada variación silenciosa posible. Un experto reconocerá que puede modificarse cada codón en un ácido nucleico (excepto AUG, que normalmente es el único codón para metionina) para proporcionar una molécula funcionalmente idéntica mediante técnicas estándar. Por consiguiente, cada "variación silenciosa" de un ácido nucleico que codifica un polipéptido está implícita en cada secuencia descrita.

Asimismo, el entendido reconocerá que las adiciones, eliminaciones y sustituciones individuales que alteran, agregan o eliminan un único aminoácido o un porcentaje pequeño de aminoácidos (por ejemplo menos de un 5%, tal como menos de un 4%, menos de un 3%, menos de un 2% o aún menos de un 1%) en una secuencia

codificada son variaciones conservadoras donde las alteraciones ocasionan la sustitución de un aminoácido con un aminoácido químicamente similar.

Las sustituciones conservadoras de aminoácidos que proporcionan aminoácidos funcionalmente similares se conocen bien en la técnica. Los siguientes seis grupos

5 contienen aminoácidos que son sustituciones conservativas entre sí:

- 1) Alanina (A), Serina (S), Treonina (T);
- 2) Ácido aspártico (D), Ácido glutámico (E);
- 3) Asparagina (N), Glutamina (Q);
- 4) Arginina (R), Lisina (K);
- 10 5) Isoleucina (I), Leucina (L), Metionina (M), Valina (V) y
- 6) Fenilalanina (F), Tirosina (Y), Triptófano (W).

No todas las posiciones de residuos dentro de una proteína tolerarán otra sustitución "conservadora". Por ejemplo, si un residuo de aminoácido es esencial para una función de la proteína, incluso otra sustitución conservadora puede interrumpir esa

15 actividad.

Resistencia a enfermedades o resistencia a plagas: La evasión de los síntomas nocivos que son el resultado de las interacciones planta-patógeno. Los genes resistentes a enfermedades y resistentes a plagas tales como lisozimas o cecropinas para la protección antibacteriana, o proteínas tales como defensinas, glucanasas o quitinasas

20 para la protección antifúngica, o endotoxinas *Bacillus thuringiensis*, inhibidores de proteasa, colagenasas, lectinas o glucosidasas para controlar nematodos o insectos son todos ejemplos de productos génicos útiles.

Tal como se utiliza en la presente, el término "plaga" incluye, a modo no taxativo, insectos, hongos, bacterias, virus, nematodos, ácaros, garrapatas y similares. Las plagas

25 de insectos incluyen insectos que se seleccionan de *Coleoptera*, *Diptera*, *Hymenoptera*, *Lepidoptera*, *Mallophaga*, *Homoptera*, *Hemiptera*, *Orthoptera*, *Thysanoptera*, *Dermaptera*, *Isoptera*, *Anoplura*, *Siphonaptera*, *Trichoptera*, etc., particularmente *Coleoptera*, *Lepidoptera*, y *Diptera*. Los virus incluyen, a modo no taxativo, virus del mosaico del tabaco o pepino, virus de las manchas anulares, virus de la necrosis, virus del mosaico enanizante del maíz, etc. Los nematodos incluyen, a modo no taxativo, nematodos

30 parásitos tales como nematodos del nudo de la raíz, quiste y lesión, incluidos *Heterodera spp.*, *Meloidogyne spp.* y *Globodera spp.*; particularmente miembros de los nematodos del quiste, incluidos, a modo no taxativo, *Heterodera glycines* (nematodo del quiste de la soja); *Heterodera schachtii* (nematodo del quiste de la remolacha); *Heterodera avenae*

35 (nematodo del quiste del cereal); y *Globodera rostochiensis* y *Globodera pallida*

(nematodo del quiste de la papa). Los nematodos de la lesión incluyen, a modo no taxativo, *Pratylenchus spp.* Las plagas fúngicas incluyen aquellas que ocasionan royas amarillas, de hojas y tallo. **ADN (ácido desoxirribonucleico):** El ADN es un polímero de cadena larga que comprende el material genético de la mayoría de los organismos (algunos virus tienen genes que comprenden ácido ribonucleico (ARN)). Las unidades repetitivas en polímeros de ADN son cuatro nucleótidos diferentes, cada uno comprende una de las cuatro bases, adenina, guanina, citosina y timina ligadas a un azúcar desoxirribosa a la cual se une un grupo de fosfato. Los tripletes de nucleótidos (conocidos como codones) codifican cada aminoácido en un polipéptido, o una señal de detención. El término codón también se utiliza para las secuencias correspondientes (y complementarias) de tres nucleótidos en el ARNm al cual se transcribe la secuencia de ADN.

A menos que se especifique lo contrario, toda referencia a una molécula de ADN incluye el complemento inverso de esa molécula de ADN. Excepto cuando el presente texto requiera que las moléculas de ADN sean monocatenarias, aunque se escriban para describir únicamente una cadena simple, estas abarcan ambas cadenas de una molécula de ADN bicatenario.

Codificación: Se dice que un polinucleótido codifica un polipéptido si, en su estado natural o cuando se manipula mediante métodos conocidos por los entendidos en la técnica, la molécula de polinucleótido puede transcribirse y/o traducirse para producir un ARNm y/o el polipéptido o un fragmento de este. La cadena antisentido es el complemento de dicho ácido nucleico y la secuencia de codificación puede deducirse a partir de ella.

Dominio de mejora: Un elemento regulador de la transcripción con actuación cis (conocido como elemento cis) que confiere un aspecto del control general de la expresión génica. Un dominio de mejora puede funcionar para unir factores de transcripción, que son factores de proteínas con actuación trans que regulan la transcripción. Algunos dominios de mejora unen más de un factor de transcripción y los factores de transcripción pueden interactuar con diferentes afinidades con más de un dominio de mejora. Los dominios de mejora pueden identificarse mediante una cantidad de técnicas, incluidos los análisis de eliminación (eliminar uno o más nucleótidos del extremo 5' o internos a un promotor); análisis de proteína de unión a ADN mediante el uso de plantillas de impresión DNasa I, interferencia de metilación, ensayos de cambio de movilidad de electroforesis, impresión de plantillas genómicas *in vivo* mediante PCR mediada por ligación y otros ensayos convencionales; o mediante comparación de la secuencia de ADN con motivos

de elementos cis conocidos mediante el uso de métodos de comparación convencionales de la secuencia de ADN. La estructura fina de un dominio de mejora puede estudiarse además mediante mutagénesis (o sustitución) de uno o más nucleótidos o mediante otros métodos convencionales. Los dominios de mejora pueden obtenerse mediante síntesis química o mediante aislamiento de promotores que incluyen dichos elementos y pueden sintetizarse con nucleótidos de flanqueo adicionales que contienen sitios de enzima de restricción útiles para facilitar la manipulación de subsecuencia.

Expresión (génica): La transcripción de una molécula de ADN en una molécula de ARN transcrita. Más generalmente, la expresión génica abarca los procesos mediante los cuales la información codificada de un gen se convierte en las estructuras presentes y que funcionan en la célula. Los genes expresados incluyen aquellos que se transcriben en ARNm y luego se traducen en proteínas y aquellos que se transcriben en ARN pero no se traducen en proteínas (por ejemplo, ARNip, ARN de transferencia y ARN ribosomal). De este modo, la expresión de una secuencia diana, tal como un gen o una región promotora de un gen, puede ocasionar la expresión de un ARNm, una proteína o ambos. La expresión de la secuencia diana puede inhibirse o potenciarse (disminuirse o aumentarse). La expresión génica puede describirse como que está relacionada a calidades temporales, espaciales, de desarrollo o morfológicas así como a indicaciones cuantitativas o cualitativas.

Actividad reguladora de genes: La capacidad de un polinucleótido de afectar la transcripción o traducción de una molécula de polinucleótido transcribible operativamente ligada, tal como un promotor inducible. Una molécula de polinucleótido aislada que tiene actividad reguladora de genes puede proporcionar una expresión espacial o temporal o modular niveles y tasas de expresión de la molécula de polinucleótido transcribible operativamente ligada. Una molécula de polinucleótido aislada que tiene actividad reguladora de genes puede incluir un promotor, intrón, líder o región de terminación de transcripción 3'.

Material genético: Una frase que pretende incluir todos los genes, ácido nucleico, ADN y ARN.

Secuencia heteróloga de nucleótidos: Una secuencia que no ocurre naturalmente con una secuencia de promotor. Mientras esta secuencia de nucleótidos es heteróloga a la secuencia de promotores, puede ser homóloga o natural o heteróloga o extraña al hospedador vegetal. La invención abarca además la expresión de las secuencias homólogas de codificación de los promotores, particularmente las secuencias de codificación relacionadas con el fenotipo de resistencia. La expresión de las

secuencias homólogas de codificación modificará el fenotipo de la célula vegetal o planta transformada.

Resistencia superior a plagas o resistencia a plagas mejorada: Una resistencia elevada o mejorada para exceder una planta normal o de control o parte de esta (por ejemplo, una planta que no se ha transformado con un ácido nucleico aislado que codifica esterasa de salicilato de metilo o carboxilo que depende de S-adenosil-L-metionina, tal como se describe en la presente). En algunos ejemplos, un aumento o mejora es una elevación de al menos aproximadamente un 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400%, 500% o más.

En cis: Indica que dos secuencias están ubicadas en la misma parte del ARN o ADN.

En trans: Indica que dos secuencias están ubicadas en partes diferentes del ARN o ADN.

ADN de inserción: El ADN heterólogo dentro de un casete de expresión, tal como el casete de expresión divulgado, utilizado para transformar el material vegetal mientras el "ADN flanqueante" puede comprender ADN genómico naturalmente presente en un organismo tal como una planta, o ADN extraño (heterólogo) introducido a través del proceso de transformación que es extraño a la molécula de ADN de inserción original, por ejemplo, fragmentos asociados con el evento de transformación. Una "región flanqueante" o "secuencia flanqueante", tal como se utiliza en la presente, se refiere a una secuencia de al menos 20, 50, 100, 200, 300, 400, 1000, 1500, 2000, 2500 o 5000 pares de bases o más que está ubicada inmediatamente corriente arriba y contigua o inmediatamente corriente abajo y contigua a la molécula de ADN de inserción extraña original.

Aislado: Un componente biológico "aislado" (tal como un ácido nucleico, péptido o proteína) ha sido sustancialmente separado, producido separado de o purificado fuera de otros componentes biológicos en la célula del organismo en el cual el componente sucede naturalmente, por ejemplo, otro ADN y ARN cromosomal y extracromosomal y proteínas. Los ácidos nucleicos, péptidos y proteínas que han sido "aislados" incluyen entonces ácidos nucleicos y proteínas purificados mediante métodos de purificación estándar. El término también abarca ácidos nucleicos, péptidos y proteínas preparados mediante la expresión recombinante en una célula hospedadora así como ácidos nucleicos sintetizados químicamente.

Ligado de forma operativa: Este término se refiere a la yuxtaposición de componentes, particularmente secuencias de nucleótidos, de modo que pueda llevarse a

cabo la función normal de los componentes. Así, una primera secuencia de ácidos nucleicos se encuentra ligada de forma operativa a una segunda secuencia de ácidos nucleicos cuando la primera secuencia de ácidos nucleicos está ubicada en una relación funcional con la segunda secuencia de ácidos nucleicos. Por ejemplo, un promotor se encuentra ligado de forma operativa a una secuencia de codificación si el promotor afecta la transcripción o expresión de la secuencia de codificación. En general, las secuencias de ADN ligadas de forma operativa son contiguas y, en caso que fuera necesario unir dos regiones de codificación de proteínas, se encuentran en el mismo marco de lectura. Una secuencia de codificación que se encuentra "ligada de forma operativa" a una o más secuencias reguladoras se refiere a una configuración de secuencias de nucleótidos donde la secuencia de codificación puede expresarse según el control de regulación (por ejemplo, control de traducción y/o transcripción) de las secuencias reguladoras.

Planta: Toda planta y progenie de esta. El término también incluye partes de plantas, incluidos semillas, esquejes, tubérculos, frutas, flores, etc. Tal como se utiliza en la presente, el término planta incluye células vegetales, órganos vegetales, protoplastos vegetales, cultivos de tejido de células vegetales a partir de los cuales se pueden regenerar plantas, callos vegetales, matas de plantas y células vegetales que están intactas en plantas o partes de plantas tales como embriones, polen, óvulos, semillas, hojas, flores, ramas, frutas, tallos, raíces, puntas de raíces, anteras y similares. Las progenies, variantes y mutaciones de las plantas regeneradas también se encuentran incluidas en el alcance de la invención. El término **célula vegetal**, tal como se utiliza en la presente, se refiere a la unidad estructural y fisiológica de plantas, que consiste en un protoplasto y la pared celular circundante, incluidos aquellos que presentan modificación genética, tal como transformación, que ha sido modificado como un gen de interés, o es una planta o célula vegetal que descende de una planta o célula así modificada y que comprende la modificación. Un "control" o "planta de control" o "célula vegetal de control" proporciona un punto de referencia para medir los cambios en el fenotipo de la célula vegetal o planta de referencia. Una célula vegetal o planta de control puede comprender, por ejemplo: (a) una célula o planta de tipo salvaje, es decir, del mismo genotipo que el material de partida para la modificación genética que dio como resultado la célula o planta de referencia; (b) una célula vegetal o planta del mismo genotipo que el material de partida pero que ha sido modificada con una construcción nula (es decir, con una construcción que no tiene efecto conocido sobre el rasgo de interés, tal como una construcción que comprende un gen marcador); (c) una célula vegetal o planta que es un segregante no transformado entre la progenie de una célula vegetal o planta de

referencia; (d) una célula vegetal o planta genéticamente idéntica a la célula vegetal o planta de referencia pero que no está expuesta a condiciones o estímulos que inducirían la expresión del gen de interés o (e) la célula vegetal o planta de referencia en sí misma, en condiciones en las que el gen de interés no se expresa. El término **órgano vegetal**, tal como se utiliza en la presente, se refiere a una parte de planta diferente y visiblemente diferenciada, tal como una raíz, tallo, hoja o embrión. Más generalmente, el término **tejido vegetal** se refiere a cualquier tejido de una planta en planta o en cultivo. Este término incluye la totalidad de una planta, célula vegetal, órgano vegetal, protoplasto, cultivo celular o cualquier grupo de células vegetales organizados en una unidad estructural y funcional.

Molécula de polinucleótido: ADN o ARN de cadena simple o doble de origen genómico o sintético; esto es, un polímero de bases de desoxirribonucleótido o ribonucleótido, respectivamente, leído desde el extremo 5' (corriente arriba) al extremo 3' (corriente abajo).

Promotor: Un arreglo de secuencias de control de ácido nucleico que dirige la transcripción de un ácido nucleico, mediante reconocimiento y aglutinación de, por ejemplo, polimerasa de ARN II y otras proteínas (factores de transcripción con actuación trans) para iniciar la transcripción. Un promotor incluye secuencias de ácido nucleico necesarias próximas al sitio de inicio de la transcripción, tal como, en el caso de un promotor de tipo II de polimerasa, un elemento TATA. Como mínimo, un promotor incluye normalmente al menos un sitio de unión a polimerasa de ARN junto con uno o más sitios de unión al factor de transcripción, lo que modula la transcripción en respuesta a la ocupación mediante factores de transcripción. Los ejemplos representativos de promotores (y elementos que pueden ensamblarse para producir un promotor) se describen en la presente. Los promotores pueden definirse mediante su patrón de expresión temporal, espacial o de desarrollo.

Un promotor vegetal es un promotor natural o no natural que es funcional en células vegetales. En un ejemplo, el promotor es un promotor constitutivo de alto nivel, tal como un promotor específico de tejido.

Proteína: Una molécula biológica, por ejemplo un polipéptido, expresada por un gen y que comprende aminoácidos.

Protoplasto: Una célula vegetal aislada sin paredes celulares, que tiene potencial de regeneración en un cultivo celular o en una planta completa.

Purificado: El término purificado no requiere una pureza absoluta; en cambio, pretende ser un término relativo. De este modo, por ejemplo, una preparación de

proteínas purificada es una donde la proteína está más enriquecida que la proteína en su ambiente de generación, por ejemplo, dentro de una célula o en una cámara de reacción bioquímica. De preferencia, la preparación de proteínas se purifica de modo que la proteína represente al menos un 50% del contenido de proteínas total de la preparación.

5 **Recombinante:** El ácido nucleico recombinante es uno que tiene una secuencia que no ocurre naturalmente o que tiene una secuencia realizada mediante combinación artificial de dos segmentos de secuencias que de otro modo estarían separados. Esta combinación artificial se logra a menudo mediante síntesis química o, más comúnmente, mediante la manipulación artificial de segmentos aislados de ácidos nucleicos, por
10 ejemplo, mediante técnicas de modificación genética. De manera similar, una proteína recombinante es una que se codifica por una molécula de ácido nucleico recombinante.

Promotor regulable o promotor inducible: La actividad que está regulada (directa o indirectamente) por un agente, tal como un factor de transcripción, un compuesto químico, una condición ambiental, o una molécula de ácido nucleico.

15 **Regulación de la expresión génica:** Los procesos para controlar la expresión de un gen mediante el aumento o la disminución de la expresión, producción o actividad de un agente que afecta la expresión génica. El agente puede ser una proteína, tal como un factor de transcripción, o una molécula de ácido nucleico, tal como una molécula de ARNm o ARNip que, cuando entra en contacto con el gen o sus secuencias de regulación
20 corriente arriba, o un ARNm codificado por el gen, aumenta o disminuye la expresión génica.

Secuencias o elementos reguladores: Estos términos se refieren normalmente a una clase de moléculas de polinucleótido (tal como moléculas de ADN, que tienen secuencias de ADN) que ejerce influencia o controla la transcripción o traducción de una
25 molécula de polinucleótido transcribible ligado de forma operativa y, de ese modo, la expresión de genes. En los términos se incluyen promotores, potenciadores, líderes, intrones, regiones de control de locus, aislantes/elementos límite, silenciadores, regiones de unión a la matriz (también conocidas como regiones de unión al andamio), represores, terminadores de transcripción (conocidos como regiones de terminación de transcripción),
30 orígenes de replicación, centrómeros y puntos de acceso de recombinación meiótica. Los promotores son secuencias de ADN próximas al extremo 5' de un gen que actúa como un sitio de unión para la polimerasa de ARN y a partir de las cuales se inicia la transcripción. Los potenciadores son elementos de control que elevan el nivel de transcripción de un promotor, normalmente independientemente de la orientación del potenciador o de la
35 distancia del promotor. Las regiones de control del locus (LCR) confieren expresión

temporalmente regulada y específica del tejido a genes a los que están unidas. Las LCR funcionan independientemente de su posición con relación al gen, pero dependen de la cantidad de copias. Se cree que funcionan para abrir la estructura de nucleosomas, de modo que otros factores se unan al ADN. Las LCR también pueden afectar el tiempo de replicación y el uso del origen. Los aislantes (también conocidos como elementos límite) son secuencias de ADN que previenen la activación (o inactivación) de la transcripción de un gen, mediante efectos de bloqueo de cromatina circundante. Los silenciadores y represores son elementos de control que suprimen la expresión génica; actúan en un gen independientemente de su orientación o distancia del gen. Las regiones de unión a la matriz (MAR), también conocidas como regiones de unión al andamiaje, son secuencias dentro del ADN que se unen al andamiaje nuclear. Pueden afectar la transcripción, posiblemente mediante la separación de los cromosomas en los dominios reguladores. Se cree que las MAR median estructuras de alto orden con bucles dentro de los cromosomas. Los terminadores de transcripción son regiones dentro de la proximidad del gen donde la polimerasa de ARN se libera de la plantilla. Los orígenes de replicación son regiones del genoma que, durante la síntesis del ADN o fases de replicación de la división celular, comienzan con el proceso de replicación de ADN. Los puntos de acceso de recombinación meiótica son regiones del genoma que se recombinan con más frecuencia que el promedio durante la meiosis. Los nucleótidos específicos dentro de una región reguladora pueden cumplir varias funciones. Por ejemplo, un nucleótido específico puede ser parte de un promotor y participar en la unión de una proteína activadora de transcripción. Los elementos reguladores aislados que funcionan en células (por ejemplo, en plantas o células vegetales) son útiles para modificar fenotipos vegetales, por ejemplo, mediante modificación genética.

ARN: Un polímero típicamente lineal de monómeros de ácido ribonucleico, unido mediante enlaces de fosfodiéster. Las moléculas de ARN de origen natural son de tres clases, mensajeras (ARNm, que codifica proteínas), ribosomales (ARNr, componentes de ribosomas) y de transferencia (ARNt, moléculas responsables de transferir monómeros de aminoácido al ribosoma durante la síntesis de proteínas). El ARN mensajero incluye ARN polisomal asociado con la membrana y heteronuclear (ARNhn) (unido al retículo endoplásmico rugoso). El ARN total se refiere a una mezcla heterogénea de todos los tipos de moléculas de ARN.

Marcador analizable: Un marcador que confiere un rasgo identificado mediante observación o análisis.

Marcador seleccionable: Un marcador que confiere un rasgo que se puede

seleccionar a través de medios químicos, por ejemplo, mediante el uso de un agente selectivo (por ejemplo, un herbicida, un antibiótico o similar). Los marcadores seleccionables incluyen, a modo no taxativo, genes resistentes a antibióticos, tales como, kanamicina (nptII), G418, bleomicina, higromicina, cloranfenicol, ampicilina, tetraciclina o similar. Los marcadores seleccionables adicionales incluyen un gen de barra que codifica resistencia a bialafós; un gen de sintasa EPSP mutante que codifica resistencia a glifosato; un gen de nitrilasa que confiere resistencia a bromoxinilo; un gen de sintasa de acetolactato mutante (ALS) que confiere resistencia a imidazolinona o sulfonilurea; o un gen de DHFR resistente a metotrexato. En un ejemplo, el marcador seleccionable es AAD1.

Identidad de secuencia: La similitud entre dos secuencias de ácido nucleico, o dos secuencias de aminoácido, se expresa en términos de la similitud entre las secuencias, de otro modo conocida como identidad de secuencia. La identidad de secuencia se mide con frecuencia en términos de identidad porcentual (o similitud u homología); cuanto más alto sea el porcentaje, más similares serán las dos secuencias. El porcentaje de identidad de secuencia se representa como la fracción de identidad multiplicada por 100. La comparación de una o más secuencias de polinucleótidos o polipéptidos puede realizarse con respecto a una secuencia de polinucleótidos o polipéptidos de longitud completa o una porción de esta, o con respecto a una secuencia de polinucleótidos de mayor longitud.

Los métodos para alinear secuencias para su comparación son bien conocidos en la técnica. Varios programas y alineación de algoritmos se describen en: Smith y Waterman (Adv. Appl. Math. 2: 482, 1981); Needleman y Wunsch (J. Mol. Biol. 48: 443, 1970); Pearson y Lipman (PNAS. USA 85: 2444, 1988); Higgins y Sharp (Gene, 73: 237-244, 1988); Higgins y Sharp (CABIOS 5: 151-153, 1989); Corpet *et al.* (Nuc. Acids Res. 16: 10881-90, 1988); Huang *et al.* (Comp. Appls Biosci. 8: 155-65, 1992); y Pearson *et al.* (*Methods in Molecular Biology* 24: 307-31, 1994). Altschul *et al.* (*Nature Genet.*, 6: 119-29, 1994) presenta una consideración detallada de métodos de alineación de secuencia y cálculos de homología.

Las herramientas de alineación ALIGN (Myers y Miller, CABIOS 4:11-17, 1989) o LFASTA (Pearson y Lipman, 1988) se pueden utilizar para realizar comparaciones de secuencias (Internet Program © 1996, W. R. Pearson y la Universidad de Virginia, "fasta20u63" versión 2.0u63, fecha de publicación Diciembre de 1996). ALIGN compara secuencias completas entre sí, mientras que LFASTA compara regiones de similitud local. Estas herramientas de alineación y sus tutoriales respectivos se encuentran disponibles

en internet con la dirección web biology.ncsa.uiuc.edu.

Los ortólogos o parálogos (más generalmente, homólogos) de una secuencia específica se caracterizan normalmente por poseer más de 75% de identidad de secuencia contada sobre la alineación de longitud completa con respecto a la secuencia de aminoácidos de una proteína específica (o la secuencia de ácidos nucleicos de una molécula de ácido nucleico específica) mediante el uso de ALIGN para establecer los parámetros estándar. Las secuencias con similitud aún mayor a las secuencias de referencia mostrarán identidades porcentuales en aumento cuando se evalúan con este método, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 92%, al menos un 95% o al menos un 98% de identidad de secuencia. En dicho caso, las identidades porcentuales serán esencialmente similares a aquellas analizadas para determinar la identidad de secuencia de longitud completa.

Una indicación alternativa de que dos moléculas de ácido nucleico se encuentran íntimamente relacionadas es que las dos moléculas se hibriden mutuamente en condiciones rigurosas. Las condiciones rigurosas dependen de la secuencia y son diferentes bajo parámetros ambientales diferentes. En general, se seleccionan condiciones rigurosas para ser de aproximadamente 5°C a 20°C más bajas que el punto de fusión térmico (T_m) para la secuencia específica a un pH y fuerza iónica definidos. La T_m es la temperatura (a un pH y una fuerza iónica definidos) en la que el 50% de la secuencia diana se hibrida para obtener una sonda perfectamente coincidente. Las condiciones de hibridación de ácido nucleico y el cálculo de la rigurosidad pueden encontrarse en Sambrook *et al.* (En *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, Nueva York, 1989) y Tijssen (*Laboratory Techniques in Biochemistry and Molecular Biology*, Parte I, Cap. 2, Elsevier, Nueva York, 1993). Las moléculas de ácido nucleico que se hibridan en condiciones rigurosas en una secuencia de proteína específica se hibridarán normalmente en una sonda con base en la secuencia que codifica la proteína, un dominio completo, u otras partes seleccionadas de la secuencia de codificación en condiciones de lavado de 0.2 x SSC, SDS al 0.1% a 65°C.

Las secuencias de ácido nucleico que no muestran un alto grado de identidad pueden, no obstante, codificar secuencias de aminoácido similares, debido a la degeneración del código genético. Se entiende que los cambios en la secuencia de ácidos nucleicos pueden llevarse a cabo mediante el uso de esta degeneración para producir varias secuencias de ácido nucleico que codifiquen cada una la misma proteína. **La identidad de secuencia porcentual sustancial** tiene al menos aproximadamente un 80% de identidad de secuencia, tal como al menos aproximadamente un 80%, al menos

aproximadamente un 85%, al menos aproximadamente un 90%, al menos aproximadamente un 95%, o aún más identidad de secuencia, tal como aproximadamente un 98% o aproximadamente un 99% de identidad de secuencia.

5 **Soja:** Glycine max e incluye todas las variedades vegetales que pueden cultivarse con la soja.

Gen resistente al nematodo del quiste de la soja: Un gen que, cuando se expresa en una planta, contribuye a la resistencia al nematodo del quiste de la soja. De manera alternativa, un gen del nematodo del quiste de la soja puede ser una variante alélica del gen resistente, particularmente variantes, que ocasionan un fenotipo susceptible.

10 El **evento transgénico** se produce mediante la transformación de células vegetales con una o más construcciones de ADN heterólogo, incluido un casete de expresión de ácido nucleico que incluye un transgen de interés, la regeneración de una población de plantas que son el resultado de la inserción del transgen en el genoma de la planta, y la selección de una planta particular caracterizada por la inserción en una ubicación de genoma particular. En algunas realizaciones de la presente divulgación, el transgen de interés se encuentra ligado de forma operativa a un promotor inducible divulgado, tal como SEQ ID NO: 1 o 2. El evento está caracterizado fenotípicamente por la expresión del o de los transgenes. A nivel genético, un evento es parte de la realización genética de una planta. El término "evento" se refiere también a la progenie producida mediante cruzamiento sexual entre el transformante y otra variedad que incluye el ADN heterólogo. Aun después del retrocruzamiento repetido a una base recurrente, el ADN insertado y el ADN flanqueante de la base transformada se encuentran presentes en la progenie del cruzamiento en la misma ubicación cromosómica. El término "evento" también se refiere al ADN del transformante original que comprende el ADN insertado y la secuencia flanqueante inmediatamente adyacente al ADN insertado que se esperaría se transfiera a una progenie que reciba el ADN insertado incluido el transgen de interés como resultado del cruzamiento sexual de una línea parental que incluya el ADN insertado (por ejemplo, la progenie y el transformante originales que resultan de la autofecundación) y una línea parental que no contenga el ADN insertado.

25 **Planta transgénica:** Una planta que contiene una secuencia de nucleótidos (heteróloga) extraña insertada en su genoma nuclear o en su genoma organular.

Transgen: Una secuencia de ácidos nucleicos que está insertada en una célula hospedadora o células hospedadoras mediante la técnica de transformación.

35 **Transgénico:** Este término se refiere a una planta/hongo/célula/otra entidad u

organismo que contiene material genético recombinante que no se encuentra normalmente en entidades de este tipo/especie (esto es, material genético heterólogo) y que ha sido introducido en la entidad en cuestión (o en progenitores de la entidad) mediante manipulación humana. De este modo, una planta que se cultiva a partir de una
 5 célula vegetal en la cual se introduce ADN recombinante mediante transformación (una célula vegetal transformada) es una planta transgénica, ya que todas descienden de esa planta que contiene el transgen introducido (ya sea producido sexual o asexualmente).

Transformación: Proceso mediante el cual el ADN exógeno ingresa y cambia una célula receptora. Puede ocurrir en condiciones naturales o en condiciones artificiales
 10 mediante el uso de varios métodos bien conocidos en la técnica. La transformación puede realizarse mediante cualquier método conocido para la inserción de secuencias de ácido nucleico extrañas en una célula hospedadora procariota o eucariota. La selección del método se ve influenciada por la célula hospedadora que se transforma y puede incluir, a modo no taxativo, infección viral, electroporación, lipofección y bombardeo de partículas.

Vector: Una molécula de ácido nucleico tal como se introduce en una célula hospedadora, que de ese modo produce una célula hospedadora transformada. Un vector puede incluir secuencias de ácido nucleico que le permiten replicarse en la célula hospedadora, tal como un origen de replicación. Un vector puede incluir también uno o
 15 más genes terapéuticos y/o genes marcadores seleccionables y otros elementos genéticos conocidos en la técnica. Un vector puede transducir, transformar o infectar una célula, haciendo así que la célula exprese ácidos nucleicos y/o proteínas en lugar de aquellas que son naturales a la célula. Un vector incluye opcionalmente materiales para ayudar a lograr el ingreso del ácido nucleico en la célula, tal como una partícula viral, liposoma, recubrimiento de proteína o similar.

A continuación se describen los métodos y materiales adecuados para la práctica o el análisis de la presente divulgación. Dichos métodos y materiales son únicamente ilustrativos y no pretenden ser limitativos. Se pueden utilizar otros métodos y materiales similares o equivalentes a aquellos descritos en la presente. Por ejemplo, los métodos convencionales bien conocidos en la técnica a los que pertenece la invención divulgada
 25 se describen en varias referencias generales y más específicas, incluidas, por ejemplo, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2da ed., Cold Spring Harbor Laboratory Press, 1989; Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 3ra ed., Cold Spring Harbor Press, 2001; Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Associates, 1992 (y Complementos a 2000); y Ausubel *et al.*,
 30 *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in*
 35

Molecular Biology, 4ta ed., Wiley & Sons, 1999.

III. Descripción de varias realizaciones

La presente divulgación describe nuevos genes resistentes a plagas, tal como una secuencia de ácidos nucleicos que codifica GmSAMT1 (tal como se estableció en la SEQ ID NO: 1) y/o GmSABP2-1 tal como se estableció en la SEQ ID NO: 2) o un fragmento funcional de este. También se proporcionan construcciones de ADN que comprenden un gen resistente a plagas descrito. En una realización, el gen resistente a plagas confiere un rasgo agronómico a una planta en la cual se expresa, por ejemplo resistencia a plagas, tal como resistencia al SCN.

También se proporcionan plantas transgénicas. En una realización, una planta transgénica se transforma de forma estable con una construcción de ADN divulgada. En algunas realizaciones, la planta transgénica es dicotiledónea. En otras realizaciones, la planta transgénica es una monocotiledónea. En una realización particular, la planta transgénica es una planta de soja. Se proporciona además una semilla de una planta transgénica divulgada. En una realización, la semilla comprende la construcción de ADN divulgada. Además se proporciona un tejido o célula vegetal transgénico. En una realización, un tejido o célula vegetal transgénica comprende un gen resistente a plagas descrito. En algunas realizaciones, el tejido o célula vegetal deriva de una dicotiledónea. En algunas realizaciones, el tejido o célula vegetal deriva de una monocotiledónea. En una realización particular, el tejido o célula vegetal deriva de una planta de soja.

También se proporcionan métodos para producir un tejido, semilla, célula vegetal o planta transgénica divulgada. En algunas realizaciones, el método comprende transformar un tejido o célula vegetal con una construcción de ADN divulgada. En algunas realizaciones, el método es un método para potenciar la resistencia a la enfermedad en una planta.

Además se proporciona un esqueje, semilla, flor, brote, raíz, hoja, fruta, célula vegetal u otro material reproductor útil en la propagación sexual o asexual, plantas de progenie incluidos híbridos de F1, plantas estériles macho y otras plantas y productos de planta que derivan de las plantas transgénicas divulgadas.

A. Genes resistentes a los patógenos de plantas

La presente divulgación proporciona genes resistentes a plagas no reconocidos anteriormente aislados de la soja. En una realización, el gen resistente a las plagas codifica una carboxil metiltransferasa dependiente de S-adenosil-L-metionina (SAM) y se conoce como *GmSAMT1*. En otra realización, el gen resistente a las plagas codifica una esterasa de salicilato de metilo y se conoce como *GmSABP2-1*.

En una realización, la secuencia de ácidos nucleicos *GmSAMT1* se proporciona como la SEQ ID NO: 1 a continuación:

ATGGAAGTAGCACAGGTACTCCACATGAACGGTGGCGTTGGAGACGCAAGCTATGCA
AACAACTCCCTTGTTTCAGCAAAAGGTGATTTGTTTGACAAAGCCCATAAGAGAGGAAG
5 CCATAAGAAGCCTCTATTGCAGCACACACCCCAAGCTTGGCAATTGCAGATTTGG
GTTGCTCTTCTGGACCAAACACTTTATTTGTTGTGTCTGAATTCATAAAAATTGTGGAG
AAGCTTTGCCGAGAGCTGAACCACAAATCTCCAGAATACAAAGTCTTTCTGAATGATC
TCCCTGGGAATGACTTCAACAACATCTTCAAGTCTCTTGACAGCGTCAAAGAGAAATT
GTGTGATGAAATGGAAGTGGGATCGGTCCATGCTACTTCTCGGGTGTTCCCGGTTT
10 TTTCTATGGAAGGGTTTTCCCATATCAAAGTCTTCATTTTGTCCATTCCTCATACAGCC
TTCAATGGCTATCTAAGGTTCTGAGGGTGTAGACAACAACAAGGGCAATGTTTACAT
AGGCAGTACGAGCCCCAAAAATGTTGTGAGAGCTTACTATGAGCAATTTTCAGAGAGAT
TTCTCTCTTTTTCTCAAGTGTCGTGCAGAGGAATTGGTTGAAGGAGGACGCATGGTTC
TCACATTTTTGGGAAGAAGAAGCGATGATCCATCTAGCAAGGATGGTTGCTACATTTG
15 GGAGCTTTTGGCTACTGCTCTTAGTGATATGGTCTTGCAGGGAATCATAAGAGAAGAG
CAATTAGATACTTTTAACATCCCTCAATACACTCCATCCCCATCTGAAGTGAAATTGGA
AGTTCTTAAAGAAGGATCATTGCGCATCAATCGTCTAGAGGTTTCTGAGGTGAATTGG
AATGCTCTCGATGAGTGGAATGCTCTAGACTTTGAATCTGAAAGGTCTGAATCACTTA
GTGATGGTGAATACAATGTGGCACAGTGCATGAGGGCTGTGGCAGAACCTATGCTGA
20 TTAGCCACTTTGGTGAAGCTATCATTGAAGAGGTTTTTTGCCGCTACCAGCAAATCTT
GGCTGAACGTATGTCCAAGGAGAAAACCAAGTTCATCAATGTTACCATATTATTGACT
AGAAAAGCATAA (SEQ ID NO: 1). Mientras se describió una secuencia de ácidos nucleicos particular para *GmSAMT1*, se entiende que una secuencia de ácidos nucleicos *GmSAMT1* incluye cualquier secuencia de ácidos nucleicos redundante en virtud de la degeneración del código genético que codifica una proteína *GmSAMT1*, tal como
25 codificada por la SEQ ID NO: 1.

En una realización, una secuencia de ácidos nucleicos *GmSABP2-1* se proporciona como la SEQ ID NO: 2 a continuación:

ATGGGTTCACAAAATTGTATGGATAGGAAGCACTATGTTCTGGTGCATGGGGCATGC
30 CATGGAGCTTGGTGTGTTGGTATAAGCTCAAGCCACGCTTGAATCTGCAGGCCATAAG
GTCACAGTACTTGACCTTGCAGCTTCTGGAACCAACATGAAGAAAATTGAAGATGTTG
ATACTTTCTCAGAGTATTCTGCGCCTTTGTTGCAGCTAATGGCCACAATTCCCTCAAAT
GAGAAGTTAGTTCTAGTTGGTCACAGCCTTGGAGGGCTGAACATAGCACTTGCAATG
GAGAAATTCCCAGAAAAGGTAGCAGTTGGTGTGTTTCTTAACAGCTTTTGCTCCAGACA
35 CTGAACACCACCCATCTTATGTCTTGGAAGTACAATGAGAGGACCCCGTTAGCTGC

ATGGTTAGACACTGAATTTGCTCCAAGCGGAAACAAAACATCAATGTTCTTCGGCCCC
AACTTCTTGTCCGACAAGCTCTACCAACTGTCCCCTATTGAGGACTTGGAATTGGCCA
AGACTTTAGCAAGGCCATCATCGCTCTTCATGGAAGACTTGACTAAACAAAAGAACTT
CTCCAAAGAGGGATATGGGTCAAGTCCACGTGCCTTTATTGTTTGCACGGAGGACCT
5 TGAATTCCATTGGAATATCAGCTCTTGATGATCCAAAATGTTGGGTTCAATGACGTT
GTAGAGGTCAAAGACGCAGATCATATGGTTATGCTTTGCAAGCCACAAGAACTATTCTG
ATTCCCTCCAGCAGATAGCGACTAAATATGCATGA (SEQ ID NO: 2). Mientras una
secuencia de ácidos nucleicos particular se demostró para *GmSABP2-1*, se entiende que
una secuencia de ácidos nucleicos *GmSABP2-1* incluye cualquier secuencia de ácidos
10 nucleicos redundante en virtud de la degeneración del código genético que codifica una
proteína *GmSABP2-1*, tal como codificada por la SEQ ID NO: 2.

Esta divulgación también contempla variantes de los genes resistentes a las
plagas divulgados aislados de la soja. Las variantes de secuencias de nucleótidos
también incluyen secuencias de nucleótidos derivadas sintéticamente, tal como aquellas
15 generadas, por ejemplo, mediante el uso de mutagénesis dirigida al sitio, pero que cuando
se expresan aún presentan actividad de resistencia a las plagas. Los métodos para la
mutagénesis y modificaciones de la secuencia de nucleótidos son bien conocidos en la
técnica. Véase, por ejemplo, Kunkel (1985) *Proc. Natl Acad. Sci.* USA
52:488-492; Kunkel *et al.* (1987) *Methods in Enzymol.* 75:367-382; Patente de Estados
20 Unidos No. 4,873,192; Walker y Gaastra, eds. (1983) *Techniques in Molecular Biology*
(MacMillan Publishing Company, Nueva York) y las referencias allí citadas. Se entenderá
además que las secuencias de aminoácidos codificadas por los ácidos nucleicos
GmSABP2-1 y los ácidos nucleicos *GmSAMT1* tolerarán normalmente las sustituciones
en la secuencia de aminoácidos y retendrán la actividad biológica de forma sustancial. La
25 identificación habitual de sustituciones de aminoácidos biológicamente activos puede
basarse en cualquier característica conocida en la técnica, incluidas las diferencias o
similitudes relativas de los sustitutos de cadena lateral de aminoácidos, por ejemplo, su
hidrofobicidad, hidrofiliidad, carga, tamaño y similares. Generalmente, las variantes de
secuencias de nucleótidos codificarán una proteína que tenga al menos un 80% de
30 identidad de secuencia con respecto a la proteína codificada por un gen resistente a
plagas divulgado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al
menos un 95% al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99%
de identidad de secuencia o aún más con respecto a la proteína codificada por su
secuencia de nucleótidos del gen resistente a las plagas de referencia respectiva. Así, se
35 divulgan ácidos nucleicos *GmSABP2-1* que codifican una proteína con aproximadamente

un 80% de homología de secuencia con respecto a la proteína codificada por la SEQ ID NO: 2 o un ácido nucleico degenerado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad de secuencia o aún más con respecto a la proteína codificada por la SEQ ID NO: 2 o un ácido nucleico degenerado. También se divulgan ácidos nucleicos *GmSAMT1* que tienen aproximadamente un 80% de homología de secuencia con respecto a la proteína codificada por la SEQ ID NO: 1 o un ácido nucleico degenerado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad de secuencia o aún más con respecto a la proteína codificada por la SEQ ID NO: 1 o un ácido nucleico degenerado. En algunas realizaciones, un ácido nucleico que tiene aproximadamente 80% de homología de secuencia con respecto a la proteína codificada por la SEQ ID NO: 1 o un ácido nucleico degenerado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad de secuencia o aún más con respecto a la proteína codificada por la SEQ ID NO: 1 o un ácido nucleico degenerado no tiene una A en la posición número 627 respecto a la SEQ ID NO: 1. También se divulgan ácidos nucleicos de *GmSABP2-1* que tienen aproximadamente un 80% de homología de secuencia con respecto a la SEQ ID NO: 2 o un ácido nucleico degenerado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad de secuencia o aún mayor con respecto a la SEQ ID NO: 2 o un ácido nucleico degenerado. También se divulgan ácidos nucleicos de *GmSAMT1* que tienen un 80% de homología de secuencia con respecto a la SEQ ID NO: 1 o un ácido nucleico degenerado, tal como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad de secuencia o aún mayor con respecto a la SEQ ID NO: 1 o un ácido nucleico degenerado. En algunas realizaciones, el ácido nucleico que tiene aproximadamente un 80% de homología secuencial con respecto a la SEQ ID NO: 1 o al ácido nucleico degenerado, como al menos un 80%, al menos un 85%, al menos un 90%, al menos un 95%, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% de identidad secuencial o incluso mayor con respecto a la SEQ ID NO: 1 o al ácido nucleico degenerado, no tiene una A en la posición número 627 con respecto a la SEQ ID NO: 1.

En algunas realizaciones, el ácido nucleico resistente a plagas divulgado codifica un fragmento funcional de un ácido nucleico *GmSABP2-1* o *GmSAMT1*. Dichos

fragmentos funcionales de GmSABP2-1 y GmSAMT1 aún muestran actividad de resistencia a plagas. Los fragmentos funcionales incluyen proteínas en las que se han eliminado los residuos en el extremo N, extremo C y/o internos con respecto a la proteína de longitud completa.

5 La proteína codificada por el gen *GmSABP2-1* es un miembro de la superfamilia de las enzimas α/β hidrolasa, con la tríada catalítica Ser, Asp, His, cuya ubicación se muestra por las flechas en la figura 15. Además, la estructura de cristal de la enzima relacionada del tabaco ha sido resuelta para la resolución atómica (véase Forouhar *et al.*, *PNAS* 102 (5) 1773-1778, 2005). Por tanto, el entendido en la técnica puede producir
10 fragmentos y/o variantes funcionales de gmSABP2 (y ácidos nucleicos que codifican dichos fragmentos) por ejemplo mientras que mantiene la tríada catalítica. Dichos fragmentos pueden probarse para detectar su actividad con experimentación mínima dada la guía presentada en la memoria descriptiva. Por ejemplo, con base en la estructura de la superfamilia de enzimas α/β hidrolasa, el entendido en la técnica puede realizar
15 eliminaciones de unos pocos aminoácidos en el extremo N, en el extremo C y/o en los bucles internos de la proteína codificada por el gen *GmSABP2-1* sin afectar la tríada catalítica del sitio activo. Por ejemplo, una eliminación menor a aproximadamente 50, 40, 30, 25, 20, 15, 10, 5, 4, 3, 2, o 1 eliminaciones del extremo N, del extremo C y/o de los bucles internos mientras que se mantiene la tríada catalítica del sitio activo con pruebas
20 y/o experimentación mínima para determinar la actividad de la proteína resultante.

La proteína codificada por el gen *GmSAMT1* es una carboxil metiltransferasa dependiente de S-adenosil-L-metionina (SAM) denominada *GmSAMT1* que pertenece a la familia de las metiltransferasas SABATH. La estructura de cristal de los miembros de la familia SABATH ha sido resuelta para la resolución atómica y el sitio activo de la enzima
25 ha sido trazado (véase Zhao *et al. Plant Physiol.* 146(2): 455–467, 2008). Por tanto, el entendido en la técnica puede producir fragmentos y/o variantes funcionales de GmSAMT1 (y ácidos nucleicos que codifican dichos fragmentos) por ejemplo mientras que mantiene la tríada catalítica. Dichos fragmentos pueden probarse para determinar su actividad con experimentación mínima dada la guía presentada en la memoria descriptiva.
30 Por ejemplo, con base en la estructura de la familia de enzimas SABATH, el entendido en la técnica puede realizar eliminaciones de unos pocos aminoácidos en el extremo N, en el extremo C y/o en los bucles internos de la proteína codificada por el gen *GmSAMT1* sin afectar la tríada catalítica del sitio activo. Por ejemplo, una eliminación menor a aproximadamente 50, 40, 30, 25, 20, 15, 10, 5, 4, 3, 2, o 1 eliminaciones del extremo N,
35 del extremo C y/o de los bucles internos mientras que se mantiene la tríada catalítica del

sitio activo con pruebas y/o experimentación mínima para determinar la actividad de la proteína resultante.

B. Casetes de expresión

Las secuencias de nucleótidos para los genes resistentes a plagas divulgados, como una secuencia de ácidos nucleicos que codifica GmSAMT1 (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o *GmSABP2-1* como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su fragmento funcional, son útiles en la manipulación genética de cualquier planta para conferir resistencia a plagas, como insectos, hongos, bacterias, virus, nematodos, ácaros, garrapatas y similares, por ejemplo, SCN cuando se encuentra unido de forma operativa a un promotor, como un promotor procesable o constitutivo. El término “unido de forma operativa” significa que la transcripción o traducción de los genes resistentes a plagas divulgados se encuentra bajo la influencia de la secuencia promotora. De esta forma, las secuencias de nucleótidos para los genes resistentes a plagas divulgados se proporcionan en casetes de expresión para la expresión en la planta de interés. Dichos casetes de expresión típicamente comprenderán una región de inicio de transcripción que comprende una secuencia de nucleótidos promotora unida de forma operativa a uno o más de los genes resistentes a plagas divulgados o sus variantes. Dicho casete de expresión puede proporcionarse con una pluralidad de sitios de restricción para la inserción de la secuencia de nucleótidos que debe encontrarse bajo regulación transcripcional de las regiones reguladoras. El casete de expresión puede contener además genes o secuencias marcadoras seleccionables. Los casetes de expresión de la presente divulgación pueden formar parte de un vector de expresión, como un plásmido.

En algunas realizaciones, el casete de transcripción incluirá en la dirección 5' a 3' de la transcripción, una región de inicio de transcripción y traducción, un gen resistente a plagas, como una secuencia de ácidos nucleicos que codifica GmSAMT1 (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o *GmSABP2-1* como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su fragmento funcional, y una región de terminación de transcripción y traducción funcional en células vegetales. La región de terminación puede ser nativa con la región de inicio de transcripción, puede ser nativa con el gen resistente a plagas o puede derivar de otra fuente. Las regiones de terminación convenientes se encuentran disponibles en el plásmido Ti de *A. tumefaciens*, como las regiones de terminación de octopina sintasa y nopalina sintasa. Véase también, Guerineau *et al.*, *Mol. Gen. Genet.* 262:141-144, 1991; Proudfoot *Cell* 64:671-674, 1991; Sanfacon *et al.*, *Genes Dev.* 5:141-149, 1991; Mogen *et al.*, *Plant Cell* 2:1261-1272, 1990;

Munroe *et al.*, *Gene* 91 :151-158, 1990; Ballas *et al.*, *Nucleic Acids Res.* 17:7891-7903, 1989; Joshi *et al.*, *Nucleic Acid Res.* 15:9627-9639, 1987.

El casete de expresión que incluye uno de los genes resistentes a plagas divulgados, como la secuencia de ácidos nucleicos que codifica GmSAMT1 (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o GmSABP2-1 como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su fragmento funcional, unido de forma operativa a una secuencia promotora también puede contener al menos una secuencia de nucleótidos adicional para que un gen se cotransforme en un organismo. De forma alternativa, la(s) secuencia(s) adicional(es) pueden proporcionarse en otro casete de expresión.

Cuando fuera adecuado, las secuencias de nucleótidos cuya expresión se desea, pueden optimizarse para una expresión superior en la planta transformada. Es decir, estas secuencias de nucleótidos pueden sintetizarse mediante el uso de codones vegetales preferidos para una expresión mejorada. Los métodos para sintetizar secuencias de nucleótidos preferidas de plantas se encuentran disponibles en la técnica. Véase, por ejemplo, las patentes de Estados Unidos No. 5,380,831 y 5,436,391 y Murray *et al.*, *Nucleic Acids Res.* 17:477-498, 1989.

Se conocen modificaciones de secuencias adicionales para mejorar la expresión de genes en un hospedador celular. Estas incluyen la eliminación de secuencias que codifican señales de poliadenilación falsas, señales de sitio de empalme exón-intrón, repeticiones similares a los transposones y otras secuencias bien caracterizadas que pueden ser perjudiciales para la expresión de genes. El contenido de G-C de la secuencia de nucleótidos heteróloga puede ajustarse a niveles promedio para un hospedador celular dado, como se calcula por referencia a genes conocidos expresados en la célula hospedadora. Cuando fuera posible, la secuencia se modifica para evitar las estructuras de ARNm secundarias en forma de horquilla predichas.

Los casetes de expresión pueden contener además secuencias líder 5' en la construcción del casete de expresión. Dichas secuencias líder pueden actuar para mejora la traducción. Los líderes de traducción son conocidos en la técnica e incluyen: líderes de picornavirus, por ejemplo, líder de EMCV (región no codificante 5' de encefalomiocarditis) (Elroy-Stein *et al. Proc. Nat. Acad. Sci.*, E.U.A. 86:6126-6130); líderes de potivirus, por ejemplo, líder de TEV (virus del grabado del tabaco); líder de MDMV (virus del mosaico enanizante del maíz); líder de la proteína de unión a inmunoglobulina de cadena pesada (BiP) (Macejak & Sarnow, *Nature* 353:90-94, 1991); líder sin traducir de la proteína de recubrimiento del ARNm del virus del mosaico de la alfalfa (AMV RNA 4; Jobling y

Gehrke, *Nature* 325:622-625, 1987); líder del mosaico del tabaco (TMV) (Gallie *et al.*, *Molecular Biology of RNA*, páginas 237 a 256, 1989); y líder del virus del moteado clorótico del maíz (MCMV) (Lommel *et al.*, *Virology* 81:382-385, 1991). Véase también, Della-Cioppa *et al.*, *Plant Physiology* 84:965-968, 1987. También pueden utilizarse otros
 5 métodos conocidos para mejorar la traducción y/o estabilidad del ARNm, por ejemplo, intrones y similares.

En los casos en que es deseable que el producto expresado del gen resistente a plagas se dirija a un organelo particular, como el cloroplasto o la mitocondria, o se secrete en la superficie de la célula o extracelularmente, el casete de expresión puede
 10 comprender además una secuencia codificante para un péptido de tránsito. Dichos péptidos de tránsito son bien conocidos en la técnica e incluyen, a modo no taxativo, el péptido de tránsito para la proteína vehículo de acilo, la subunidad pequeña de RUBISCO, la sintasa EPSP vegetal y similares.

En la preparación del casete de expresión, los varios fragmentos de ADN pueden
 15 manipularse mediante métodos conocidos en la técnica, de forma tal de proporcionar secuencias de ADN en la orientación adecuada y, según sea adecuado, en el marco de lectura adecuado. Para ello, pueden emplearse adaptadores o enlazadores para unir los fragmentos de ADN o pueden realizarse otras manipulaciones para prever sitios de restricción convenientes, eliminación de ADN superfluo, eliminación de sitios de
 20 restricción o similares. A estos efectos, puede realizarse mutagénesis *in vitro*, reparación de cebadores, restricción, recocido, resustituciones, por ejemplo, transiciones y transversiones.

Los casetes de expresión pueden incluir genes reporteros o genes marcadores seleccionables. Los ejemplos de genes reporteros adecuados conocidos en la técnica
 25 pueden encontrarse en, por ejemplo, Jefferson *et al.*, en *Plant Molecular Biology Manual*, edición de Gelvin *et al.* (Kluwer Academic Publishers), páginas 1 a 33, 1991; DeWet *et al.*, *Mol. Cell. Biol.* 7:725-737, 1987; Goff *et al.*, *EMBO J.* 9:2517-2522, 1990; y Kain *et al.*, *BioTechniques* 19:650- 655, 1995; y Chiu *et al.*, *Current Biology* 6:325-330, 1996. Los genes marcadores seleccionables para la selección de células o tejidos transformados
 30 pueden incluir genes que confieren resistencia a antibióticos o resistencia a herbicidas. Los ejemplos de genes marcadores seleccionables adecuados incluyen, a modo no taxativo, genes que codifican la resistencia al cloramfenicol (Herrera Estrella *et al.*, *EMBO J.* 2:987-992, 1983); metotrexato (Herrera Estrella *et al.*, *Nature* 303:209-213, 1983; Meijer *et al.*, *Plant Mol. Biol.* 16:807-820, 1991); higromicina (Waldron *et al.*, *Plant Mol. Biol.*
 35 5:103-108, 1985; Zhijian *et al.*, *Plant Science* 108:219-227, 1995); estreptomycin (Jones

et al., *Mol. Gen. Genet.* 210:86-91, 1987); espectinomicina (Bretagne-Sagnard et al., *Transgenic Res.* 5:131-137, 1996); bleomicina (Hille et al., *Plant Mol. Biol.* 7:171-176, 1990); sulfonamida (Guerineau et al., *Plant Mol. Biol.* 15:127-136, 1990); bromoxinil (Stalker et al., *Science* 242:419-423,1988); glifosato (Shaw et al., *Science* 233:478-481, 1986); y fosfinotricina (DeBlock et al., *EMBO J.* 6:2513-2518, 1987).

Otros genes que pueden servir de utilidad en la recuperación de eventos transgénicos pero que puede que no sean necesarios en el producto final incluirían, a modo no taxativo, ejemplos como GUS (b-glucoronidasa; Jefferson *Plant Mol. Biol. Rep.* 5:387, 1987), GFP y otras proteínas fluorescentes relacionadas y luciferasa. En algunas realizaciones, el casete de expresión es pJL-OFP-35S::GmSAMT1 o pJL-OFP-35S::GmSABP2-1. La siguiente es la secuencia del vector binario pJL-OFP-35S::GmSAMT1 que muestra la secuencia de GmSAMT1 complementada de forma inversa en negrita.

TCATGGCTCTTTCAAAGCAAAGTGGGGTCAAAGATGTAATGAACACCGAGCTTCATAT
15 GGACGGGATCGTCAATGGACACCCCTTTGAGATAAAAGGAAAAGGAAAGGGAAACCC
GTACAAGGGTGTGCAGACCATGAACTTACAGTCATTAAGGGTGCGCCTTTGCCATTT
TCTATTGACATTTTGCTGCCTCAACACATGTATGGAAGCAAGCCATTTATTAAGTATCC
TGAGAGTATCCCAGACTACATCAAGTTGTCATTTCCCGAGGGAATCACATGGGAAAG
GTCCATGACCTTTGAAGATGGTGCAGTGTGCACTGTCTCTAACGACTCCAGGCTCGA
20 TGGCGACTCTTTCATCTACGAAGTCAGGTTTCTTGGCGTGAACCTTTCCCGAGATGGA
CCTGTTATGCAGAAGAAGACGCGAGGCTGGGACCCGTCCACAGAGAGACTGTATGA
GTGTGGTGGGTGGCAGAGAGGAGACGTCCACATGGCCTTGAAGTTGGAGAACGGTG
GCCATTATACGTGCGACTTCAAACTACTTACAAATCAAAGAAGGGCTTGAAGGTGCC
ACCGTATCACTTCGTTGACCACAACTAGATTTACTGAGCCACAACACCGATGGTGCT
25 ACCTTTGAAGAGTTTGAACAACGAGAAATTGCACATGCACATCTTTCTAACTTACCGG
TAGCCCATCACCATCACCATCACTAAGATCGTTCAAACATTTGGCAATAAAGTTTCTTA
AGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGT
TAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGA
TTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAA
30 CTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCCCATGGCTACTACT
AAGCATTTGGCTCTTGCCATCCTTGTCCTCCTTAGCATTGGTATGACCACCAGTGCAA
GAACCCTCCTAGATCTGAGGGTAAATTTCTAGTTTTTCTCCTTCATTTTCTTGGTTAGG
ACCCTTTTCTCTTTTTATTTTTTTGAGCTTTGATCTTTCTTTAACTGATCTATTTTTTAA
TTGATTGGTTATGGTGTAATATTACATAGCTTTAACTGATAATCTGATTACTTTATTTT
35 GTGTGTCTATGATGATGATGATAGTTACAGAACCGACGAAGTCTGTACCCGATCA

ACACCGAGACCCGTGGCGTCTTCGACCTCAATGGCGTCTGGAACCTCAAGCTGGACT
 ACGGGAAAGGACTGGAAGAGAAGTGGTACGAAAGCAAGCTGACCGACACTATTAGTA
 TGGCCGTCCCAAGCAGTTACAATGACATTGGCGTGACCAAGGAAATCCGCAACCATA
 TCGGATATGTCTGGTACGAACGTGAGTTCACGGTGCCGGCCTATCTGAAGGATCAGC
 5 GTATCGTGCTCCGCTTCGGCTCTGCAACTCACAAAGCAATTGTCTATGTCAATGGTGA
 GCTGGTTCGTGGAGCACAAAGGGCGGATTCCTGCCATTCTGAAGCGGAAATCAACAACCTC
 GCTGCGTGATGGCATGAATCGCGTCACCGTCGCCGTGGACAACATCCTCGACGATA
 GCACCCTCCCGGTGGGGCTGTACAGCGAGCGCCACGAAGAGGGCCTCGGAAAAGT
 CATTCGTAACAAGCCGAACCTTCGACTTCTTCAACTATGCAGGCCTGCACCGTCCGGT
 10 GAAAATCTACACGACCCCGTTTACGTACGTGAGGACATCTCGGTTGTGACCGACTT
 CAATGGCCCAACCGGGACTGTGACCTATACGGTGGACTTTCAAGGCAAAGCCGAGAC
 CGTGAAAGTGTTCGGTCGTGGATGAGGAAGGCAAAGTGGTCGCAAGCACCGAGGGCC
 TGAGCGGTAACGTGGAGATTCCGAATGTCATCCTCTGGGAACCACTGAACACGTATC
 TCTACCAGATCAAAGTGGAACCTGGTGAACGACGGACTGACCATCGATGTCTATGAAG
 15 AGCCGTTTCGGCGTGCGGACCGTGGAAGTCAACGACGGCAAGTTCCTCATCAACAAC
 AAACCGTTCTACTTCAAGGGCTTTGGCAAACATGAGGACACTCCTATCAACGGCCGT
 GGCTTTAACGAAGCGAGCAATGTGATGGATTTCAATATCCTCAAATGGATCGGCGCC
 AACAGCTTCCGGACCGCACACTATCCGTA CTCTGAAGAGTTGATGCGTCTTGCGGAT
 CGCGAGGGTCTGGTCGTGATCGACGAGACTCCGGCAGTTGGCGTGACCTCAACTT
 20 CATGGCCACCACGGGACTCGGCGAAGGCAGCGAGCGCGTCAGTACCTGGGAGAAG
 ATTCGGACGTTTGAGCACCATCAAGACGTTCTCCGTGAACTGGTGTCTCGTGACAAG
 AACCATCCAAGCGTCGTGATGTGGAGCATCGCCAACGAGGCGGCGACTGAGGAAGA
 GGGCGCGTACGAGTACTTCAAGCCGTTGGTGGAGCTGACCAAGGAACTCGACCCAC
 AGAAGCGTCCGGTCACGATCGTGCTGTTTGTGATGGCTACCCCGGAGACGGACAAA
 25 GTCGCCGAACTGATTGACGTCATCGCGCTCAATCGCTATAACGGATGGTACTTCGAT
 GGCGGTGATCTCGAAGCGGCCAAAGTCCATCTCCGCCAGGAATTTACGCGTGGA
 CAAGCGTTGCCAGGAAAGCCGATCATGATCACTGAGTACGGCGCAGACACCGTTG
 CGGGCTTTACGACATTGATCCAGTGATGTTACCGAGGAATATCAAGTCGAGTACTA
 CCAGGCGAACCACGTCGTGTTGATGAGTTTGAGAACTTCGTGGGTGAGCAAGCGTG
 30 GAACTTCGCGGACTTCGCGACCTCTCAGGGCGTGATGCGCGTCCAAGGAAACAAGA
 AGGGCGTGTTCACTCGTGACCGCAAGCCGAAGCTCGCCGCGCACGTCTTTCGCGAG
 CGCTGGACCAACATTCCAGATTTCCGGCTACAAGAACGCTAGCCATCACCATCACCAT
 CACGTGTGAATTGGTGACCAGCTCGAATTTCCCGATCGTTCAAACATTTGGCAATAA
 AGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
 35 GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGG

TTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAG
 CGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGGGAAT
 TAAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTT
 TATTAGAATAACGGATATTTAAAAGGGCGTGAAAAGGTTTATCCGTTTCGTCCATTTGTA
 5 TGTGCATGCCAACACAGGGTTCCCCTCGGGATCAAAGTACTTTGATCCAACCCCTC
 CGCTGCTATAGTGCAGTCGGCTTCTGACGTTCAGTGCAGCCGTCTTCTGAAAACGAC
 ATGTCGCACAAGTCCTAAGTTACGCGACAGGCTGCCGCCCTGCCCTTTTCCTGGCGT
 TTTCTTGTCGCGTGTTTTAGTCGCATAAAGTAGAATACTTGCGACTAGAACCGGAGAC
 ATTACGCCATGAACAAGAGCGCCGCCGCTGGCCTGCTGGGCTATGCCCCGCGTCAGC
 10 ACCGACGACCAGGACTTGACCAACCAACGGGCCGAACCTGCACGCGGCCGGCTGCAC
 CAAGCTGTTTTCCGAGAAGATCACCGGCACCAGGCGCGACCGCCCGGAGCTGGCCA
 GGATGCTTGACCACCTACGCCCTGGCGACGTTGTGACAGTGACCAGGCTAGACCGC
 CTGGCCCGCAGCACCCGCGACCTACTGGACATTGCCGAGCGCATCCAGGAGGCCG
 GCGCGGGCCTGCGTAGCCTGGCAGAGCCGTGGGCCGACACCACCACGCCGGCCGG
 15 CCGCATGGTGTTGACCGTGTTCCGCCGGCATTGCCGAGTTCGAGCGTTCCTAATCAT
 CGACCGCACCCGGAGCGGGCGCGAGGCCGCCAAGGCCCGAGGCGTGAAGTTTGGC
 CCCCCGCCCTACCCTCACCCCGGCACAGATCGCGCACGCCCGCGAGCTGATCGACCA
 GGAAGGCCGCACCGTGAAAGAGGCGGCTGCACTGCTTGGCGTGTCATCGCTCGACCC
 TGTACCGCGCACTTGAGCGCAGCGAGGAAGTGACGCCACCGAGGCCAGGCGGGC
 20 CGGTGCCTTCCGTGAGGACGCATTGACCGAGGCCGACGCCCTGGCGGCCGCCGAG
 AATGAACGCCAAGAGGAACAAGCATGAAACCGCACCAAGGACGGCCAGGACGAACCG
 TTTTTCATTACCGAAGAGATCGAGGCGGAGATGATCGCGGCCGGGTACGTGTTTCGAG
 CCGCCCGCGCACGTCTCAACCGTGCGGCTGCATGAAATCCTGGCCGGTTTGTCTGAT
 GCCAAGCTGGCGGCCTGGCCGGCCAGCTTGGCCGCTGAAGAAACCGAGCGCCGCC
 25 GTCTAAAAAGGTGATGTGTATTTGAGTAAACAGCTTGCGTCATGCGGTCGCTGCGTA
 TATGATGCGATGAGTAAATAAACAATACGCAAGGGGAACGCATGAAGGTTATCGCT
 GTACTTAACCAGAAAGGCGGGTCAGGCAAGACGACCATCGCAACCCATCTAGCCCG
 CGCCCTGCAACTCGCCGGGGCCGATGTTCTGTAGTCGATTCCGATCCCCAGGGCA
 GTGCCCCGCGATTGGGCGGCCGTGCGGGAAGATCAACCGCTAACCGTTGTCGGCATC
 30 GACCGCCCGACGATTGACCGCGACGTGAAGGCCATCGGCCGGCGCGACTTCGTAGT
 GATCGACGGAGCGCCCCAGGCGGCGGACTTGGCTGTGTCCGCGATCAAGGCAGCC
 GACTTCGTGCTGATTCCGGTGCAGCCAAGCCCTTACGACATATGGGCCACCGCCGAC
 CTGGTGGAGCTGGTTAAGCAGCGCATTGAGGTCACGGATGGAAGGCTACAAGCGGC
 CTTTGTCTGTGTCGCGGGCGATCAAAGGCACGCGCATCGGCGGTGAGGTTGCCGAGG
 35 CGCTGGCCGGGTACGAGCTGCCATTCTTGAGTCCCGTATCACGCAGCGCGTGAGC

TACCCAGGCACTGCCGCCGCCGGCACAACCGTTCTTGAATCAGAACCCGAGGGCGA
 CGCTGCCCCGCGAGGTCCAGGCGCTGGCCGCTGAAATTAAATCAAACTCATTTGAGT
 TAATGAGGTAAAGAGAAAAATGAGCAAAAGCACAAACACGCTAAGTGCCGGCCGTCCG
 AGCGCACGCAGCAGCAAGGCTGCAACGTTGGCCAGCCTGGCAGACACGCCAGCCAT
 5 GAAGCGGGTCAACTTTTCAGTTGCCGGCGGAGGATCACACCAAGCTGAAGATGTACG
 CGGTACGCCAAGGCAAGACCATTACCGAGCTGCTATCTGAATACATCGCGCAGCTAC
 CAGAGTAAATGAGCAAATGAATAAATGAGTAGATGAATTTTAGCGGCTAAAGGAGGCG
 GCATGGAAAATCAAGAACAACCAGGCACCGACGCCGTGGAATGCCCCATGTGTGGA
 GGAACGGGCGGTTGGCCAGGCGTAAGCGGCTGGGTTGTCTGCCGGCCCTGCAATG
 10 GCACTGGAACCCCCAAGCCCGAGGAATCGGCGTGACGGTCGCAAACCATCCGGCCC
 GGTACAAATCGGCGCGGCGCTGGGTGATGACCTGGTGGAGAAGTTGAAGGCCGCGC
 AGGCCGCCCAGCGGCAACGCATCGAGGCAGAAGCACGCCCCGGTGAATCGTGGA
 AGCGGCCGCTGATCGAATCCGCAAAGAATCCCGGCAACCGCCGGCAGCCGGTGCG
 CCGTCGATTAGGAAGCCGCCCAAGGGCGACGAGCAACCAGATTTTTTCGTTCCGATG
 15 CTCTATGACGTGGGCACCCGCGATAGTCGCAGCATCATGGACGTGGCCGTTTTCCGT
 CTGTGAAGCGTGACCGACGAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGACGG
 GCACGTAGAGGTTTCCGCAGGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACC
 TGGTACTGATGGCGGTTTTCCCATCTAACCGAATCCATGAACCGATACCGGGAAGGGA
 AGGGAGACAAGCCCGGCCGCGTGTTCCGTCCACACGTTGCGGACGTACTCAAGTTC
 20 TGCCGGCGAGCCGATGGCGGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTG
 GTTAAACACCACGCACGTTGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCC
 TGGTGACGGTATCCGAGGGTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCG
 AAACCGGGCGGCCGGAGTACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAG
 ATCACAGAAGGCAAGAACCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTGATC
 25 GATCCCGGCATCGGCCGTTTTCTCTACCGCCTGGCACGCCGCGCCGCAGGCAAGGC
 AGAAGCCAGATGGTTGTTCAAGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGT
 TCAAGAAGTTCTGTTTCACCGTGCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGT
 ACGATTTGAAGGAGGAGGCGGGGCAGGCTGGCCCGATCCTAGTCATGCGCTACCGC
 AACCTGATCGAGGGCGAAGCATCCGCCGTTTCTAATGTACGGAGCAGATGCTAGG
 30 GCAAATTGCCCTAGCAGGGGAAAAAGGTGCAAAAGGTCTCTTTCCTGTGGATAGCAC
 GTACATTGGGAACCCAAAGCCGTACATTGGGAACCGGAACCCGTACATTGGGAACCC
 AAAGCCGTACATTGGGAACCGGTCACACATGTAAGTGACTGATATAAAAGAGAAAAAA
 GGCGATTTTTCCGCCTAAACTCTTTAAACTTATTAAACTCTTAAACCCGCCTGGC
 CTGTGCATAACTGTCTGGCCAGCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCT
 35 TCGGTCGCTGCGCTCCCTACGCCCCGCCGCTTCGCGTCGGCCTATCGCGGCCGCTG

GCCGCTCAAAAATGGCTGGCCTACGGCCAGGCAATCTACCAGGGCGCGGACAAGCC
 GCGCCGTCGCCACTCGACCGCCGGCGCCACATCAAGGCACCCTGCCTCGCGCGTT
 TCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTT
 GTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTT
 5 GCGGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAC
 TGGCTTAACTATGCGGCATCAGAGCAGATTGTAAGTGCACCATATGCGGTGT
 GAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCC
 TCGCTCACTGACTCGCTGCGCTCGGTGCTTCGGCTGCGGCGAGCGGTATCAGCTCA
 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACAT
 10 GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT
 TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGA
 GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC
 TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC
 CTTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
 15 AGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTTCAGCCCGACCGC
 TCGGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
 CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTA
 TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
 20 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGC
 GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
 GTGGAACGAAAACCTCACGTAAAGGATTTTGGTCATGCATTCTAGGTACTAAAACAAT
 TCATCCAGTAAAATATAATATTTTATTTTCTCCCAATCAGGCTTGATCCCCAGTAAGTC
 AAAAAATAGCTCGACATACTGTTCTTCCCGATATCCTCCCTGATCGACCGGACGCAG
 25 AAGGCAATGTCATACCACTTGTCGCGCCCTGCCGCTTCTCCCAAGATCAATAAAGCCAC
 TTAATTTGCCATCTTTCACAAAGATGTTGCTGTCTCCAGGTGCGCGTGGGAAAAGAC
 AAGTTCCTCTTCGGGCTTTTCCGTCTTTAAAAAATCATAACAGCTCGCGCGGATCTTTAA
 ATGGAGTGTCTTCTTCCAGTTTTTCGCAATCCACATCGGCCAGATCGTTATTCAGTAA
 GTAATCCAATTCGGCTAAGCGGCTGTCTAAGCTATTCGTATAGGGACAATCCGATATG
 30 TCGATGGAGTGAAAGAGCCTGATGCACTCCGCATACAGCTCGATAATCTTTTCAGGG
 CTTTGTTTCATCTTCATACTCTTCCGAGCAAAGGACGCCATCGGCCTCACTCATGAGCA
 GATTGCTCCAGCCATCATGCCGTTCAAAGTGCAGGACCTTTGGAACAGGCAGCTTTC
 CTTCCAGCCATAGCATCATGTCCTTTTCCCGTTCCACATCATAGGTGGTCCCTTTATA
 CCGGCTGTCCGTCAATTTTAAATATAGGTTTTTCAATTTTCTCCACCAGCTTATATACCT
 35 TAGCAGGAGACATTCCTTCCGTATCTTTTACGCAGCGGTATTTTTTCGATCAGTTTTTTC

AATTCGGTGATATTCTCATTTTAGCCATTTATTATTTCTTCCTCTTTTCTACAGTATT
 TAAAGATACCCCAAGAAGCTAATTATAACAAGACGAACTCCAATTCAGTTCCTTGCA
 TTCTAAACCTTAAATACCAGAAAACAGCTTTTTCAAAGTTGTTTTCAAAGTTGGCGTA
 TAACATAGTATCGACGGAGCCGATTTTGAAACCGCGGTGATCACAGGCAGCAACGCT
 5 CTGTCATCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCCG
 GCAGCTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCC
 TTACAACGGCTCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGT
 GGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGAT
 ATATTGTGGTGTAACAAATTGACGCTTAGACAACTTAATAACACATTGCGGACGTTTT
 10 TAATGTACTGAATTAACGCCGAATTAATTCGGGGGATCTGGATTTTAGTACTGGATTTT
 GGTTTTAGGAATTAGAAATTTTATTGATAGAAGTATTTTACAAATACAAATACATACTAA
 GGGTTTCTTATATGCTCAACACATGAGCGAAACCCTATAGGAACCCTAATTCCCTTAT
 CTGGGAACTACTCACACATTATTATGGAGAACTCGAGCTTGTCGATCGACAGATCCG
 GTCGGCATCTACTCTATTTCTTTGCCCTCGGACGAGTGCTGGGGCGTCGGTTTCCAC
 15 TATCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCGCTTCTGCGGGCG
 ATTTGTGTACGCCCCGACAGTCCCGGCTCCGGATCGGACGATTGCGTCGCATCGACC
 CTGCGCCCAAGCTGCATCATCGAAATTGCCGTCAACCAAGCTCTGATAGAGTTGGTC
 AAGACCAATGCGGAGCATATACGCCCCGAGTCGTGGCGATCCTGCAAGCTCCGGAT
 GCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAACCACGGCCTCCAG
 20 AAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCGCCTCGCTCCAGTCA
 ATGACCGCTGTTATGCGGCCATTGTCCGTCAGGACATTGTTGGAGCCGAAATCCGCG
 TGCACGAGGTGCCGGACTTCGGGGCAGTCCTCGGCCCAAGCATCAGCTCATCGAG
 AGCCTGCGCGACGGACGCACTGACGGTGTCGTCCATCACAGTTTGCCAGTGATACAC
 ATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCT
 25 TGCGGTCCGAATGGGCCGAACCCGCTCGTCTGGCTAAGATCGGCCGCAGCGATCGC
 ATCCATAGCCTCCGCGACCGGTTGTAGAACAGCGGGCAGTTCGGTTTCAGGCAGGT
 CTTGCAACGTGACACCCTGTGCACGGCGGGAGATGCAATAGGTCAGGCTCTCGCTAA
 ACTCCCCAATGTCAAGCACTTCCGGAATCGGGAGCGCGGCCGATGCAAAGTGCCGA
 TAAACATAACGATCTTTGTAGAAACCATCGGCGCAGCTATTTACCCGCAGGACATATC
 30 CACGCCCTCCTACATCGAAGCTGAAAGCACGAGATTCTTCGCCCTCCGAGAGCTGCA
 TCAGGTCGGAGACGCTGTCGAACTTTTCGATCAGAACTTCTCGACAGACGTCGCGG
 TGAGTTCAGGCTTTTTTCATATCTCATTGCCCCCGGGATCTGCGAAAGCTCGAGAGA
 GATAGATTTGTAGAGAGAGACTGGTGATTTAGCGTGTCTCTCAAATGAAATGAAC
 TTCCTTATATAGAGGAAGGTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTACG
 35 TCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTT

TTTCCACGATGCTCCTCGTGGGTGGGGGTCCATCTTTGGGACCACTGTCGGCAGAG
 GCATCTTGAACGATAGCCTTTCCTTTATCGCAATGATGGCATTGTAGGTGCCACCTT
 CCTTTTCTACTGTCCTTTTGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAG
 GTTCCCGATATTACCCTTTGTTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTG
 5 TATCTTTGATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTATCACATCA
 ATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTG
 GGTGGGGGTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTGAACGATAGCCTTT
 CCTTTATCGCAATGATGGCATTGTAGGTGCCACCTTCCTTTTCTACTGTCCTTTTGAT
 GAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTG
 10 TTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGT
 AGACGAGAGTGTCGTGCTCCACCATGTTGGCAAGCTGCTCTAGCCAATACGCAAACC
 GCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGA
 CTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGC
 ACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGA
 15 TAACAATTTACACAGGAAACAGCTATGACCATGATTACGAATCCCGATCTAGTAAC
 ATAGATGACACCGCGCGCGATAATTTATCCTAGTTTGCGCGCTATATTTTGTCTTCTAT
 CGCGTATTAAATGTATAATTGCGGGACTCTAATCATAAAAACCCATCTCATAAATAACG
 TCATGCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAATTATATGATAA
 TCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAACTTTATTGCCAAATGTTTGA
 20 ACGATCGGGGAAATTCGAGCTCTTATGCTTTTCTAGTCAATAATATGGTAACATTGAT
GAACTTGGTTTTCTCCTTGGACATACGTTACAGCCAAGATTTGCTGGTAGCGGCAAAA
AACCTCTTCAATGATAGCTTCACCAAAGTGGCTAATCAGCATAGGTTCTGCCACAGC
CCTCATGCACTGTGCCACATTGTATCCACCATCACTAAGTGATTACAGACCTTTCAGA
TTCAAAGTCTAGAGCATTCCACTCATCGAGAGCATTCCAATTCACCTCAGAAACCTC
 25 TAGACGATTGATGGCGAATGATCCTTCTTTAAGAACTTCCAATTTCACTTCAGATGG
 GGATGGAGTGTATTGAGGGATGTTAAAAGTATCTAATTGCTCTTCTCTTATGATTCC
 CTGCAAGACCATATCACTAAGAGCAGTAGCCAAAAGCTCCCAAATGTAGCAACCAT
 CCTTGCTAGATGGATCATCGCTTCTTCTTCCCAAAAATGTGAGAACCATGCGTCCTC
 CTTCAACCAATTCCTCTGCACGACACTTGAGAAAAAGAGAGAAATCTCTCTGAAATT
 30 GCTCATAGTAAGCTCTCACAACATTTTTGGGGCTCGTACTGCCTATGTAAACATTGC
 CCTTGTTGTTGTCTACACCCTCAGGAACCTTAGATAGCCATTGAAGGCTGTATGAGG
 AATGGACAAAATGAAGACTTTGATATGGGAAAACCCTTCATAGAAAGAACCGGGA
 ACACCCGAGAAGTAGCATGGACCGATCCCCTTTCCATTTATCACACAATTTCTCT
 TTGAAGCTGTCAAGAGACTTGAAGATGTTGTTGAAGTCATTCCCAGGGAGATCATTC
 35 AGAAAGACTTTGTATTCTGGAGATTTGTGGTTCAGCTCTCGGCAAAGCTTCTCCACA

ATTTTTATGAATTCAGACACAACAAATAAAGTGTTTGGTCCAGAAGAGCAACCCAAA
TCTGCAATTGCCAAGCTTCTGGGGTGTGTGCTGCAATAGAGGCTTCTTATGGCTTCC
TCTCTTATGGGCTTTGTCAAACAAATCACCTTTTGCTGAACAAGGGAGTTGTTTGCA
TAGCTTGCGTCTCCAACGCCACCGTTCATGTGGAGTACCTGTGCTACTTCCATGGAT
5 CCTCTAGAGTCCCCCGTGTTCTCTCAAATGAAATGAACTTCCTTATATAGAGGAAGG
GTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTACGTCAGTGGAGATATCACA
TCAATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTC
GTGGGTGGGGGTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTCAACGATGGC
CTTTCCTTTATCGCAATGATGGCATTGTAGGAGCCACCTTCCTTTTCCACTATCTTCA
10 CAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGGATATTACCC
TTTGTTGAAAAGTCTCAATTGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTTTG
GAGTAGACAAGTGTGTCGTGCTCCACCATGTTGACGAAGATTTTCTTCTTGTCATTGA
GTCGTAAGAGACTCTGTATGAACTGTTGCCAGTCTTTACGGCGAGTTCTGTTAGGTC
CTCTATTTGAATCTTTGACTCCATGGCCTTTGATTCAGTGGGAACTACCTTTTTAGAGA
15 CTCCAATCTCTATTACTTGCTTGGTTTGTGAAGCAAGCCTTGAATCGTCCATACTGG
AATAGTACTTCTGATCTTGAGAAATATATCTTTCTCTGTGTTCTTGATGCAGTTAGTCC
TGAATCTTTTGACTGCATCTTTAACCTTCTTGGGAAGGTATTTGATTTCTGGAGATTA
TTGCTCGGGTAGATCGTCTTGATGAGACCTGCTGCGTAAGCCTCTCTAACCATCTGTG
GGTTAGCATTCTTTCTGAAATTGAAAAGGCTAATCTGGGGACCTGCAGGCATGCAAG
20 CTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCA
ACTTAATCGCCTTGCAGCACATCCCCCTTCGCCAGCTGGCGTAATAGCGAAGAGGC
CCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGCTAGAGCA
GCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTAGCTTCATGGAGTCAAA
GATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGACTGGCGAACAGTTCATACA
25 GAGTCTCTTACGACTCAATGACAAGAAGAAAATCTTCGTCAACATGGTGGAGCACGAC
ACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGACCAAAGGGCAATTG
AGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGATTCCATTGCCAGCTAT
CTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTCCTACAAATGCCATCAT
TGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGACAGTGGTCCCAAAGAT
30 GGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACGTTCCAACCACGTCTTCA
AAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGATGACGCACAATCCCAC
TATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTTCATTTGGAGAGAACAC
GGGGGACTCTTGAC (SEQ ID NO: 3)

La siguiente es la secuencia del vector binario pJL-OFP-35S:: *GmSABP2-1* que
35 muestra la secuencia de *GmSABP2-1* complementada de forma inversa en negrita.

