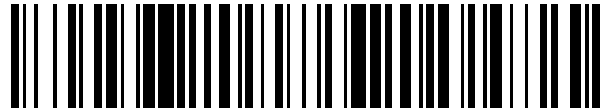


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 537 342**

51 Int. Cl.:

C12N 15/29 (2006.01)

C12N 15/52 (2006.01)

C07K 14/415 (2006.01)

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.06.2009 E 09758583 (0)**

97 Fecha y número de publicación de la concesión europea: **04.03.2015 EP 2294200**

54 Título: **Nuevos genes implicados en biosíntesis**

30 Prioridad:

06.06.2008 US 59691 P
06.06.2008 NZ 56892808

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
05.06.2015

73 Titular/es:

GRASSLANZ TECHNOLOGY LIMITED (100.0%)
East Street Ruakura Campus
Hamilton 3214, NZ

72 Inventor/es:

HANCOCK, KERRY, RUTH y
GREIG, MARGARET

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 537 342 T3

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

DESCRIPCIÓN

Nuevos genes implicados en biosíntesis

CAMPO TÉCNICO

5 La invención se refiere a un nuevo gen o genes implicados en la biosíntesis. En particular, la presente invención se refiere a gen o genes que codifican un factor regulador que controla la expresión de genes claves implicados en la producción de taninos condensados en plantas.

TÉCNICA ANTECEDENTE

La ruta molecular de fenilpropanoides

10 La ruta de fenilpropanoides (mostrada en la Figura 1) produce un conjunto de metabolitos secundarios que incluyen flavonas, antocianinas, flavonoides, taninos condensados e isoflavonoides (Dixon et al., 1996; 2005). En particular, la ruta biosintética de taninos condensados (CT) comparte sus etapas tempranas con la ruta de antocianinas antes de divergir hacia la biosíntesis de proantocianidinas.

15 Las antocianidinas son precursores de flavan-3-oles (por ejemplo, (-)-epicatequina), que son bloques de construcción importantes para los CTs. Estos cis-flavan-3-oles se forman a partir de antocianidinas mediante antocianidin reductasa (ANR), que se ha clonado para muchas especies, incluyendo *A. thaliana* y *M. truncatula* (Xie et al., 2003; 2004). En *A. thaliana*, (-)-epicatequina es el monómero de CT exclusivo (Abrahams et al., 2002), pero en muchas otras especies, incluyendo legumbres, tanto (+)- como (-)-flavan-3-oles se polimerizan a CTs. La biosíntesis de estos (+)-flavan-3-oles (catequinas) alternativos está catalizada por leucoantocianidin reductasa (LAR). Esta enzima se ha clonado y caracterizado de legumbres, incluyendo el árbol de leguminosa rico en CTs *Desmodium uncinatum* (Tanner et al., 2003), así como de otras especies tales como uvas y manzanas (Pfeiffer et al., 2006). La enzima cataliza la reducción de leucopelargonidina, leucocianidina, y leucodelfinidina a afcelequina, catequina, y galocatequina, respectivamente. No se han encontrado homólogos de LAR en *A. thaliana*, consistente con la presencia exclusiva de bloques de construcción de CTs derivados de (-)-epicatequina en esta planta.

25 Mientras que la información sobre la regulación de TFs de esta ruta en semillas de *Arabidopsis* está bien definida, todavía se han de identificar TFs que controlan la biosíntesis de CTs de las hojas dentro de la tribu de *Trifolieae*. Una familia importante de proteínas TF, la familia MYB, controla un intervalo diverso de funciones, incluyendo la regulación del metabolismo secundario, tal como las rutas de antocianinas y de CTs en plantas. La expresión de la MYB TF *AtTT2* apaga o enciende coordinadamente los genes estructurales tardíos en *Arabidopsis thaliana*, controlando finalmente la expresión de la ruta de CTs.

30 Se ha obtenido un conjunto de mutantes testa transparente (TT) de *Arabidopsis thaliana* (Winkel-Shirley, 2002; Debeaujon et al., 2001) y mutantes de semillas deficientes en taninos (TDS) (Abrahams et al. 2002; 2003) – todos ellos deficientes en la acumulación de CTs en el tegumento. Los estudios genéticos moleculares de estos mutantes ha permitido la identificación de un número de genes estructurales y factores de transcripción (TFs) que regulan la expresión y especificidad tisular de la síntesis de tanto antocianinas como de CTs en *A. thaliana* (Walker et al., 1999; Nesi et al., 2000; 2002).

40 Aunque la mayoría de los genes estructurales en la ruta de CTs se han identificado en un intervalo de legumbres, hasta ahora han fallado los intentos para manipular la biosíntesis de CTs en hojas manipulando mediante ingeniería la expresión de estos genes individuales. La razón principal para esto es que ninguna (o unas pocas) enzima o enzimas son limitantes de la velocidad, sino que la actividad de virtualmente todas las enzimas en una ruta se ha de incrementar para lograr un flujo incrementado global en productos finales específicos tales como taninos condensados.

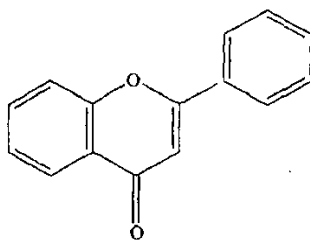
45 Los factores de transcripción (TFs) son proteínas reguladoras que actúan como represoras o activadoras de rutas metabólicas. Por lo tanto, los TFs se pueden usar como una herramienta poderosa para la manipulación de todas las rutas metabólicas en plantas. Muchos TFs de MYB son reguladores importantes de la ruta de fenilpropanoides, incluyendo tanto la síntesis de antocianinas como de taninos condensados (Debeaujon et al., 2003; Davies y Schwinn, 2003). Por ejemplo, el gen TT2 (*AtTT2*) de *A. thaliana* codifica un factor TF R2R3-MYB que es expresado solamente en el tegumento durante etapas tempranas de la embriogénesis, cuando se produce la biosíntesis de taninos condensados (Nesi et al., 2001). Se ha mostrado que TT2 regula la expresión de los genes estructurales biosintéticos tardíos de flavonoides TT3 (DFR), TT18, TT12 (proteína MATE) y ANR durante la biosíntesis y almacenamiento de CTs. *AtTT2* determina parcialmente la expresión espacial y temporal restringida de genes, en combinación con otros dos TFs; a saber, TT8 (proteína bHLH) y TTG1 (proteína de repetición WD-40; Baudry et al., 2004).

55 También se han dado a conocer recientemente otros TFs de MYB en uva *Vitis vinifera*; (VvMYBPA1), loto y *Brassica napus* (BnTT2), que están implicados en la regulación de la biosíntesis de CTs (Wei et al., 2007; Bogs et al., 2007; Yoshida et al., 2008).

También se ha mostrado que el gen AtTT2 comparte un grado de similitud con el OsMYB3 del arroz (*Oryza sativa*), el ZmC1 del maíz (*Zea mays*), AmMYBROSEA de *Antirrhinum majus*, y PhMYBAN2 de *Petunia hybrida*, genes los cuales se ha mostrado que regulan la biosíntesis de antocianinas (Stracke et al., 2001; Mehrtens et al., 2005).

Taninos condensados

- 5 Los taninos condensados (CTs), también denominados proantocianidinas (Pas), son polímeros coloreados, uno de varios metabolitos vegetales secundarios. Los CTs son polímeros de 2 a 50 (o más) unidades de flavonoides (véase el compuesto (I) más abajo) que están unidas mediante enlaces carbono-carbono que no son susceptibles de ser escindidos por hidrólisis. La estructura base de flavonoides es:



10

COMPUESTO (I)

Los taninos condensados están situados en un intervalo de partes vegetales, por ejemplo las hojas, tallos, flores, raíces, productos de la madera, corteza, capullos. Los CTs se encuentran generalmente en vacuolas o en la epidermis superficial de la planta.

Taninos condensados en plantas forrajeras

- 15 Las plantas forrajeras, tales como legumbres forrajeras, son beneficiosas en sistemas de ganado a base de pastura, debido a que mejoran tanto la ingesta como la calidad de la dieta del animal. También, su valor con respecto a la economía de nitrógeno (N) de pasturas con respecto a la producción de los rumiantes es considerable (Caradus et al., 2000). Sin embargo, aunque produce una fuente de pienso eficaz desde el punto del coste para rumiantes de pastoreo, la pastura es a menudo subóptima cuando se trata de satisfacer los requisitos nutricionales tanto de la microflora de la panza como del propio animal. De este modo, el potencial genético de rumiantes de pastoreo para la producción de carne, lana o leche raramente se logra en una dieta forrajera.

- 20 Las pasturas de Nueva Zelanda contienen hasta 20% de trébol blanco; mientras que el incremento de los niveles de trébol blanco en pasturas ayuda a abordar este déficit, también exagera la incidencia de meteorismo. El trébol blanco (*Trifolium repens*), el trébol rojo (*Trifolium pratense*) y la alfalfa (*Medicago sativa*) son causas bien documentadas de meteorismo, debido a la deficiencia de compuestos polifenólicos de la planta, tales como CT, en estas especies. Por lo tanto, el desarrollo de variedades de cultivo forrajeras que produzcan mayores niveles de taninos en tejido vegetal sería un desarrollo importante en la industria de agricultura para reducir la incidencia de meteorismo (Burggraaf et al., 2006).

- 30 En particular, los taninos condensados, si están presentes en cantidades suficientes, no solo ayudan a eliminar el meteorismo, sino también influyen fuertemente sobre la calidad de la planta, apetecibilidad y el valor nutritivo de legumbres forrajeras, y por lo tanto pueden ayudar a mejorar el rendimiento del animal. Los beneficios de salud animal y productividad dados a conocer a partir de mayores niveles de CTs incluyen mayores tasas de ovulación en ovejas, mayor ganancia de peso vivo, crecimiento de lana y producción de leche, cambio de la composición de la leche y efectos antihelmínticos mejorados sobre parásitos gastrointestinales (Rumbaugh, 1985; Marten et al., 1987; Niezen et al., 1993; 1995; Tanner et al., 1994; McKenna, 1994; Douglas et al., 1995; Waghorn et al., 1998; Aerts et al., 1999; McMahan et al., 2000; Molan et al., 2001; Sykes y Coop, 2001).

- 40 Un mayor nivel de tanino condensado también representa una solución viable para reducir los gases de efecto invernadero (metano, óxido nitroso) liberados al medio ambiente por rumiantes de pastoreo (Kingston-Smith y Thomas, 2003). El ganado rumiante produce al menos 88% de las emisiones totales de metano de Nueva Zelanda, y es un contribuyente importante de las emisiones de gases de efecto invernadero (Clark, 2001). La principal fuente de metano ganadero es la fermentación entérica en el tubo digestivo de los rumiantes. La producción de metano, que representa una pérdida de energía para los rumiantes de alrededor de 3 a 9% de la ingesta total de energía (Blaxter y Clapperton, 1965), se puede reducir tanto como 5% mejorando la calidad del forraje. Se ha mostrado que el forraje rico en CT reduce la emisión de metano de animales de pastoreo (Woodward, et al 2001; Puchala, et al., 2005). Por lo tanto, el incremento del contenido de CTs de plantas de pastura puede contribuir directamente a niveles reducidos de emisión de metano procedente de ganado.

- 45 Por lo tanto, los beneficios medioambientales y agronómicos que se podrían derivar de desencadenar la acumulación de una cantidad incluso moderada de taninos condensados en plantas forrajeras, incluyendo trébol

blanco, son de importancia considerable en la protección y nutrición de rumiantes (Damiani et al., 1999).

Legumbres

5 Los inventores entienden que la regulación de la ruta específica foliar de CTs en legumbres de *Trifolium*, que implica la interacción de factores de transcripción reguladores (TFs) con la ruta, sigue siendo desconocida. Se ha explorado la modificación o manipulación de esta ruta para influir en la cantidad de CT, pero, puesto que el proceso no es directo, ha habido poco éxito firme en la comprensión de esta ruta.

10 El género del trébol, *Trifolium*, por ejemplo, es uno de los géneros más grandes en la familia Leguminosae (D Fabaceae), con aprox. 255 especies (Ellison et al., 2006). Se sabe que solamente dos especies de *Trifolium*, *T. affine* (también conocida como *Trifolium preslianum* Boiss. Is) y *T. arvense* (también conocido como clavo de pie de liebre), acumulan niveles elevados de CTs foliares (Fay y Dale, 1993). Aunque existen niveles significativos de CTs en inflorescencias de trébol blanco (Jones et al., 1976), solo se pueden detectar cantidades en trazas en tricomas de las hojas (Woodfield et al., 1998). Varios enfoques, incluyendo el cribado de conjuntos génicos y mutagénesis al azar, han fracasado a la hora de proporcionar plantas de trébol blanco o rojo con mayores niveles de CTs foliares (Woodfield et al., 1998).