TCATGGCTCTTTCAAAGCAAAGTGGGGTCAAAGATGTAATGAACACCGAGCTTCATAT
 GGACGGGATCGTCAATGGACACCCCTTTGAGATAAAAGGAAAAGGAAAGGGAAACCC
 GTACAAGGGTGTGCAGACCATGAACTTACAGTCATTAAGGGTGCGCCTTTGCCATTT
 TCTATTGACATTTTGCTGCCTCAACACATGTATGGAAGCAAGCCATTTATTAAGTATCC
 5 TGAGAGTATCCCAGACTACATCAAGTTGTCATTTCCCGAGGGAATCACATGGGAAAG
 GTCCATGACCTTTGAAGATGGTGCAGTGTGCACTGTCTCTAACGACTCCAGGCTCGA
 TGGCGACTCTTTCATCTACGAAGTCAGGTTTCTTGGCGTGAACCTTTCCCGAGATGGA
 CCTGTTATGCAGAAGAAGACGCGAGGCTGGGACCCGTCCACAGAGAGACTGTATGA
 GTGTGGTGGGTGGCAGAGAGGAGACGTCCACATGGCCTTGAAGTTGGAGAACGGTG
 10 GCCATTATACGTGCGACTTCAAACACTTACAAATCAAAGAAGGGCTTGAAGGTGCC
 ACCGTATCACTTCGTTGACCACAACTAGATTTACTGAGCCACAACACCGATGGTGCT
 ACCTTTGAAGAGTTTGAACAACGAGAAATTGCACATGCACATCTTTCTAACTTACCGG
 TAGCCCATCACCATCACCATCACTAAGATCGTTCAAACATTTGGCAATAAAGTTTCTTA
 AGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTTGAATTACGT
 15 TAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGGTTTTTATGA
 TTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAGCGCGCAAA
 CTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCCCATGGCTACTACT
 AAGCATTTGGCTCTTGCCATCCTTGTCCTCCTTAGCATTGGTATGACCACCAGTGCAA
 GAACCCTCCTAGATCTGAGGGTAAATTTCTAGTTTTTCTCCTTCATTTTCTTGGTTAGG
 20 ACCCTTTTCTCTTTTTATTTTTTTGAGCTTTGATCTTTCTTTAACTGATCTATTTTTTAA
 TTGATTGGTTATGGTGTAATATTACATAGCTTTAACTGATAATCTGATTACTTTATTTT
 GTGTGTCTATGATGATGATGATAGTTACAGAACCGACGAAGTCTGTACCCGATCA
 ACACCGAGACCCGTGGCGTCTTCGACCTCAATGGCGTCTGGAACCTTCAAGCTGGACT
 ACGGGAAAGGACTGGAAGAGAAGTGGTACGAAAGCAAGCTGACCGACACTATTAGTA
 25 TGGCCGTCCCAAGCAGTTACAATGACATTGGCGTGACCAAGGAAATCCGCAACCATA
 TCGGATATGTCTGGTACGAACGTGAGTTCACGGTGCCGGCCTATCTGAAGGATCAGC
 GTATCGTGCTCCGCTTCGGCTCTGCAACTCACAAAGCAATTGTCTATGTCAATGGTGA
 GCTGGTTCGTGGAGCACAAGGGCGGATTCCTGCCATTCTGAAGCGGAAATCAACAACCTC
 GCTGCGTGATGGCATGAATCGCGTCACCGTCGCCGTGGACAACATCCTCGACGATA
 30 GCACCCTCCCGGTGGGGCTGTACAGCGAGCGCCACGAAGAGGGCCTCGGAAAAGT
 CATTCTGTAACAAGCCGAACCTTCGACTTCTTCAACTATGCAGGCCTGCACCGTCCGGT
 GAAAATCTACACGACCCCGTTTACGTACGTGAGGACATCTCGGTTGTGACCGACTT
 CAATGGCCCAACCGGGACTGTGACCTATACGGTGGACTTTCAAGGCAAAGCCGAGAC
 CGTGAAAGTGTGCGTCGTGGATGAGGAAGGCAAAGTGGTCGCAAGCACCGAGGGCC
 35 TGAGCGGTAACGTGGAGATTCCGAATGTCATCCTCTGGGAACCACTGAACACGTATC

TCTACCAGATCAAAGTGGAAGTGGTGAACGACGGACTGACCATCGATGTCTATGAAG
 AGCCGTTTCGGCGTGCGGACCGTGGAAGTCAACGACGGCAAGTTCCTCATCAACAAC
 AAACCGTTCTACTTCAAGGGCTTTGGCAAACATGAGGACACTCCTATCAACGGCCGT
 GGCTTTAACGAAGCGAGCAATGTGATGGATTTCAATATCCTCAAATGGATCGGCGCC
 5 AACAGCTTCCGGACCGCACACTATCCGTACTCTGAAGAGTTGATGCGTCTTGCGGAT
 CGCGAGGGTCTGGTCGTGATCGACGAGACTCCGGCAGTTGGCGTGACCTCAACTT
 CATGGCCACCACGGGACTCGGCGAAGGCAGCGAGCGCGTCAGTACCTGGGAGAAG
 ATTCGGACGTTTGAGCACCATCAAGACGTTCTCCGTGAACTGGTGTCTCGTGACAAG
 AACCATCCAAGCGTCGTGATGTGGAGCATCGCCAACGAGGCGGCGACTGAGGAAGA
 10 GGGCGCGTACGAGTACTTCAAGCCGTTGGTGGAGCTGACCAAGGAACTCGACCCAC
 AGAAGCGTCCGGTCACGATCGTGCTGTTTGTGATGGCTACCCCGGAGACGGACAAA
 GTCGCCGAACTGATTGACGTCATCGCGCTCAATCGCTATAACGGATGGTACTTCGAT
 GGCGGTGATCTCGAAGCGGCCAAAGTCCATCTCCGCCAGGAATTTACGCGTGGA
 CAAGCGTTGCCCAGGAAAGCCGATCATGATCACTGAGTACGGCGCAGACACCGTTG
 15 CGGGCTTTCACGACATTGATCCAGTGATGTTACCGAGGAATATCAAGTCGAGTACTA
 CCAGGCGAACCACGTCGTGTTTCGATGAGTTTGAGAACTTCGTGGGTGAGCAAGCGTG
 GAACTTCGCGGACTTCGCGACCTCTCAGGGCGTGATGCGCGTCCAAGGAAACAAGA
 AGGGCGTGTTCACTCGTGACCGCAAGCCGAAGCTCGCCGCGCACGTCTTTCGCGAG
 CGCTGGACCAACATTCCAGATTTCCGGCTACAAGAACGCTAGCCATCACCATCACCAT
 20 CACGTGTGAATTGGTGACCAGCTCGAATTTCCCGATCGTTCAAACATTTGGCAATAA
 AGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTATCATATAATTTCTGTT
 GAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGACGTTATTTATGAGATGGG
 TTTTATGATTAGAGTCCCGCAATTATACATTTAATACGCGATAGAAAACAAAATATAG
 CGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTCATCTATGTTACTAGATCGGGAAT
 25 TAACTATCAGTGTTTGACAGGATATATTGGCGGGTAAACCTAAGAGAAAAGAGCGTT
 TATTAGAATAACGGATATTTAAAAGGGCGTGAAAAGGTTTATCCGTTTCGTCCATTTGTA
 TGTGCATGCCAACCACAGGGTTCCCCTCGGGATCAAAGTACTTTGATCCAACCCCTC
 CGCTGCTATAGTGACGTCGGCTTCTGACGTTCAGTGCAGCCGTCTTCTGAAAACGAC
 ATGTCGCACAAGTCCTAAGTTACGCGACAGGCTGCCGCCCTGCCCTTTTCTGGCGT
 30 TTTCTTGTCGCGTGTTTTAGTCGCATAAAGTAGAATACTTGCGACTAGAACCGGAGAC
 ATTACGCCATGAACAAGAGCGCCGCCGCTGGCCTGCTGGGCTATGCCC GCGTCAGC
 ACCGACGACCAGGACTTGACCAACCAACGGGCCGAACCTGCACGCGGCCGGCTGCAC
 CAAGCTGTTTTCCGAGAAGATCACCGGCACCAGGCGCGACCGCCCGGAGCTGGCCA
 GGATGCTTGACCACCTACGCCCTGGCGACGTTGTGACAGTGACCAGGCTAGACCGC
 35 CTGGCCCGCAGCACCCGCGACCTACTGGACATTGCCGAGCGCATCCAGGAGGCCG

GCGCGGGCCTGCGTAGCCTGGCAGAGCCGTGGGCGGACACCACCACGCCGGCCGG
 CCGCATGGTGTGACCGTGTTGCGCCGGCATTGCCGAGTTCGAGCGTTCCTAATCAT
 CGACCGCACCCGGAGCGGGCGCGAGGCCGCCAAGGCCCGAGGCGTGAAGTTTGGC
 CCCCGCCCTACCCTCACCCCGGCACAGATCGCGCACGCCCGCGAGCTGATCGACCA
 5 GGAAGGCCGCACCGTGAAAGAGGCGGCTGCACTGCTTGCGGTGCATCGCTCGACCC
 TGTACCGCGCACTTGAGCGCAGCGAGGAAGTGACGCCACCGAGGCCAGGCGGCG
 CGGTGCCTTCCGTGAGGACGCATTGACCGAGGCCGACGCCCTGGCGGGCCGCCGAG
 AATGAACGCCAAGAGGAACAAGCATGAAACCGCACCCAGGACGGCCAGGACGAACCG
 TTTTTCATTACCGAAGAGATCGAGGCGGAGATGATCGCGGCCGGGTACGTGTTGAG
 10 CCGCCCGCGCACGTCTCAACCGTGCGGCTGCATGAAATCCTGGCCGGTGTGTCTGAT
 GCCAAGCTGGCGGCCTGGCCGGCCAGCTTGCCCGCTGAAGAAACCGAGCGCCGCC
 GTCTAAAAAGGTGATGTGTATTTGAGTAAAACAGCTTGCGTCATGCGGTGCTGCGTA
 TATGATGCGATGAGTAAATAAACAATACGCAAGGGGAACGCATGAAGGTTATCGCT
 GACTTAACCAGAAAGGCGGGTCAGGCAAGACGACCATCGCAACCCATCTAGCCCG
 15 CGCCCTGCAACTCGCCGGGGCCGATGTTCTGTAGTCGATTCCGATCCCCAGGGCA
 GTGCCCGCGATTGGGCGGCCGTGCGGGAAGATCAACCGCTAACCGTTGTCGGCATC
 GACCGCCCGACGATTGACCGCGACGTGAAGGCCATCGGCCGGCGCGACTTCGTAGT
 GATCGACGGAGCGCCCCAGGCGGCGGACTTGGCTGTGTCCGCGATCAAGGCAGCC
 GACTTCGTGCTGATTCCGGTGCAGCCAAGCCCTTACGACATATGGGCCACCGCCGAC
 20 CTGGTGGAGCTGGTTAAGCAGCGCATTGAGGTCACGGATGGAAGGCTACAAGCGGC
 CTTTGTGCTGTGCGGGGCGATCAAAGGCACGCGCATCGGCGGTGAGGTTGCCGAGG
 CGCTGGCCGGGTACGAGCTGCCCATTCTTGAGTCCCGTATCACGCAGCGCGTGAGC
 TACCCAGGCACTGCCGCCGCCGGCACAACCGTTCTTGAATCAGAACCCGAGGGCGA
 CGCTGCCCCGCGAGGTCCAGGCGCTGGCCGCTGAAATTAATCAAACTCATTTGAGT
 25 TAATGAGGTAAAGAGAAAATGAGCAAAAGCACAAACACGCTAAGTGCCGGCCGTCCG
 AGCGCACGCAGCAGCAAGGCTGCAACGTTGGCCAGCCTGGCAGACACGCCAGCCAT
 GAAGCGGGTCAACTTTCAGTTGCCGGCGGAGGATCACACCAAGCTGAAGATGTACG
 CGGTACGCCAAGGCAAGACCATTACCGAGCTGCTATCTGAATACATCGCGCAGCTAC
 CAGAGTAAATGAGCAAATGAATAAATGAGTAGATGAATTTTAGCGGCTAAAGGAGGCG
 30 GCATGGAAAATCAAGAACAACCAGGCACCGACGCCGTGGAATGCCCCATGTGTGGA
 GGAACGGGCGGTTGGCCAGGCGTAAGCGGCTGGGTGTCTGCCGGCCCTGCAATG
 GCACTGGAACCCCCAAGCCCGAGGAATCGGCGTGACGGTCGCAAACCATCCGGCCC
 GGTACAAATCGGCGCGGCGCTGGGTGATGACCTGGTGGAGAAGTTGAAGGCCGCGC
 AGGCCGCCAGCGGCAACGCATCGAGGCAGAAGCACGCCCCGGTGAATCGTGGA
 35 AGCGGCCGCTGATCGAATCCGCAAAGAATCCCGGCAACCGCCGGCAGCCGGTGCG

CCGTCGATTAGGAAGCCGCCCAAGGGCGACGAGCAACCAGATTTTTTCGTTCCGATG
 CTCTATGACGTGGGCACCCGCGATAGTCGCAGCATCATGGACGTGGCCGTTTTCCGT
 CTGTGGAAGCGTGACCGACGAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGACGG
 GCACGTAGAGGTTTCCGCAGGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACC
 5 TGGTACTGATGGCGGTTTTCCCATCTAACCGAATCCATGAACCGATACCGGGAAGGA
 AGGGAGACAAGCCCGGCCGCGTGTTCCGTCCACACGTTGCGGACGTACTCAAGTTC
 TGCCGGCGAGCCGATGGCGGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTCCG
 GTTAAACACCACGCACGTTGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCC
 TGGTGACGGTATCCGAGGGTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCG
 10 AAACCGGGCGGCCGGAGTACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAG
 ATCACAGAAGGCAAGAACCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTGATC
 GATCCCGGCATCGGCCGTTTTCTCTACCGCCTGGCACGCCGCGCCGCAGGCAAGGC
 AGAAGCCAGATGGTTGTTCAAGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGT
 TCAAGAAGTTCTGTTTCACCGTGCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGT
 15 ACGATTTGAAGGAGGAGGCGGGGCAGGCTGGCCCGATCCTAGTCATGCGCTACCGC
 AACCTGATCGAGGGCGAAGCATCCGCCGGTTCCTAATGTACGGAGCAGATGCTAGG
 GCAAATTGCCCTAGCAGGGGAAAAAGGTGCAAAAGGTCTCTTTCCTGTGGATAGCAC
 GTACATTGGGAACCCAAAGCCGTACATTGGGAACCGGAACCCGTACATTGGGAACCC
 AAAGCCGTACATTGGGAACCGGTACACATGTAAGTGACTGATATAAAAGAGAAAAAA
 20 GCGGATTTTTCCGCCTAAACTCTTTAAACTTATTAAACTCTTAAACCCGCCTGGC
 CTGTGCATAACTGTCTGGCCAGCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCCT
 TCGGTCGCTGCGCTCCCTACGCCCCGCGCTTCGCGTCGGCCTATCGCGGCCGCTG
 GCCGCTCAAAAATGGCTGGCCTACGGCCAGGCAATCTACCAGGGCGCGGACAAGCC
 GCGCCGTCGCCACTCGACCGCCGGCGCCACATCAAGGCACCCTGCCTCGCGCGTT
 25 TCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTT
 GTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTT
 GGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATAC
 TGGCTTAACTATGCGGCATCAGAGCAGATTGTAAGTGCAGAGTGCACCATATGCGGTGT
 GAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCC
 30 TCGCTCACTGACTCGCTGCGCTCGGTGTTTCGGCTGCGGCGAGCGGTATCAGCTCA
 CTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACAT
 GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGT
 TTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGA
 GGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCC
 35 TCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC

CTTCCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGT
 AGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCACCGC
 TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGC
 CACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCT
 5 ACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTA
 TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCG
 GCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAGATTACGC
 GCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
 GTGGAACGAAAACCTCACGTAAAGGGATTTTGGTCATGCATTCTAGGTACTAAAACAAT
 10 TCATCCAGTAAATATAATATTTTATTTTCTCCCAATCAGGCTTGATCCCCAGTAAGTC
 AAAAAATAGCTCGACATACTGTTCTTCCCCGATATCCTCCCTGATCGACCGGACGCAG
 AAGGCAATGTCATACCACTTGTCGCCCTGCCGCTTCTCCCAAGATCAATAAAGCCAC
 TTACTTTGCCATCTTTCACAAAGATGTTGCTGTCTCCAGGTCGCCGTGGGAAAAGAC
 AAGTTCCTCTTCGGGCTTTTCCGTCTTTAAAAAATCATACAGCTCGCGCGGATCTTTAA
 15 ATGGAGTGTCTTCTTCCAGTTTTTCGCAATCCACATCGGCCAGATCGTTATTCAGTAA
 GTAATCCAATTCGGCTAAGCGGCTGTCTAAGCTATTCGTATAGGGACAATCCGATATG
 TCGATGGAGTGAAAGAGCCTGATGCACTCCGCATACAGCTCGATAATCTTTTCAGGG
 CTTTGTTTCATCTTCATACTCTTCCGAGCAAAGGACGCCATCGGCCTCACTCATGAGCA
 GATTGCTCCAGCCATCATGCCGTTCAAAGTGCAGGACCTTTGGAACAGGCAGCTTTC
 20 CTTCCAGCCATAGCATCATGTCCTTTTCCCGTTCCACATCATAGGTGGTCCCTTTATA
 CCGGCTGTCCGTCATTTTTAAATATAGGTTTTTCATTTTCTCCCACCAGCTTATATACCT
 TAGCAGGAGACATTCCCTCCGTATCTTTACGCAGCGGTATTTTTCGATCAGTTTTTTC
 AATTCCGGTGATATTCTCATTTTAGCCATTTATTATTTCTTCTCTTTTCTACAGTATT
 TAAAGATACCCCAAGAAGCTAATTATAACAAGACGAACTCCAATTCAGTTCCTTGCA
 25 TTCTAAACCTTAAATACCAGAAAACAGCTTTTTCAAAGTTGTTTTCAAAGTTGGCGTA
 TAACATAGTATCGACGGAGCCGATTTTGAAACCGCGGTGATCACAGGCAGCAACGCT
 CTGTCATCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCCG
 GCAGCTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCC
 TTACAACGGCTCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGT
 30 GGTGATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGAT
 ATATTGTGGTGTAACAAATTGACGCTTAGACAACCTAATAACACATTGCGGACGTTTT
 TAATGTACTGAATTAACGCCGAATTAATTCGGGGGATCTGGATTTTAGTACTGGATTTT
 GGTTTTAGGAATTAGAAATTTTATTGATAGAAGTATTTTACAAATACAAATACATACTAA
 GGGTTTCTTATATGCTCAACACATGAGCGAAACCCTATAGGAACCCTAATTCCCTTAT
 35 CTGGGAACCTACTCACACATTATTATGGAGAACTCGAGCTTGTCGATCGACAGATCCG

GTCGGCATCTACTCTATTTCTTTGCCCTCGGACGAGTGCTGGGGCGTCGGTTTCCAC
 TATCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCGCTTCTGCGGGCG
 ATTTGTGTACGCCCCGACAGTCCCGGCTCCGGATCGGACGATTGCGTCGCATCGACC
 CTGCGCCCAAGCTGCATCATCGAAATTGCCGTCAACCAAGCTCTGATAGAGTTGGTC
 5 AAGACCAATGCGGAGCATATACGCCCCGAGTCGTGGCGATCCTGCAAGCTCCGGAT
 GCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAACCACGGCCTCCAG
 AAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCGCCTCGCTCCAGTCA
 ATGACCGCTGTTATGCGGCCATTGTCCGTGAGGACATTGTTGGAGCCGAAATCCGCG
 TGCACGAGGTGCCGGACTTCGGGGCAGTCCTCGGCCCAAAGCATCAGCTCATCGAG
 10 AGCCTGCGCGACGGACGCACTGACGGTGTGCTCCATCACAGTTTGCCAGTGATACAC
 ATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTAGTGTATTGACCGATTCTT
 TCGGGTCCGAATGGGCCGAACCCGCTCGTCTGGCTAAGATCGGCCGCGAGCGATCGC
 ATCCATAGCCTCCGCGACCGGTTGTAGAACAGCGGGCAGTTCGGTTTCAGGCAGGT
 CTTGCAACGTGACACCCTGTGCACGGCGGGAGATGCAATAGGTCAGGCTCTCGCTAA
 15 ACTCCCCAATGTCAAGCACTTCCGGAATCGGGAGCGCGGCCGATGCAAAGTGCCGA
 TAAACATAACGATCTTTGTAGAAACCATCGGCGCAGCTATTTACCCGCAGGACATATC
 CACGCCCTCCTACATCGAAGCTGAAAGCACGAGATTCTTCGCCCTCCGAGAGCTGCA
 TCAGGTGCGGAGACGCTGTGCAACTTTTCGATCAGAACTTCTCGACAGACGTCGCGG
 TGAGTTCAGGCTTTTTTCATATCTCATTGCCCCCGGGATCTGCGAAAGCTCGAGAGA
 20 GATAGATTTGTAGAGAGAGACTGGTGATTTGAGCGTGTCTCTCCAAATGAAATGAAC
 TTCCTTATATAGAGGAAGGTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTACG
 TCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTT
 TTTCCACGATGCTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGCAGAG
 GCATCTTGAACGATAGCCTTTTCTTTATCGCAATGATGGCATTGTAGGTGCCACCTT
 25 CCTTTTCTACTGTCCTTTTGATGAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAG
 GTTCCCGATATTACCCTTTGTTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTG
 TATCTTTGATATTCTTGGAGTAGACGAGAGTGTCGTGCTCCACCATGTTATCACATCA
 ATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTG
 GGTGGGGGTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTGAACGATAGCCTTT
 30 CCTTTATCGCAATGATGGCATTGTAGGTGCCACCTTCTTTTCTACTGTCCTTTTGAT
 GAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTG
 TTGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGAGT
 AGACGAGAGTGTCGTGCTCCACCATGTTGGCAAGCTGCTCTAGCCAATACGCAAACC
 GCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGA
 35 CTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGC

ACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGA
 TAACAATTTACACAGGAAACAGCTATGACCATGATTACGAATCCCGATCTAGTAAC
 ATAGATGACACCGCGCGCGATAATTTATCCTAGTTTGCGCGCTATATTTGTTTTCTAT
 CGCGTATTAAATGTATAATTGCGGGACTCTAATCATAAAAACCCATCTCATAAATAACG
 5 TCATGCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAATTATATGATAA
 TCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAACTTTATTGCCAAATGTTTGA
 ACGATCGGGGAAATTCGAGCTCT**CATGCATATTTAGTCGCTATCTGCTGGAGGGAAT**
CGAATAGTTCTTGTGGCTTGCAAAGCATAACCATATGATCTGCGTCTTTGACCTCTA
CAACGTCATTGAACCCAACATTTTGGATCATCAAGAGCTGATATTCCAATGGAATTC
 10 **CAAGGTCCTCCGTGCAAACAATAAAGGCACGTGGAAGTACCCATATCCCTCTTTG**
GAGAAGTTCTTTTGTAGTCAAGTCTTCCATGAAGAGCGATGATGGCCTTGCTAAA
GTCTTGGCCAATTCCAAGTCCTCAATAGGGGACAGTTGGTAGAGCTTGTGGGACAA
GAAGTTGGGGCCGAAGAACATTGATGTTTTGTTCCGCTTGGAGCAAATTCAGTGTC
TAACCATGCAGCTAACGGGGTCCTCTCATTGTACTTTTCCAAGACATAAGATGGGTG
 15 **GTGTTCAAGTGTCTGGAGCAAAAGCTGTTAAGAAAACACCAACTGCTACCTTTTCTGG**
GAATTTCTCCATTGCAAGTGCTATGTTCAAGCCTCCAAGGCTGTGACCAACTAGAAC
TAACTTCTCATTGAGGGAATTGTGGCCATTAGCTGCAACAAAGGCGCAGAATACT
CTGAGAAAGTATCAACATCTTCAATTTTCTTCATGTTGGTTCCAGAAGCTGCAAGGT
CAAGTACTGTGACCTTATGGCCTGCAGATTCCAAGCGTGGCTTGAGCTTATACCAA
 20 **CACCAAGCTCCATGGCATGCCCCATGCACCAGAACATAGTGCTTCCTATCCATACA**
ATTTTGTGAACCCATGGATCCTCTAGAGTCCCCCGTGTTCTCTCAAATGAAATGAAC
 TTCCTTATATAGAGGAAGGGTCTTGCGAAGGATAGTGGGATTGTGCGTCATCCCTTAC
 GTCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGACGTGGTTGGAACGTCTTCT
 TTTTCCACGATGCTCCTCGTGGGTGGGGTCCATCTTTGGGACCACTGTCGGCAGAG
 25 GCATCTTCAACGATGGCCTTTCCTTTATCGCAATGATGGCATTGTAGGAGCCACCTT
 CCTTTTCCACTATCTTCACAATAAAGTGACAGATAGCTGGGCAATGGAATCCGAGGAG
 GTTCCGGATATTACCCTTTGTTGAAAAGTCTCAATTGCCCTTTGGTCTTCTGAGACTG
 TATCTTTGATATTTTGGAGTAGACAAGTGTGTCGTGCTCCACCATGTTGACGAAGATT
 TTCTTCTTGTCATTGAGTCGTAAGAGACTCTGTATGAACTGTTGCCAGTCTTTACGG
 30 CGAGTTCTGTTAGGTCCTCTATTTGAATCTTTGACTCCATGGCCTTTGATTCAGTGGG
 AACTACCTTTTTAGAGACTCCAATCTCTATTACTTGCCTTGGTTTGTGAAGCAAGCCTT
 GAATCGTCCATACTGGAATAGTACTTCTGATCTTGAGAAATATATCTTTCTCTGTGTTT
 TTGATGCAGTTAGTCCTGAATCTTTGACTGCATCTTTAACCTTCTTGGGAAGGTATTT
 GATTTCTGAGATTATTGCTCGGGTAGATCGTCTTGATGAGACCTGCTGCGTAAGC
 35 CTCTCTAACCATCTGTGGGTTAGCATTCTTTCTGAAATTGAAAAGGCTAATCTGGGGA

CCTGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAA
ACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGC
GTAATAGCGAAGAGGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAAT
GGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTTTA
5 GCTTCATGGAGTCAAAGATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGACTG
GCGAACAGTTCATACAGAGTCTCTTACGACTCAATGACAAGAAGAAAATCTTCGTCAA
CATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAA
GACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGCTC
10 CTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGCCGA
CAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAGACG
TTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTAAGGGA
TGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAGTTCATTT
CATTTGGAGAGAACACGGGGGACTCTTGAC (SEQ ID NO: 4)

15 **C. Transgénicos**

Puede utilizarse un casete de expresión que incluya una secuencia de ácidos nucleicos que codifique el GmSAMT1 (como se divulga en la SEQ ID NO: 1 o su variante degenerada) y/o *GmSABP2-1* como se expresa en la SEQ ID NO: 2) o su fragmento funcional, unido de forma operativa a un promotor y opcionalmente otros ácidos nucleicos heterólogos, para transformar cualquier planta, por ejemplo, como un vector, como un plásmido. De esta forma, pueden obtenerse plantas, células vegetales, tejido vegetal, semillas y similares modificados genéticamente. Por lo tanto, se divulgan métodos para producir una planta resistente a una plaga. Dichos métodos incluyen introducir en una planta el *GmSAMT1* (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o el *GmSABP2-1* como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su fragmento funcional, unido de forma operativa a un promotor y opcionalmente a otros ácidos nucleicos heterólogos. Asimismo se divulgan métodos para aumentar la resistencia a plagas en una planta. Dichos métodos incluyen introducir en una planta el *GmSAMT1* (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o el *GmSABP2-1* como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su fragmento funcional, unido de forma operativa a un promotor y opcionalmente a otros ácidos nucleicos heterólogos, por lo que aumenta la resistencia a plagas en la planta. La planta puede ser una planta transformada de forma transitoria o estable. La resistencia a plagas, como un aumento en la resistencia a plagas puede determinarse con respecto a una planta de control relevante, como una planta que no expresa un polipéptido resistente a plagas

divulgado en la presente, una planta que no transformó con un ácido nucleico que codifica un polipéptido resistente a plagas como se divulga en la presente, una planta que se transformó con un ácido nucleico irrelevante y similares. La planta de control generalmente coincide en lo que respecta a la especie, variedad, edad y similar y se
 5 somete a las mismas condiciones de cultivo, por ejemplo, temperatura, suelo, luz solar, pH, agua y similares. La selección de una planta de control adecuada es rutina para los entendidos en la técnica.

Los protocolos de transformación así como los protocolos para introducir secuencias de nucleótidos en plantas pueden variar en función del tipo de planta o célula
 10 vegetal, por ejemplo, monocotiledónea, dicotiledónea, que se va a transformar. Los métodos adecuados para introducir secuencias de nucleótidos en células vegetales y su posterior inserción en el genoma de la planta incluyen microinyección (Crossway *et al.*, *Biotechniques* 4:320-334, 1986), electroporación (Riggs *et al.*, *Proc. Natl. Acad. Sci. USA* 53:5602-5606, 1986), transformación mediada por *Agrobacterium* (Patente de Estados
 15 Unidos No. 5,563,055), transferencia directa de genes (Paszkowski *et al.*, *EMBO J.* 3:2717-2722, 1984) y aceleración balística de partículas (véase, por ejemplo, la patente de Estados Unidos No. 4,945,050; Tomes *et al.* "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," en *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, edición de Gamborg y Phillips (Springer- Verlag, Berlin), 1995; y
 20 McCabe *et al.*, *Biotechnology* 5:923-926, 1988). Véase además Weissinger *et al.*, *Ann. Rev. Genet.* 22:421-477, 1988; Sanford *et al.*, *Paniculate Science and Technology* 5:27-37, 1987; Christou *et al.*, *Plant Physiol* 57:671-674, 1988; McCabe *et al.*, *Bio/Technology* 5:923-926, 1988; Finer y McMullen, *In Vitro Cell Dev. Biol.* 27P:175-182, 1991; Singh *et al.*, *Theor. Appl Genet.* 95:319-324, 1998; Datta *et al.*, *Biotechnology* 5:736-740, 1990;
 25 Klein *et al.*, *Proc. Natl. Acad. Sci. USA* 55:4305-4309, 1988; Klein *et al.*, *Biotechnology* 5:559-563, 1988; patentes de Estados Unidos No. 5,240,855, 5,322,783 y 5,324,646; Klein *et al.*, *Plant Physiol* 97:440-444, 1988; Fromm *et al.* *Biotechnology* 5:833-839, 1990; Hooykaas-Van Slogteren *et al.*, *Nature* 377:763-764, 1984; Bytebier *et al.*, *Proc. Natl. Acad. Sci. USA* 54:5345-5349, 1987; De Wet *et al.* (1985) en *The Experimental*
 30 *Manipulation of Ovule Tissues*, edición de Chapman *et al.* (Longman, Nueva York), páginas 197 a 209; Kaeppeler *et al.*, *Plant Cell Reports* 9:415-418, 1990; Kaeppeler *et al.*, *Theor. Appl. Genet.* 54:560-566, 1992; D'Halluin *et al.*, *Plant Cell* 4:1495-1505 1992; Li *et al.*, *Plant Cell Reports* 72:250-255, 1993; Christou y Ford *Annals of Botany* 75:407-413, 1995; Osjoda *et al.*, *Nature Biotechnology* 74:745-750, 1996; y similares. "Introducir" en el
 35 contexto de una célula vegetal, tejido vegetal, parte de planta y/o planta significa poner en

contacto una molécula de ácido nucleico con la célula vegetal, tejido vegetal, parte de planta y/o planta de forma tal que la molécula de ácido nucleico tenga acceso al interior de la célula vegetal o célula del tejido vegetal, parte de planta o planta. Cuando debe introducirse más de una molécula de ácido nucleico, estas moléculas de ácido nucleico
5 pueden ensamblarse como parte de un único polinucleótido o construcción de ácido nucleico o como polinucleótidos o construcciones de ácidos nucleicos separadas y pueden ubicarse en la misma construcción de ácido nucleico o en construcciones diferentes. Por consiguiente, estos polinucleótidos pueden introducirse en las células vegetales en un único evento de transformación, en eventos de transformación separados
10 o, por ejemplo, como parte de un protocolo de reproducción.

Las células que han sido transformadas pueden cultivarse en plantas de conformidad con las formas convencionales. Véase, por ejemplo, McCormick *et al.* *Plant Cell Reports* 5:81-84, 1986. Estas plantas a continuación pueden cultivarse y polinarse con la misma cepa transformada o con cepas diferentes y puede identificarse el híbrido
15 resultante que tiene la expresión de la característica fenotípica deseada. Pueden cultivarse dos o más generaciones para asegurar que la expresión de la característica fenotípica deseada se mantenga de forma estable y sea heredada y después las semillas pueden cosecharse para asegurarse que se ha logrado la expresión de la característica fenotípica deseada.

Los genes resistentes a plagas divulgados en la presente, como las secuencias de ácidos nucleicos que codifican el GmSAMT1 (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o el *GmSABP2-1* como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su variante o fragmentos activos, pueden utilizarse para la transformación de cualquier especie de planta, incluidas, a modo no taxativo, las
20 monocotiledóneas y las dicotiledóneas. Los ejemplos de especies de plantas de interés incluyen, a modo no taxativo, maíz (*Zea mays*), *Brassica sp.* (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), particularmente las especies de *Brassica* útiles como fuentes de aceites de semillas, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), mijo (por ejemplo, mijo perla (*Pennisetum glaucum*), mijo proso (*Panicum miliaceum*), mijo cola de zorro (*Setaria italica*), mijo africano (*Eleusine coracana*), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), papa (*Solanum tuberosum*), cacahuete (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), boniato (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea spp.*),
30 coco (*Cocos nucifera*), ananá (*Ananas comosus*), árboles cítricos (*Citrus spp.*), cacao

(*Theobroma cacao*), té (*Camellia sinensis*), banana (*Musa spp.*), aguacate (*Persea americana*), higo (*Ficus casica*), guayava (*Psidium guajava*), mango (*Mangifera indica*), olivo (*Olea europaea*), papaya (*Carica papaya*), anacardo (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendro (*Prunus amygdalus*), remolacha azucarera
 5 (*Beta vulgaris*), caña de azúcar (*Saccharum spp.*), avena, cebada, vegetales, plantas decorativas y coníferas. En realizaciones particulares, los genes resistentes a plagas divulgados en la presente, como las secuencias de ácidos nucleicos que codifican el GmSAMT1 (como se expresa en la SEQ ID NO: 1 o su variante degenerada) y/o el GmSABP2-1 como se expresa en la SEQ ID NO: 2 o su variante degenerada) o su
 10 variante y fragmentos activos, pueden utilizarse para la transformación de soja (*Glycine max*).

Los vegetales incluyen tomates (*Lycopersicon esculentum*), lechuga (por ejemplo, *Lactuca sativa*), frijoles verdes (*Phaseolus vulgaris*), frijoles de media luna (*Phaseolus limensis*), guisantes (*Lathyrus spp.*) y miembros del género *Cucumis* como pepino (*C.*
 15 *sativus*), cantalupo (*C. cantalupensis*) y melón amizcleño (*C. melo*).

Las plantas decorativas incluyen azalea (*Rhododendron spp.*), hortensia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasanensis*), rosas (*Rosa spp.*), tulipanes (*Tulipa spp.*), narcisos (*Narcissus spp.*), petunias (*Petunia hybrida*), claveles (*Dianthus caryophyllus*), flor de Pascua (*Euphorbia pulcherrima*) y crisantemo.

Las coníferas que pueden emplearse en la práctica de la presente invención incluyen, por ejemplo, pinos como pino taeda (*Pinus taeda*), pino elliotti (*Pinus elliottii*), pino ponderoso (*Pinus ponderosa*), pino contorto (*Pinus contorta*) y pino de Monterey (*Pinus radiata*); abeto de Douglas (*Pseudotsuga menziesii*); cicuta maculata (*Tsuga canadensis*); píceas de sika (*Picea glauca*); secuoya (*Sequoia sempervirens*); abetos como
 20 abeto del Pacífico (*Abies amabilis*) y abeto balsámico (*Abies balsamea*); y cedros como cedro rojo del Pacífico (*Thuja plicata*) y cedro de Alaska (*Chamaecyparis nootkatensis*) y álamo y eucalipto.

En realizaciones específicas, las plantas de la presente invención son plantas de cultivo (por ejemplo, maíz, alfalfa, girasol, Brassica, soja, algodón, cártamo, cacahuete,
 30 sorgo, trigo, mijo, tabaco, etc.). En otras realizaciones, las plantas de maíz y soja son óptimas y en otras realizaciones las plantas de soja son óptimas. Otras plantas de interés incluyen plantas de granos que proporcionan semillas de interés, plantas de semillas oleaginosas y plantas leguminosas. Las semillas de interés incluyen semillas de granos como maíz, algodón, cebada, arroz, sorgo, centeno, etc. Las plantas de semilla
 35 oleaginosa incluyen algodón, soja, cártamo, Brassica, maíz, alfalfa, palma, coco, etc. Las

plantas leguminosas incluyen frijoles y guisantes. Los frijoles incluyen guar, garrofín, fenogreco, soja, alubias, caupí, frijol mungo, frijol de media luna, habas, lentejas, garbanzos, etc.

En algunas realizaciones, los polinucleótidos que comprenden el gen resistente a plagas divulgado se diseñan como una agrupación molecular. Por tanto, las varias plantas, células vegetales y semillas divulgadas en la presente pueden comprender además uno o más rasgos de interés y en realizaciones más específicas, la planta, parte de planta o célula vegetal se agrupa con cualquier combinación de secuencias de polinucleótidos de interés para crear plantas con una combinación de rasgos deseada. Como se utiliza en la presente, el término “apilados” incluye el hecho de tener múltiples rasgos presentes en la misma planta.

Estas combinaciones apiladas pueden crearse mediante cualquier método que incluye, a modo no taxativo, la reproducción de plantas mediante cualquier metodología convencional o transformación genética. Si las secuencias se apilan mediante la transformación genética de las plantas, las secuencias de polinucleótidos de interés pueden combinarse en cualquier momento y en cualquier orden. Los rasgos pueden introducirse simultáneamente en un protocolo de cotransformación con los polinucleótidos de interés proporcionados en cualquier combinación de casetes de expresión. Por ejemplo, si van a introducirse dos secuencias, las dos secuencias pueden estar contenidas en casetes de transformación separados (trans) o contenidas en el mismo casete de transformación (cis). La expresión de las secuencias puede estar dirigida por el mismo promotor o por promotores diferentes. En ciertos casos, puede ser deseable introducir un casete de transformación que suprimirá la expresión del polinucleótido de interés. Este puede combinarse con cualquier combinación de otros casetes de supresión o casetes de sobreexpresión para generar la combinación de rasgos deseada en la planta. Se reconoce además que las secuencias de polinucleótidos pueden aplanarse en una ubicación genómica deseada mediante el uso de un sistema de recombinación específico para el sitio. Véase, por ejemplo, W099/25821, W099/25854, W099/25840, W099/25855 y W099/25853.

Las plantas transformadas pueden analizarse para determinar la presencia de gen(es) de interés y el nivel de expresión. Numerosos métodos se encuentran disponibles para los entendidos en la técnica para el análisis de plantas transformadas. Por ejemplo, los métodos para el análisis de plantas incluyen análisis de inmunotransferencia Southern y Northern, enfoques a base de PCR (o a base de otra amplificación de ácidos nucleicos), análisis bioquímicos, métodos de análisis fenotípico, evaluaciones de campo y ensayos

inmunodiagnósticos (por ejemplo, para la detección, localización y/o cuantificación de las proteínas).

EJEMPLOS

Ejemplo 1

5 Aislamiento y caracterización de GmSAMT1

Fuentes vegetales, nematódicas, bacterianas y químicas: En los estudios descritos en la presente, se utilizaron dos líneas de soja genéticamente relacionadas, TN02-226 y TN02-275, para la clonación de genes y la generación de raíces velludas de soja. Los detalles de reproducción para estas dos líneas de semillas de soja han sido descritos en Mazarei
 10 *et al. Theor. Appl. Genet.* 123: 1193-1206, 2011, que se incorpora específicamente en la presente a modo de referencia en su totalidad. Estas dos líneas hermanas derivadas de F₆, TN02-226 y TN02-275, son resistentes y susceptibles, respectivamente, al SCN raza 2 (HG tipo 1.2.5.7), lo que ha sido confirmado en las pruebas regionales del sur de la USDA del 2004 al 2007. En los estudios descritos en la presente, los huevos de SCN raza
 15 2 se utilizaron como los patógenos en el bioensayo de las raíces velludas transgénicas dado que las dos líneas de soja anteriores mostraron diferentes respuestas al SCN raza 2. El mantenimiento de SCN siguió el método descrito por Arelli *et al.*, *Heterodera glycines*. *Crop Sci.* 40: 824-826, 2000. La cepa BL21 de *Escherichia coli* (DE3) CODONPLUS® (Stratagene, La Jolla, CA) se utilizó para la expresión de la proteína recombinante. Los cultivos de *Agrobacterium*, *A. tumefaciens* GV3101 y la cepa K599 de *A. rhizogenes*, se utilizaron para la transformación transitoria del tabaco y la generación de las raíces velludas de soja. Todos los químicos se obtuvieron en Sigma-Aldrich (St. Louis, MO). *Inoculación de virus, genotipos de soja, y detección de SMV:* En los estudios descritos en la presente, se utilizó un mutante del virus del mosaico de la soja
 20 SMV-N (SMV-N25) para la infección con el virus en la soja. Se utilizaron dos genotipos de soja Williams 82, susceptible al SMV-N25 y L78-375, resistente al SMV-N25. Las plantas inoculadas mecánicamente se mantuvieron en una cámara de cultivo que operaba a 22 °C con un período de luz de 16 horas. Las plántulas de soja a las que se les inoculó el SMV se cultivaron por otros 21 días y sus hojas trifoliadas se cosecharon, pulverizaron
 25 en la presencia de nitrógeno líquido y se almacenaron a -80 °C hasta que se procesaron para el análisis de PCR cuantitativa con transcripción inversa.

Análisis de secuencias: La secuencia de la sonda de Affymetrix, Gma.12911.1.A1_S_AT, que identificó nuestro gen diana en un experimento de microensayo, se utilizó como una consulta a la búsqueda de BLAST contra la base de
 35 datos de soja GmGDB/GMtranscript (disponible en el sitio web plantgdb.org/GmGDB/cgi-

bin/blastGDB.pl). Se utilizó el programa BLASTN con un valor E de 1e-20. Cinco genes de soja mostraron una similitud alta (FIG. 2). Se realizó una alineación múltiple de secuencias de proteínas SABATH seleccionadas y Glyma02g06070 mediante el uso del programa ClustalX y se observó con el uso del software TreeView (disponible en el sitio web taxonomy.zoology.gla.ac.uk/rod/treeview.html). Se construyó un árbol filogenético sin raíces con el método Neighbor-Joining (FIG. 3).

Aislamiento del gen Glyma02g06070 de soja: Se aisló la totalidad del ARN del tejido de la raíz infectado con SCN de la línea de soja TN02-226, mediante el uso del mini kit vegetal RNEASY® (QIAGEN®, Valencia, CA, E.U.A.) y se eliminó la contaminación de ADN con un tratamiento con DNasa en columna (QIAGEN®, Valencia, CA, E.U.A.). A continuación se sintetizó el ADNc del ARN total en un volumen de reacción de 15 µl mediante el uso del kit de síntesis de la primera cadena de ADNc (GE Healthcare, Piscataway, NJ, USA) como se describió anteriormente. Con base en un análisis de microensayo, se seleccionó un gen candidato de soja *Glyma02g06070* para el estudio adicional. La secuencia de ADNc de longitud completa de *Glyma02g06070* se obtuvo de la base de datos de soja pública (disponible en el sitio web phytozome.net/soybean) y los cebadores se diseñaron de la siguiente forma ATGGAAGTAGCACAGGTACTCCACATG (SEQ ID NO: 5) y TGCTTTTCTAGTCAATAATATGGTAAC (SEQ ID NO: 6). Las condiciones de PCR fueron las siguientes: 94 °C durante 2 minutos seguidos por 35 ciclos a 94 °C durante 30 segundos, 57 °C durante 30 segundos y 72 °C durante 1 minuto 30 segundos y una extensión final a 72 °C durante 10 minutos. Los productos de la PCR se clonaron en el vector pEXP5/CT-TOPO®. Por secuenciado, el producto de PCR de soja TN02-226 fue el mismo que el predicho en la base de datos de soja.

Para verificar su función bioquímica, el gen candidato a continuación se clonó en un vector de construcción de expresión de proteínas, vector pET100/D-TOPO® (Invitrogen, Carlsband, CA) y se expresó en la cepa BL21 de *E. coli* (DE3) CODONPLUS® (Stratagene, La Jolla, CA). La expresión de proteínas se indujo con isopropil β-D-1-tiogalactopiranosida (IPTG) a una concentración de 500 µM durante 18 horas a 25 °C, donde las células se lisaron mediante sonicación. El *Glyma02g06070* expresado por *E. coli* con un marcador His se purificó a partir del lisado celular de *E. coli* mediante el uso de Ni-NTA agarosa siguiendo las instrucciones del fabricante (Invitrogen). La pureza de las proteínas se verificó mediante SDS-PAGE y las concentraciones de proteínas se determinaron con el ensayo de Bradford.

Ensayo radioquímico de actividad de SAMT: Se utilizó un volumen de 50 µl que contenía 50 mM Tris-HCl, pH 7.5, 1 mM SA y 3 µM de ¹⁴C-SAM con una actividad

específica de 51.4 mCi/mmol (Perkin-Elmer, Boston, MA). El ensayo se inició mediante la adición de SAM, se mantuvo a 25 °C durante 30 minutos y se detuvo mediante la adición de acetato de etilo (150 µl). Después de la separación de fases con 1 minuto de centrifugación a 14000 g, la fase orgánica superior se contó mediante el uso de un
 5 contador de centelleo líquido (Beckman Coulter, Fullerton, CA). Los conteos de radiación en la fase orgánica indicaron la cantidad de MeSA sintetizado. También se realizó un ensayo de especificidad de sustrato para el GmSAMT1 recombinante, con una variedad de sustratos que incluyó SA, ácido benzoico, ácido antranílico, ácido 3-hidroxibenzoico, ácido 4-hidroxibenzoico, ácido nicotínico, ácido p-coumárico, ácido cafeico, ácido
 10 cinámico, ácido vanílico, ácido jasmónico (JA), ácido giberélico, ácido indol-3-acético y ácido 2,4-diclorofenoxiacético. La actividad con SA se ajustó a un 100%. Se llevaron a cabo tres ensayos independientes para cada compuesto.

Determinación de los parámetros cinéticos de GmSAMT1: Las concentraciones de enzimas y el tiempo de incubación adecuados se determinaron en ensayos temporales
 15 para asegurarse de que la velocidad de reacción fuera lineal durante el período de ensayo. Las gráficas de Lineweaver-Burk proporcionaron valores de Km aparentes. Para determinar el Km para SA, las concentraciones de SA variaron independientemente de 2 a 150 µM, mientras que SAM se mantuvo constante a 200 µM. Los ensayos se llevaron a cabo a 25 °C durante 30 minutos, como se describe en el ensayo de actividad de SAMT. El pH óptimo para GmSAMT1 se probó de pH 6.0 a pH 9.0 y la reacción óptima de temperatura se probó entre 0 °C y 60 °C. Los efectos de los iones metálicos sobre la actividad de GmSAMT1 se probaron con una variedad de iones metálicos, incluidos K⁺, Cu²⁺, Zn²⁺, Fe²⁺, Fe³⁺, Mn²⁺, Ca²⁺, Na⁺, NH₄⁺ y Mg²⁺, en la forma de sales de cloruro a concentraciones finales de 5 mM.

Creación de las construcciones con sobreexpresión de GmSAMT1 y genes reporteros de OFP: El vector pCAMBIA 1305.2 se utilizó como un vector binario primario. A continuación se encuentra el procedimiento de construcción del vector que porta el GmSAMT1 sobreexpresado (FIG. 6). Como herramienta de análisis, se introdujo un gen de proteína anaranjada fluorescente (OFP), *ppor RFP* con cebadores que contenían el
 25 fragmento NcoI, en el sitio NcoI del vector pCAMBIA 1305.2. Este vector pCAMBIA 1305.2 modificado se llamó pJL-OFP y se utilizó para probar la expresión de OFP a través de la transformación transitoria del tabaco. A continuación, el casete del finalizador 35S::GUS/NOS se tomó del vector pBI121 mediante digestión con *EcoRI* y *HindIII* y se insertó en los sitios *EcoRI* y *HindIII* del vector pJL-OFP para formar el vector pJL-OFP-
 30 35S::GUS, que se utilizó para probar la expresión de genes de OFP y GUS en un vector

binario. El gen *GmSAMT1* se amplificó a partir del vector pEXP5/CT-TOPO® mencionado anteriormente con cebadores que contenían el fragmento *BamHI* y *SacI* y se reemplazó con el gen GUS en el vector pJL-OFP-35S::GUS mediante digestión y ligación. Finalmente, se construyó el vector *GmSAMT1* sobreexpresado y se llamó vector pJL-OFP-35S::GmSAMT1. Los vectores de pJL-OFP-35S::GmSAMT1 y pJL-OFP se transformaron en *A. rhizogenes* como gen diana y vector de control, respectivamente. La cepa K599 de *A. tumefaciens* GV3101 y *A. rhizogenes*, se transformó mediante un método de congelamiento y descongelamiento. Las células que no contenían vector binario se colocaron en placas en un medio LB sin antibiótico. Las células con los vectores binarios se colocaron en placas sobre medio LB selectivo con kanamicina y se incubaron a 28°C durante 2 días.

Prueba de la expresión del gen reportero a través de la transformación transitoria del tabaco: Antes de generar las raíces velludas, las construcciones que contenían el gen de OFP se probaron con ensayos de expresión transitoria en plantas de tabaco de conformidad con Sparkes *et al.*, *Nat. Protocols* 1: 2019-2025, 2006. Los plásmidos, pJL-OFP y pJL-OFP-35S::GmSAMT1, se probaron para determinar la función de OFP mediante agroinfiltración de *A. tumefaciens* GV3101 en las hojas de tabaco, junto con *A. tumefaciens* GV3101 sin plásmido binario como control negativo. Las suspensiones bacterianas que portaban los plásmidos diana se infiltraron con presión mediante el uso de una jeringa sin aguja a las hojas de tabaco de 5 semanas a través de las superficies abaxiales de los segmentos de hoja que se encontraban entre las venas mayores que salen hacia la nervadura media de la hoja. Se investigó la fluorescencia de OFP 2 días después de la infiltración en las secciones de hoja extirpada que cubrían la parte de la zona infiltrada del tejido circundante no infiltrado, mediante el uso de un microscopio de fluorescencia Olympus SZX12 y un filtro OFP.

Generación de raíces velludas transgénicas de soja: El sistema de transformación de raíces velludas de soja mediante infección con *A. rhizogenes* que porta los genes diana se modificó y se utilizó para estudiar la resistencia de la soja a SCN. Se utilizó la cepa K599 de *A. rhizogenes* que portaba el gen diana, un vector de control y la cepa K599 de *A. rhizogenes* sola para inocular las semillas de soja TN02-275 que son susceptibles al SCN raza 2. Para evaluar mejor la resistencia del gen diana, también se utilizó *A. rhizogenes* que portaba el vector de control y la cepa K599 de *A. rhizogenes* sola para inocular las semillas de soja TN02-226 que son resistentes al SCN raza 2. La generación de raíces velludas de soja se modificó de la siguiente forma: En primer lugar, las semillas de soja se esterilizaron con gas de cloro. En segundo lugar, una única capa

de semillas de soja maduras con cubierta de semilla intacta se colocó en una placa Petri abierta, que se colocó en un desecador de campana de vidrio dentro de una campana de extracción de vapores. También se colocó un vaso de precipitación con 100 ml de lejía en el desecador y se agregaron 5 ml de HCl concentrado (12N) a lo largo del costado del vaso de precipitación. A continuación las semillas se mantuvieron en un desecador cerrado inmediatamente y se mantuvieron en una atmósfera de gas de cloro durante la noche (aproximadamente 16 horas). Después de la esterilización, las semillas se transfirieron al papel de filtro esterilizado en autoclave, se humedecieron con agua destilada esterilizada en las placas Petri para la germinación en una campana de flujo laminar. Las semillas de soja germinadas se sembraron en vermiculita esterilizada y se cultivaron seis semillas en cada célula de la bandeja de germinación de 18 células. La colonia de bacterias cultivada durante la noche de la cepa K599 de *A. rhizogenes* con el vector binario pJL-OFP, pJL-OFP-35S::GUS o pJL-OFP-35S::GmSAMT1 de la placa de medio LB que contenía 50 mg ml⁻¹ de kanamicina, o un cultivo de la cepa K599 de *A. rhizogenes* que carecía del vector binario cultivado en el medio LB sin antibióticos se recolectó y suspendió en 1 ml de agua destilada estéril, respectivamente. Los nudos cotiledóneos de soja y los hipocotilos superiores de las plántulas de 5 días de edad con cotiledones desplegados se penetraron tres veces con una jeringa que contenía la suspensión bacteriana. En la primera semana, las bandejas se cubrieron con tapas transparentes estériles para mantener la humedad alta. En las tres semanas siguientes, las plantas se transfirieron a macetas y los sitios con heridas por *A. rhizogenes* se cubrieron con vermiculita estéril húmeda para mantener de forma continua la humedad alta. Todos los días las plantas de soja se regaron con una solución B&D estéril que contenía CaCl₂ 1 mM, KH₂PO₄ 0.5 mM, citrato de Fe 10 µM, MgSO₄ 0.25 mM, K₂SO₄ 0.25 mM, MnSO₄ 1 µM, H₃BO₄ 2 µM, 0.5 µM ZnSO₄, CuSO₄ 2 µM, CoSO₄ 0.1 µM, Na₂MoO₄ 0.1 µM, y KNO₃ 1 mM. Las condiciones de cultivo para las plantas de soja fueron 12 horas de luz/12 horas de oscuridad a 28 °C /25 °C con irradiancia de bulbos fluorescentes a 150 a 200 µmol m⁻²·s⁻¹. Después de aproximadamente cuatro semanas, las raíces velludas crecieron a aproximadamente 10cm de longitud. Las raíces transgénicas de soja se detectaron con base en la expresión de GFP, mediante el uso de un estereomicroscopio de fluorescencia (microscopio de fluorescencia Olympus SZX12 y filtro GFP). La raíz principal y las raíces no transgénicas se extirparon debajo del sitio de la herida. Las raíces transgénicas que portaban el pJL-OFP-35S::GUS se analizaron mediante señal de GFP, se sumergieron en solución GUS (1 mM X-Gluc), ferrocianuro de potasio 0.5 mM, TRITON™ X-100 al 0.1% (v/v), tampón fosfato de sodio 100 mM, pH 7.0) y se incubaron

a 37 °C durante la noche. Al día siguiente, la solución de X-Gluc se desechó. Las raíces se lavaron repetidamente con etanol al 70% para eliminar la clorofila. Se registraron las raíces transgénicas con tinte GUS visual.

Infección por nematodos y ensayo demográfico: Para comparar fácilmente la infección por SCN, se dejó una raíz transgénica para la inoculación de SCN. Se recolectaron otras raíces transgénicas para la extracción de ARN y la detección de la expresión de genes. Las partes de la raíz velluda de veinte plantas de soja que contenían la misma construcción se cargaron de forma horizontal en una bandeja de inoculación esterilizada de 13x 9 x 2 cm que contenía una capa fina de mezcla de arena estéril y mantillo (1:1) y las partes con brotes se mantuvieron fuera de la bandeja de inoculación. Se agregaron aproximadamente 0.75 ml de inóculo que contenía aproximadamente 5600 huevos de SCN a cada sistema radicular. Los nematodos se dejaron infectar las raíces durante 7 días en condiciones húmedas. A continuación, todas las raíces se lavaron para eliminar los huevos de SCN extra y los nematodos juveniles que no habían penetrado el tejido de la raíz. Se utilizó la vermiculita estéril dentro de los recipientes para cultivar las quimeras infectadas en la cámara de cultivo. A las dos semanas de la inoculación, las muestras de raíces infectadas se lavaron con agua corriente para eliminar la vermiculita y se aclararon con lejía al 20% (v/v) durante no más de 4 minutos y después se tiñeron con fucsina ácida. Se registraron los nematodos de diferentes etapas dentro de cada muestra de raíz infectada. Dado que las respuestas entre las líneas de soja susceptibles y resistentes ocurren en la etapa J3 de SCN, la relación de la cantidad de J3+J4 con respecto a la cantidad total de nematodos por planta se utilizó como un índice para determinar la diferencia de resistencia entre las raíces velludas que contienen las construcciones respectivas. Se analizó un mínimo de 10 plantas transformadas para cada vector.

Reacción en cadena de la polimerasa cuantitativa con transcripción inversa (qRT-PCR): La PCR cuantitativa con transcripción inversa se llevó a cabo para verificar el nivel de transcripción de *GmSAMT1* y de algunos genes relacionados con la defensa de las raíces velludas transgénicas de soja de conformidad con Mazarei *et al.*, 2007. Para llevar a cabo los ensayos de qRT-PCR, se diseñaron cebadores específicos de genes de *GmSAMT1*, genes relacionados con la defensa, *GmPR-1* (banco de genes, acceso BU577813), *GmPR-3* (banco de genes, acceso AF202731) y el gen de referencia, gen ubiquitina3 de soja (*GmUBI-3*, banco de genes, acceso D28123) (**Tabla 1**).

Tabla 1. Cebadores para el análisis de qRT-PCR del gen SA metiltransferasa de soja y genes relevantes relacionados con la patogénesis.

Nombre del gen	Orientación	Secuencia (5' a 3')
<i>GmSAMT1</i>	Directa	GTGAAATTGGAAGTTCTTAAAGAAGGA (SEQ ID NO: 7)
	Inversa	CAGATTCAAAGTCTAGAGCATTCCAC (SEQ ID NO: 8)
<i>GmUBI-3</i>	Directa	GTGTAATGTTGGATGTGTTCCC (SEQ ID NO: 9)
	Inversa	ACACAATTGAGTTCAACACAAACCG (SEQ ID NO: 10)
<i>GmPR-1</i>	Directa	AACTATGCTCCCCCTGGCAACTATATTG (SEQ ID NO: 11)
	Inversa	TCTGAAGTGGTAGCTTCTACATCGAAACAA (SEQ ID NO: 12)
<i>GmPR-3</i>	Directa	AACTACAATTACGGGCAAGCTGGCAA (SEQ ID NO: 13)
	Inversa	TTGATGGCTTGTTTCCCTGTGCAGT (SEQ ID NO: 14)
<i>GmNPR1-1</i>	Directa	AATTGACCAAGAGCTTCCGC (SEQ ID NO: 26)
	Inversa	CTACAGAAGCATCATTTTCAACATCTT (SEQ ID NO: 27)
<i>GmNPR1-2</i>	Directa	CAATTGACCAAGAGCTTCCAA (SEQ ID NO: 28)
	Inversa	CTATAGAAGCATCGTTTTCAACATCTC (SEQ ID NO: 29)
<i>GmICS1</i>	Directa	GGCCATTTTCGGAGCTGG (SEQ ID NO: 30)
	Inversa	AGGAGAAGGTGGTTCTGTGAGAGA (SEQ ID NO: 31)
<i>GmICS2</i>	Directa	CATGGCCATTTTCGGAGCTTA (SEQ ID NO: 32)
	Inversa	CAGAAGATGGTTCTGCGAATGG (SEQ ID NO: 33)

Se aisló la totalidad del ARN de los tejidos de raíces respectivos de las raíces velludas transgénicas y de las raíces velludas de control mediante el uso de columnas RNEASY® (QIAGEN®) siguiendo las instrucciones del fabricante para plantas. El ARN se trató con el conjunto DNasa libre de RNasa (QIAGEN®) para eliminar el ADN genómico y

el ARN total (aproximadamente 1.5 µg) se transcribió de forma inversa para sintetizar el ADNc mediante el uso del kit de transcripción inversa de ADNc de alta capacidad (APPLIED BIOSYSTEMS®) para la qRT-PCR de conformidad con las instrucciones del fabricante. La primera cadena de ADNc se diluyó y colocó en cada reacción de qRT -
 5 PCR. La acumulación de ADN se midió mediante el uso de color verde SYBR® como el tinte de referencia. La contaminación con ADN genómico se controló mediante PCR con el uso de una muestra de ARN como plantilla. Solamente un producto se encontró presente en cada reacción según se indicó por la curva de disociación del tinte de referencia de los productos amplificados. Se utilizaron las siguientes condiciones de PCR:
 10 50 °C durante 2 minutos; 95 °C durante 10 minutos; seguidos por 40 ciclos de 95 °C durante 15 segundos; 60 °C durante 1 minuto; 72 °C durante 30 segundos. Todos los ensayos de qRT-PCR se realizaron por triplicado. Las eficiencias de PCR para los genes diana y de referencia fueron iguales entre las muestras diana y de control. Los valores de Ct y abundancia relativa se calcularon mediante el uso del software suministrado con el
 15 sistema de PCR en tiempo real 7900 HT Fast de APPLIED BIOSYSTEMS®.

Análisis estadístico: La diferencia de resistencia entre las raíces velludas transgénicas de soja con sobreexpresión de *GmSAMT1* y las raíces velludas de soja de control se analizó con ANOVA de una vía mediante el uso del modelo mixto de SAS (SAS versión 9.2). Los promedios de mínimos cuadrados del índice de resistencia se separaron
 20 mediante la prueba de diferencia mínima significativa (LSD) de Fisher. La diferencia de expresión del gen diana y de los genes relacionados con la defensa analizada mediante qRT-PCR también se analizó con ANOVA de una vía mediante el uso del modelo mixto de SAS. Los promedios de mínimos cuadrados del nivel de expresión relativa se separaron mediante el uso de la prueba de LSD de Fisher. El $P < 0.05$ se consideró significativo.

25 RESULTADOS

Glyma02g06070 es un miembro de la familia SABATH: El análisis de microensayo sobre las líneas de soja recombinantes híbridas (RIL), TN02-226 y TN02-275, reveló algunos genes con cambios de expresión significativos durante las reacciones resistentes y susceptibles al SCN. Uno de estos genes mostró regulación positiva
 30 únicamente en las raíces de soja resistentes inoculadas con SCN mientras que no ocurrió ningún cambio significativo de su expresión de genes en las raíces de soja susceptibles inoculadas con SCN (Mazarei *et al.*, *Theor. Appl. Genet.* 123: 1193-1206, 2011). Para obtener el marco de lectura abierto de este gen candidato resistente, la secuencia de la sonda Affymetrix SoyChip, Gma.12911.1.A1_S_AT, que identificó nuestro gen diana con
 35 microensayos, se utilizó como una consulta a la búsqueda de BLAST contra la base de

datos de soja GmGDB/GMtranscript (disponible en el sitio web plantgdb.org/GmGDB/cgi-bin/blastGDB.pl)y en la base de soja (disponible en el sitio web soybase.org/gbrowse/cgi-bin/gbrowse/gmax1.01/), se encontró que el gen de soja, *Glyma02g06070*, tiene la mayor similitud (99%) con respecto a la secuencia de Gma.12911.1.A1_S_AT y se anotó para
5 codificar la carboximetiltransferasa dependiente de SAM (**FIG. 2**).

Mediante comparación adicional con las metiltransferasas conocidas de otras especies, se determinó que esta carboximetiltransferasa dependiente de SAM codificada por *Glyma02g06070* es un miembro de la familia SABATH. La familia SABATH se nombra después de las primeras enzimas identificadas (SAMT, BAMT y teobromo sintasa) en esta
10 familia (Chen et al. 2003a). Para investigar las relaciones evolutivas entre *Glyma02g06070* y los miembros conocidos de SABATH de otras especies, se construyó un árbol filogenético de miembros seleccionados de SABATH caracterizados funcionalmente con *Glyma02g06070*. El resultado mostró que *Glyma02g06070* pertenecía al conjunto que contenía las enzimas con actividad de metiltransferasa de ácido salicílico
15 (FIG. 3).

***Glyma02g06070* codifica SAMT de soja:** Una búsqueda de proteínas con BLAST reveló que *Glyma02g06070* contiene 370 aminoácidos y es un 4.01% idéntico a la SAMT del arroz identificada anteriormente, OsBSMT1, a nivel de secuencia de aminoácidos. También se predijo que *Glyma02g06070* contiene el dominio de unión a SAM. Cuando la
20 secuencia de *Glyma02g06070* se buscó en la base de datos EST, también mostró que este gen es inducido por SCN y SA. Para confirmar la predicción de las bases de datos, amplificamos y secuenciamos el marco de lectura abierto (ORF) completo de este gen a partir del ADNc obtenido del tejido de la raíz infectada con SCN de la línea de soja TN02-226.

25 Dado que *Glyma02g06070* codifica un miembro de la familia SABATH, cuyas enzimas miembro pueden transferir un grupo metilo a partir de SAM donante de metilo a una variedad de sustratos, la actividad enzimática de la proteína recombinante de *Glyma02g06070* se examinó mediante el uso de varios sustratos posibles (Tabla 2).