15 Manipulación genética de taninos condensados

20 Los inventores en relación con el documento US2006/012508 crearon una planta de alfalfa transgénica usando el gen regulador de MYB TT2, y fueron capaces de producir sorprendentemente CTs constitutivamente en todos los tejidos de la raíz. Sin embargo, de forma importante, los inventores fueron incapaces de lograr la acumulación de CTs en las hojas de esta legumbre forrajera. Se ha dado a conocer previamente que no existen circunstancias conocidas que puedan inducir proantocianidinas (CTs) en forraje de alfalfa (Ray et al., 2003). Los autores de este artículo evaluaron entre otros aspectos si el gen regulador similar a myc LC (TF) del maíz o el gen regulador de myb C1 (TF) del maíz podrían estimular la ruta de flavonoides en forraje de alfalfa y tegumento. Los autores de este artículo encontraron que solo el gen LC, y no C1, podría estimular la biosíntesis de antocianinas y proantocianidinas en forraje de alfalfa, pero la estimulación solo ocurrió en presencia de un factor de alfalfa desconocido sensible al estrés.

25 Los estudios que evalúan la producción de taninos condensados en plantas de *Lotus* usando un gen regulador bHLH (TF) de maíz encontraron que la transformación de este TF en plantas de *Lotus* dio como resultado CTs solamente en un incremento muy pequeño (1%) en niveles de taninos condensados en hojas (Robbins et al., 2003).

30 Los intentos previos para alterar y potenciar compuestos agrícolaemente importantes en trébol blanco implicaron alterar la biosíntesis de antocianinas derivada de la ruta de fenilpropanoides. A pesar de los intentos para activar esta ruta usando varios TFs de myc y MYB heterólogos, solo se ha dado a conocer un éxito, usando el TF de myc de maíz TF B-Peru (de Majnik et al., 2000). Otros TFs investigados dieron como resultado regenerantes pobres o ningún regenerante, implicando un efecto pernicioso debido a su sobreexpresión.

35 Más recientemente, se han dado a conocer homólogos de TT2 derivados de la legumbre rica en CTs, *Lotus japonicus* (Yoshida et al., 2008). El bombardeo de estos genes en células de hojas de *A. thaliana* ha mostrado expresión transitoria que da como resultado expresión detectable de ANR y acumulación limitada de CTs según se detecta mediante DMACA. Sin embargo, estos genes no se han transformado ni analizado en ninguna especie de legumbre.

40 La expresión del gen Lc del maíz dio como resultado la acumulación de compuestos similares a PAs en alfalfa solamente si las plantas estaban bajo estrés abiótico (Ray et al., 2003). La coexpresión de tres factores de transcripción, TT2, PAP1 y Lc en *Arabidopsis* fue necesaria para superar la expresión de PAs específica del tipo celular, pero esta acumulación constitutiva de PAs estuvo acompañada de la muerte de las plantas (Sharma y Dixon, 2005).

45 La introducción de PAs en plantas por la expresión combinada de un factor de transcripción de la familia MYB y antocianidina reductasa para la conversión de antocianidina en (epi)-flavan-3-ol se ha intentado por Xie et al. (2006).

50 Este intento para incrementar los niveles de proantocianidinas (PAs) en las hojas de tabaco coexpresando PAP1 (un TF de MYB) y ANR dio a conocer que hay niveles de PAs en tabaco que, si se traduce a alfalfa, pueden proporcionar potencialmente protección frente a meteorismo (Xie et al., 2006). Las hojas que contienen antocianinas de *M. truncatula* transgénica que expresa constitutivamente *MtANR* contenían hasta tres veces más PAs que aquellas de plantas de tipo salvaje en la misma etapa de desarrollo, y estos compuestos fueron de un subconjunto específico de oligómeros de PAs. Adicionalmente, estos niveles de PA producidos en *M. truncatula* estuvieron muy lejos de los necesarios para un beneficio agronómico mejorado. Los autores afirman que siguió siendo incierto qué genes biosintéticos y no biosintéticos adicionales serán necesarios para la manipulación genética de PAs en cualquier tejido vegetal específico que acumule de forma natural los compuestos.

55 También se encontraron dificultades similares a la hora de expresar CTs o PAs en hojas cuando los genes TT2 y/o BAN se transformaron en alfalfa – refiérase a los documentos US 2004/0093632 y US 2006/0123508.

Taninos condensados útiles en productos naturales para la salud

También es ampliamente conocido el uso de cualquier flavonoide, incluyendo proantocianidinas, para formar suplementos alimentarios, composiciones o medicamentos. Por ejemplo:

- 5 • La solicitud de patente US nº 2003/0180406 describe un método que usa composiciones polifenólicas derivadas específicamente de cacao para mejorar la función cognitiva.
- La publicación de patente WO 2005/044291 describe el uso de semilla de uva (género *Vitis*) para prevenir enfermedades cerebrales degenerativas que incluyen apoplejía, conmoción cerebral, enfermedad de Huntington, CJD, Alzheimer, Parkinson, y demencia senil.
- 10 • La publicación de patente WO 2005/067915 describe una combinación sinérgica de flavonoides e hidroxiestilbenos (sintéticos o procedentes de té verde) combinados con flavonas, flavonoides, proantocianidinas y antocianidinas (sintéticos o procedente de extracto de corteza) para reducir la degeneración neuronal asociada con estados mórbidos tales como demencia, Alzheimer, enfermedad cerebrovascular, alteración cognitiva relacionada con la edad, y depresión.
- El documento US 5.719.178 describe el uso de extracto de proantocianidinas para tratar ADHD.
- 15 • La publicación PCT número 06/126895 describe una composición que contiene extracto de corteza del género *Pinus* para mejorar o prevenir un deterioro en capacidades cognitivas humanas, o mejorar o prevenir síntomas de trastornos neurológicos en un ser humano.

Ninguno de los anteriores considera el uso de legumbres como fuente de materia prima de CT.

20 Por lo tanto, sería útil si se pudiese proporcionar moléculas de ácido nucleico y polipéptidos útiles en el estudio de las rutas metabólicas implicadas en la biosíntesis de flavonoides y/o taninos condensados.

También sería útil si se pudiese proporcionar moléculas de ácido nucleico y polipéptidos que sean capaces de alterar los niveles de flavonoides y/o taninos condensados en plantas o sus partes.

En particular, sería útil si se pudiese proporcionar moléculas de ácido nucleico que se pueden usar para producir flavonoides y/o taninos condensados en plantas o sus partes *de novo*.

25 Por lo tanto, es un objeto de la invención proporcionar un método para incrementar los niveles de CTs en las hojas de especies de legumbres forrajeras. La identificación del gen también proporciona un método para prevenir la acumulación de CTs en especies de legumbres que produce niveles elevados perjudiciales de CT en hojas o semillas.

30 También sería útil si se pudiese proporcionar moléculas de ácido nucleico que se puedan usar solas o junto con otras moléculas de ácido nucleico para producir plantas, particularmente forrajes y legumbres, con niveles potenciados de flavonoides y/o taninos condensados.

Es un objeto de la presente invención abordar los problemas anteriores o al menos proporcionar al público una elección útil.

SUMARIO DE LA INVENCION

35 La presente invención se refiere a la identificación y usos de un nuevo gen de MYB y polipéptido asociado que se ha denominado por los inventores "MYB14", que se ha aislado por los solicitantes y que se ha mostrado que está implicado en la producción de compuestos flavonoides, incluyendo taninos condensados.

A lo largo de esta memoria descriptiva, las moléculas de ácido nucleico y polipéptidos de la presente invención se pueden denominar por el descriptor MYB14.

40 La presente invención contempla el uso de MYB14 independientemente o junto con otras moléculas de ácido nucleico para manipular la ruta biosintética de taninos condensados en plantas.

Polinucleótidos que codifican polipéptidos

En un aspecto, la invención proporciona una molécula de ácido nucleico aislada que codifica:

- 45 a) un polipéptido MYB14 que comprende una secuencia con al menos 70% de identidad con SEC ID NO: 14, en el que el % de identidad se calcula a lo largo de toda la longitud de SEC ID NO: 14; y
- b) un fragmento funcional de SEC ID NO: 14;

en la que el polipéptido MYB14 en a) y el fragmento funcional en b) son capaces de incrementar la producción de taninos condensados en plantas.

En una realización, el polipéptido MYB14 comprende la secuencia de SEC ID NO: 14.

En una realización, el polipéptido MYB14 o fragmento funcional comprende la secuencia de SEC ID NO: 17.

En una realización adicional, el polipéptido MYB14 regula la producción de taninos condensados en una planta.

El polipéptido MYB14 regula al menos un gen en la ruta biosintética de flavonoides en una planta.

- 5 El polipéptido MYB14 regula al menos un gen en la ruta biosintética de taninos condensados en una planta.

En una realización adicional, el fragmento funcional tiene sustancialmente la misma actividad que el polipéptido MYB14.

En un aspecto adicional, la invención proporciona una molécula de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos sustancialmente como se muestra en SEC ID NO: 17.

- 10 En un aspecto adicional, la invención proporciona una molécula de ácido nucleico que codifica un polipéptido que tiene una secuencia de aminoácidos sustancialmente como se muestra en SEC ID NO: 17.

En un aspecto adicional, la invención proporciona una molécula de ácido nucleico que codifica un polipéptido que comprende una secuencia de aminoácidos sustancialmente como se muestra en SEC ID NO: 14.

- 15 En un aspecto adicional, la invención proporciona una molécula de ácido nucleico que codifica un polipéptido que tiene una secuencia de aminoácidos sustancialmente como se muestra en SEC ID NO: 14.

En un aspecto adicional, la invención proporciona una molécula de ácido nucleico aislada que codifica un polipéptido que comprende un motivo de secuencia de aminoácidos de 3' como se expone en SEC ID NO: 17.

Polinucleótidos

- 20 En un aspecto adicional, la invención proporciona una molécula de ácido nucleico aislada que tiene una secuencia nucleotídica seleccionada del grupo que consiste en:

a) SEC ID NO: 1, 2 ó 55;

b) un complemento de la secuencia o secuencias en a);

c) un fragmento funcional de la secuencia o secuencias en a) o b);

- 25 d) una secuencia con al menos 70% de identidad con una secuencia en a) o b) a lo largo de toda la longitud de la secuencia;

en la que la secuencia de d) y el fragmento funcional de c) codifican un polipéptido que es capaz de incrementar la producción de taninos condensados en plantas.

En una realización adicional, el fragmento comprende la secuencia codificante de la secuencia especificada.

En una realización adicional, la molécula de ácido nucleico aislada comprende la secuencia de SEC ID NO: 2.

- 30 En una realización adicional, la molécula de ácido nucleico aislada comprende la secuencia de SEC ID NO: 1.

En una realización adicional, la molécula de ácido nucleico aislada comprende la secuencia de SEC ID NO: 55.

Sondas

También se describe una sonda capaz de unirse a un ácido nucleico de la invención.

- 35 También se describe una sonda capaz de unirse a un dominio de 3' de la molécula de ácido nucleico MYB14 sustancialmente como se describe anteriormente.

En un caso, la sonda es capaz de unirse a una molécula de ácido nucleico que codifica la secuencia de aminoácidos de SEC ID NO: 17, o a un complemento de la molécula de ácido nucleico.

En un caso, la sonda es capaz de unirse a la molécula de ácido nucleico, o a su complemento, en condiciones de hibridación restrictivas.

- 40 También se describe una sonda para una secuencia de 3' que codifica el motivo como se expone en SEC ID NO: 17.

Cebadores

También se describe un cebador capaz de unirse a un ácido nucleico de la invención.

También se describe un cebador capaz de unirse a un dominio de 3' de la molécula de ácido nucleico MYB14 sustancialmente como se describe anteriormente.

En un caso, la sonda es capaz de unirse a una molécula de ácido nucleico que codifica la secuencia de aminoácidos de SEC ID NO: 15, o a un complemento de la molécula de ácido nucleico.

- 5 En un caso, la sonda es capaz de unirse a la molécula de ácido nucleico, o a su complemento, en condiciones de PCR.

También se describe un cebador para un ácido nucleico que codifica una secuencia de 3' que codifica el motivo como se expone en SEC ID NO: 17.

Polipéptidos

- 10 En un aspecto, la invención proporciona un polipéptido MYB14 aislado, o un fragmento funcional del mismo, como se define en las reivindicaciones.

En un caso, el polipéptido MYB14 comprende la secuencia de SEC ID NO: 15 y SEC ID NO: 17, pero carece de la secuencia de SEC ID NO: 16.

- 15 En un aspecto adicional, la invención proporciona un polipéptido MYB14 aislado que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:

a) una secuencia con al menos 70% de identidad con SEC ID NO: 14, en el que el % de identidad se calcula a lo largo de toda la longitud de SEC ID NO: 14; y

b) un fragmento funcional de la secuencia dada en a), en el que la secuencia en a) y el fragmento funcional en b) son capaces de incrementar la producción de taninos condensados en plantas.

- 20 En una realización adicional, el polipéptido MYB14 comprende la secuencia de SEC ID NO: 14.

El polipéptido MYB14 regula la producción de taninos condensados en una planta.

En un caso adicional, el polipéptido MYB14 regula al menos un gen en la ruta biosintética de flavonoides en una planta.

En una realización adicional, el polipéptido MYB14 regula la ruta biosintética de taninos condensados en una planta.

- 25 En una realización adicional, el polipéptido MYB14 regula al menos un gen en la ruta biosintética de taninos condensados en una planta.

En una realización adicional, el fragmento funcional tiene sustancialmente la misma actividad que el polipéptido MYB14.

- 30 Según otro aspecto de la invención, se proporciona un polipéptido aislado que tiene una secuencia de aminoácidos seleccionada del grupo que consiste en:

a) SEC ID NO: 14;

b) un fragmento funcional de la secuencia dada en a), en el que la secuencia en a) y el fragmento en b) son capaces de incrementar la producción de taninos condensados en una planta.

- 35 También se describe un polipéptido aislado que comprende un motivo de secuencia de aminoácidos de 3' como se expone en SEC ID NO: 17.

También se describe un polipéptido aislado que tiene un motivo de secuencia de aminoácido de 3' como se expone en SEC ID NO: 17.

- 40 También se describe un polipéptido MYB14 aislado o un fragmento funcional del mismo en el que dicho polipéptido MYB14 incluye un motivo de secuencia de aminoácidos de subgrupo 5 como se muestra en SEC ID NO: 15, así como un motivo de 3' de secuencia de aminoácidos como se muestra en SEC ID NO: 17, pero que carece de un motivo de secuencia de aminoácidos de subgrupo 6 como se muestra en SEC ID NO: 16.

También se describe un polipéptido aislado codificado por una molécula de ácido nucleico que tiene una secuencia nucleotídica seleccionada de aquellas expuestas en una cualquiera de SEC ID NO: 1 a 13 y 55 a 64.

- 45 También se describe un polipéptido aislado codificado por una molécula de ácido nucleico que tiene una secuencia nucleotídica como se expone en SEC ID NO: 1, 2 ó 55.

También se describe una molécula de ácido nucleico que comprende una secuencia que codifica un polipéptido de

la invención.

Constructos

Según un aspecto adicional de la presente invención, se proporciona un constructo que incluye una secuencia nucleotídica sustancialmente como se describe anteriormente.

5 Según un aspecto adicional de la presente invención, se proporciona un constructo que incluye:

- al menos un promotor; y
- una molécula de ácido nucleico sustancialmente como se describe anteriormente;

en el que el promotor está enlazado operativamente a la molécula de ácido nucleico para controlar la expresión de la molécula de ácido nucleico.

10 Preferiblemente, el constructo puede incluir una o más moléculas de ácido nucleico adicionales de interés y/o una o más secuencias reguladoras adicionales, tales como, entre otras, secuencias terminadoras.

Lo más preferible, la molécula de ácido nucleico en el constructo puede tener una secuencia nucleotídica seleccionada de SEC ID NO: 1, 2 ó 55.

Células hospedantes

15 Según un aspecto adicional de la presente invención, se proporciona una célula hospedante que se ha alterado a partir del tipo salvaje al transformarla para incluir una molécula de ácido nucleico sustancialmente como se describe anteriormente.

En una realización, el ácido nucleico es parte de un constructo genético de la invención.

En una realización, la célula hospedante no forma parte de un ser humano.

20 En una realización adicional, la célula hospedante es una célula vegetal.

Células vegetales y plantas

Según un aspecto adicional de la presente invención, se proporciona una planta o célula vegetal transformada con un constructo sustancialmente como se describe anteriormente.

25 Según un aspecto adicional de la presente invención, se proporciona una planta transformada con un constructo sustancialmente como se describe anteriormente.

Según un aspecto adicional de la presente invención, se proporciona una planta o parte de la misma que se ha alterado a partir del tipo salvaje para incluir una molécula de ácido nucleico sustancialmente como se describe anteriormente.

30 Según un aspecto adicional de la presente invención, se proporciona una célula vegetal, planta o parte de la misma que se ha manipulado vía expresión alterada de un gen MYB14 para que tenga niveles de taninos condensados mayores que una planta de tipo salvaje correspondiente o parte de la misma.

Según un aspecto adicional de la presente invención, se proporciona una célula vegetal, célula vegetal la cual se ha manipulado vía expresión alterada de un gen MYB14 para que tenga niveles de taninos condensados mayores que una célula vegetal de tipo salvaje correspondiente.

35 Según un aspecto adicional de la presente invención, se proporciona una hoja de una planta que, vía expresión alterada de un gen MYB14, tiene niveles de taninos condensados mayores que una planta de tipo salvaje correspondiente o parte de la misma.

40 Según un aspecto adicional de la presente invención, se proporciona la progenie de una célula vegetal o una planta sustancialmente como se describe anteriormente, la cual, vía expresión alterada de un gen MYB14, tiene niveles de taninos condensados mayores que una célula vegetal de tipo salvaje o planta correspondiente.

Según un aspecto adicional de la presente invención, se proporciona la semilla de una planta transgénica sustancialmente como se describe anteriormente.

Composiciones

45 También se describe una composición que incluye un ingrediente que es, o se obtiene de, una planta y/o parte de la misma, en el que dicha planta o parte de la misma se ha manipulado vía expresión alterada de un gen MYB14 para que tenga niveles mayores o menores de flavonoides y/o taninos condensados en comparación con aquellos de una

planta de tipo salvaje correspondiente o parte de la misma.

Métodos que usan polinucleótidos

Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico sustancialmente como se describe anteriormente para alterar una planta o célula vegetal.

- 5 Según un aspecto adicional de la presente invención, se proporciona un método para producir una planta alterada o célula vegetal usando una molécula de ácido nucleico sustancialmente como se describe anteriormente.

En una realización, la planta o célula vegetal se altera en la producción de al menos un tanino condensado, o un intermedio en la producción de al menos un tanino condensado.

- 10 En una realización adicional, el tanino condensado se selecciona de catequina, epicatequina, epigallocatequina y galocatequina.

En la invención, la alteración es un incremento.

En una realización adicional, la planta o célula vegetal se altera en la expresión de al menos una enzima en una ruta biosintética de taninos condensados.

En la invención, la expresión alterada es expresión incrementada.

- 15 En una realización adicional, la enzima es LAR o ANR.

En una realización adicional, la planta se altera en la expresión tanto de LAR como de ANR.

La planta puede ser cualquier planta, y la célula vegetal puede ser de cualquier planta.

En una realización, la planta es una planta de cultivo forrajera.

En una realización adicional, la planta es una planta leguminosa.

- 20 En una realización, la producción o expresión alterada, descrita anteriormente, es en sustancialmente todos los tejidos de la planta.

En una realización, la producción o expresión alterada, descrita anteriormente, es en el tejido foliar de la planta.

En una realización, la producción o expresión alterada, descrita anteriormente, es en las porciones vegetativas de la planta.

- 25 En una realización, la producción o expresión alterada, descrita anteriormente, es en los tejidos epidérmicos de la planta.

Para los fines de esta memoria descriptiva, el tejido epidérmico se refiere al grupo de monocapa exterior de células, incluyendo la hoja, tallos, y raíces y tejidos jóvenes de una planta vascular.

- 30 En una realización, la producción alterada de taninos condensados descrita anteriormente es en un tejido de la planta que está sustancialmente desprovisto de los taninos condensados.

Por lo tanto, en algunas realizaciones de la invención, la producción de taninos condensados es producción *de novo*.

En una realización, el ácido nucleico codifica una proteína MYB14 como se define aquí.

En un caso, el ácido nucleico codifica una proteína que comprende una secuencia de aminoácidos como se expone en una cualquiera de SEC ID NOs 1-13 y 55 a 64, o fragmento o variante de la misma.

- 35 En un caso, el ácido nucleico comprende una secuencia sustancialmente como se expone en una cualquiera de SEC ID NOs 1-13 y 55 a 64, o fragmento o variante de la misma.

En una realización adicional, el ácido nucleico comprende una secuencia sustancialmente como se expone en SEC ID NOs 1, 2 ó 55, o fragmento o variante de la misma.

- 40 En una realización adicional, el ácido nucleico es parte de un constructo sustancialmente como se describe anteriormente.

En una realización, la planta se altera transformando la planta con el ácido nucleico o constructo.

En una realización adicional, la planta se altera manipulando el genoma de una planta para expresar niveles incrementados o disminuidos del ácido nucleico, o fragmento o variante del mismo, en la planta en comparación con el producido en una planta de tipo salvaje correspondiente o parte de la misma.

También se describe el uso de una molécula de ácido nucleico o polipéptido de la presente invención para identificar otros genes/polipéptidos reguladores de flavonoides y/o taninos condensados relacionados.

5 Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico sustancialmente como se describe anteriormente para alterar una planta o célula vegetal, en el que dicha planta es, o dicha célula vegetal procede de, un cultivo forrajero.

En la invención, la planta tiene producción incrementada de taninos condensados.

Preferiblemente, el cultivo forrajero puede ser una legumbre forrajera.

10 Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico sustancialmente como se describe anteriormente para alterar los niveles de taninos condensados en plantas leguminosas o células de plantas leguminosas.

Preferiblemente, los niveles de taninos condensados se alteran en tejido foliar.

También se describe el uso de información de secuencia de ácido nucleico sustancialmente como se expone en una cualquiera de SEC ID NO: 1-13 y 55 a 64 para alterar la ruta biosintética de flavonoides o de taninos condensados *in planta*.

15 Según un aspecto adicional de la presente invención, se proporciona el uso de información de secuencia de ácido nucleico sustancialmente como se expone en una cualquiera de SEC ID NO: 1, 2 y 55 para alterar la ruta biosintética de taninos condensados *in planta*.

20 Según un aspecto adicional de la presente invención, se proporciona el uso de un constructo sustancialmente como se describe anteriormente para transformar una planta leguminosa o célula vegetal para alterar los niveles de taninos condensados en las porciones vegetativas de la planta leguminosa o célula vegetal.

Según un aspecto adicional de la presente invención, se proporciona un método para alterar la producción de taninos condensados en una planta leguminosa o parte de la misma, que incluye la etapa de manipular el genoma de una planta para expresar niveles crecientes de un gen MYB14 de leguminosa, o fragmento o variante del mismo, en la planta en comparación con el producido en una planta de tipo salvaje correspondiente o parte de la misma.

25 Según un aspecto adicional de la presente invención, se proporciona un método para alterar la producción de taninos condensados en una planta leguminosa o parte de la misma, que incluye la etapa de manipular el genoma de una planta para expresar niveles crecientes o decrecientes de un gen MYB14 de leguminosa, o fragmento o variante del mismo, en la planta en comparación con el producido en una planta de tipo salvaje correspondiente o parte de la misma.

30 Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico para producir taninos condensados *in planta* en una planta leguminosa o parte de la misma *de novo*.

Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico sustancialmente como se describe anteriormente para manipular en una planta leguminosa o parte de la misma la ruta biosintética de taninos condensados *in planta*.

35 Según un aspecto adicional de la presente invención, se proporciona el uso de un constructo sustancialmente como se describe anteriormente para manipular la ruta biosintética de taninos condensados *in planta*.

Según un aspecto adicional de la presente invención, se proporciona el uso de un gen MYB14 que tiene una secuencia de ácido nucleico sustancialmente que corresponde a una molécula de ácido nucleico de la presente invención para manipular la ruta biosintética *in planta*.

40 Según un aspecto adicional de la presente invención, se proporciona el uso de una molécula de ácido nucleico sustancialmente como se describe anteriormente para producir un tanino condensado, enzima, intermedio u otro compuesto químico asociado con la ruta biosintética de flavonoides y/o de taninos condensados.

45 También se describe un método para manipular la ruta biosintética de flavonoides y/o de taninos condensados, caracterizado por la etapa de alterar un ácido nucleico sustancialmente como se describe anteriormente para producir un gen que codifica un péptido no funcional.

También se describe el uso de una molécula de ácido nucleico aislada de la presente invención *in planta* para manipular los niveles de LAR y/o ANR en una planta leguminosa o célula de la planta.

50 Según otro aspecto, se proporciona el uso de una molécula de ácido nucleico aislada de la presente invención *in planta* para manipular los niveles de catequina y/o epicatequina u otro monómero de tanino (epigalocatequina o galocatequina) en una planta leguminosa o célula de la planta.

También se describe el uso de una molécula de ácido nucleico o polipéptido para identificar otros genes/polipéptidos reguladores de flavonoides y/o de taninos condensados relacionados.

- 5 En una realización, se puede manipular el conjunto del tejido vegetal. En una realización alternativa, se puede manipular el tejido epidérmico de la planta. Para los fines de esta memoria descriptiva, el tejido epidérmico se refiere al grupo de monocapa exterior de células, la hoja, los tallos, y raíces y tejidos jóvenes de una planta vascular.

Más preferiblemente, los niveles de taninos condensados alterados por la presente invención son suficientes para proporcionar un beneficio terapéutico o agronómico a un sujeto que consuma la planta con niveles alterados de flavonoides y/o taninos condensados.

Plantas producidas vía los métodos

- 10 En una realización adicional, la invención proporciona una planta producida mediante un método de la invención.

En una realización adicional, la invención proporciona una parte, semilla, fruta, material cosechado, propágulo o progenie de una planta de la invención.