30 Tabla 2. Actividad relativa de GmSAMT1 con ácido salicílico y otros sustratos relacionados

Sustrato	Actividad relativa ^a (%)
Ácido salicílico	100
Ácido benzoico	16.9
Ácido antranílico	9.3
Ácido 3-hidroxibenzoico	1.8

Ácido 4-hidroxibenzoico	0
Ácido nicotínico	0
Ácido p-cumárico	0
Ácido cafeico	0
Ácido cinámico	0
Ácido vanílico	0
Ácido indol-3-ácetico	0
Ácido 2,4-diclorofenoxiacético	0
Ácido jasmónico	0
Ácido giberélico	0

a: Los valores son promedios de tres mediciones independientes. Todos los sustratos se probaron a una concentración de 1 mM. La actividad del GmSAMT1 recombinante con ácido salicílico se ajustó arbitrariamente a un 100%.

Se encontró que los aumentos en la velocidad de reacción mediante el aumento de las concentraciones de SAM y SA obedecen a la cinética de Michaelis-Menten. Entre los 14 sustratos examinados, SA, ácido benzoico, ácido antranílico, y ácido 3-hidroxibenzoico sirvieron como sustratos *in vitro* en las reacciones catalizadas por la proteína recombinante codificada por *Glyma02g06070*, a pesar de que mostraron diferencias obvias en las actividades. En las concentraciones de sustrato de 1 mM, la proteína recombinante mostró la actividad más alta con SA (100%) en lugar de ácido benzoico (16.9%), ácido antranílico (9.3%) y ácido 3-hidroxibenzoico (1.8%). Para otros sustratos que incluyeron JA, ácido indol-3-ácetico y ácido giberélico, no se detectó actividad en las mismas condiciones de reacción. *Glyma02g06070* catalizó la metilación de SA con un valor de Km aparente de 32.9±1.7 µM. Los valores finales representan el promedio de tres mediciones independientes. Dado que SA sirve como su sustrato suficiente, el *Glyma02g06070* se denominó *GmSAMT1*.

Propiedades bioquímicas de SAMT de soja: El GmSAMT1 recombinante con marcador His expresado en *E. coli* se purificó para lograr una homogeneidad electroforética y se sometió a caracterización bioquímica detallada. La masa molecular de GmSAMT1 con proteína marcadora His estimada sobre gel fue de 43 kDa, lo que fue cercano al tamaño esperado (FIG. 4). El GmSAMT1 recombinante purificado mostró un valor de Km aparente de 32.9±1.7 µM para el ácido salicílico. Cuando los ensayos enzimáticos se realizaron en tampones de diferentes valores de pH, el GmSAMT1 mostró el nivel más alto de actividad catalítica a pH 7.5. (FIG. 5A). La temperatura óptima para la actividad de GmSAMT1 fue de 25 °C (FIG. 5B). Se midieron los efectos de varios iones

metálicos sobre la actividad de GmSAMT1. En comparación con el control, el GmSAMT1 mostró actividad doble en la presencia de K^+ , mientras que no mostró actividad alguna en la presencia de Cu^{2+} , Zn^{2+} , Fe^{2+} y Fe^{3+} . La actividad de GmSAMT1 se inhibió de forma moderada con Mn^{2+} y Ca^{2+} . Otros iones metálicos, incluidos Na^+ , NH_4^+ y Mg^{2+} , tuvieron un efecto mínimo sobre la actividad de GmSAMT1 (FIG. 5C). El GmSAMT1 mostró una cinética pseudo Michaelis-Menton, con un valor de K_m aparente de $46.2 \pm 4.2 \mu M$ para el ácido salicílico (FIG. 5D). Todos los resultados anteriores sugirieron que el GmSAMT1 recombinante mostró actividad de metiltransferasa de ácido salicílico. En conjunto, la evidencia filogenética y bioquímica sugirió que GmSAMT1 es una enzima conservada entre varias especies de plantas.

Prueba de la expresión del gen reportero a través de la transformación transitoria del tabaco: Antes de llevar a cabo la transformación mediada por *A. rhizogenes*, se llevó a cabo la expresión transitoria de pJL-OFP-35S::GUS en la hoja de tabaco a través de la transformación mediada por *A. tumefaciens* para verificar la señal de OFP con un filtro OFP. Las hojas de tabaco transgénicas que portaban el gen diana mostraron fluorescencia de OFP, mientras que la planta simulada no mostró señal de OFP (FIG. 7A). Este resultado ayudó a confirmar rápidamente la expresión de OFP.

Después de la generación mediada por *A. rhizogenes* de las raíces velludas de soja transformadas con genes diana de 10 cm de longitud, las raíces velludas transgénicas que portaban el *GmSAMT1* sobreexpresado se analizaron con filtro OFP. Este método se demostró mediante la verificación de la expresión de GUS mediante tinción de las raíces velludas transgénicas que portaban el plásmido pJL-OFP-35S::GUS. Todas las raíces velludas transgénicas positivas analizadas mediante señal de OFP mostraron expresión de GUS uniforme en el tejido de la raíz, mientras que las raíces velludas de control negativo analizadas mediante señal de OFP no mostraron expresión de GUS (FIG. 7B).

Expresión de genes de *GmSAMT1* y varios genes relacionados con la patogénesis (PR): El nivel de expresión génica de *GmSAMT1* y los genes PR seleccionados en las raíces velludas transgénicas se examinó mediante qRT-PCR (FIG. 8 y FIG. 9). El nivel de expresión de *GmSAMT1* en las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* fue aproximadamente 25 veces más alto que el de las plantas de control (FIG. 8B). *GmPR-1* mostró una expresión significativamente más alta en las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* en comparación con algunas líneas de raíces velludas de control. Pero no hubo diferencia significativa de la expresión de genes de *GmPR-1* entre las raíces velludas transgénicas con

sobreexpresión de *GmSAMT1* y el vector de control en el ámbito de la soja susceptible (FIG. 9A). Esto puede sugerir que la soja susceptible puede responder a proteínas exógenas cuando no ocurre infección por SCN. Cuando ocurre el parasitismo por SCN, el sistema de defensa de la soja susceptible puede cambiarse o inhibirse por los SCN, que producen y secretan proteínas parasíticas en la soja para ayudarse a sí mismos en la supresión de la defensa de la planta hospedadora. Se encontró que la corismato mutasa es una de las proteínas parasíticas de SCN que puede manipular la ruta del sikimato de la planta. Dado que la biosíntesis de SA se encuentra estrechamente asociada con la ruta del sikimato, la vía de señalización de SA de la soja susceptible se suprimiría por SCN. Se encontró que la supresión local de la señalización de SA en las raíces de *Arabidopsis* era necesaria para el parasitismo del nematodo del quiste de la remolacha. En el futuro, será necesario examinar la expresión de genes de PR para determinar las raíces velludas infectadas con SCN. Dado que el gen *PR-1* se consideró un gen marcador de la vía de señalización de SA, la sobreexpresión de *GmSAMT1* puede aumentar la transducción de señal de la vía de SA mediante el suministro de la molécula móvil abundante MeSA para responder a la infección por SCN. De forma interesante, se observó que la expresión de *PR-3*, que codifica la quitinasa de soja, es superior en las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* que en otras líneas de control (FIG. 9B). Se encontró que JA induce la expresión de *PR-3* en *Arabidopsis*.

El SCN es la plaga de la soja que produce el mayor daño a nivel económico en Estados Unidos, por lo que se ha estudiado el mecanismo de interacción entre la soja y SCN para controlar mejor esta plaga. Con el desarrollo genómico, se han llevado a cabo varios análisis de microensayo para estudiar los cambios de expresión de genes de la soja cuando se encuentra infectada por SCN. No obstante, los varios experimentos no son congruentes debido a los diferentes tipos de muestras de soja y SCN. En el estudio actual, se identificó el gen candidato con base en el análisis de microensayo de dos líneas hermanas de soja relacionadas genéticamente TN02-226 y TN02-275, que son resistentes y susceptibles, respectivamente, a la infección por SCN raza 2.

Durante la inoculación de SCN, la actividad de *GmSAMT1* también puede manipularse mediante el cambio de la concentración del potasio de las células de soja. El síntoma más común por encima del suelo de la infección por SCN es un síntoma de deficiencia de potasio con amarillecimiento de los márgenes de las hojas. No obstante, dichos síntomas de deficiencia de potasio inducida por SCN son el resultado de la incapacidad de la planta de obtener potasio suficiente debido a la alimentación del nematodo. Se demostró que la concentración del potasio en las raíces de soja cultivadas

en un nivel medio de condiciones de fertilidad de potasio disminuyó debido al SCN (Smith et al. 2001). El resultado indicó que el *GmSAMT1* *in vitro* mostró el doble de actividad en la presencia de potasio, lo que sugiere que el SCN también puede inhibir la actividad de *GmSAMT1* mediante la disminución de la concentración de potasio además de la supresión del nivel de transcripción de *GmSAMT1*.

El *GmSAMT1* puede jugar un papel en la interacción entre las vías de señalización de SA y JA en la resistencia de soja contra SCN. Se demostró que la expresión de genes de SAMT del arroz puede inducirse por JA. También se detectó que MeSA era altamente elevado después que *Arabidopsis* de tipo salvaje se trató con MeJA. Como se divulgó en la presente, la sobreexpresión de *GmSAMT1* en la línea celular de soja susceptible provocó la regulación positiva de *GmPR-1* y *GmPR-3*, que tiene el efecto integrante de las vías de señalización de SA y JA. Los análisis de microensayo también indican que la vía de JA se activó en la respuesta de defensa de soja contra SCN. Por consiguiente, la vía dependiente de SA, combinada con la vía de JA a través de los puntos reguladores, incluidos NPR1, factor de transcripción de la familia WRKY y LOX, juegan un papel de represión del desarrollo de SCN. Los resultados presentados en la presente indicaron que el *GmSAMT1* es otro punto de regulación en la interacción entre las vías de SA y JA.

Resistencia a SCN: La resistencia de la soja a SCN se evaluó mediante el análisis del desarrollo de SCN, es decir, nematodos, quistes o huevos por planta. En este estudio, este método se modificó de conformidad con Melito *et al.*, *BMC Plant Biology* 10:104, 2010. La relación de la cantidad de nematodos en la etapa J3 y J4 con respecto a la cantidad total de SCN se ajustó como el índice, $(J3+J4)/SCN$ totales, para evaluar la resistencia entre las diferentes raíces velludas dos semanas después de la inoculación con huevos de SCN.

Se observaron los nematodos hembra en las etapas J2 y J4 en la mayoría de las secciones de raíz velluda transgénica y los nematodos macho se observaron de forma ocasional (FIG. 11). La cantidad de SCN varió ampliamente entre las plantas individuales transgénicas y de control. La cantidad promedio de SCN establecidos no fue significativamente diferente entre la línea de raíz velluda transgénica y los controles (datos no mostrados). No obstante, la relación $(J3+J4)/SCN$ totales, mostró una diferencia significativa entre las raíces velludas transgénicas y la línea de control susceptible. Las relaciones de $(J3+J4)/SCN$ totales para las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* (3.4% para TN02-275, 5.3% para Williams 82) fueron comparables con el vector de control resistente (2.5% para TN02-226) y fueron significativamente diferentes con el vector de control susceptible (23.6% para TN02-275,

18.4% para Williams 82). Las raíces velludas de soja transgénicas con sobreexpresión de *GmSAMT1* presentaron una cantidad significativamente menor de SCN juveniles que se desarrollaron en las etapas J3 y J4 en comparación con el vector de control en el ámbito de soja susceptible. Cuando el nivel de susceptibilidad del vector de control en el ámbito susceptible se ajustó de forma arbitraria a un 100%, las raíces velludas transgénicas de soja con sobreexpresión de *GmSAMT1* presentaron una reducción significativa de la susceptibilidad (un 85.6% de reducción para TN02-275, un 71.2% de reducción para Williams 82).

Es de mayor interés el hecho de que las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* muestran una resistencia superior a SCN que las raíces velludas transgénicas de control (FIG. 11). La resistencia mediada por *GmSAMT1* no es específica para la línea, dado que dos líneas susceptibles mostraron el mismo fenotipo (FIG. 11). El papel biológico de *GmSAMT1* en la resistencia a SCN puede evaluarse de forma adicional mediante el uso de transformación estable.

Expresión de los genes relacionados con la vía de SA en las raíces velludas transgénicas: Se evaluó la abundancia de transcripción para un gen que codifica la isocorismato sintasa (ICS) mediante la biosíntesis de proteínas de SA, en raíces velludas transgénicas de SAMT (FIG. 12). En *Arabidopsis*, fue necesario el *AtICS1* para la acumulación de SA inducida por patógenos. A pesar de que se propuso que el gen *ICS* era un gen crítico para la biosíntesis de SA, se sabe muy poco acerca del papel del *ICS* en la soja. También se examinó la expresión de *GmICS1* (*Glyma01g25690*) (FIG. 12A) y *GmICS2* (*Glyma03g17420*) (FIG. 12B) mediante qRT-PCR. Sin tratamiento con SCN, tanto el *GmICS1* como el *GmICS2* en las raíces velludas transgénicas con sobreexpresión de *GmSAMT1* mostraron una abundancia de transcripción superior (5.3 veces y 8.2 veces) a la de los controles. En el tratamiento con el estímulo de SCN no hubo diferencia significativa en la expresión de *GmICS1* entre la raíz velluda transgénica y la línea de control. Para *GmICS2*, las raíces velludas transgénicas presentaron una abundancia de transcripción levemente mayor a la de los controles.

Para explorar de forma adicional la vía de transducción de señal dependiente de SA, se probó el gen regulador clave de la vía de señalización de SA, dos homólogos del no expresor de los genes *PR1* (*NPR1*), *GmNPR1-1* (*Glyma09g02430*) y *GmNPR1-2* (*Glyma15g13320*) en soja mediante qRT-PCR (FIG. 13). Para *GmNPR1-1*, no hubo diferencia significativa en el nivel de expresión entre las raíces velludas transgénicas y las líneas de control sin tratamiento con SCN. No obstante, con el tratamiento con SCN, el *GmNPR1-1* mostró una expresión superior a la del control (FIG. 13A). Para *GmNPR1-2*,

no hubo diferencia significativa en el nivel de expresión entre las raíces velludas transgénicas y la línea de control con tratamiento con SCN. No obstante, sin tratamiento con SCN, el *GmNPR1-2* mostró una expresión superior a la de los controles. En conjunto, el *GmNPR1-2* mostró una expresión relativamente superior en la raíz velluda transgénica en respuesta al tratamiento con SCN (FIG. 13B).

Papeles biológicos de *GmSAMT1* en la defensa de la soja contra el virus del mosaico de la soja: Para probar si el *GmSAMT1* también se encontraba implicado en la resistencia contra SMV, también se llevó a cabo una qRT-PCR al medir la expresión de *GmSAMT1* and *GmPR-1* durante la interacción entre SMV y soja. En este ensayo, se utilizaron dos líneas de soja L78-379 (*Rsv1*) y Williams 82 (*rsv1*). El genotipo de soja L78-379 es una isolínea de Williams, pero contiene el gen de resistencia a *Rsv1* de PI 96983 (*Rsv1*). La inoculación de L78-379 (*Rsv1*) con un mutante derivado de SMV-N (N25) resulta en la inducción de una respuesta hipersensible retrasada, mientras que la infección de Williams 82 (*rsv1*) con el mismo virus provoca el mosaico.

Cuando la respuesta resistente ocurrió para L78 contra el virus de la soja 25 (21 días después de la inoculación, la expresión de *GmSAMT1* de las hojas superiores de soja fue aproximadamente 7.3 veces la de la planta simulada y, de forma interesante, también se observó un alto nivel de expresión (44.6 veces) de *GmPR-1*. Mientras que en la respuesta susceptible de la soja Williams 82, la expresión de *GmSAMT1* y *GmPR-1* se redujo a la mitad de la de la planta simulada (FIG. 14). Por lo que la correlación fuertemente positiva entre la expresión de *GmSAMT1* y *GmPR-1* indica que *GmSAMT1* juega un papel en la defensa de la soja contra el virus del mosaico de la soja.

En conjunto, las diferencias en la expresión de genes de *GmSAMT1* y los genes de *PR* seleccionados se regularon positivamente en los dos tejidos inoculados de las raíces y las hojas de la línea de soja resistente pero se regularon negativamente o no se cambiaron en la línea de soja susceptible, lo que demuestra las respuestas de la soja contra los patógenos.

Ejemplo 2

Aislamiento y caracterización de *GmSABP2-1*

Fuentes vegetales, nematódicas, bacterianas y químicas: En el presente estudio, se utilizaron dos líneas de soja genéticamente relacionadas, TN02-226 y TN02-275, para la clonación de genes y la generación de raíces velludas de soja. Se han descrito los detalles de reproducción para estas dos líneas de soja. Se utilizaron huevos de SCN raza 2 como los patógenos en el bioensayo de las raíces velludas, con el mismo método de inoculación descrito anteriormente. Se utilizó la cepa BL21 (DE3) de *Escherichia coli*

CODONPLUS® para la expresión de la proteína recombinante. Se utilizó la cepa K599 de *A. rhizogenes* para portar el *GmSABP2-1* sobreexpresado y generar raíces velludas de soja. Los químicos utilizados en este estudio se obtuvieron en Sigma-Aldrich.

Búsqueda en la base de datos y análisis de secuencias: Ambas secuencias de las sondas de Affymetrix GmaAffx.92649.1.S1_at y Gma.5867.1.A1_s_at mostraron un 100% de similitud con un gen de soja, *Glyma16g26060*. Con base en la anotación en fitozoma (disponible en el sitio web phytozome.net/soybean), este gen codifica una α/β hidrolasa. La alineación de Glyma16g26060 con otras dos metil éster estearasas caracterizadas se llevó a cabo para estudiar los dominios conservados de esta proteína candidata.

Aislamiento del gen Glyma16g26060 de soja: La extracción total del ARN y la eliminación de la contaminación de ADN se realizó en la forma descrita. Para estudiar de forma adicional la función de Glyma16g26060, la secuencia de ADNc de longitud completa Glyma16g26060 se obtuvo de la base de datos de soja pública (disponible en el sitio web phytozome.net/soybean). Con el uso de los cebadores específicos ATGGGTTCACAAAATTGTATGGATAGG (SEQ ID NO: 15) y TCATGCATATTTAGTCGCTATCTGCTG (SEQ ID NO: 16), las condiciones de ciclo térmico fueron las siguientes: 94 °C durante 2 minutos seguidos por 35 ciclos a 94 °C durante 30 segundos, 57 °C durante 30 segundos y 72 °C durante 1 minuto y una extensión final a 72 °C durante 10 minutos. Los productos de la PCR se clonaron en el vector pEXP5/CT-TOPO®. El resultado del secuenciamiento del producto de la PCR de la soja TN02-226 fue coherente con el predicho en la base de datos de la soja Williams 82.

Para verificar su función bioquímica, el gen candidato a continuación se clonó en un vector de purificación de proteínas, vector pET100/D-TOPO® y se expresó en la cepa BL21 de *E. coli* (DE3) CODONPLUS®. La expresión de proteínas se indujo con isopropil β -D-1-tiogalactopiranosida (IPTG) a una concentración de 500 μ M durante 18 horas a 25°C, donde las células se lisaron mediante sonicación. El Glyma16g26060 expresado por *E. coli* con un marcador His se purificó a partir del lisado celular de *E. coli* mediante el uso de Ni-NTA agarosa siguiendo las instrucciones del fabricante (Invitrogen). La pureza de las proteínas se verificó mediante SDS-PAGE y las concentraciones de proteínas se determinaron con el ensayo de Bradford.

Ensayo de la actividad de metil éster estearasa: Se realizó un ensayo radioquímico de dos etapas con estearasa para determinar la actividad de Glyma16g26060 expresado por *E. coli* siguiendo el protocolo informado anteriormente por (Forouhar *et al.*, *Proc. Natl. Acad. Sci. USA* 102: 1773-1778, 2005). La reacción de la primera etapa fue la reacción del sustrato, 10 μ M de MeSA y Glyma16g26060 recombinante purificado expresado por *E.*

coli durante 30 minutos a 25°C. Después de ello, la muestra se hirvió para detener la reacción y desnaturalizar la enzima. La segunda etapa de reacción comenzó con la adición de 3 µM de S-adenosil-L-metionina (SAM) marcada con ¹⁴C con una actividad específica de 51.4 mCi/mmol (Perkin Elmer, Boston, MA, E.U.A.) y la respectiva metiltransferasa purificada conocida con actividad alta. Aquí, el GmSAMT1 mencionado anteriormente se utilizó para la enzima metiltransferasa de SA. La reacción estuvo precedida por 30 minutos a 25 °C. Se extrajeron los productos radiomarcados y se contó la radioactividad mediante el uso de un contador de centelleo líquido.

Determinación de los parámetros cinéticos de SABP2 de soja: Se evaluó el aumento en la velocidad de reacción con concentraciones crecientes de MeSA con el ensayo radioquímico descrito anteriormente y se encontró que obedecía a la cinética de Michaelis–Mention. Las concentraciones de enzimas y el tiempo de incubación adecuados se determinaron en ensayos temporales para asegurarse de que la velocidad de reacción fuera lineal durante el período de ensayo. Para determinar el valor de Km para MeSA, las concentraciones de MeSA variaron de forma independiente en el intervalo de 5 µM a 100 µM. Se realizaron las gráficas de Lineweaver–Burk para obtener un valor de Km aparente. Se analizó el pH óptimo para GmSABP2-1 de pH 6.0 a pH 9.0. Los valores finales son un promedio de tres mediciones independientes.

Evaluación de la resistencia a SCN de las raíces velludas transgénicas con sobreexpresión de GmSABP2-1: Las raíces velludas de soja transgénicas con sobreexpresión de *GmSABP2-1* se utilizaron para evaluar la función biológica de *GmSABP2-1*. Las etapas que incluían la creación de la construcción con sobreexpresión de *GmSABP2-1* y el gen reportero de GFP, la generación de las raíces velludas transgénicas de soja que portaban el *GmSABP2-1* sobreexpresado, la inoculación de SCN y el ensayo demográfico de los nematodos, se llevaron a cabo en la forma descrita, pero se utilizó el gen *GmSABP2-1* en lugar del gen *GmSAMT1*.

Análisis estadístico: La diferencia de resistencia entre las raíces velludas de soja con sobreexpresión de *GmSABP2-1* y las raíces velludas de soja de control se analizó con ANOVA de una vía mediante el uso de un modelo mixto de SAS (SAS versión 9.2) y los promedios de mínimos cuadrados del índice de resistencia se separaron mediante la prueba de LSD de Fisher. El P<0.05 se consideró significativo.

RESULTADOS

Identificación y análisis de secuencia de un gen putativo resistente a SCN en soja: El análisis de microensayo reveló genes que muestran cambios de expresión significativos durante las reacciones resistentes y susceptibles a SCN. Uno de estos

genes se reguló positivamente de forma significativa únicamente en las raíces de soja resistentes a las que se les inoculó SCN, mientras que se reguló negativamente de forma significativa en las raíces de soja susceptibles a las que se les inoculó SCN mediante análisis de microensayo y PCR cuantitativa en tiempo real.

5 Para obtener los marcos de lectura abiertos completos de este gen relacionado con la resistencia, las secuencias de la sonda de Affymetrix SoyChip, GmaAffx.92649.1.S1_at y Gma.5867.1.A1_s_at que identificaron el gen diana en análisis de microensayo, se utilizaron como una consulta a la búsqueda de BLAST contra la base de datos de soja de GmGDB/GMtranscript (disponible en el sitio web
10 plantgdb.org/GmGDB/cgi-bin/blastGDB.pl), se encontró que el gen *Glyma16g26060* mostró la mayor similitud (100%) con respecto a la secuencia de las dos sondas anteriores y se predijo que pertenecía a la superfamilia de genes α/β hidrolasa.

***Glyma16g26060* codifica la metil éster estearasa de soja:** Una búsqueda de BLAST reveló que *Glyma16g26060* es un 61.2% idéntico con respecto al NtSABP2
15 (Banco de Genes, acceso AAR87711.1) caracterizado anteriormente a nivel de secuencias de aminoácidos y su identidad con respecto a NtSABP2 es la mayor entre los genes homólogos de soja. Se realizó la alineación de las secuencias peptídicas entre *Glyma16g26060* y NtSABP2, At4g37150 (FIG. 15). *Glyma16g26060* contiene la tríada catalítica (Ser-86, His-239 y Asp-211) que es altamente conservada entre los ortólogos
20 SABP2 del tabaco, *Arabidopsis* (AtMES), álamo y papa (StMES1) así como los miembros estrechamente relacionados de la superfamilia de α/β hidrolasa. *Glyma16g26060* también contiene la totalidad de los 15 aminoácidos identificados en NtSABP2 que interactúa con SA. En contraste, AtMES9 contiene solamente 5 y StMES1 comparte 12 de estos 15 residuos. La alta similitud y conservación secuencial de los residuos de unión a SA
25 críticos entre *Glyma16g26060* y NtSABP2 indicó que estas dos proteínas compartían propiedades bioquímicas similares. Para confirmar la predicción de las bases de datos, el ADNc de longitud completa de *Glyma16g26060* se clonó con la reacción en cadena de la polimerasa de transcripción inversa (RT-PCR) a partir del tejido de la raíz infectado con SCN de la línea de soja TN02-226 y se secuenció. El gen secuenciado contenía 786
30 pares de bases y codificó una proteína de 261 aminoácidos.

Propiedades bioquímicas de SABP2 de soja: Para explorar de forma adicional las propiedades bioquímicas de la enzima codificada por *Glyma16g26060*, este gen se clonó en un vector de expresión de proteínas. El *Glyma16g26060* recombinante con marcador His expresado en *E. coli* se purificó para lograr una homogeneidad
35 electroforética y se sometió a caracterización bioquímica detallada. La masa molecular de

Glyma16g26060 con proteína marcadora His estimada sobre gel fue de 33 kDa (FIG. 16), lo que fue cercano al tamaño esperado de nuestra enzima diana. Glyma16g26060 recombinante purificado mostró la actividad de conversión de MeSA en SA. En condiciones de estado estacionario, Glyma16g26060 hidrolizó MeSA con un valor de Km
 5 aparente de $46.2 \pm 2.2 \mu\text{M}$ (FIG. 17). Por tanto, Glyma16g26060 se diseñó como la soja SABP2 (GmSABP2-1). Cuando los ensayos enzimáticos se realizaron en tampones de diferentes condiciones de pH, el GmSABP2-1 mostró el nivel más alto de actividad catalítica a un pH de aproximadamente 7.0 (FIG. 18).

Papel biológico de soja SABP2 en la resistencia a SCN: El estudio del papel
 10 biológico de SABP2 en otras especies de plantas se realizó mediante la observación de la resistencia de las hojas a la infección por virus o bacterias. Dado que la resistencia de la soja a SCN ocurre en el tejido de la raíz, se registró la cantidad de SCN que existen en las raíces velludas transgénicas en las diferentes etapas para analizar el nivel de resistencia y evaluar la función biológica de GmSABP2-1. Como se mencionó, las diferentes
 15 respuestas resistentes a SCN raza 2 entre las líneas de soja susceptibles (TN02-275) y resistentes (TN02-226) ocurrió en la etapa J3 de SCN, se eligió el punto de tiempo después de dos semanas de la inoculación con SCN para analizar el perfil demográfico de SCN en todas las raíces velludas con las construcciones respectivas. Se utilizó el índice de J3+J4/cantidad total de SCN para estimar la resistencia entre la línea susceptible con
 20 sobreexpresión de *GmSABP2-1* y otras líneas de control, esto con el ensayo descrito. El nivel de susceptibilidad del vector de control en el ámbito susceptible se ajustó arbitrariamente en un 100%. Los resultados mostraron que las raíces velludas transgénicas de soja con sobreexpresión de *GmSABP2-1* tenían una cantidad significativamente menor de SCN hembra juveniles que se desarrollaron en la etapa J3, en comparación con la línea del vector de control en el ámbito de soja susceptible (84.5%
 25 de reducción). Y no hubo diferencia significativa en el nivel de resistencia entre las raíces velludas transgénicas con sobreexpresión de *GmSABP2-1* y el vector de control en el ámbito de soja resistente, que también tenía una cantidad significativamente menor de SCN hembra juveniles que se desarrollaron en la etapa J3, en comparación con el vector
 30 de control negativo en el ámbito de soja susceptible (94.8% de reducción) (FIG. 19).

Se describió la identificación de GmSABP2-1 de soja y se demostró que compartía una alta similitud y unidades conservadas con respecto a sus ortólogos de tabaco y *Arabidopsis*. Mediante ensayo bioquímico adicional, se indicó que el GmSABP2-1 recombinante mostró actividad de estearasa hacia MeSA. El valor de Km de GmSABP2-1
 35 ($46.2 \mu\text{M}$) fue mayor al valor de Km de NtSABP2 ($8.6 \mu\text{M}$) PtSABP2-2 ($24.6 \mu\text{M}$) y menor

que el valor de Km de AtMES1 (57.3 μ M), AtMES9 (147.1 μ M), PtSABP2-1(68.2 μ M) y StMES1 (57.9 μ M). Las diferencias en estos parámetros cinéticos pueden reflejar niveles de MeSA endógeno de especies de plantas.

5 Las raíces velludas transgénicas con sobreexpresión de *GmSABP2-1* en la línea de soja susceptible, mostraron una resistencia superior a SCN. Este resultado indicó que era necesario un cierto nivel de expresión de *GmSABP2-1* para la defensa de la soja contra SCN. El GmSABP2-1 es un componente importante de la recepción de la señal de MeSA para inducir SAR en la defensa de soja contra SCN, consideradas las diferencias temporales y espaciales de SCN en la llegada a los sitios de alimentación de los tejidos vasculares de soja dentro de las poblaciones de SCN. El nivel superior de la proteína GmSABP2-1 mediante la regulación positiva de su expresión de genes bajo la infección por SCN puede ser beneficioso para la captura de MeSA móvil. La desmetilación de MeSA catalizada por GmSABP2-1 puede ser capaz de aumentar el nivel de SA en los tejidos distales no dañados de la raíz de soja dado que su ortólogo NtSABP2 tiene el rol de inducir SAR en la hoja del tabaco no dañada.

10

15

A pesar de que la presente divulgación ha sido descrita con énfasis en realizaciones particulares, resultará obvio para los entendidos en la técnica que pueden utilizarse variaciones de las realizaciones particulares y se pretende que la divulgación pueda realizarse de una forma diferente a la descrita específicamente en la presente.

20 Debe entenderse que las características, compuestos, porciones químicas o ejemplos descritos junto con un aspecto, realización o ejemplo particular de la invención son aplicables a cualquier otro aspecto, realización o ejemplo de la invención. Por consiguiente, la presente divulgación incluye todas las modificaciones comprendidas dentro del espíritu y alcance de la divulgación como se define por las siguientes reivindicaciones.

25

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

REIVINDICACIONES

- 5 1. Una molécula de ácido nucleico aislada que comprende una secuencia de ácidos nucleicos que codifica una proteína resistente a plagas que es al menos un 80% idéntica a una proteína codificada por la secuencia de ácidos nucleicos de conformidad con la SEQ ID NO: 1 o SEQ ID NO: 2 o su variante degenerada o su fragmento funcional.
2. Una molécula de ácido nucleico aislada que comprende un promotor unido de forma operativa a la molécula de ácido nucleico de la reivindicación 1.
- 10 3. Un vector de expresión que comprende la molécula de ácido nucleico de la reivindicación 2.
4. El vector de la reivindicación 3, donde el vector es pJL-OFP-35S::GmSAMT1 (SEQ ID NO: 3) o pJL-OFP-35S::GmSABP2-1 (SEQ ID NO: 4).
- 15 5. Una célula hospedadora aislada transformada con el vector de la reivindicación 3 o la reivindicación 4.
6. Una construcción que comprende una molécula de ácido nucleico aislada de la reivindicación 1 unida de forma operativa a un promotor.
7. La construcción de la reivindicación 6, donde la construcción confiere un rasgo agronómico a una planta en la que se expresa.
- 20 8. La construcción de la reivindicación 7, donde el rasgo agronómico comprende la resistencia a plagas.
9. La construcción de la reivindicación 8, donde la resistencia a plagas comprende la resistencia al nematodo del quiste de la soja (*Heterodera glycines*, SCN).
- 25 10. Una planta transgénica transformada de forma estable con la construcción de cualquiera de las reivindicaciones 6 a 9.
11. Una semilla de la planta transgénica de la reivindicación 10, donde la semilla comprende la construcción.
12. La planta transgénica de la reivindicación 10, donde la planta es una
- 30 planta de soja.
13. Un método para producir una planta transgénica que comprende transformar una célula o tejido vegetal con la construcción de cualquiera de las reivindicaciones 6 a 9.
- 35 14. Un método para mejorar la resistencia a plagas en una planta que comprende transformar una célula o tejido vegetal con la construcción de

cualquiera de las reivindicaciones 6 a 9, lo que de ese modo mejora la resistencia a plagas en la planta.

15.El método de la reivindicación 13 o 14, donde la planta es una dicotiledónea.

5 16.El método de la reivindicación 13 o 14, donde la planta es una monocotiledónea.

17.El método de la reivindicación 14, donde la resistencia a plagas comprende la resistencia al nematodo del quiste de la soja (*Heterodera glycines*, SCN).

10 18.La planta transgénica de la reivindicación 14, donde la planta es una planta de soja.

19.Una célula o tejido vegetal transformado con la construcción de cualquiera de las reivindicaciones 6 a 9.

15 20.La célula o tejido vegetal de la reivindicación 19, donde la célula o el tejido vegetal es de una dicotiledónea.

21.La célula o tejido vegetal de la reivindicación 19, donde la célula o el tejido vegetal se deriva de una monocotiledónea.

20 22.Una célula vegetal, fruta, hoja, raíz, brote, flor, semilla, esqueje y otro material reproductivo útil para la propagación sexual o asexual, plantas de pro genie incluidos los híbridos de F1, plantas estériles macho y otras plantas y productos vegetales derivados de la planta transgénica de la reivindicación 10.

25 23.Un método para producir un producto vegetal de consumo, que comprende obtener la planta de la reivindicación 10 o una parte de esta, donde el producto vegetal de consumo es un concentrado de proteínas, un aislado de proteínas, cáscaras, material triturado, harina o aceite de soja y producir el producto vegetal de consumo a partir de estos.

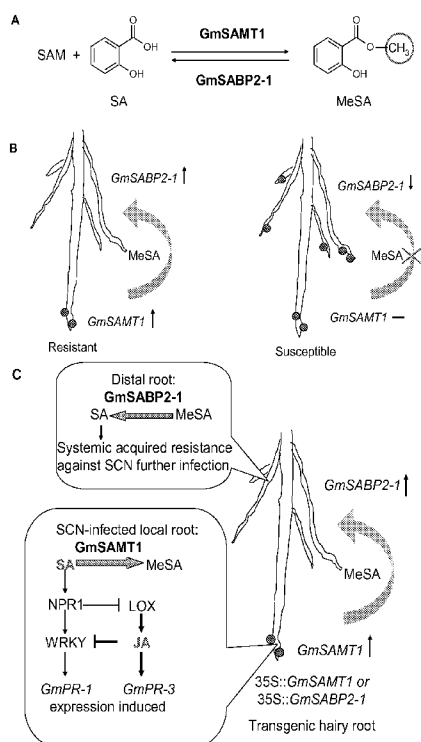


- (51) International Patent Classification:
C07H 21/04 (2006.01) *C12N 15/82* (2006.01)
- (21) International Application Number:
PCT/US2012/062759
- (22) International Filing Date:
31 October 2012 (31.10.2012)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
61/581,039 28 December 2011 (28.12.2011) US
- (71) Applicant: UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION [US/US]; UT Conference Center, 600 Henley Street, Suite 211, Knoxville, TN 37996 (US).
- (72) Inventors: STEWART, C., Neal; 1600 Wheaton Place, Knoxville, TN 37919 (US). LIN, Jingyu; 1504 Sutters Mill Lane, Knoxville, TN 37909 (US). MAZAREI, Mitra; 421 Anteeleh Trail, Knoxville, TN 37919 (US). CHEN, Feng; 1162 Keowee Ave., Knoxville, TN 37919 (US).
- (74) Agent: GRISWOLD, Ian, James; Johnson, Marcou & Isaacs, LLC, 317A East Liberty Street, Savannah, GA 31401 (US).
- (81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on next page]

(54) Title: INCREASING SOYBEAN DEFENSE AGAINST PESTS

FIG. 1



(57) Abstract: Disclosed are pest resistance genes isolated from soybean. In several embodiments, a pest resistance gene isolated from soybean includes a nucleic acid sequence at least 80% identical to SEQ ID NO: 1 or SEQ ID NO: 2 or a degenerate variant thereof, or a fragment thereof. Also disclosed are expression vectors and constructs that include such nucleic acids. Methods are disclosed of producing a transgenic plant that has enhanced resistance to pests. Transgenic plants, plant cells or tissue (such as a dicotyledon or a monocotyledon plants, plant cells or tissue) transformed with the disclosed constructs are also disclosed.



TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

— *with international search report (Art. 21(3))*

— *with sequence listing part of description (Rule 5.2(a))*

INCREASING SOYBEAN DEFENSE AGAINST PESTS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims the priority benefit of U.S. Provisional Application
5 61/581,039 filed on December 28, 2011, which is incorporated by reference herein
in its entirety.

FIELD OF THE DISCLOSURE

This disclosure concerns the field of plant molecular biology. More
10 specifically, this disclosure pertains to tolerance to pathogens conferred by transgene
expression of genes in soybeans and other plants.

BACKGROUND

Soybean [*Glycine max* (L.) Merr.] is the most widely grown legume in the
15 world, providing an important source of protein and oil. Soybean can be utilized in
many ways such as ingredients in the formulation of a multitude of human foods,
animal feed, and industrial products. As the dominant oil-seed in world trade,
soybean contributes about 56% of global oilseed production. The estimated value of
soybean products is about \$48.6 billion world-wide and about \$18.7 billion in the
20 U.S. alone. Therefore, soybean is considered as one of the most important economic
crops. Although global soybean production has increased steadily over the past
century, future demand for soybean still cannot be satisfied due to a growing world
population and limited land resources.

During its entire life cycle, soybean may be attacked by many pathogens,
25 such as fungi, bacteria, viruses, and nematodes, and suffer many diseases in any
tissue. For instance, the bacterial pathogen *Pseudomonas syringae* pv. *glycinea*, can
cause bacterial blight, overwinter in crop residue, and be transmitted by seed. In
laboratory studies, *P. syringae* is usually chosen as the model of bacterial pathogen.
Fungal pathogens can also cause soybean diseases on both above- and below-ground
30 tissues, accounting for approximately 50% of all soybean disease losses in the
United States and around the world. Soybean mosaic virus (SMV) detected on seeds
and on plants in the field, is the major virus causing soybean disease. Since this viral

disease is seed-borne, it may cause problems in the next growing season. Its effect on seed quality is the main concern. Virus infected soybean may produce fewer, smaller, and often mottled seed. Another concern is that SMV infected soybean tends to be dually infected by other viruses, which magnifies the risk of yield loss and reduces seed quality even further. In addition, nematodes and soybean insect pests such as beetles, caterpillars, aphids, and spider mites can also cause yield loss of soybean.

The soybean cyst nematode (*Heterodera glycines*, SCN), is the pest that causes the most economic damage of soybean in the United States, causing nearly \$1 billion in yield losses. SCN is a small plant-parasitic roundworm and most stages of SCN cannot be seen by unaided eyes. SCN feeds on the soybean roots and robs nutrients from the soybean. One important feature, genetic heterogeneity of SCN populations, endows this pathogen to readily overcome resistance. To complete their life cycle, infective second-stage juveniles (J2) enter host roots and migrate intracellularly within the cortical tissue to the vascular cylinder. The juveniles then initiate formation of specialized feeding sites called syncytia, which function as metabolic sinks to nourish the nematodes. The nematodes feed exclusively from their syncytia as they develop into adult males and females. Once fertilized, the female soybean nematode produces up to several hundred eggs that, for the most part, are retained within the nematode uterus. After the female's death, her body develops into a protective cyst around the eggs, giving the nematode its name.

SCN infestations are difficult to identify in the field because the above-ground symptoms of SCN infestations are easily confused with nutrient deficiency symptoms. When soybean plants are severely damaged by nematodes, they become stunted and turn chlorotic. These symptoms can be mistaken for nutrient deficiency and water stress. Until the population of SCN reaches a certain number, its damaging effect can not be clearly identified by farmers. Currently, the most accurate way of diagnosing SCN infection is to analyze a soil sample and observe the pest directly (Opperman and Bird 1998). Controlling SCN in commercial soybean productions still remains difficult because SCN has a short life cycle and populations can build rapidly. Frequent changes in population virulence of SCN also contribute to the difficulty in the management of this pest. In addition, the cysts of

SCN can survive in the soil for up to nine years and then break to release the eggs under proper conditions, increasing the probability of the nematodes' dispersing via infested soil.

The methods used to control and manage SCN in soybean production include crop rotation, the use of SCN-resistant cultivars, and the application of nematicides, which are often used in an integrated manner. However, these approaches always face economic restrictions, are time-consuming, and use of nematicides can result in environmental problems. Additionally, for some areas, economic factors may limit the use of crop rotation. Considering the long-term demands of soybean, it is critical to develop resistant soybean cultivars through breeding for managing soybean diseases. Even for crop rotation, resistant soybean cultivars are also required. Thus, the need exists for the creation of plants, such as soybeans, that are pest resistance, such as resistant to SCN infestations. This disclosure meets that need.

SUMMARY OF THE DISCLOSURE

Disclosed are previously unrecognized pest resistance genes isolated from soybean. In one embodiment, a pest resistance gene encodes a S-adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase and is designated *GmSAMT1*. In another embodiment, a pest resistance gene encodes a methyl salicylate esterase and is designated *GmSABP2-1*.

In several embodiments, a pest resistance gene isolated from soybean includes a nucleic acid sequence at least 80% identical to SEQ ID NO: 1 or SEQ ID NO: 2 or a degenerate variant thereof, or a fragment thereof that encodes an active S-adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase designated *GmSAMT1* or methyl salicylate esterase designated *GmSABP2-1*. Isolated nucleic acids encoding such genes are disclosed as are proteins encoded by such genes. The isolated nucleic acids can be operably linked to promoter. Also disclosed are expression vectors and constructs that include such nucleic acids. In some embodiments an expression vector is pJL-OFP-35S::*GmSAMT1* or pJL-OFP-35S::*GmSABP2-1*.

The constructs disclosed provide for enhanced resistance to pests, such as soybean cyst nematode infestation. Thus, methods are disclosed for producing a

transgenic plant that has enhanced resistance to pests, such as soybean cyst nematode infestation. Transgenic plants, plant cells or tissue (such as a dicotyledon or a monocotyledon plants, plant cells or tissue) transformed with the disclosed constructs are also disclosed. Also provided is a plant seed, fruit, leaf, root, shoot,
5 flower, cutting and other reproductive material useful in sexual or asexual propagation, progeny plants inclusive of F1 hybrids, male-sterile plants and all other plants and plant products derivable from the disclosed transgenic plant. Methods of producing the disclosed transgenic plants, plant cells or tissue are also provided herein.

10 The foregoing and other features and advantages will become more apparent from the following detailed description of several embodiments, which proceeds with reference to the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

15 **FIGS. 1A-1C** are a set of schematics showing a proposed model of soybean resistance against SCN and the functions of GmSAMT1 and GmSABP2-1. **1A** shows the reactions between SA and MeSA that are catalyzed by GmSAMT1 and GmSABP2-1. SAM represents S-adenosyl-L-methionine functioning as the donor of a methyl group. **1B** shows that in the resistant soybean root, the gene expression of
20 *GmSAMT1* and *GmSABP2-1* is up-regulated when soybean is infected by SCN. The efficient transportation of mobile signal MeSA increases the soybean resistance against SCN. In the susceptible soybean root, the gene expression of GmSAMT1 does not change and the gene expression of GmSABP2-1 is down-regulated, when soybean is infected by SCN. The transportation of mobile signal MeSA is blocked,
25 which decreases the soybean resistance against SCN. There are more SCNs completing their life cycle in the susceptible than the resistant soybean root. **1C** shows that in SCN-infected local soybean root, SA is biosynthesized and the level of SA increases. The level of SA can be manipulated by GmSAMT1 through converting to mobile signal MeSA. The residual level of SA is sufficient for
30 inducing *GmPR-1* gene expression. The decreased level of SA benefits jasmonic acid (JA)-dependent defense pathway, which induces *GmPR-3* gene expression. In undamaged soybean root, MeSA can be catalyzed by GmSABP2-1 to form SA,

which plays a role of systemic acquired resistance against further SCN infection. Non-expressors of pathogenesis-related genes 1 (NPR1), WRKY family transcription factor, and lipoxygenase (LOX) may function as the regulatory points. The dots represent the adult female SCNs within soybean roots.

5 **FIG. 2** shows the alignment of Affymetrix probe Gma.12911.1.A1_S_AT sequence (SEQ ID NO: 17) with five soybean candidate genes (SEQ ID NOS: 18-22). The sequence of Affymetrix probe Gma.12911.1.A1_S_AT aligned to the end of open reading frame sequence of the soybean genes and covers the part of 3' UTR sequence (TAA in the box is the stop codon). Based on the alignment, only
10 *Glyma02g06070* has the identical sequence with the probe (indicated in the box).

FIG. 3 shows a phylogenetic tree of *Glyma02g06070* and the functionally characterized SABATH family proteins from other plants. A neighbor-joining tree based on the degree of sequence similarity between *Glyma02g06070* and selected functionally characterized SABATH from other plant species using the ClustalX
15 program. *Glyma02g06070* is in the SAMT clade shown in the oval. CbSAMT, *Clarkia breweri* SAMT (AF133053); AmSAMT, *Antirrhinum majus* (snapdragon) SAMT (AF515284); SfSAMT, *Stephanotis floribunda* SAMT (AJ308570); HcSAMT, *Hoya carnos*
20 *nicotiana suaveolens* BSMT (AJ628349); AtBSMT, *Arabidopsis thaliana* BSMT (BT022049); AlBSMT, *A. lyrata* BSMT (AY224596); AbSAMT, *Atropa belladonna* SAMT (AB049752); OsBSMT1, *Oryza sativa* BSMT1 (XM467504); PhBSMT, *Petunia hybrida* BSMT (AY233465); OsIAMT1, *O. sativa* IAMT1 (EU375746); PtiIAMT1, *Populus trichocarpa* IAMT1 (XP_002298843); AtIAMT, *A. thaliana* IAMT (AK175586); AtGAMT1, *A. thaliana* GAMT1 (At4g26420);
25 AtGAMT2, *A. thaliana* GAMT2 (At5g56300); CaCaS1, *Coffea arabica* caffeine synthase 1 (AB086414); CaXMT1, *C. arabica* XMT1 (AB048793); CaDXMT1, *C. arabica* DXMT1 (AB084125). Branches were drawn to scale with the bar indicating 0.1 substitutions per site.

30 **FIG. 4** is a digital image of an SDS-PAGE gel of purified recombinant GmSAMT1 protein. His-tagged GmSAMT1 expressed in *E. coli* was purified as described in Example 1. Lane M contained protein molecular weight markers. Lane

1 contained crude extract and lane 2 contained about 2 µg of purified GmSAMT1 protein. The gel was stained with Coomassie blue.

FIGS. 5A-5D are set of graphs showing the biochemical properties of GmSAMT1. **FIG. 5A** shows the pH effect on GmSAMT1 activity. The level of GmSAMT1 activity in 50 mM Tris-HCl buffer, pH 7.5, was arbitrarily set at 1.0. **FIG. 5B** shows the GmSAMT1 activity in response to temperature ranging from 0 °C to 60 °C. The level of GmMAN1 activity at 25 °C was arbitrarily set at 1.0. **FIG. 5C** shows the effects of metal ions on activity of GmSAMT1. Metal ions were added to reactions in the form of chloride salts at 5 mM final concentrations. The level of GmSAMT1 activity without any metal ion added as control (Ctr) was arbitrarily set at 1.0. **FIG. 5D** shows GmSAMT1 displayed pseudo Michaelis-Menton kinetics, with apparent Km value of 46.2 ± 4.2 µM for salicylic acid.

FIG. 6 is a schematic showing the construction of pJL-OFP-35S::GmSAMT1, a plasmid for over-expression of *GmSAMT1* gene in soybean roots. pJL-OFP was modified from pCAMBIA 1305.2 by inserting a 35S promoter-*ppor RFP*-NOS terminator at the *NocI* site. *ppor RFP* encodes a orange fluorescent protein. pJL-OFP is an Agrobacterium compatible binary plasmid for plant transformation and used for a vector control plasmid. pJL-OFP-35S::GUS is an intermediate vector and used for testing whether *OFP* and *GUS* within two cassettes co-express in soybean roots. pJL-OFP-35S::GmSAMT1 is obtained by replacing *GUS* gene with *GmSAMT1* by digestion and ligation between the *BamH*I and *Sac*I sites.

FIGS. 7A and 7B are digital images showing a test for reporter gene expression in tobacco leaves and soybean hairy roots. **FIG. 7A** shows the transient expression of pJL-OFP and empty vector in tobacco leaves. **FIG. 7B** shows coexpression of two reporter genes in transgenic hairy roots. All the tested transgenic hairy roots with OFP signal showed GUS staining. A schematic representation of the construct used for coexpression of an orange fluorescent protein (OFP) reporter gene and a GUS reporter gene is shown on the top. “35S-Pro” and “NOS-ter” represent the CaMV 35S promoter and the NOS terminator, respectively. Analysis of transgenic hairy roots transformed with the construct (top). OFP-positive transgenic hairy roots were separated from OFP-negative hairy roots

under OFP filter. OFP-positive hairy roots were further demonstrated to be positive for GUS staining.

FIGS. 8A and 8B show the production of transgenic hairy roots of soybean overexpressing *GmSAMT1*. **FIG. 8A.** is a schematic representation of the construct used for coexpression of an orange fluorescent protein (OFP) reporter gene and *GmSAMT1*. “35S-Pro” and “NOS-ter” represent the CaMV 35S promoter and the NOS terminator, respectively. **FIG. 8B.** shows expression of *GmSAMT1* in five types of transgenic hairy roots from three genetic backgrounds TN02-226, TN02-275 and Williams 82. “*GmSAMT1*” represents transgenic hairy roots containing *GmSAMT1* transgene. “*VC*” represents transgenic hairy roots produced using a control vector containing OFP marker but not *GmSAMT1* gene. qRT-PCR was performed with *GmSAMT1*-specific primers. Expression values were normalized to the expression levels of the soybean ubiquitin 3 gene (*GmUBI3*) in respective samples. The level of *GmSAMT1* expression in the “*VC*” transgenic hairy roots of TN02-275 was arbitrarily set at 1.0. Each bar represents the mean relative expression level of three independent experiments with the standard errors, each containing a pool of five plants. The bars with asterisks are significantly different (LSD $p < 0.01$).

FIGS. 9A and 9B are a set of bar graphs showing qRT-PCR analysis on transgenic hairy roots for *GmPR-1*, and *GmPR-3*. “*GmSAMT1*” represents transgenic hairy roots containing *GmSAMT1* transgene. “*VC*” represents transgenic hairy roots produced using a control vector containing OFP marker but not *GmSAMT1* gene. Each bar represents the mean relative expression level of three independent experiments with the standard errors. In the graphs shown in **FIG. 9A** and **FIG. 9B**, the bars with different letters are significantly different (LSD $p < 0.05$).

FIG. 10 is a set of digital images showing the J2, J3, and J4 stages of soybean cyst nematodes shown in transgenic hairy roots by clearing and staining with acid fuchsin.

FIG. 11 is a bar graph and a set of digital images showing a SCN assay between transgenic hairy roots with over-expression of *GmSAMT1* and control hairy roots. Each bar represents the mean of resistance index with the standard error from 10 soybean hairy roots. The resistance index is the ratio of the sum of J3 and J4

nematodes to total infecting female nematodes shown in *A. rhizogenes*-transformed hairy roots after two-week post inoculation by SCN. VS and VR represent the hairy root lines harboring plasmid pJL-OFP in the susceptible (TN02-275 and Williams 82) and resistant (TN02-226) background, which serve as susceptible and resistant control lines respectively. SAMT represents the hairy root line with overexpression of *GmSAMT1* and OFP reporter gene in the susceptible background. Bars with asterisks are significantly different (LSD $p < 0.01$). The bottom photograph showed the representative stages of SCN in the responding transgenic hairy root lines.

FIGS. 12A and 12B are bar graphs showing relative transcript abundance of *GmICS1* (12A) and *GmICS2* (12B) in transgenic hairy roots with or without SCN infection. Two types of transgenic hairy roots of the Williams 82 background were produced and analyzed. “*GmSAMT1*” represents transgenic hairy roots containing *GmSAMT1* transgene. “*VC*” represents transgenic hairy roots produced using a control vector containing *OFP* marker gene but not *GmSAMT1*. “SCN” represents the administration of SCN infection. “Ctr” represents the uninfected control. qRT-PCR was performed with *GmSAMT1*-specific primers. Expression values were normalized to the expression levels of the soybean ubiquitin 3 gene (*GmUBI3*) in respective samples. The level of *GmSAMT1* expression in the “*VC*” transgenic hairy roots without SCN infection (“Ctr”) was arbitrarily set at 1.0. Each bar represents the mean relative expression level of three independent experiments with the standard errors, each containing a pool of five plants. Bars with different letters are significantly different (LSD $p < 0.05$).

FIGS. 13A and 13B are bar graphs showing relative transcript abundance of *GmNPR1-1* (13A) and *GmNPR1-2* (13B) in transgenic hairy roots with or without SCN infection. Two types of transgenic hairy roots of the Williams 82 background were produced and analyzed. “*GmSAMT1*” represents transgenic hairy roots containing *GmSAMT1* transgene. “*VC*” represents transgenic hairy roots produced using a control vector containing *OFP* marker gene but not *GmSAMT1*. “SCN” represents the administration of SCN infection. “Ctr” represents the uninfected control. qRT-PCR was performed with *GmSAMT1*-specific primers. Expression values were normalized to the expression levels of the soybean ubiquitin 3 gene (*GmUBI3*) in respective samples. The level of *GmSAMT1* expression in the “*VC*”

transgenic hairy roots without SCN infection (“Ctr”) was arbitrarily set at 1.0. Each bar represents the mean relative expression level of three independent experiments with the standard errors, each containing a pool of five plants. Bars with different letters are significantly different (LSD $p < 0.05$).

5 **FIG. 14** is a bar graph showing qRT-PCR analysis on soybean mosaic virus infected soybean leaves for *GmSAMT1* and *GmPR-1*. The grey and white bars represent the relative gene expression from the soybean plants with the resistant and susceptible responses, respectively. Each bar represents the mean relative expression level of three independent experiments with the standard errors. The bars with
10 star(s) mean the gene expression are significantly different with the bars without star(s) (LSD $p < 0.05$).

FIG. 15 shows a multiple sequence alignment of soybean Glyma16g26060 (SEQ ID NO: 23), tobacco SABP2 (NtSABP2) (SEQ ID NO: 24), and an Arabidopsis ortholog of SABP2 (At4g37150) (SEQ ID NO: 25). Identical residues
15 are shaded in black and similar residues in gray. The catalytic triad residues are indicated by arrows, and residues that contact salicylic acid are indicated with triangles.

FIG. 16 is a digital image of an SDS-PAGE gel of purified recombinant GmSABP2-1 protein. His-tagged GmSABP2-1 expressed in *E. coli* was purified as
20 described in Example 2. Lane M contained protein molecular weight markers. Lane 1 contained crude extract and lane 2 contained about 2 μ g of purified GmSABP2-1 protein. The gel was stained with Coomassie blue.

FIG. 17 is a graph showing the calculation of the K_m value of GmSABP2-1 using MeSA as substrate at concentrations of 5 μ M to 100 μ M. The apparent K_m
25 value was obtained through Lineweaver-Burk plots.

FIG. 18 is a graph showing the pH effect on GmSABP2-1 activity. Activity of the purified GmSABP2-1 was assayed from pH 6.0-9.0. The level of GmSABP2-1 activity in 50 mM Tris-HCl buffer, pH 7.0, was arbitrarily set at 1.0.

FIG. 19 is a bar graph showing a nematode demographic assay for SCN
30 resistance between transgenic hairy roots with over-expression of *GmSABP2-1* and control soybean lines. Ratio of the sum of J3 and J4 nematodes to the total infecting nematodes shown in hairy roots is used as the resistance index. The resistance index,

WS, WR, VS and VR have been described. WS and WR are wild type hairy root lines in the susceptible and resistant background, respectively. VS and VR are hairy root lines harboring OFP reporter gene in the susceptible and resistant background, respectively as the vector control lines. SABP2 represents the hairy root line with over-expressed *GmSABP2-1* and OFP reporter gene in TN02-275 background. Bars with different letters are significantly different (LSD $p < 0.05$).

FIG. 20 shows a plasmid map of the pJL-OFP-35S::*GmSAMT1* vector.

FIG. 21 shows a plasmid map of the pJL-OFP-35S::*GmSABP2-1* vector.

10 BRIEF DESCRIPTION OF THE SEQUENCE LISTING

The nucleic and amino acid sequences shown herein are shown using standard letter abbreviations for nucleotide bases and amino acids, as defined in 37 C.F.R. 1.822. Only one strand of each nucleic acid sequence is shown, but the complementary strand is understood as included by any reference to the displayed strand. The Sequence Listing is submitted as an ASCII text file in the form of the file named UTK_0141WP_ST25.txt, which was created on October 30, 2012, is 56 kilobytes, and is incorporated by reference herein.

SEQ ID NO: 1 is an exemplary nucleic acid sequence of *GmSAMT1*.

SEQ ID NO: 2 is an exemplary nucleic acid sequence of *GmSABP2-1*.

20 SEQ ID NO: 3 is the nucleic acid sequence of pJL-OFP-35S::*GmSAMT1*.

SEQ ID NO: 4 is the nucleic acid sequence of pJL-OFP-35S::*GmSABP2-1*.

SEQ ID NOS: 5-16 and 26-33 are the nucleic acid sequences of primers.

SEQ ID NO: 17 is the nucleic acid sequence of Affymetrix probe
Gma.12911.1.A1_S_AT.

25 SEQ ID NOS: 18-22 are the nucleic acid sequences of five soybean candidate genes.

SEQ ID NO: 23 is an amino acid sequence of Glyma16g26060.

SEQ ID NO: 24 is an amino acid sequence of tobacco SABP2

30 SEQ ID NO: 25 is an amino acid sequence of Arabidopsis ortholog of
SABP2.

DETAILED DESCRIPTION

I. Introduction

Soybean cyst nematode (*Heterodera glycines*, SCN) is the most severe pest of soybean [*Glycine max* (L.) Merr.]. As disclosed herein, through microarray
5 analysis on the gene expression of two genetically-related soybean lines (recombinant inbred lines, RILs) with different responses to SCN, the soybean genes *Glyma02g06070* and *Glyma16g26060* were identified to be SCN-induced in the resistant line but not in the susceptible line.

From sequence data, *Glyma02g06070* was predicted to encode an S-
10 adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase. Based on biochemical activity of *Escherichia coli*-expressed protein encoded by this gene, *Glyma02g06070* was identified to be a salicylic acid methyltransferase gene and designated *GmSAMT1*. *GmSAMT1* catalyzes the conversion of salicylic acid to methyl salicylate (MeSA). To determine the biological role of *GmSAMT1* in
15 soybean resistance against SCN, the inventors generated transgenic hairy roots in the susceptible soybean background with a binary vector harboring cassettes for the over-expression of *GmSAMT1* and orange fluorescent protein (OFP) gene. After two weeks post inoculation by SCNs the demographics of SCN in over-expressed *GmSAMT1* transgenic hairy roots and control hairy roots were compared. The result
20 showed the transgenic hairy roots with over-expression of *GmSAMT1* in the susceptible soybean line exhibited increased resistance, similar to that of the resistant soybean line. To further understand its defense role in soybean, the expression level of *GmSAMT1* was also compared in the compatible and incompatible reactions between soybean and soybean mosaic virus. The expression
25 of *GmSAMT1* was found to be up-regulated in the incompatible reaction, and not the compatible reaction. Taken together, the results disclosed herein indicate that *GmSAMT1* plays a critical role in soybean defense against SCN and soybean mosaic virus.

From sequence data, *Glyma16g26060* was predicted to encode an α/β fold
30 hydrolase. After cloning this gene and performing biochemical activity assay, this gene was identified to encode a methyl salicylate esterase. Since this gene was the ortholog gene of methyl salicylate esterase from tobacco, which was named

NtSABP2, it was designated the soybean gene *GmSABP2-1*. The recombinant enzyme encoded by *GmSABP2-1* can catalyze the conversion from MeSA to salicylic acid (SA). To explore the biological role of *GmSABP2-1* in soybean resistance against SCN, the inventors generated transgenic hairy roots with over-expression of *GmSABP2-1* and orange fluorescent protein (OFP) gene in the susceptible soybean background. After two weeks post inoculation by SCN eggs, the demographics of SCN in over-expressed *GmSABP2-1* transgenic hairy roots and control hairy roots were compared. The results showed that the transgenic hairy roots with over-expressed *GmSABP2-1* in the susceptible soybean line had increased resistance, which was similar to the resistant soybean control line. Therefore it was determined that *GmSABP2-1* plays a role in soybean resistance against SCN through SA-dependent signaling pathway.

Taken together, the two components, *GmSAMT1* and *GmSABP2-1*, work for the transportation of mobile signal MeSA and play an important role in soybean defense against SCN. While not being bound by theory, the functions of *GmSAMT1* and *GmSABP2-1* and a model of soybean resistance against SCN is proposed (see **FIG. 1**). In SCN-infected local root, *GmSAMT1* plays a critical role of regulating the levels of basal SA and MeSA. In the root knot nematode resistance research, the residual level of free SA was found to be sufficient to confer basal resistance. This indicates that the level of SA required for basal defense against SCN may be low. Biosynthesized SA can be changed to other forms, such as MeSA, to induce SAR in the distal root tissue. Relative high level of basal SA appear inhibit the activity of *GmSABP2-1* and therefore inhibit the conversation of MeSA to SA at the local site. In the distal root tissue, the relative high level of MeSA and low level of SA favors the activity of *GmSABP2-1*, leading to the conversion of MeSA to SA. Thus, as disclosed herein through suitable specific and inducible promoters driving *GmSAMT1* and *GmSABP2-1* in transgenic plants soybean resistance can be achieved against SCN by manipulating SA and MeSA levels.

II. Summary of Terms

Unless otherwise noted, technical terms are used according to conventional usage. Definitions of common terms in molecular biology may be found in

Benjamin Lewin, Genes IX, published by Jones and Bartlet, 2008 (ISBN 0763752223); Kendrew et al. (eds.), The Encyclopedia of Molecular Biology, published by Blackwell Science Ltd., 1994 (ISBN 0632021829); and Robert A. Meyers (ed.), Molecular Biology and Biotechnology: a Comprehensive Desk
5 Reference, published by VCH Publishers, Inc., 1995 (ISBN 9780471185710).

The singular terms “a,” “an,” and “the” include plural referents unless context clearly indicates otherwise. Similarly, the word “or” is intended to include “and” unless the context clearly indicates otherwise. The term “comprises” means “includes.” In case of conflict, the present specification, including explanations of
10 terms, will control.

To facilitate review of the various embodiments of this disclosure, the following explanations of terms are provided:

5' and/or 3': Nucleic acid molecules (such as, DNA and RNA) are said to have “**5' ends**” and “**3' ends**” because mononucleotides are reacted to make
15 polynucleotides in a manner such that the 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor in one direction via a phosphodiester linkage. Therefore, one end of a polynucleotide is referred to as the “5' end” when its 5' phosphate is not linked to the 3' oxygen of a mononucleotide pentose ring. The other end of a polynucleotide is referred to as the “3' end” when its
20 3' oxygen is not linked to a 5' phosphate of another mononucleotide pentose ring. Notwithstanding that a 5' phosphate of one mononucleotide pentose ring is attached to the 3' oxygen of its neighbor, an internal nucleic acid sequence also may be said to have 5' and 3' ends.

In either a linear or circular nucleic acid molecule, discrete internal elements
25 are referred to as being “**upstream**” or 5' of the “**downstream**” or 3' elements. With regard to DNA, this terminology reflects that transcription proceeds in a 5' to 3' direction along a DNA strand. Promoter and enhancer elements, which direct transcription of a linked gene, are generally located 5' or upstream of the coding region. However, enhancer elements can exert their effect even when located 3' of
30 the promoter element and the coding region. Transcription termination and polyadenylation signals are located 3' or downstream of the coding region.

Agronomic trait: Characteristic of a plant, which characteristics include, but are not limited to, plant morphology, physiology, growth and development, yield, nutritional enhancement, disease or pest resistance, or environmental or chemical tolerance. An “enhanced agronomic trait” refers to a measurable improvement in an agronomic trait including, but not limited to, yield increase, including increased yield under non-stress conditions and increased yield under environmental stress conditions. Stress conditions may include, for example, drought, shade, fungal disease, viral disease, bacterial disease, insect infestation, nematode infestation, cold temperature exposure, heat exposure, osmotic stress, reduced nitrogen nutrient availability, reduced phosphorus nutrient availability and high plant density. “Yield” can be affected by many properties including without limitation, plant height, pod number, pod position on the plant, number of internodes, incidence of pod shatter, grain size, efficiency of nodulation and nitrogen fixation, efficiency of nutrient assimilation, resistance to biotic and abiotic stress, carbon assimilation, plant architecture, resistance to lodging, percent seed germination, seedling vigor, and juvenile traits. In some embodiments, an agronomic trait is resistance to a pest, such as soybean resistance to SCN.

Altering level of production or expression: Changing, either by increasing or decreasing, the level of production or expression of a nucleic acid molecule or an amino acid molecule (for example a gene, a polypeptide, a peptide), as compared to a control level of production or expression.

Amplification: When used in reference to a nucleic acid, this refers to techniques that increase the number of copies of a nucleic acid molecule in a sample or specimen. An example of amplification is the polymerase chain reaction, in which a biological sample collected from a subject is contacted with a pair of oligonucleotide primers, under conditions that allow for the hybridization of the primers to nucleic acid template in the sample. The primers are extended under suitable conditions, dissociated from the template, and then re-annealed, extended, and dissociated to amplify the number of copies of the nucleic acid. The product of in vitro amplification can be characterized by electrophoresis, restriction endonuclease cleavage patterns, oligonucleotide hybridization or ligation, and/or nucleic acid sequencing, using standard techniques. Other examples of in vitro

amplification techniques include strand displacement amplification (see U.S. Patent No. 5,744,311); transcription-free isothermal amplification (see U.S. Patent No. 6,033,881); repair chain reaction amplification (see WO 90/01069); ligase chain reaction amplification (see EP-A-320 308); gap filling ligase chain reaction
5 amplification (see U.S. Patent No. 5,427,930); coupled ligase detection and PCR (see U.S. Patent No. 6,027,889); and NASBA™ RNA transcription-free amplification (see U.S. Patent No. 6,025,134).

Cassette: A manipulable fragment of DNA carrying (and capable of expressing) one or more genes of interest between one or more sets of restriction
10 sites. A cassette can be transferred from one DNA sequence (usually on a vector) to another by “cutting” the fragment out using restriction enzymes and “pasting” it back into the new context. In disclosed embodiments, a cassette includes or more of the disclosed inducible promoters.

cDNA (complementary DNA): A piece of DNA lacking internal, non-
15 coding segments (introns) and transcriptional regulatory sequences. cDNA may also contain untranslated regions (UTRs) that are responsible for translational control in the corresponding RNA molecule. cDNA is usually synthesized in the laboratory by reverse transcription from messenger RNA extracted from cells or other samples. In some examples cDNA is used as a source of a gene of interest, such as a gene
20 encoding GmSAMT1 or GmSABP2-1.

Construct: Any recombinant polynucleotide molecule such as a plasmid, cosmid, virus, autonomously replicating polynucleotide molecule, phage, or linear or circular single-stranded or double-stranded DNA or RNA polynucleotide molecule, derived from any source, capable of genomic integration or autonomous
25 replication, comprising a polynucleotide molecule where one or more transcribable polynucleotide molecule has been operably linked.

Control plant: A plant that does not contain a recombinant DNA that confers (for instance) an enhanced or altered agronomic trait in a transgenic plant, is used as a baseline for comparison, for instance in order to identify an enhanced or
30 altered agronomic trait in the transgenic plant. A suitable control plant may be a non-transgenic plant of the parental line used to generate a transgenic plant, or a plant that at least is non-transgenic for the particular trait under examination (that is,

the control plant may have been engineered to contain other heterologous sequences or recombinant DNA molecules). Thus, a control plant may in some cases be a transgenic plant line that comprises an empty vector or marker gene, but does not contain the recombinant DNA, or does not contain all of the recombinant DNAs, in
5 the test plant.

Degenerate variant and conservative variant: A polynucleotide encoding a polypeptide or an antibody that includes a sequence that is degenerate as a result of the genetic code. There are 20 natural amino acids, most of which are specified by more than one codon. Therefore, all degenerate nucleotide sequences are included as
10 long as the amino acid sequence of the polypeptide encoded by the nucleotide sequence is unchanged. Because of the degeneracy of the genetic code, a large number of functionally identical nucleic acids encode any given polypeptide. For instance, the codons CGU, CGC, CGA, CGG, AGA, and AGG all encode the amino acid arginine. Thus, at every position where an arginine is specified within a protein
15 encoding sequence, the codon can be altered to any of the corresponding codons described without altering the encoded protein. Such nucleic acid variations are “silent variations,” which are one species of conservative variations. Each nucleic acid sequence herein that encodes a polypeptide also describes every possible silent variation. One of skill will recognize that each codon in a nucleic acid (except AUG,
20 which is ordinarily the only codon for methionine) can be modified to yield a functionally identical molecule by standard techniques. Accordingly, each “silent variation” of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

Furthermore, one of ordinary skill will recognize that individual
25 substitutions, deletions or additions which alter, add or delete a single amino acid or a small percentage of amino acids (for instance less than 5%, such as less than 4%, less than 3%, less than 2%, or even less than 1%) in an encoded sequence are conservative variations where the alterations result in the substitution of an amino acid with a chemically similar amino acid.

30 Conservative amino acid substitutions providing functionally similar amino acids are well known in the art. The following six groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Serine (S), Threonine (T);
- 2) Aspartic acid (D), Glutamic acid (E);
- 3) Asparagine (N), Glutamine (Q);
- 4) Arginine (R), Lysine (K);
- 5 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V); and
- 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W).

Not all residue positions within a protein will tolerate an otherwise “conservative” substitution. For instance, if an amino acid residue is essential for a function of the protein, even an otherwise conservative substitution may disrupt that activity.

Disease resistance or pest resistance: The avoidance of the harmful symptoms that are the outcome of the plant-pathogen interactions. Disease resistance and pest resistance genes such as lysozymes or cecropins for antibacterial protection, or proteins such as defensins, glucanases or chitinases for antifungal protection, or *Bacillus thuringiensis* endotoxins, protease inhibitors, collagenases, lectins, or glycosidases for controlling nematodes or insects are all examples of useful gene products.

As used herein, the term “pest” includes, but is not limited to, insects, fungi, bacteria, viruses, nematodes, mites, ticks, and the like. Insect pests include insects selected from the orders *Coleoptera*, *Diptera*, *Hymenoptera*, *Lepidoptera*, *Mallophaga*, *Homoptera*, *Hemiptera*, *Orthoptera*, *Thysanoptera*, *Dermaptera*, *Isoptera*, *Anoplura*, *Siphonaptera*, *Trichoptera*, etc., particularly *Coleoptera*, *Lepidoptera*, and *Diptera*. Viruses include but are not limited to tobacco or cucumber mosaic virus, ringspot virus, necrosis virus, maize dwarf mosaic virus, etc. Nematodes include but are not limited to parasitic nematodes such as root knot, cyst, and lesion nematodes, including *Heterodera* spp., *Meloidogyne* spp., and *Globodera* spp.; particularly members of the cyst nematodes, including, but not limited to, *Heterodera glycines* (soybean cyst nematode); *Heterodera schachtii* (beet cyst nematode); *Heterodera avenae* (cereal cyst nematode); and *Globodera rostochiensis* and *Globodera pallida* (potato cyst nematodes). Lesion nematodes include but are not limited to *Pratylenchus* spp. Fungal pests include those that cause leaf, yellow, stripe and stem rusts. **DNA (deoxyribonucleic acid):** DNA is a

long chain polymer which comprises the genetic material of most organisms (some viruses have genes comprising ribonucleic acid (RNA)). The repeating units in DNA polymers are four different nucleotides, each of which comprises one of the four bases, adenine, guanine, cytosine and thymine bound to a deoxyribose sugar to which a phosphate group is attached. Triplets of nucleotides (referred to as codons) code for each amino acid in a polypeptide, or for a stop signal. The term codon is also used for the corresponding (and complementary) sequences of three nucleotides in the mRNA into which the DNA sequence is transcribed.

Unless otherwise specified, any reference to a DNA molecule includes the reverse complement of that DNA molecule. Except where single-strandedness is required by the text herein, DNA molecules, though written to depict only a single strand, encompass both strands of a double-stranded DNA molecule.

Encode: A polynucleotide is said to encode a polypeptide if, in its native state or when manipulated by methods known to those skilled in the art, the polynucleotide molecule can be transcribed and/or translated to produce a mRNA for and/or the polypeptide or a fragment thereof. The anti-sense strand is the complement of such a nucleic acid, and the encoding sequence can be deduced therefrom.

Enhancer domain: A cis-acting transcriptional regulatory element (a.k.a. cis-element) that confers an aspect of the overall control of gene expression. An enhancer domain may function to bind transcription factors, which are trans-acting protein factors that regulate transcription. Some enhancer domains bind more than one transcription factor, and transcription factors may interact with different affinities with more than one enhancer domain. Enhancer domains can be identified by a number of techniques, including deletion analysis (deleting one or more nucleotides from the 5' end or internal to a promoter); DNA binding protein analysis using DNase I foot printing, methylation interference, electrophoresis mobility-shift assays, in vivo genomic foot printing by ligation-mediated PCR, and other conventional assays; or by DNA sequence comparison with known cis-element motifs using conventional DNA sequence comparison methods. The fine structure of an enhancer domain can be further studied by mutagenesis (or substitution) of one or more nucleotides or by other conventional methods. Enhancer domains can be

obtained by chemical synthesis or by isolation from promoters that include such elements, and they can be synthesized with additional flanking nucleotides that contain useful restriction enzyme sites to facilitate subsequence manipulation.