- 15 En una realización adicional, la parte, semilla, fruta, material cosechado, propágulo o progenie de la planta se modifica genéticamente para comprender al menos una molécula de ácido nucleico de la invención, o un constructo de la invención.

Fuente de ácidos nucleicos y proteínas de la invención

Los ácidos nucleicos y proteínas de la invención pueden derivar de cualquier planta, como se describe más abajo, o se pueden producir sintética o recombinantemente.

Plantas

- 20 Las células vegetales y plantas de la invención, o aquellas transformadas o manipuladas en métodos y usos de la invención, pueden proceder de cualquier especie.

En una realización, la célula vegetal o planta deriva de especies de planta gimnosperma.

En una realización adicional, la célula vegetal o planta deriva de una especie de planta angiosperma.

En una realización adicional, la célula vegetal o planta deriva de una especie de planta dicotiledónea.

- 25 En una realización adicional, la célula vegetal o planta deriva de una especie de planta monocotiledónea.

Preferiblemente, las plantas proceden de especies dicotiledóneas.

Otras plantas preferidas son especies de plantas forrajeras de un grupo que comprende pero no se limita a los siguientes géneros: *Lolium*, *Festuca*, *Dactylis*, *Bromus*, *Thinopyrum*, *Trifolium*, *Medicago*, *Pheleum*, *Phalaris*, *Holcus*, *Lotus*, *Plantago* y *Cichorium*.

- 30 Otras plantas preferidas son plantas leguminosas. La planta leguminosa o parte de la misma puede englobar cualquier planta en la familia de plantas *Leguminosae* o *Fabaceae*. Por ejemplo, las plantas se pueden seleccionar de legumbres forrajeras que incluyen alfalfa, trébol; leucaena; legumbres de grano que incluyen habas, lentejas, altramuces, guisantes, cacahuets, haba de soja; legumbres de pruína que incluyen altramuces, legumbres farmacéuticas o industriales; y especies de legumbres de barbecho o de abono natural.

- 35 Un género particularmente preferido es *Trifolium*.

Las especies de *Trifolium* preferidas incluyen *Trifolium repens*; *Trifolium arvense*; *Trifolium affine*; y *Trifolium occidentale*.

Una especie de *Trifolium* particularmente preferida es *Trifolium repens*.

Otro género preferido es *Medicago*.

- 40 Especies de *Medicago* preferidas incluyen *Medicago sativa* y *Medicago truncatula*.

Una especie de *Medicago* particularmente preferida es *Medicago sativa*, conocida habitualmente como alfalfa.

Otro género preferido es *Glycine*.

Las especies de *Glycine* preferidas incluyen *Glycine max* y *Glycine wightii* (también conocida como *Neonotonia wightii*).

- 45 Una especie de *Glycine* particularmente preferida es *Glycine max*, conocida habitualmente como haba de soja.

Una especie de *Glycine* particularmente preferida es *Glycine wightii*, conocida habitualmente como haba de soja perenne.

Otro género preferido es *Vigna*.

La especie de *Vigna* preferida incluye *Vigna unguiculata*.

- 5 Una especie de *Vigna* particularmente preferida es *Vigna unguiculata*, habitualmente conocida como frijol de carita.

Otro género preferido es *Mucana*.

La especie de *Mucana* preferida incluye *Mucana pruniens*.

Una especie de *Mucana* particularmente preferida es *Mucana pruniens*, habitualmente conocida como grano de terciopelo.

- 10 Otro género preferido es *Arachis*.

Una especie de *Arachis* preferida incluye *Arachis glabrata*.

Una especie de *Arachis* particularmente preferida es *Arachis glabrata*, habitualmente conocida como cacahuete perenne.

Otro género preferido es *Pisum*.

- 15 Una especie de *Pisum* preferida incluye *Pisum sativum*.

Una especie de *Pisum* particularmente preferida es *Pisum sativum*, habitualmente conocido como guisante.

Otro género preferido es *Lotus*.

Especies de *Lotus* preferidas incluyen *Lotus corniculatus*, *Lotus pedunculatus*, *Lotus glabar*, *Lotus tenuis* y *Lotus uliginosus*.

- 20 Una especie de *Lotus* particularmente preferida es *Lotus corniculatus*, habitualmente conocido como loto corniculado.

Una especie de *Lotus* particularmente preferida es *Lotus glabar*, habitualmente conocido como loto corniculado de hoja estrecha.

Una especie de *Lotus* particularmente preferida es *Lotus pedunculatus*, habitualmente conocido como lotera.

- 25 Una especie de *Lotus* particularmente preferida es *Lotus tenuis*, habitualmente conocido como trébol esbelto.

Otro género preferido es *Brassica*.

Una especie de *Brassica* preferida incluye *Brassica oleracea*.

Una especie de *Brassica* particularmente preferida es *Brassica oleracea*, habitualmente conocida como berza o repollo forrajero.

- 30 El término "planta", como se usa aquí, se refiere a la planta en su totalidad, y cualquier parte de la misma, puede incluir pero no se limita a: porciones seleccionadas de la planta durante el ciclo de vida de la planta, tales como semillas vegetales, brotes, hojas, corteza, vainas, raíces, flores, fruto, tallos, y similares. Una "parte de la misma" preferida son las hojas.

DESCRIPCIÓN DETALLADA DE LA INVENCION

- 35 En esta memoria descriptiva en la que se ha hecho referencia a memorias descriptivas de patentes, otros documentos externos u otras fuentes de información, esto es generalmente con el fin de proporcionar un contexto para discutir las características de la invención. Excepto que se señale específicamente de otro modo, la referencia a tales documentos externos no se ha de interpretar como una admisión de que tales documentos, o tales fuentes de información, en cualquier jurisdicción, son técnica anterior, o forman parte del conocimiento general común en la técnica.

- 40 El término "que comprende", como se usa en esta memoria descriptiva y en las reivindicaciones, significa "que consiste al menos en parte en"; es decir, cuando se interpretan afirmaciones en esta memoria descriptiva y en las reivindicaciones que incluyen "que comprende", las características prologadas por este término en cada afirmación necesitan estar todas ellas presentes, pero también pueden estar presentes otras características. Términos relacionados, tales como "comprender" y "comprendido", se han de interpretar de manera similar. Sin embargo, en
- 45

realizaciones preferidas, “que comprende” se puede sustituir por “que consiste”.

La expresión “polipéptido MYB14” se refiere a un factor de transcripción MYB de clase R2R3.

Preferiblemente, el polipéptido MYB14 comprende una secuencia con al menos 70% de identidad con SEC ID NO: 14.

5 El polipéptido MYB14 puede comprender el motivo de secuencia de SEC ID NO: 15.

El polipéptido MYB14 puede comprender el motivo de secuencia de SEC ID NO: 17.

El polipéptido MYB14 puede comprender la secuencia de SEC ID NO: 15 y SEC ID NO: 17, pero carece de la secuencia de SEC ID NO: 16.

Un “gen MYB14” es un gen, por la definición estándar de gen, que codifica un polipéptido MYB14.

10 La expresión “factor de transcripción MYB” es un término bien entendido por los expertos en la técnica para referirse a una clase de factores de transcripción caracterizada por un dominio de unión a ADN estructuralmente conservado que consiste en repeticiones imperfectas individuales o múltiples.

15 La expresión “factor de transcripción R2R3” o “transcripción MYB con un dominio de unión a ADN R2R3” es una expresión bien entendida por los expertos en la técnica para referirse a factores de transcripción MYB de la clase de dos repeticiones.

Los términos “proantocianidinas” y “taninos condensados” se pueden usar de forma intercambiable a lo largo de la memoria descriptiva.

La expresión “motivo de secuencia”, como se usa aquí, significa un tramo de aminoácidos o nucleótidos. Preferiblemente, el tramo de aminoácidos o nucleótidos es contiguo.

20 El término “alterada”, con respecto a una planta con “producción alterada” o “expresión alterada”, significa alterada con respecto a la misma planta, o planta del mismo tipo, en el estado no transformado.

El término “alterada” puede significar incrementada o disminuida. Preferiblemente, alterada es incrementada.

Polinucleótidos y fragmentos

25 El término “polinucleótido o polinucleótidos”, como se usa aquí, significa un polímero desoxirribonucleotídico o ribonucleotídico monocatenario o bicatenario de cualquier longitud pero preferiblemente al menos 15 nucleótidos, e incluye, como ejemplos no limitantes, secuencias codificantes y no codificantes de un gen, complementos de secuencias sentido y antisentido, exones, intrones, ADN genómico, ADNC, pre-ARNm, ARNm, ARNr, ARNis, miARN, ARNt, ribozimas, polipéptidos recombinantes, secuencias de ADN o ARN de origen natural aisladas y purificadas, secuencias de ARN y ADN sintéticas, sondas, cebadores y fragmentos de ácidos nucleicos.

30 El término “polinucleótido” se puede usar de forma intercambiable con “molécula de ácido nucleico”.

35 Un “fragmento” de una secuencia polinucleotídica proporcionada aquí es una subsecuencia de nucleótidos contiguos que tiene una longitud preferiblemente de al menos 15 nucleótidos. Los fragmentos de la invención comprenden preferiblemente al menos 20 nucleótidos, más preferiblemente al menos 30 nucleótidos, más preferiblemente al menos 40 nucleótidos, más preferiblemente al menos 50 nucleótidos, y lo más preferible al menos 60 nucleótidos contiguos de un polinucleótido de la invención. Un fragmento de una secuencia polinucleotídica se puede usar en tecnología antisentido, de silenciamiento génico, de triple hélice o de ribozima, o como un cebador, una sonda, incluido en una micromatriz, o usado en métodos de selección a base de polinucleótidos.

40 Preferiblemente, los fragmentos de secuencias polinucleotídicas de la invención comprenden al menos 25, más preferiblemente al menos 50, más preferiblemente al menos 75, más preferiblemente al menos 100, más preferiblemente al menos 150, más preferiblemente al menos 200, más preferiblemente al menos 300, más preferiblemente al menos 400, más preferiblemente al menos 500, más preferiblemente al menos 600, más preferiblemente al menos 700, más preferiblemente al menos 800, más preferiblemente al menos 900, más preferiblemente al menos 1000 nucleótidos contiguos del polinucleótido especificado.

45 El término “cebador” se refiere a un polinucleótido corto, que tiene habitualmente un grupo libre 3'OH, que se hibrida a un molde y se usa para cebar la polimerización de un polinucleótido complementario al molde. Tal cebador tiene preferiblemente al menos 5, más preferiblemente al menos 6, más preferiblemente al menos 7, más preferiblemente al menos 9, más preferiblemente al menos 10, más preferiblemente al menos 11, más preferiblemente al menos 12, más preferiblemente al menos 13, más preferiblemente al menos 14, más preferiblemente al menos 15, más preferiblemente al menos 16, más preferiblemente al menos 17, más preferiblemente al menos 18, más preferiblemente al menos 19, más preferiblemente al menos 20 nucleótidos de longitud.

50

El término “sonda” se refiere a un polinucleótido corto que se usa para detectar una secuencia polinucleotídica, que es complementaria a la sonda, en un ensayo a base de hibridación. La sonda puede consistir en un “fragmento” de un polinucleótido como se define aquí. Preferiblemente, tal sonda tiene al menos 5, más preferiblemente al menos 10, más preferiblemente al menos 20, más preferiblemente al menos 30, más preferiblemente al menos 40, más preferiblemente al menos 50, más preferiblemente al menos 100, más preferiblemente al menos 200, más preferiblemente al menos 300, más preferiblemente al menos 400 y lo más preferible al menos 500 nucleótidos de longitud.

Polipéptidos y fragmentos

El término “polipéptido”, como se usa aquí, engloba cadenas de aminoácidos de cualquier longitud, pero preferiblemente al menos 5 aminoácidos, incluyendo proteínas de longitud completa, en las que los restos de aminoácidos están enlazados mediante enlaces peptídicos covalentes. Los polipéptidos pueden ser productos naturales purificados, o se pueden producir parcial o totalmente usando técnicas recombinantes o sintéticas. El término puede referirse a un polipéptido, un agregado de un polipéptido tal como un dímero u otro multímero, un polipéptido de fusión, un fragmento polipeptídico, una variante polipeptídica, o derivado del mismo.

Un “fragmento” de un polipéptido es una subsecuencia del polipéptido que realiza una función que se requiere para la actividad biológica y/o proporciona estructura tridimensional del polipéptido. El término se puede referir a un polipéptido, un agregado de un polipéptido tal como un dímero u otro multímero, un polipéptido de fusión, un fragmento polipeptídico, una variante polipeptídica, o derivado del mismo, capaz de llevar a cabo la actividad anterior.

El término “aislada”, como se aplica a las secuencias polinucleotídicas o polipeptídicas descritas aquí, se usa para referirse a secuencias que están eliminadas de su entorno celular natural. Una molécula aislada se puede obtener mediante cualquier método o combinación de métodos incluyendo técnicas bioquímicas, recombinantes, y sintéticas.

La expresión “derivada de”, con respecto a una secuencia polinucleotídica o polipeptídica que deriva de un género o especie particular, significa que la secuencia tiene la misma secuencia que una secuencia polinucleotídica o polipeptídica encontrada de forma natural en ese género o especie. La secuencia, derivada de un género o especie particular, se puede producir por lo tanto de forma sintética o recombinante.