(Gene) Expression: Transcription of a DNA molecule into a transcribed RNA molecule. More generally, gene expression encompasses the processes by which a gene's coded information is converted into the structures present and operating in the cell. Expressed genes include those that are transcribed into mRNA and then translated into protein and those that are transcribed into RNA but not translated into protein (for example, siRNA, transfer RNA and ribosomal RNA). Thus, expression of a target sequence, such as a gene or a promoter region of a gene, can result in the expression of an mRNA, a protein, or both. The expression of the target sequence can be inhibited or enhanced (decreased or increased). Gene expression may be described as related to temporal, spatial, developmental, or morphological qualities as well as quantitative or qualitative indications.

Gene regulatory activity: The ability of a polynucleotide to affect transcription or translation of an operably linked transcribable polynucleotide molecule, such as an inducible promoter. An isolated polynucleotide molecule having gene regulatory activity may provide temporal or spatial expression or modulate levels and rates of expression of the operably linked transcribable polynucleotide molecule. An isolated polynucleotide molecule having gene regulatory activity may include a promoter, intron, leader, or 3' transcription termination region.

Genetic material: A phrase meant to include all genes, nucleic acid, DNA and RNA.

Heterologous nucleotide sequence: A sequence that is not naturally occurring with a promoter sequence. While this nucleotide sequence is heterologous to the promoter sequence, it may be homologous, or native, or heterologous, or foreign, to the plant host. The invention additionally encompasses expression of the homologous coding sequences of the promoters, particularly the coding sequences related to the resistance phenotype. The expression of the homologous coding sequences will alter the phenotype of the transformed plant or plant cell.

Increasing pest resistance or enhancing pest resistance: An enhanced or elevated resistance to a pest over a normal or control plant or part thereof (for example a plant that has not been transformed with an isolated nucleic acid encoding S-adenosyl-L-methionine dependent carboxyl or methyl salicylate esterase, such as disclosed herein). In some examples, an increase or enhancement is an elevation of
5 at least about 25%, 50%, 75%, 100%, 150%, 200%, 300%, 400%, 500% or more.

In *cis*: Indicates that two sequences are positioned on the same piece of RNA or DNA.

In *trans*: Indicates that two sequences are positioned on different pieces of
10 RNA or DNA.

Insert DNA: Heterologous DNA within an expression cassette, such as the disclosed expression cassette, used to transform the plant material while “flanking DNA” can comprise either genomic DNA naturally present in an organism such as a plant, or foreign (heterologous) DNA introduced via the transformation process
15 which is extraneous to the original insert DNA molecule, e.g. fragments associated with the transformation event. A “flanking region” or “flanking sequence” as used herein refers to a sequence of at least 20, 50, 100, 200, 300, 400, 1000, 1500, 2000, 2500, or 5000 base pair or greater which is located either immediately upstream of and contiguous with or immediately downstream of and contiguous with the original
20 foreign insert DNA molecule.

Isolated: An “isolated” biological component (such as a nucleic acid, peptide or protein) has been substantially separated, produced apart from, or purified away from other biological components in the cell of the organism in which the component naturally occurs, e.g., other chromosomal and extrachromosomal DNA
25 and RNA, and proteins. Nucleic acids, peptides and proteins which have been “isolated” thus include nucleic acids and proteins purified by standard purification methods. The term also embraces nucleic acids, peptides and proteins prepared by recombinant expression in a host cell as well as chemically synthesized nucleic acids.

Operably linked: This term refers to a juxtaposition of components, particularly nucleotide sequences, such that the normal function of the components can be performed. Thus, a first nucleic acid sequence is operably linked with a
30

second nucleic acid sequence when the first nucleic acid sequence is placed in a functional relationship with the second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Generally, operably linked DNA
5 sequences are contiguous and, where necessary to join two protein-coding regions, in the same reading frame. A coding sequence that is "operably linked" to regulatory sequence(s) refers to a configuration of nucleotide sequences wherein the coding sequence can be expressed under the regulatory control (e.g., transcriptional and/or translational control) of the regulatory sequences.

10 **Plant:** Any plant and progeny thereof. The term also includes parts of plants, including seed, cuttings, tubers, fruit, flowers, etc. As used herein, the term plant includes plant cells, plant organs, plant protoplasts, plant cell tissue cultures from which plants can be regenerated, plant calli, plant clumps, and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves,
15 flowers, branches, fruit, stalks, roots, root tips, anthers, and the like. Progeny, variants, and mutants of the regenerated plants are also included within the scope of the invention. The term **plant cell**, as used herein, refers to the structural and physiological unit of plants, consisting of a protoplast and the surrounding cell wall, including those with genetic alteration, such as transformation, has been affected as
20 to a gene of interest, or is a plant or plant cell which is descended from a plant or cell so altered and which comprises the alteration. A "control" or "control plant" or "control plant cell" provides a reference point for measuring changes in phenotype of the subject plant or plant cell. A control plant or plant cell may comprise, for example: (a) a wild-type plant or cell, i.e., of the same genotype as the starting
25 material for the genetic alteration which resulted in the subject plant or cell; (b) a plant or plant cell of the same genotype as the starting material but which has been transformed with a null construct (i.e. with a construct which has no known effect on the trait of interest, such as a construct comprising a marker gene); (c) a plant or plant cell which is a non-transformed segregant among progeny of a subject plant or
30 plant cell; (d) a plant or plant cell genetically identical to the subject plant or plant cell but which is not exposed to conditions or stimuli that would induce expression of the gene of interest; or (e) the subject plant or plant cell itself, under conditions in

which the gene of interest is not expressed. The term **plant organ**, as used herein, refers to a distinct and visibly differentiated part of a plant, such as root, stem, leaf or embryo. More generally, the term **plant tissue** refers to any tissue of a plant in planta or in culture. This term includes a whole plant, plant cell, plant organ,
5 protoplast, cell culture, or any group of plant cells organized into a structural and functional unit.

Polynucleotide molecule: Single- or double-stranded DNA or RNA of genomic or synthetic origin; that is, a polymer of deoxyribonucleotide or ribonucleotide bases, respectively, read from the 5' (upstream) end to the 3'
10 (downstream) end.

Promoter: An array of nucleic acid control sequences which direct transcription of a nucleic acid, by recognition and binding of e.g., RNA polymerase II and other proteins (trans-acting transcription factors) to initiate transcription. A promoter includes necessary nucleic acid sequences near the start site of
15 transcription, such as, in the case of a polymerase II type promoter, a TATA element. Minimally, a promoter typically includes at least an RNA polymerase binding site together with one or more transcription factor binding sites, which modulate transcription in response to occupation by transcription factors. Representative examples of promoters (and elements that can be assembled to
20 produce a promoter) are described herein. Promoters may be defined by their temporal, spatial, or developmental expression pattern.

A plant promoter is a native or non-native promoter that is functional in plant cells. In one example, a promoter is a high level constitutive promoter, such as a tissue specific promoter.

25 **Protein:** A biological molecule, for example a polypeptide, expressed by a gene and comprised of amino acids.

Protoplast: An isolated plant cell without cell walls, having the potential for regeneration into cell culture or a whole plant.

Purified: The term purified does not require absolute purity; rather, it is
30 intended as a relative term. Thus, for example, a purified protein preparation is one in which the protein is more enriched than the protein is in its generative environment, for instance within a cell or in a biochemical reaction chamber.

Preferably, a preparation of protein is purified such that the protein represents at least 50% of the total protein content of the preparation.

Recombinant: A recombinant nucleic acid is one that has a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques. Similarly, a recombinant protein is one encoded for by a recombinant nucleic acid molecule.

Regulatable promoter or inducible promoter: A promoter the activity of which is regulated (directly or indirectly) by an agent, such as a transcription factor, a chemical compound, an environmental condition, or a nucleic acid molecule.

Regulating gene expression: Processes of controlling the expression of a gene by increasing or decreasing the expression, production, or activity of an agent that affects gene expression. The agent can be a protein, such as a transcription factor, or a nucleic acid molecule, such as a miRNA or an siRNA molecule, which when in contact with the gene or its upstream regulatory sequences, or a mRNA encoded by the gene, either increases or decreases gene expression.

Regulatory sequences or elements: These terms refer generally to a class of polynucleotide molecules (such as DNA molecules, having DNA sequences) that influence or control transcription or translation of an operably linked transcribable polynucleotide molecule, and thereby expression of genes. Included in the term are promoters, enhancers, leaders, introns, locus control regions, boundary elements/insulators, silencers, Matrix attachment regions (also referred to as scaffold attachment regions), repressor, transcriptional terminators (a.k.a. transcription termination regions), origins of replication, centromeres, and meiotic recombination hotspots. Promoters are sequences of DNA near the 5' end of a gene that act as a binding site for RNA polymerase, and from which transcription is initiated. Enhancers are control elements that elevate the level of transcription from a promoter, usually independently of the enhancer's orientation or distance from the promoter. Locus control regions (LCRs) confer tissue-specific and temporally regulated expression to genes to which they are linked. LCRs function

independently of their position in relation to the gene, but are copy-number dependent. It is believed that they function to open the nucleosome structure, so other factors can bind to the DNA. LCRs may also affect replication timing and origin usage. Insulators (also known as boundary elements) are DNA sequences that
5 prevent the activation (or inactivation) of transcription of a gene, by blocking effects of surrounding chromatin. Silencers and repressors are control elements that suppress gene expression; they act on a gene independently of their orientation or distance from the gene. Matrix attachment regions (MARs), also known as scaffold attachment regions, are sequences within DNA that bind to the nuclear scaffold.
10 They can affect transcription, possibly by separating chromosomes into regulatory domains. It is believed that MARs mediate higher-order, looped structures within chromosomes. Transcriptional terminators are regions within the gene vicinity that RNA polymerase is released from the template. Origins of replication are regions of the genome that, during DNA synthesis or replication phases of cell division, begin
15 the replication process of DNA. Meiotic recombination hotspots are regions of the genome that recombine more frequently than the average during meiosis. Specific nucleotides within a regulatory region may serve multiple functions. For example, a specific nucleotide may be part of a promoter and participate in the binding of a transcriptional activator protein. Isolated regulatory elements that function in cells
20 (for instance, in plants or plant cells) are useful for modifying plant phenotypes, for instance through genetic engineering.

RNA: A typically linear polymer of ribonucleic acid monomers, linked by phosphodiester bonds. Naturally occurring RNA molecules fall into three general classes, messenger (mRNA, which encodes proteins), ribosomal (rRNA,
25 components of ribosomes), and transfer (tRNA, molecules responsible for transferring amino acid monomers to the ribosome during protein synthesis). Messenger RNA includes heteronuclear (hnRNA) and membrane-associated polysomal RNA (attached to the rough endoplasmic reticulum). Total RNA refers to a heterogeneous mixture of all types of RNA molecules.

30 **Screenable Marker:** A marker that confers a trait identified through observation or testing.

Selectable Marker: A marker that confers a trait that one can select for by

chemical means, e.g., through the use of a selective agent (e.g., an herbicide, antibiotic, or the like). Selectable markers include but are not limited to antibiotic resistance genes, such as, kanamycin (nptII), G418, bleomycin, hygromycin, chloramphenicol, ampicillin, tetracycline, or the like. Additional selectable markers
5 include a bar gene which codes for bialaphos resistance; a mutant EPSP synthase gene which encodes glyphosate resistance; a nitrilase gene which confers resistance to bromoxynil; a mutant acetolactate synthase gene (ALS) which confers imidazolinone or sulphonylurea resistance; or a methotrexate resistant DHFR gene. In one example, the selectable marker is AAD1.

10 **Sequence identity:** The similarity between two nucleic acid sequences, or two amino acid sequences, is expressed in terms of the similarity between the sequences, otherwise referred to as sequence identity. Sequence identity is frequently measured in terms of percentage identity (or similarity or homology); the higher the percentage, the more similar the two sequences are. Percent sequence
15 identity is represented as the identity fraction multiplied by 100. The comparison of one or more polynucleotide or polypeptide sequences may be to a full-length polynucleotide or polypeptide sequence or a portion thereof, or to a longer polynucleotide sequence.

Methods of alignment of sequences for comparison are well known in the
20 art. Various programs and alignment algorithms are described in: Smith and Waterman (Adv. Appl. Math. 2: 482, 1981); Needleman and Wunsch (J. Mol. Biol. 48: 443, 1970); Pearson and Lipman (PNAS. USA 85: 2444, 1988); Higgins and Sharp (Gene, 73: 237-244, 1988); Higgins and Sharp (CABIOS 5: 151-153, 1989); Corpet et al. (Nuc. Acids Res. 16: 10881-90, 1988); Huang et al. (Comp. Appls
25 Biosci. 8: 155-65, 1992); and Pearson et al. (Methods in Molecular Biology 24: 307-31, 1994). Altschul et al. (Nature Genet., 6: 119-29, 1994) presents a detailed consideration of sequence alignment methods and homology calculations.

The alignment tools ALIGN (Myers and Miller, CABIOS 4:11-17, 1989) or LFASTA (Pearson and Lipman, 1988) may be used to perform sequence
30 comparisons (Internet Program © 1996, W. R. Pearson and the University of Virginia, "fasta20u63" version 2.0u63, release date December 1996). ALIGN compares entire sequences against one another, while LFASTA compares regions of

local similarity. These alignment tools and their respective tutorials are available on the Internet at with a web address of biology.ncsa.uiuc.edu.

Orthologs or paralogs (more generally, homologs) of a specified sequence are typically characterized by possession of greater than 75% sequence identity
5 counted over the full-length alignment with the amino acid sequence of a specified protein (or the nucleic acid sequence of a specified nucleic acid molecule) using ALIGN set to default parameters. Sequences with even greater similarity to the reference sequences will show increasing percentage identities when assessed by this method, such as at least 80%, at least 85%, at least 90%, at least 92%, at least
10 95%, or at least 98% sequence identity. In such an instance, percentage identities will be essentially similar to those discussed for full-length sequence identity.

An alternative indication that two nucleic acid molecules are closely related is that the two molecules hybridize to each other under stringent conditions. Stringent conditions are sequence-dependent and are different under different
15 environmental parameters. Generally, stringent conditions are selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Conditions for nucleic acid hybridization and calculation of
20 stringencies can be found in Sambrook et al. (In Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York, 1989) and Tijssen (Laboratory Techniques in Biochemistry and Molecular Biology Part I, Ch. 2, Elsevier, New York, 1993). Nucleic acid molecules that hybridize under stringent conditions to a specified protein sequence will typically hybridize to a probe based on either the protein
25 encoding sequence, an entire domain, or other selected portions of the encoding sequence under wash conditions of 0.2 x SSC, 0.1% SDS at 65°C.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences, due to the degeneracy of the genetic code. It is understood that changes in nucleic acid sequence can be made
30 using this degeneracy to produce multiple nucleic acid sequences that each encode substantially the same protein. **Substantial percent sequence identity** is at least about 80% sequence identity, such as at least about 80%, at least about 85%, at least

about 90%, at least about 95%, or even greater sequence identity, such as about 98% or about 99% sequence identity.

Soybean: Glycine max and includes all plant varieties that can be bred with soybean.

- 5 **Soybean cyst nematode resistance gene:** A gene that, when expressed in a plant contributes to resistance to soybean cyst nematode. Alternatively, a soybean cyst nematode gene may be an allelic variant of the resistance gene particularly variants resulting in a susceptible phenotype.

- 10 A **transgenic event** is produced by transformation of plant cells with a heterologous DNA construct(s), including a nucleic acid expression cassette that includes a transgene of interest, the regeneration of a population of plants resulting from the insertion of the transgene into the genome of the plant, and selection of a particular plant characterized by insertion into a particular genome location. In some embodiments of this disclosure, the transgene of interest is operable linked to a
- 15 disclosed inducible promoter, such as SEQ ID NO: 1 or 2. An event is characterized phenotypically by the expression of the transgene(s). At the genetic level, an event is part of the genetic makeup of a plant. The term “event” also refers to progeny produced by a sexual outcross between the transformant and another variety that include the heterologous DNA. Even after repeated back-crossing to a recurrent
- 20 parent, the inserted DNA and flanking DNA from the transformed parent is present in the progeny of the cross at the same chromosomal location. The term “event” also refers to DNA from the original transformant comprising the inserted DNA and flanking sequence immediately adjacent to the inserted DNA that would be expected to be transferred to a progeny that receives inserted DNA including the transgene of
- 25 interest as the result of a sexual cross of one parental line that includes the inserted DNA (e.g., the original transformant and progeny resulting from selfing) and a parental line that does not contain the inserted DNA.

Transgenic plant: A plant that contains a foreign (heterologous) nucleotide sequence inserted into either its nuclear genome or organellar genome.

- 30 **Transgene:** A nucleic acid sequence that is inserted into a host cell or host cells by a transformation technique.

Transgenic: This term refers to a plant/fungus/cell/other entity or organism

that contains recombinant genetic material not normally found in entities of this type/species (that is, heterologous genetic material) and which has been introduced into the entity in question (or into progenitors of the entity) by human manipulation. Thus, a plant that is grown from a plant cell into which recombinant DNA is
5 introduced by transformation (a transformed plant cell) is a transgenic plant, as are all offspring of that plant that contain the introduced transgene (whether produced sexually or asexually).

Transformation: Process by which exogenous DNA enters and changes a recipient cell. It may occur under natural conditions, or artificial conditions using
10 various methods well known in the art. Transformation may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. Selection of the method is influenced by the host cell being transformed and may include, but is not limited to, viral infection, electroporation, lipofection, and particle bombardment.

Vector: A nucleic acid molecule as introduced into a host cell, thereby
15 producing a transformed host cell. A vector may include nucleic acid sequences that permit it to replicate in the host cell, such as an origin of replication. A vector may also include one or more therapeutic genes and/or selectable marker genes and other genetic elements known in the art. A vector can transduce, transform or infect a cell,
20 thereby causing the cell to express nucleic acids and/or proteins other than those native to the cell. A vector optionally includes materials to aid in achieving entry of the nucleic acid into the cell, such as a viral particle, liposome, protein coating or the like.

Suitable methods and materials for the practice or testing of this disclosure
25 are described below. Such methods and materials are illustrative only and are not intended to be limiting. Other methods and materials similar or equivalent to those described herein can be used. For example, conventional methods well known in the art to which a disclosed invention pertains are described in various general and more specific references, including, for example, Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 2d ed., Cold Spring Harbor Laboratory Press, 1989; Sambrook
30 *et al.*, *Molecular Cloning: A Laboratory Manual*, 3d ed., Cold Spring Harbor Press, 2001; Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing

Associates, 1992 (and Supplements to 2000); and Ausubel *et al.*, *Short Protocols in Molecular Biology: A Compendium of Methods from Current Protocols in Molecular Biology*, 4th ed., Wiley & Sons, 1999.

5 **III. Description of Several Embodiments**

 The present disclosure describes novel pest resistance genes, such a nucleic acid sequence encoding GmSAMT1 (such as set forth in SEQ ID NO: 1) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2) or functional fragment thereof. Also provided are DNA constructs comprising a described pest resistance gene. In
10 one embodiment, the pest resistance gene confers an agronomic trait to a plant in which it is expressed, for example pest resistance, such as resistance to SCN.

 Also provided are transgenic plants. In one embodiment, a transgenic plant is stably transformed with a disclosed DNA construct. In some embodiments, the transgenic plant is a dicotyledon. In other embodiments, the transgenic plant is a
15 monocotyledon. In one particular embodiment, the transgenic plant is a soybean plant. Further provided is a seed of a disclosed transgenic plant. In one embodiment, the seed comprises the disclosed DNA construct. Even further provided is a transgenic plant cell or tissue. In one embodiment, a transgenic plant cell or tissue comprises a disclosed pest resistance gene. In some embodiments, the plant cell or
20 tissue is derived from a dicotyledon. In other embodiments, the plant cell or tissue is from a monocotyledon. In one particular embodiment, the plant cell or tissue is from a soybean plant.

 Also provided are methods of producing a disclosed transgenic plant, plant cell, seed or tissue. In some embodiments, the method comprises transforming a
25 plant cell or tissue with a disclosed DNA construct. In some embodiments, the method is a method of enhancing disease resistance in a plant.

 Further provided are a plant cell, fruit, leaf, root, shoot, flower, seed, cutting and other reproductive material useful in sexual or asexual propagation, progeny plants inclusive of F1 hybrids, male-sterile plants and all other plants and plant
30 products derivable from the disclosed transgenic plants.

A. Plant Pathogen Resistance Genes

The present disclosure provides previously unrecognized pest resistance genes isolated from soybean. In one embodiment, a pest resistance gene encodes a S-adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase and is designated *GmSAMT1*. In another embodiment, a pest resistance gene encodes a methyl salicylate esterase and is designated *GmSABP2-1*.

In one embodiment, a *GmSAMT1* nucleic acid sequence is provided as SEQ ID NO: 1 below:

```

ATGGAAGTAGCACAGGTACTCCACATGAACGGTGGCGTTGGAGACGCAAGCTATGC
10 AAACAACCTCCCTTGTTTCAGCAAAAGGTGATTTGTTTGACAAAGCCCATAAGAGAGG
AAGCCATAAGAAGCCTCTATTGCAGCACACACCCCAGAAGCTTGGCAATTGCAGAT
TTGGGTGCTCTTCTGGACCAACACCTTTATTTGTTGTGTCTGAATTCATAAAAAT
TGTGGAGAAGCTTTGCCGAGAGCTGAACCACAAATCTCCAGAATACAAAGTCTTTC
TGAATGATCTCCCTGGGAATGACTTCAACAACATCTTCAAGTCTCTTGACAGCGTC
15 AAAGAGAAATTGTGTGATGAAATGGAAAGTGGGATCGGTCCATGCTACTTCTCGGG
TGTTCCCGGTTCTTTCTATGGAAGGGTTTTCCCATATCAAAGTCTTCATTTTGTCC
ATTCTTCATACAGCCTTCAATGGCTATCTAAGGTTCTTGAGGGTGTAGACAACAAC
AAGGGCAATGTTTACATAGGCAGTACGAGCCCCAAAATGTTGTGAGAGCTTACTA
TGAGCAATTTTCAGAGAGATTTCTCTCTTTTTCTCAAGTGTCGTGCAGAGGAATTGG
20 TTGAAGGAGGACGCATGGTTCTCACATTTTTGGGAAGAAGAAGCGATGATCCATCT
AGCAAGGATGGTTGCTACATTTGGGAGCTTTTGGCTACTGCTCTTAGTGATATGGT
CTTGCAGGGAATCATAAGAGAAGAGCAATTAGATACTTTTAACATCCCTCAATACA
CTCCATCCCCATCTGAAGTGAAATTGGAAGTTCTTAAAGAAGGATCATTCGCCATC
AATCGTCTAGAGGTTTCTGAGGTGAATTGGAATGCTCTCGATGAGTGGAATGCTCT
25 AGACTTTGAATCTGAAAGGTCTGAATCACTTAGTGATGGTGAATACAATGTGGCAC
AGTGCATGAGGGCTGTGGCAGAACCTATGCTGATTAGCCACTTTGGTGAAGCTATC
ATTGAAGAGGTTTTTTTGCCGCTACCAGCAAATCTTGGCTGAACGTATGTCCAAGGA
GAAAACCAAGTTCATCAATGTTACCATATTATTGACTAGAAAAGCATAA (SEQ ID
30 NO: 1). While a particular nucleic acid sequence has been shown for GmSAMT1, it
is understood that a GmSAMT1 nucleic acid sequence includes any nucleic acid
sequence redundant by virtue of the degeneracy of genetic code that encodes a
GmSAMT1 protein, such as encoded by SEQ ID NO: 1.

```

In one embodiment, a *GmSABP2-1* nucleic acid sequence is provided as

SEQ ID NO: 2 below:

ATGGGTTTCACAAAATTGTATGGATAGGAAGCACTATGTTCTGGTGCATGGGGCATG
 CCATGGAGCTTGGTGTGGTATAAGCTCAAGCCACGCTTGAATCTGCAGGCCATA
 AGGTCACAGTACTTGACCTTGACGCTTCTGGAACCAACATGAAGAAAATTGAAGAT
 5 GTTGATACTTTCTCAGAGTATTCTGCGCCTTTGTTGCAGCTAATGGCCACAATTCC
 CTCAAATGAGAAGTTAGTTCTAGTTGGTCACAGCCTTGGAGGGCTGAACATAGCAC
 TTGCAATGGAGAAATTCCCAGAAAAGGTAGCAGTTGGTGTTTTCTTAACAGCTTTT
 GCTCCAGACACTGAACACCACCCATCTTATGTCTTGGAAAAGTACAATGAGAGGAC
 CCCGTTAGCTGCATGGTTAGACACTGAATTTGCTCCAAGCGGAAACAAAACATCAA
 10 TGTTCTTCGGCCCCCACTTCTTGTCGACAAGCTCTACCAACTGTCCCCTATTGAG
 GACTTGGAATTGGCCAAGACTTTAGCAAGGCCATCATCGCTCTTCATGGAAGACTT
 GACTAAACAAAAGAACTTCTCCAAAGAGGGATATGGGTCAGTTCCACGTGCCTTTA
 TTGTTTGCACGGAGGACCTTGAATTCCATTGGAATATCAGCTCTTGATGATCCAA
 AATGTTGGGTTCAATGACGTTGTAGAGGTCAAAGACGCAGATCATATGGTTATGCT
 15 TTGCAAGCCACAAGAACTATTCGATTCCCTCCAGCAGATAGCGACTAAATATGCAT
 GA (SEQ ID NO: 2).

While a particular nucleic acid sequence has been shown for *GmSABP2-1*, it is understood that a *GmSABP2-1* nucleic acid sequence includes any nucleic acid sequence redundant by virtue of the degeneracy of genetic code that encodes a GmSABP2-1 protein, such as encoded by SEQ ID NO: 2.

20 Variants of the disclosed pest resistance genes isolated from soybean are also contemplated by this disclosure. Variant nucleotide sequences also include synthetically derived nucleotide sequences, such as those generated, for example, by using site- directed mutagenesis, but which when expressed still exhibit pest resistance activity. Methods for mutagenesis and nucleotide sequence alterations are well known in the art. See, for example, Kunkel (1985) Proc. Natl Acad. Sci. USA 25 52:488-492; Kunkel et al. (1987) Methods in Enzymol. 75:367-382; US Patent No. 4,873,192; Walker and Gaastra, eds. (1983) Techniques in Molecular Biology (MacMillan Publishing Company, New York) and the references cited therein. It will further be understood that amino acid sequences encoded by *GmSABP2-1* 30 nucleic acids and *GmSAMT1* nucleic acids will typically tolerate substitutions in the amino acid sequence and substantially retain biological activity. To routinely identify biologically active amino acid substitutions may be based on any characteristic known in the art, including the relative similarity or differences of the

amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Generally, nucleotide sequence variants will encode a protein have at least 80% sequence identity to the protein encoded by a disclosed pest resistance gene, such as at least at 80%, at least 85%, at least 90%,
5 at least, 95% at least 96%, at least 97%, at least 98% at least 99% sequence identity or even greater to the protein encoded by its respective reference pest resistance gene nucleotide sequence. Thus, disclosed are *GmSABP2-1* nucleic acids that encode a protein with about 80% sequence homology to the protein encoded by SEQ ID NO: 2 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least
10 90%, at least, 95% at least 96%, at least 97%, at least 98% at least 99% sequence identity or even greater to the protein encoded by SEQ ID NO: 2 or a degenerate nucleic acid. Also disclosed are *GmSAMT1* nucleic acids that have about 80% sequence homology to the protein encoded by SEQ ID NO: 1 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least 90%, at least, 95% at least
15 96%, at least 97%, at least 98% at least 99% sequence identity or even greater to the protein encoded by SEQ ID NO: 1 or a degenerate nucleic acid. In some embodiments, a nucleic acid that have about 80% sequence homology to the protein encoded by SEQ ID NO: 1 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least 90%, at least, 95% at least 96%, at least 97%, at least 98% at least
20 99% sequence identity or even greater to the protein encoded by SEQ ID NO: 1 or a degenerate nucleic acid does not have an A at position number 627 relative to SEQ ID NO: 1. Also disclosed are *GmSABP2-1* nucleic acids that have about 80% sequence homology to SEQ ID NO: 2 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least 90%, at least, 95% at least 96%, at least 97%, at least
25 98% at least 99% sequence identity or even greater to SEQ ID NO: 2 or a degenerate nucleic acid. Also disclosed are *GmSAMT1* nucleic acids that have about 80% sequence homology to SEQ ID NO: 1 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least 90%, at least, 95% at least 96%, at least 97%, at least 98% at least 99% sequence identity or even greater to SEQ ID NO: 1 or a
30 degenerate nucleic acid. In some embodiments, a nucleic acid that have about 80% sequence homology to SEQ ID NO: 1 or a degenerate nucleic acid, such as at least at 80%, at least 85%, at least 90%, at least, 95% at least 96%, at least 97%, at least

98% at least 99% sequence identity or even greater to SEQ ID NO: 1 or a degenerate nucleic acid does not have an A at position number 627 relative to SEQ ID NO: 1.

In some embodiments, a disclosed pest resistance nucleic acid encodes a functional fragment of a GmSABP2-1 or GmSAMT1 nucleic acids. Such functional fragments of the GmSABP2-1 and GmSAMT1 still exhibit pest resistance activity. Functional fragments include proteins in which residues at the N-terminus, C-terminus and/or internal to the full length protein have been deleted.

The protein encoded by the *GmSABP2-1* gene is a member of the α/β hydrolase superfamily of enzymes, with Ser, Asp His catalytic triad, the location of which is shown by the arrows in FIG. 15. In addition, the crystal structure of the related enzyme from tobacco has been solved to atomic resolution (see Forouhar *et al.*, *PNAS* 102 (5) 1773-1778, 2005). Thus, one of skill in the art can produce functional fragments and/or variants of gmSABP2 (and nucleic acids encoding such fragments) for example while maintaining the catalytic triad. Such fragments can be tested for activity with minimal experimentation given the guidance presented in the specification. For example based on the structure of the α/β hydrolase superfamily of enzymes, one of skill in the art can make deletions of a few amino acids on the N-terminus, C-terminus and/or internal loops of the protein encoded by the *GmSABP2-1* gene without affecting the active site catalytic triad. For example a deletion of less than about 50, 40, 30, 25, 20, 15, 10, 5, 4, 3, 2, or 1 deletions from the N-terminus, C-terminus and/or internal loops while maintaining the active site catalytic triad with minimal testing and/or experimentation to determine the activity of the resultant protein.

The protein encoded by the *GmSAMT1* gene is a S-adenosyl-L-methionine (SAM) dependent carboxyl methyltransferase designated *GmSAMT1* belonging to the SABATH methyltransferases family. The crystal structure of members of the SABATH family have been solved to atomic resolution and the active site of the enzyme mapped (see Zhao *et al.* *Plant Physiol.* 146(2): 455–467, 2008). Thus, one of skill in the art can produce functional fragments and/or variants of GmSAMT1 (and nucleic acids encoding such fragments) for example while maintaining the catalytic triad. uch fragments can be tested for activity with minimal experimentation given

the guidance presented in the specification. For example based on the structure of the SABATH family of enzymes, one of skill in the art can make deletions of a few amino acids on the N-terminus, C-terminus and/or internal loops of the protein encoded by the *GmSAMT1* gene without affecting the active site c. For example a
5 deletion of less than about 50, 40, 30, 25, 20, 15, 10, 5, 4, 3, 2, or 1 deletions from the N-terminus, C-terminus and/or internal loops while maintaining the active site catalytic triad with minimal testing and/or experimentation to determine the activity of the resultant protein.

10 **B. Expression Cassettes**

The nucleotide sequences for the disclosed pest resistance genes, such a nucleic acid sequence encoding GmSAMT1 (such as set forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2 or a degenerate variant thereof) or functional fragment thereof, are useful in the
15 genetic manipulation of any plant to confer resistance to pests, such as insects, fungi, bacteria, viruses, nematodes, mites, ticks, and the like, for example SCN when operably linked with a promoter, such as an inducible or constitutive promoter. By “operably linked” is intended the transcription or translation of the disclosed pest resistance genes is under the influence of the promoter sequence. In
20 this manner, the nucleotide sequences for the disclosed pest resistance genes are provided in expression cassettes for expression in the plant of interest. Such expression cassettes will typically comprise a transcriptional initiation region comprising a promoter nucleotide sequence operably linked to one or more of the disclosed pest resistance genes or variants thereof. Such an expression cassette can
25 be provided with a plurality of restriction sites for insertion of the nucleotide sequence to be under the transcriptional regulation of the regulatory regions. The expression cassette may additionally contain selectable marker genes or sequences. The expression cassettes of this disclosure can be part of and an expression vector, such as a plasmid.

30 In some embodiments, the transcriptional cassette will include in the 5'-to-3' direction of transcription, a transcriptional and translational initiation region, a pest resistance gene, such a nucleic acid sequence encoding GmSAMT1 (such as set

forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2 or a degenerate variant thereof) or functional fragment thereof, and a transcriptional and translational termination region functional in plant cells. The termination region may be native with the transcriptional initiation region, may be native with the pest resistance gene, or may be derived from another source. Convenient termination regions are available from the Ti-plasmid of *A. tumefaciens*, such as the octopine synthase and nopaline synthase termination regions. See also, Guerineau *et al.*, *Mol. Gen. Genet.* 262:141-144, 1991; Proudfoot *Cell* 64:671-674, 1991; Sanfacon *et al.*, *Genes Dev.* 5:141-149, 1991; Mogen *et al.*, *Plant Cell* 2:1261-1272, 1990; Munroe *et al.*, *Gene* 91 :151-158, 1990; Ballas *et al.*, *Nucleic Acids Res.* 17:7891-7903, 1989; Joshi *et al.*, *Nucleic Acid Res.* 15:9627-9639, 1987.

An expression cassette including a disclosed a pest resistance gene, such a nucleic acid sequence encoding GmSAMT1 (such as set forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2 or a degenerate variant thereof) or functional fragment thereof, operably linked to a promoter sequence may also contain at least additional nucleotide sequence for a gene to be cotransformed into the organism. Alternatively, the additional sequence(s) can be provided on another expression cassette.

Where appropriate, the nucleotide sequences whose expression is desired may be optimized for increased expression in the transformed plant. That is, these nucleotide sequences can be synthesized using plant preferred codons for improved expression. Methods are available in the art for synthesizing plant-preferred nucleotide sequences. See, for example, U.S. Patent Nos. 5,380,831 and 5,436,391, and Murray *et al.*, *Nucleic Acids Res.* 17:477-498, 1989.

Additional sequence modifications are known to enhance gene expression in a cellular host. These include elimination of sequences encoding spurious polyadenylation signals, exon-intron splice site signals, transposon-like repeats, and other such well-characterized sequences that may be deleterious to gene expression. The G-C content of the heterologous nucleotide sequence may be adjusted to levels average for a given cellular host, as calculated by reference to known genes expressed in the host cell. When possible, the sequence is modified to avoid predicted hairpin secondary mRNA structures.

The expression cassettes may additionally contain 5' leader sequences in the expression cassette construct. Such leader sequences can act to enhance translation. Translation leaders are known in the art and include: picornavirus leaders, for example, EMCV leader (Encephalomyocarditis 5' noncoding region) (Elroy-Stein *et al.*, *Proc. Nat. Acad. Sci. USA* 86:6126-6130, 1989); potyvirus leaders, for example, TEV leader (Tobacco Etch Virus); MDMV leader (Maize Dwarf Mosaic Virus); human immunoglobulin heavy-chain binding protein (BiP) (Macejak and Sarnow *Nature* 353:90-94, 1991); untranslated leader from the coat protein mRNA of alfalfa mosaic virus (AMV RNA 4) (Jobling and Gehrke *Nature* 325:622-625, 1987); tobacco mosaic virus leader (TMV) (Gallie *et al.* *Molecular Biology of RNA*, pages 237-256, 1989; and maize chlorotic mottle virus leader (MCMV) (Lommel *et al.*, *Virology* 81:382-385, 1991). See also Della-Cioppa *et al.*, *Plant Physiology* 84:965-968, 1987. Other methods known to enhance translation and/or mRNA stability can also be utilized, for example, introns, and the like.

In those instances where it is desirable to have the expressed product of the a pest resistance gene directed to a particular organelle, such as the chloroplast or mitochondrion, or secreted at the cell's surface or extracellularly, the expression cassette may further comprise a coding sequence for a transit peptide. Such transit peptides are well known in the art and include, but are not limited to, the transit peptide for the acyl carrier protein, the small subunit of RUBISCO, plant EPSP synthase, and the like.

In preparing the expression cassette, the various DNA fragments may be manipulated by methods known in the art, so as to provide for the DNA sequences in the proper orientation and, as appropriate, in the proper reading frame. Toward this end, adapters or linkers may be employed to join the DNA fragments or other manipulations may be involved to provide for convenient restriction sites, removal of superfluous DNA, removal of restriction sites, or the like. For this purpose, in vitro mutagenesis, primer repair, restriction, annealing, resubstitutions, for example, transitions and transversions, may be involved.

The expression cassettes may include reporter genes or selectable marker genes. Examples of suitable reporter genes known in the art can be found in, for example, Jefferson *et al.* in *Plant Molecular Biology Manual*, ed. Gelvin *et al.*

(Kluwer Academic Publishers), pp. 1-33, 1991; DeWet *et al.*, *Mol. Cell. Biol.* 7:725-737, 1987; Goff *et al.*, *EMBO J.* 9:2517-2522, 1990; and Kain *et al.*, *BioTechniques* 19:650- 655, 1995; and Chiu *et al.*, *Current Biology* 6:325-330, 1996. Selectable marker genes for selection of transformed cells or tissues can include genes that

5 confer antibiotic resistance or resistance to herbicides. Examples of suitable selectable marker genes include, but are not limited to, genes encoding resistance to chloramphenicol (Herrera Estrella *et al.*, *EMBO J.* 2:987-992, 1983); methotrexate (Herrera Estrella *et al.*, *Nature* 303:209-213, 1983; Meijer *et al.*, *Plant Mol. Biol.* 16:807-820, 1991); hygromycin (Waldron *et al.*, *Plant Mol. Biol.* 5:103-108, 1985;

10 Zhijian *et al.*, *Plant Science* 108:219-227, 1995); streptomycin (Jones *et al.*, *Mol. Gen. Genet.* 210:86-91, 1987); spectinomycin (Bretagne-Sagnard *et al.*, *Transgenic Res.* 5:131-137, 1996); bleomycin (Hille *et al.*, *Plant Mol. Biol.* 7:171-176, 1990); sulfonamide (Guerineau *et al.*, *Plant Mol. Biol.* 15:127-136, 1990); bromoxynil (Stalker *et al.*, *Science* 242:419-423, 1988); glyphosate (Shaw *et al.*, *Science*

15 233:478-481, 1986); and phosphinothricin (DeBlock *et al.*, *EMBO J.* 6:2513-2518, 1987).

Other genes that could serve utility in the recovery of transgenic events but might not be required in the final product would include, but are not limited to, such examples as GUS (b-glucuronidase; Jefferson *Plant Mol. Biol. Rep.* 5:387, 1987),

20 GFP and other related fluorescent proteins, and luciferase. In some embodiments, an expression cassette is pJL-OFP-35S::GmSAMT1 or pJL-OFP-35S::GmSABP2-1. The following is the sequence of binary vector pJL-OFP-35S::GmSAMT1 showing reverse complemented sequence of GmSAMT1 in bold.

TCATGGCTCTTTCAAAGCAAAGTGGGGTCAAAGATGTAATGAACACCGAGCTTCAT

25 ATGGACGGGATCGTCAATGGACACCCCTTTGAGATAAAAGGAAAAGGAAAGGGAAA
CCCGTACAAGGGTGTGCAGACCATGAACTTACAGTCATTAAGGGTGC GCCTTTGC
CATTTTCTATTGACATTTTGCTGCCTCAACACATGTATGGAAGCAAGCCATTTATT
AAGTATCCTGAGAGTATCCCAGACTACATCAAGTTGTCATTTCCCGAGGGAATCAC
ATGGGAAAGGTCCATGACCTTTGAAGATGGTGCAGTGTGCACTGTCTCTAACGACT

30 CCAGGCTCGATGGCGACTCTTTTCATCTACGAAGTCAGGTTTCTTGGCGTGAACCTT
CCCCGAGATGGACCTGTTATGCAGAAGAAGACGCGAGGCTGGGACCCGTCCACAGA
GAGACTGTATGAGTGTGGTGGGTGGCAGAGAGGAGACGTCCACATGGCCTTGAAGT

TGGAGAACGGTGGCCATTATACGTGCGACTTCAAACTACTTACAAATCAAAGAAG
GGCTTGAAGGTGCCACCGTATCACTTCGTTGACCACAACTAGATTTACTGAGCCA
CAACACCGATGGTGCTACCTTTGAAGAGTTTGAACAACGAGAAATTGCACATGCAC
ATCTTTCTAACTTACCGGTAGCCCATCACCATCACCATCACTAAGATCGTTCAAAC
5 ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGA
CGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTC
ATCTATGTTACTAGATCCCATGGCTACTACTAAGCATTTGGCTCTTGCCATCCTTG
10 TCCTCCTTAGCATTGGTATGACCACCAGTGCAAGAACCCTCCTAGATCTGAGGGTA
AATTTCTAGTTTTTCTCCTTCATTTTCTTGGTTAGGACCCTTTTCTCTTTTTATTT
TTTTGAGCTTTGATCTTTCTTTAACTGATCTATTTTTTAATTGATTGGTTATGGT
GTAAATATTACATAGCTTTAACTGATAATCTGATTACTTTATTTCTGTGTCTATG
ATGATGATGATAGTTACAGAACCGACGAACTAGTCTGTACCCGATCAACACCGAGA
15 CCCGTGGCGTCTTCGACCTCAATGGCGTCTGGAACCTCAAGCTGGACTACGGGAAA
GGACTGGAAGAGAAGTGGTACGAAAGCAAGCTGACCGACACTATTAGTATGGCCGT
CCCAAGCAGTTACAATGACATTGGCGTGACCAAGGAAATCCGCAACCATATCGGAT
ATGTCTGGTACGAACGTGAGTTCACGGTGCCGGCCTATCTGAAGGATCAGCGTATC
GTGCTCCGCTTCGGCTCTGCAACTCACAAAGCAATTGTCTATGTCAATGGTGAGCT
20 GGTCGTGGAGCACAAGGGCGGATTCCTGCCATTCTGAAGCGGAAATCAACAACTCGC
TGCGTGATGGCATGAATCGCGTCACCGTCGCCGTGGACAACATCCTCGACGATAGC
ACCCTCCCGGTGGGGCTGTACAGCGAGCGCCACGAAGAGGGCCTCGGAAAAGTCAT
TCGTAACAAGCCGAACCTTCGACTTCTTCAACTATGCAGGCCTGCACCGTCCGGTGA
AAATCTACACGACCCCGTTTACGTACGTGAGGACATCTCGGTTGTGACCGACTTC
25 AATGGCCCAACCGGGACTGTGACCTATACGGTGGACTTTCAAGGCAAAGCCGAGAC
CGTGAAAGTGTCGGTCGTGGATGAGGAAGGCAAAGTGGTCGCAAGCACCGAGGGCC
TGAGCGGTAACGTGGAGATTCCGAATGTCATCCTCTGGGAACCACTGAACACGTAT
CTCTACCAGATCAAAGTGGAACCTGGTGAACGACGGACTGACCATCGATGTCTATGA
AGAGCCGTTTCGGCGTGCGGACCGTGGAAGTCAACGACGGCAAGTTCCTCATCAACA
30 ACAAACCGTTCTACTTCAAGGGCTTTGGCAAACATGAGGACACTCCTATCAACGGC
CGTGGCTTTAACGAAGCGAGCAATGTGATGGATTTCAATATCCTCAAATGGATCGG
CGCCAACAGCTTCCGGACCGCACACTATCCGTACTCTGAAGAGTTGATGCGTCTTG
CGGATCGCGAGGGTCTGGTCGTGATCGACGAGACTCCGGCAGTTGGCGTGCACCTC
AACTTCATGGCCACCACGGGACTCGGCGAAGGCAGCGAGCGCGTCAGTACCTGGGA

GAAGATTCGGACGTTTGTGAGCACCATCAAGACGTTCTCCGTGAACTGGTGTCTCGTG
ACAAGAACCATCCAAGCGTCGTGATGTGGAGCATCGCCAACGAGGCGGCGACTGAG
GAAGAGGGCGCGTACGAGTACTTCAAGCCGTTGGTGGAGCTGACCAAGGAACTCGA
CCCACAGAAGCGTCCGGTCACGATCGTGCTGTTTGTGATGGCTACCCCGGAGACGG
5 ACAAAGTCGCCGAACTGATTGACGTCATCGCGCTCAATCGCTATAACGGATGGTAC
TTCGATGGCGGTGATCTCGAAGCGGCCAAAGTCCATCTCCGCCAGGAATTTACGC
GTGGAACAAGCGTTGCCCAGGAAAGCCGATCATGATCACTGAGTACGGCGCAGACA
CCGTTGCGGGCTTTCACGACATTGATCCAGTGATGTTACCGAGGAATATCAAGTC
GAGTACTACCAGGCGAACCACGTCGTGTTTCGATGAGTTTGAGAACTTCGTGGGTGA
10 GCAAGCGTGGAACCTTCGCGGACTTCGCGACCTCTCAGGGCGTGATGCGCGTCCAAG
GAAACAAGAAGGGCGTGTTCACTCGTGACCGCAAGCCGAAGCTCGCCGCGCACGTC
TTTCGCGAGCGCTGGACCAACATTCCAGATTTTCGGCTACAAGAACGCTAGCCATCA
CCATCACCATCACGTGTGAATTGGTGACCAGCTCGAATTTCCCGATCGTTCAAAC
ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
15 CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGA
CGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAAATTATCGCGCGCGGTGTC
ATCTATGTTACTAGATCGGGAATTAACTATCAGTGTTTGACAGGATATATTGGCG
GGTAAACCTAAGAGAAAAGAGCGTTTATTAGAATAACGGATATTTAAAGGGCGTG
20 AAAAGGTTTATCCGTTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCTC
GGGATCAAAGTACTTTGATCCAACCCCTCCGCTGCTATAGTGCACTCGGCTTCTGA
CGTTCACTGACGCGTCTTCTGAAAACGACATGTCGCACAAGTCCTAAGTTACGCG
ACAGGCTGCCGCCCTGCCCTTTTCTGGCGTTTTCTTGTCGCGTGTTTTAGTCGCA
TAAAGTAGAATACTTGCGACTAGAACCGGAGACATTACGCCATGAACAAGAGCGCC
25 GCCGCTGGCCTGCTGGGCTATGCCCCGCTCAGCACCGACGACCAGGACTTGACCAA
CCAACGGGCCGAACTGCACGCGGCCGGCTGCACCAAGCTGTTTTCCGAGAAGATCA
CCGGCACCAGGCGCGACCGCCCGGAGCTGGCCAGGATGCTTGACCACCTACGCCCT
GGCGACGTTGTGACAGTGACCAGGCTAGACCGCCTGGCCCGCAGCACCCGCGACCT
ACTGGACATTGCCGAGCGCATCCAGGAGGCCGGCGCGGGCCTGCGTAGCCTGGCAG
30 AGCCGTGGGCCGACACCACCACGCCGGCCGGCCGCATGGTGTTGACCGTGTTTCGCC
GGCATTGCCGAGTTCGAGCGTTCCCTAATCATCGACCGCACCCGGAGCGGGCGCGA
GGCCGCCAAGGCCCGAGGCGTGAAGTTTGGCCCCGCCCTACCTCACCCGGCAC
AGATCGCGCACGCCCGCGAGCTGATCGACCAGGAAGGCCGCACCGTGAAAGAGGCG
GCTGCACTGCTTGGCGTGATCGCTCGACCCTGTACCGCGCACTTGAGCGCAGCGA

GGAAGTGACGCCCACCGAGGCCAGGCGGCGCGGTGCCTTCCGTGAGGACGCATTGA
CCGAGGCCGACGCCCTGGCGGCCGCCGAGAATGAACGCCAAGAGGAACAAGCATGA
AACCGCACCAGGACGGCCAGGACGAACCGTTTTTCATTACCGAAGAGATCGAGGCG
GAGATGATCGCGGCCGGGTACGTGTTTCGAGCCGCCCGCGCACGTCTCAACCGTGCG
5 GCTGCATGAAATCCTGGCCGGTTTTGTCTGATGCCAAGCTGGCGGCCTGGCCGGCCA
GCTTGGCCGCTGAAGAAACCGAGCGCCGCCGTCTAAAAAGGTGATGTGTATTTGAG
TAAACAGCTTGCGTCATGCGGTGCTGCGTATATGATGCGATGAGTAAATAAACA
AATACGCAAGGGGAACGCATGAAGGTTATCGCTGTACTTAACCAGAAAGGCGGGTC
AGGCAAGACGACCATCGCAACCCATCTAGCCCGCGCCCTGCAACTCGCCGGGGCCG
10 ATGTTCTGTTAGTCGATTCCGATCCCCAGGGCAGTGCCCGCGATTGGGCGGCCGTG
CGGGAAGATCAACCGCTAACCGTTGTTCGGCATCGACCGCCCGACGATTGACCGCGA
CGTGAAGGCCATCGGCCGGCGCGACTTCGTAGTGATCGACGGAGCGCCCCAGGCGG
CGGACTTGGCTGTGTCCGCGATCAAGGCAGCCGACTTCGTGCTGATTCCGGTGCGAG
CCAAGCCCTTACGACATATGGGCCACCGCCGACCTGGTGAGCTGGTTAAGCAGCG
15 CATTGAGGTCACGGATGGAAGGCTACAAGCGGCCCTTTGTCTGTGTCGCGGGCGATCA
AAGGCACGCGCATCGGCGGTGAGGTTGCCGAGGCGCTGGCCGGGTACGAGCTGCCC
ATTCTTGAGTCCCGTATCACGCAGCGCGTGAGCTACCCAGGCACTGCCGCCGCCGG
CACAACCGTTCTTTGAATCAGAACCCGAGGGCGACGCTGCCCCGCGAGGTCCAGGCGC
TGGCCGCTGAAATTAAATCAAACTCATTTGAGTTAATGAGGTAAAGAGAAAATGA
20 GCAAAAGCACAAACACGCTAAGTGCCGGCCGTCCGAGCGCACGCAGCAGCAAGGCT
GCAACGTTGGCCAGCCTGGCAGACACGCCAGCCATGAAGCGGGTCAACTTTTCAGTT
GCCGGCGGAGGATCACACCAAGCTGAAGATGTACGCGGTACGCCAAGGCAAGACCA
TTACCGAGCTGCTATCTGAATACATCGCGCAGCTACCAGAGTAAATGAGCAAATGA
ATAAATGAGTAGATGAATTTTAGCGGCTAAAGGAGGCGGCATGGAAAATCAAGAAC
25 AACCAGGCACCGACGCCGTGGAATGCCCCATGTGTGGAGGAACGGGCGGTTGGCCA
GGCGTAAGCGGCTGGGTTGTCTGCCGGCCCTGCAATGGCACTGGAACCCCCAAGCC
CGAGGAATCGGCGTGACGGTCGCAAACCATCCGGCCCCGTACAAATCGGCGCGGGC
CTGGGTGATGACCTGGTGGAGAAGTTGAAGGCCGCGCAGGCCGCCAGCGGCAACG
CATCGAGGCAGAAGCACGCCCCGGTGAATCGTGGCAAGCGGCCGCTGATCGAATCC
30 GCAAAGAATCCCGGCAACCGCCGGCAGCCGGTGCGCCGTGATTAGGAAGCCGCC
AAGGGCGACGAGCAACCAGATTTTTTCGTTCCGATGCTCTATGACGTGGGCACCCG
CGATAGTCGCAGCATCATGGACGTGGCCGTTTTCCGTCTGTGGAAGCGTGACCGAC
GAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGACGGGCACGTAGAGGTTTCCGCA
GGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACCTGGTACTGATGGCGGTTTTC

CCATCTAACCGAATCCATGAACCGATACCGGGAAGGGAAGGGAGACAAGCCCCGGCC
GCGTGTTCCGTCCACACGTTGCGGACGTACTCAAGTTCTGCCGGCGAGCCGATGGC
GGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTTCGGTTAAACACCACGCACGT
TGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCCTGGTGACGGTATCCGAGG
5 GTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCGAAACCGGGCGGCCGGAG
TACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAGATCACAGAAGGCAAGAA
CCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTTGATCGATCCCGGCATCGGCC
GTTTTCTCTACCGCCTGGCACGCCGCGCCGCAGGCAAGGCAGAAGCCAGATGGTTG
TTCAAGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGTTCAAGAAGTTCTGTTT
10 CACCGTGCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGTACGATTTGAAGGAGG
AGGCGGGGCGAGGCTGGCCCCGATCCTAGTCATGCGCTACCGCAACCTGATCGAGGGC
GAAGCATCCGCCGGTTCCTAATGTACGGAGCAGATGCTAGGGCAAATTGCCCTAGC
AGGGGAAAAAGGTGCGAAAGGTCTCTTTTCTGTGGATAGCACGTACATTGGGAACC
CAAAGCCGTACATTGGGAACCGGAACCCGTACATTGGGAACCCAAAGCCGTACATT
15 GGGAACCGGTCACACATGTAAGTGACTGATATAAAAGAGAAAAAAGGCGATTTTTTC
CGCCTAAAACTCTTTAAAACTTATTAAACTCTTAAACCCGCCTGGCCTGTGCAT
AACTGTCTGGCCAGCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCCCTTCGGTCG
CTGCGCTCCCTACGCCCCGCGCTTCGCGTCGGCCTATCGCGGCCGCTGGCCGCTC
AAAAATGGCTGGCCTACGGCCAGGCAATCTACCAGGGCGCGGACAAGCCGCGCCGT
20 CGCCACTCGACCGCCGGCGCCACATCAAGGCACCCTGCCTCGCGCGTTTCGGTGA
TGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGT
AAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGG
TGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTT
AACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAAT
25 ACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCTGCTCGGCTGCGGCGAGCGGTATCAGCTCACTCA
AAGGCGGTAAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGG
30 TGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC
CTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
TAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTACGCCCCGACCG
CTGCGCCTTATCCGGTAACCTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT

CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAG
5 ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGCATTCTAGGT
ACTAAAACAATTTCATCCAGTAAATATAATATTTTATTTTCTCCCAATCAGGCTTG
ATCCCCAGTAAGTCAAAAAATAGCTCGACATACTGTTCTTCCCCGATATCCTCCCT
GATCGACCGGACGCAGAAGGCAATGTCATACCACTTGTCCGCCCTGCCGCTTCTCC
10 CAAGATCAATAAAGCCACTTACTTTGCCATCTTTCACAAAGATGTTGCTGTCTCCC
AGGTGCGCGTGGGAAAAGACAAGTTCCTCTTCGGGCTTTTCCGTCTTTAAAAAATC
ATACAGCTCGCGCGGATCTTTAAATGGAGTGTCTTCTTCCCAGTTTTTCGCAATCCA
CATCGGCCAGATCGTTATTTCAGTAAGTAATCCAATTCGGCTAAGCGGCTGTCTAAG
CTATTCGTATAGGGACAATCCGATATGTCGATGGAGTGAAAGAGCCTGATGCACTC
15 CGCATACAGCTCGATAATCTTTTCAGGGCTTTGTTTCATCTTCATACTCTTCCGAGC
AAAGGACGCCATCGGCCTCACTCATGAGCAGATTGCTCCAGCCATCATGCCGTTCA
AAGTGCAGGACCTTTGGAACAGGCAGCTTTCCTTCCAGCCATAGCATCATGTCCTT
TTCCCGTTCCACATCATAGGTGGTCCCTTTATACCGGCTGTCCGTCATTTTTTAAAT
ATAGGTTTTTCATTTTCTCCCACCAGCTTATATACCTTAGCAGGAGACATTCCTTCC
20 GTATCTTTTACGCAGCGGTATTTTTCGATCAGTTTTTTCAATTCCGGTGATATTCT
CATTTTAGCCATTTATTATTTCTTCTCTTTTCTACAGTATTTAAAGATAACCCCA
AGAAGCTAATTATAACAAGACGAACTCCAATTCAGTTCCTTGCAATTTCTAAAACC
TTAAATACCAGAAAACAGCTTTTTTCAAAGTTGTTTTCAAAGTTGGCGTATAACATA
GTATCGACGGAGCCGATTTTGAAACCGCGGTGATCACAGGCAGCAACGCTCTGTCA
25 TCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCCGGCAG
CTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCCTTA
CAACGGCTCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGTGGT
GATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGATATA
TTGTGGTGTAAACAAATTGACGCTTAGACAACCTTAATAACACATTGCGGACGTTTT
30 TAATGTACTGAATTAACGCCGAATTAATTCGGGGGATCTGGATTTTACTAGTGGAT
TTTGGTTTTTAGGAATTAGAAATTTTATTGATAGAAGTATTTTACAAATACAAATAC
ATACTAAGGGTTTCTTATATGCTCAACACATGAGCGAAACCCTATAGGAACCCTAA
TTCCCTTATCTGGGAACCTACTCACACATTATTATGGAGAACTCGAGCTTGTCGAT
CGACAGATCCGGTCGGCATCTACTCTATTTCTTTGCCCTCGGACGAGTGCTGGGGC

GTCGGTTTTCCACTATCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCG
CTTCTGCGGGCGATTTGTGTACGCCCCAGTCCCGGCTCCGGATCGGACGATTGC
GTCGCATCGACCCTGCGCCCAAGCTGCATCATCGAAATTGCCGTCAACCAAGCTCT
GATAGAGTTGGTCAAGACCAATGCGGAGCATATACGCCCCGAGTCGTGGCGATCCT
5 GCAAGCTCCGGATGCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAA
CCACGGCCTCCAGAAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCG
CCTCGCTCCAGTCAATGACCGCTGTTATGCGGCCATTGTCCGTCAGGACATTGTTG
GAGCCGAAATCCGCGTGCACGAGGTGCCGGACTTCGGGGCAGTCCTCGGCCCAAAG
CATCAGCTCATCGAGAGCCTGCGCGACGGACGCACTGACGGTGTGCTCCATCACAG
10 TTTGCCAGTGATACACATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTA
GTGTATTGACCGATTCCCTTGCGGTCCGAATGGGCCGAACCCGCTCGTCTGGCTAAG
ATCGGCCGCAGCGATCGCATCCATAGCCTCCGCGACCGGTTGTAGAACAGCGGGCA
GTTCCGTTTTAGGCAGGTCTTGCAACGTGACACCCTGTGCACGGCGGGAGATGCAA
TAGGTCAGGCTCTCGCTAAACTCCCCAATGTCAAGCACTTCCGGAATCGGGAGCGC
15 GGCCGATGCAAAGTGCCGATAAACATAACGATCTTTGTAGAAACCATCGGCGCAGC
TATTTACCCGCAGGACATATCCACGCCCTCCTACATCGAAGCTGAAAGCACGAGAT
TCTTCGCCCTCCGAGAGCTGCATCAGGTCCGAGACGCTGTCGAACTTTTCGATCAG
AAACTTCTCGACAGACGTGCGGGTGAGTTCAGGCTTTTTTCATATCTCATTGCCCCC
CGGGATCTGCGAAAGCTCGAGAGAGATAGATTTGTAGAGAGAGACTGGTGATTTCA
20 GCGTGTCTCTCCAAATGAAATGAACTTCCTTATATAGAGGAAGGTCTTGCGAAGG
ATAGTGGGATTGTGCGTCATCCCTTACGTCAGTGGAGATATCACATCAATCCACTT
GCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGG
GGTCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTGAACGATAGCCTTTCTTTT
ATCGCAATGATGGCATTGTAGGTGCCACCTTCCTTTTCTACTGTCCTTTTGATGA
25 AGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTGT
TGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGA
GTAGACGAGAGTGTGCTGCTCCACCATGTTATCACATCAATCCACTTGCTTTGAAG
ACGTGGTTGGAACGTCTTCTTTTTTCCACGATGCTCCTCGTGGGTGGGGGTCCATCT
TTGGGACCACTGTCGGCAGAGGCATCTTGAACGATAGCCTTTCTTTTATCGCAATG
30 ATGGCATTGTAGGTGCCACCTTCCTTTTCTACTGTCCTTTTGATGAAGTGACAGA
TAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTC
TCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGAGTAGACGAG
AGTGTGCTGCTCCACCATGTTGGCAAGCTGCTCTAGCCAATACGCAAACCGCCTCT
CCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA

AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCC
CAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATA
ACAATTTACACAGGAAACAGCTATGACCATGATTACGAATTCCCGATCTAGTAAC
ATAGATGACACCGCGCGCGATAATTTATCCTAGTTTGCGCGCTATATTTTGTTTTC
5 TATCGCGTATTAAATGTATAATTGCGGGACTCTAATCATAAAAACCCATCTCATAA
ATAACGTCATGCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAAT
TATATGATAATCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAACTTTATTG
CCAAATGTTTGAACGATCGGGGAAATTCGAGCTCTTATGCTTTTCTAGTCAATAAT
ATGGTAACATTGATGAACTTGGTTTTCTCCTTGGACATACGTTTCAGCCAAGATTTG
10 CTGGTAGCGGCAAAAACCTCTTCAATGATAGCTTCACCAAAGTGGCTAATCAGCA
TAGGTTCTGCCACAGCCCTCATGCACTGTGCCACATTGTATCCACCATCACTAAGT
GATTCAGACCTTTTCAGATTCAAAGTCTAGAGCATTCCACTCATCGAGAGCATTCCA
ATTCACCTCAGAAACCTCTAGACGATTGATGGCGAATGATCCTTCTTTAAGAACTT
CCAATTTCACTTCAGATGGGGATGGAGTGTATTGAGGGATGTTAAAAGTATCTAAT
15 TGCTCTTCTCTTATGATTCCCTGCAAGACCATATCACTAAGAGCAGTAGCCAAAAG
CTCCCAAATGTAGCAACCATCCTTGCTAGATGGATCATCGCTTCTTCTTCCCAAAA
ATGTGAGAACCATGCGTCCTCCTTCAACCAATTCCTCTGCACGACACTTGAGAAAA
AGAGAGAAATCTCTCTGAAATTGCTCATAGTAAGCTCTCACAACATTTTTTGGGGCT
CGTACTGCCTATGTAAACATTGCCCTTGTTGTTGTCTACACCCTCAGGAACCTTAG
20 ATAGCCATTGAAGGCTGTATGAGGAATGGACAAAATGAAGACTTTGATATGGGAAA
ACCTTTCATAGAAAGAACCGGGAACACCCGAGAAGTAGCATGGACCGATCCCACT
TTCCATTTTCATCACACAATTTCTCTTTGAAGCTGTCAAGAGACTTGAAGATGTTGT
TGAAGTCATTCCCAGGGAGATCATTCAGAAAGACTTTGTATTCTGGAGATTTGTGG
TTCAGCTCTCGGCAAAGCTTCTCCACAATTTTTATGAATTCAGACACAACAAATAA
25 AGTGTTTGGTCCAGAAGAGCAACCCAAATCTGCAATTGCCAAGCTTCTGGGGTGTG
TGCTGCAATAGAGGCTTCTTATGGCTTCCTCTCTTATGGGCTTTGTCAAACAAATC
ACCTTTTGCTGAACAAGGGAGTTGTTTGCATAGCTTGCGTCTCCAACGCCACCGTT
CATGTGGAGTACCTGTGCTACTTCCATGGATCCTCTAGAGTCCCCCGTGTTCTCTC
CAAATGAAATGAACTTCCTTATATAGAGGAAGGGTCTTGCGAAGGATAGTGGGATT
30 GTGCGTCATCCCTTACGTCAGTGGAGATATCACATCAATCCACTTGCTTTGAAGAC
GTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGGGTCCATCTTT
GGGACCACTGTGCGCAGAGGCATCTTCAACGATGGCCTTTCCTTTATCGCAATGAT
GGCATTGTAGGAGCCACCTTCCTTTTCCACTATCTTCACAATAAAGTGACAGATA
GCTGGGCAATGGAATCCGAGGAGGTTTCCGGATATTACCCTTTGTTGAAAAGTCTC

AATTGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTTTGGAGTAGACAAGTG
 TGTCGTGCTCCACCATGTTGACGAAGATTTTCTTCTTGTCATTGAGTCGTAAGAGA
 CTCTGTATGAACTGTTCCGAGTCTTTACGGCGAGTTCTGTTAGGTCCTCTATTTG
 AATCTTTGACTCCATGGCCTTTGATTCAGTGGGAACCTTTTTAGAGACTCCAA
 5 TCTCTATTACTTGCCCTTGGTTTGTGAAGCAAGCCTTGAATCGTCCATACTGGAATA
 GTACTTCTGATCTTGAGAAATATATCTTTCTCTGTGTTCTTGATGCAGTTAGTCCT
 GAATCTTTTGACTGCATCTTTAACCTTCTTGGGAAGGTATTTGATTTCTGAGAT
 TATTGCTCGGGTAGATCGTCTTGATGAGACCTGCTGCGTAAGCCTCTCTAACCATC
 TGTGGGTTAGCATTCTTTCTGAAATTGAAAAGGCTAATCTGGGGACCTGCAGGCAT
 10 GCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGT
 TACCCAACCTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCG
 AAGAGGCCCCGACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGC
 TAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTTCCCGCCTTCAGTTTAGCTTCAT
 GGAGTCAAAGATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGACTGGCGAAC
 15 AGTTCATACAGAGTCTCTTACGACTCAATGACAAGAAGAAAATCTTCGTCAACATG
 GTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGTCTCAGAAGA
 CCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACCTCCTCGGAT
 TCCATTGCCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAGGAAGGTGGC
 TCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGATGCCTCTGC
 20 CGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGGAAAAAGAAG
 ACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCCACTGACGTA
 AGGGATGACGCACAATCCCCTATCCTTCGCAAGACCCTTCCTCTATATAAGGAAG
 TTCATTTTCATTTGGAGAGAACACGGGGGACTCTTGAC (SEQ ID NO: 3)

The following is the sequence of binary vector pJL-OFP-35S::*GmSABP2-1*,
 25 showing reverse complemented sequence of *GmSABP2-1* in bold.

TCATGGCTCTTTCAAAGCAAAGTGGGGTCAAAGATGTAATGAACACCGAGCTTCAT
 ATGGACGGGATCGTCAATGGACACCCCTTTGAGATAAAAGGAAAAGGAAAGGGAAA
 CCCGTACAAGGGTGTGCAGACCATGAACTTACAGTCATTAAGGGTGCGCCTTTGC
 CATTTTCTATTGACATTTTGCTGCCTCAACACATGTATGGAAGCAAGCCATTTATT
 30 AAGTATCCTGAGAGTATCCCAGACTACATCAAGTTGTCATTTCCCGAGGGAATCAC
 ATGGGAAAGGTCCATGACCTTTGAAGATGGTGCAGTGTGCACTGTCTCTAACGACT
 CCAGGCTCGATGGCGACTCTTTCATCTACGAAGTCAGGTTTCTTGGCGTGAACCTT
 CCCCAGATGGACCTGTTATGCAGAAGAAGACGCGAGGCTGGGACCCGTCCACAGA

GAGACTGTATGAGTGTGGTGGGTGGCAGAGAGGAGACGTCCACATGGCCTTGAAGT
TGGAGAACGGTGGCCATTATACGTGCGACTTCAAACTACTTACAAATCAAAGAAG
GGCTTGAAGGTGCCACCGTATCACTTCGTTGACCACAACTAGATTTACTGAGCCA
CAACACCGATGGTGCTACCTTTGAAGAGTTTGAACAACGAGAAATTGCACATGCAC
5 ATCTTTCTAACTTACCGGTAGCCCATCACCATCACCATCACTAAGATCGTTCAAAC
ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGA
CGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACTAGGATAAATTATCGCGCGCGGTGTC
10 ATCTATGTTACTAGATCCCATGGCTACTACTAAGCATTTGGCTCTTGCCATCCTTG
TCCTCCTTAGCATTTGGTATGACCACCAGTGCAAGAACCCTCCTAGATCTGAGGGTA
AATTTCTAGTTTTTCTCCTTCATTTTCTTGTTAGGACCCTTTTCTCTTTTTATTT
TTTTGAGCTTTGATCTTTCTTTAACTGATCTATTTTTTAATTGATTGGTTATGGT
GTAAATATTACATAGCTTTAACTGATAATCTGATTACTTTATTTCTGTGTCTATG
15 ATGATGATGATAGTTACAGAACCGACGAACTAGTCTGTACCCGATCAACACCGAGA
CCCGTGGCGTCTTCGACCTCAATGGCGTCTGGAAC TTCAAGCTGGACTACGGGAAA
GGACTGGAAGAGAAGTGGTACGAAAGCAAGCTGACCGACACTATTAGTATGGCCGT
CCCAAGCAGTTACAATGACATTGGCGTGACCAAGGAAATCCGCAACCATATCGGAT
ATGTCTGGTACGAACGTGAGTTCACGGTGCCGGCCTATCTGAAGGATCAGCGTATC
20 GTGCTCCGCTTCGGCTCTGCAACTCACAAAGCAATTGTCTATGTCAATGGTGAGCT
GGTCGTGGAGCACAAGGGCGGATTCCTGCCATTCTGAAGCGGAAATCAACAACTCGC
TGCGTGATGGCATGAATCGCGTCACCGTCGCCGTGGACAACATCCTCGACGATAGC
ACCTCCCGGTGGGGCTGTACAGCGAGCGCCACGAAGAGGGCCTCGGAAAAGTCAT
TCGTAACAAGCCGAAC TTCTCAACTATGCAGGCCTGCACCGTCCGGTGA
25 AAATCTACACGACCCCGTTTACGTACGTGAGGACATCTCGGTTGTGACCGACTTC
AATGGCCCAACCGGGACTGTGACCTATACGGTGGACTTTCAAGGCAAAGCCGAGAC
CGTGAAAGTGTGCGTCGTGGATGAGGAAGGCAAAGTGGTCGCAAGCACCGAGGGCC
TGAGCGGTAACGTGGAGATTCCGAATGTCATCCTCTGGGAACCACTGAACACGTAT
CTCTACCAGATCAAAGTGGAAGTGGTGAACGACGGACTGACCATCGATGTCTATGA
30 AGAGCCGTTTCGGCGTGCGGACCGTGGAAGTCAACGACGGCAAGTTCCTCATCAACA
ACAAACCGTTCTACTTCAAGGGCTTTGGCAAACATGAGGACACTCCTATCAACGGC
CGTGGCTTTAACGAAGCGAGCAATGTGATGGATTTCAATATCCTCAAATGGATCGG
CGCCAACAGCTTCCGGACCGCACACTATCCGTACTCTGAAGAGTTGATGCGTCTTG
CGGATCGCGAGGGTCTGGTCGTGATCGACGAGACTCCGGCAGTTGGCGTGACCTC

AACTTCATGGCCACCACGGGACTCGGCGAAGGCAGCGAGCGCGTCAGTACCTGGGA
GAAGATTCGGACGTTTGAGCACCATCAAGACGTTCTCCGTGAACTGGTGTCTCGTG
ACAAGAACCATCCAAGCGTCGTGATGTGGAGCATCGCCAACGAGGCGGCGACTGAG
GAAGAGGGCGCGTACGAGTACTTCAAGCCGTTGGTGGAGCTGACCAAGGAACTCGA
5 CCCACAGAAGCGTCCGGTCACGATCGTGCTGTTTGTGATGGCTACCCCGGAGACGG
ACAAAGTCGCCGAACTGATTGACGTCATCGCGCTCAATCGCTATAACGGATGGTAC
TTCGATGGCGGTGATCTCGAAGCGGCCAAAGTCCATCTCCGCCAGGAATTTACGC
GTGGAACAAGCGTTGCCCAGGAAAGCCGATCATGATCACTGAGTACGGCGCAGACA
CCGTTGCGGGCTTTACGACATTGATCCAGTGATGTTACCGAGGAATATCAAGTC
10 GAGTACTACCAGGCGAACCACGTCGTGTTGATGAGTTTGAGAACTTCGTGGGTGA
GCAAGCGTGGAACCTTCGCGGACTTCGCGACCTCTCAGGGCGTGATGCGCGTCCAAG
GAAACAAGAAGGGCGTGTTCACTCGTGACCGCAAGCCGAAGCTCGCCGCGCACGTC
TTTCGCGAGCGCTGGACCAACATTCCAGATTTCCGGCTACAAGAACGCTAGCCATCA
CCATCACCATCACGTGTGAATTGGTGACCAGCTCGAATTTCCCGATCGTTCAAAC
15 ATTTGGCAATAAAGTTTCTTAAGATTGAATCCTGTTGCCGGTCTTGCGATGATTAT
CATATAATTTCTGTTGAATTACGTTAAGCATGTAATAATTAACATGTAATGCATGA
CGTTATTTATGAGATGGGTTTTTATGATTAGAGTCCCGCAATTATACATTTAATAC
GCGATAGAAAACAAAATATAGCGCGCAAACCTAGGATAAATTATCGCGCGCGGTGTC
ATCTATGTTACTAGATCGGGAATTAACTATCAGTGTTTGACAGGATATATTGGCG
20 GGTAACCTAAGAGAAAAGAGCGTTTATTAGAATAACGGATATTTAAAAGGGCGTG
AAAAGGTTTATCCGTTTCGTCCATTTGTATGTGCATGCCAACCACAGGGTTCCCTC
GGGATCAAAGTACTTTGATCCAACCCCTCCGCTGCTATAGTGCACTCGGCTTCTGA
CGTTCAGTGCAGCCGTCTTCTGAAAACGACATGTGCGACAAGTCCTAAGTTACGCG
ACAGGCTGCCGCCCTGCCCTTTTCTTGCGGTTTTCTTGTCGCGTGTTTTAGTCGCA
25 TAAAGTAGAATACTTGCGACTAGAACCGGAGACATTACGCCATGAACAAGAGCGCC
GCCGCTGGCCTGCTGGGCTATGCCCCGCGTCAGCACCGACGACCAGGACTTGACCAA
CCAACGGGGCCGAACCTGCACGCGGCGGCTGCACCAAGCTGTTTTCCGAGAAGATCA
CCGGCACCAGGCGCGACCGCCCGGAGCTGGCCAGGATGCTTGACCACCTACGCCCT
GGCGACGTTGTGACAGTGACCAGGCTAGACCGCCTGGCCCGCAGCACCCGCGACCT
30 ACTGGACATTGCCGAGCGCATCCAGGAGGCCGGCGCGGGCCTGCGTAGCCTGGCAG
AGCCGTGGGCCGACACCACCACGCCGGCCGGCCGCATGGTGTGACCGTGTTTCGCC
GGCATTGCCGAGTTCGAGCGTTCCCTAATCATCGACCGCACCCGGAGCGGGCGCGA
GGCCGCCAAGGCCCGAGGCGTGAAGTTTGGCCCCCGCCCTACCCCTACCCCGGCAC
AGATCGCGCACGCCCGCGAGCTGATCGACCAGGAAGGCCGCACCGTGAAAGAGGCG

GCTGCACTGCTTGGCGTGCATCGCTCGACCCTGTACCGCGCACTTGAGCGCAGCGA
GGAAGTGACGCCCACCGAGGCCAGGCGGCGCGGTGCCTTCCGTGAGGACGCATTGA
CCGAGGCCGACGCCCTGGCGGCCGCCGAGAATGAACGCCAAGAGGAACAAGCATGA
AACCGCACCCAGGACGGCCAGGACGAACCGTTTTTCATTACCGAAGAGATCGAGGCG
5 GAGATGATCGCGGCCGGGTACGTGTTTCGAGCCGCCCGCGCACGTCTCAACCGTGCG
GCTGCATGAAATCCTGGCCGGTTTTGTCTGATGCCAAGCTGGCGGCCCTGGCCGGCCA
GCTTGGCCGCTGAAGAAACCGAGCGCCGCCGTCTAAAAAGGTGATGTGTATTTGAG
TAAACAGCTTGCGTCATGCGGTTCGCTGCGTATATGATGCGATGAGTAAATAACA
AATACGCAAGGGGAACGCATGAAGGTTATCGCTGTACTTAACCAGAAAGGCGGGTC
10 AGGCAAGACGACCATCGCAACCCATCTAGCCCGCGCCCTGCAACTCGCCGGGGCCG
ATGTTCTGTTAGTCGATTCCGATCCCCAGGGCAGTGCCCGCGATTGGGCGGCCGTG
CGGGAAGATCAACCGCTAACCGTTGTCGGCATCGACCGCCCGACGATTGACCGCGA
CGTGAAGGCCATCGGCCGGCGCGACTTCGTAGTGATCGACGGAGCGCCCCAGGCGG
CGGACTTGGCTGTGTCCGCGATCAAGGCAGCCGACTTCGTGCTGATTCCGGTGCAG
15 CCAAGCCCTTACGACATATGGGCCACCGCCGACCTGGTGGAGCTGGTTAAGCAGCG
CATTGAGGTCACGGATGGAAGGCTACAAGCGGCCCTTTGTCGTGTCGCGGGCGATCA
AAGGCACGCGCATCGGCGGTGAGGTTGCCGAGGCGCTGGCCGGGTACGAGCTGCCC
ATTCTTGAGTCCCGTATCACGCAGCGCGTGAGCTACCCAGGCACTGCCGCCGCCGG
CACAACCGTTCTTGAATCAGAACCCGAGGGCGACGCTGCCCCGCGAGGTCCAGGCGC
20 TGGCCGCTGAAATTAAATCAAACTCATTGAGTTAATGAGGTAAAGAGAAAATGA
GCAAAAGCACAAACACGCTAAGTGCCGGCCGTCCGAGCGCACGCAGCAGCAAGGCT
GCAACGTTGGCCAGCCTGGCAGACACGCCAGCCATGAAGCGGGTCAACTTTTCAGTT
GCCGGCGGAGGATCACACCAAGCTGAAGATGTACGCGGTACGCCAAGGCAAGACCA
TTACCGAGCTGCTATCTGAATACATCGCGCAGCTACCAGAGTAAATGAGCAAATGA
25 ATAAATGAGTAGATGAATTTTAGCGGCTAAAGGAGGCGGCATGGAAAATCAAGAAC
AACCAGGCACCGACGCCGTGGAATGCCCCATGTGTGGAGGAACGGGCGGTTGGCCA
GGCGTAAGCGGCTGGGTTGTCTGCCGGCCCTGCAATGGCACTGGAACCCCCAAGCC
CGAGGAATCGGCGTGACGGTCGCAAACCATCCGGCCCCGTACAAATCGGCGCGGGCG
CTGGGTGATGACCTGGTGGAGAAGTTGAAGGCCGCGCAGGCCGCCAGCGGCAACG
30 CATCGAGGCAGAAGCACGCCCCGGTGAATCGTGGCAAGCGGCCGCTGATCGAATCC
GCAAAGAATCCCGGCAACCGCCGGCAGCCGGTGCGCCGTCGATTAGGAAGCCGCC
AAGGGCGACGAGCAACCAGATTTTTTCGTTCCGATGCTCTATGACGTGGGCACCCG
CGATAGTCGCAGCATCATGGACGTGGCCGTTTTCCGTCTGTCTGAAGCGTGACCGAC
GAGCTGGCGAGGTGATCCGCTACGAGCTTCCAGACGGGCACGTAGAGGTTTCCGCA

GGGCCGGCCGGCATGGCCAGTGTGTGGGATTACGACCTGGTACTGATGGCGGTTTC
CCATCTAACCGAATCCATGAACCGATAACCGGAAGGGAAGGGAGACAAGCCCCGGCC
GCGTGTTCGGTCCACACGTTGCGGACGTACTCAAGTTCTGCCGGCGAGCCGATGGC
GGAAAGCAGAAAGACGACCTGGTAGAAACCTGCATTCGGTTAAACACCACGCACGT
5 TGCCATGCAGCGTACGAAGAAGGCCAAGAACGGCCGCCTGGTGACGGTATCCGAGG
GTGAAGCCTTGATTAGCCGCTACAAGATCGTAAAGAGCGAAACCGGGCGGCCGGAG
TACATCGAGATCGAGCTAGCTGATTGGATGTACCGCGAGATCACAGAAGGCAAGAA
CCCGGACGTGCTGACGGTTCACCCCGATTACTTTTTTGATCGATCCCGGCATCGGCC
GTTTTCTCTACCGCCTGGCACGCCGCGCCGAGGCAAGGCAGAAGCCAGATGGTTG
10 TTCAAGACGATCTACGAACGCAGTGGCAGCGCCGGAGAGTTCAAGAAGTTCTGTTT
CACCGTGCGCAAGCTGATCGGGTCAAATGACCTGCCGGAGTACGATTTGAAGGAGG
AGGCGGGGCAGGCTGGCCCGATCCTAGTCATGCGCTACCGCAACCTGATCGAGGGC
GAAGCATCCGCCGGTTCCTAATGTACGGAGCAGATGCTAGGGCAAATTGCCCTAGC
AGGGGAAAAAGGTCGAAAAGGTCTCTTTCTGTGGATAGCACGTACATTGGGAACC
15 CAAAGCCGTACATTGGGAACCGGAACCCGTACATTGGGAACCCAAAGCCGTACATT
GGGAACCGGTCACACATGTAAGTGACTGATATAAAAGAGAAAAAAGGCGATTTTTTC
CGCCTAAAACTCTTTAAAACTTATTAAAACTCTTAAACCCGCCTGGCCTGTGCAT
AACTGTCTGGCCAGCGCACAGCCGAAGAGCTGCAAAAAGCGCCTACCCTTCGGTTCG
CTGCGCTCCCTACGCCCCGCGCTTCGCGTCGGCCTATCGCGGCCGCTGGCCGCTC
20 AAAAATGGCTGGCCTACGGCCAGGCAATCTACCAGGGCGCGGACAAGCCGCGCCGT
CGCCACTCGACCGCCGGCGCCACATCAAGGCACCCTGCCTCGCGCGTTTCGGTGA
TGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGT
AAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTGGCGGG
TGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTT
25 AACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAAT
ACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCA
AAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTG
AGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTT
30 TCCATAGGCTCCGCCCCCTGACGAGCATCACAAAATCGACGCTCAAGTCAGAGG
TGCGGAAACCCGACAGGACTATAAAGATAACAGGCGTTTCCCCCTGGAAGCTCCCT
CGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCC
CTTCGGGAAGCGTGGCGCTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTG
TAGGTCGTTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCG

CTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTAT
CGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGT
GCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATT
TGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTT
5 GATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTTGTTTGCAAGCAGCAG
ATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTC
TGACGCTCAGTGGAACGAAAACCTCACGTTAAGGGATTTTGGTCATGCATTCTAGGT
ACTAAAACAATTCATCCAGTAAATATAATATTTTATTTTCTCCCAATCAGGCTTG
ATCCCCAGTAAGTCAAAAAATAGCTCGACATACTGTTCTTCCCGATATCCTCCCT
10 GATCGACCGGACGCAGAAGGCAATGTCATACCACCTTGTCGCCCTGCCGCTTCTCC
CAAGATCAATAAAGCCACTTACTTTGCCATCTTTCACAAAGATGTTGCTGTCTCCC
AGGTCGCCGTGGGAAAAGACAAGTTCCTCTTCCGGCTTTTCCGTCTTTAAAAAATC
ATACAGCTCGCGCGGATCTTTAAATGGAGTGTCTTCTTCCAGTTTTTCGCAATCCA
CATCGGCCAGATCGTTATTCAGTAAGTAATCCAATTCGGCTAAGCGGCTGTCTAAG
15 CTATTCGTATAGGGACAATCCGATATGTCGATGGAGTGAAAGAGCCTGATGCACTC
CGCATACAGCTCGATAATCTTTTCAGGGCTTTGTTTCATCTTCATACTCTTCCGAGC
AAAGGACGCCATCGGCCTCACTCATGAGCAGATTGCTCCAGCCATCATGCCGTTCA
AAGTGCAGGACCTTTGGAACAGGCAGCTTTCCTTCCAGCCATAGCATCATGTCCTT
TTCCCGTTCCACATCATAGGTGGTCCCTTTATACCGGCTGTCCGTCATTTTTTAAAT
20 ATAGGTTTTTCATTTTCTCCCACCAGCTTATATACCTTAGCAGGAGACATTCCTTCC
GTATCTTTTACGCAGCGGTATTTTTTCGATCAGTTTTTTTCAATTCCGGTGATATTCT
CATTTTAGCCATTTATTATTTCCTTCTCTTTTCTACAGTATTTAAAGATAACCCCA
AGAAGCTAATTATAACAAGACGAACTCCAATTCAGTGTTCCTTGCATTCTAAAACC
TTAAATACCAGAAAACAGCTTTTTTCAAAGTTGTTTTCAAAGTTGGCGTATAACATA
25 GTATCGACGGAGCCGATTTTGAAACCGCGGTGATCACAGGCAGCAACGCTCTGTCA
TCGTTACAATCAACATGCTACCCTCCGCGAGATCATCCGTGTTTCAAACCCGGCAG
CTTAGTTGCCGTTCTTCCGAATAGCATCGGTAACATGAGCAAAGTCTGCCGCCTTA
CAACGGCTCTCCCGCTGACGCCGTCCCGGACTGATGGGCTGCCTGTATCGAGTGGT
GATTTTGTGCCGAGCTGCCGGTCGGGGAGCTGTTGGCTGGCTGGTGGCAGGATATA
30 TTGTGGTGTAACAAATTGACGCTTAGACAACCTAATAACACATTGCGGACGTTTT
TAATGTACTGAATTAACGCCGAATTAATTCGGGGGATCTGGATTTTAGTACTGGAT
TTTGGTTTTAGGAATTAGAAATTTTATTGATAGAAGTATTTTACAAATACAAATAC
ATACTAAGGGTTTTCTTATATGCTCAACACATGAGCGAAACCCTATAGGAACCCTAA
TTCCCTTATCTGGGAACCTACTCACACATTATTATGGAGAACTCGAGCTTGTGCGAT

CGACAGATCCGGTCGGCATCTACTCTATTTCTTTGCCCTCGGACGAGTGCTGGGGC
GTCGGTTTCCACTATCGGCGAGTACTTCTACACAGCCATCGGTCCAGACGGCCGCG
CTTCTGCGGGCGATTTGTGTACGCCCAGAGTCCCGGCTCCGGATCGGACGATTGC
GTCGCATCGACCCTGCGCCCAAGCTGCATCATCGAAATTGCCGTCAACCAAGCTCT
5 GATAGAGTTGGTCAAGACCAATGCGGAGCATATACGCCCAGAGTCGTGGCGATCCT
GCAAGCTCCGGATGCCTCCGCTCGAAGTAGCGCGTCTGCTGCTCCATACAAGCCAA
CCACGGCCTCCAGAAGAAGATGTTGGCGACCTCGTATTGGGAATCCCCGAACATCG
CCTCGCTCCAGTCAATGACCGCTGTTATGCGGGCCATTGTCCGTCAGGACATTGTTG
GAGCCGAAATCCGCGTGACAGAGGTGCCGGACTTCGGGGCAGTCCTCGGCCCAAAG
10 CATCAGCTCATCGAGAGCCTGCGCGACGGACGCACTGACGGTGTCGTCCATCACAG
TTTGCCAGTGATACACATGGGGATCAGCAATCGCGCATATGAAATCACGCCATGTA
GTGTATTGACCGATTCCCTTGGCGTCCGAATGGGCCGAACCCGCTCGTCTGGCTAAG
ATCGGCCGCAGCGATCGCATCCATAGCCTCCGCGACCGGTTGTAGAACAGCGGGCA
GTTCCGTTTTAGGCAGGTCTTGCAACGTGACACCCTGTGCACGGCGGGAGATGCAA
15 TAGGTCAGGCTCTCGCTAAACTCCCCAATGTCAAGCACTTCCGGAATCGGGAGCGC
GGCCGATGCAAAGTGCCGATAAACATAACGATCTTTGTAGAAACCATCGGCGCAGC
TATTTACCCGCAGGACATATCCACGCCCTCCTACATCGAAGCTGAAAGCACGAGAT
TCTTCGCCCTCCGAGAGCTGCATCAGGTCGGAGACGCTGTCGAACTTTTCGATCAG
AAACTTCTCGACAGACGTGCGGGTGAGTTCAGGCTTTTTTCATATCTCATTTGCCCCC
20 CGGGATCTGCGAAAGCTCGAGAGAGATAGATTTGTAGAGAGAGACTGGTGATTTCA
GCGTGTCCTCTCCAAATGAAATGAACTTCCTTATATAGAGGAAGGTCTTGCGAAGG
ATAGTGGGATTGTGCGTCATCCCTTACGTCAGTGGAGATATCACATCAATCCACTT
GCTTTGAAGACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGG
GGTCCATCTTTGGGACCACTGTGCGGCAGAGGCATCTTGAACGATAGCCTTTCCCTTT
25 ATCGCAATGATGGCATTGTAGGTGCCACCTTCCTTTTTCTACTGTCCTTTTGATGA
AGTGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTGT
TGAAAAGTCTCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGGA
GTAGACGAGAGTGTCGTGCTCCACCATGTTATCACATCAATCCACTTGCTTTGAAG
ACGTGGTTGGAACGTCTTCTTTTTCCACGATGCTCCTCGTGGGTGGGGGTCCATCT
30 TTGGGACCACTGTGCGCAGAGGCATCTTGAACGATAGCCTTTCCCTTTATCGCAATG
ATGGCATTGTAGGTGCCACCTTCCTTTTTCTACTGTCCTTTTGATGAAGTGACAGA
TAGCTGGGCAATGGAATCCGAGGAGGTTTCCCGATATTACCCTTTGTTGAAAAGTC
TCAATAGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTCTTGAGTAGACGAG
AGTGTCGTGCTCCACCATGTTGGCAAGCTGCTCTAGCCAATACGCAAACCGCCTCT

CCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGA
AAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAGGCACCC
CAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATA
ACAATTTACACACAGGAAACAGCTATGACCATGATTACGAATTCCCGATCTAGTAAC
5 ATAGATGACACCGCGCGCGATAATTTATCCTAGTTTGC GCGCTATATTTTGT TTTTC
TATCGCGTATTAAATGTATAATTGCGGGACTCTAATCATAAAAACCCATCTCATAA
ATAACGTCATGCATTACATGTTAATTATTACATGCTTAACGTAATTCAACAGAAAT
TATATGATAATCATCGCAAGACCGGCAACAGGATTCAATCTTAAGAACTTTATTG
CCAAATGTTTGAACGATCGGGGAAATTCGAGCTCT**CATGCATATTTAGTCGCTATC**
10 **TGCTGGAGGGAATCGAATAGTTCTTGTGGCTTGCAAAGCATAACCATATGATCTGC**
GTCTTTGACCTCTACAACGTCATTGAACCCAACATTTTGGATCATCAAGAGCTGAT
ATTCCAATGGAATTCCAAGGTCCTCCGTGCAAACAATAAAGGCACGTGGAAC TGAC
CCATATCCCTCTTTGGAGAAGTTCTTTTGT TTAGTCAAGTCTTCCATGAAGAGCGA
TGATGGCCTTGCTAAAGTCTTGGCCAATTCCAAGTCCTCAATAGGGGACAGTTGGT
15 **AGAGCTTGTCGGACAAGAAGTTGGGGCCGAAGAACATTGATGTTTTGTTTCCGCTT**
GGAGCAAATTCAGTGTCTAACCATGCAGCTAACGGGGTCCTCTCATTGTACTTTTC
CAAGACATAAGATGGGTGGTGTTCAGTGTCTGGAGCAAAAGCTGTTAAGAAAACAC
CAACTGCTACCTTTTCTGGGAATTTCTCCATTGCAAGTGCTATGTT CAGCCCTCCA
AGGCTGTGACCAACTAGAACTAACTTCTCATTTGAGGGAATTGTGGCCATTAGCTG
20 **CAACAAAGGCGCAGAATACTCTGAGAAAGTATCAACATCTTCAATTTTCTTCATGT**
TGGTTCCAGAAGCTGCAAGGTCAAGTACTGTGACCTTATGGCCTGCAGATTCCAAG
CGTGGCTTGAGCTTATACCAACACCAAGCTCCATGGCATGCCCCATGCACCAGAAC
ATAGTGCTTCCTATCCATAACAATTTTGTGAACCCATGGATCCTCTAGAGTCCCCCG
TGTTCTCTCCAAATGAAATGAACTTCCTTATATAGAGGAAGGGTCTTGCGAAGGAT
25 **AGTGGGATTGTGCGTCATCCCTTACGTCAGTGGAGATATCACATCAATCCACTTGC**
TTTGAAGACGTGGTTGGAACGTCTTCTTTTCCACGATGCTCCTCGTGGGTGGGGG
TCCATCTTTGGGACCACTGTCGGCAGAGGCATCTTCAACGATGGCCTTTCCTTTAT
CGCAATGATGGCATTGTAGGAGCCACCTTCCTTTTCCACTATCTTCACAATAAAG
TGACAGATAGCTGGGCAATGGAATCCGAGGAGGTTTCCGGATATTACCCTTTGTTG
30 **AAAAGTCTCAATTGCCCTTTGGTCTTCTGAGACTGTATCTTTGATATTTTGGAGT**
AGACAAGTGTGTCGTGCTCCACCATGTTGACGAAGATTTTCTTCTTGTCATTGAGT
CGTAAGAGACTCTGTATGAACTGTTGCGCCAGTCTTTACGGCGAGTTCTGTTAGGTC
CTCTATTTGAATCTTTGACTCCATGGCCTTTGATT CAGTGGGAACTACCTTTT TAG
AGACTCCAATCTCTATTACTTGCCCTTGGTTTGTGAAGCAAGCCTTGAATCGTCCAT

ACTGGAATAGTACTTCTGATCTTGAGAAATATATCTTTCTCTGTGTTCTTGATGCA
 GTTAGTCCTGAATCTTTTGACTGCATCTTTAACCTTCTTGGAAGGTATTTGATTT
 CCTGGAGATTATTGCTCGGGTAGATCGTCTTGATGAGACCTGCTGCGTAAGCCTCT
 CTAACCATCTGTGGGTAGCATTCTTTCTGAAATTGAAAAGGCTAATCTGGGGACC
 5 TGCAGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAA
 CCCTGGCGTTACCCAACCTTAATCGCCTTGACGACATCCCCCTTTCGCCAGCTGGC
 GTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAAT
 GGCGAATGCTAGAGCAGCTTGAGCTTGGATCAGATTGTCGTTTCCCGCCTTCAGTT
 TAGCTTCATGGAGTCAAAGATTCAAATAGAGGACCTAACAGAACTCGCCGTAAAGA
 10 CTGGCGAACAGTTCATACAGAGTCTCTTACGACTCAATGACAAGAAGAAAATCTTC
 GTCAACATGGTGGAGCACGACACACTTGTCTACTCCAAAAATATCAAAGATACAGT
 CTCAGAAGACCAAAGGGCAATTGAGACTTTTCAACAAAGGGTAATATCCGGAAACC
 TCCTCGGATTCCATTGCCAGCTATCTGTCACTTTATTGTGAAGATAGTGGAAAAG
 GAAGGTGGCTCCTACAAATGCCATCATTGCGATAAAGGAAAGGCCATCGTTGAAGA
 15 TGCCCTGCGGACAGTGGTCCCAAAGATGGACCCCCACCCACGAGGAGCATCGTGG
 AAAAAGAAGACGTTCCAACCACGTCTTCAAAGCAAGTGGATTGATGTGATATCTCC
 ACTGACGTAAGGGATGACGCACAATCCCACTATCCTTCGCAAGACCCTTCCTCTAT
 ATAAGGAAGTTCATTTTCAATTTGGAGAGAACACGGGGGACTCTTGAC (SEQ ID NO:
 4)
 20

C. *Transgenics*

An expression cassette including a disclosed a nucleic acid sequence
 encoding GmSAMT1 (such as set forth in SEQ ID NO: 1 or a degenerate variant
 thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2) or functional
 25 fragment thereof, operably linked to promoter and optionally other heterologous
 nucleic acids can be used to transform any plant, for example as a vector, such as a
 plasmid. In this manner, genetically modified plants, plant cells, plant tissue, seed,
 and the like can be obtained. Thus, methods of producing a plant resistant to a pest
 are disclosed. Such methods, include introducing into a plant *GmSAMT1* (such as set
 30 forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or *GmSABP2-1* such as
 set forth in SEQ ID NO: 2 or a degenerate variant thereof) or functional fragment
 thereof, operably linked to promoter and optionally other heterologous nucleic acids.
 Also disclosed are methods increasing pest resistance in a plant. Such methods,

include introducing into a plant *GmSAMT1* (such as set forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or *GmSABP2-1* such as set forth in SEQ ID NO: 2 or a degenerate variant thereof) or functional fragment thereof, operably linked to promoter and optionally other heterologous nucleic acids, thereby increasing pest
5 resistance in the plant. The plant can be transiently or stably transformed. Pest resistance, such as an increase in pest resistance can be determined relative to a relevant control plant, such as a plant that does not express a pest resistance polypeptide disclosed herein, a plant that has not been transformed with a nucleic acid encoding a pest resistance polypeptide as disclosed herein, a plant that is
10 transformed with an irrelevant nucleic acid, and the like. The control plant is generally matched for species, variety, age, and the like and is subjected to the same growing conditions, for example temperature, soil, sunlight, pH, water, and the like. The selection of a suitable control plant is routine for those skilled in the art.

Transformation protocols as well as protocols for introducing nucleotide
15 sequences into plants may vary depending on the type of plant or plant cell, for example, monocot or dicot, targeted for transformation. Suitable methods of introducing nucleotide sequences into plant cells and subsequent insertion into the plant genome include microinjection (Crossway *et al.*, *Biotechniques* 4:320-334, 1986), electroporation (Riggs *et al.*, *Proc. Natl. Acad. Sci. USA* 53:5602-5606, 1986), Agrobacterium-mediated transformation (U.S. Pat No. 5,563,055), direct
20 gene transfer (Paszkowski *et al.*, *EMBO J.* 3:2717-2722, 1984), and ballistic particle acceleration (see, for example, U.S. Patent No. 4,945,050; Tomes *et al.* "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," in *Plant Cell, Tissue, and Organ Culture: Fundamental Methods*, ed. Gamborg and Phillips
25 (Springer- Verlag, Berlin), 1995; and McCabe *et al.*, *Biotechnology* 5:923-926, 1988). Also see Weissinger *et al.*, *Ann. Rev. Genet.* 22:421-477, 1988; Sanford *et al.*, *Paniculate Science and Technology* 5:27-37, 1987; Christou *et al.*, *Plant Physiol* 57:671-674, 1988; McCabe *et al.*, *Bio/Technology* 5:923-926, 1988; Finer and McMullen, *In Vitro Cell Dev. Biol.* 27P:175-182, 1991; Singh *et al.*, *Theor. Appl*
30 *Genet.* 95:319-324, 1998; Datta *et al.*, *Biotechnology* 5:736-740, 1990; Klein *et al.*, *Proc. Natl. Acad. Sci. USA* 55:4305-4309, 1988; Klein *et al.*, *Biotechnology* 5:559-563, 1988; U.S. Patent Nos. 5,240,855, 5,322,783 and 5,324,646; Klein *et al.*, *Plant*

Physiol 97:440-444, 1988; Fromm *et al. Biotechnology* 5:833-839, 1990; Hooykaas-
 Van Slogteren *et al., Nature* 377:763-764, 1984; Bytebier *et al., Proc. Natl. Acad.*
Sci. USA 54:5345-5349, 1987; De Wet *et al. (1985) in The Experimental*
 Manipulation of Ovule Tissues, ed. Chapman *et al. (Longman, New York), pp. 197-*
 5 209; Kaeppler *et al., Plant Cell Reports* 9:415-418, 1990; Kaeppler *et al., Theor.*
Appl. Genet. 54:560-566, 1992; D'Halluin *et al., Plant Cell* 4:1495-1505 1992; Li *et*
al., Plant Cell Reports 72:250-255, 1993; Christou and Ford *Annals of Botany*
 75:407-413, 1995; Osjoda *et al., Nature Biotechnology* 74:745-750, 1996; and the
 like. "Introducing" in the context of a plant cell, plant tissue, plant part and/or plant
 10 means contacting a nucleic acid molecule with the plant cell, plant tissue, plant part,
 and/or plant in such a manner that the nucleic acid molecule gains access to the
 interior of the plant cell or a cell of the plant tissue, plant part or plant. Where more
 than one nucleic acid molecule is to be introduced, these nucleic acid molecules can
 be assembled as part of a single polynucleotide or nucleic acid construct, or as
 15 separate polynucleotide or nucleic acid constructs, and can be located on the same or
 different nucleic acid constructs. Accordingly, these polynucleotides can be
 introduced into plant cells in a single transformation event, in separate
 transformation events, or, for example as part of a breeding protocol.

The cells that have been transformed may be grown into plants in accordance
 20 with conventional ways. See, for example, McCormick *et al. Plant Cell Reports*
 5:81-84, 1986. These plants may then be grown, and either pollinated with the same
 transformed strain or different strains, and the resulting hybrid having expression of
 the desired phenotypic characteristic identified. Two or more generations may be
 grown to ensure that expression of the desired phenotypic characteristic is stably
 25 maintained and inherited and then seeds harvested to ensure expression of the
 desired phenotypic characteristic has been achieved.

The pest resistance genes disclosed herein, such as nucleic acid sequences
 encoding GmSAMT1 (such as set forth in SEQ ID NO: 1 or a degenerate variant
 thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2 or a degenerate
 30 variant thereof) or active variant and fragments thereof may be used for
 transformation of any plant species, including, but not limited to, monocots and
 dicots. Examples of plant species of interest include, but are not limited to, corn

(*Zea mays*), *Brassica* sp. (e.g., *B. napus*, *B. rapa*, *B. juncea*), particularly those *Brassica* species useful as sources of seed oil, alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor*, *Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense*, *Gossypium hirsutum*), sweet potato (*Ipomoea batatas*), cassava (*Manihot esculenta*), coffee (*Coffea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), oats, barley, vegetables, ornamentals, and conifers. In particular embodiments, the pest resistance genes disclosed herein, such as nucleic acid sequences encoding GmSAMT1 (such as set forth in SEQ ID NO: 1 or a degenerate variant thereof) and/or GmSABP2-1 such as set forth in SEQ ID NO: 2 or a degenerate variant thereof) or active variant and fragments thereof may be used for transformation of soybean (*Glycine max*).

Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Lathyrus* spp.), and members of the genus *Cucumis* such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*).

Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum.

Conifers that may be employed in practicing the present invention include, for example, pines such as loblolly pine (*Pinus taeda*), slash pine (*Pinus elliotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*); Douglas-fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga*

canadensis); Sitka spruce (*Picea glauca*); redwood (*Sequoia sempervirens*); true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*); and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow-cedar (*Chamaecyparis nootkatensis*), and Poplar and Eucalyptus.

5 In specific embodiments, plants of the present invention are crop plants (for example, corn, alfalfa, sunflower, Brassica, soybean, cotton, safflower, peanut, sorghum, wheat, millet, tobacco, etc.). In other embodiments, corn and soybean plants are optimal, and in yet other embodiments soybean plants are optimal. Other plants of interest include grain plants that provide seeds of interest, oilseed plants,
10 and leguminous plants. Seeds of interest include grain seeds, such as corn, wheat, barley, rice, sorghum, rye, etc. Oil-seed plants include cotton, soybean, safflower, sunflower, Brassica, maize, alfalfa, palm, coconut, etc. Leguminous plants include beans and peas. Beans include guar, locust bean, fenugreek, soybean, garden beans, cowpea, mungbean, lima bean, fava bean, lentils, chickpea, etc.

15 In some embodiments, the polynucleotides comprising disclosed pest resistance gene are engineered into a molecular stack. Thus, the various plants, plant cells and seeds disclosed herein can further comprise one or more traits of interest, and in more specific embodiments, the plant, plant part or plant cell is stacked with any combination of polynucleotide sequences of interest in order to create plants
20 with a desired combination of traits. As used herein, the term "stacked" includes having the multiple traits present in the same plant.

 These stacked combinations can be created by any method including, but not limited to, breeding plants by any conventional methodology, or genetic transformation. If the sequences are stacked by genetically transforming the plants,
25 the polynucleotide sequences of interest can be combined at any time and in any order. The traits can be introduced simultaneously in a co-transformation protocol with the polynucleotides of interest provided by any combination of transformation cassettes. For example, if two sequences will be introduced, the two sequences can be contained in separate transformation cassettes (trans) or contained on the same
30 transformation cassette (cis). Expression of the sequences can be driven by the same promoter or by different promoters. In certain cases, it may be desirable to introduce a transformation cassette that will suppress the expression of the polynucleotide of

interest. This may be combined with any combination of other suppression cassettes or overexpression cassettes to generate the desired combination of traits in the plant. It is further recognized that polynucleotide sequences can be stacked at a desired genomic location using a site-specific recombination system. See, for example,
5 W099/25821, W099/25854, W099/25840, W099/25855, and W099/25853.

The transformed plants may be analyzed for the presence of the gene(s) of interest and the expression level. Numerous methods are available to those of ordinary skill in the art for the analysis of transformed plants. For example, methods for plant analysis include Southern and northern blot analysis, PCR-based (or other
10 nucleic acid amplification-based) approaches, biochemical analyses, phenotypic screening methods, field evaluations, and immunodiagnostic assays (e.g., for the detection, localization, and/or quantification of proteins).

EXAMPLES

15 Example 1

Isolation and Characterization of GmSAMT1

Plant, Nematode, Bacterium, and Chemical Sources: In the studies described herein, two genetically-related soybean lines, TN02-226 and TN02-275, were used for gene cloning and generating soybean hairy roots. The breeding details for these
20 two soybean lines have been described in Mazarei *et al. Theor. Appl. Genet.* 123: 1193-1206, 2011, which is specifically incorporated herein by reference in its entirety. These two F₆-derived sister lines, TN02-226 and TN02-275, are resistant and susceptible, respectively, to the SCN race 2 (HG type 1.2.5.7), which has been confirmed in USDA Southern Regional Tests from 2004 through 2007. In the
25 studies described herein, the SCN race 2 eggs were utilized as the pathogens in the bioassay of the transgenic hairy roots since the above two soybean lines exhibited different responses to SCN race 2. The maintenance of SCN followed the method described by Arelli *et al., Heterodera glycines*. Crop Sci. 40: 824-826, 2000.

Escherichia coli strain BL21 (DE3) CODONPLUS® (Stratagene, La Jolla, CA) was
30 used for the expression of recombinant protein. *Agrobacterium* cultures, *A. tumefaciens* GV3101 and *A. rhizogenes* strain K599, were utilized for tobacco transient transformation and the generation of the soybean hairy roots. All chemicals

were obtained from Sigma-Aldrich (St. Louis, MO).

Viruses, Soybean Genotypes, Inoculation, and SMV Detection: In the studies described herein, a mutant of soybean mosaic virus SMV-N (SMV-N25) was used for virus infection on soybean. Two soybean genotypes Williams 82, susceptible to SMV-N25, and L78-375, resistant to SMV-N25, were used. The mechanically inoculated plants were maintained in a growth chamber operating at 22 °C with a photoperiod of 16 h. SMV inoculated soybean seedlings were grown for another 21 days and their trifoliolate leaves were harvested, pulverized in the presence of liquid nitrogen and stored at -80 °C until processed for quantitative reverse-transcription PCR analysis.

Sequence Analysis: The sequence of Affymetrix probe Gma.12911.1.A1_S_AT, which identified our target gene in microarray experiment, was used as a query to BLAST search against the soybean database GmGDB/GMtranscript (as available on the world wide web at plantgdb.org/GmGDB/cgi-bin/blastGDB.pl). Program blastn was performed with E-value of 1e-20. Five soybean genes showed high similarity (**FIG. 2**). A multiple sequence alignment of selected SABATH proteins and Glyma02g06070 was performed using ClustalX program and viewed using TreeView software (as available on the world wide web at taxonomy.zoology.gla.ac.uk/rod/treeview.html). A neighbor-joining unrooted phylogenetic tree was constructed (**FIG. 3**).

Isolation of Glyma02g06070 Gene from Soybean: Total RNA was isolated from the SCN-infected the root tissue of soybean line TN02-226, using the RNEASY® Plant Mini Kit (QIAGEN®, Valencia, CA, USA) and DNA contamination was removed with an on-column DNase (QIAGEN®, Valencia, CA, USA) treatment. Then cDNA was synthesized from total RNA in a 15 µl reaction volume using the First-Strand cDNA Synthesis Kit (GE Healthcare, Piscataway, NJ, USA) as previously described. Based on microarray analysis, one soybean candidate gene *Glyma02g06070* was chosen for further study. The full length cDNA sequence of Glyma02g06070 was gained from public soybean database (as available on the world wide web at phytozome.net/soybean), and primers were designed as following ATGGAAGTAGCACAGGTACTCCACATG (SEQ ID NO: 5) and TGCTTTTCTAGTCAATAATATGGTAAC (SEQ ID NO: 6). The PCR conditions

were as follows: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min 30 s, and a final extension at 72 °C for 10 min. PCR products were cloned into vector pEXP5/CT-TOPO®. By sequencing, the PCR product from TN02-226 soybean was the same as the one predicted in the soybean database.

To verify its biochemical function, the candidate gene was then cloned into a protein expression construct vector, pET100/D-TOPO® vector (Invitrogen, Carlsband, CA) and expressed in the *E. coli* strain BL21 (DE3) CODONPLUS® (Stratagene, La Jolla, CA). Protein expression was induced by isopropyl β -D-1-thiogalactopyranoside (IPTG) at a concentration of 500 μ M for 18 h at 25 °C, with cells lysed by sonication. *E. coli*-expressed Glyma02g06070 with a His-tag was purified from the *E. coli* cell lysate using Ni-NTA agarose following the manufacturer instructions (Invitrogen). Protein purity was verified by SDS-PAGE and protein concentrations were determined by the Bradford assay.

Radiochemical SAMT Activity Assay: A 50 μ l volume containing 50 mM Tris-HCl, pH 7.5, 1 mM SA, and 3 μ M 14 C-SAM with a specific activity of 51.4 mCi/mmol (Perkin-Elmer, Boston, MA) was used. The assay was initiated by addition of SAM, maintained at 25 °C for 30 min, and stopped by the addition of ethyl acetate (150 μ l). After phase separation by 1 min centrifugation at 14,000 g, the upper organic phase was counted using a liquid scintillation counter (Beckman Coulter, Fullerton, CA). The radiation counts in the organic phase indicated the amount of synthesized MeSA. Substrate specificity assay was also performed for recombinant GmSAMT1, with a range of substrates including SA, benzoic acid, anthranilic acid, 3-hydroxybenzoic acid, 4-hydroxybenzoic acid, nicotinic acid, p-coumaric acid, caffeic acid, cinnamic acid, vanillic acid, jasmonic acid (JA), gibberellic acid, indole-3-acetic acid, and 2,4-dichlorophenoxyacetic acid. The activity with SA was set at 100%. Three independent assays were performed for each compound.

Determination of Kinetic Parameters of GmSAMT1: Appropriate enzyme concentrations and incubation time were determined in time course assays to make sure that the reaction velocity was linear during the assay period. Lineweaver-Burk plots yielded apparent K_m values. To determine the K_m for SA, concentrations of

SA were independently varied from 2 to 150 μ M, while SAM was held constant at 200 μ M. Assays were conducted at 25 °C for 30 min, as described in SAMT activity assay. The optimal pH for GmSAMT1 was assayed from pH 6.0 to pH 9.0, and optimal reaction of temperature was assayed ranging from 0 °C to 60 °C. Effects of metal ions on activity of GmSAMT1 were assayed with a range of metal ions, including K^+ , Cu^{2+} , Zn^{2+} , Fe^{2+} , Fe^{3+} , Mn^{2+} , Ca^{2+} , Na^+ , NH_4^+ , and Mg^{2+} , in the form of chloride salts at 5 mM final concentrations.

Creating the Constructs with Over-expressed GmSAMT1 and OFP Reporter

Genes: The pCAMBIA 1305.2 vector was used as a primary binary vector. The following is the procedures of constructing the vector harboring over-expressed *GmSAMT1* (**FIG. 6**). As a screening tool, a orange fluorescent protein (OFP) gene, *ppor RFP* with primers containing *NcoI* fragment, was introduced into *NcoI* site of pCAMBIA 1305.2 vector. This modified pCAMBIA 1305.2 vector was called pJL-OFP vector and used to test OFP expression through transient transforming tobacco. Then the cassette of 35S::GUS/NOS terminator was taken from a pBI121 vector by digestion with *EcoRI* and *HindIII* and inserted into the *EcoRI* and *HindIII* sites of pJL-OFP vector to form pJL-OFP-35S::GUS vector, which was used for testing the OFP and GUS gene expression in one binary vector. *GmSAMT1* gene was amplified from the above mentioned pEXP5/CT-TOPO® vector with primers containing *BamHI* and *SacI* fragment and replaced with GUS gene in pJL-OFP-35S::GUS vector by digestion and ligation. Finally the over-expressed *GmSAMT1* vector was constructed and called pJL-OFP-35S::GmSAMT1 vector. The vectors of pJL-OFP-35S::GmSAMT1 and pJL-OFP were transformed into *A. rhizogenes* as target gene and vector control, respectively. *A. tumefaciens* GV3101 and *A. rhizogenes* strain K599, were transformed by freeze-thaw method. The cells with no binary vector were plated on LB media with no antibiotic. The cells with the binary vectors were plated on selective LB media with kanamycin and incubated at 28°C for 2 days.

Testing the Reporter Gene Expression Through Tobacco Transient

Transformation: Before generating hairy roots, the constructs containing OFP gene were tested by transient expression assays in tobacco plants according to Sparkes *et al.*, *Nat. Protocols* 1: 2019-2025, 2006. The plasmids, pJL-OFP and pJL-OFP-35S::GmSAMT1, were tested for OFP function by agroinfiltration of *A. tumefaciens*

GV3101 into tobacco leaves, along with *A. tumefaciens* GV3101 without binary plasmid as a negative control. Bacterial suspensions carrying target plasmids were pressure-infiltrated using a syringe with no needle into the 5-week tobacco leaves through the abaxial surfaces over leaf segments lying between the major veins that branch off from the mid-rib of the leaf. OFP fluorescence was investigated 2 days after infiltration in excised leaf sections that covered the infiltrated zone part of surrounding non-infiltrated tissue, using an Olympus SZX12 fluorescence microscope and an OFP filter.

Generation of Transgenic Soybean Hairy Roots: The soybean hairy root transformation system by infection with *A. rhizogenes* harboring the target genes were modified and used to study soybean resistance against SCN. *A. rhizogenes* K599 harboring the target gene, vector control and *A. rhizogenes* K599 alone were used to inoculate the TN02-275 soybeans, which are susceptible to SCN race 2. To better evaluate the resistance of the target gene, *A. rhizogenes* harboring the vector control and *A. rhizogenes* K599 alone were also used to inoculate the TN02-226 soybeans, which are resistant to SCN race 2. Generation of soybean hairy roots was modified as follows: First, soybean seeds were sterilized with chlorine gas. Second, a single layer of mature soybean seeds with intact seed coats were placed in open Petri dish, which was put in a bell jar desiccator within a fume hood. A beaker with 100 ml bleach was also placed in the desiccator and added 5 ml concentrated (12N) HCl along the side of the beaker. Then the seeds were kept in the closed desiccator immediately and maintained in the chlorine gas atmosphere overnight (about 16 h). After sterilization, seeds were transferred to the autoclaved filter paper moistened with sterilized distilled water in the Petri dishes for germination in a laminar flow hood. The germinated soybean seeds were sowed in sterilized vermiculite and six seeds were grown in each cell of 18-cell germination tray. The overnight cultured bacteria lawn of *A. rhizogenes* K599 with binary vector pJL-OFP, pJL-OFP-35S::GUS or pJL-OFP-35S::GmSAMT1 from LB medium plate containing 50 mg ml⁻¹ kanamycin, or a culture of *A. rhizogenes* K599 lacking the binary vector grown in LB medium without any antibiotics was collected and suspended in 1 ml of sterile distilled water, respectively. Soybean cotyledonary nodes and upper hypocotyls of 5-day-old seedlings with unfolded cotyledons were stabbed three

times by the syringe containing the bacterial suspension. In the first week, the trays were covered with the sterile transparent lids to retain high humidity. In the subsequent three weeks, plants were transferred into pots and the *A. rhizogenes* wounding sites were covered by wet sterile vermiculite to continually keep high

5 humidity. Each day soybean plants were watered using sterile B&D solution containing 1 mM CaCl₂, 0.5 mM KH₂PO₄, 10 µM Fe-citrate, 0.25 mM MgSO₄, 0.25 mM K₂SO₄, 1 µM MnSO₄, 2 µM H₃BO₄, 0.5 µM ZnSO₄, 2 µM CuSO₄, 0.1 µM CoSO₄, 0.1 µM Na₂MoO₄, and 1 mM KNO₃. The growth conditions for soybean plants were 12 h light/12 h dark at 28 °C /25 °C with irradiance from

10 fluorescent bulbs at 150-200 µmol m⁻²·s⁻¹. After about four weeks, the hairy roots grew to approximately 10 cm in length. Soybean transgenic roots were detected based on GFP expression, using a fluorescence stereomicroscope (Olympus SZX12 fluorescence microscope and GFP filter). The tap root and non-transgenic roots were excised under the wounding site. Transgenic roots harboring pJL-GFP-

15 35S::GUS screened by GFP signal were immersed in GUS solution (1 mM X-Gluc), 0.5 mM potassium ferrocyanide, 0.1% (v/v) TRITON™ X-100, 100 mM sodium phosphate buffer pH 7.0), and incubated at 37 °C overnight. In the following day, the X-Gluc solution was discarded. The roots were washed repeatedly with 70% ethanol to remove chlorophyll. Transgenic roots with visual

20 GUS stain were recorded.

Nematode Infection and Demographics Assay: To easily compare the SCN infection, one transgenic root was left for the SCN inoculation. Other transgenic roots were collected for RNA extraction and detection of the gene expression. The hairy root parts of twenty soybean plants harboring the same construct were loaded

25 horizontally in a 13x 9 x 2 cm sterilized inoculating tray containing one thin layer mixture of sterile sand and top soil (1:1) and the shoot parts were left out of inoculating tray. About 0.75 ml of inoculum, which contained about 5600 SCN eggs, was added to each root system. Nematodes were allowed to infect roots for 7 days under humid conditions. Then all the roots were washed to remove extra SCN eggs

30 and juvenile nematodes that had not penetrated the root tissue. Sterile vermiculite within cone-tainers was used for growing the infected chimeras in the growth chamber. In two weeks post inoculation, infected root samples were washed to

remove vermiculite by tap water and cleared by 20% (v/v) bleach for no more than 4 min, and then stained by acid fuchsin. Nematodes of different stages within each infected root sample were recorded. Since the responses between the susceptible and resistant soybean lines occur on the J3 stage of SCN, the ratio of the number of J3+J4 to total number of nematodes per plant was used as an index to determine the resistance difference among the hairy roots harboring respective constructs. A minimum of 10 transformed plants were analyzed for each vector.

Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR):

Quantitative reverse transcription PCR was performed to check *GmSAMT1* and some defense-related genes' transcription level of transgenic soybean hairy roots according to Mazarei *et al.*, 2007. In order to conduct qRT-PCR assays, gene specific primers of *GmSAMT1*, defense-related genes, *GmPR-1* (GenBank accession BU577813), *GmPR-3* (GenBank accession AF202731), and reference gene, soybean ubiquitin3 gene (*GmUBI-3*, GenBank accession D28123), were designed (**Table 1**).

Table 1. Primers for qRT-PCR analysis of soybean SA methyltransferase gene and relevant pathogenesis-related genes.

Gene Name	Orientation	Sequence (5'-3')
<i>GmSAMT1</i>	Forward	GTGAAATTGGAAGTTCTTAAAGAAGGA (SEQ ID NO: 7)
	Reverse	CAGATTCAAAGTCTAGAGCATTCAC (SEQ ID NO: 8)
<i>GmUBI-3</i>	Forward	GTGTAATGTTGGATGTGTTCCC (SEQ ID NO: 9)
	Reverse	ACACAATTGAGTTCAACACAAACCG (SEQ ID NO: 10)
<i>GmPR-1</i>	Forward	AACTATGCTCCCCCTGGCAACTATATTG (SEQ ID NO: 11)
	Reverse	TCTGAAGTGGTAGCTTCTACATCGAAACAA (SEQ ID NO: 12)
<i>GmPR-3</i>	Forward	AACTACAATTACGGGCAAGCTGGCAA (SEQ ID NO: 13)

		ID NO: 13)
	Reverse	TTGATGGCTTGTTTCCCTGTGCAGT (SEQ ID NO: 14)
<i>GmNPR1-1</i>	Forward	AATTGACCAAGAGCTTCCGC (SEQ ID NO: 26)
	Reverse	CTACAGAAGCATCATTTTCAACATCTT (SEQ ID NO: 27)
<i>GmNPR1-2</i>	Forward	CAATTGACCAAGAGCTTCCAA (SEQ ID NO: 28)
	Reverse	CTATAGAAGCATCGTTTTCAACATCTC (SEQ ID NO: 29)
<i>GmICS1</i>	Forward	GGCCATTTTCGGAGCTGG (SEQ ID NO: 30)
	Reverse	AGGAGAAGGTGGTTCTGTGAGAGA (SEQ ID NO: 31)
<i>GmICS2</i>	Forward	CATGGCCATTTTCGGAGCTTA (SEQ ID NO: 32)
	Reverse	CAGAAGATGGTTCTGCGAATGG (SEQ ID NO: 33)

- Total RNA was isolated from the respective root tissues of transgenic hairy roots and control hairy roots using RNEASY® columns (QIAGEN®) following the manufacturer's instructions for plants. RNA was treated with RNase-Free DNase set (QIAGEN®) to remove genomic DNA and total RNA (about 1.5 µg) was reverse-transcribed to synthesize cDNA using High-Capacity cDNA Reverse Transcription Kit (APPLIED BIOSYSTEMS®) for qRT-PCR according to manufacturer's instructions. First-strand cDNA was diluted and placed in each qRT -PCR reaction. DNA accumulation was measured using SYBR® Green as the reference dye.
- Genomic DNA contamination was monitored by PCR using RNA sample as template. Only one product was present in each reaction as indicated by the reference dye's dissociation curve of amplified products. The following PCR conditions were used: 50 °C for 2 min; 95 °C for 10 min; followed by 40 cycles of

95 °C for 15 s; 60 °C for 1 min; 72 °C for 30 s. All qRT-PCR assays were conducted in triplicate. PCR efficiencies for target and reference genes were equal between the target and control samples. Ct values and relative abundance were calculated using software supplied with the APPLIED BIOSYSTEMS® 7900 HT Fast Real-Time
5 PCR system.

Statistical Analysis: The resistance difference among the transgenic soybean hairy roots with over-expression of *GmSAMT1* and control soybean hairy roots was analyzed by one way ANOVA using the mixed model of SAS (SAS 9.2 version). Least squares means of the resistance index were separated using Fisher's least
10 significant difference (LSD) test. The expression difference of target gene and defense-related genes examined by qRT-PCR was also analyzed by one way ANOVA using the mixed model of SAS. Least squares means of relative expression level were separated using Fisher's LSD test. A $P < 0.05$ was considered significant.

15 RESULTS

Glyma02g06070 Is a Member of the SABATH Family: Microarray analysis on soybean recombinant inbred lines (RILs), TN02-226 and TN02-275, revealed some genes with significant expression changes during resistant and susceptible reactions to SCN. One of these genes showed significant up-regulation only in SCN
20 inoculated resistant soybean roots while no significant change of its gene expression occurred in SCN-inoculated susceptible soybean roots (Mazarei *et al.*, *Theor. Appl. Genet.* 123: 1193-1206, 2011). To obtain the complete open reading frame of this resistance candiadate gene, the sequence of Affymetrix SoyChip probe Gma.12911.1.A1_S_AT, which identified our target gene from microarrays, was
25 used as a query to BLAST search against the soybean database GmGDB/GMtranscript (as available on the world wide web at plantgdb.org/GmGDB/cgi-bin/blastGDB.pl) and SoyBase (as available on the world wide web at soybase.org/gbrowse/cgi-bin/gbrowse/gmax1.01/), the soybean gene, *Glyma02g06070* was found to have the highest similarity (99%) with the sequence
30 of Gma.12911.1.A1_S_AT and was annotated for encoding SAM-dependent carboxyl methyltransferase (**FIG. 2**).

By further comparison with the known methyltransferases from other species, this SAM-dependent carboxyl methyltransferase encoded by *Glyma02g06070* was determined to be a member of the SABATH family. The SABATH family is named after the first identified enzymes (SAMT, BAMT and theobromine synthase) in this family (Chen et al. 2003a). To investigate the evolutionary relationships between *Glyma02g06070* and the known SABATH members from other species, a phylogenetic tree of selected functionally characterized SABATH members with *Glyma02g06070* was constructed. The result showed that *Glyma02g06070* belonged to the cluster containing the enzymes with salicylic acid methyltransferase activity (**FIG. 3**).

***Glyma02g06070* Encodes Soybean SAMT:** A protein BLAST search reveals that *Glyma02g06070* contains 370 amino acids and is 40.1% identical to the previously identified rice SAMT, OsBSMT1, at the amino acid sequence level. *Glyma02g06070* was also predicted to contain an SAM-binding domain. When the sequence of *Glyma02g06070* was searched in EST database, it also showed this gene is induced by SCN and SA. To confirm the prediction from the databases, we amplified and sequenced the complete open reading frame (ORF) of this gene from cDNA attained from SCN-infected root tissue of soybean TN02-226 line.

Since *Glyma02g06070* encodes a member of the SABATH family, whose member enzymes can transfer a methyl group from methyl-donor SAM to a range of substrates, the enzyme activity of the recombinant protein of *Glyma02g06070* was examined using various potential substrates (**Table 2**).

Table 2 Relative activity of GmSAMT1 with salicylic acid and other related substrates

Substrate	Relative activity ^a (%)
Salicylic acid	100
Benzoic acid	16.9
Anthranilic acid	9.3
3-Hydroxybenzoic acid	1.8
4-Hydroxybenzoic acid	0
Nicotinic acid	0
<i>p</i> -Coumaric acid	0

Caffeic acid	0
Cinnamic acid	0
Vanillic acid	0
Indole-3-acetic acid	0
2,4-Dichlorophenoxyacetic acid	0
Jasmonic acid	0
Gibberellic acid	0

a: Values are averages of three independent measurements. All substrates were tested at 1 mM concentration. The activity of the recombinant GmSAMT1 with salicylic acid was set arbitrarily at 100%.

Increases in reaction rate by increasing concentrations of SAM and SA were found to obey Michaelis-Menten kinetics. Among the 14 examined substrates, SA, benzoic acid, anthranilic acid and 3-hydroxybenzoic acid served as *in vitro* substrates in the reactions catalyzed by the recombinant protein encoded by *Glyma02g06070*, although they showed obvious differences in activities. At substrate concentrations of 1 mM, the recombinant protein showed the highest activity with SA (100%) rather than benzoic acid (16.9%), anthranilic acid (9.3%) and 3-hydroxybenzoic acid (1.8%). For other substrates including JA, indole-3-acetic acid, and gibberellic acid, no activity was detected under the same reaction conditions. *Glyma02g06070* catalyzed the methylation of SA with an apparent K_m value of $32.9 \pm 1.7 \mu\text{M}$. Final values represent the average of three independent measurements. Due to SA serving as its sufficient substrate, *Glyma02g06070* was designated as *GmSAMT1*.

Biochemical Properties of Soybean SAMT: Recombinant GmSAMT1 with a His-tag expressed in *E. coli* was purified to electrophoretic homogeneity and subjected to detailed biochemical characterization. The molecular mass of the GmSAMT1 with His-tag protein estimated on the gel was 43 kDa, which was close to the expected size (FIG. 4). Purified recombinant GmSAMT1 exhibited an apparent K_m value of $32.9 \pm 1.7 \mu\text{M}$ for salicylic acid. When enzyme assays were performed in buffers of different pH values, GmSAMT1 showed the highest level of catalytic activity at pH 7.5. (FIG. 5A). The optimal temperature for GmSAMT1 activity was 25 °C (FIG. 5B). The effects of various metal ions on GmSAMT1

activity were measured. Compared with control, GmSAMT1 exhibited double activity in presence of K^+ , while it showed no activity in presence of Cu^{2+} , Zn^{2+} , Fe^{2+} and Fe^{3+} . GmSAMT1 activity was moderately inhibited by Mn^{2+} and Ca^{2+} . Other metal ions, including Na^+ , NH_4^+ , and Mg^{2+} , had minimal effect on

- 5 GmSAMT1 activity (**FIG. 5C**). GmSAMT1 displayed pseudo Michaelis-Menton kinetics, with apparent K_m value of $46.2 \pm 4.2 \mu M$ for salicylic acid (**FIG. 5D**). All the above results suggested the recombinant GmSAMT1 showed salicylic acid methyltransferase activity. Taken together, the phylogenetic and biochemical evidence suggested that GmSAMT1 is a conserved enzyme among a number of
- 10 plant species.

Reporter Gene Expression through Tobacco Transient Transformation:

Before performing *A. rhizogenes*-mediated transformation, transient expression of pJL-OFP-35S::GUS in tobacco leaf through *A. tumefaciens*-mediated transformation was conducted to check OFP signal under OFP filter. Transgenic tobacco leaves

15 harboring the target gene displayed OFP fluorescence, while the mock plant showed no OFP signal (**FIG. 7A**). This result helped to rapidly confirm the OFP expression.

After *A. rhizogenes*-mediated transformed soybean hairy roots with target genes were generated to be about 10 cm long, transgenic hairy roots harboring over-expressed *GmSAMT1* were screened under OFP filter. This method was

20 demonstrated by checking GUS expression by staining the transgenic hairy roots harboring plasmid pJL-OFP-35S::GUS. All the positive transgenic hairy roots screened by OFP signal showed uniform GUS expression in the root tissue, while negative control hairy roots screened by OFP signal did not show GUS expression (**FIG. 7B**).

Gene Expression of GmSAMT1 and Several Pathogenesis-Related (PR)

Genes: The gene expression level of *GmSAMT1* and selected PR genes in transgenic hairy roots were examined by qRT-PCR (**FIG. 8 and FIG. 9**). The expression level of *GmSAMT1* in over-expressed *GmSAMT1* transgenic hairy roots was about 25 times higher than the control plants (**FIG. 8B**). *GmPR-1* showed significant higher

30 expression in over-expressed *GmSAMT1* transgenic hairy roots compared with some control hairy root lines. But there was no significant difference of the gene expression of *GmPR-1* between the transgenic hairy roots with over-expression of

GmSAMT1 and vector control in the susceptible soybean background (**FIG. 9A**). This may suggest that the susceptible soybean can respond to exogenous proteins when SCN infection does not occur. When SCN parasitism happens, the defense system of the susceptible soybean may be changed or inhibited by SCNs, which
5 produce and secrete parasitism proteins into soybean to aid themselves in the suppression of host plant defense. Chorismate mutase was found to be one of the SCN parasitism proteins, which can manipulate the plant's shikimate pathway. Since the SA biosynthesis is closely associated with the shikimate pathway, the SA signaling pathway of the susceptible soybean would be suppressed by SCN. It was
10 already found that a local suppression of SA signaling in Arabidopsis roots was required for beet cyst nematode parasitism. In future, the expression of PR genes needs to be examined for the SCN-infected transgenic hairy roots. Since *PR-1* gene was considered as a marker gene of SA signaling pathway, over-expressed *GmSAMT1* might increase the signal transduction of SA pathway by providing
15 abundant mobile molecule MeSA for responding SCN infection. Interestingly, the expression of *PR-3*, which encodes soybean chitinase, was observed to be higher in over-expressed *GmSAMT1* transgenic hairy roots than other control lines (**FIG. 9B**). JA was found to induce *PR-3* expression in Arabidopsis.

SCN is the most economically damaging pest of soybean in the U.S., so the
20 mechanism of the interaction between soybean and SCN has been studied to better control this pest. With the development of genomics, a number of microarray analyses have been performed to study gene expression changes of soybean when infected by SCN. However, the various experiments are not congruent owing to different types of soybean and SCN samples. In the present study, the candidate
25 gene was identified based on microarray analysis of two genetically-related soybean sister lines TN02-226 and TN02-275, which are resistant and susceptible, respectively, to the SCN race 2 infection.

During SCNs' inoculation, *GmSAMT1* activity may also be manipulated through changing the concentration of potassium of soybean cells. The most
30 common aboveground symptom of SCN infection is a potassium deficiency symptom with a bright yellowing of the margins of leaves. However, such SCN-induced potassium deficiency symptoms are the result of the plant's inability to raise

sufficient potassium due to nematode feeding. It was demonstrated that potassium concentration in soybean roots grown at the medium level of potassium fertility condition was decreased by SCN (Smith et al. 2001). The result indicated that *in vitro* GmSAMT1 showed double activity in the presence of potassium, which
5 suggests that SCN may also inhibit the activity of GmSAMT1 by decreasing the concentration of potassium in addition to suppressing the transcriptional level of *GmSAMT1*.

GmSAMT1 may have a role in crosstalk between SA and JA signaling pathway in soybean resistance against SCN. It was demonstrated that the gene
10 expression of rice *SAMT* can be induced by JA. MeSA was also detected to be highly elevated after the wild-type Arabidopsis was treated by MeJA. As disclosed herein, the over-expression of *GmSAMT1* in the susceptible soybean line caused up-regulation of *GmPR-1* and *GmPR-3*, which has the integrative effect of SA and JA signaling pathways. Microarray analyses also indicate that the JA pathway was
15 activated in the soybean defense response against SCN. Therefore, the SA-dependent pathway, combined with JA pathway through the regulatory points, including NPR1, WRKY family transcription factor, and LOX, play a role of restraining the development of SCN. The results presented herein indicated that
20 *GmSAMT1* is another regulatory point in the crosstalk between SA and JA pathways.

Resistance to SCN: The resistance of soybean to SCN has been assessed by assaying SCN development, i.e., nematodes, cysts, or eggs per plant. In this study, this method was modified according to Melito *et al.*, *BMC Plant Biology* **10**:104, 2010. The ratio of the number of J3 and J4 stage nematodes to the total SCN number
25 was set as the index, (J3+J4)/total SCN, to evaluate the resistance among the different hairy roots two weeks after inoculation with SCN eggs.

Female nematodes from J2 to J4 stages were observed for most transgenic hairy root sections and male nematodes were occasionally observed (**FIG. 11**). The number of SCN varied widely among individual transgenic and control plants. The
30 mean number of established SCN was not significantly different between the transgenic hairy root line and the controls (data not shown). However, the ratio, (J3+J4)/total SCN, showed a significant difference between transgenic hairy roots

and susceptible control line. The (J3+J4)/total SCN ratios for the transgenic hairy roots overexpressing *GmSAMT1* (3.4% for TN02-275, 5.3% for Williams 82) were comparable to the resistant vector control (2.5% for TN02-226), and were significantly different with the susceptible vector control (23.6% for TN02-275, 18.4% for Williams 82). The transgenic soybean hairy roots with overexpressing *GmSAMT1* had significantly fewer female juvenile SCNs developing to J3 and J4 stages compared with vector control in the susceptible soybean background. When the susceptibility level of vector control in the susceptible background was arbitrarily set at 100%, the transgenic soybean hairy roots overexpressing *GmSAMT1* has significant reduction of susceptibility (85.6% reduction for TN02-275, 71.2% reduction for Williams 82).

Of most interest is that transgenic hairy roots with overexpressing *GmSAMT1* exhibit higher resistance to SCN than control transgenic hairy roots (**FIG. 11**). The resistance mediated by *GmSAMT1* is not line-specific, because two susceptible lines showed the same phenotype (**FIG. 11**). The biological role of *GmSAMT1* in SCN resistance can be further evaluated using stable transformation.

Expression of SA pathway related genes in transgenic hairy roots:

Transcript abundance for an SA-biosynthesis protein, isochorismate synthase (ICS)-encoding gene was assayed in SAMT-transgenic hairy roots (**FIG. 12**). In Arabidopsis, *AtICS1* was required for the pathogen-induced accumulation of SA. Although *ICS* gene was proposed to be a critical gene for SA biosynthesis, little is known about the role of ICS in soybean. The expression of *GmICS1* (*Glyma01g25690*) (**FIG. 12A**) and *GmICS2* (*Glyma03g17420*) (**FIG. 12B**) was examined by qRT-PCR as well. Without SCN treatment, both *GmICS1* and *GmICS2* in the *GmSAMT1* overexpression transgenic hairy roots showed higher transcript abundance (5.3 times and 8.2 times) than controls. Under SCN challenge treatment there was no significant difference on the expression of *GmICS1* between the transgenic hairy root and control line. For *GmICS2*, the transgenic hairy roots had slightly greater transcript abundance than controls.

To further explore the SA dependent signal transduction pathway, the key regulatory gene of the SA signal pathway, two homolog of *non-expressor of PRI* (*NPRI*) genes, *GmNPRI-1* (*Glyma09g02430*) and *GmNPRI-2* (*Glyma15g13320*) in

soybean were tested by qRT-PCR (**FIG. 13**). For *GmNPR1-1*, there was no significant difference on the expression level between the transgenic hairy root and control lines without SCN treatment. However, with SCN treatment, *GmNPR1-1* showed higher expression than the control (**FIG. 13A**). For *GmNPR1-2*, there was no significant difference in the expression level between the transgenic hairy root and control line with SCN treatment. However, without SCN treatment, *GmNPR1-2* showed higher expression than controls. Taken together, *GmNPR1-2* showed relatively higher expression in the transgenic hairy root in response to SCN treatment (**FIG. 13B**).

10 ***Biological Roles of GmSAMT1 in Soybean Defense Against Soybean Mosaic Virus:*** To test whether *GmSAMT1* is also involved in resistance against SMV, qRT-PCR was also conducted on measuring the expression of *GmSAMT1* and *GmPR-1* during interaction between SMV and soybean. In this assay two soybean lines L78-379 (*Rsv1*) and Williams 82 (*rsv1*) were utilized. Soybean genotype L78-379 is an isolate of Williams, but contains the *Rsv1* resistance gene from PI 96983 (*Rsv1*). Inoculation of L78-379 (*Rsv1*) with a mutant derived from SMV-N (N25) results in induction of delayed hypersensitive response, whereas infection of Williams 82 (*rsv1*) with the same virus causes mosaic.

When the resistant response occurred for L78 against soybean virus 25 (21 days post inoculation), *GmSAMT1* expression of the soybean top leaves is around 7.3 times that in the mock plant, and interestingly, a high expression level (44.6 times) of *GmPR-1* were also observed. While in the susceptible response of soybean Williams 82, the expression of *GmSAMT1* and *GmPR-1* was reduced to half of that in the mock plant (**FIG. 14**). So the strongly positive correlation between the expression of *GmSAMT1* and *GmPR-1* indicates that *GmSAMT1* plays a role in soybean defense against soybean mosaic virus.

Taken together, the gene expression differences of *GmSAMT1* and selected *PR* genes, were up-regulated in the two inoculated tissues of the resistant soybean line, roots and leaves but down-regulated or not changed in the susceptible soybean line, demonstrates soybean responses against pathogens.

Example 2

Isolation and Characterization of GmSABP2-1

Plant, Nematode, Bacteria, and Chemical Sources: In this study, two genetically-related soybean lines, TN02-226 and TN02-275, were used for gene cloning and generating soybean hairy roots. The breeding details for these two soybean lines and have been described. SCN race 2 eggs were used as the pathogens in the bioassay of the hairy roots, with the same inoculation method described above. *Escherichia coli* strain BL21 (DE3) CODONPLUS® was used for expressing the recombinant protein. *A. rhizogenes* strain K599 was utilized for carrying the over-expressed *GmSABP2-1* and generating soybean hairy roots. The chemicals used in this study were obtained from Sigma-Aldrich.

Database Search and Sequence Analysis: Both the sequences of Affymetrix probes GmaAffx.92649.1.S1_at and Gma.5867.1.A1_s_at showed 100% similarity with one soybean gene, *Glyma16g26060*. Based on the annotation in phytozome (as available on the world wide web at phytozome.net/soybean), this gene encodes an α/β fold hydrolase. Alignment of *Glyma16g26060* with two other characterized methyl ester esterase was performed to study the conserved domains of this candidate protein.

Isolation of Glyma16g26060 Gene from Soybean: Total RNA extraction and DNA contamination removal were performed as described. To further study the function of *Glyma16g26060*, the full length cDNA sequence of *Glyma16g26060* was obtained from public soybean database (as available on the world wide web at phytozome.net/soybean). Using the specific primers ATGGGGTTCACAAAATTGTATGGATAGG (SEQ ID NO: 15) and TCATGCATATTTAGTCGCTATCTGCTG (SEQ ID NO: 16), thermal cycling conditions were: 94 °C for 2 min followed by 35 cycles at 94 °C for 30 s, 57 °C for 30 s and 72 °C for 1 min, and a final extension at 72 °C for 10 min. The PCR product was cloned into vector pEXP5/CT-TOPO®. The sequencing result of the PCR product from TN02-226 soybean was consistent with that predicted in the soybean database on Williams 82 soybean.

To verify its biochemical function, the candidate gene was then cloned into a protein purification vector, pET100/D-TOPO® vector and expressed in the *E. coli*

strain BL21 (DE3) CODONPLUS®. Protein expression was induced by isopropyl β -D-1-thiogalactopyranoside (IPTG) at a concentration of 500 μ M for 18 h at 25°C, with cells lysed by sonication. *E. coli*-expressed Glyma16g26060 with a His-tag was purified from the *E. coli* cell lysate using Ni-NTA agarose following the
5 manufacturer instructions (Invitrogen). Protein purity was verified by SDS-PAGE and protein concentrations were determined by the Bradford assay.

Methyl Ester Esterase Activity Assay: A two-step radiochemical esterase assay was performed to determine the activity of *E. coli*-expressed Glyma16g26060 following a protocol previously reported by (Forouhar *et al.*, *Proc. Natl. Acad. Sci. USA* 102: 1773-1778, 2005). The reaction of the first step was the reaction of
10 substrate 10 μ M MeSA and purified recombinant *E. coli*-expressed Glyma16g26060 for 30 min at 25°C. After that, the sample was boiled to stop the reaction and denature the enzyme. The second reaction step was started with the addition of 3 μ M 14 C labeled S-adenosyl-L-methionine (SAM) with a specific activity of 51.4
15 mCi/mmol (Perkin Elmer, Boston, MA, USA), and respective known purified methyltransferase with high activity. Here, GmSAMT1 mentioned previously was used for SA methyltransferase enzyme. The reaction was preceded for 30 min at 25 °C. Then radiolabeled products were extracted and the radioactivity was counted using a liquid scintillation counter.

20 *Determination of Kinetic Parameters of Soybean SABP2:* The increase in reaction rate with increasing concentrations of MeSA was evaluated with the radiochemical assay described above and was found to obey Michaelis-Mention kinetics. Appropriate enzyme concentrations and incubation time were determined in time-course assays so that the reaction velocity was linear during the assay
25 period. To determine the K_m for MeSA, the concentrations of MeSA were independently varied in the range from 5 μ M to 100 μ M. Lineweaver-Burk plots were made to obtain apparent K_m value. Optimal pH for GmSABP2-1 was assayed from pH 6.0 to pH 9.0. Final values are an average of three independent measurements.

30 *Evaluation of SCN Resistance of Transgenic Hairy Roots with Over-expression of GmSABP2-1:* Transgenic soybean hairy roots with over-expressed *GmSABP2-1* were utilized to evaluate the biological function of GmSABP2-1. The

steps including creating the construct with over-expressed *GmSABP2-1* and OFP reporter gene, generating the transgenic soybean hairy roots harboring over-expressed *GmSABP2-1*, the SCN inoculation, and demographics assay of nematodes were performed as described, but *GmSABP2-1* gene was used in place of *GmSAMT1* gene.

Statistical Analysis: The resistance difference between soybean hairy roots with over-expressing *GmSABP2-1* and control soybean hairy roots was analyzed by one way ANOVA using the mixed model of SAS (SAS 9.2 version), and least squares means of the resistance index were separated using Fisher's LSD test. A $P < 0.05$ was considered significant.

RESULTS

Identification and Sequence Analysis of One Putative SCN-Resistant Gene in Soybean: Microarray analysis revealed genes showing significant expression changes during resistant and susceptible reactions to SCN. One of these genes was significantly up-regulated only in SCN inoculated resistant soybean roots whereas it was significantly down-regulated in SCN inoculated susceptible soybean roots through the microarray analysis and quantitative real time PCR.

To obtain the complete open reading frames of this resistance related gene, the sequences of Affymetrix SoyChip probe GmaAffx.92649.1.S1_at and Gma.5867.1.A1_s_at which identified the target gene in microarray analysis, were used as a query to BLAST search against the soybean database GmGDB/GMtranscript (as available on the world wide web at plantgdb.org/GmGDB/cgi-bin/blastGDB.pl), soybean gene *Glyma16g26060* was found to show the highest similarity (100%) with the sequence of the above two probes and predicted to belong to α/β hydrolase gene superfamily.

Glyma16g26060 Encodes Soybean Methyl Ester Esterase: A BLAST search revealed that Glyma16g26060 is 61.2% identical to the previously characterized NtSABP2 (GenBank accession AAR87711.1) at the amino acid sequence level, and its identity with the NtSABP2 is the highest among the soybean homolog genes. The alignment of peptide sequences among Glyma16g26060 and NtSABP2, At4g37150 was performed (**FIG. 15**). Glyma16g26060 contains the

catalytic triad (Ser-86, His-239, and Asp-211) that is highly conserved among SABP2 orthologs from tobacco, Arabidopsis (AtMES), poplar and potato (StMES1) as well as the closely related members of the α/β hydrolase superfamily.

Glyma16g26060 also contains all the 15 amino acids identified in NtSABP2 that
5 interacts with SA. In contrast, AtMES9 contains only 5 and StMES1 shares 12 of these 15 residues. The high sequence similarity and conservation of critical SA-binding residues between *Glyma16g26060* and NtSABP2 indicated that these two proteins shared similar biochemical properties. To confirm the prediction from the databases, the full-length cDNA of the *Glyma16g26060* was cloned via reverse-
10 transcription polymerase chain reaction (RT-PCR) from SCN-infected root tissue of soybean TN02-226 line and sequenced. The sequenced gene contained 786 bp, encoding a protein of 261 amino acids.

Biochemical Properties of Soybean SABP2: To further explore the biochemical properties of the enzyme encoded by *Glyma16g26060*, this gene was
15 cloned into a protein expression vector. Recombinant *Glyma16g26060* with a His-tag expressed in *E. coli* was purified to electrophoretic homogeneity and subjected to detailed biochemical characterization. The molecular mass of the *Glyma16g26060* with His-tag protein estimated on the gel was 33 kDa (**FIG. 16**), which was close to the expected size of our target enzyme. Purified recombinant
20 *Glyma16g26060* showed the activity of converting MeSA to SA. Under steady-state conditions, *Glyma16g26060* hydrolyzed MeSA with an apparent K_m value of $46.2 \pm 2.2 \mu\text{M}$ (**FIG. 17**). So *Glyma16g26060* was designated as soybean SABP2 (GmSABP2-1). When enzyme assays were performed in buffers with different pH conditions, GmSABP2-1 showed the highest level of catalytic activity at around pH
25 7.0 (**FIG. 18**).