Variantes

Como se usa aquí, el término “variante” se refiere a secuencias polinucleotídicas o polipeptídicas diferentes de las secuencias identificadas específicamente, en las que uno o más nucleótidos o restos de aminoácidos están suprimidos, sustituidos, o añadidos. Las variantes pueden ser variantes alélicas de origen natural, o variantes de origen no natural. Las variantes pueden proceder de la misma especie o de otra especie, y pueden englobar homólogos, parálogos y ortólogos. En ciertas realizaciones, las variantes de los polinucleótidos y polipéptidos de la invención poseen actividades biológicas que son las mismas o similares a las de los polinucleótidos o polipéptidos de la invención. El término “variante”, con referencia a polinucleótidos y polipéptidos, engloba todas las formas de polinucleótidos y polipéptidos como se definen aquí.

Variantes polinucleotídicas

Las secuencias polinucleotídicas variantes exhiben preferiblemente al menos 50%, más preferiblemente al menos 51%, más preferiblemente al menos 52%, más preferiblemente al menos 53%, más preferiblemente al menos 54%, más preferiblemente al menos 55%, más preferiblemente al menos 56%, más preferiblemente al menos 57%, más preferiblemente al menos 58%, más preferiblemente al menos 59%, más preferiblemente al menos 60%, más preferiblemente al menos 61%, más preferiblemente al menos 62%, más preferiblemente al menos 63%, más preferiblemente al menos 64%, más preferiblemente al menos 65%, más preferiblemente al menos 66%, más preferiblemente al menos 67%, más preferiblemente al menos 68%, más preferiblemente al menos 69%, más preferiblemente al menos 70%, más preferiblemente al menos 71%, más preferiblemente al menos 72%, más preferiblemente al menos 73%, más preferiblemente al menos 74%, más preferiblemente al menos 75%, más preferiblemente al menos 76%, más preferiblemente al menos 77%, más preferiblemente al menos 78%, más preferiblemente al menos 79%, más preferiblemente al menos 80%, más preferiblemente al menos 81%, más preferiblemente al menos 82%, más preferiblemente al menos 83%, más preferiblemente al menos 84%, más preferiblemente al menos 85%, más preferiblemente al menos 86%, más preferiblemente al menos 87%, más preferiblemente al menos 88%, más preferiblemente al menos 89%, más preferiblemente al menos 90%, más preferiblemente al menos 91%, más preferiblemente al menos 92%, más preferiblemente al menos 93%, más preferiblemente al menos 94%, más preferiblemente al menos 95%, más preferiblemente al menos 96%, más preferiblemente al menos 97%, más preferiblemente al menos 98%, y lo más preferible al menos 99% de identidad con una secuencia polinucleotídica especificada. La identidad se encuentra a lo largo de una ventana de comparación de al menos 20 posiciones nucleotídicas, más preferiblemente al menos 50 posiciones nucleotídicas, más preferiblemente al menos 100 posiciones nucleotídicas, más preferiblemente al menos 200 posiciones nucleotídicas, más preferiblemente al menos 300 posiciones nucleotídicas, más preferiblemente al menos 400 posiciones nucleotídicas, más preferiblemente al menos 500 posiciones nucleotídicas, más preferiblemente al menos

600 posiciones nucleotídicas, más preferiblemente al menos 700 posiciones nucleotídicas, más preferiblemente al menos 800 posiciones nucleotídicas, más preferiblemente al menos 900 posiciones nucleotídicas, más preferiblemente al menos 1000 posiciones nucleotídicas, y lo más preferible a lo largo de toda la longitud de la secuencia polinucleotídica especificada.

- 5 La identidad de la secuencia polinucleotídica se puede determinar de la siguiente manera. La secuencia polinucleotídica objeto se compara con una secuencia polinucleotídica candidato usando BLASTN (de la colección de programas BLAST, versión 2.2.5 [noviembre 2002]) en bl2seq (Tatiana A. Tatusova, Thomas L. Madden (1999), "Blast 2 sequences - a new tool for comparing protein and nucleotide sequences", FEMS Microbiol Lett. 174:247-250), que está públicamente disponible de NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). Se utilizan los parámetros por defecto de bl2seq, excepto que se debería apagar el filtrado de las partes de baja complejidad.

La identidad de las secuencias polinucleotídicas se puede examinar usando los siguientes parámetros de línea de comando de unix:

```
bl2seq -i nucleotideseq1 -j nucleotideseq2 -F F -p blastn
```

- 15 El parámetro -F F apaga el filtrado de secciones de baja complejidad. El parámetro -p selecciona el algoritmo apropiado para el par de secuencias. El programa bl2seq da a conocer la identidad de secuencia como tanto el número como el porcentaje de nucleótidos idénticos en una línea "Identities =".

- 20 La identidad de las secuencias polinucleotídicas también se puede calcular a lo largo de toda la longitud del solapamiento entre una secuencia polinucleotídica candidato y una secuencia polinucleotídica objeto, usando programas de alineamiento global de secuencias (por ejemplo, Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). Una implementación completa del algoritmo de alineamiento global de Needleman-Wunsch se encuentra en el programa needle en el paquete EMBOSS (Rice, P. Longden, I. y Bleasby, A. EMBOSS: The European Molecular Biology Open Software Suite, Trends in Genetics junio 2000, vol. 16, nº 6. p. 276-277) que se puede obtener de <http://www.hgmp.mrc.ac.uk/Software/EMBOSS/>. El servidor del European Bioinformatics Institute también proporciona el servicio para llevar a cabo alineamientos globales de EMBOSS-needle entre dos secuencias en línea en <http://www.ebi.ac.uk/emboss/align/>.

Como alternativa, para calcular la identidad de secuencia, se puede usar el programa GAP, que calcula un alineamiento global óptimo de dos secuencias sin penalizar saltos terminales. GAP se describe en el siguiente documento: Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227-235.

- 30 La identidad de secuencia también se puede calcular alineando secuencias a comparar usando Vector NTI versión 9.0, que usa un algoritmo Clustal W (Thompson et al., 1994, Nucleic Acids Research 24, 4876-4882), calculando entonces el porcentaje de identidad de secuencia entre las secuencias alineadas usando Vector NTI versión 9.0 (septiembre 02, 2003 ©1994-2003 InforMax, con licencia a Invitrogen).

- 35 Las variantes polinucleotídicas de la presente invención también engloban aquellas que exhiben una similitud con una o más de las secuencias específicamente identificadas que es probable que conserven la equivalencia funcional de esas secuencias y que no se podría esperar razonablemente que se produjeran por casualidad al azar. Tal similitud de secuencias con respecto a polinucleótidos se puede determinar usando el programa bl2seq públicamente disponible de la colección de programas BLAST (versión 2.2.5 [noviembre 2002]) de NCBI (<ftp://ftp.ncbi.nih.gov/blast/>).

- 40 La similitud de secuencias polinucleotídicas se puede examinar usando los siguientes parámetros de línea de comando de unix:

```
bl2seq -i nucleotideseq1 -j nucleotideseq2 -F F -p tblastx
```

- 45 El parámetro -F F apaga el filtrado de secciones de baja complejidad. El parámetro -p selecciona el algoritmo apropiado para el par de secuencias. Este programa encuentra regiones de similitud entre las secuencias, y para cada una de tales regiones da a conocer un "valor E" que es el número esperado de veces que se podría esperar observar tal emparejamiento por casualidad en una base de datos de un tamaño de referencia fijo que contiene secuencias aleatorias. El tamaño de esta base de datos se establece por defecto en el programa bl2seq. Para valores E pequeños, mucho menores que uno, el valor E es aproximadamente la probabilidad de tal emparejamiento al azar.

- 50 Las secuencias polinucleotídicas variantes exhiben preferiblemente un valor E menor que menor que 1×10^{-10} más preferiblemente menor que 1×10^{-20} , más preferiblemente menor que 1×10^{-30} , más preferiblemente menor que 1×10^{-40} , más preferiblemente menor que 1×10^{-50} , más preferiblemente menor que 1×10^{-60} , más preferiblemente menor que 1×10^{-70} , más preferiblemente menor que 1×10^{-80} , más preferiblemente menor que 1×10^{-90} y lo más preferible menor que 1×10^{-100} cuando se comparan con una cualquiera de las secuencias específicamente identificadas.

55

Como alternativa, los polinucleótidos variantes de la presente invención se hibridan a una secuencia polinucleotídica especificada, o a sus complementos, en condiciones restrictivas.

La expresión "hibridar en condiciones restrictivas", y sus equivalentes gramaticales, se refiere a la capacidad de una molécula polinucleotídica para hibridarse a una molécula polinucleotídica diana (tal como una molécula polinucleotídica diana inmovilizada en una mancha de ADN o ARN, tal como una transferencia Southern o una transferencia Northern) en condiciones definidas de temperatura y concentración salina. La capacidad para hibridarse en condiciones de hibridación restrictivas se puede determinar hibridando inicialmente en condiciones menos restrictivas e incrementando después la restricción hasta la restricción deseada.

Con respecto a moléculas polinucleotídicas mayores que alrededor de 100 bases de longitud, las condiciones de hibridación restrictivas típicas son no más de 25 a 30°C (por ejemplo, 10°C) por debajo de la temperatura de fusión (T_m) del dúplex nativo (véanse generalmente, Sambrook et al., Eds, 1987, Molecular Cloning, A Laboratory Manual, 2ª Ed. Cold Spring Harbor Press; Ausubel et al., 1987, Current Protocols in Molecular Biology, Greene Publishing). La T_m para moléculas polinucleotídicas mayores que alrededor de 100 bases se puede calcular mediante la fórmula $T_m = 81,5 + 0,41\% (G + C - \log (Na^+))$. (Sambrook et al., Eds, 1987, Molecular Cloning, A Laboratory Manual, 2ª Ed. Cold Spring Harbor Press; Bolton y McCarthy, 1962, PNAS 84:1390). Las condiciones restrictivas típicas para polinucleótido de más de 100 bases de longitud serían condiciones de hibridación tales como prelavado en una disolución de 6X SSC, 0,2% de SDS; hibridación a 65°C, 6X SSC, 0,2% de SDS toda la noche; seguido de dos lavados de 30 minutos cada uno en 1X SSC, 0,1% de SDS a 65°C y dos lavados de 30 minutos cada uno en 0,2X SSC, 0,1% de SDS a 65°C.

Con respecto a las moléculas polinucleotídicas que tienen una longitud menor que 100 bases, las condiciones de hibridación restrictivas ejemplares son 5 a 10°C por debajo de T_m . De media, la T_m de una molécula polinucleotídica de longitud menor que 100 pb se reduce en aproximadamente (500/longitud del oligonucleótido)°C.

Con respecto a los miméticos de ADN conocidos como ácidos peptidonucleicos (PNAs) (Nielsen et al., Science. 6 de diciembre de 1991; 254(5037):1497-500), los valores de T_m son mayores que aquellos para híbridos de ADN-ADN o de ADN-ARN, y se pueden calcular usando la fórmula descrita en Giesen et al., Nucleic Acids Res. 1 de noviembre de 1998; 26(21):5004-6. Las condiciones de hibridación restrictivas ejemplares para un híbrido de ADN-PNA que tiene una longitud menor que 100 bases son 5 a 10°C por debajo de la T_m .

Los polinucleótidos variantes tales como aquellos en constructos de la invención que codifican proteínas a expresar también engloban polinucleótidos que difieren de las secuencias especificadas, pero que, como consecuencia de la degeneración del código genético, codifican un polipéptido que tiene actividad similar a un polipéptido codificado por un polinucleótido de la presente invención. Una alteración de secuencia que no cambie la secuencia de aminoácidos del polipéptido es una "variación silenciosa". Excepto para ATG (metionina) y TGG (triptófano), se pueden cambiar otros codones para el mismo aminoácido mediante técnicas reconocidas en la técnica, por ejemplo para optimizar la expresión de codones en un organismo hospedante particular.

También se contemplan alteraciones de secuencias polinucleotídicas que dan como resultado sustituciones conservativas de uno o varios aminoácidos en la secuencia polipeptídica codificada sin alterar significativamente su actividad biológica. Un experto estará al tanto de métodos para obtener sustituciones de aminoácidos fenotípicamente silenciosas (véase, por ejemplo, Bowie et al., 1990, Science 247, 1306).

Los polinucleótidos variantes debidos a variaciones silenciosas y sustituciones conservativas en la secuencia polipeptídica codificada se puede determinar usando el programa `bl2seq` públicamente disponible de la colección de programas BLAST (versión 2.2.5 [noviembre 2002]) de NCBI (<ftp://ftp.ncbi.nih.gov/blast/>) vía el algoritmo `tblastx` como se describe previamente.

Variantes polipeptídicas

El término "variante", con referencia a polipéptidos, engloba polipéptidos de origen natural, producidos recombinantemente y producidos sintéticamente. Las secuencias polipeptídicas variantes exhiben preferiblemente al menos 50%, más preferiblemente al menos 51%, más preferiblemente al menos 52%, más preferiblemente al menos 53%, más preferiblemente al menos 54%, más preferiblemente al menos 55%, más preferiblemente al menos 56%, más preferiblemente al menos 57%, más preferiblemente al menos 58%, más preferiblemente al menos 59%, más preferiblemente al menos 60%, más preferiblemente al menos 61%, más preferiblemente al menos 62%, más preferiblemente al menos 63%, más preferiblemente al menos 64%, más preferiblemente al menos 65%, más preferiblemente al menos 66%, más preferiblemente al menos 67%, más preferiblemente al menos 68%, más preferiblemente al menos 69%, más preferiblemente al menos 70%, más preferiblemente al menos 71%, más preferiblemente al menos 72%, más preferiblemente al menos 73%, más preferiblemente al menos 74%, más preferiblemente al menos 75%, más preferiblemente al menos 76%, más preferiblemente al menos 77%, más preferiblemente al menos 78%, más preferiblemente al menos 79%, más preferiblemente al menos 80%, más preferiblemente al menos 81%, más preferiblemente al menos 82%, más preferiblemente al menos 83%, más preferiblemente al menos 84%, más preferiblemente al menos 85%, más preferiblemente al menos 86%, más preferiblemente al menos 87%, más preferiblemente al menos 88%, más preferiblemente al menos 89%, más

preferiblemente al menos 90%, más preferiblemente al menos 91%, más preferiblemente al menos 92%, más preferiblemente al menos 93%, más preferiblemente al menos 94%, más preferiblemente al menos 95%, más preferiblemente al menos 96%, más preferiblemente al menos 97%, más preferiblemente al menos 98%, y lo más preferible al menos 99% de identidad con una secuencia de la presente invención. La identidad se encuentra a lo largo de una ventana de comparación de al menos 20 posiciones de aminoácidos, preferiblemente al menos 50 posiciones de aminoácidos, más preferiblemente al menos 100 posiciones de aminoácidos, y lo más preferible a lo largo de toda la longitud de un polipéptido de la invención.

La identidad de secuencia polipeptídica se puede determinar de la siguiente manera. La secuencia polipeptídica objeto se compara con una secuencia polipeptídica candidato usando BLASTP (de la colección de programas BLAST, versión 2.2.5 [noviembre 2002]) en bl2seq, que está públicamente disponible de NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). Se utilizan los parámetros por defecto de bl2seq, excepto que se debería de apagar el filtrado de regiones de baja complejidad.

La identidad de secuencias polipeptídica también se puede calcular a lo largo de toda la longitud del solapamiento entre una secuencia polinucleotídica candidato y una secuencia polinucleotídica objeto, usando programas de alineamiento global de secuencias. EMBOSS-needle (disponible en <http://www.ebi.ac.uk/emboss/align/>) y GAP (Huang, X. (1994) On Global Sequence Alignment. Computer Applications in the Biosciences 10, 227-235), como se explican anteriormente, también son programas adecuados de alineamiento global de secuencias para calcular la identidad de secuencias polipeptídicas.

La identidad de secuencia también se puede calcular alineando secuencias a comparar usando Vector NTI versión 9.0, que usa un algoritmo Clustal W (Thompson et al., 1994, Nucleic Acids Research 24, 4876-4882), calculando entonces el porcentaje de identidad de secuencia entre las secuencias polipeptídicas alineadas usando Vector NTI versión 9.0 (sept. 02, 2003 ©1994-2003 InforMax, con licencia a Invitrogen).

Las variantes polipeptídicas de la presente invención también engloban aquellas que muestran una similitud con una o más de las secuencias específicamente identificadas que es probable que conserven la equivalencia funcional de aquellas secuencias y que razonablemente no se podría esperar que se produjeran por casualidad al azar. Tal similitud de secuencia con respecto a polipéptidos se puede determinar usando el programa bl2seq públicamente disponible de la colección de programas BLAST (versión 2.2.5 [noviembre 2002]) de NCBI (<ftp://ftp.ncbi.nih.gov/blast/>). La similitud de secuencias polipeptídicas se puede examinar usando los siguientes parámetros de línea de comando de unix:

```
bl2seq -i nucleotideseq1 -j nucleotideseq2 -F F -p blastp
```

Las secuencias polipeptídicas variantes exhiben preferiblemente un valor E menor que menor que 1×10^{-6} más preferiblemente menor que 1×10^{-9} , más preferiblemente menor que 1×10^{-12} , más preferiblemente menor que 1×10^{-15} , más preferiblemente menor que 1×10^{-18} , más preferiblemente menor que 1×10^{-21} , más preferiblemente menor que 1×10^{-30} , más preferiblemente menor que 1×10^{-40} , más preferiblemente menor que 1×10^{-50} , más preferiblemente menor que 1×10^{-60} , más preferiblemente menor que 1×10^{-70} , más preferiblemente menor que 1×10^{-80} , más preferiblemente menor que 1×10^{-90} y lo más preferible menor que 1×10^{-100} cuando se comparan con una cualquiera de las secuencias específicamente identificadas.

El parámetro -FF apaga el filtrado de secciones de baja complejidad. El parámetro -p selecciona el algoritmo apropiado para el par de secuencias. Este programa encuentra regiones de similitud entre las secuencias, y para cada una de tales regiones da a conocer un "valor E" que es el número de veces esperado que se podría esperar observar tal emparejamiento por casualidad en una base de datos de un tamaño de referencia fijo que contiene secuencias aleatorias. Para valores E pequeños, mucho menores que uno, éste es aproximadamente la probabilidad de tal emparejamiento aleatorio.

También se incluyen en la invención sustituciones conservativas de uno o varios aminoácidos de una secuencia polipeptídica descrita sin alterar significativamente su actividad biológica. Un experto estará al tanto de métodos para obtener sustituciones de aminoácidos fenotípicamente silenciosas (véase, por ejemplo, Bowie et al., 1990, Science 247, 1306).

Constructos, vectores y sus componentes

La expresión "constructo genético" se refiere a una molécula polinucleotídica, habitualmente ADN bicatenario, que puede tener insertada en ella otra molécula polinucleotídica (la molécula polinucleotídica inserta) tal como, pero sin limitarse a, una molécula de ADNc. Un constructo genético puede contener un polinucleótido promotor que incluye los elementos necesarios que permiten transcribir la molécula polinucleotídica inserta, y, opcionalmente, traducir el transcrito a un polipéptido. La molécula polinucleotídica inserta puede derivar de la célula hospedante, o puede derivar de una célula u organismo diferente, y/o puede ser un polinucleótido sintético o recombinante. Una vez dentro de la célula hospedante, el constructo genético puede integrarse en el ADN cromosómico del hospedante. El constructo genético puede estar enlazado a un vector.

El término "vector" se refiere a una molécula polinucleotídica, habitualmente ADN bicatenario, que se usa para

transportar el constructo genético a una célula hospedante. El vector puede ser capaz de replicarse en al menos un sistema hospedante adicional, tal como *E. coli*.

La expresión “constructo de expresión” se refiere a un constructo genético que incluye los elementos necesarios que permiten transcribir la molécula polinucleotídica inserta, y, opcionalmente, traducir el transcrito a un polipéptido.

5 Un constructo de expresión comprende típicamente, en una dirección de 5' a 3':

- a) un promotor funcional en la célula hospedante en la que el constructo se transformará,
- b) el polinucleótido a expresar, y
- c) un terminador funcional en la célula hospedante en la que se transformará el constructo.

10 La expresión “región codificante” o “marco de lectura abierto” (ORF) se refiere a la hebra sentido de una secuencia de ADN genómico o una secuencia de ADNc que es capaz de producir un producto de transcripción y/o un polipéptido bajo el control de secuencias reguladoras apropiadas. La secuencia codificante se identifica mediante la presencia de un codón de inicio de la traducción de 5' y un codón de parada de la traducción de 3'. Cuando se inserta en un constructo genético, una “secuencia codificante” es capaz de ser expresada cuando está operativamente enlazada a secuencias promotoras y terminadoras.

15 La expresión “operativamente enlazada” significa que la secuencia a expresar se coloca bajo el control de elementos reguladores que incluyen promotores, elementos reguladores específicos de tejido, elementos reguladores temporales, potenciadores, represores y terminadores.

20 La expresión “región no codificante” incluye secuencias no traducidas que están en dirección 5' del sitio de inicio traduccional, y en dirección 3' del sitio de parada traduccional. Estas secuencias también se denominan respectivamente como la 5' UTR y la 3' UTR. Estas secuencias pueden incluir elementos requeridos para la iniciación y terminación de la transcripción y para la regulación de la eficiencia de la traducción. La expresión “no codificante” también incluye secuencias intrónicas en clones genómicos.

25 Los terminadores son secuencias que terminan la transcripción, y se encuentran en los extremos no traducidos de 3' de genes en dirección 3' de la secuencia traducida. Los terminadores son determinantes importantes de la estabilidad del ARNm, y en algunos casos se ha encontrado que tienen funciones reguladoras espaciales.

30 El término “promotor” se refiere a una secuencia polinucleotídica capaz de regular o conducir la expresión de una secuencia polinucleotídica a la que el promotor está operativamente enlazado en una célula, o célula libre del sistema de transcripción. Los promotores pueden comprender elementos iniciadores en cis que especifican el sitio de iniciación de la transcripción, y cajas conservadas tales como la caja TATA, y motivos que están unidos mediante factores de transcripción.

Métodos para aislar o producir polinucleótidos

35 Las moléculas polinucleotídicas de la invención se pueden aislar usando una variedad de técnicas conocidas por aquellos de pericia normal en la técnica. A título de ejemplo, tales polinucleótidos se pueden aislar mediante el uso de la reacción en cadena de la polimerasa (PCR) descrita en Mullis et al., Eds. 1994 The Polymerase Chain Reaction, Birkhauser, incorporada aquí como referencia. Los polinucleótidos de la invención se pueden amplificar usando cebadores, como se definen aquí, derivados de las secuencias polinucleotídicas de la invención.

40 Otros métodos para aislar polinucleótidos de la invención, o útiles en los métodos de la invención, incluyen el uso de todo o de porciones de los polinucleótidos expuestos aquí, como sondas de hibridación. La técnica de hibridar sondas polinucleotídicas marcadas a polinucleótidos inmovilizados sobre soportes sólidos, tales como filtros de nitrocelulosa o membranas de nailon, se puede usar para identificar el genómico. Las condiciones de hibridación y lavado ejemplares son: hibridación durante 20 horas a 65°C en 5,0 X SSC, 0,5% de dodecilsulfato de sodio, 1 X disolución de Denhardt; lavado (tres lavados durante veinte minutos cada uno a 55°C) en 1,0 X SSC, 1% (p/v) de dodecilsulfato de sodio, y opcionalmente un lavado (durante veinte minutos) en 0,5 X SSC, 1% (p/v) de dodecilsulfato de sodio, a 60°C. Se puede realizar un lavado adicional opcional (durante veinte minutos) en condiciones de 0,1 X SSC, 1% (p/v) de dodecilsulfato de sodio, a 60°C.

45 Los fragmentos polinucleotídicos de la invención se pueden producir mediante técnicas bien conocidas en la técnica, tales como digestión con endonucleasas de restricción, síntesis oligonucleotídica y amplificación mediante PCR.

50 Una secuencia polinucleotídica parcial se puede usar en los métodos bien conocidos en la técnica para identificar la secuencia polinucleotídica de longitud completa correspondiente y/o el gen completo y/o el promotor. Tales métodos incluyen métodos a base de PCR, 5'RACE (Frohman MA, 1993, Methods Enzymol. 218: 340-56) y método a base de hibridación, métodos a base de ordenador/base de datos. Además, a título de ejemplo, la PCR inversa permite la adquisición de secuencias desconocidas, que flanquean las secuencias polinucleotídicas descritas aquí, partiendo de cebadores basados en una región conocida (Triglia et al., 1998, Nucleic Acids Res 16, 8186, incorporado aquí como referencia). El método usa varias enzimas de restricción para generar un fragmento adecuado en la región

conocida de un polinucleótido. El fragmento se circulariza entonces mediante ligación intramolecular y se usa como un molde de la PCR. Los cebadores divergentes se diseñan a partir de la región conocida. Las secuencias promotoras y de flanqueo también se pueden aislar mediante paseo genómico mediante PCR usando un kit GenomeWalker™ (Clontech, Mountain View, California), siguiendo las instrucciones del fabricante. A fin de ensamblar físicamente clones de longitud completa, se pueden utilizar enfoques estándar de biología molecular (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2ª Ed. Cold Spring Harbor Press, 1987).

Puede ser beneficioso, cuando se produce una planta transgénica a partir de una especie particular, transformar tal planta con una secuencia o secuencias derivadas de esa especie. El beneficio puede ser para aliviar preocupaciones públicas con respecto a la transformación entre especies en la generación de organismos transgénicos. Adicionalmente, cuando el resultado deseado es la disminución de un gen, puede ser necesario utilizar una secuencia idéntica (o al menos muy similar) a aquella en la planta, para la que se desea la expresión reducida. Por estas razones, entre otras, es deseable ser capaces de identificar y aislar ortólogos de un gen particular en varias especies vegetales diferentes. Las variantes (incluyendo ortólogos) se pueden identificar mediante los métodos descritos.