Biological Role of Soybean SABP2 in SCN Resistance: The study of the biological role of SABP2 in other plant species was performed through observing the leaves' resistance under virus or bacteria infection. Since the soybean resistance against SCN occurs in the root tissue, the number of SCNs that exist in the
30 transgenic hairy roots at different stages were recorded to analyze the resistance level and evaluate the biological function of GmSABP2-1. As mentioned, the different resistant responses to SCN race 2 between the susceptible (TN02-275) and

resistant (TN02-226) soybean lines occurred on the J3 stage of SCN, the time point after two-week inoculation by SCNs was chosen for analyzing the demographics of SCN in all the hairy roots with respective constructs. The index of J3+J4/total SCN number was used to estimate the resistance between the susceptible line with over-expressed *GmSABP2-1* and other control lines, the same with the assay described. The susceptibility level of vector control in the susceptible background was arbitrarily set at 100%. The results showed the transgenic soybean hairy roots with over-expressed *GmSABP2-1* had significantly fewer female SCN juvenile developed to J3 stage, compared with vector control line in the susceptible soybean background (84.5% reduction). And there was no significant difference on the resistance level between the transgenic hairy roots with over-expressed *GmSABP2-1* and vector control in the resistant soybean background, which also had significant fewer female SCN juvenile developed to J3 stage, compared with negative vector control in the susceptible soybean background (94.8% reduction) (**FIG. 19**).

Described is the identification of GmSABP2-1 from soybean and demonstrated that it shared high similarity and conserved motifs with its tobacco and Arabidopsis orthologs. Through further biochemical assay, it was indicated that recombinant GmSABP2-1 exhibited esterase activity toward MeSA. The K_m of GmSABP2-1 (46.2 μM) was higher than the K_m of NtSABP2 (8.6 μM) PtSABP2-2 (24.6 μM), and lower than the K_m of AtMES1 (57.3 μM), AtMES9 (147.1 μM), PtSABP2-1 (68.2 μM), and StMES1 (57.9 μM). The differences in these kinetic parameters may reflect endogenous MeSA levels of plant species.

The transgenic hairy roots with over-expressed *GmSABP2-1* in the susceptible soybean line, showed increased resistance against SCN. This result indicted that a certain expression level of *GmSABP2-1* was required for soybean defense against SCN. GmSABP2-1 is an important component of receiving MeSA signal to induce SAR in soybean defense against SCN, considering the temporal and spatial differences of SCN on reaching the feeding sites of soybean vascular tissues within the SCN populations. The increased level of GmSABP2-1 protein by up-regulating its gene expression under SCN infection may benefit for capture of the mobile MeSA. Demethylation of MeSA catalyzed by GmSABP2-1 may be able to increase the SA level in the distal undamaged soybean root tissues since its ortholog

NtSABP2 has the role of inducing SAR in the undamaged tobacco leaf.

While this disclosure has been described with an emphasis upon particular embodiments, it will be obvious to those of ordinary skill in the art that variations of the particular embodiments may be used, and it is intended that the disclosure may
5 be practiced otherwise than as specifically described herein. Features, characteristics, compounds, chemical moieties, or examples described in conjunction with a particular aspect, embodiment, or example of the invention are to be understood to be applicable to any other aspect, embodiment, or example of the invention. Accordingly, this disclosure includes all modifications encompassed
10 within the spirit and scope of the disclosure as defined by the following claims.

We claim:

1. An isolated nucleic acid molecule comprising a nucleic acid sequence that encodes a pest resistance protein at least 80% identical to a protein encoded by the nucleic acid sequence according to SEQ ID NO: 1 or SEQ ID NO: 2 or a degenerate variant thereof or a functional fragment thereof.
2. An isolated nucleic acid molecule comprising a promoter operably linked to the isolated nucleic acid molecule of claim 1.
3. An expression vector comprising the nucleic acid molecule of claim 2.
4. The vector of claim 3, wherein the vector is pJL-OFP-35S::*GmSAMT1* (SEQ ID NO: 3) or pJL-OFP-35S::*GmSABP2-1* (SEQ ID NO: 4).
5. An isolated host cell transformed with the vector of claim 3 or claim 4.
6. A construct comprising isolated nucleic acid molecule claim 1 operably linked to a promoter.
7. The construct of claim 6, wherein the construct confers an agronomic trait to a plant in which it is expressed.
8. The construct of claim 7, wherein the agronomic trait comprises pest resistance.
9. The construct of claim 8, wherein pest resistance comprises resistance to soybean cyst nematode (*Heterodera glycines*, SCN).

10. A transgenic plant stably transformed with the construct of any one of claims 6-9.

11. A seed of the transgenic plant of claim 10, wherein the seed
5 comprises the construct.

12. The transgenic plant of claim 10, which plant is a soybean plant.

13. A method of producing a transgenic plant comprising transforming a
10 plant cell or tissue with the construct of any one of claims 6-9.

14. A method of enhancing pest resistance in a plant comprising
transforming a plant cell or tissue with the construct of any one of claims 6-9,
thereby enhancing pest resistance in the plant.

15

15. The method of claim 13 or 14, wherein the plant is a dicotyledon.

16. The method of claim 13 or 14, wherein the plant is a monocotyledon.

17. The method of claim 14, wherein pest resistance comprises resistance
20 to soybean cyst nematode (*Heterodera glycines*, SCN).

18. The transgenic plant of claim 14, which plant is a soybean plant.

19. A plant cell or tissue transformed with the construct of any one of
25 claims 6-9.

20. The plant cell or tissue of claim 19, wherein the plant cell or tissue is
from a dicotyledon.

30

21. The plant cell or tissue of claim 19, wherein the plant cell or tissue is
derived from a monocotyledon.

22. A plant cell, fruit, leaf, root, shoot, flower, seed, cutting and other reproductive material useful in sexual or asexual propagation, progeny plants inclusive of F1 hybrids, male-sterile plants and all other plants and plant products
5 derivable from the transgenic plant of claim 10.

23. A method of producing a commodity plant product, comprising obtaining the plant of 10 or a part thereof, wherein the commodity plant product is protein concentrate, protein isolate, soybean hulls, meal, flour or oil and producing
10 the commodity plant product therefrom.

FIG. 1

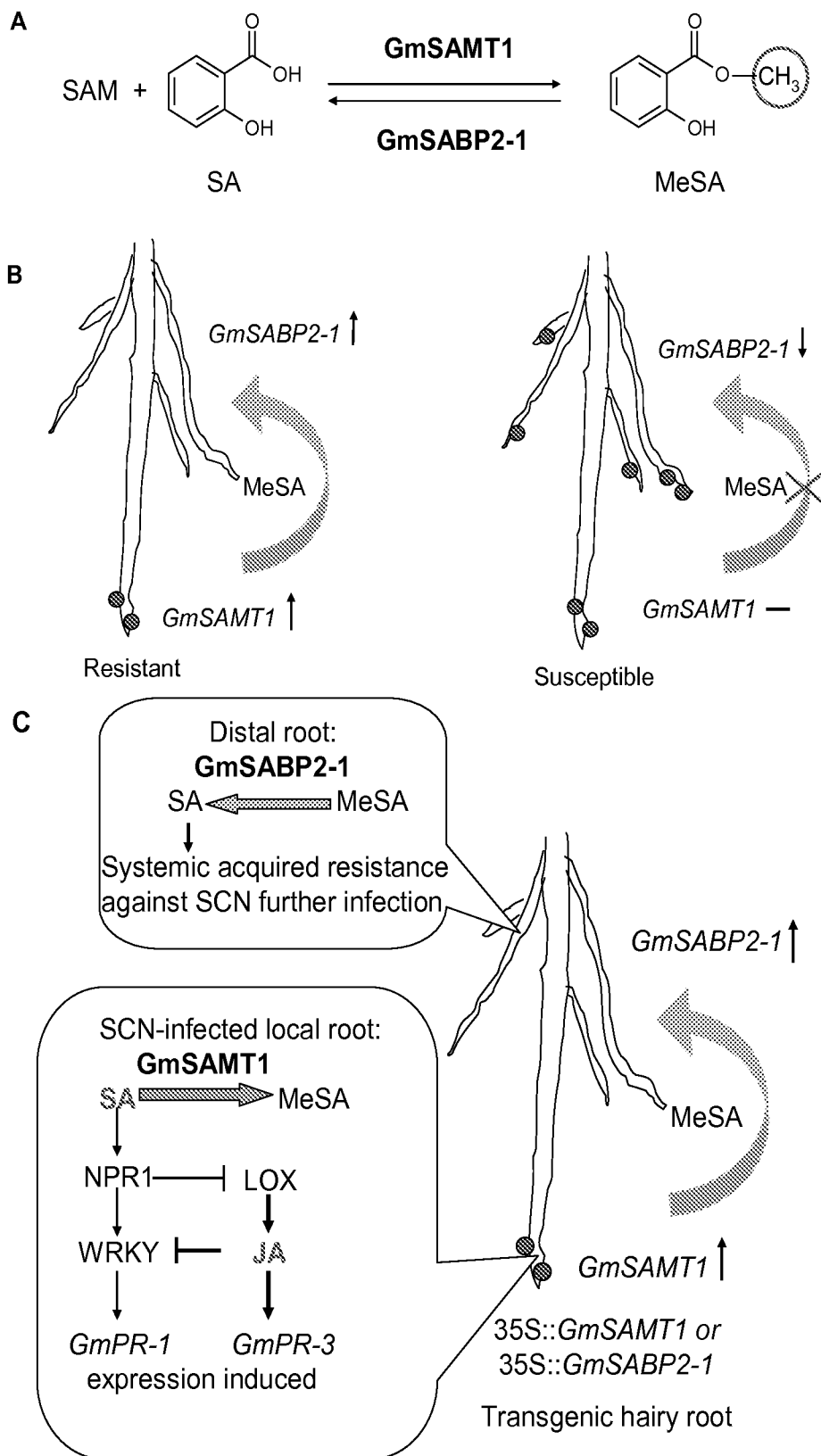


FIG. 3

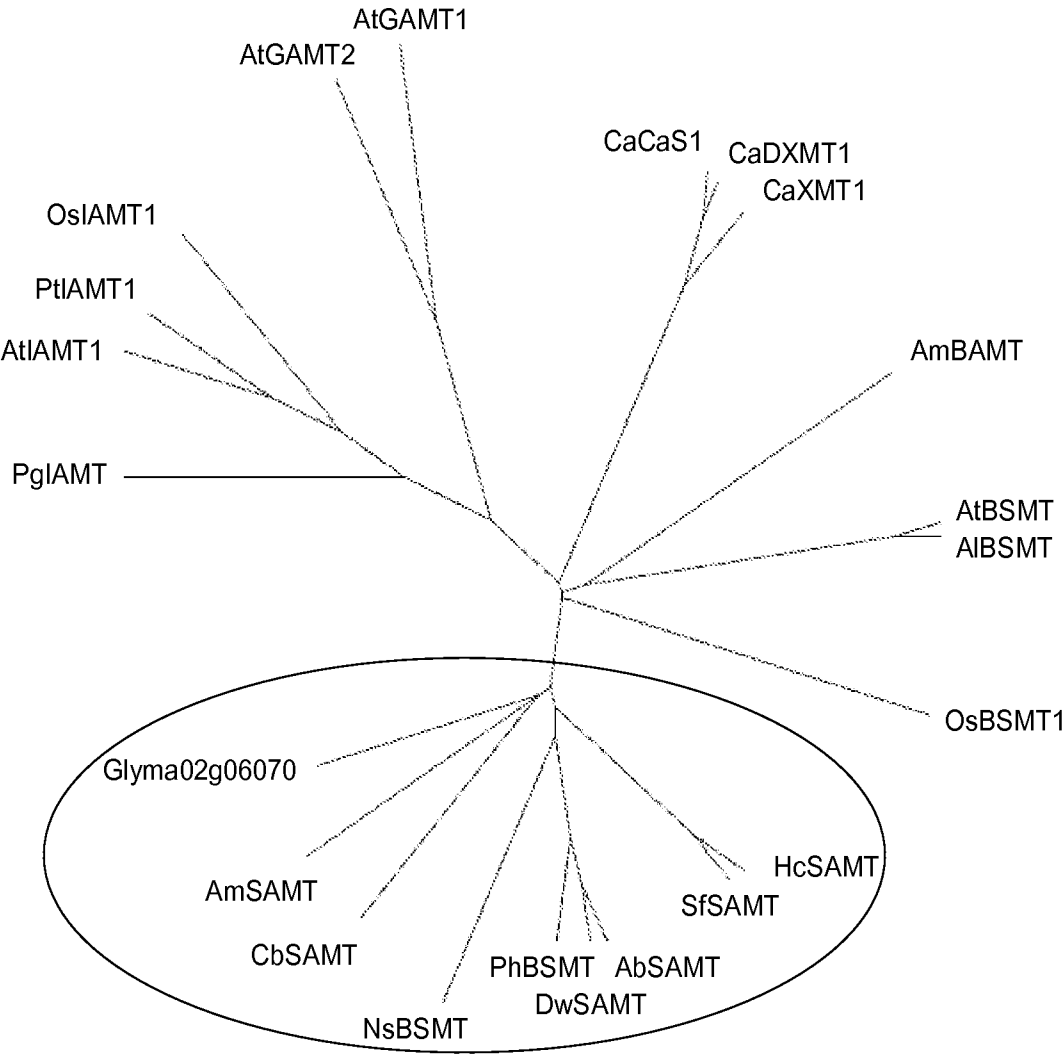


FIG. 4

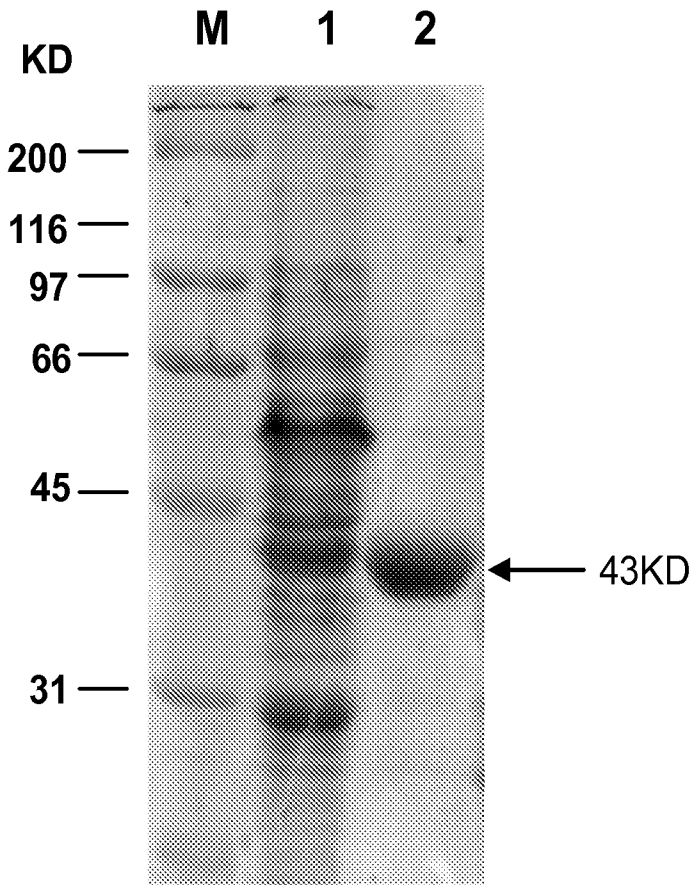


FIG. 5

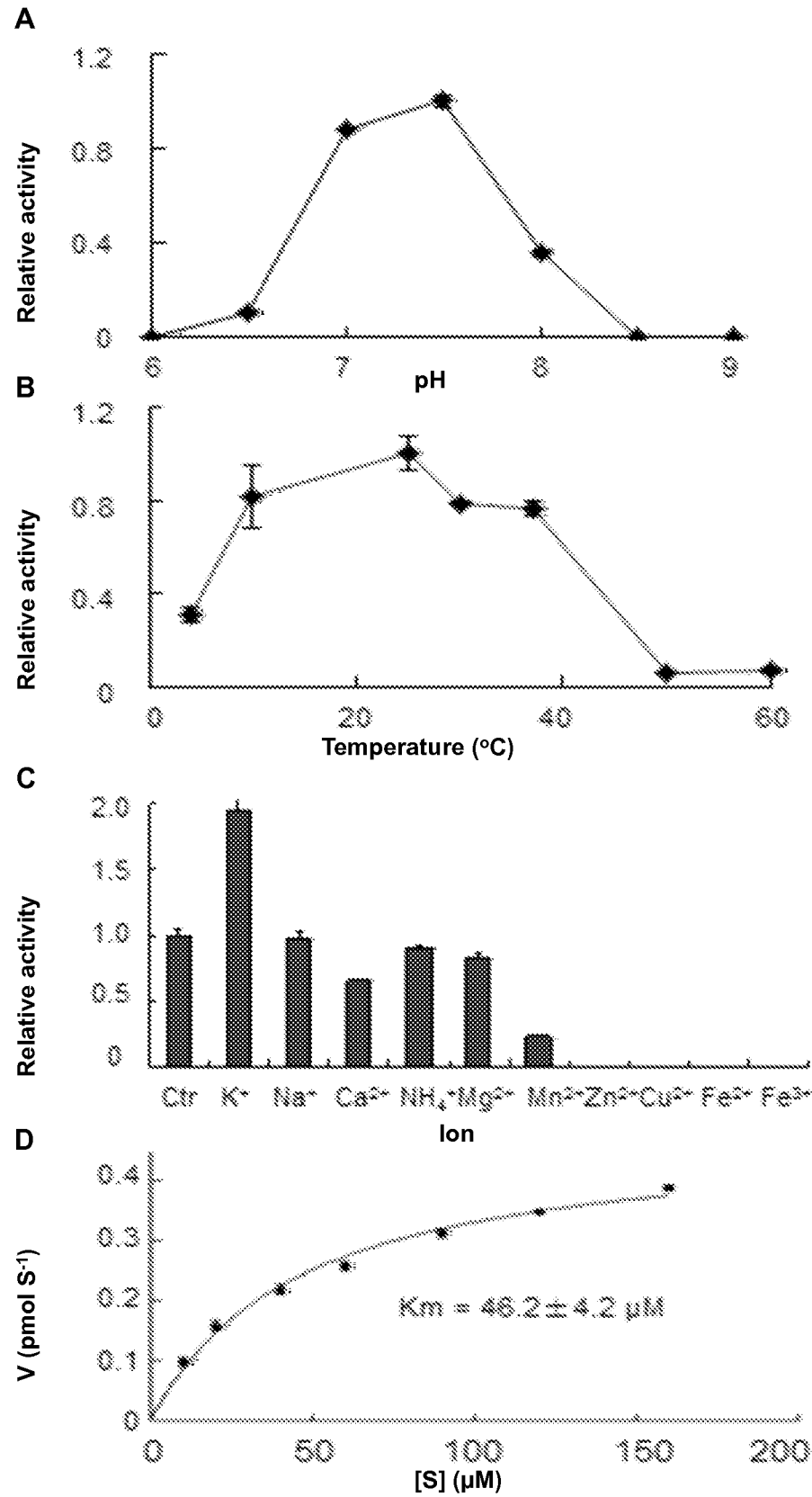


FIG. 6

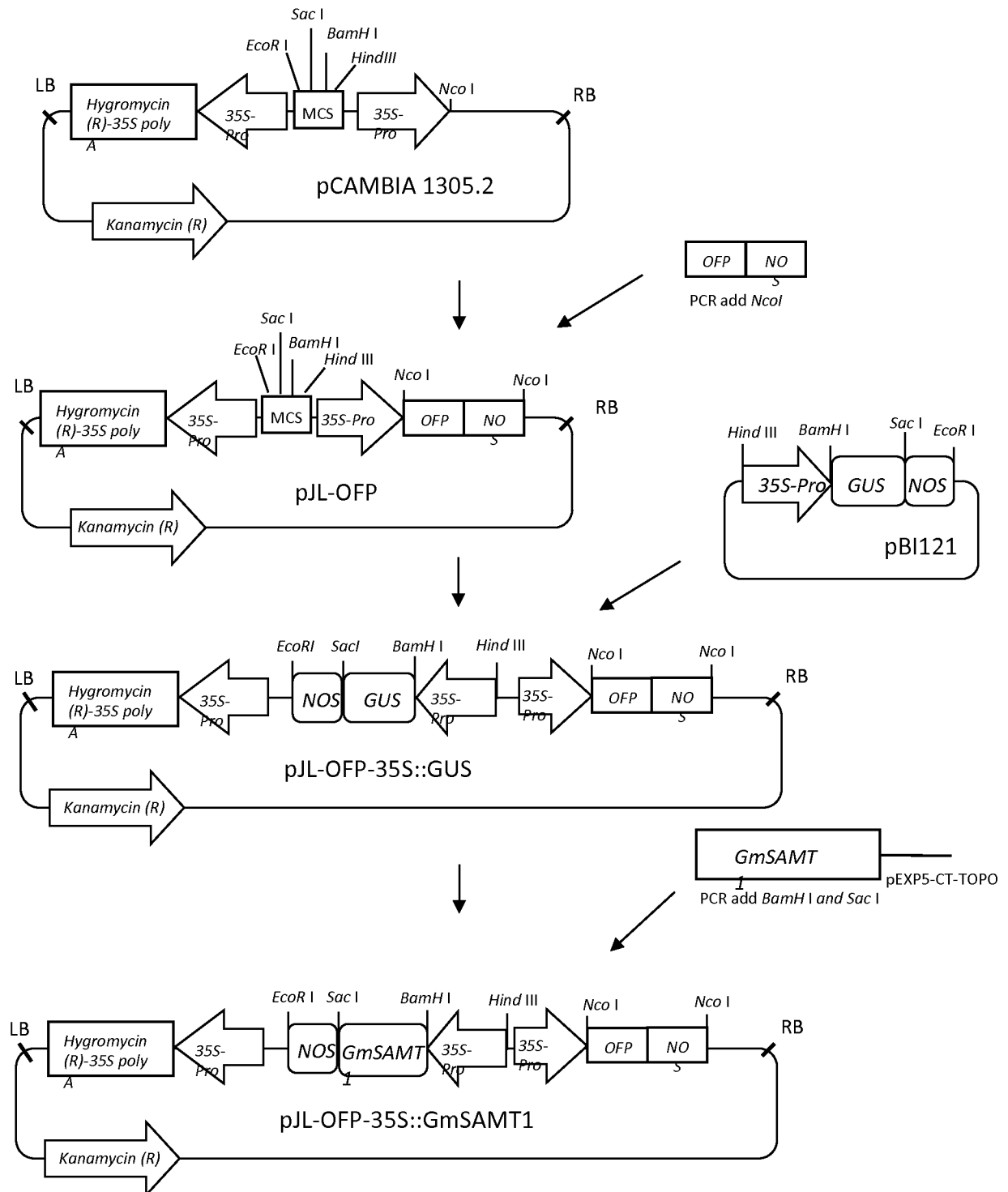


FIG. 7

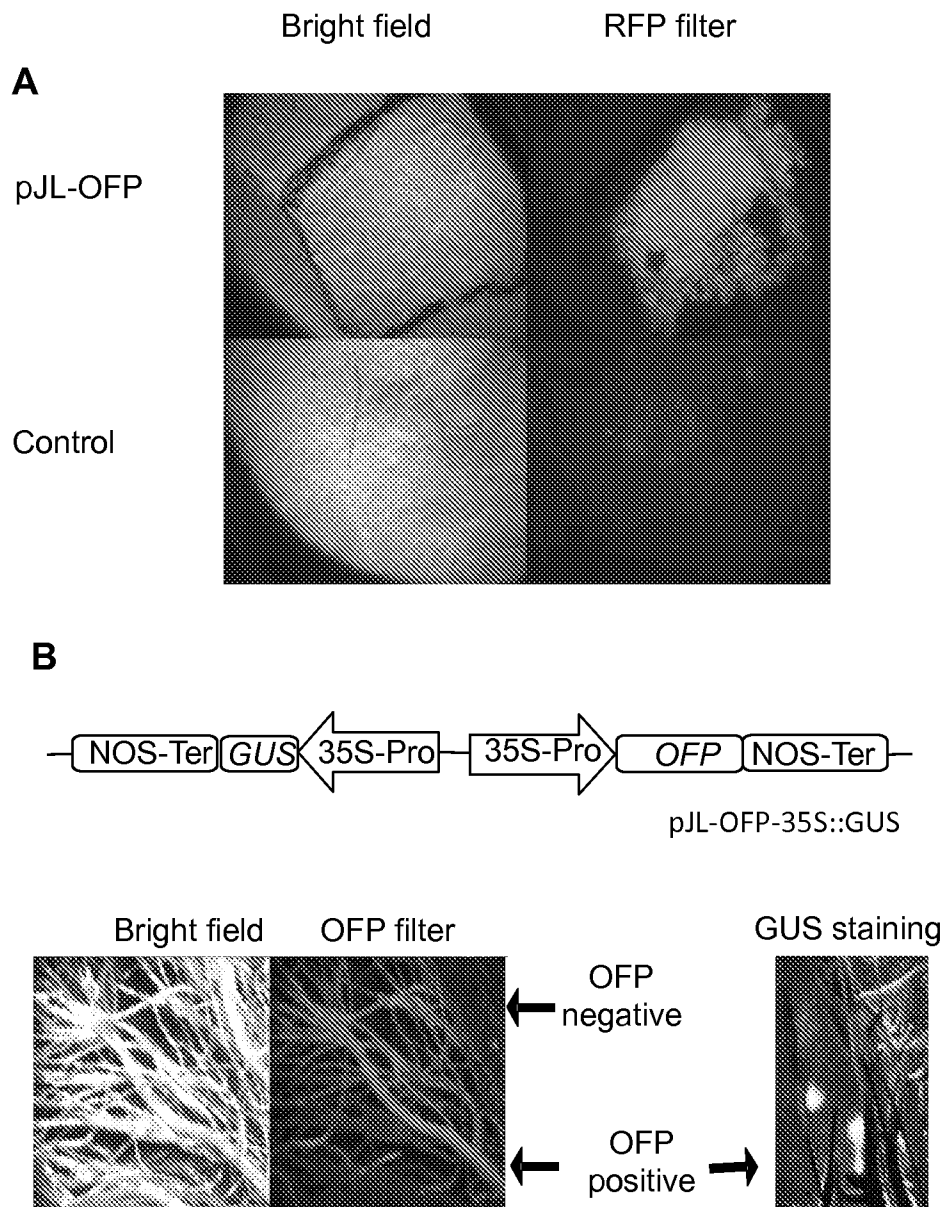


FIG. 8

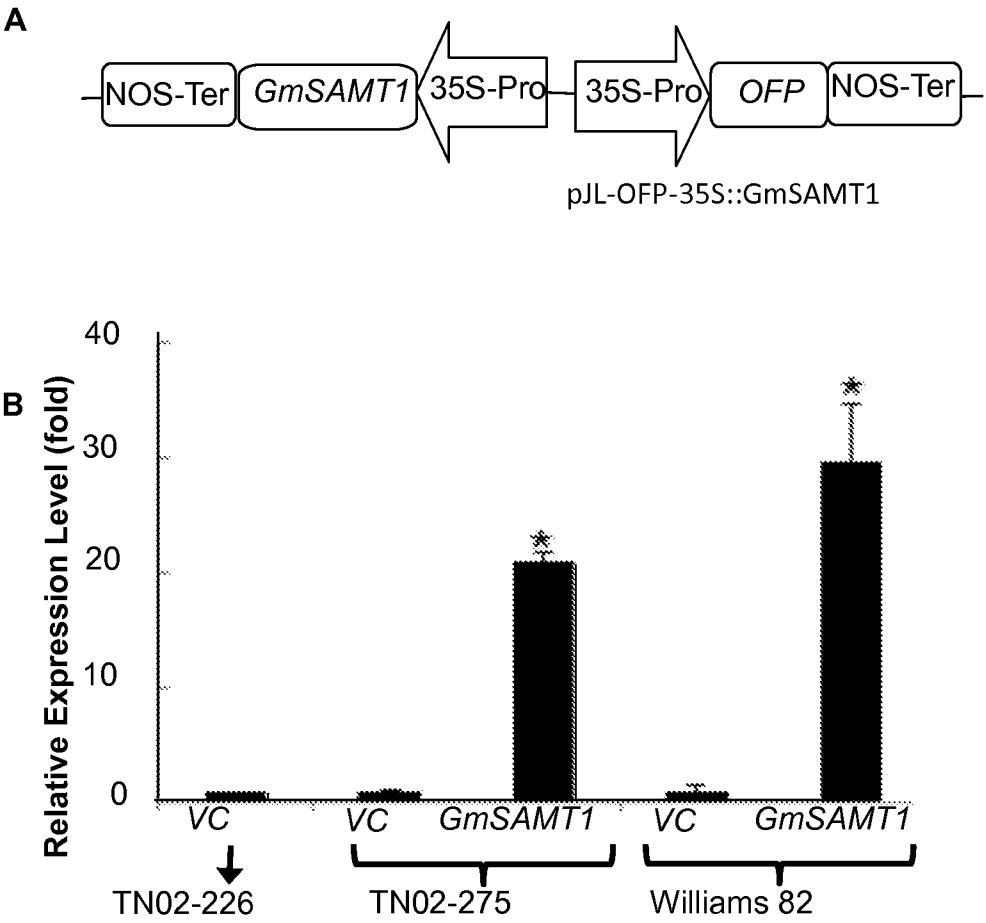
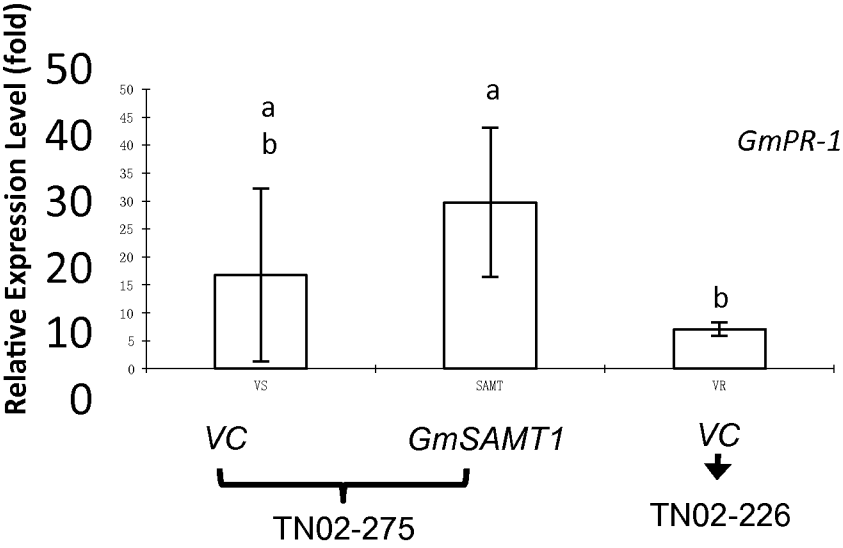


FIG. 9

A



B

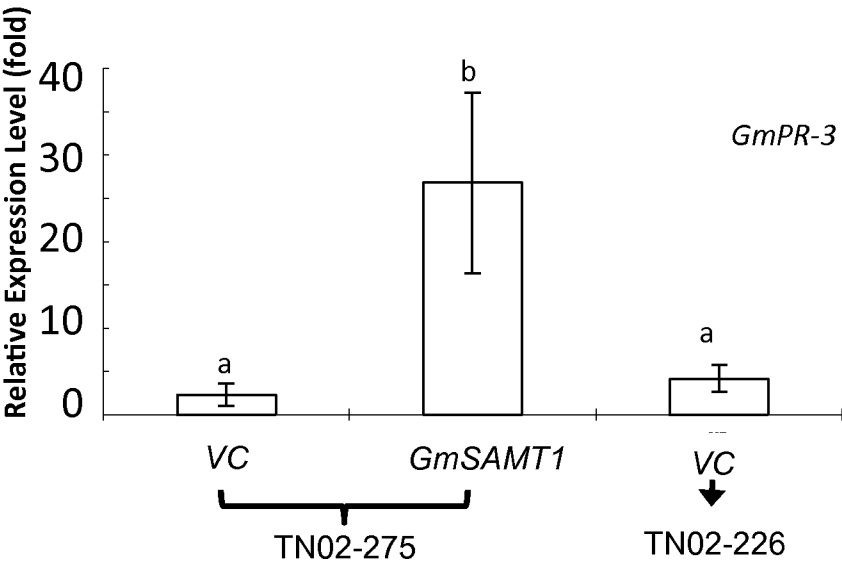


FIG. 10

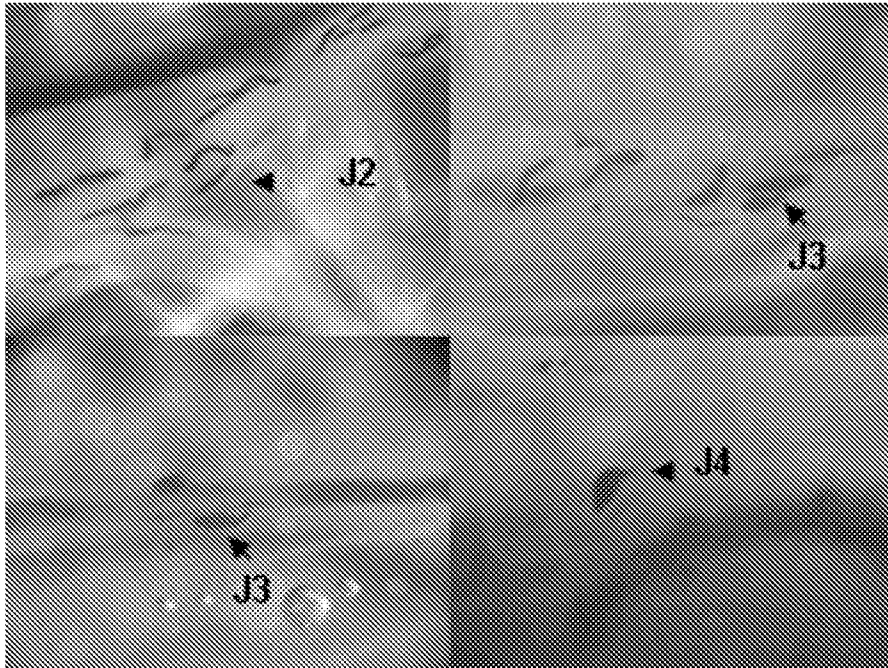


FIG. 11

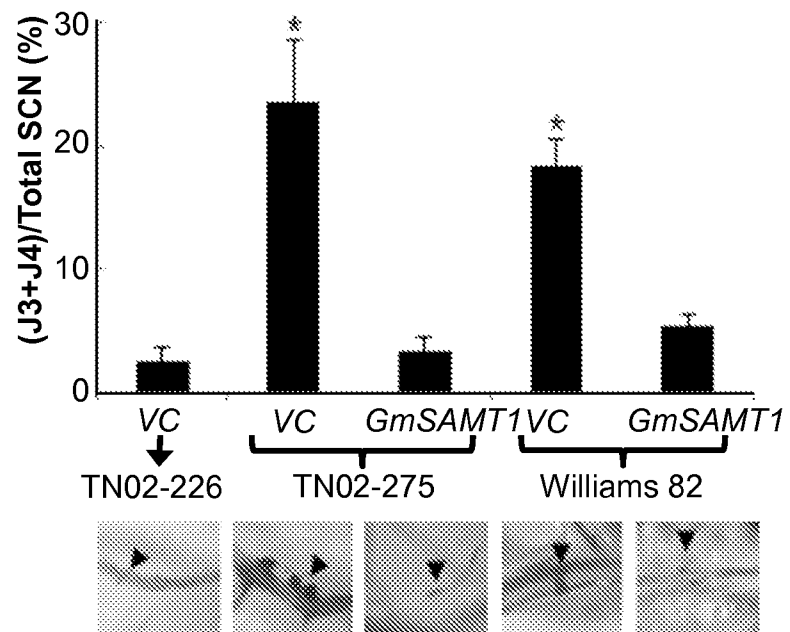


FIG. 12

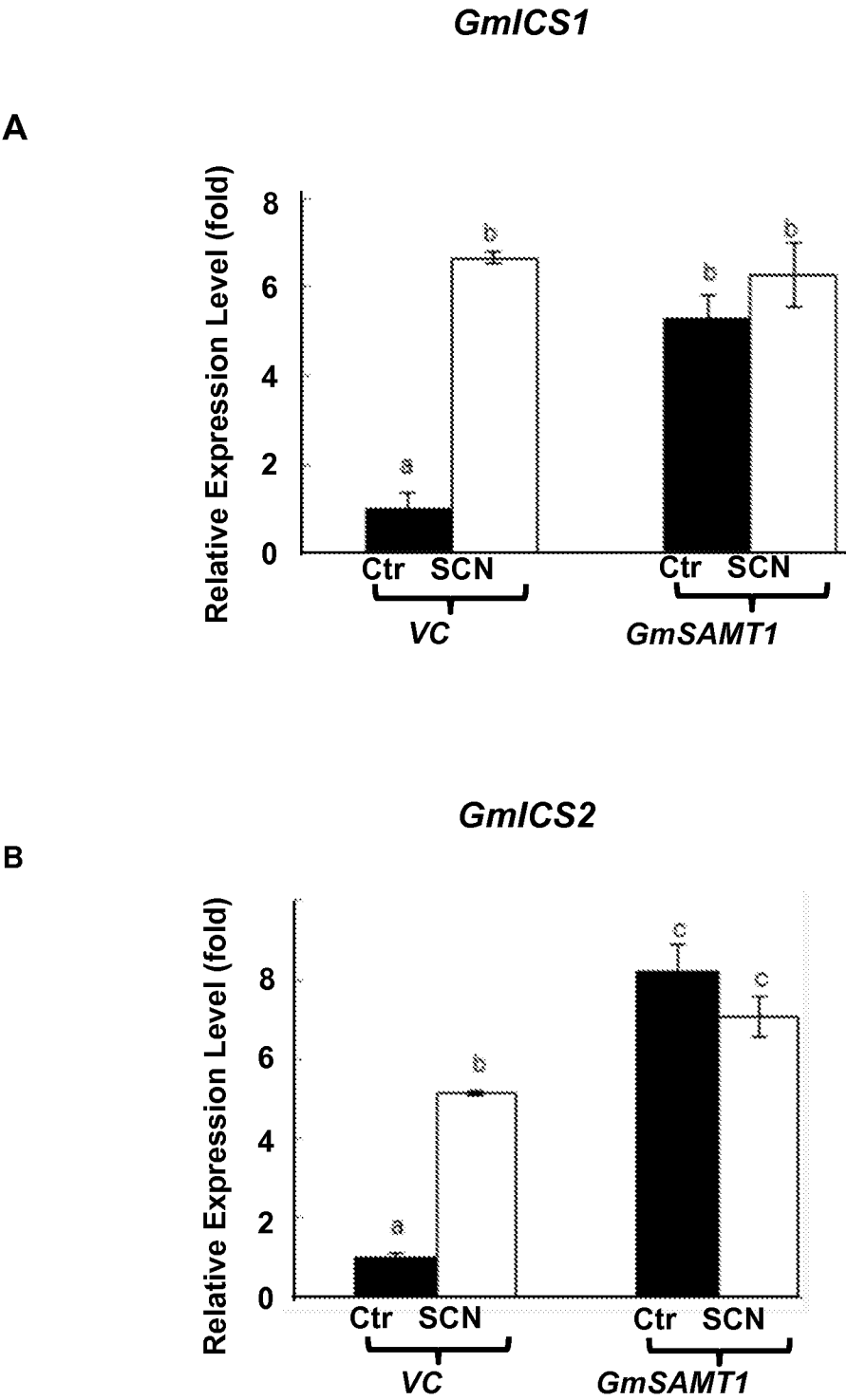


FIG. 13

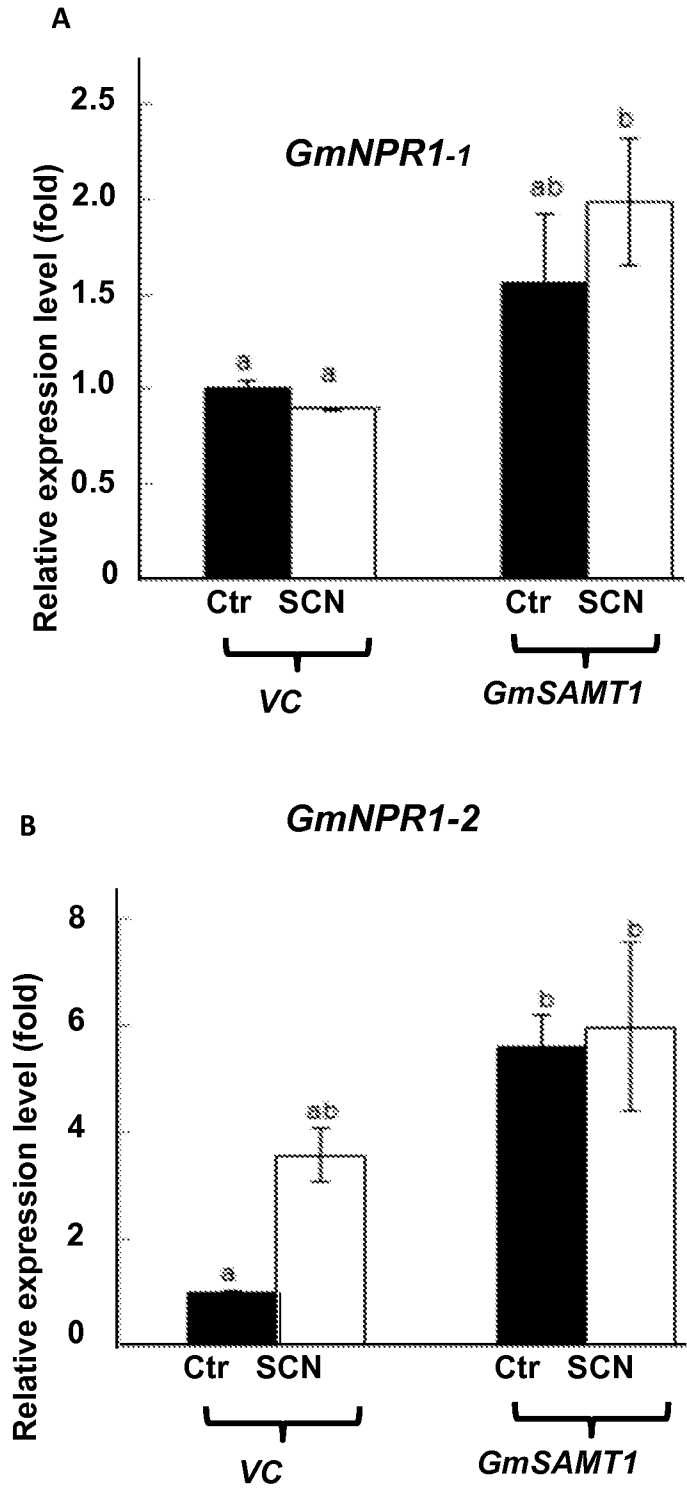


FIG. 14

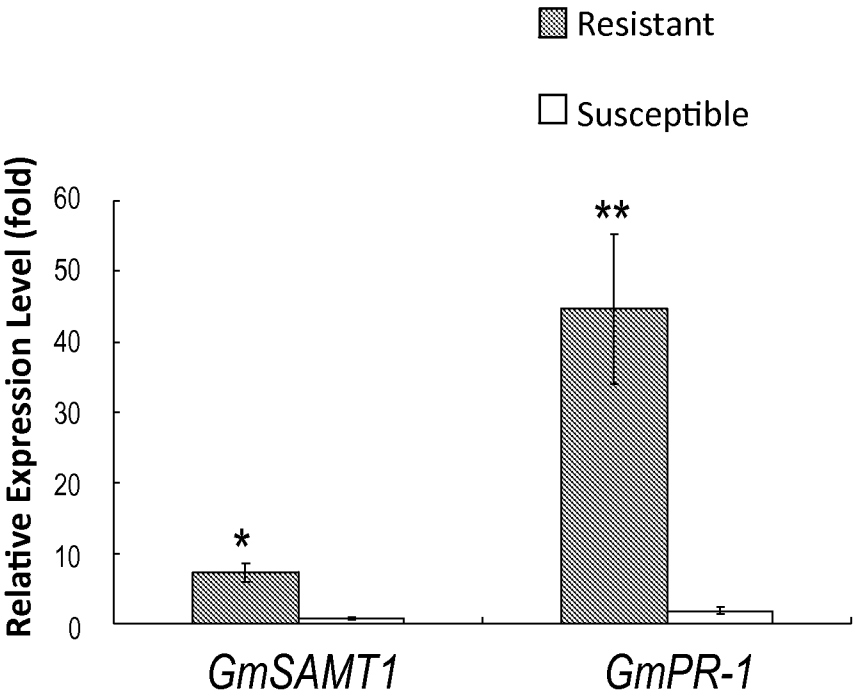


FIG. 15

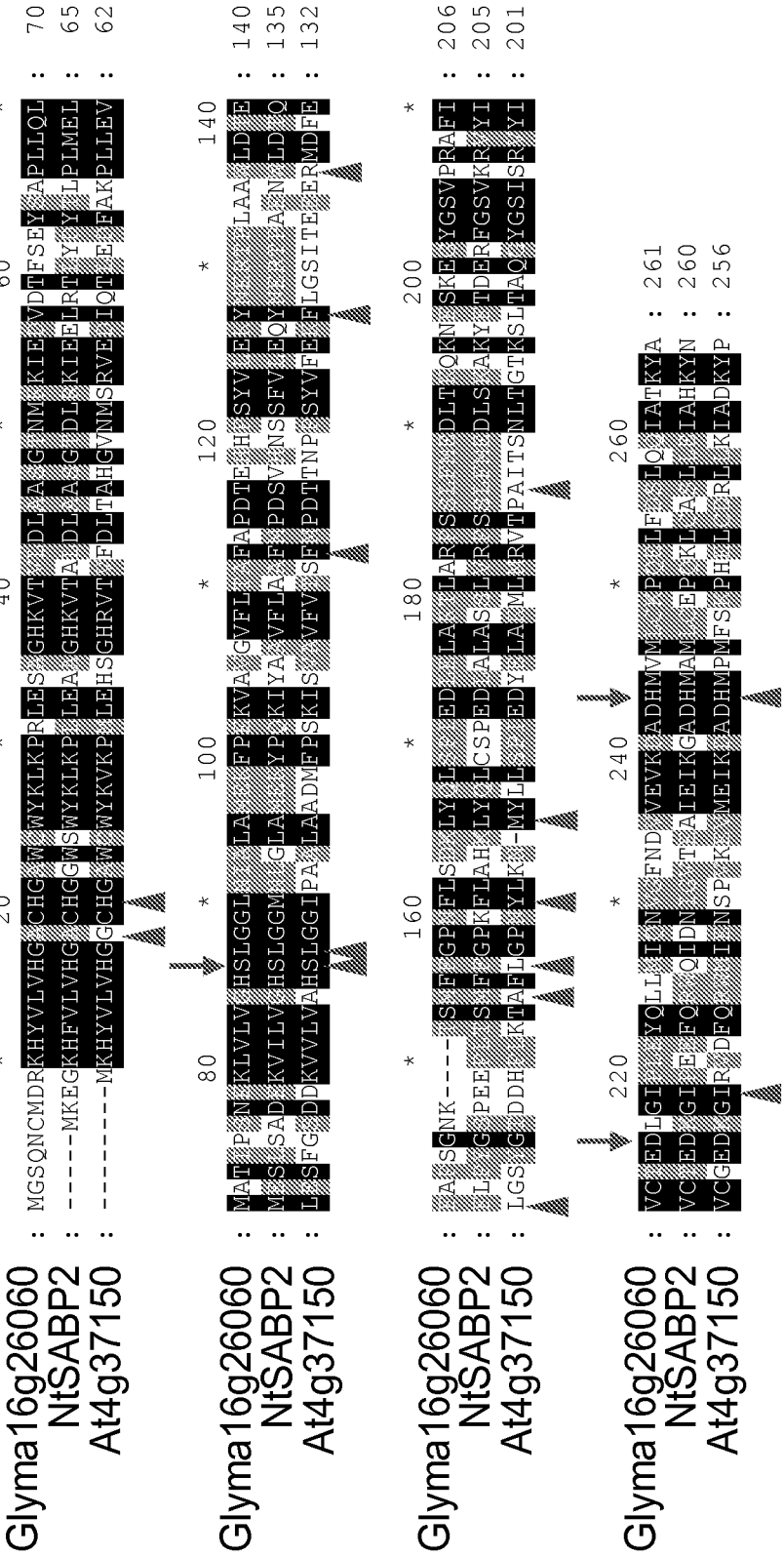


FIG. 16

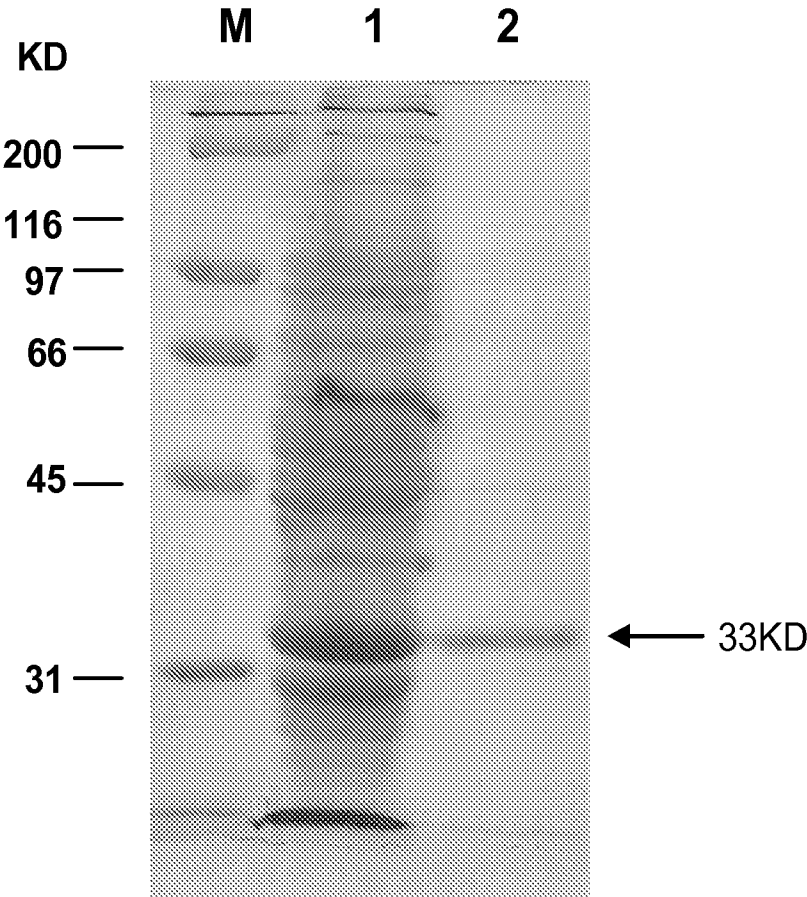


FIG. 17

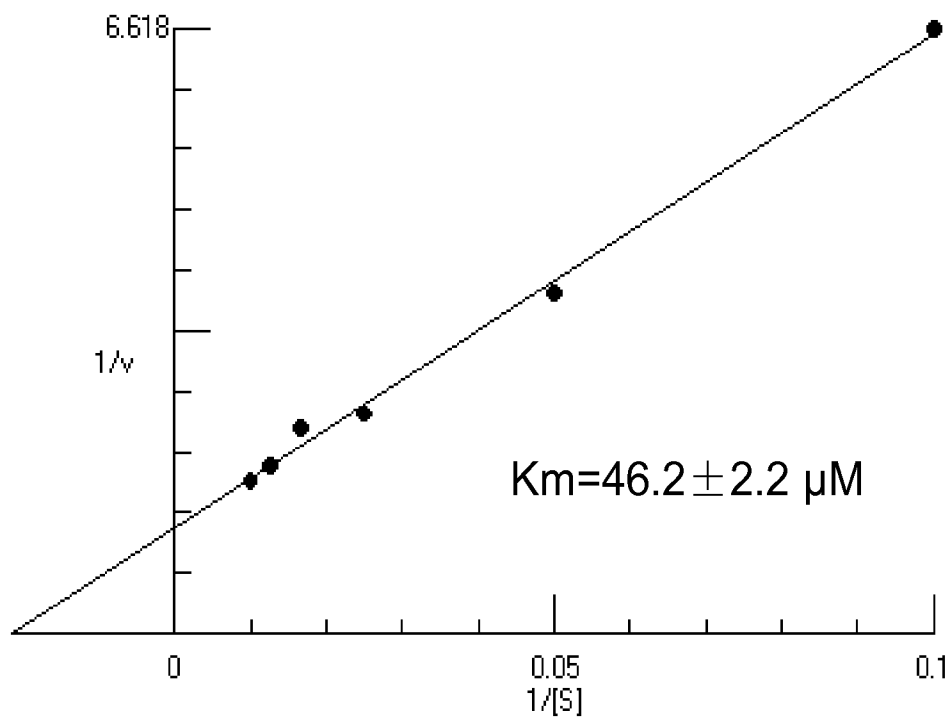


FIG. 18

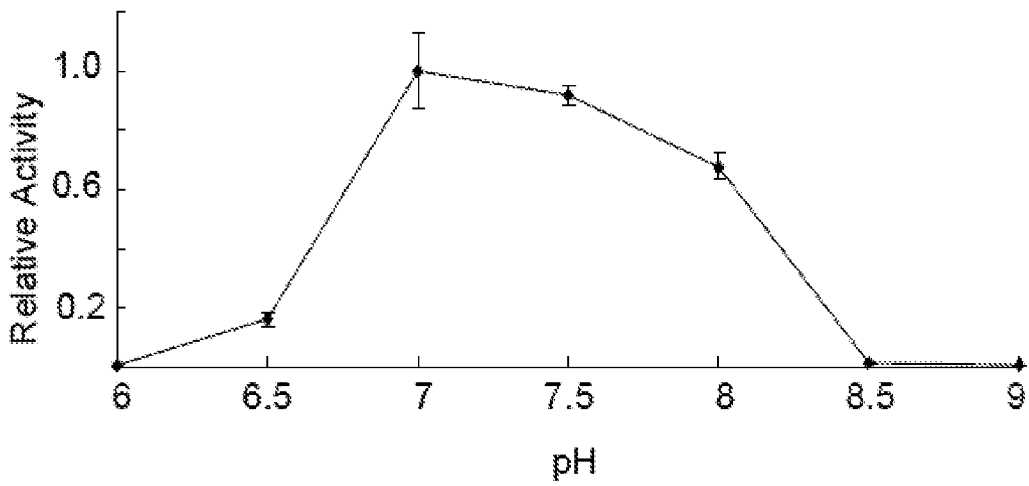


FIG. 19

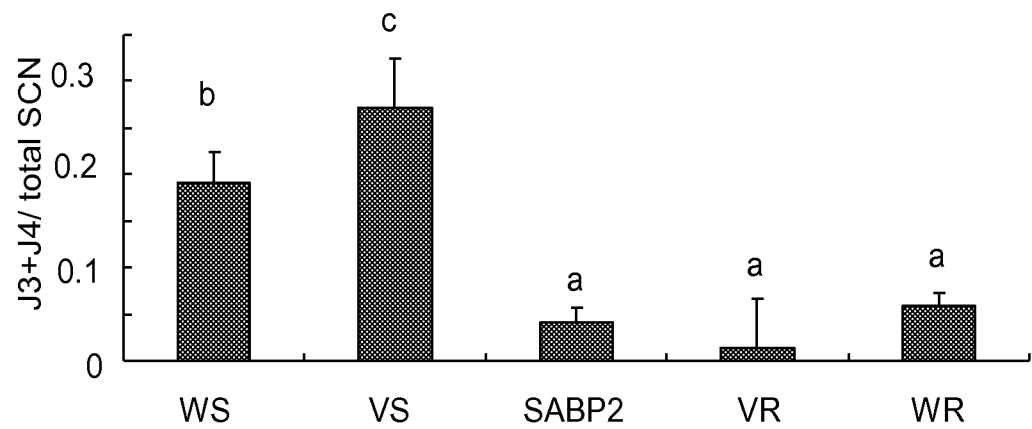


FIG. 20

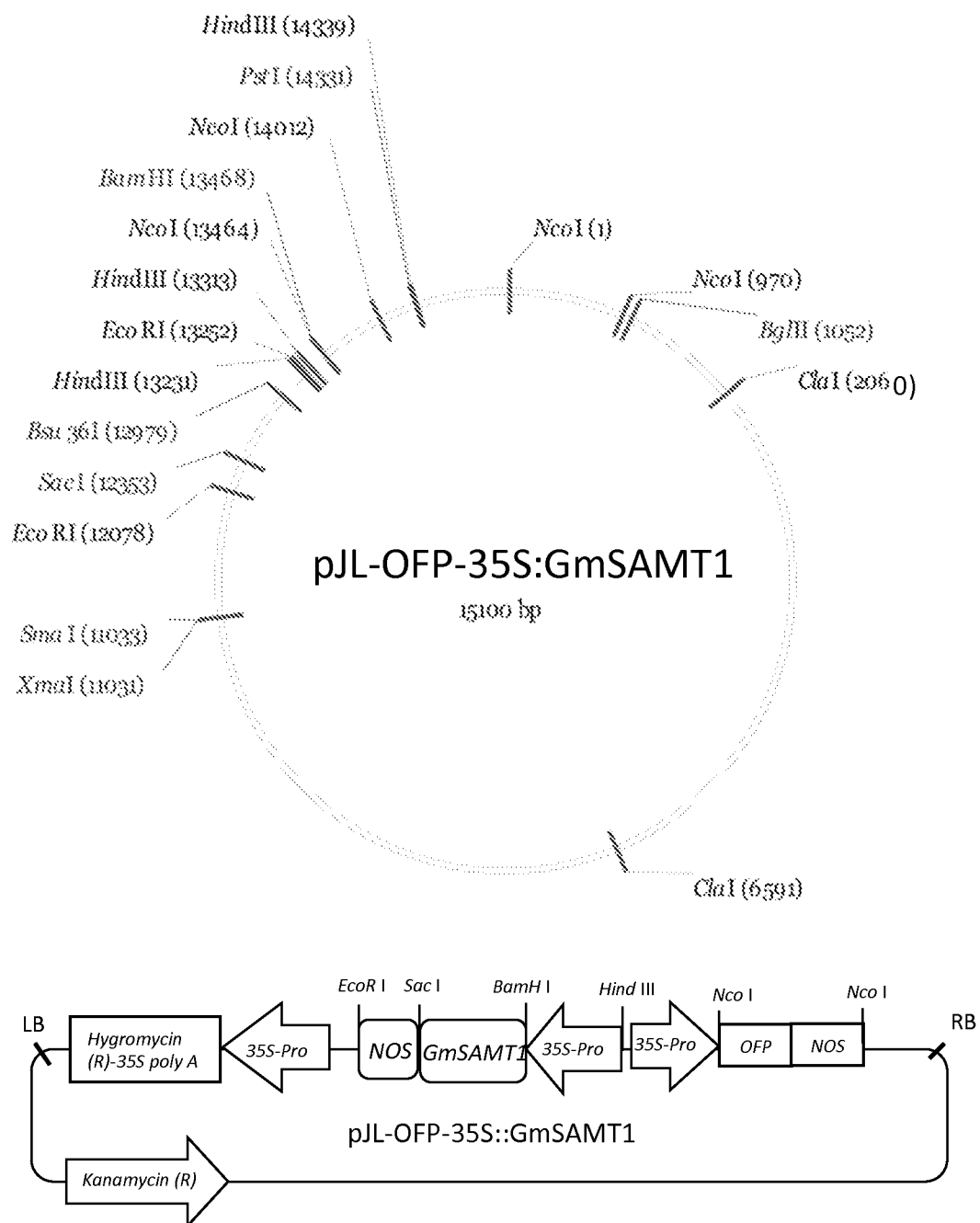
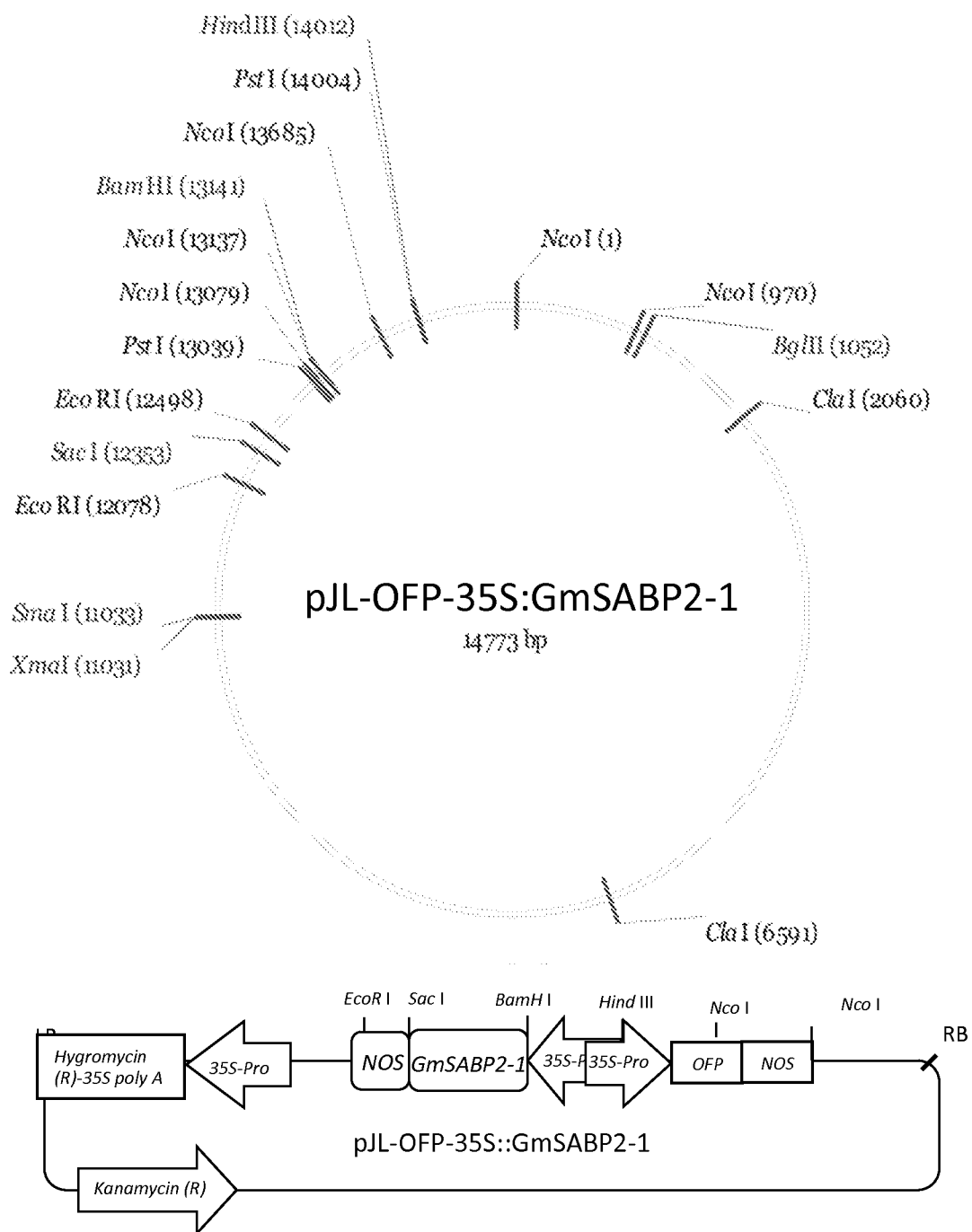


FIG. 21



INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/62759

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C07H 21/04, C12N 15/82 (2012.01)

USPC - 536/23.2, 536/23.6, 435/320

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

USPC -- 536/23.2, 536/23.6, 435/320, 536/23.1, 536/22.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Pat Base; PubWest -- PGPB,USPT,USOC,EPAB,JPAB; USPTO Web Page; Google Scholar; Search terms -- polynucleotide sequence, pest resistance, soybean cyst nematode, transgenic, dicot, monocot, transgenic seed, asexual propagation, stable transformation, pJL-OGP-35S::GMS ABP2-1 expression vector, promoter, progeny, transgenic plant tissues

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X -- Y	US 2003/0233670 A1 (EDGERTON et al.) 18 December 2003 (18.12.2003) para [0007], [0008], [0018], [0019], [0021], [0023], [0026], [0051]-[0053], [0075], [0084], [0090], [0091], [0095], [0096], [0423], [0578], [0603], [0610], [0628], [0630], [0659], [0670], [0672]; abstract SEQ ID NO: 160	1-3, 5/(3), 6-8, 10/(6-8), 11-12/(10/(6-8)), 13-14/(6-8), 18/(14/6-8), 19/(6-8), 20-21/(19/(6-8)) 4, 5/(4), 9, 10/(9), 11-12/(10/(9)), 13-14/(9), 17, 18/(14/9), 19/(9), 20-23
Y	US 2004/0248304 A1 (HU et al.) 09 December 2004 (09.12.2004) para [0010], [0013], [0100], [0135], [0136], [0138], [0151], [0184], [0218]	9, 10/(9), 11-12/(10/(9)), 13-14/(9), 17, 18/(14/9), 19/(9), 20-21/(19/(9)), 22, 23
Y	EMBL Accession Number AF354046, Binary vector pCAMBIA-1305.2, complete sequence. 17 April 2001 <http://www.ebi.ac.uk/Tools/dbfetch/dbfetch?db=embl&id=AF354046&f>, pages 1-5; pg 1, 2	4, 5/(4)
Y	US 2010/0154075 A1 (McCARTHY et al.) 17 June 2010 (17.06.2010) para [0014], [0015], [0087], [0097]	4, 5/(4)
Y	EMBL Accession Number EU818946, Glycine max salicylic acid methyl transferase-like protein (SABATH2) mRNA, complete cds. 16 July 2008, <http://www.ebi.ac.uk/Tools/dbfetch/dbfetch?db=embl&id=EU818946&f>, pages 1-3, pg 1	4, 5/(4)



Further documents are listed in the continuation of Box C.



* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

26 December 2012 (26.12.2012)

Date of mailing of the international search report

25 JAN 2013

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents
P.O. Box 1450, Alexandria, Virginia 22313-1450

Facsimile No. 571-273-3201

Authorized officer:

Lee W. Young

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/62759

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010/0064389 A1 (ASCENZI et al.) 11 March 2010 (11.03.2010) para [0002], [0044], [0047], [0055], abstract	1-14, 19-23

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 12/62759

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 15 and 16
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

FIG. 1

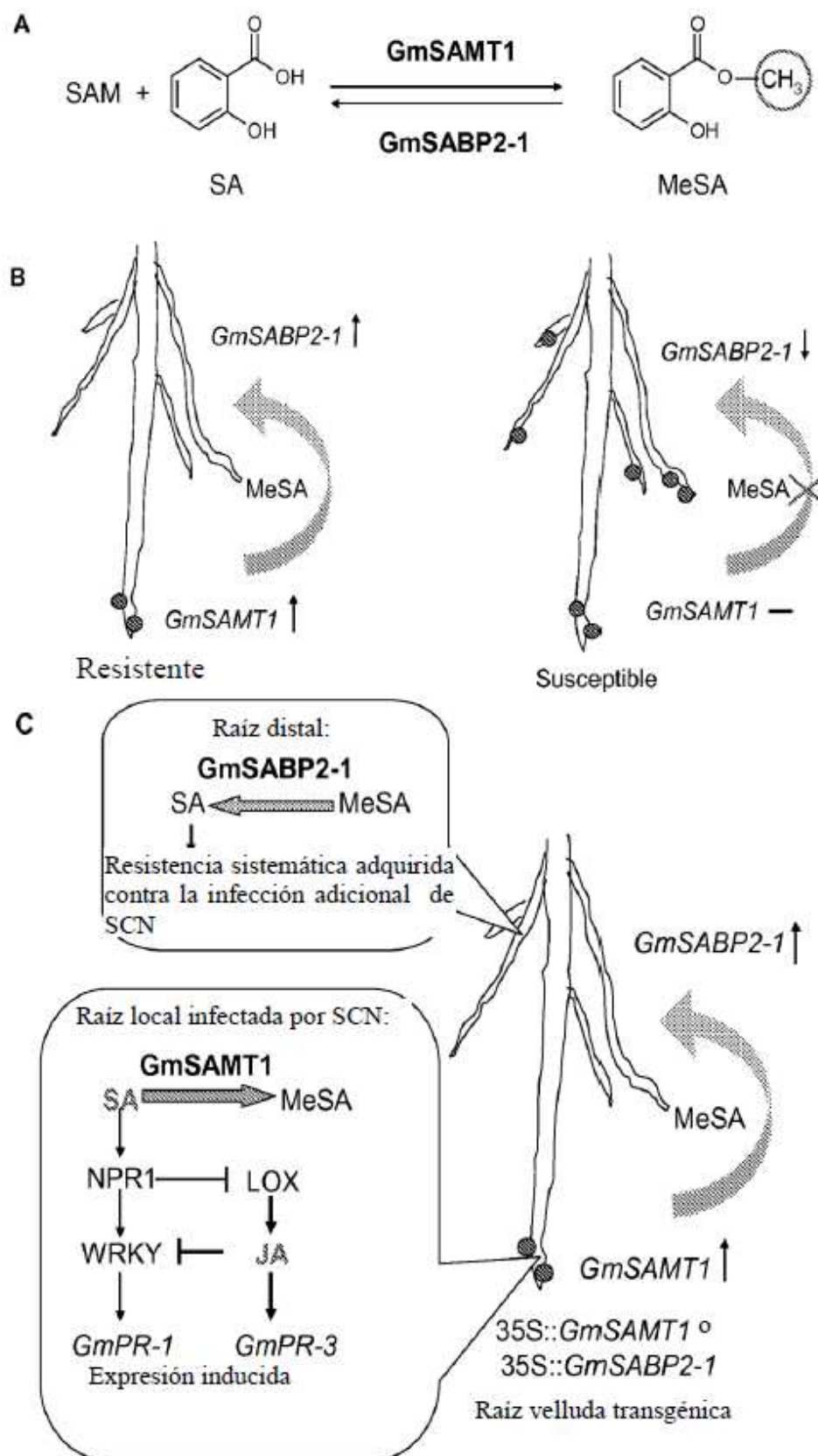


FIG. 2

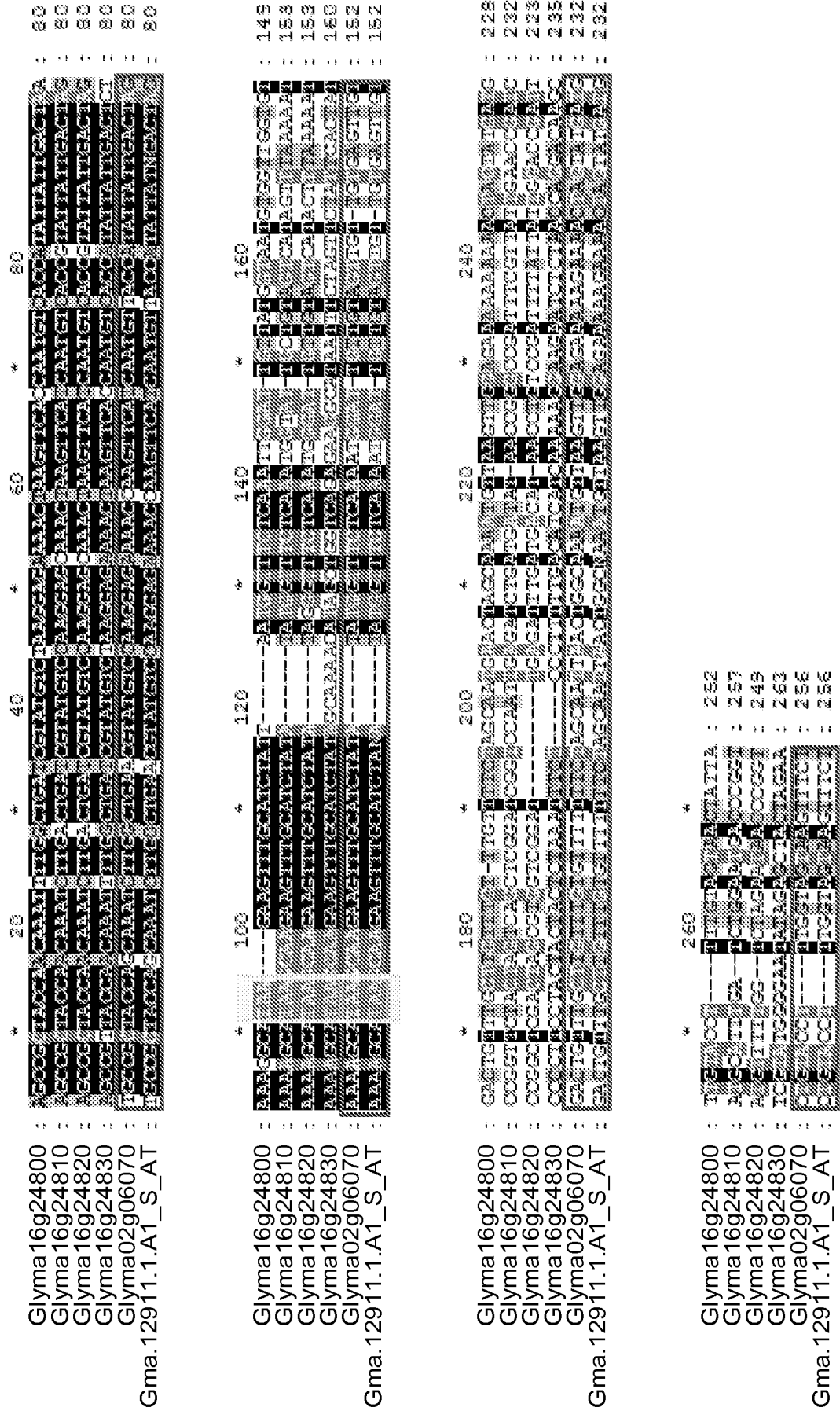


FIG. 3

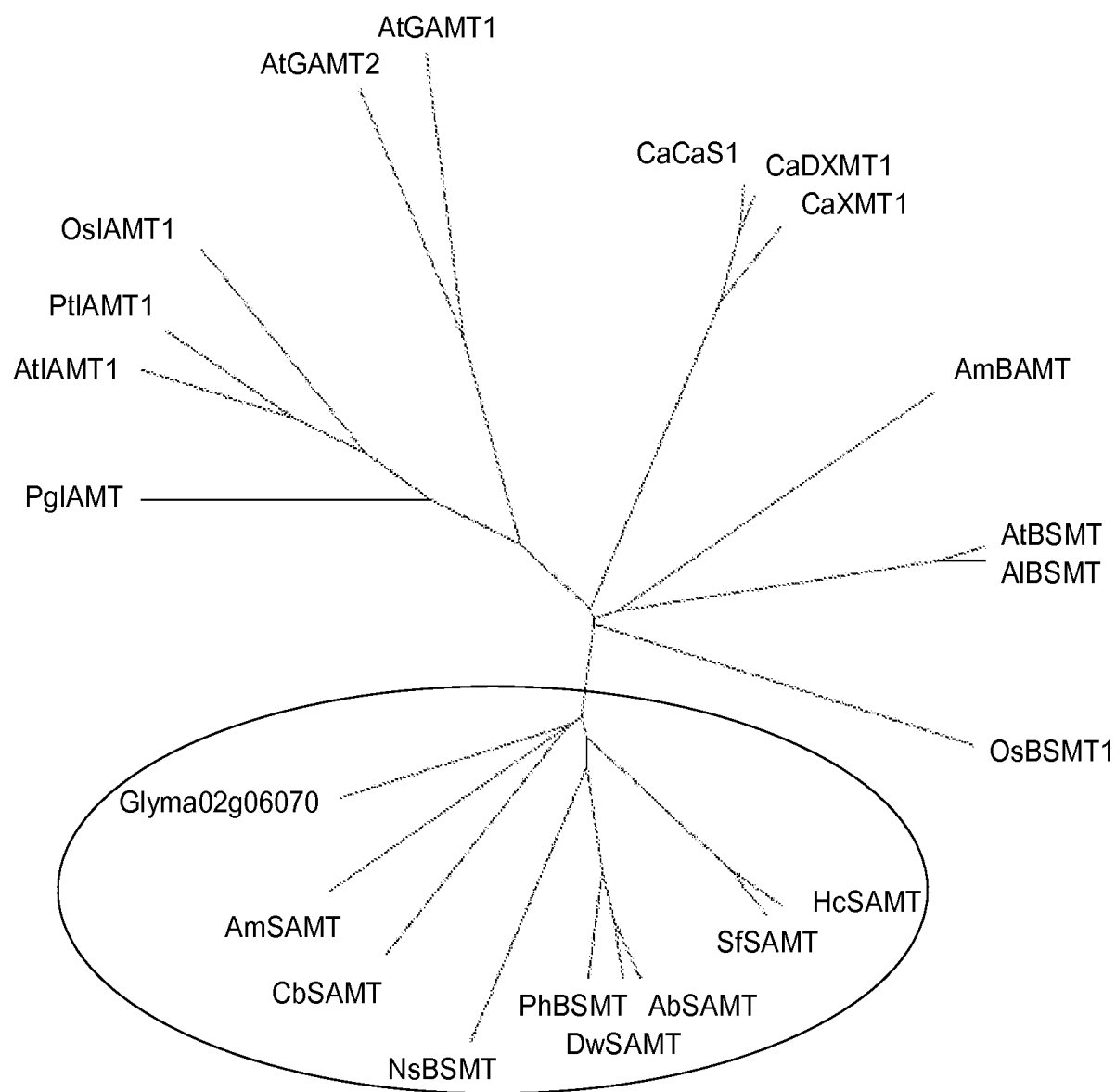


FIG. 4

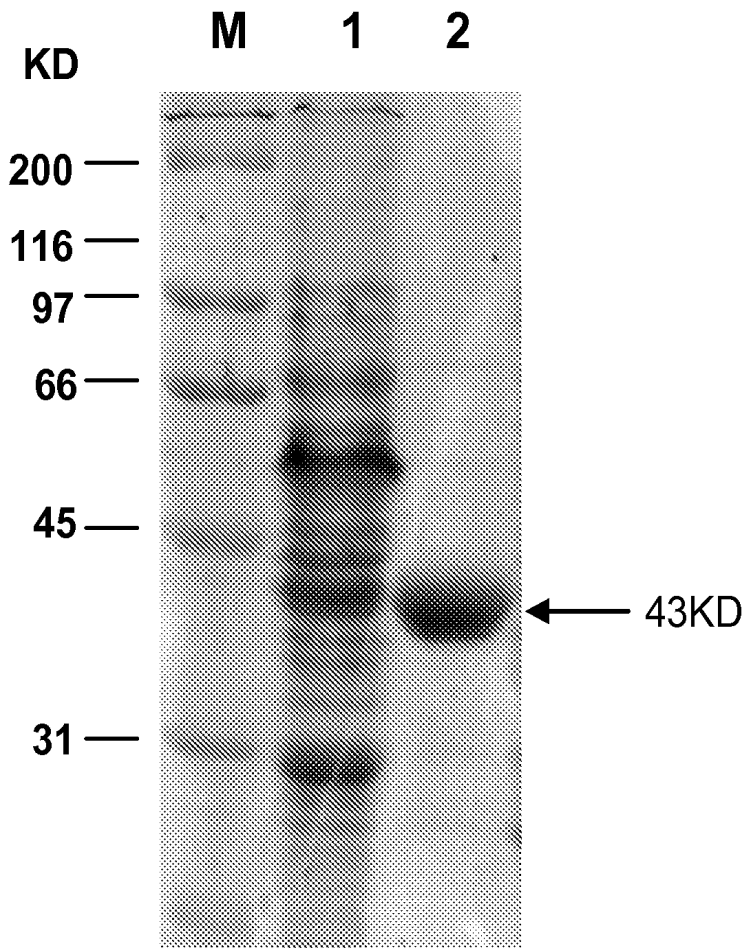


FIG. 5

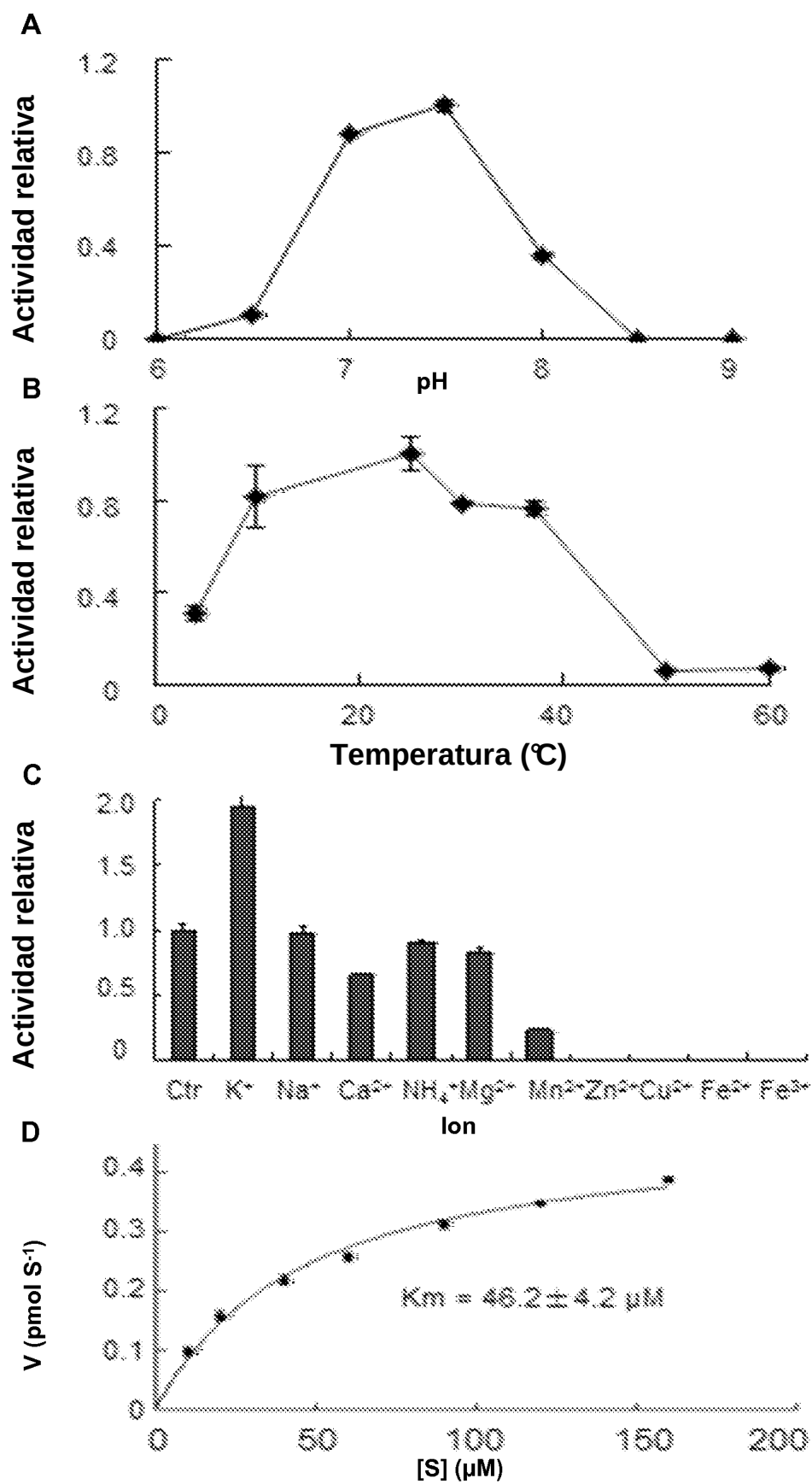


FIG. 6

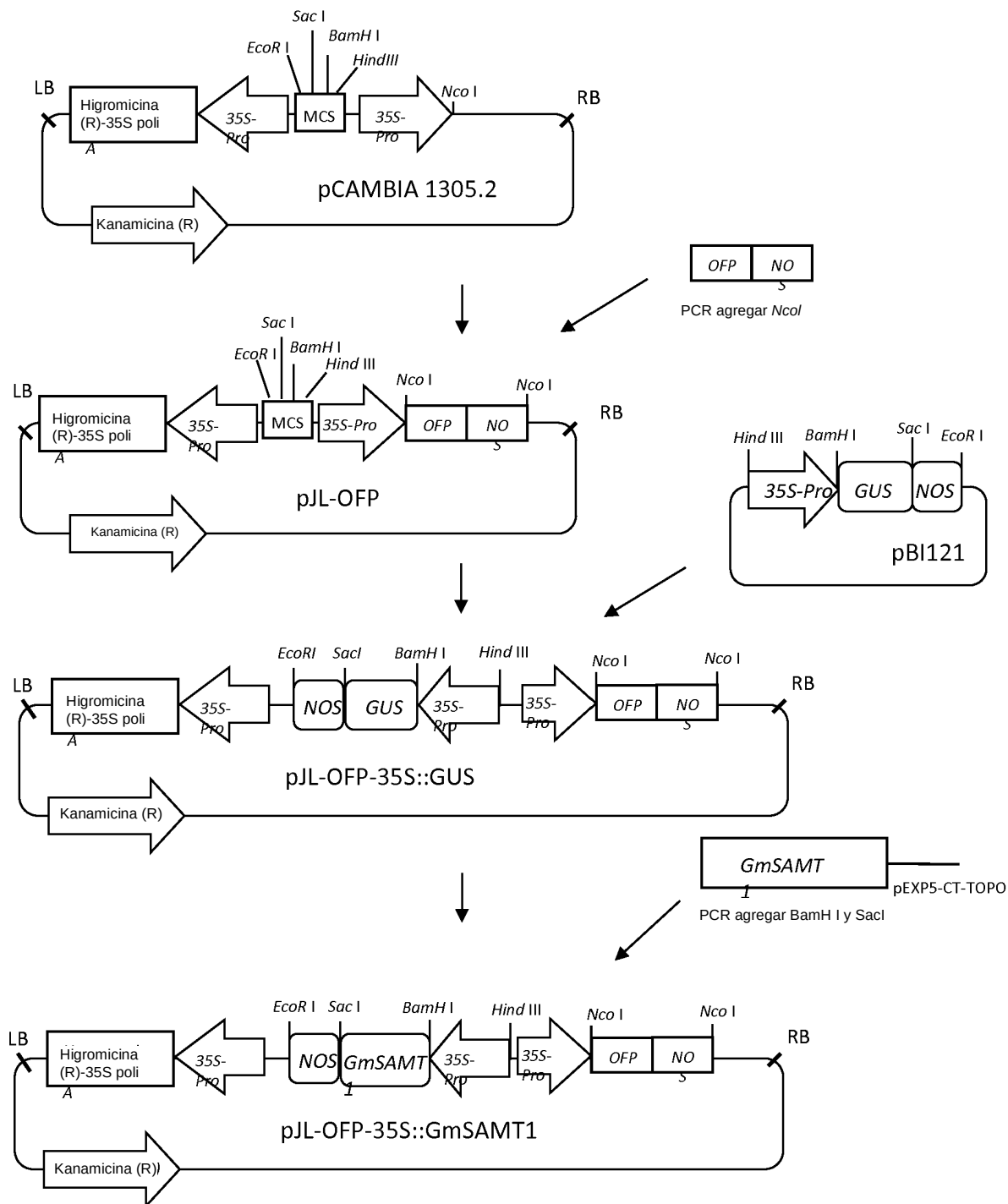


FIG. 7

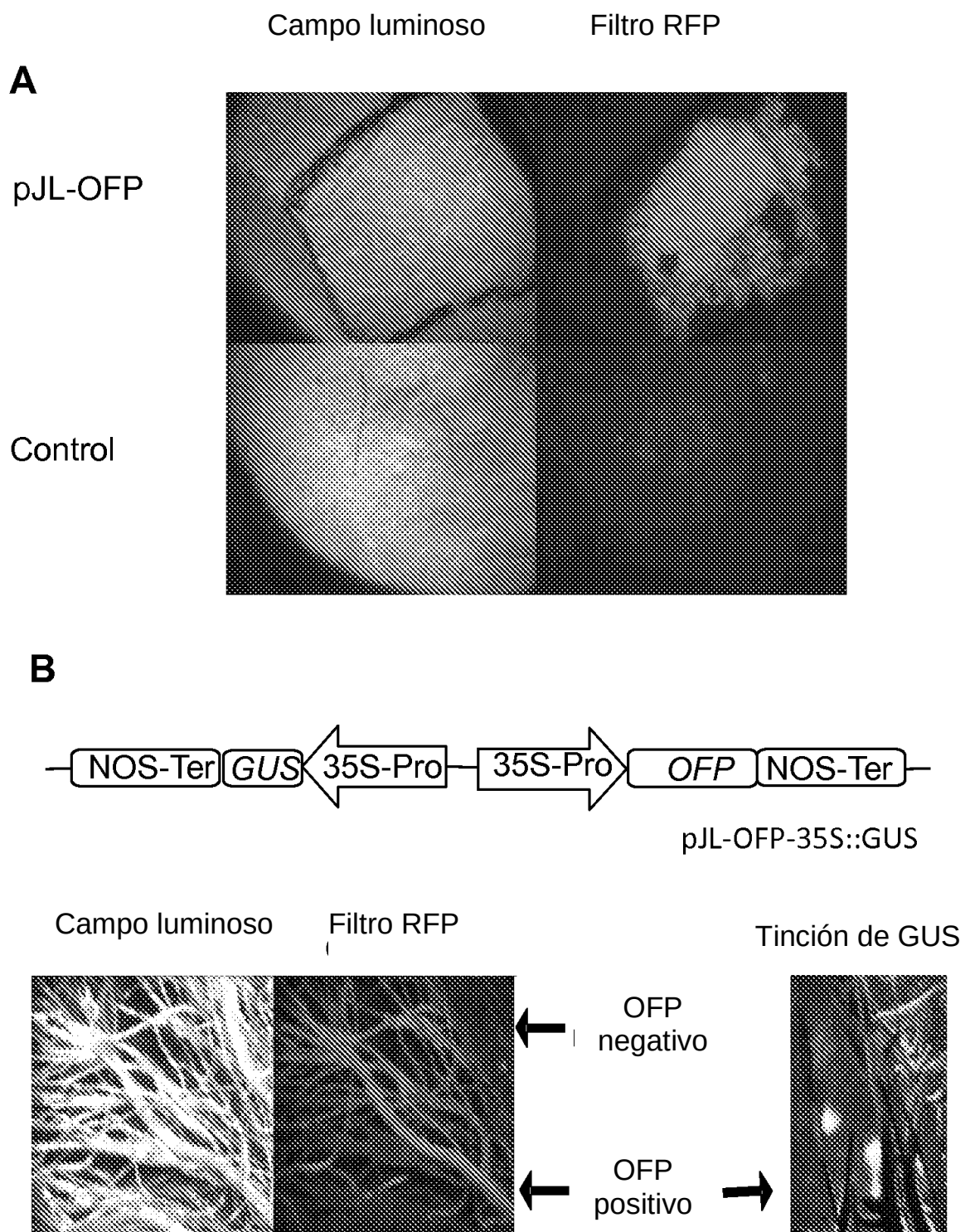


FIG. 8

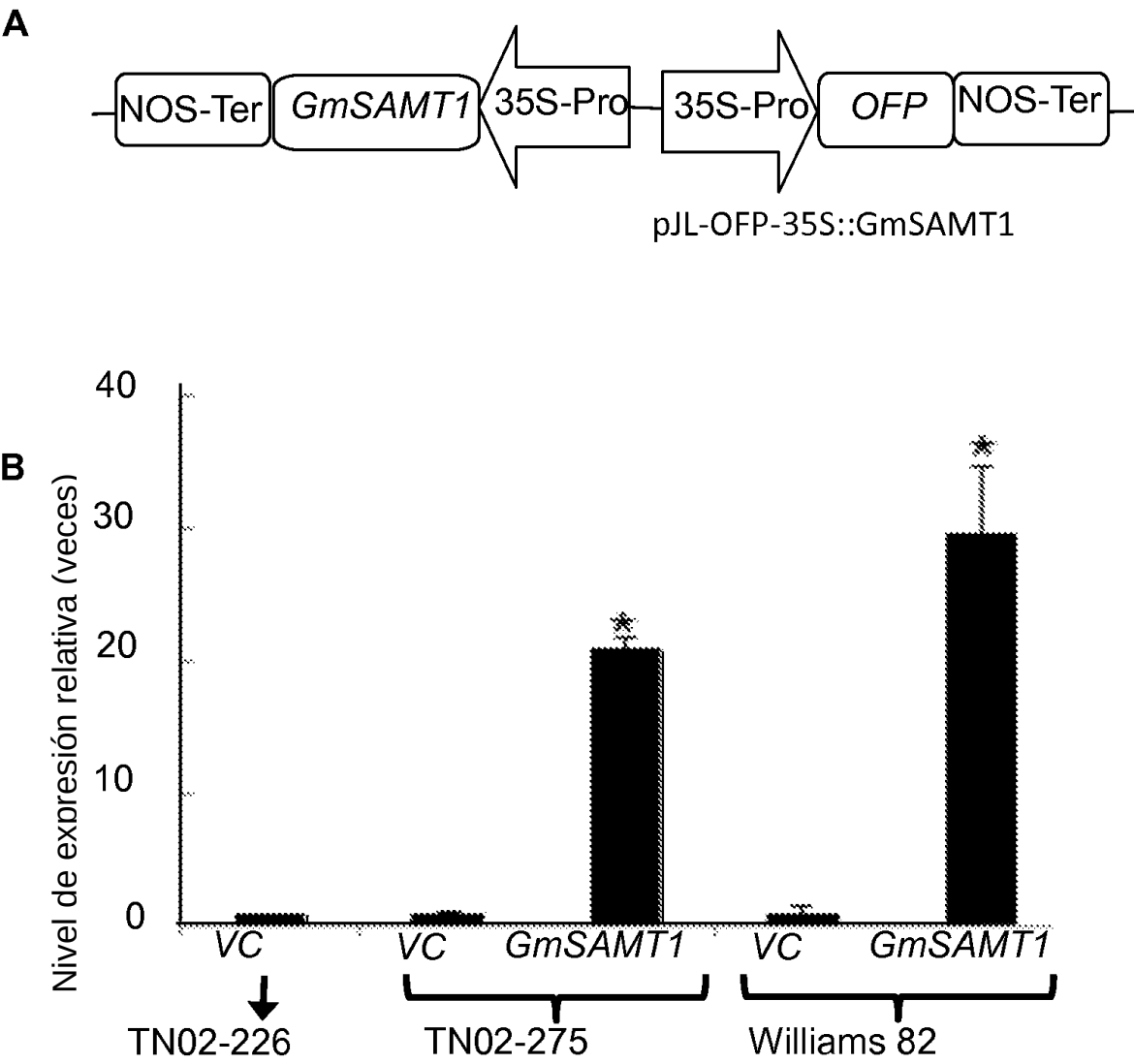


FIG. 9

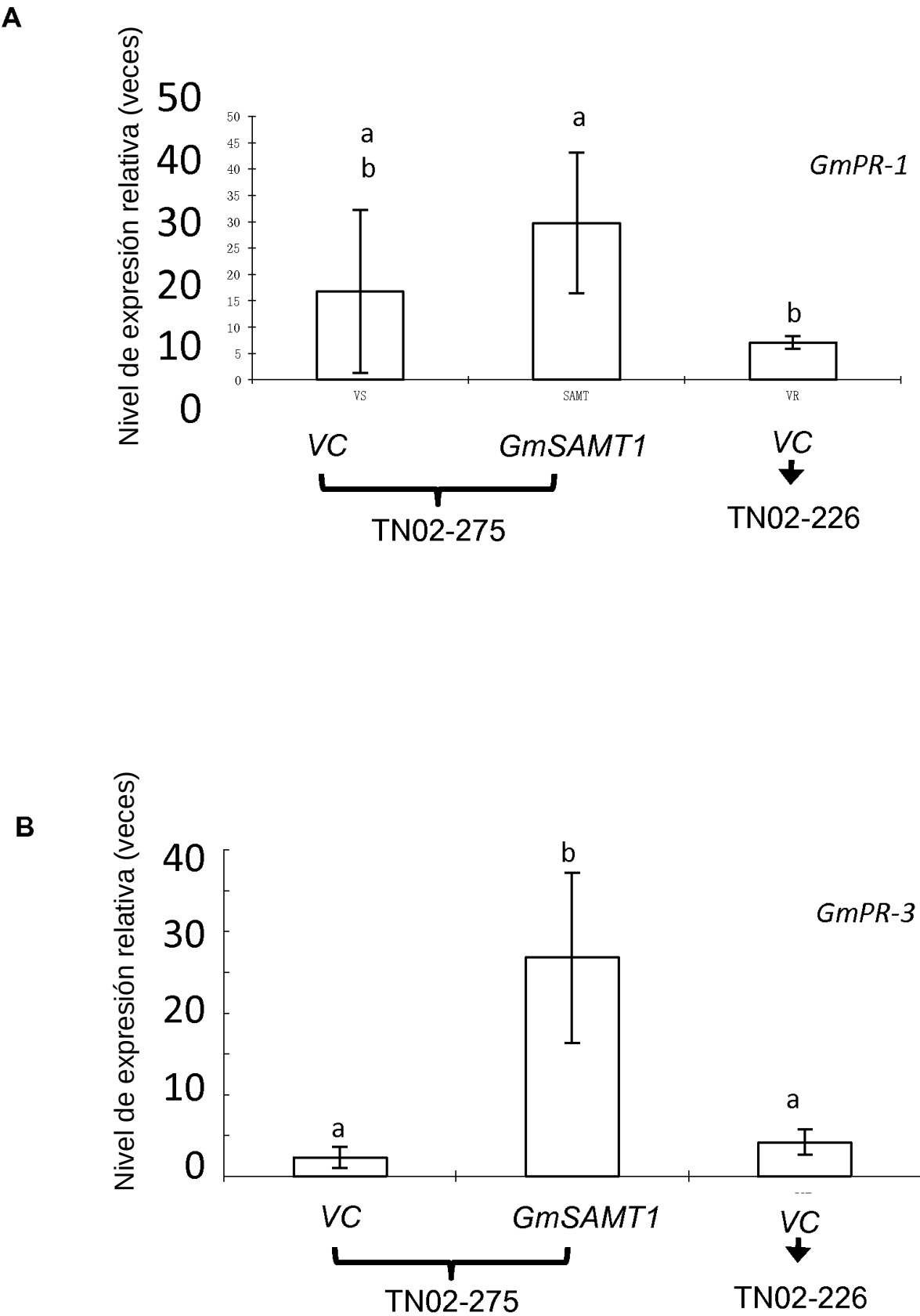


FIG. 10

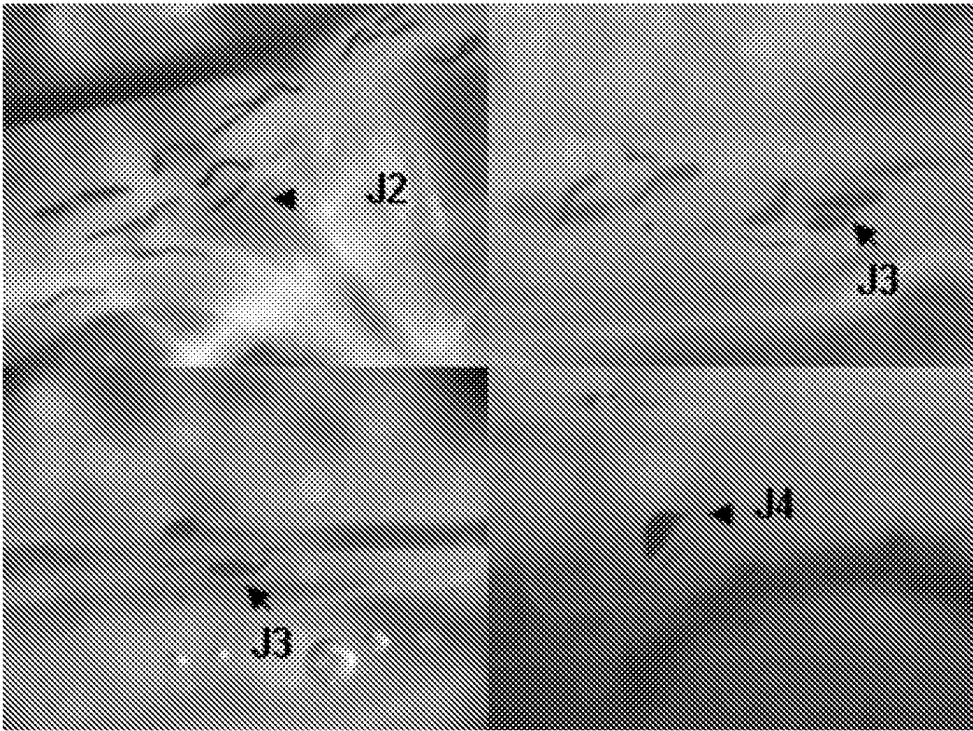


FIG. 11

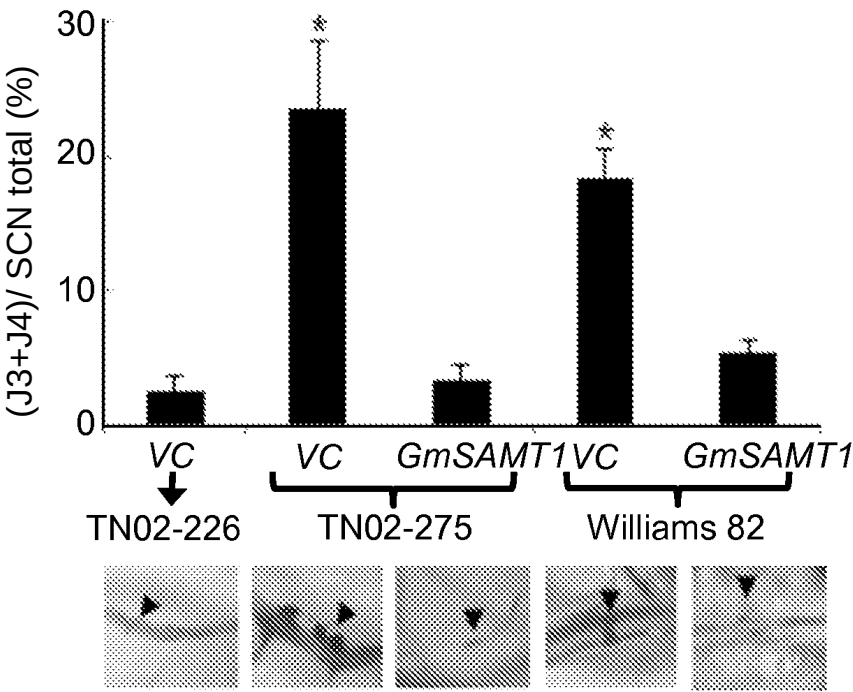
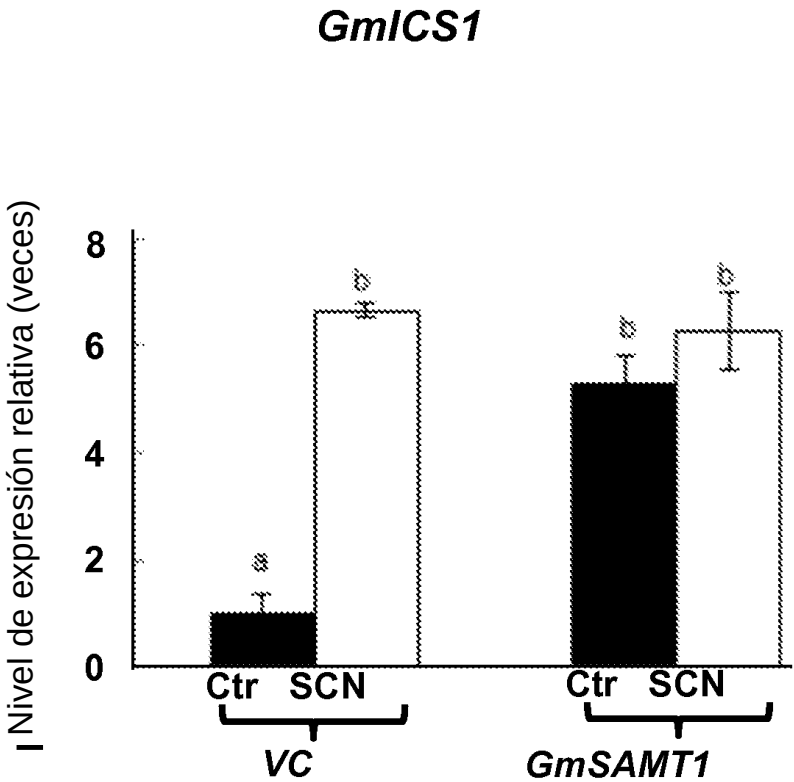


FIG. 12

A



B

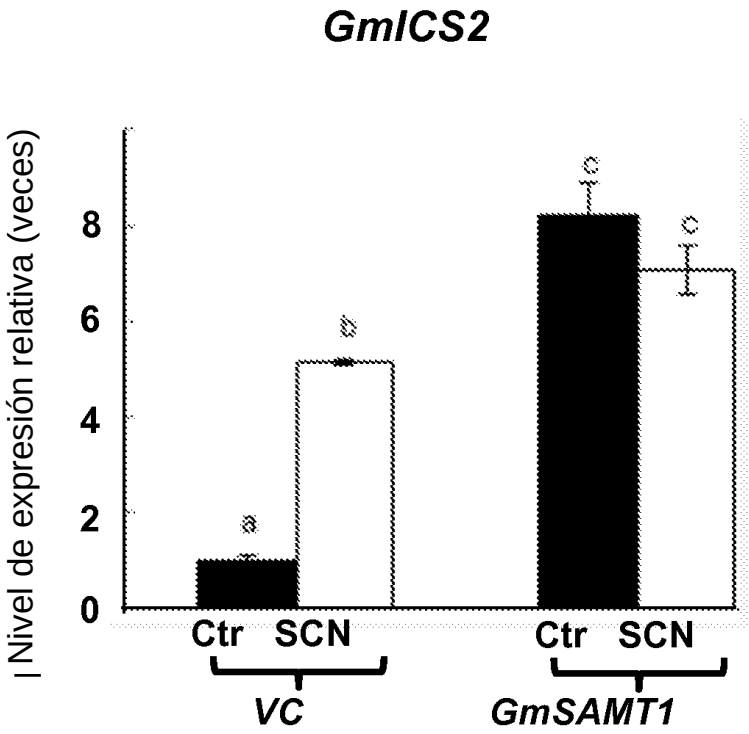


FIG. 13

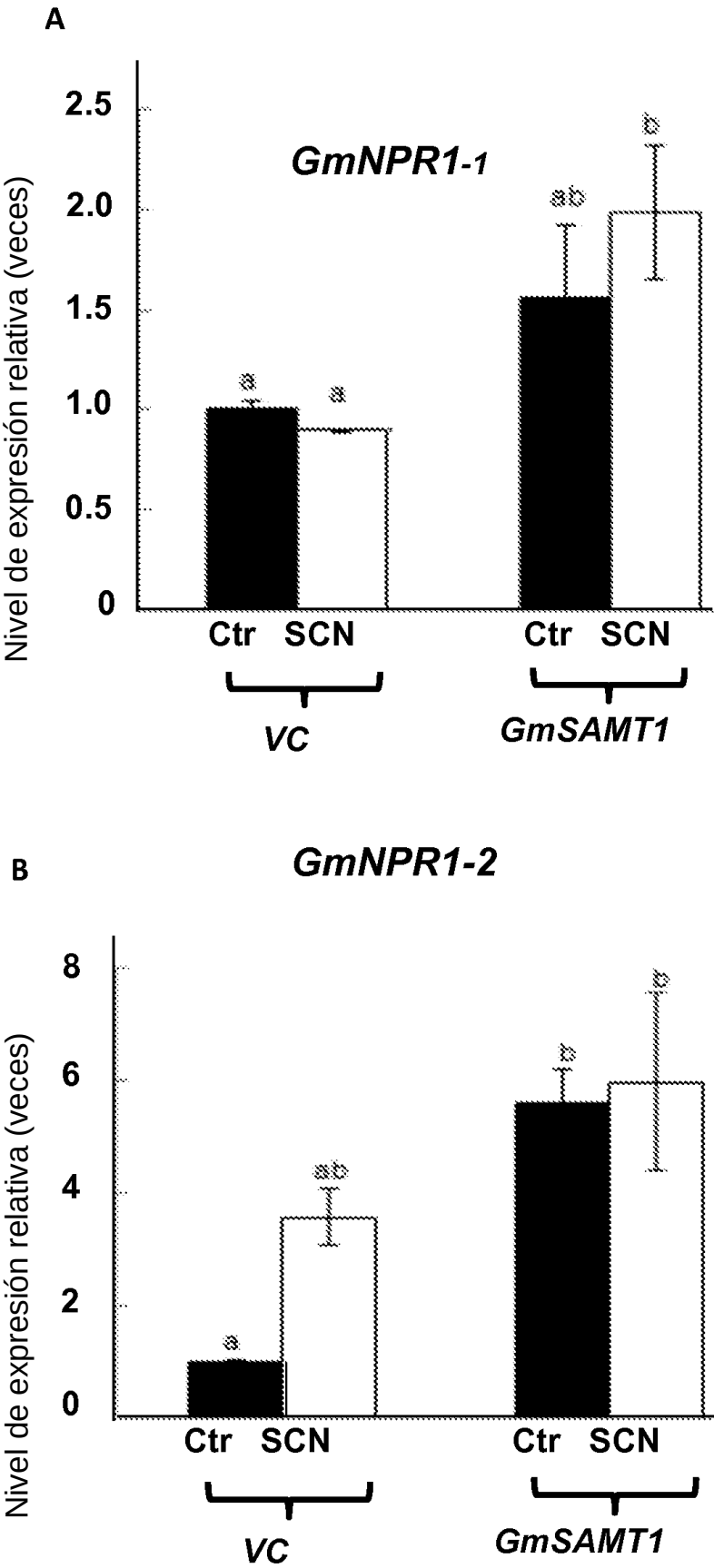


FIG. 14

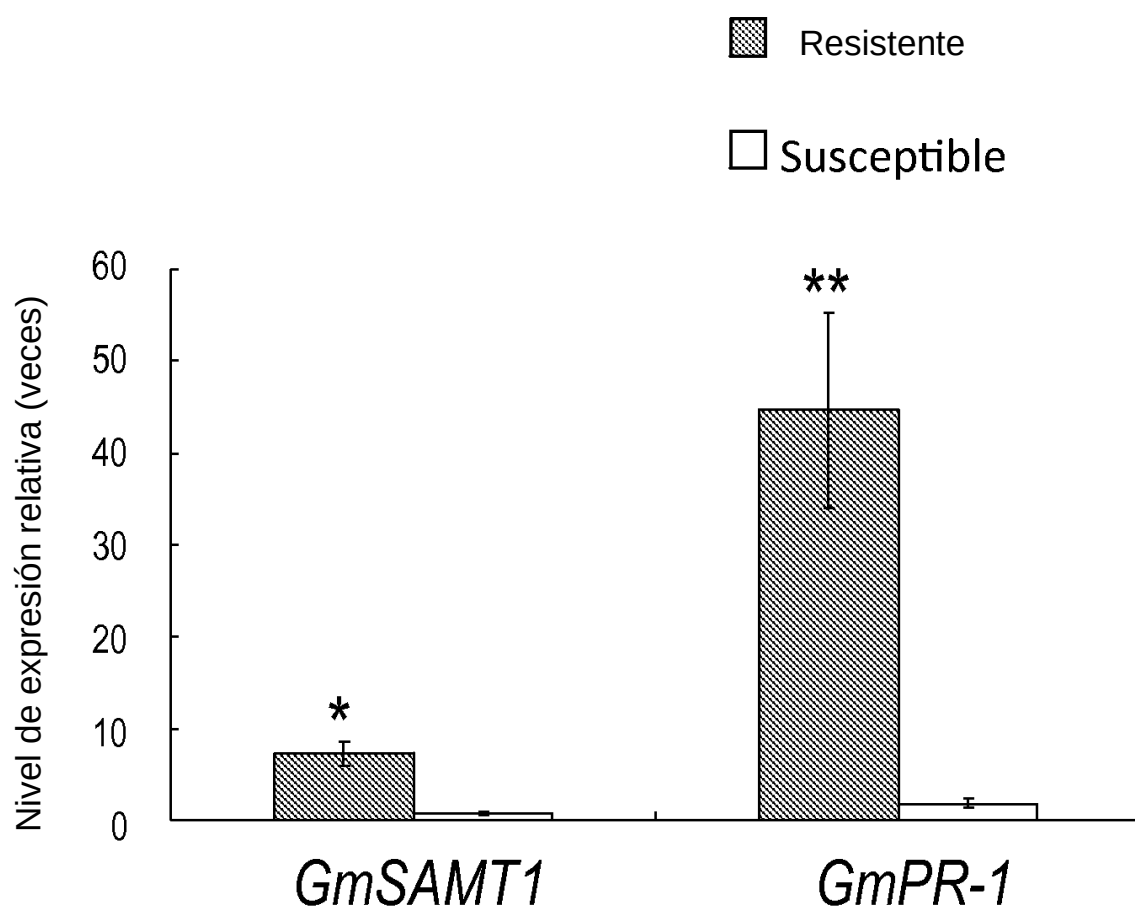


FIG. 15

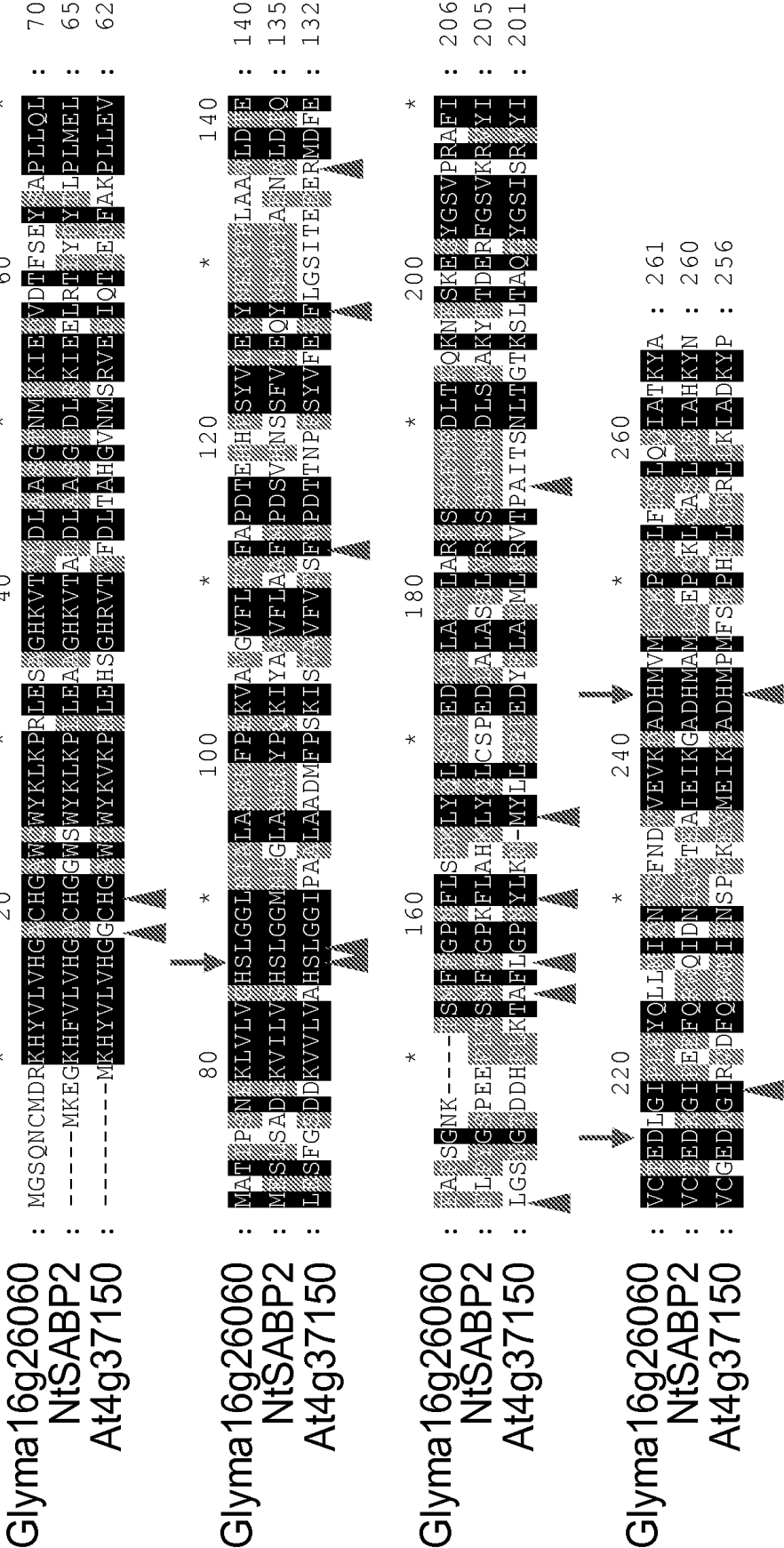


FIG. 16

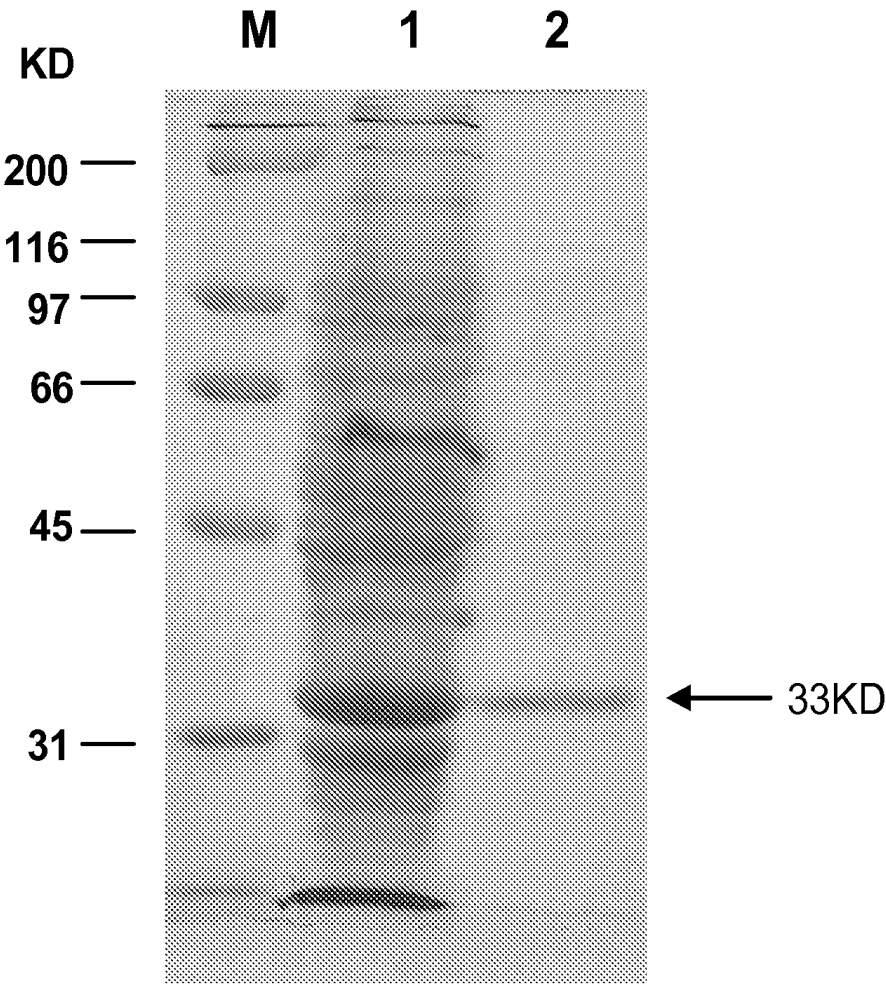


FIG. 17

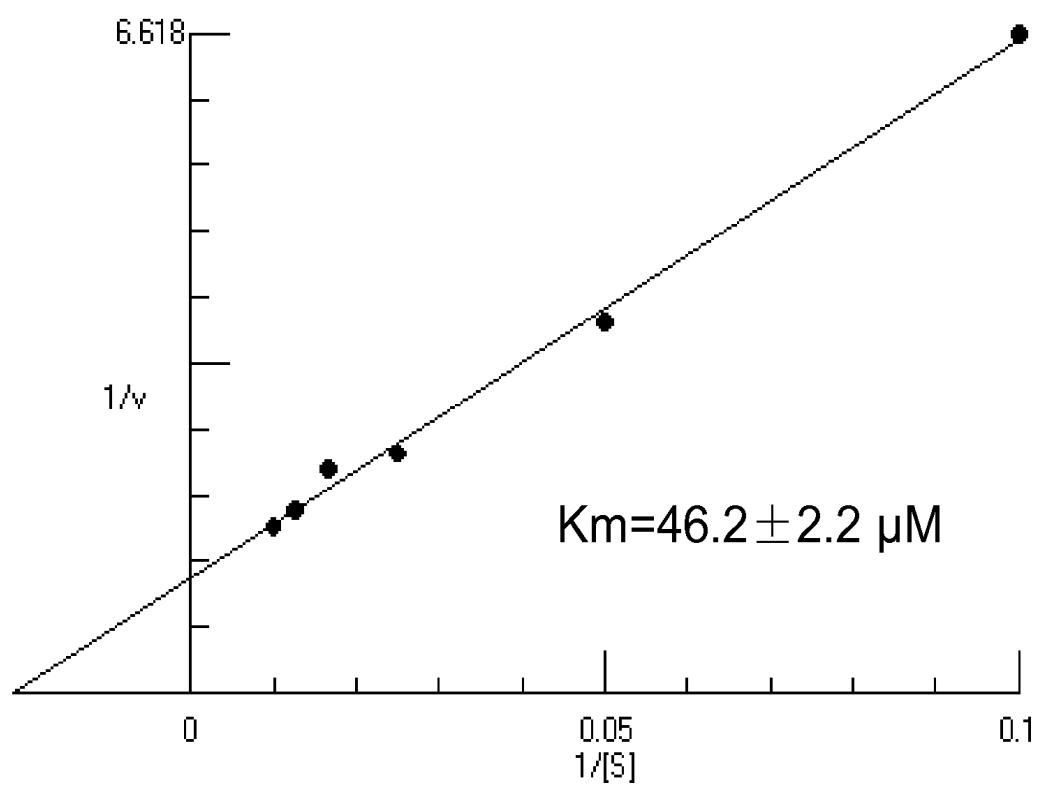


FIG. 18

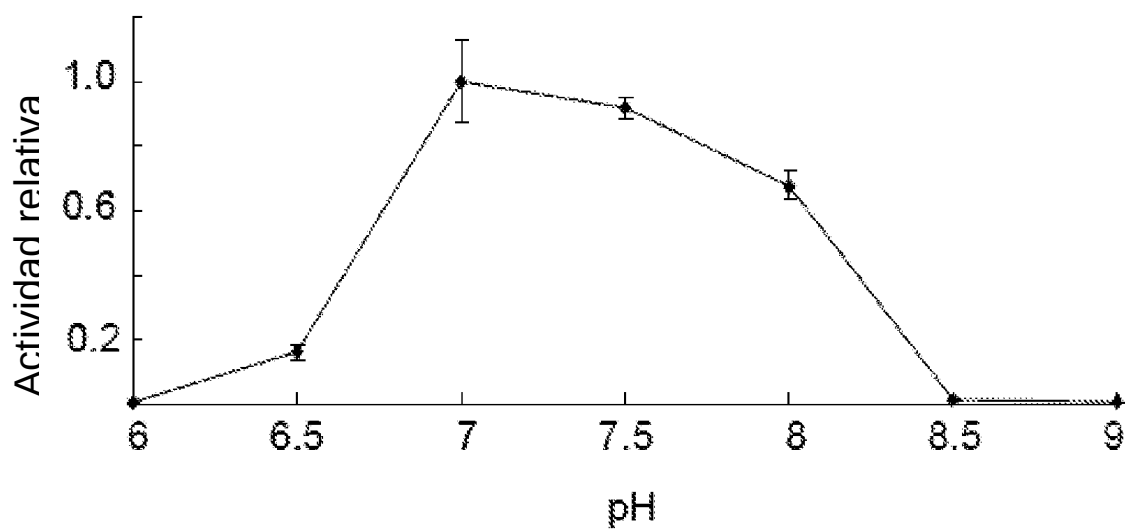


FIG. 19

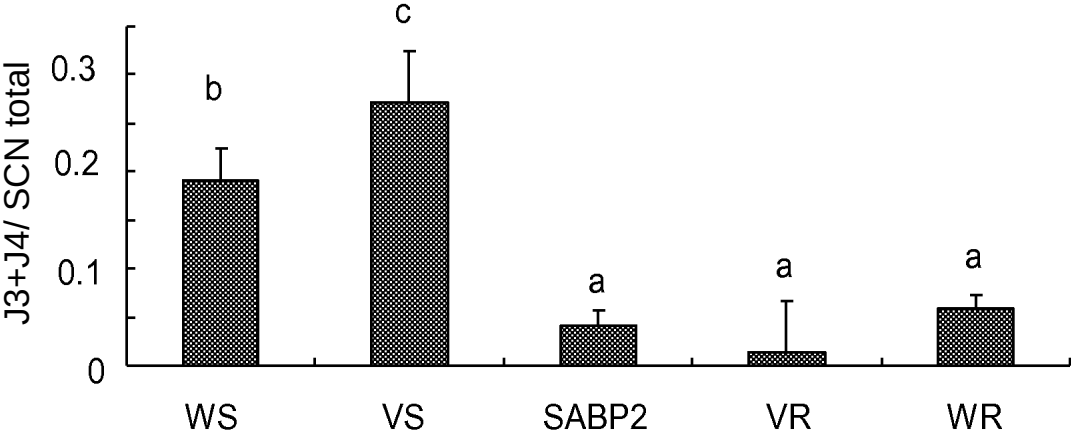


FIG. 20

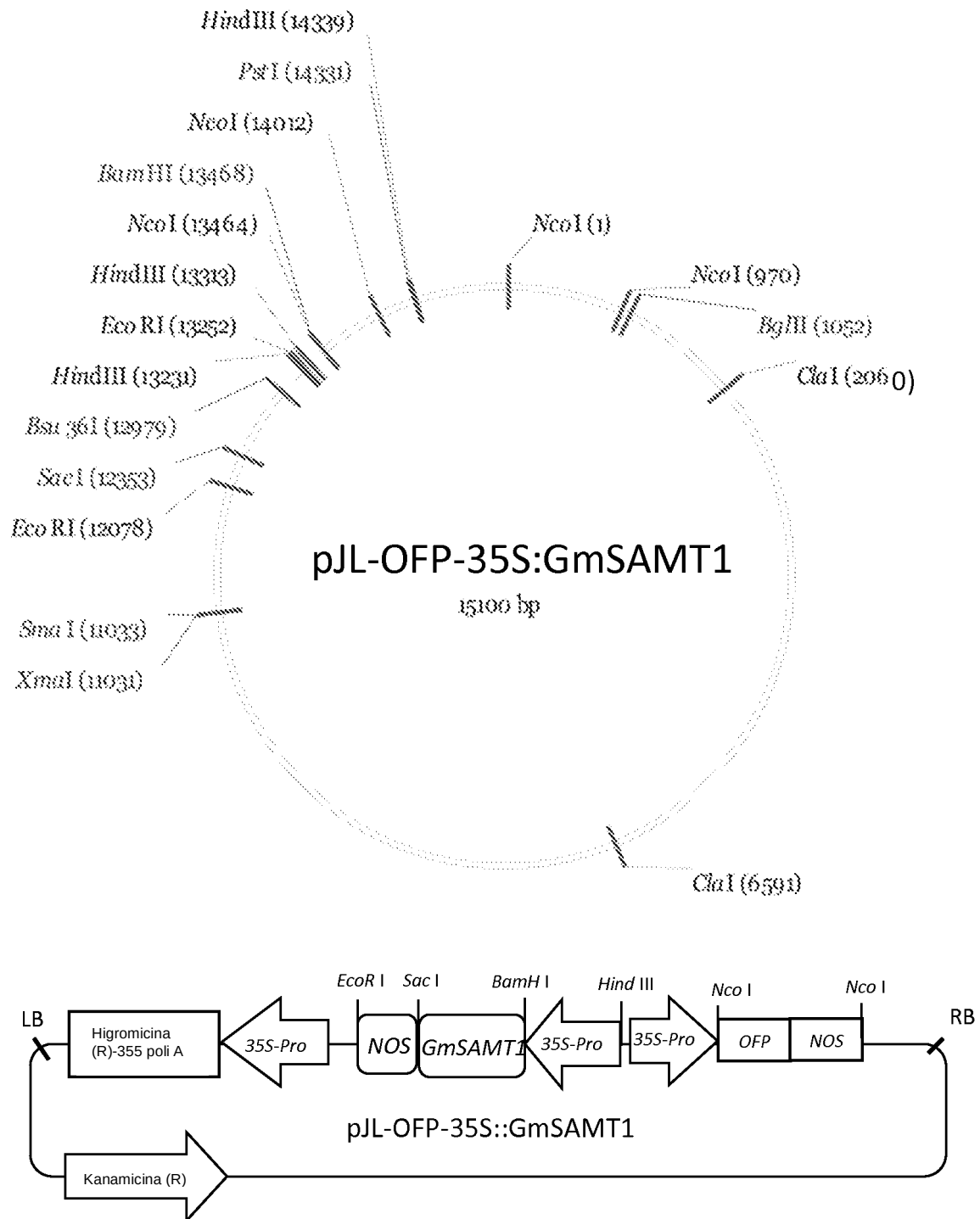
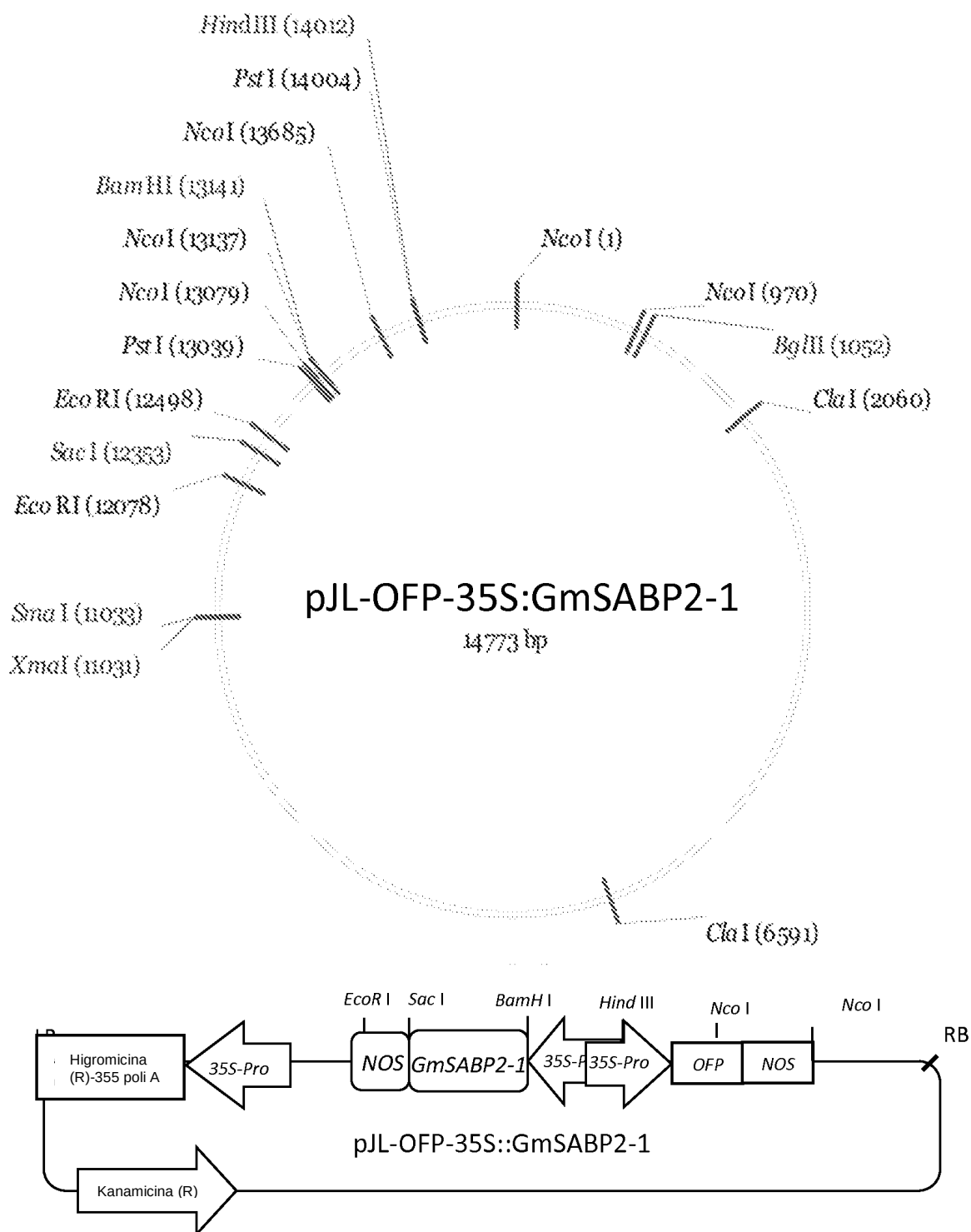


FIG. 21





UNITED STATES PATENT AND TRADEMARK OFFICE

UNDER SECRETARY OF COMMERCE FOR INTELLECTUAL PROPERTY AND
DIRECTOR OF THE UNITED STATES PATENT AND TRADEMARK OFFICE

NOVEMBER 29, 2012

PTAS

IAN J. GRISWOLD
317A E. LIBERTY ST.
JOHNSON, MARCOU & ISAACS, LLC
SAVANNAH, GA 31401

502142049

UNITED STATES PATENT AND TRADEMARK OFFICE NOTICE OF RECORDATION OF ASSIGNMENT DOCUMENT

THE ENCLOSED DOCUMENT HAS BEEN RECORDED BY THE ASSIGNMENT RECORDATION BRANCH OF THE U.S. PATENT AND TRADEMARK OFFICE. A COMPLETE COPY IS AVAILABLE AT THE ASSIGNMENT SEARCH ROOM ON THE REEL AND FRAME NUMBER REFERENCED BELOW.

PLEASE REVIEW ALL INFORMATION CONTAINED ON THIS NOTICE. THE INFORMATION CONTAINED ON THIS RECORDATION NOTICE REFLECTS THE DATA PRESENT IN THE PATENT AND TRADEMARK ASSIGNMENT SYSTEM. IF YOU SHOULD FIND ANY ERRORS OR HAVE QUESTIONS CONCERNING THIS NOTICE, YOU MAY CONTACT THE ASSIGNMENT RECORDATION BRANCH AT 571-272-3350. PLEASE SEND REQUEST FOR CORRECTION TO: U.S. PATENT AND TRADEMARK OFFICE, MAIL STOP: ASSIGNMENT RECORDATION BRANCH, P.O. BOX 1450, ALEXANDRIA, VA 22313.

RECORDATION DATE: 11/28/2012

REEL/FRAME: 029362/0977
NUMBER OF PAGES: 4

BRIEF: NUNC PRO TUNC ASSIGNMENT (SEE DOCUMENT FOR DETAILS).

DOCKET NUMBER: UTK-0141WP

ASSIGNOR:
STEWART, C. NEAL, JR.

DOC DATE: 11/26/2012

ASSIGNOR:
LIN, JINGYU

DOC DATE: 11/26/2012

ASSIGNOR:
MAZAREI, MITRA

DOC DATE: 11/27/2012

ASSIGNOR:
CHEN, FENG

DOC DATE: 11/27/2012

ASSIGNEE:
UNIVERSITY OF TENNESSEE RESEARCH
FOUNDATION
600 HENLEY STREET
SUITE 211
KNOXVILLE, TENNESSEE 37996

APPLICATION NUMBER:
PATENT NUMBER:
PCT NUMBER: US1262759
TITLE:

FILING DATE:
ISSUE DATE:

ASSIGNMENT RECORDATION BRANCH
PUBLIC RECORDS DIVISION

PATENT ASSIGNMENT

Electronic Version v1.1

Stylesheet Version v1.1

SUBMISSION TYPE:	NEW ASSIGNMENT												
NATURE OF CONVEYANCE:	NUNC PRO TUNC ASSIGNMENT												
EFFECTIVE DATE:	10/31/2012												
CONVEYING PARTY DATA													
<table border="1"> <thead> <tr> <th>Name</th> <th>Execution Date</th> </tr> </thead> <tbody> <tr> <td>C. Neal Stewart Jr.</td> <td>11/26/2012</td> </tr> <tr> <td>Jingyu Lin</td> <td>11/26/2012</td> </tr> <tr> <td>Mitra Mazarei</td> <td>11/27/2012</td> </tr> <tr> <td>Feng Chen</td> <td>11/27/2012</td> </tr> </tbody> </table>		Name	Execution Date	C. Neal Stewart Jr.	11/26/2012	Jingyu Lin	11/26/2012	Mitra Mazarei	11/27/2012	Feng Chen	11/27/2012		
Name	Execution Date												
C. Neal Stewart Jr.	11/26/2012												
Jingyu Lin	11/26/2012												
Mitra Mazarei	11/27/2012												
Feng Chen	11/27/2012												
RECEIVING PARTY DATA													
<table border="1"> <tr> <td>Name:</td> <td>University of Tennessee Research Foundation</td> </tr> <tr> <td>Street Address:</td> <td>600 Henley Street</td> </tr> <tr> <td>Internal Address:</td> <td>Suite 211</td> </tr> <tr> <td>City:</td> <td>Knoxville</td> </tr> <tr> <td>State/Country:</td> <td>TENNESSEE</td> </tr> <tr> <td>Postal Code:</td> <td>37996</td> </tr> </table>		Name:	University of Tennessee Research Foundation	Street Address:	600 Henley Street	Internal Address:	Suite 211	City:	Knoxville	State/Country:	TENNESSEE	Postal Code:	37996
Name:	University of Tennessee Research Foundation												
Street Address:	600 Henley Street												
Internal Address:	Suite 211												
City:	Knoxville												
State/Country:	TENNESSEE												
Postal Code:	37996												
PROPERTY NUMBERS Total: 1													
<table border="1"> <thead> <tr> <th>Property Type</th> <th>Number</th> </tr> </thead> <tbody> <tr> <td>PCT Number:</td> <td>US1262759</td> </tr> </tbody> </table>		Property Type	Number	PCT Number:	US1262759								
Property Type	Number												
PCT Number:	US1262759												
CORRESPONDENCE DATA													
Fax Number:	2403594879												
<i>Correspondence will be sent via US Mail when the fax attempt is unsuccessful.</i>													
Phone:	912.257.4864												
Email:	csullivan@jmi-iplaw.com												
Correspondent Name:	Ian J. Griswold												
Address Line 1:	317A E. Liberty St.												
Address Line 2:	Johnson, Marcou & Isaacs, LLC												
Address Line 4:	Savannah, GEORGIA 31401												
ATTORNEY DOCKET NUMBER:	UTK-0141WP												

ASSIGNMENT NUNC PRO TUNC

WHEREAS, we, **C. NEAL STEWART, JR.**, a citizen of the United States of America, residing at 1600 Wheaton Place, Knoxville, Tennessee 37919; **JINGYU LIN**, a citizen of the People's Republic of China residing at 1504 Sutters Mill Lane, Knoxville, Tennessee, 37909; **MITRA MAZAREI**, a citizen of the United States residing at 421 Anteelah Trail, Knoxville, Tennessee 37919; and, **FENG CHEN**, a citizen of the People's Republic of China residing at 1162 Keowee Avenue, Knoxville, Tennessee 37919, have invented certain inventions and improvements disclosed and claimed in International Patent Application No. PCT/US12/62759, filed October 31, 2012, entitled "**INCREASING SOYBEAN DEFENSE AGAINST PESTS**"; and

WHEREAS, the **UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION ("UTRF")** a corporation organized under the laws of the state of Tennessee and having an address at 600 Henley Street, Suite 211, Knoxville, Tennessee 37996, is desirous of acquiring an interest in the same.

NOW, THEREFORE, for good and valuable consideration, the receipt and sufficiency of which we hereby acknowledge, we, **C. NEAL STEWART, JR.**, **JINGYU LIN**, **MITRA MAZAREI** and **FENG CHEN**, by these presents, do sell, assign, and transfer to the **UTRF** our full, exclusive, and entire right, title, and interest in and to said patent application, in and to any non-provisionals, divisions, continuations, reexaminations, and reissues thereof, and in and to all inventions and improvements disclosed and described in said patent application;

For the same consideration, we, by these presents, do sell, assign, and transfer to the **UTRF** our full, exclusive, and entire right, title, and interest in and to any foreign application or applications corresponding to said application, in whole or in part, in countries other than the United States, in and to any Letters Patent and similar protective rights granted on said foreign applications, and in and to the right to claim any applicable priority rights arising from or required for said foreign applications under the terms of any applicable conventions, treaties,

statutes, or regulations; said foreign applications to be filed and issued in the name of the **UTRF** or its designee insofar as permitted by applicable law;

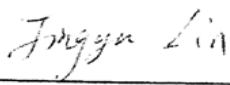
AND, for the same consideration, we agree to sign all lawful papers, execute all division, continuing, reexamination, reissue, and other applications, make all assignments and rightful oaths, and generally do everything possible to aid the **UTRF**, its successors, assigns, and nominees to obtain and enforce proper protection for all said inventions and improvements in all countries throughout the world.

This Assignment is expressly made NUNC PRO TUNC to have the same legal force and effect as if executed on October 31, 2012 (the filing date of the above-referenced international application)

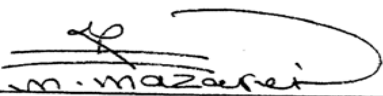
Date: Nov 26 2012


C. Neal Stewart, Jr.

Date: 11/26/2012


Jingyu Lin

Date: 11/27/12


Mitra Mazarei

Date: 11/27/2012


Feng Chen

Docket No. UTK-0141WP

TRADUCCIÓN.

[Aparece estampado un sello que reza]: Oficina de Patentes y Marcas de los Estados Unidos de América. Departamento de Comercio.

Oficina de Patentes y Marcas de los EE.UU. Subsecretario de Comercio de
5 Propiedad Intelectual y Director de la Oficina de Patentes y Marcas de los Estados Unidos de América.

29 de noviembre de 2012. Sistema de Cesión de Patentes y Marcas *[PTAS]*.

IAN J. GRISWOLD, 317A E. LIBERTY ST. JOHNSON, MARCOU & ISAACS, LLC.
SAVANNAH, GA 31401.

10 **502142049.**

Oficina de Patentes y Marcas de los Estados Unidos de América. Aviso de registro de documento de cesión.

La división de Registro de Cesiones de la Oficina de Patentes y Marcas de los EE.UU. ha registrado el documento adjunto. En la oficina de búsqueda de
15 cesiones se encuentra disponible una copia completa en el número de rollo y fotograma que se mencionan a continuación.

Sírvase leer la información completa contenida en este aviso. La información contenida en este aviso de registro refleja los datos que obran actualmente en el sistema de cesiones de la Oficina de Patentes y Marcas. Si detectara algún error
20 o tuviera alguna consulta en relación con este aviso, puede contactarse con la división de Registro de Cesiones al 571-272-3350. Sírvase enviar la solicitud de corrección a: Oficina de Patentes y Marcas de los EE.UU. Dirección interna: División de Registro de Cesiones, Casilla de Correo 1450. Alexandria, VA 22313.
Fecha de registro: 28 de noviembre de 2012. Rollo/Fotograma: 029362/0977.
25 Cantidad de páginas: cuatro.

Resumen: Cesión con efecto retroactivo (*nunc pro tunc*) (véase el documento para mayores detalles). Número de expediente: UTK-0141WP.

Cedente: STEWART, C. NEAL, JR. Fecha del documento: 26/11/2012.

Cedente: LIN, JINGYU. Fecha del documento: 26/11/2012.

5 Cedente: MAZAREI, MITRA. Fecha del documento: 27/11/2012.

Cedente: CHEN, FENG. Fecha del documento: 27/11/2012.

Cesionaria: UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION. 600 HENLEY STREET. SUITE 211. KNOXVILLE, TENNESSEE 37996.

Casilla de Correo 1450, Alexandria, Virginia 22313-1450 – WWW.USPTO.GOV.

10 [En la página siguiente, sobre el margen izquierdo]. 029362/0977 PÁGINA 2.

NÚMERO DE SOLICITUD: [EN BLANCO]. FECHA DE PRESENTACIÓN: [EN BLANCO]. NÚMERO DE PATENTE: [EN BLANCO]. FECHA DE EMISIÓN: [EN BLANCO]. NÚMERO PCT: US1262759. TÍTULO: [EN BLANCO].

SECCIÓN REGISTRO DE CESIONES. DIVISIÓN DE REGISTROS PÚBLICOS.

15 [En la página siguiente].

502142049. 28/11/2012.

CESIÓN DE PATENTE DE INVENCION.

Versión electrónica v1.1.

Versión hoja de estilo v1.1.

20 Tipo de presentación: Cesión nueva.

Naturaleza de la transmisión: Cesión con efecto retroactivo (*nunc pro tunc*).

Fecha de entrada en vigor: 31 de octubre de 2012.

Datos de la parte transmisora.

Nombre; Fecha de firma.

25 C. Neal Stewart Jr.; 26 de noviembre de 2012.

Jingyu Lin; 26 de noviembre de 2012.

Mitra Mazarej; 27 de noviembre de 2012.

Feng Chen; 27 de noviembre de 2012.

Datos de la parte receptora.

5 Nombre: University of Tennessee Research Foundation.

Domicilio: 600 Henley Street.

Dirección interna: Suite 211.

Ciudad: Knoxville.

Estado/País: TENNESSEE.

10 Código postal: 37996.

Cantidad total de bienes: 1.

Tipo de bien; Número.

Número PCT (*Acuerdo de Cooperación en Materia de Patentes*): US1262759.

Datos de correspondencia.

15 Número de fax: 2403594879.

La correspondencia se enviará mediante el correo postal de los EE.UU. cuando el intento por fax resultare infructuoso.

Teléfono: 912.257.4864.

Correo electrónico: csullivan@jmi-iplaw.com.

20 Nombre del correspondiente: Ian J. Griswold.

Línea 1 del domicilio: 317A E. Liberty St.

Línea 2 del domicilio: Johnson, Marcou & Isaacs, LLC.

Línea 4 del domicilio: Savannah, GEORGIA 31401.

Nro. de expediente del agente: UTK-0141WP.

25 *[En el margen lateral derecho se lee]:* OP US\$ 40,00 US1262759.

[En la página siguiente].

CESIÓN CON EFECTO RETROACTIVO (NUNC PRO TUNC)

POR CUANTO, nosotros, C. NEAL STEWART JR., con ciudadanía de los EE.UU., con domicilio en 1600 Wheaton Place, Knoxville, Tennessee 37919;

5 JINGYU LIN, con ciudadanía de la República Popular China, con domicilio en 1504 Sutters Mill Lane, Knoxville, Tennessee 37909; MITRA MAZAREI, con ciudadanía de los EE.UU., con domicilio en 421 Anteelah Trail, Knoxville, Tennessee 37919; y FENG CHEN, con ciudadanía de la República Popular China, con domicilio en 1162 Keowee Avenue, Tennessee 37919, somos los

10 autores de ciertas invenciones y mejoras divulgadas y reivindicadas en la Solicitud de Patente Internacional Nro. PCT/US12/62759, presentada el 31 de octubre de 2012, titulada “DEFENSA AUMENTADA DEL POROTO DE SOJA CONTRA LAS PLAGAS”; y

POR CUANTO, la UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION

15 (“UTRF”), una sociedad constituida de conformidad con las leyes del estado de Tennessee y con domicilio en 600 Henley Street, Suite 211, Knoxville, Tennessee 37996 tiene deseos de adquirir el derecho sobre dichas invenciones y mejoras.