Métodos para identificar variantes

Métodos físicos

Los polinucleótidos variantes se pueden identificar usando métodos a base de PCR (Mullis et al., Eds. 1994 The Polymerase Chain Reaction, Birkhauser).

Como alternativa, se pueden emplear métodos de cribado de librerías, bien conocidos por los expertos en la técnica (Sambrook et al., Molecular Cloning: A Laboratory Manual, 2ª Ed. Cold Spring Harbor Press, 1987). Cuando se identifican variantes de la secuencia sonda, la restricción de la hibridación y/o lavado se reducirá típicamente con respecto a cuando se buscan emparejamientos exactos de secuencias.

Métodos a base de ordenador

Las variantes polinucleotídicas y polipeptídicas también se pueden identificar por métodos a base de ordenador bien conocidos por los expertos en la técnica, usando algoritmos de alineamiento de secuencias de dominio público y herramientas de búsqueda de similitud de secuencias para buscar bases de datos de secuencias (las bases de datos de dominio público incluyen Genbank, EMBL, Swiss-Prot, PIR y otras). Para ejemplos de fuentes en línea, véase, por ejemplo, Nucleic Acids Res. 29: 1-10 y 11-16, 2001. Las búsquedas de similitud recuperan y alinean secuencias diana para la comparación con una secuencia a analizar (es decir, una secuencia problema). Los algoritmos de comparación de secuencias usan matrices de puntuación para asignar una puntuación global a cada uno de los alineamientos.

Una familia de programas ejemplar útil para identificar variantes en bases de datos de secuencias es la colección de programas BLAST (versión 2.2.5 [noviembre 2002] que incluye BLASTN, BLASTP, BLASTX, tBLASTN y tBLASTX, que están públicamente disponibles de (<ftp://ftp.ncbi.nih.gov/blast/>) o del National Center for Biotechnology Information (NCBI), National Library of Medicine, Building 38A, Room 8N805, Bethesda, MD 20894 USA. El servidor de NCBI también proporciona la herramienta para facilitar el uso de los programas para identificar un número de bases de datos de secuencias públicamente disponibles. BLASTN compara una secuencia problema nucleotídica frente a una base de datos de secuencias nucleotídicas. BLASTP compara una secuencia problema de aminoácidos frente a una base de datos de secuencias proteicas. BLASTX compara una secuencia problema nucleotídica traducida en todos los marcos de lectura frente a una base de datos de secuencias proteicas. tBLASTN compara una secuencia problema proteica frente a una base de datos de secuencias nucleotídicas traducida dinámicamente en todos los marcos de lectura. tBLASTX compara las traducciones en los seis marcos de una secuencia problema nucleotídica frente a las traducciones en los seis marcos de una base de datos de secuencias nucleotídicas. Los programas BLAST se pueden usar con parámetros por defecto, o los parámetros se pueden alterar según se necesite para refinar el cribado.

El uso de la familia de algoritmos BLAST, incluyendo BLASTN, BLASTP, y BLASTX, se describe en la publicación de Altschul et al., Nucleic Acids Res. 25: 3389-3402, 1997.

Los “aciertos” con respecto a una o más secuencias de la base de datos por una secuencia problema producidos por BLASTN, BLASTP, BLASTX, tBLASTN, tBLASTX, o un algoritmo similar, alinean e identifican porciones similares de secuencias. Los aciertos se disponen en orden del grado de similitud y la longitud de solapamiento de las secuencias. Los aciertos con respecto a una secuencia de base de datos generalmente representan un solapamiento a lo largo de solamente una fracción de la longitud de secuencia de la secuencia problema.

Los algoritmos BLASTN, BLASTP, BLASTX, tBLASTN y tBLASTX también producen valores “esperados” para los alineamientos. El valor Esperado (E) indica el número de aciertos que se pueden “esperar” observar por casualidad cuando se busca una base de datos del mismo tamaño que contiene secuencias contiguas aleatorias. El valor Esperado se usa como un umbral de significancia para determinar si el acierto con respecto a la base de datos indica similitud verdadera. Por ejemplo, un valor E de 0,1 asignado a un acierto polinucleotídico se interpreta como

que en una base de datos del tamaño de la base de datos cribada, se puede esperar observar 0,1 emparejamientos a lo largo de la porción alineada de la secuencia con una puntuación similar simplemente por casualidad. Para secuencias que tienen un valor E de 0,01 o menos a lo largo de porciones alineadas y emparejadas, la probabilidad de encontrar un emparejamiento por casualidad en esa base de datos es 1% o menos usando el algoritmo BLASTN, BLASTP, BLASTX, tBLASTN o tBLASTX.

Los alineamiento múltiples de secuencias de un grupo de secuencias relacionadas se puede llevar a cabo con CLUSTALW (Thompson, J.D., Higgins, D.G. y Gibson, T.J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22:4673-4680, <http://www-igbmc.u-strasbg.fr/BioInfo/ClustalW/Top.html>) o T-COFFEE (Cedric Notredame, Desmond G. Higgins, Jaap Heringa, T-Coffee: A novel method for fast and accurate multiple sequence alignment, *J. Mol. Biol.* (2000) 302: 205-217) o PILEUP, que usa alineamientos progresivos por parejas (Feng y Doolittle, 1987, *J. Mol. Evol.* 25, 351). Hay disponibles aplicaciones de software de reconocimiento de patrones para encontrar motivos o secuencias signatura. Por ejemplo, MEME (Múltiple Em para Extracción de Motivos) encuentra motivos y secuencias signatura en un conjunto de secuencias, y MAST (Herramienta de Alineamiento y Búsqueda de Motivos) usa estos motivos para identificar motivos similares o los mismos motivos en secuencias problema. Los resultados de MAST se proporcionan como una serie de alineamientos con datos estadísticos apropiados y un resumen visual de los motivos encontrados. MEME y MAST se desarrollaron en la University of California, San Diego.

PROSITE (Bairoch y Bucher, 1994, *Nucleic Acids Res.* 22, 3583; Hofmann et al., 1999, *Nucleic Acids Res.* 27, 215) es un método para identificar las funciones de proteínas no caracterizadas traducidas a partir de secuencias genómicas o de ADNc. La base de datos PROSITE (www.expasy.org/prosite) contiene patrones y perfiles biológicamente significativos, y está diseñada de manera que se puede usar con herramientas computacionales apropiadas para asignar una nueva secuencia a una familia conocida de proteínas o para determinar qué dominio o dominios conocidos están presentes en la secuencia (Falquet et al., 2002, *Nucleic Acids Res.* 30, 235). Prosearch es una herramienta que puede realizar búsquedas en bases de datos SWISS-PROT y EMBL con un patrón o signatura de secuencia dada.

Función de las variantes

La función de los polinucleótidos/polipéptidos de la invención se puede evaluar usando métodos proporcionados aquí. En particular, véase el Ejemplo 7.

Métodos para producir constructos y vectores

Los constructos genéticos de la presente invención comprenden una o más secuencias polinucleotídicas de la invención y/o polinucleótidos que codifican polipéptidos descritos, y pueden ser útiles para transformar, por ejemplo, organismos bacterianos, fúngicos, de insectos, de mamíferos o particularmente vegetales. Los constructos genéticos de la invención están destinados a incluir constructos de expresión como se definen aquí.

Los métodos para producir y usar constructos genéticos y vectores son bien conocidos en la técnica, y se describen generalmente en Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2ª Ed. Cold Spring Harbor Press, 1987; Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing, 1987).

Métodos para producir células hospedantes que comprenden constructos y vectores

La invención proporciona una célula hospedante que comprende un constructo genético o vector de la invención. Las células hospedantes pueden derivar de, por ejemplo, organismos bacterianos, fúngicos, de insectos, de mamíferos o vegetales.

Las células hospedantes que comprenden constructos genéticos, tales como constructos de expresión, de la invención son útiles en métodos bien conocidos en la técnica (por ejemplo Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2ª Ed. Cold Spring Harbor Press, 1987; Ausubel et al., *Current Protocols in Molecular Biology*, Greene Publishing, 1987) para la producción recombinante de polipéptidos. Tales métodos pueden implicar el cultivo de células hospedantes en un medio apropiado en condiciones adecuadas para o que conduzcan a la expresión de un polipéptido de la invención. El polipéptido recombinante expresado, que puede segregarse opcionalmente en el cultivo, se puede separar entonces del medio, células hospedantes o medio de cultivo mediante métodos bien conocidos en la técnica (por ejemplo Deutscher, Ed, 1990, *Methods in Enzymology*, Vol 182, *Guide to Protein Purification*).

Métodos para producir células vegetales y plantas que comprenden constructos y vectores

La invención proporciona además células vegetales que comprenden un constructo genético de la invención, y células vegetales modificadas para alterar la expresión de un polinucleótido o polipéptido. Las plantas que comprenden tales células también forman un aspecto de la invención.

Los métodos para transformar células vegetales, plantas y sus porciones con polinucleótidos se describen en Draper

et al., 1988, *Plant Genetic Transformation and Gene Expression. A Laboratory Manual*. Blackwell Sci. Pub. Oxford, p. 365; Potrykus y Spangenburg, 1995, *Gene Transfer to Plants*. Springer-Verlag, Berlin.; y Gelvin et al., 1993, *Plant Molecular Biol. Manual*. Kluwer Acad. Pub. Dordrecht. En Galun y Breiman, 1997, *Transgenic Plants*. Imperial College Press, Londres, se proporciona un repaso de plantas transgénicas, incluyendo técnicas de transformación.

5 Las siguientes son publicaciones representativas que describen protocolos de transformación genética que se pueden usar para transformar genéticamente las siguientes especies vegetales: arroz (Alam et al., 1999, *Plant Cell Rep.* 18, 572); manzana (Yao et al., 1995, *Plant Cell Reports* 14, 407-412); maíz (Patente US Serie n^{os} 5.177.010 y 5.981.840); trigo (Ortiz et al., 1996, *Plant Cell Rep.* 15, 1996, 877); tomate (Patente US Serie n^o 5.159.135); patata (Kumar et al., 1996 *Plant J.* 9: 821); casave (Li et al., 1996 *Nat. Biotechnology* 14, 736); lechuga (Michelmores et al., 10 1987, *Plant Cell Rep.* 6, 439); tabaco (Horsch et al., 1985, *Science* 227, 1229); algodón (Patentes US Serie n^{os} 5.846.797 y 5.004.863); ballico (Bajaj et al., 2006, *Plant Cell Rep.* 25, 651); gramas (Patentes US n^{os} 5.187.073, 6.020.539) y menta piperita (Niu et al., 1998, *Plant Cell-Rep.* 17, 165); cítricos (Pena et al., 1995, *Plant Sci.* 104, 183); alcaravea (Krens et al., 1997, *Plant Cell Rep.* 17, 39); plátano (Patente US Serie n^o 5.792.935); haba de soja (Patente US Serie n^{os} 5.416.011; 5.569.834; 5.824.877; 5.563.04455 y 5.968.830); piña (Patente US Serie n^o 5.952.543); álamo (Patente US n^o 4.795.855); monocotiledóneas en general (Patente US n^{os} 5.591.616 y 6.037.522); crucífera (Patente US n^{os} 5.188.958; 5.463.174 y 5.750.871); y cereales (Patente US n^o 6.074.877); pera (Matsuda et al., 2005, *Plant Cell Rep.* 24(1):45-51); Prunus (Ramesh et al., 2006, *Plant Cell Rep.* 25(8):821-8; Song y Sink 2005, *Plant Cell Rep.* 2006; 25(2):117-23; Gonzalez Padilla et al., 2003, *Plant Cell Rep.* 22(1):38-45); fresa (Oosumi et al., 2006, *Planta.*; 223(6):1219-30; Folta et al., 2006, *Planta.* 14 de abril de 2006; PMID: 16614818); rosa (Li et al., 2003, *Planta.* 218(2):226-32); Rubus (Graham et al., 1995, *Methods Mol Biol.* 1995; 44:129-33); clavo (Voisey et al., 1994, *Plant Cell Reports* 13: 309-314, y alfalfa (Bingham, 1991, *Crop Science* 31: 1098). También se contempla por la invención la transformación de otras especies. Los métodos y protocolos adecuados para la transformación de otras especies están disponibles en la bibliografía científica.

Métodos para la manipulación genética de plantas

25 Existe un número de estrategias para manipular genéticamente las plantas (por ejemplo Birch, 1997, *Ann Rev Plant Phys Plant Mol Biol.* 48, 297). Por ejemplo, se pueden diseñar estrategias para incrementar la expresión de un polinucleótido/polipéptido en una célula vegetal, órgano y/o en una etapa de desarrollo particular en la que/cuando se expresa normalmente, o para expresar ectópicamente un polinucleótido/polipéptido en una célula, tejido, órgano y/o en una etapa de desarrollo particular que/cuando no se expresa normalmente. También se pueden diseñar 30 estrategias para incrementar la expresión de un polinucleótido/polipéptido en respuesta a estímulos externos, tales como estímulos medioambientales. Los estímulos medioambientales pueden incluir estreses medioambientales tales como estreses mecánicos (tal como actividad herbívora), deshidratación, salinidad y temperatura. El polinucleótido/polipéptido expresado puede derivar de la especie vegetal a transformar, o puede derivar de una especie vegetal diferente.