POR LO TANTO, por contraprestación a título oneroso, cuyo recibo se acredita por la presente, nosotros, C. NEAL STEWART JR., JINGYU LIN, MITRA

20 MAZAREI y FENG CHEN, vendemos, cedemos y transferimos a UTRF nuestro pleno y exclusivo derecho, título y participación en relación con dicha solicitud de patente, con todas sus solicitudes no provisionales, divisiones, continuaciones, reexaminaciones y reemisiones, y en relación con todas las invenciones y mejoras divulgadas y descriptas en dicha solicitud de patente.

25 En virtud de la misma contraprestación, vendemos, cedemos y transferimos a

UTRF nuestro pleno y exclusivo derecho, título y participación en relación con las solicitudes extranjeras relacionadas con dicha solicitud, en forma parcial o total, en países fuera de los EE.UU., y en relación con las patentes y derechos de protección similares que se concedan mediante dichas solicitudes extranjeras, y
5 en relación con los derechos a reivindicar derechos de prioridad que surjan o se requieran por dichas solicitudes extranjeras en virtud de las estipulaciones de los convenios, tratados,

Página 1 de 2.

Solicitud de Patente Internacional PCT Nro. PCT/US12/62759.

10 *[En la página siguiente].*

leyes o normas de aplicación; dichas solicitudes extranjeras se presentarán y emitirán a nombre de UTRF o de quien esta designe, dentro del alcance permitido por la ley de aplicación.

Y, por la misma contraprestación, convenimos en firmar todos los escritos
15 legítimos, en firmar las solicitudes de división, continuación, reexaminación, reemisión y otras solicitudes, realizar cesiones, prestar legítimo juramento, y en general, realizar todo lo que esté a nuestro alcance para colaborar con UTRF, sus sucesores a título universal, a título particular y sus representantes a fin de obtener y poner en vigor la protección adecuada para dichas invenciones y
20 mejoras en todos los países del mundo.

La presente Cesión se realiza expresamente con efecto retroactivo (*nunc pro tunc*) y tendrá la misma vigencia y validez desde el punto de vista legal tal como si se hubiera firmado el 31 de octubre de 2012 (la fecha de presentación de la solicitud de patente internacional citada anteriormente).

25 Fecha: 26 de noviembre de 2012. *[Aparece una firma]*. C. Neal Stewart, Jr.

Fecha: 26 de noviembre de 2012. *[Aparece una firma]*. Jingyu Lin.

Fecha: 27 de noviembre de 2012. *[Aparece una firma]*. Mitra Mazarei.

Fecha: 27 de noviembre de 2012. *[Aparece una firma]*. Feng Chen.

Expediente Nro. UTK-0141WP.

COLOMBIA

PODER

El (los) **UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION**

domiciliados en **600 Henley Street, Suite 211 Knoxville (37996-4122) (US)**

nombran por el presente a **XIMENA CASTELLANOS A. y/o MARGARITA CASTELLANOS A.** abogadas con oficinas en la Calle 94A No. 7A-09, Bogotá-Colombia para

y para obtener la protección de mi (nuestra) propiedad Industrial e intelectual y con este fin autorizamos a dichos apoderados para que nos representen ante cualesquiera autoridades o personas de los órdenes administrativo, contencioso – administrativo y judicial en todo lo relacionado con los siguientes actos y gestiones:

- a) Obtención de patentes de invención, modelos y dibujos industriales; registro de marcas de fabrica, de servicio, colectivas o defensivas y de denominaciones de origen, depósito de nombres y enseñas comerciales; su renovación, prórroga, inscripción de traspaso y cambio de nombre o domicilio, cancelación voluntaria y el registro de licencias de uso de esos derechos;
- b) Para la obtención de registros sanitarios;
- c) Para la obtención de registros de derechos de autor.
- c) Acciones de oposición, observación, cancelación, nulidad, medidas cautelares y en general los juicios civiles, penales, administrativos o contencioso - administrativos que promovamos o se nos promuevan.

Pudiendo en todos los actos y gestiones arriba mencionados sustituir éste mandato entre los apoderados, transigir, desistir, recibir, cancelar, renunciar, aceptar notificaciones y nombrar apoderados judiciales y extra-judiciales. .

Finalmente, este poder ratifica la validez de todos los actos y solicitudes que, a nombre de la poderdante hubieren realizado los apoderados en relación con las facultades anteriormente enumeradas, antes de la suscripción del presente poder

Dado y firmado hoy

en

por

POWER OF ATTORNEY

I (we) the undersigned **UNIVERSITY OF TENNESSEE RESEARCH FOUNDATION**

of **600 Henley Street, Suite 211 Knoxville (37996-4122) (US)**

hereby grant(s) Power of Attorney to **XIMENA CASTELLANOS A. and /or MARGARITA CASTELLANOS** a. attorneys residing at Calle 94A No. 7A-09, Bogotá- Colombia, to

and to obtain the protection of my (our) Intellectual Property, for which purpose we authorize the said attorneys to represent us before any authorities, or persons, specially those, administrative, contentious and judicial, in all matters concerning the following:

- a) Obtainment of patents, models and designs, registration of trademarks, service, collective and defensive marks, and of designations of origin; the deposit of trade names and ensgins; its renewal, extension, recordal of assignments, change of name or domicile, voluntary cancellation, and the registration of licenses of use of the above rights;
- b) For the obtainment of Sanitary registrations;
- c) For the obtainment of copyright registrations

Actions of opposition, cancellation, nullity, infringement, granting of licenses; and in general the civil, criminal and administrative suits related to those rights; actions and suits instituted by us or against us;

In all the above matters they may substitute this power of attorney within the appointed attorneys, compromise, cancel, receive, withdraw, renounce, accept notifications and appoint attorneys in fact in or out of court.

Finally, this power of attorney ratifies the validity of all acts and applications made by the representatives in the name of the signatory of this power of attorney, with regard to the faculties above given, before the execution of the present power of attorney.

Given and signed this April 24, 2014

in Tennessee, U.S.A.

by



Dr. Stacey Patterson, Vice President of the University of Tennessee Research Foundation

CASTELLANOS & Co.
INTELLECTUAL PROPERTY LAW FIRM

CALLE 94 A No. 7 A – 09 / P.O. BOX 6349
PHONE (57 1) 756 07 70 / FAX (57 1) 755 01 58
BOGOTÁ D.C. COLOMBIA
info@castellanosyco.co

LISTADO DE SECUENCIAS

<110> University of Tennessee Research Foundation

<120> DEFENSA SUPERIOR DE SOJA CONTRA PLAGAS

<130> UTK-0141WP

<150> US 61/581,039

<151> 2011-12-28

<160> 33

<170> PatentIn versión 3.5

<210> 1

<211> 1113

<212> ADN

<213> Glycine Max

<400> 1

atggaagtag cacaggtact ccacatgaac ggtggcgttg gagacgcaag ctatgcaaac 60

aactcccttg ttcagcaaaa ggtgatttgt ttgacaaagc ccataagaga ggaagccata 120

agaagcctct attgcagcac acaccccaga agcttggaag ttgcagattt gggttgctct 180

tctggaccaa acactttatt tgttgtgtct gaattcataa aaattgtgga gaagctttgc 240

cgagagctga accacaaatc tccagaatac aaagtcttc tgaatgatct ccctgggaat 300

gacttcaaca acatcttcaa gtctcttgac agcgtcaaag agaaattgtg tgatgaaatg 360

gaaagtggga tcggtccatg ctacttctcg ggtgttcccg gttctttcta tggaaggggt 420

ttcccatatc aaagtcttca tttgtccat tcctcataca gccttcaatg gctatctaag 480

gttcctgagg gtgtagacaa caacaagggc aatgtttaca taggcagtac gagccccaaa 540

aatgttgtga gagcttacta tgagcaattt cagagagatt tctctcttt tctcaagtgt 600

cgtgcagagg aattggttga aggaggacgc atggttctca cattttggg aagaagaagc 660

gatgatccat ctagcaagga tggttgctac attggggagc tttggctac tgctcttagt 720

gatatggtct tgcaggggaat cataagagaa gagcaattag atacttttaa catccctcaa 780

tacactccat ccccatctga agtgaaattg gaagtctta aagaaggatc attcgccatc 840

aatcgtctag aggtttctga ggtgaattgg aatgctctcg atgagtggaa tgctctagac 900

ttgaatctg aaaggtctga atcacttagt gatggtgaat acaatgtggc acagtgcattg 960

agggctgtgg cagaacctat gctgattagc cactttggtg aagctatcat tgaagaggtt 1020

tttgccgct accagcaaat ctgggctgaa cgtatgtcca aggagaaaac caagttcatc 1080

aatgttacca tattattgac tagaaaagca taa 1113

<210> 2

<211> 786

<212> ADN

<213> Glycine Max

<400> 2

atgggttcac aaaattgtat ggataggaag cactatgttc tggatcatgg ggcattgcat 60

ggagcttggg gttggtataa gctcaagcca cgcttggaa ctgcaggcca taaggtcaca 120

gtacttgacc ttgcagcttc tggaaccaac atgaagaaaa ttgaagatgt tgatacttc 180

tcagagtatt ctgcgccttt gttgcagcta atggccacaa ttcctcaaa tgagaagtta 240

gttctagtgt gtcacagcct tggagggctg aacatagcac ttgcaatgga gaaattcca 300

gaaaaggtag cagttggtgt ttcttaaca gctttgctc cagacactga acaccacca 360

tctatgtct tggaagagta caatgagagg acccgtag ctgcatggt agacactgaa 420

ttgctcaa gcggaacaa aacatcaatg ttctcgcc ccaactctt gtccgacaag 480

ctctaccaac tgcccctat tgaggacttg gaattggcca agacttagc aaggccatca 540

tcgctctca tggaagactt gactaaacaa aagaactct ccaaagaggg atatgggtca 600

gtccacgtg ctttattgt ttgcacggag gaccttgaa ttccattgga atatcagctc 660

ttgatgatcc aaaatgttg gttcaatgac gttgtagagg tcaaagacgc agatcatatg 720

gttatgctt gcaagccaca agaactattc gattccctcc agcagatagc gactaaatat 780

gcatga 786

<210> 3

<211> 15101

<212> ADN

<213> Secuencia Artificial

<220>

<223> Constructo de plásmido sintético pJL-OFP-35S::GmSAMT1

<400> 3

tcattggctct ttcaaagcaa agtgggggtca aagatgtaat gaacaccgag cttcatatgg	60
acgggatcgt caatggacac ccctttgaga taaaaggaaa aggaaaggga aaccctgaca	120
agggtgtgca gaccatgaaa cttacagtca ttaaggggtgc gcctttgcca ttttctattg	180
acattttgct gcctcaacac atgtatggaa gcaagccatt tattaagtat cctgagagta	240
tcccagacta catcaagttg tcatttcccg aggggaatcac atgggaaagg tccatgacct	300
ttgaagatgg tgcaagtgtc actgtctcta acgactccag gctcgaaggc gactctttca	360
tctacgaagt caggtttctt ggcgatgaact ttccccgaga tggacctgtt atgcagaaga	420
agacgcgagg ctgggacctg tccacagaga gactgtatga gtgtgggtggg tggcagagag	480
gagacgtcca catggccttg aagttggaga acggtggcca ttatacgtgc gacttcaaaa	540
ctacttaca atcaaagaag ggcttgaagg tgccaccgta tcacttcgtt gaccacaaac	600
tagatttact gagccacaac accgatgggtg ctacctttga agagtttgaa caacgagaaa	660
ttgcacatgc acatctttct aacttaccgg tagcccatca ccatcaccat cactaagatc	720
gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt cttgcgatga	780
ttatcatata atttctgttg aattacgtta agcatgtaat aattaacatg taatgcatga	840
cgttatttat gagatgggtt ttatgatta gagtcccga attatacatt taatacgcga	900
tagaaaacaa aatatagcgc gcaaactagg ataaattatc gcgcgcggtg tcattctatgt	960

tactagatcc catggctact actaagcatt tggctcttgc catcctgtc ctccttagca 1020

ttggtatgac caccagtgc agaaccctcc tagatctgag ggtaaatttc tagttttct 1080

ccttcattt ctgggttagg accctttct cttttattt tttgagctt tgatcttct 1140

ttaaactgat ctattttta attgattgg tttggtgtaa atattacata gctttaactg 1200

ataatctgat tactttattt cgtgtgtcta tgatgatgat gatagttaca gaaccgacga 1260

actagtctgt acccgatcaa caccgagacc cgtggcgtct tcgacctcaa tggcgtctgg 1320

aactcaagc tggactacgg gaaaggactg gaagagaagt ggtacgaaag caagctgacc 1380

gacactatta gtatggccgt cccaagcagt tacaatgaca ttggcgtgac caaggaaatc 1440

cgcaaccata tcggatatgt ctggtacgaa cgtgagttca cggtgccggc ctatctgaag 1500

gatcagcgta tcgtgctccg cttcggctct gcaactcaca aagcaattgt ctatgtcaat 1560

ggtgagctgg tcgtggagca caagggcgga ttctgccat tcgaagcgga aatcaacaac 1620

tcgtgctgtg atggcatgaa tcgcgtcacc gtcgccgtgg acaacatcct cgacgatagc 1680

accctcccgg tggggctgta cagcgagcgc cacgaagagg gcctcgaaa agtcattcgt 1740

aacaagccga acttcgactt ctcaactat gcaggcctgc accgtccggg gaaaatctac 1800

acgaccccg ttacgtacgt cgaggacatc tcggttgta ccgactcaa tggcccaacc 1860

gggactgtga cctatacggg ggactttcaa ggcaaagccg agaccgtgaa agtgtcggtc 1920

gtggatgagg aaggcaaagt ggtcgcaagc accgagggcc tgagcggtaa cgtggagatt 1980

ccgaatgtca tcctctggga accactgaac acgtatctct accagatcaa agtggaaactg 2040

gtgaacgacg gactgaccat cgatgtctat gaagagccgt tcggcgtgcg gaccgtggaa 2100

gtcaacgacg gcaagttcct catcaacaac aaaccgttct acttcaaggg ctttggcaaa 2160

catgaggaca ctctatcaa cggccgtggc ttaacgaag cgagcaatgt gatggatttc 2220

aatatcctca aatggatcgg cgccaacagc ttccggaccg cacactatcc gtactctgaa 2280

gagttgatgc gtcttcgga tcgcgagggc ctggctgta tcgacgagac tccggcagtt 2340

ggcgtgcacc tcaactcat ggccaccacg ggactcggcg aaggcagcga gcgcgtcagt 2400

acctgggaga agattcggac gttgagcac catcaagacg ttctccgtga actggtgtct 2460

cgtgacaaga accatccaag cgtcgtgatg tggagcatcg ccaacgaggc ggcgactgag 2520

gaagagggcg cgtacgagta ctcaagccg ttggtggagc tgaccaagga actcgacca 2580

cagaagcgtc cggtcacgat cgtgctgtt gtgatggcta ccccgagac ggacaaagtc 2640

gccgaactga ttgacgtcat cgcgctcaat cgctataacg gatggtactt cgatggcggc 2700

gatctgaag cggccaaagt ccatctccgc caggaatttc acgcgtggaa caagcgttc 2760

ccaggaaagc cgatcatgat cactgagtac ggcgagaca ccgttcggg ctttcacgac 2820

attgatccag tgatgttcac cgaggaatat caagtcgagt actaccaggc gaaccacgtc 2880

gtgttcgatg agtttgagaa cttcgtgggt gagcaagcgt ggaacttcgc ggacttcgcg 2940

acctctcagg gcgtgatgcg cgtccaagga aacaagaagg gcgtgttcac tcgtgaccgc 3000

aagccgaagc tcgccgcgca cgtcttctgc gagcgctgga ccaacattcc agatttcggc 3060

tacaagaacg ctagccatca ccataccat cacgtgtgaa ttggtgacca gctcgaattt 3120

ccccgatcgt tcaaacattt ggcaataaag ttcttaaga ttgaatcctg ttgccggtct 3180

tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta 3240

atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta 3300

atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc 3360

atctatgta ctagatcggg aattaaacta tcagtgtttg acaggatata ttggcgggta 3420

aacctaagag aaaagagcgt ttattagaat aacggatatt taaaagggcg tgaaaaggtt 3480

tatccgttcg tccatttgta tgtgcatgcc aaccacaggg tcccctcgg gatcaaagta 3540

cttgatcca acccctccgc tgctatagtg cagtcggctt ctgacgttca gtgcagccgt 3600

cttctgaaaa cgacatgtcg cacaagtcct aagttacgcg acaggctgcc gccctgccct 3660

tttctggcg tttcttctg gcgtgtttta gtcgcataaa gtagaatact tgcgactaga 3720

accggagaca ttacgcatg aacaagagcg ccgccgctgg cctgctgggc tatgcccgcg 3780

tcagcaccga cgaccaggac ttgaccaacc aacgggccga actgcacgcg gccggctgca 3840

ccaagctgtt ttccgagaag atcaccggca ccaggcgcga ccgcccggag ctggccagga 3900

tgcttgacca cctacgccct ggcgacgttg tgacagtgc caggctagac cgcttgccc 3960

gcagcaccgc cgacctactg gacattgccg agcgcaccca ggaggccggc gcgggcctgc 4020

gtagcctggc agagccgtgg gccgacacca ccacgccggc cggccgcatg gtgttgaccg 4080

tgttcgccgg cattgccgag ttcgagcgtt ccctaatacat cgaccgcacc cggagcgggc 4140

gcgaggccgc caaggcccga ggcgtgaagt ttggccccg ccctaccctc accccggcac 4200

agatcgcgca cgcccgag ctgatcgacc aggaaggccg caccgtgaaa gaggcggctg 4260

cactgcttgg cgtgcatcgc tcgaccctgt accgcgcact tgagcgcagc gaggaagtga 4320

cgcccaccga ggccaggcgg cgcggtgcct tccgtgagga cgcattgacc gaggccgacg 4380

ccctggcggc cgccgagaat gaacgccaag aggaacaagc atgaaaccgc accaggacgg 4440

ccaggacgaa ccgttttca ttaccgaaga gatcgaggcg gagatgatcg cggccgggta 4500

cggttcgag ccgcccgcgc acgtctcaac cgtgcggctg catgaaatcc tggccggtt 4560

gtctgatgcc aagctggcgg cctggccggc cagcttgcc gctgaagaaa ccgagcgccg 4620

ccgtctaaaa aggtgatgtg tatttgagta aaacagcttg cgtcatcgcg tcgctgcgta 4680

tatgatcgga tgagtaaata aacaaatacg caaggggaac gcatgaaggc tatcgctgta 4740

ctaaccaga aaggcgggtc aggcaagacg accatcgcaa cccatctagc ccgcgccctg 4800

caactcgccg gggccgatgt tctgttagtc gattccgac cccagggcag tgcccgcgat 4860

tgggcggccg tgcgggaaga tcaaccgta accgtgtcg gcatcgaccg cccgacgatt 4920

gaccgcgacg tgaaggccat cggccggcgc gacttcgtag tgatcgacgg agcgccccag 4980

gcggcggact tggctgtgtc cgcgatcaag gcagccgact tcgtgctgat tccggtgcag 5040

ccaagccctt acgacatatg ggccaccgcc gacctggtgg agctgggttaa gcagcgcatt 5100

gaggtcacgg atggaaggct acaagcggcc ttgtcgtgt cgcgggcgat caaaggcacg 5160

cgcatcggcg gtgaggttgc cgaggcgctg gccgggtacg agctgccccat tcttgagtcc 5220

cgatcacgc agcgcgtgag ctaccaggc actgccgccg ccggcacaac cgttcttgaa 5280

tcagaaccgg agggcgacgc tgcccgcgag gtccaggcgc tggccgctga aattaaatca 5340

aaactcattt gagttaatga ggtaaagaga aaatgagcaa aagcacaac acgctaagt 5400

ccggccgtcc gagcgcacgc agcagcaagg ctgcaacgtt ggccagcctg gcagacacgc 5460

cagccatgaa gcgggtcaac ttccagttgc cgccggagga tcacaccaag ctgaagatgt 5520

acgcggtacg ccaaggcaag accattaccg agctgctatc tgaatacatc gcgcagctac 5580

cagagtaa at gagcaaatga ataatgagt agatgaattt tagcggctaa aggaggcggc 5640

atggaaaatc aagaacaacc aggcaccgac gccgtggaat gccccatgtg tggaggaacg 5700

ggcggttggc caggcgtaag cggtgggtt gtctgccggc cctgcaatgg cactggaacc 5760

ccaagcccg aggaatcggc gtgacggtcg caaacatcc ggcccgttac aaatcggcg 5820

ggcgctgggt gatgacctgg tggagaagtt gaaggccgcg caggccgccc agcggcaacg 5880

catcgaggca gaagcacgcc ccggtgaatc gtggcaagcg gccgctgac gaatccgcaa 5940

agaatcccg caaccgccg cagccggtgc gccgtcgatt aggaagccgc ccaagggcga 6000

cgagcaacca gatttttcg ttccgatgct ctatgacgtg ggcacccgcg atagtcgcag 6060

catcatggac gtggccgttt tccgtctgtc gaagcgtgac cgacgagctg gcgaggatgat 6120

ccgctacgag ctccagacg ggcacgtaga ggttccgca gggccggccg gcatggccag 6180

tgtgtgggat tacgacctgg tactgatggc ggttcccat ctaaccgaat ccatgaaccg 6240

ataccgggaa gggaagggag acaagcccgg ccgctgtgtc cgtccacacg ttgcggacgt 6300

actcaagttc tgccggcgag ccgatggcgg aaagcagaaa gacgacctgg tagaaacctg 6360

cattcggtta aacaccacgc acgttgccat gcagcgtacg aagaaggcca agaacggccg 6420

cctggtgacg gtatccgagg gtgaagcctt gattagccgc tacaagatcg taaagagcga 6480

aaccgggcgg ccggagtaca tcgagatcga gctagctgat tggatgtacc gcgagatcac 6540

agaaggcaag aaccgggacg tgctgacggt tcacccgat tacttttga tcgatcccgg 6600

catcggccgt ttctctacc gcctggcacg ccgcgccgca ggcaaggcag aagccagatg 6660

gttgtcaag acgatctacg aacgcagtgg cagcgcggga gagttcaaga agttctgttt 6720

caccgtgcgc aagctgatcg ggtcaaatga cctgccggag tacgattga aggaggaggc 6780

ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac ctgatcgagg gcgaagcatc 6840

cgccggttc taatgtacgg agcagatgct agggcaaatt gccctagcag gggaaaaagg 6900

tcgaaaagggt ctcttcctg tggatagcac gtacattggg aacccaaagc cgtacattgg 6960

gaaccggaac ccgtacattg ggaacccaaa gccgtacatt ggaaccgggt cacacatgta 7020

agtgactgat ataaaagaga aaaaaggcga ttttccgcc taaaactctt taaaacttat 7080

taaaactctt aaaacccgcc tggcctgtgc ataactgtct ggccagcgca cagccgaaga 7140

gctgcaaaaa ggcctaccc ttcggtcgct gcgtcccta cgccccgccg ctgcgctcg 7200

gcctatcgcg gccgctggcc gctcaaaaat ggctggccta cggccaggca atctaccagg 7260

gcgcggacaa gccgcgccgt cgccactcga ccgccggcgc ccacatcaag gcaccctgcc 7320

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 7380

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 7440

ttggcgggtg tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg 7500

gcttaactat gcggcatcag agcagattgt actgagagtg caccatatgc ggtgtgaaat 7560

accgcacaga tgcgtaagga gaaaataccg catcaggcgc tctccgctt cctcgctcac 7620

tgactcgctg cgctcggtcg ttcggctcgc gcgagcggta tcagctcact caaaggcgtt 7680

aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 7740

gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg ttttcata ggctccgccc 7800

ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 7860

ataaagatac caggcgttc ccctggaag ctccctcgtg cgctctcctg ttccgacct 7920

gccgcttacc ggatacctgt ccgcctttct ccctcggga agcgtggcgc ttctcatag 7980

ctcacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca 8040

cgaaccccc gttcagccc accgctcgc cttatccggt aactatcgtc ttgagtcaa 8100

cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 8160

gaggatgta ggcggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag 8220

aaggacagta ttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 8280

tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggttttttg ttgcaagca 8340

gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatcttt ctacgggggc 8400

tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgcatt ctaggtacta 8460

aaacaattca tccagtaaaa tataatattt tttttctcc caatcaggct tgatccccag 8520

taagtcaaaa aatagctcga catactgttc tccccgata tcctccctga tcgaccggac 8580

gcagaaggca atgtcatacc actgtccgc cctgccgctt ctccaagat caataaagcc 8640

acttactttg ccatctttca caaagatgtt gctgtctccc aggtcgccgt gggaaaagac 8700

aagttcctct tcgggctttt ccgtctttaa aaaatcatc agctcgcgcg gatctttaa 8760

tggagtgtct tcttccagtt ttgcgaatc cacatcggcc agatcggtat tcagtaagta 8820

atccaattcg gctaagcggc tgtctaagct attcgatat ggacaatccg atatgtcgat 8880

ggagtgaaag agcctgatgc actccgcata cagctcgata atctttcag ggctttgtc 8940

atcttcatac tcttccgagc aaaggacgcc atcggcctca ctcagagca gattgtcca 9000

gccatcatgc cgttcaaagt gcaggacctt tggaacaggc agcttcctt ccagccatag 9060

catcatgtcc tttcccggtt ccacatcata ggtggtcctt ttataccggc tgtccgtcat 9120

tttaaatat aggttttcat ttctccac cagcttatat accttagcag gagacattcc 9180

ttccgtatct ttacgcagc ggtattttc gatcagttt ttcaattccg gtgatattct 9240

catttagcc atttattatt tccttctct ttctacagt atttaaagat accccaagaa 9300

gctaattata acaagacgaa ctccaattca ctgttcttg cattctaaaa ccttaaatac 9360

cagaaaacag cttttcaaa gttgtttca aagtggcgt ataacatagt atcgacggag 9420

ccgattttga aaccgcggtg atcacaggca gcaacgctct gtcacgtta caatcaacat 9480

gctaccctcc gcgagatcat ccgtgttca aaccggcag ctagttgcc gttctccga 9540

atagcatcgg taacatgagc aaagtctgcc gcctacaac ggctctccg ctgacgccgt 9600

cccggactga tgggctgcct gtatcgagtg gtgattttgt gccgagctgc cggtcgggga 9660

gctgtggct ggctggggc aggatattt gtggtgtaa caaattgacg ctagacaac 9720

ttaataacac attgcggacg ttttaatgt actgaattaa cgccgaatta attcggggga 9780

tctggatttt agtactggat ttggttta ggaattagaa atttattga tagaagtatt 9840

ttacaaatac aaatacatac taagggtttc ttatatgctc aacacatgag cgaaacccta 9900

taggaaccct aattcccta tctgggaact actcacacat tattatggag aaactcgagc 9960

ttgtcgatcg acagatccgg tcggcatcta ctctatttct ttgccctcg acgagtctg 10020

gggcgtcggg ttccactatc ggcgagtact tctacacagc catcggcca gacggccgcg 10080

cttctcggg cgatttgtg acgcccgaac gtcccggtc cggatcggac gattgcgtcg 10140

catcgaccct gcgccaagc tgcatcatcg aaattgccgt caaccaagct ctgatagagt 10200

tggtaagac caatgcggag catatacgcc cggagtcgtg gcgacctgc aagctccgga 10260

tgctccgct cgaagtagcg cgtctgctgc tccatacaag ccaaccacgg cctccagaag 10320

aagatgttgg cgacctgta ttgggaatcc ccgaacatcg cctcgctcca gtcaatgacc 10380

gctgttatgc ggccattgtc cgtcaggaca ttgtggagc cgaaatccgc gtgcacgagg 10440

tgccggactt cggggcagtc ctcggcccaa agcatcagct catcgagagc ctgcgcgacg 10500

gacgcactga cgggtgctgc catcacagtt tgccagtgat acacatgggg atcagcaatc 10560

gcgcatatga aatcacgcca ttagtgtat tgaccgattc cttgcggtcc gaatgggccg 10620

aaccgctcg tctggctaag atcggccgca gcgatcgcat ccatagcctc cgcgaccggt 10680

tgtagaacag cgggcagttc gggttcaggc aggtcttgca acgtgacacc ctgtgcacgg 10740

cgggagatgc aataggtcag gctctgcta aactcccaa tgtcaagcac ttccggaatc 10800

gggagcgcg ccatgcaaa gtgccgataa acataacgat cttttagaa accatcggcg 10860

cagctattta cccgcaggac atatccacgc cctcctacat cgaagctgaa agcacgagat 10920

tcttcgcct cggagagctg catcaggctg gagacgctgt cgaactttc gatcagaaac 10980

ttctcgacag acgtcgcggt gagttcaggc ttttcatat ctattgccc cccgggatct 11040

gcgaaagctc gagagagata gattttaga gagagactgg tgattcagc gtgtcctctc 11100

caaatgaaat gaacttcctt atatagagga aggtcttgcg aaggatagtg ggattgtgcg 11160

tcatccctta cgtcagtga gatatcacat caatccactt gctttgaaga cgtggttga 11220

acgtcttctt ttccacgat gctcctcgtg ggtgggggtc catctttggg accactgtcg 11280

gcagaggcat ctgaacgat agcctttcct ttatcgcaat gatggcattt gtaggtgcca 11340

ccttcctttt ctactgtcct ttgatgaag tgacagatag ctgggcaatg gaatccgagg 11400

aggtttcccg atattaccct ttgtgaaaa gtctcaatag ccctttggtc ttctgagact 11460

gtatctttga tattcttga gtagacgaga gtgtcgtgct ccaccatgtt atcacatcaa 11520

tccacttgct ttgaagacgt ggttgaacg tcttctttt ccacgatgct cctcgtgggt 11580

gggggtccat cttgggacc actgtcggca gaggcattt gaacgatagc ctttcctta 11640

tcgcaatgat ggcatttga ggtgccacct tccttttcta ctgtccttt gatgaagtga 11700

cagatagctg ggcaatggaa tccgaggagg ttcccgata ttacccttg ttgaaaagtc 11760

tcaatagccc ttggtcttc tgagactgta tcttgatat tcttgagta gacgagagt 11820

tcgtgtcca ccatgttggc aagctgtct agccaatacg caaaccgcct ctccccgcgc 11880

gttgccgat tcattaatgc agctggcacg acaggttcc cgactggaaa gcgggcagt 11940

agcgcaacgc aattaatgtg agttagtca ctattaggc acccaggct ttacactta 12000

tgctccggc tcgtatgtg tgtggaattg tgagcggata acaattcac acaggaaaca 12060

gctatgacca tgattacgaa ttccgatct agtaacatag atgacaccgc gcgcgataat 12120

ttatcctagt ttgcgcgcta tatttggtt tctatcgcgt attaaatga taattgcggg 12180

actctaataca taaaaaccca ttcataaat aacgtcatgc attacatgtt aattattaca 12240

tgcttaacgt aattcaacag aaattatatg ataatcatcg caagaccggc aacaggattc 12300

aatcttaaga aactttattg ccaaattgtt gaacgatcgg ggaaattcga gctcttatgc 12360

tttctagtc aataatatgg taacattgat gaacttggtt ttctcctgg acatacggtc 12420

agccaagatt tgctggtagc ggcaaaaaac ctctcaatg atagcttcac caaagtggct 12480

aatcagcata ggttctgcca cagccctcat gcactgtgcc acattgtatc caccatcact 12540

aagtgattca gaccttcag attcaaagtc tagagcattc cactcatcga gagcattcca 12600

attcacctca gaaacctcta gacgattgat ggcgaatgat ccttcttaa gaactccaa 12660

tttcacttca gatggggatg gagtgattg agggatgtta aaagtatcta attgctcttc 12720

tcttatgatt ccctgcaaga ccatatcact aagagcagta gccaaaagct cccaaatgta 12780

gcaaccatcc ttgctagatg gatcatcgct tcttctccc aaaaatgtga gaaccatgcg 12840

tcctcttca accaattcct ctgcacgaca ctgagaaaa agagagaaat ctctctgaaa 12900

ttgctcatag taagcttca caacatttt ggggctcgta ctgcctatgt aaacattgcc 12960

ctgtgttg tctacacct caggaacctt agatagccat tgaaggctgt atgaggaatg 13020

gacaaaatga agactttgat atgggaaaac cctccatag aaagaaccgg gaacaccga 13080

gaagtagcat ggaccgatcc cactttccat tcatcacac aatttctt tgaagctgtc 13140

aagagacttg aagatgtgt tgaagtcatt ccagggaga tcattcagaa agactttgta 13200

ttctggagat ttgtggttca gctctcggca aagcttctcc acaattttta tgaattcaga 13260

cacaacaaat aaagtgttg gtccagaaga gcaacccaaa tctgcaattg ccaagcttct 13320

ggggtgtgtg ctgcaataga ggcttcttat ggcttcctct cttatgggct ttgtcaaaca 13380

aatcaccttt tgctgaacaa gggagtgtt tgcatagctt gcgtctcaa cgccaccgtt 13440

catgtggagt acctgtgcta ctccatgga tctctagag tccccgtgt tctctccaaa 13500

tgaaatgaac ttcctatat agaggaaggg tcttgcaag gatagtggga ttgtgctca 13560

tccctacgt cagtggagat atcacatcaa tccactgct tgaagacgt ggttgaacg 13620

tctcttttt ccacgatgct cctcgtgggt ggggttccat cttgggacc actgtcggca 13680

gaggcatctt caacgatggc ctttcttta tcgcaatgat ggcatttgta ggagccacct 13740

tcctttcca ctatctcac aataaagtga cagatagctg ggcaatggaa tccgaggagg 13800

ttccggata ttacccttg ttgaaaagtc tcaattgcc tttgtcttc tgagactgta 13860

tcttgatat tttggagta gacaagtgtg tcgtgctcca ccatgtgac gaagatttc 13920

ttctgtcat tgagtcgtaa gagactctgt atgaactgtt cgccagctt tacggcgagt 13980

tctgttaggt cctctatttg aatcttgac tccatggcct ttgattcagt gggaactacc 14040

ttttagaga ctccaatctc tattacttgc cttggtttgt gaagcaagcc ttgaatcgtc 14100

catactggaa tagtacttct gatcttgaga aatatatctt tctctgtgtt cttgatgcag 14160

ttagtcctga atctttgac tgcatctta accttctgg gaaggattt gatttcctgg 14220

agattattgc tcgggtagat cgtcttgatg agacctgctg cgtaagcctc tctaaccatc 14280

tgtgggtag cattcttct gaaattgaaa aggctaactt ggggacctgc aggcattgcaa 14340

gcttggcact ggccgtcgtt ttacaacgtc gtgactggga aaacctggc gttaccaac 14400

ttaatgcct tgcagacat cccctttcg ccagctggcg taatagcgaa gagggccgca 14460

ccgatcgccc ttccaacag ttgcgcagcc tgaatggcg atgctagagc agcttgagct 14520

tggatcagat tgtcgtttcc cgccttcagt ttagcttcat ggagtcaaag attcaaatag 14580

aggacctaac agaactcgcc gtaaagactg gcgaacagtt catacagagt ctctacgac 14640

tcaatgacaa gaagaaaatc ttcgtcaaca tggaggagca cgacacactt gtctactcca 14700

aaaatatcaa agatacagtc tcagaagacc aaagggcaat tgagacttt caacaaaggg 14760

taatatccgg aaacctctc ggattccatt gccagctat ctgtcactt attgtgaaga 14820

tagtggaaaa ggaagggtgc tcctacaaat gccatcattg cgataaagga aaggccatcg 14880

ttgaagatgc ctctgccgac agtgggtcca aagatggacc cccaccacg aggagcatcg 14940

tggaaaaaga agacgttcca accacgtctt caaagcaagt ggattgatgt gatattcca 15000

ctgacgtaag ggatgacgca caatccact atccttcgca agacccttc tctatataag 15060

gaagttcatt tcatttgag agaacacggg ggactctga c 15101

<210> 4

<211> 14774

<212> ADN

<213> Secuencia Artificial

<220>

<223> Constructo de plásmido sintético pJL-OFP-35S::GmSABP2-1

<400> 4

tcatggctct ttcaaagcaa agtgggggtca aagatgtaat gaacaccgag cttcatatgg 60

acgggatcgt caatggacac cccttgaga taaaaggaaa aggaaaggga aaccgtaca 120

agggtgtgca gaccatgaaa cttacagtca ttaagggtgc gccttgcca tttctattg 180

acatttgct gcctcaacac atgtatggaa gcaagccatt tattaagtat cctgagagta 240

tccagacta catcaagtg tcatttccg agggaatcac atgggaaagg tccatgacct 300

ttgaagatgg tgcagtgtgc actgtctcta acgactccag gtcgatggc gactcttca 360

tctacgaagt caggtttctt ggcgtgaact ttccccgaga tggacctgtt atgcagaaga 420

agacgcgagg ctgggaccg tccacagaga gactgtatga gtgtggtggg tggcagagag 480

gagacgtcca catggccttg aagtggaga acggtggcca ttatactgc gacttcaaaa 540

ctacttaca atcaaagaag ggcttgaagg tgccaccgta tcacttcgtt gaccacaaac 600

tagatttact gagccacaac accgatggg ctaccttga agagttgaa caacgagaaa 660

ttgcacatgc acatcttct aactaccgg tagcccatca ccatcaccat cactaagatc 720

gttcaaacat ttggcaataa agtttcttaa gattgaatcc tgttgccggt cttgcatga 780

ttatcatata atttctgtg aattacgtta agcatgtaat aattaacatg taatgcatga 840

cggtatttat gagatgggtt ttatgatta gagtcccga attatacatt taatacgca 900

tagaaaacaa aatatagcgc gcaaaactagg ataaattatc gcgcgcggtg tcatctatgt 960

tactagatcc catggctact actaagcatt tggctctgc catcctgtc ctccttagca 1020

ttggtatgac caccagtga agaaccctcc tagatctgag ggtaaatttc tagttttct 1080

ccttcatttt ctgggttagg accctttct cttttattt tttgagctt tgatcttct 1140

ttaaactgat ctattttta attgattggt tatggtgtaa atattacata gctttaactg 1200

ataatctgat tactttattt cgtgtgtcta tgatgatgat gatagttaca gaaccgacga 1260

actagtctgt acccgatcaa caccgagacc cgtggcgtct tcgacctcaa tggcgtctgg 1320

aactcaagc tggactacgg gaaaggactg gaagagaagt ggtacgaaag caagctgacc 1380

gacactatta gtagggcgt cccaagcagt tacaatgaca ttggcgtgac caaggaaatc 1440

cgcaaccata tcggatatgt ctggtacgaa cgtgagttca cggtgccggc ctatctgaag 1500

gatcagcgta tcgtgctccg cttcggctct gcaactcaca aagcaattgt ctatgtcaat 1560

ggtgagctgg tcgtggagca caagggcgga ttctgcat tcgaagcgga aatcaacaac 1620

tcgtgcgtg atggcatgaa tcgcgtcacc gtcgccgtgg acaacatcct cgacgatagc 1680

accctcccgg tggggctgta cagcgagcgc cacgaagagg gcctcgaaa agtcattcgt 1740

aacaagccga acttcgactt ctcaactat gcaggcctgc accgtccgtt gaaaatctac 1800

acgaccccggt ttacgtacgt cgaggacatc tcggttgta ccgactcaa tggcccaacc 1860

gggactgtga cctatacggg ggactttcaa ggcaaagccg agaccgtgaa agtgtcggtc 1920

gtggatgagg aaggcaaagt ggtcgcaagc accgagggcc tgagcggtaa cgtggagatt 1980

ccgaatgtca tcctctggga accactgaac acgtatctct accagatcaa agtggaaactg 2040

gtgaacgacg gactgacat cgatgtctat gaagagccgt tcggcgtgcg gaccgtggaa 2100

gtcaacgacg gcaagttcct catcaacaac aaaccgttct acttaaggg ctttggcaaa 2160

catgaggaca ctctatcaa cggccgtggc tttaacgaag cgagcaatgt gatggatttc 2220

aatatcctca aatggatcgg cgccaacagc ttccggaccg cacactatcc gtactctgaa 2280

gagttgatgc gtcttcgga tcgcgagggg ctggctgta tcgacgagac tccggcagtt 2340

ggcgtgcacc tcaactcat ggccaccagc ggactcggcg aaggcagcga gcgcgtcagt 2400

acctgggaga agattcggac gtttgagcac catcaagacg ttctccgtga actggtgtct 2460

cgtgacaaga accatccaag cgtcgtgatg tggagcatcg ccaacgaggc ggcgactgag 2520

gaagagggcg cgtacgagta cttaagccg ttggtggagc tgaccaagga actcgacca 2580

cagaagcgtc cggtcacgat cgtgctgttt gtgatggcta ccccgagac ggacaaagtc 2640

gccgaactga ttgacgtcat cgcgtcaat cgctataacg gatggtactt cgatggcgg 2700

gatctgaag cggccaaagt ccattccgc caggaatttc acgcgtggaa caagcgttgc 2760

ccaggaaagc cgatcatgat cactgagtac ggcgcagaca ccgttcggg ctttcacgac 2820

attgatccag tgatgttcac cgaggaatat caagtcgagt actaccaggc gaaccacgtc 2880

gtgttcgatg agtttgagaa cttcgtgggt gagcaagcgt ggaacttcgc ggacttcgcg 2940

acctctcagg gcgtgatgcg cgtccaagga aacaagaagg gcgtgttcac tcgtgaccgc 3000

aagccgaagc tcgccgcgca cgtctttcgc gagcgctgga ccaacattcc agatttcggc 3060

tacaagaacg ctagccatca ccatcaccat cacgtgtgaa ttggtgacca gctcgaattt 3120

ccccgatcgt tcaaacattt ggcaataaag ttcttaaga ttgaatcctg ttgccggtct 3180

tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta 3240

atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgcaat tatacattta 3300

atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc 3360

atctatgtta ctagatcggg aattaaacta tcagtgtttg acaggatata ttggcgggta 3420

aacctaagag aaaagagcgt ttattagaat aacggatatt taaaagggcg tgaaaaggtt 3480

tatccgttcg tccatttgta tgtgcatgcc aaccacaggg ttcccctcgg gatcaaagta 3540

ctttgatcca acccctccgc tgctatagtg cagtcggctt ctgacgttca gtgcagccgt 3600

cttctgaaaa cgacatgtcg cacaagtcct aagttacgcg acaggctgcc gccctgccct 3660

ttcctggcg tttcttgc gcgtgttta gtcgcataaa gtagaatact tgcgactaga 3720

accggagaca ttacgcatg aacaagagcg ccgccgctgg cctgctgggc tatgccgcg 3780

tcagcaccga cgaccaggac ttgaccaacc aacgggccga actgcacgcg gccggctgca 3840

ccaagctgtt ttccgagaag atcaccggca ccaggcgcga ccgccggag ctggccagga 3900

tgcttgacca cctacgccct ggcgacgttg tgacagtac caggctagac cgcttgccc 3960

gcagaccccg cgacctactg gacattgccg agcgcatcca ggaggccggc gcgggcctgc 4020

gtagcctggc agagccgtgg gccgacacca ccacgccggc cggccgcatg gtgttgaccg 4080

tgctgcggc cattgccgag ttcgagcgtt ccctaactat cgaccgcacc cggagcgggc 4140

gcgaggccgc caaggcccga ggcgtgaagt ttggcccccg ccctaccctc acccggcac 4200

agatcgcgca cgcccgcgag ctgatcgacc aggaaggccg caccgtgaaa gaggcggctg 4260

cactgcttg cgtgcatgc tcgacctgt accgcgact tgagcgagc gaggaagtga 4320

cgccaccga ggccaggcgg cgcggtgcct tccgtgagga cgcattgacc gaggccgacg 4380

ccctggcggc cgccgagaat gaacgccaag aggaacaagc atgaaaccgc accaggacgg 4440

ccaggacgaa ccgttttca ttaccgaaga gatcgaggcg gagatgatcg cggccgggta 4500

cgtgtcgag ccgcccgcg acgtctaac cgtcggtg catgaaatcc tggccggtt 4560

gtctgatgcc aagctggcgg cctggccggc cagcttgcc gctgaagaaa ccgagcgccg 4620

ccgtctaaaa aggtgatgtg tatttgagta aaacagcttg cgtcatgcgg tcgctgcgta 4680

tatgatcgga tgagtaaata acaaatacg caaggggaac gcatgaaggt tatcgctgta 4740

cttaaccaga aaggcgggtc aggcaagacg accatcgcaa cccatctagc ccgcccctg 4800

caactcgccg gggccgatgt tctgttagtc gattccgatc ccagggcag tgcccgcgat 4860

tgggcggccg tgcgggaaga tcaaccgcta accgtgtcg gcatcgaccg cccgacgatt 4920

gaccgcgacg tgaaggccat cggccggcgc gacttcgtag tgatcgacgg agcgccccag 4980

gcggcggact tggctgtgtc cgcatcaag gcagccgact tcgtgctgat tccggtgcag 5040

ccaagccctt acgacatatg ggccaccgcc gacctgggtg agctgggtta gcagcgcat 5100

gaggtcacgg atggaaggct acaagcggcc ttgtcgtgt cggggcat caaaggcacg 5160

cgcatcggcg gtgaggttgc cgaggcgtg gccgggtacg agctgcccat tctgagtcc 5220

cgtatcacgc agcgctgag ctaccaggc actgccgccg ccggcacaac cgttctgaa 5280

tcagaacccg agggcgacgc tgcccgag gtccaggcgc tggccgctga aattaaatca 5340

aaactcattt gagttaatga ggtaaagaga aaatgagcaa aagcacaac acgctaagt 5400

ccggccgtcc gagcgcacgc agcagcaagg ctgcaacgtt ggccagcctg gcagacacgc 5460

cagccatgaa gcgggtcaac tttagttgc cggcggagga tcacaccaag ctgaagatgt 5520

acgcggtacg ccaaggcaag accattaccg agctgctatc tgaatacatc gcgcagctac 5580

cagagtaa at gagcaaatga ataatgagt agatgaatt tagcggctaa aggaggcggc 5640

atggaaaatc aagaacaacc aggcaccgac gccgtggaat gcccattgtg tggaggaacg 5700

ggcggttggc caggcgtaag cggctgggtt gtctgccggc cctgcaatgg cactggaacc 5760

ccaagccc aggaatcggc gtgacggtcg caaacatcc ggcccgttac aaatcggcgc 5820

ggcgtgggt gatgacctg tggagaagt gaaggccgc caggccgcc agcggcaacg 5880

catcgaggca gaagcacgcc ccggtgaatc gtggcaagcg gccgtgatc gaatccgcaa 5940

agaatcccgg caaccgccgg cagccgggtgc gccgtcgatt aggaagccgc ccaagggcga 6000

cgagcaacca gatttttcg ttccgatgct ctatgacgtg ggcacccgcg atagtcgcag 6060

catcatggac gtggccgttt tccgtctgtc gaagcgtgac cgacgagctg gcgaggtgat 6120

ccgctacgag ctccagacg ggcacgtaga ggttccgca gggccggccg gcatggccag 6180

tgtgtgggat tacgacctgg tactgatggc ggtttcccat ctaaccgaat ccatgaaccg 6240

ataccgggaa gggaagggag acaagcccgg ccgcgtgttc cgtccacacg ttgcggacgt 6300

actcaagttc tgccggcgag ccgatggcgg aaagcagaaa gacgacctgg tagaaacctg 6360

cattcggtta aacaccacgc acgttgccat gcagcgtacg aagaaggcca agaacggccg 6420

cctggtgacg gtatccgagg gtgaagcctt gattagccgc tacaagatcg taaagagcga 6480

aaccgggcgg ccggagtaca tcgagatcga gctagctgat tggatgtacc gcgagatcac 6540

agaaggcaag aaccgggacg tgctgacggt tcaccccgat tacttttga tcgatcccgg 6600

catcggccgt ttctctacc gcctggcacg ccgcgccgca ggcaaggcag aagccagatg 6660

gttgtcaag acgatctacg aacgcagtgg cagcgccgga gagtcaaga agttctgttt 6720

caccgtgcgc aagctgatcg ggtcaaatga cctgccggag tacgatttga aggaggaggc 6780

ggggcaggct ggcccgatcc tagtcatgcg ctaccgcaac ctgatcgagg gcgaagcatc 6840

cgccggttcc taatgtacgg agcagatgct agggcaaatt gccctagcag gggaaaaagg 6900

tcgaaaaggt ctcttcctg tggatagcac gtacattggg aacccaaagc cgtacattgg 6960

gaaccggaac ccgtacattg ggaacccaaa gccgtacatt gggaaccggt cacacatgta 7020

agtgactgat ataaaagaga aaaaaggcga ttttccgcc taaaactctt taaaacttat 7080

taaaactctt aaaacccgcc tggcctgtgc ataactgtct ggccagcgca cagccgaaga 7140

gctgcaaaaa gcgctaccc ttcggtcgct gcgctcccta cgccccgccg ctcgcgctg 7200

gcctatcgcg gccgctggcc gctcaaaaat ggctggccta cggccaggca atctaccagg 7260

gcgcggaaca gccgcgccgt cgccactga ccgccggcg ccacatcaag gcaccctgcc 7320

tcgcgcggtt cggatgatgac ggtgaaaacc tctgacacat gcagctccc gagacggta 7380

cagcttgtct gtaagcggat gccgggagca gacaagccc tcagggcgcg tcagcgggtg 7440

ttggcgggtg tcggggcgca gccatgaccc agtcacgtag cgatagcgga gtgtatactg 7500

gcttaactat gcggcatcag agcagattgt actgagagt caccatatgc ggtgtgaaat 7560

accgcacaga tgcgtaagga gaaaataccg catcaggcg tcttccgctt cctcgctcac 7620

tgactcgctg cgctcggtcg ttcggctgcg gcgagcggtc tcagctcact caaaggcgt 7680

aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca 7740

gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttccata ggctccgcc 7800

ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact 7860

ataaagatac caggcgttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct 7920

gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcg tttctcatag 7980

ctcacgctgt aggtatctca gttcgggtga ggtcgttcgc tccaagctgg gctgtgtgca 8040

cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa 8100

cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc 8160

gaggtatgta ggcgggtgcta cagagttctt gaagtgggtg cctaactacg gctacactag 8220

aaggacagta ttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg 8280

tagctcttga tccggcaaac aaaccaccgc tggtagcggg ggttttttg ttgcaagca 8340

gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatcttt ctacgggggc 8400

tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgcatt ctaggtacta 8460

aaacaattca tccagtaaaa tataatattt tattttctcc caatcaggct tgatccccag 8520

taagtcaaaa aatagctcga catactgttc tccccgata tcctccctga tcgaccggac 8580

gcagaaggca atgtcatacc actgtccgc cctgccgctt ctccaagat caataaagcc 8640

acttactttg ccatctttca caaagatgtt gctgtctccc aggtcgccgt gggaaaagac 8700

aagttcctct tcgggctttt ccgtcttaa aaaatcatac agctcgcgcg gatctttaa 8760

tggagtgtct tctcccagt ttgcgaatc cacatcggcc agatcggtat tcagtaagta 8820

atccaattcg gctaagcggc tgtctaagct attcgatat ggacaatccg atatgtcgat 8880

ggagtgaaag agcctgatgc actccgcata cagctcgata atcttttcag ggctttgttc 8940

atcttcatac tctccgagc aaaggacgcc atcggcctca ctcagagca gattgtcca 9000

gccatcatgc cgttcaaagt gcaggacctt tggaacaggc agcttcctt ccagccatag 9060

catcatgtcc tttcccggtt ccacatcata ggtggtcctt ttataccggc tgtccgtcat 9120

tttaaatat aggttttcat ttctcccac cagcttatat accttagcag gagacattcc 9180

ttccgtatct ttacgcagc ggtattttc gatcagttt tcaattccg gtgatattct 9240

catttagcc atttattatt tccttcctct ttctacagt atttaaagat accccaagaa 9300

gctaattata acaagacgaa ctccaattca ctgttccttg cattctaaaa ccttaaatac 9360

cagaaaacag cttttcaaa gtgttttca aagttggcgt ataacatagt atcgacggag 9420

ccgattttga aaccgcggtg atcacaggca gcaacgctct gtcacgtta caatcaacat 9480

gctaccctcc gcgagatcat ccgtgtttca aaccggcag cttagttgcc gttcttccga 9540

atagcatcgg taacatgagc aaagtctgcc gccttacaac ggctctcccg ctgacgccgt 9600

cccggactga tgggctgcct gtatcgagtg gtgattttgt gccgagctgc cggtcgggga 9660

gctgttggtt ggctggtggc aggatatatt gtggtgtaaa caaattgacg ctagacaac 9720

ttaataacac attgcggacg ttttaatgt actgaattaa cgccgaatta attcggggga 9780

tctggatttt agtactggat ttgggttta ggaattagaa atttattga tagaagtatt 9840

ttacaaatac aaatacatat taagggtttc ttatatgtc aacacatgag cgaaacccta 9900

taggaaccct aattccctta tctgggaact actcacacat tattatggag aaactcgagc 9960

ttgtcgatcg acagatccgg tcggcatcta ctctatttct ttgccctcgg acgagtgctg 10020

gggcgtcggg ttccactatc ggcgagtact tctacacagc catcgggtcca gacggccgcg 10080

cttctcgggg cgatttgtgt acgcccgaca gtcccggctc cggatcggac gattgcgtcg 10140

catcgaccct gcgcccgaagc tgcacatcg aaattgccgt caaccaagct ctgatagagt 10200

tggtaagac caatgcggag catatacgcc cggagtcgtg gcgacacctg aagctccgga 10260

tgctccgct cgaagtagcg cgtctgctgc tccatacaag ccaaccacgg cctccagaag 10320

aagatgttg cgacctgta ttgggaatcc ccgaacatcg cctcgtcca gtcaatgacc 10380

gctgttatgc ggccattgtc cgtcaggaca ttgtggagc cgaaatccgc gtgcacgagg 10440

tgccggactt cggggcagtc ctgggcccga agcatcagct catcgagagc ctgcgcgacg 10500

gacgcactga cgggtgctgc catcacagtt tgccagtgat acacatgggg atcagcaatc 10560

gcgcatatga aatcacgcca ttagtgtat tgaccgattc cttgcgggcc gaatgggccc 10620

aaccgctcg tctggctaag atcgggccga gcgacgcat ccatagcctc cgcgaccggt 10680

tgtagaacag cgggcagttc ggttcaggc aggtcttgca acgtgacacc ctgtgcacgg 10740

cgggagatgc aataggtagc gctctcgta aactcccaa tgtcaagcac ttccggaatc 10800

gggagcgcgg ccgatgcaaa gtgccgataa acataacgat cttttagaa accatcggcg 10860

cagctattta cccgcaggac atatccacgc cctctacat cgaagctgaa agcacgagat 10920

tcttcgcct cggagagctg catcaggctg gagacgctgt cgaactttc gatcagaaac 10980

ttctcgacag acgtcgcggg gagttcaggc ttttcatat ctattgccc cccgggatct 11040

gcgaaagctc gagagagata gattttaga gagagactgg tgatttcagc gtgtcctctc 11100

caaatgaaat gaacttcctt atatagagga aggtcttgcg aaggatagtg ggattgtgcg 11160

tcatccctta cgtcagtgga gatatcacat caatccactt gcttgaaga cgtgggtgga 11220

acgtctctt ttccacgat gctcctcg tggtgggggtc catcttggg accactgtcg 11280

gcagaggcat ctgaacgat agccttctt ttatcgcaat gatggcattt gtaggtgcca 11340

ccttccttt ctactgtcct ttgatgaag tgacagatag ctgggcaatg gaatccgagg 11400

aggttcccg atattaccct ttgtgaaaa gtctcaatag cccttggtc ttctgagact 11460

gtatcttga tattcttga gtagacgaga gtgtcgtgct ccaccatgtt atcacatcaa 11520

tccacttgc ttgaagacgt ggttgaacg tctctttt ccacgatgct cctcgtgggt 11580

gggggtccat cttgggacc actgtcggca gaggcattt gaacgatagc cttccttta 11640

tcgcaatgat ggcatttga ggtgccacct tcctttcta ctgtccttt gatgaagtga 11700

cagatagctg ggcaatggaa tccgaggagg ttcccgata ttacccttg ttgaaaagtc 11760

tcaatagccc ttggtcttc tgagactgta tcttgatat tcttgagta gacgagagt 11820

tcgtgctcca ccatgttggc aagctgctct agccaatacg caaaccgcct ctccccgcg 11880

gttggccgat tcattaatgc agctggcacg acaggttcc cgactggaaa gcgggcagt 11940

agcgcaacgc aattaatgtg agttagctca ctcataggc accccaggct ttacacttta 12000

tgctccggc tcgtatgtg tgtggaattg tgagcggata acaattcac acaggaaaca 12060

gctatgacca tgattacgaa ttcccgatct agtaacatag atgacaccgc gcgcgataat 12120

ttatcctagt ttgcgcgcta tattttgtt tctatcgct attaaatgta taattgcggg 12180

actctaata taaaaacca tctcataaat aacgtcatgc attacatgtt aattattaca 12240

tgcttaacgt aattcaacag aaattatatg ataatcatcg caagaccggc aacaggattc 12300

aatcttaaga aactttattg ccaaattgtt gaacgatcgg ggaaattcga gctctcatgc 12360

atatttagtc gctatctgct ggagggaatc gaatagtct tgtggcttgc aaagcataac 12420

catatgatct gcgtcttga cctctacaac gtcattgaac ccaacattt ggatcatcaa 12480

gagctgatat tccaatggaa ttccaaggtc ctccgtgcaa acaataaagg cacgtggaac 12540

tgaccatat ccctcttgg agaagttct ttgttagtc aagtctcca tgaagagcga 12600

tgatggcctt gctaaagtct tggccaattc caagtcctca ataggggaca gttggtagag 12660

ctgtcggac aagaagttgg ggccgaagaa cattgatgtt ttgttccgc ttggagcaaa 12720

ttcagtgtct aaccatgcag ctaacggggc cctctcattg tactttcca agacataaga 12780

tggttggtgt tcagtgtctg gagcaaaagc tgtaagaaa acaccaactg ctacctttc 12840

tgggaatttc tccattgcaa gtgctatgtt cagccctcca aggctgtgac caactagaac 12900

taacttctca ttgagggaa ttgtggccat tagctgcaac aaaggcgcag aatactctga 12960

gaaagtatca acatcttcaa tttcttcat gttggttcca gaagctgcaa ggtcaagtac 13020

tgtgacctta tggcctgcag attccaagcg tggcttgagc ttataccaac accaagctcc 13080

atggcatgcc ccatgcacca gaacatagtg cttcctatcc atacaatttt gtgaacccat 13140

ggatcctcta gagtcccccg tgttctctcc aaatgaaatg aacttcctta tatagaggaa 13200

gggtcttgcg aaggatagtg ggattgtgcg tcatccctta cgtcagtgga gatatcacat 13260

caatccactt gcttgaaga cgtgggtgga acgtctctt ttccacgat gtcctctgtg 13320

ggtgggggtc catctttggg accactgtcg gcagaggcat cttcaacgat ggcctttcct 13380

ttatcgcaat gatggcattt gtaggagcca ccttcctttt ccactatctt cacaataaag 13440

tgacagatag ctgggcaatg gaatccgagg aggtttccgg atattaccct ttgtgaaaa 13500

gtctcaattg ccttttggc ttctgagact gtatctttga tatttttga gtagacaagt 13560

gtgtcgtgct ccaccatgtt gacgaagatt ttctcttgt cattgagtcg taagagactc 13620

tgtatgaact gttcgccagt ctttacggcg agttctgtta ggtcctctat ttgaatcttt 13680

gactccatgg cctttgattc agtgggaact accttttag agactccaat ctctattact 13740

tgccctgggt tgtgaagcaa gccttgaatc gtccatactg gaatagtact tctgatcttg 13800

agaaatatat ctttctctgt gttcttgatg cagttagtcc tgaatctttt gactgcatct 13860

ttaaccttct tgggaaggta ttgatttcc tggagattat tgctcgggta gatcgtcttg 13920

atgagacctg ctgcgtaagc ctctctaacc atctgtgggt tagcattctt tctgaaattg 13980

aaaaggctaa tctggggacc tgcaggcatg caagcttggc actggccgct gttttacaac 14040

gtcgtgactg ggaaaaccct ggcgttacct aacttaatcg ccttcagca catccccctt 14100

tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca 14160

gcctgaatgg cgaatgctag agcagcttga gcttggatca gattgtcgtt tcccgccttc 14220

agtttagctt catggagtca aagattcaaa tagaggacct aacagaactc gccgtaaaga 14280

ctggcgaaca gttcatacag agtctcttac gactcaatga caagaagaaa atcttcgtca 14340

acatggtgga gcacgacaca ctgtctact ccaaaaatat caaagataca gtctcagaag 14400

accaaagggc aattgagact tttaacaaa gggtaatatc cggaacctc ctcggattcc 14460

attgccagc tatctgtcac ttattgtga agatagtga aaaggaaggt ggctcctaca 14520

aatgccatca ttgcgataaa ggaaaggcca tcgttgaaga tgcctctgcc gacagtggtc 14580

caaagatgg acccccaccc acgaggagca tcgtggaaaa agaagacgtt ccaaccacgt 14640

cttcaaagca agtggattga tgtgatatct cactgacgt aagggatgac gcacaatccc 14700

actatccttc gcaagaccct tcctctatat aaggaagttc atttcattg gagagaacac 14760

gggggactct tgac 14774

- <210> 5
- <211> 27
- <212> ADN
- <213> Secuencia Artificial
- <220>
- <223> Cebador de oligonucleótido sintético
- <400> 5

atggaagtag cacaggtact ccacatg 27

<210> 6

<211> 27

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 6

tgcttttcta gtcaataata tggtaac 27

<210> 7

<211> 27

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 7

gtgaaattgg aagttcttaa agaagga 27

<210> 8

<211> 26

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 8
cagattcaaa gtctagagca ttccac 26

<210> 9
<211> 22
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 9
gtgtaatgtt ggatgtgttc cc 22

<210> 10
<211> 25
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 10
acacaattga gttcaacaca aaccg 25

<210> 11
<211> 28
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 11
aactatgctc cccctggcaa ctatattg 28

<210> 12
<211> 30
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 12
tctgaagtgg tagcttctac atcgaaacaa 30

<210> 13
<211> 26
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 13
aactacaatt acgggcaagc tggcaa 26

<210> 14
<211> 25
<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 14

ttgatggctt gtttcctgt gcagt 25

<210> 15

<211> 27

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 15

atgggttcac aaaattgtat ggatagg 27

<210> 16

<211> 27

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 16

tcatgcatat ttagtcgcta tctgctg 27

<210> 17

<211> 255
<212> ADN
<213> Secuencia Artificial

<220>
<223> Sonda de oligonucleótido sintético Gma.12911.1.A1_S_AT

<220>
<221> misc_feature
<222> (108)..(108)
<223> n es a, c, g, o t

<400> 17
tgccgctacc agcaaattct ggctgaacgt atgtccaagg agaaaaccaa gttcatcaat 60

gttaccatat tattgactag aaaagcataa gcaagagttt gcatctanat atggtttctc 120

aaaatggatt gttgtagttg ttgtgagttg tgactgttg cttgtttgtg ttttttcta 180

gcaattaact ggcaaaatgt taagttggag aaaaagaaaa ctagtatgac gctgtacctt 240

tggtagtaag ttct 255

<210> 18
<211> 252
<212> ADN
<213> Glycine max

<400> 18
agccgctacc aacaaatctt ggctgacgt atgtctaagg agaaaactaa gttcaccaat 60

gtcaccatat tattgactaa aaaggcataa gaagtttgca tctattaatg gtttctcaaa 120

ttggattgtt atggtaatgt ggttggtgtg actgttgct tgttgttg tttctagca 180

atgaactagc aaaatgttaa gttggagaaa aaaaaaacta gtatgatgtt gtacctttg 240

tagtaactat ta 252

<210> 19

<211> 257

<212> ADN

<213> Glycine max

<400> 19

agccgctacc aacaaatctt gactgacgt atgtccaagg agcaaactaa gttcatcaat 60

gtcaccgtat tattgactag aaaagcataa gcaagaagtt tgcactata tatggtttct 120

caaatggtat tgctgtagtc atagttaaaa atccggttct actagtcagc tcggatcgggt 180

ccaattgaga tctgatgata taacctttcc gatttcgtta ttgaaccgat catgcattg 240

atctggaatg acccgggt 257

<210> 20

<211> 249

<212> ADN

<213> Glycine max

<400> 20

agccgctacc aacaaatctt gactgacgt atgtccaagg agcaaactaa gttcatcaat 60

gtcaccgtat tattgactag aaaagcataa gcaagaagtt tgcactata tagggtttct 120

caaatgggat tgtttagtc atactttaa aatccggctc gactagcgtg gtcggattga 180

gatttgatga cataacctgt ccgatttat tattggaccg attatgttt ttgtctagaa 240

taacccggt 249

<210> 21

<211> 233

<212> ADN

<213> Glycine max

<400> 21

agccgttacc aacaaatttt ggctgatcgt atgtctaagg agaaaactaa gttcaccaat 60

gcaagaagtt tgcatttata gcaaaacata gctggtcaga gaaggcataa ttctagttc 120

tattcactat ccccttccta ctactactct aaatttctcc cttttgaca tcatcaaaaa 180

ggaagaatct ctaaccagga caagctcgta tggggaatat agagctacta gaa 233

<210> 22

<211> 255

<212> ADN

<213> Glycine max

<400> 22

tgccgctacc agcaaattct ggctgaacgt atgtccaagg agaaaaccaa gttcatcaat 60

gttaccatat tattgactag aaaagcataa gcaagagttt gcatctatat atggtttctc 120

aaaatggatt gttgtagttg ttgtgagttg tgactgtttg cttgtttgtg ttttttcta 180

gcaattaact ggcaaaatgt taagttggag aaaaagaaaa ctagtatgac gctgtacctt 240

tggtagtaag tttct 255

<210> 23

<211> 261

<212> PRT

<213> Glycine Max

<400> 23

Met Gly Ser Gln Asn Cys Met Asp Arg Lys His Tyr Val Leu Val His
1 5 10 15

Gly Ala Cys His Gly Ala Trp Cys Trp Tyr Lys Leu Lys Pro Arg Leu
20 25 30

Glu Ser Ala Gly His Lys Val Thr Val Leu Asp Leu Ala Ala Ser Gly
35 40 45

Thr Asn Met Lys Lys Ile Glu Asp Val Asp Thr Phe Ser Glu Tyr Ser
50 55 60

Ala Pro Leu Leu Gln Leu Met Ala Thr Ile Pro Ser Asn Glu Lys Leu
65 70 75 80

Val Leu Val Gly His Ser Leu Gly Gly Leu Asn Ile Ala Leu Ala Met

85 90 95

Glu Lys Phe Pro Glu Lys Val Ala Val Gly Val Phe Leu Thr Ala Phe
100 105 110

Ala Pro Asp Thr Glu His His Pro Ser Tyr Val Leu Glu Lys Tyr Asn
115 120 125

Glu Arg Thr Pro Leu Ala Ala Trp Leu Asp Thr Glu Phe Ala Pro Ser
130 135 140

Gly Asn Lys Thr Ser Met Phe Phe Gly Pro Asn Phe Leu Ser Asp Lys
145 150 155 160

Leu Tyr Gln Leu Ser Pro Ile Glu Asp Leu Glu Leu Ala Lys Thr Leu
165 170 175

Ala Arg Pro Ser Ser Leu Phe Met Glu Asp Leu Thr Lys Gln Lys Asn
180 185 190

Phe Ser Lys Glu Gly Tyr Gly Ser Val Pro Arg Ala Phe Ile Val Cys
195 200 205

Thr Glu Asp Leu Gly Ile Pro Leu Glu Tyr Gln Leu Leu Met Ile Gln
210 215 220

Asn Val Gly Phe Asn Asp Val Val Glu Val Lys Asp Ala Asp His Met
225 230 235 240

Val Met Leu Cys Lys Pro Gln Glu Leu Phe Asp Ser Leu Gln Gln Ile
 245 250 255

Ala Thr Lys Tyr Ala
 260

<210> 24
<211> 260
<212> PRT
<213> Nicotiana tabacum

<400> 24

Met Lys Glu Gly Lys His Phe Val Leu Val His Gly Ala Cys His Gly
1 5 10 15

Gly Trp Ser Trp Tyr Lys Leu Lys Pro Leu Leu Glu Ala Ala Gly His
 20 25 30

Lys Val Thr Ala Leu Asp Leu Ala Ala Ser Gly Thr Asp Leu Arg Lys
 35 40 45

Ile Glu Glu Leu Arg Thr Leu Tyr Asp Tyr Thr Leu Pro Leu Met Glu
 50 55 60

Leu Met Glu Ser Leu Ser Ala Asp Glu Lys Val Ile Leu Val Gly His
65 70 75 80

Ser Leu Gly Gly Met Asn Leu Gly Leu Ala Met Glu Lys Tyr Pro Gln
 85 90 95

Lys Ile Tyr Ala Ala Val Phe Leu Ala Ala Phe Met Pro Asp Ser Val
 100 105 110

His Asn Ser Ser Phe Val Leu Glu Gln Tyr Asn Glu Arg Thr Pro Ala
 115 120 125

Glu Asn Trp Leu Asp Thr Gln Phe Leu Pro Tyr Gly Ser Pro Glu Glu
130 135 140

Pro Leu Thr Ser Met Phe Phe Gly Pro Lys Phe Leu Ala His Lys Leu
145 150 155 160

Tyr Gln Leu Cys Ser Pro Glu Asp Leu Ala Leu Ala Ser Ser Leu Val
 165 170 175

Arg Pro Ser Ser Leu Phe Met Glu Asp Leu Ser Lys Ala Lys Tyr Phe
 180 185 190

Thr Asp Glu Arg Phe Gly Ser Val Lys Arg Val Tyr Ile Val Cys Thr
195 200 205

Glu Asp Lys Gly Ile Pro Glu Glu Phe Gln Arg Trp Gln Ile Asp Asn
210 215 220

Ile Gly Val Thr Glu Ala Ile Glu Ile Lys Gly Ala Asp His Met Ala
225 230 235 240

Met Leu Cys Glu Pro Gln Lys Leu Cys Ala Ser Leu Leu Glu Ile Ala
245 250 255

His Lys Tyr Asn
260

<210> 25
<211> 256
<212> PRT
<213> Arabidopsis thaliana

<400> 25

Met Lys His Tyr Val Leu Val His Gly Gly Cys His Gly Ala Trp Cys
1 5 10 15

Trp Tyr Lys Val Lys Pro Met Leu Glu His Ser Gly His Arg Val Thr
20 25 30

Val Phe Asp Leu Thr Ala His Gly Val Asn Met Ser Arg Val Glu Asp
35 40 45

Ile Gln Thr Leu Glu Asp Phe Ala Lys Pro Leu Leu Glu Val Leu Glu
50 55 60

Ser Phe Gly Ser Asp Asp Lys Val Val Leu Val Ala His Ser Leu Gly
65 70 75 80

Gly Ile Pro Ala Ala Leu Ala Ala Asp Met Phe Pro Ser Lys Ile Ser
85 90 95

Val Ala Val Phe Val Thr Ser Phe Met Pro Asp Thr Thr Asn Pro Pro
100 105 110

Ser Tyr Val Phe Glu Lys Phe Leu Gly Ser Ile Thr Glu Glu Glu Arg
115 120 125

Met Asp Phe Glu Leu Gly Ser Tyr Gly Thr Asp Asp His Pro Leu Lys
130 135 140

Thr Ala Phe Leu Gly Pro Asn Tyr Leu Lys Asn Met Tyr Leu Leu Ser
145 150 155 160

Pro Ile Glu Asp Tyr Glu Leu Ala Lys Met Leu Met Arg Val Thr Pro
165 170 175

Ala Ile Thr Ser Asn Leu Thr Gly Thr Lys Ser Leu Thr Ala Gln Gly
180 185 190

Tyr Gly Ser Ile Ser Arg Val Tyr Ile Val Cys Gly Glu Asp Lys Gly
195 200 205

Ile Arg Val Asp Phe Gln Arg Trp Met Ile Glu Asn Ser Pro Val Lys
210 215 220

Glu Val Met Glu Ile Lys Asp Ala Asp His Met Pro Met Phe Ser Lys
225 230 235 240

Pro His Glu Leu Cys Asp Arg Leu Leu Lys Ile Ala Asp Lys Tyr Pro
245 250 255

- <210> 26
- <211> 20
- <212> ADN
- <213> Secuencia Artificial
- <220>
- <223> Cebador de oligonucleótido sintético

<400> 26
aattgaccaa gagcttcgc 20

<210> 27
<211> 27
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 27
ctacagaagc atcattttca acatctt 27

<210> 28
<211> 21
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 28
caattgacca agagcttcca a 21

<210> 29
<211> 27
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 29
ctatagaagc atcgtttca acatctc 27

<210> 30
<211> 17
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 30
ggccatttcg gagctgg 17

<210> 31
<211> 24
<212> ADN
<213> Secuencia Artificial

<220>
<223> Cebador de oligonucleótido sintético

<400> 31
aggagaaggt ggttctgtga gaga 24

<210> 32
<211> 20
<212> ADN
<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 32

catggccatt tcggagctta 20

<210> 33

<211> 22

<212> ADN

<213> Secuencia Artificial

<220>

<223> Cebador de oligonucleótido sintético

<400> 33

cagaagatgg ttctgcgaat gg 22

RESUMEN

Se divulgan genes resistentes a plagas aislados de la soja. En varias realizaciones, el gen resistente a las plagas aislado de la soja incluye una secuencia de ácidos nucleicos al menos 80% idéntica a la SEQ ID NO:1 o SEQ ID NO: 2 o su variante degenerada o su fragmento. También se divulgan vectores de expresión y construcciones que incluyen dichos ácidos nucleicos. Se divulgan métodos para producir una planta transgénica que presenta una resistencia mejorada a las plagas. También se divulgan plantas transgénicas, células o tejidos vegetales (tales como plantas dicotiledóneas o monocotiledóneas, células o tejidos vegetales) transformados con s construcciones divulgadas.