35 Se pueden diseñar estrategias de transformación para reducir la expresión de un polinucleótido/polipéptido en una célula vegetal, tejido, órgano o en una etapa particular del desarrollo que/cuando se expresa normalmente, o para reducir la expresión de un polinucleótido/polipéptido en respuesta a estímulos externos. Tales estrategias son conocidas como estrategias de silenciamiento génico.

40 Los constructos genéticos para la expresión de genes en plantas transgénicas incluyen típicamente promotores, tales como polinucleótidos promotores de la invención, para conducir la expresión de uno o más de polinucleótido clonado, terminadores y secuencias marcadoras seleccionables para detectar la presencia del constructo genético en la planta transformada.

45 Los terminadores ejemplares que se usan habitualmente en constructo genético para la transformación vegetal incluyen, por ejemplo, el terminador del virus del mosaico de la coliflor (CaMV) 35S, los terminadores de nopalina sintasa u octopina sintasa de *Agrobacterium tumefaciens*, el terminador del gen zin de *Zea mays*, el terminador de ADP-glucosa pirofosforilasa de *Oryza sativa*, y el terminador PI-II de *Solanum tuberosum*.

50 Los marcadores seleccionables usados habitualmente en la transformación vegetal incluyen el gen de neomicina fosfotransferasa II (NPT II) que confiere resistencia a kanamicina, el gen *aadA*, que confiere resistencia a espectinomicina y estreptomycin, el gen de fosfinotricin acetil transferasa (gen *bar*) para resistencia a Ignite (AgrEvo) y Basta (Hoechst), y el gen de higromicina fosfotransferasa (*hpt*) para resistencia a higromicina.

55 También se contempla el uso de constructos genéticos que comprenden genes informadores (secuencias codificantes que expresan una actividad que es extraña al hospedante, habitualmente una actividad enzimática y/o una señal visible (por ejemplo, luciferasa, GUS, GFP) que se pueden usar para el análisis de la expresión del promotor en plantas y tejidos vegetales. La bibliografía sobre genes informadores se repasa en Herrera-Estrella et al., 1993, *Nature* 303, 209, y Schrott, 1995, en: *Gene Transfer to Plants* (Potrykus, T., Spangenberg. Eds) Springer Verlag. Berlín, p. 325-336.

Las estrategias de silenciamiento génico se pueden centrar sobre el propio gen o sobre elementos reguladores que afectan a la expresión del polipéptido codificado. "Elementos reguladores" se usa aquí en el sentido más amplio

posible, e incluye otros genes que interactúan con el gen de interés.

Los constructos genéticos diseñados para disminuir o silenciar la expresión de un polinucleótido/polipéptido pueden incluir una copia antisentido de un polinucleótido. En tales constructos, el polinucleótido se coloca en una orientación antisentido con respecto al promotor y terminador.

- 5 Un polinucleótido “antisentido” se obtiene invirtiendo un polinucleótido o un fragmento del polinucleótido de manera que el transcrito producido será complementario al transcrito de ARNm del gen, por ejemplo

5' GATCTA 3' (hebra codificante)	3'CTAGAT 5' (hebra antisentido)
3'CUAGAU 5' ARNm	5'GAUCUCG 3' ARN antisentido

- 10 Los constructos genéticos diseñados para el silenciamiento génico también pueden incluir una repetición invertida. Una “repetición invertida” es una secuencia que está repetida, en la que la segunda mitad de la repetición está en la hebra complementaria, por ejemplo

5'-GATCTA..... TAGATC-3'

3'-CTAGAT..... ATCTAG-5'

- 15 El transcrito formado puede sufrir emparejamiento de bases complementarias para formar una estructura de horquilla de pelo. Para permitir la formación de la horquilla de pelo, habitualmente se requiere un espaciador de al menos 3-5 pb entre las regiones repetidas.

Otro enfoque de silenciamiento implica el uso de un ARN antisentido pequeño dirigido al transcrito equivalente a un miARN (Llave et al., 2002, Science 297, 2053). El uso de tal ARN antisentido pequeño que corresponde al polinucleótido de la invención está expresamente contemplado.

- 20 La expresión constructo genético, como se usa aquí, también incluye ARN antisentido pequeños y otros polinucleótidos útiles para efectuar el silenciamiento génico.

- 25 La transformación con un constructo de expresión, como se define aquí, también puede dar como resultado el silenciamiento génico a través de un proceso conocido como supresión de sentido (por ejemplo Napoli et al., 1990, Plant Cell 2, 279; de Carvalho Niebel et al., 1995, Plant Cell, 7, 347). En algunos casos, la supresión de sentido puede implicar la sobreexpresión de toda la secuencia codificante o de una secuencia codificante parcial, pero también puede implicar la expresión de una región no codificante del gen, tal como un intrón o un intrón no traducido de 5' o 3' (UTR). Para silenciar coordinadamente múltiples genes, se pueden usar constructos sentido parciales quiméricos (Abbott et al., 2002, Plant Physiol. 128(3): 844-53; Jones et al., 1998, Planta 204: 499-505). El uso de tales estrategias de supresión de sentido para silenciar la expresión de una secuencia enlazada operativamente al promotor de la invención también está contemplado.

- 30 Los insertos polinucleotídicos en constructos genéticos diseñados para el silenciamiento génico pueden corresponder a secuencia codificante y/o a secuencia no codificante, tal como una secuencia promotora y/o intrónica y/o 5' o 3' UTR, o el gen correspondiente.

Otras estrategias de silenciamiento génico incluyen enfoques dominantes negativos y el uso de constructos ribozímicos (McIntyre, 1996, Transgenic Res, 5, 257).

- 35 El silenciamiento pre-transcripcional se puede provocar a través de la mutación del propio gen o sus elementos reguladores. Tales mutaciones pueden incluir mutaciones de punto, desplazamientos del marco, inserciones, supresiones y sustituciones.

Plantas

- 40 El término “planta” está destinado a incluir una planta completa o cualquier parte de una planta, propágulos y progenie de una planta.

El término “progenie”, como se usa aquí, se refiere a cualquier célula, planta o parte de la misma que se ha obtenido o deriva de una célula o planta transgénica de la presente invención. De este modo, el término progenie incluye, pero no se limita a, semillas, plantas obtenidas de semillas, plantas o partes de las mismas, o derivadas de técnicas de cultivo tisular vegetal, o de clonación.

- 45 El término “propágulo” significa cualquier parte de una planta que se puede usar en reproducción o propagación, ya sea sexual o asexual, incluyendo semillas y esquejes.

Una planta “transgénica” o “transformada” se refiere a una planta que contiene nuevo material genético como

resultado de manipulación o transformación genética. El nuevo material genético puede derivar de una planta de la misma especie como la planta transgénica o transformada resultante, o de una especie diferente. Una planta transformada incluye una planta que está transformada ya sea de forma estable o transitoriamente con nuevo material genético.

- 5 Las plantas de la invención se pueden hacer crecer y se pueden autofecundar o se pueden cruzar con una estirpe vegetal diferente, y se pueden identificar los híbridos resultantes, con las características fenotípicas deseadas. Se pueden hacer crecer dos o más generaciones. Las plantas que resultan de tales enfoques de reproducción estándar también forman parte de la presente invención.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

- 10 Otros aspectos de la presente invención serán manifiestos a partir de la siguiente descripción que se da a título de ejemplo solamente y con referencia a los dibujos que se acompañan, en los que:

La Figura 1 muestra la ruta general de taninos condensados;

La Figura 2(A) ilustra la secuencia de ADNc que representa la secuencia de ADNc de longitud completa de TaMYB14, clonado a partir de tejido de hoja madura de *T. arvense*.

- 15 La Figura 2(B) ilustra la traducción de aminoácidos de TaMYB14.

La Figura 3 muestra los niveles de transcrito de TaMYB14 en tejidos variados de la especie *Trifolium* y variedades de cultivo que se hicieron crecer en condiciones de invernadero idénticas. Línea 1, (escalera); Línea 2, librería de ADNc de hoja madura de *T. repens* (variedad de cultivo Huia); Línea 3, librería de ADNc de raíz madura de *T. repens* (variedad de cultivo Huia); Línea 4, librería de ADNc de estolón maduro de *T. repens* (variedad de cultivo Huia); Línea 5, librería de ADNc floral madura de *T. repens* (variedad de cultivo DC111); Línea 6, ADNc de hoja emergente de *T. repens* (variedad de cultivo Huia); Línea 7, ADNc de hoja madura de *T. repens* (variedad de cultivo Isabelle rica en antocianinas); Línea 8, ADNc de hoja inmadura de *T. arvense* (variedad de cultivo AZ2925); Línea 9, ADNc de hoja madura de *T. arvense* (variedad de cultivo AZ2925); Línea 10, ADNc floral de meristemo de *T. repens* (variedad de cultivo Huia); Línea 11, ADNc de hoja de meristemo de *T. repens* (variedad de cultivo Huia); Línea 12, ADNc solamente de tricoma de meristemo de *T. repens* (variedad de cultivo Huia); Línea 13, librería de ADNc de planta (hoja, raíz y estolón) madura de *T. occidentale* (variedad de cultivo Huia); Línea 14, librería de ADNc de nodo maduro de *T. repens* (variedad de cultivo Huia); Línea 15, clon de MYB14cDNA de *T. arvense* clonado en TOPO, Línea 16, clon genómico de MYB14 de *T. arvense* clonado en TOPO, Línea 17, ADN genómico de *T. occidentale*; Línea 17, ADN genómico de *T. repens*; Línea 17, ADN genómico de *T. arvense*; Línea 20, (escalera).

La Figura 4 muestra los niveles de transcrito de BANYULS (A) y LAR (B) en tejidos variados de la especie *Trifolium* y variedades de cultivo que se hicieron crecer en condiciones de invernadero idénticas. Línea 1, (escalera); Línea 2, librería de ADNc de hoja madura de *T. repens* (variedad de cultivo Huia); Línea 3, librería de ADNc de raíz madura de *T. repens* (variedad de cultivo Huia); Línea 4, librería de ADNc de estolón maduro de *T. repens* (variedad de cultivo Huia); Línea 5, librería de ADNc floral madura de *T. repens* (variedad de cultivo DC111); Línea 6, ADNc de hoja emergente de *T. repens* (variedad de cultivo Huia); Línea 7, ADNc de hoja madura de *T. repens* (variedad de cultivo Isabelle rica en antocianinas); Línea 8, ADNc de hoja inmadura de *T. arvense* (variedad de cultivo AZ2925); Línea 9, ADNc de hoja madura de *T. arvense* (variedad de cultivo AZ2925); Línea 10, ADNc floral de meristemo de *T. repens* (variedad de cultivo Huia); Línea 11, ADNc de hoja de meristemo de *T. repens* (variedad de cultivo Huia); Línea 12, ADNc solamente de tricoma de meristemo de *T. repens* (variedad de cultivo Huia); Línea 13, librería de ADNc de planta (hoja, raíz y estolón) madura de *T. occidentale* (variedad de cultivo Huia); Línea 14, librería de ADNc de nodo maduro de *T. repens* (variedad de cultivo Huia); Línea 15, clon de BAN o LAR de ADNc de *T. arvense* clonado en TOPO, Línea 16, clon genómico de BAN o LAR de *T. arvense* clonado en TOPO, Línea 17, ADN genómico de *T. occidentale*; Línea 17, ADN genómico de *T. repens*; Línea 17, ADN genómico de *T. arvense*; Línea 20, (escalera).

La Figura 5 muestra los resultados de la tinción con DMACA de tejido de hoja madura de trébol blanco transformado. La tinción con DMACA (color gris claro/oscura) de tejido de hoja de trébol blanco madura que identifica taninos condensados en (A) tipo salvaje y (B) transformado con el gen TaMYB14.

- 50 La Figura 6 muestra el vector plasmídico M14AphZBarP, usado para la transformación vegetal. E1, E2 y E3 indican los 3 exones del alelo genómico TaMYB14-1.

La Figura 7 muestra el alineamiento de las secuencias de ADNc de longitud completa de MYB14 de *Trifolium*, los aciertos de BLASTN principales y *AtTT2* con similitudes resaltados en gris claro.

- 55 La Figura 8 muestra el alineamiento de los marcos de lectura abiertos traducidos de TaMYB14 de *Trifolium arvense*, los aciertos de BLASTP principales y *AtTT2* con similitudes resaltados en gris claro, y los motivos en cajas.

La Figura 9 muestra el alineamiento de las secuencias proteicas de longitud completa de TaMYB14 (TaMYB14FTa expresado y TaMYB14-2S silencioso), alelo de ToMYB14, y alelos de TrMYB14 con las diferencias resaltadas en regiones de gris oscuro/blanco, y las áreas de supresión/inserción resaltadas en cajas.

5 La Figura 10 muestra el alineamiento de las secuencias de ADN genómico de longitud completa de los alelos de TrMYB14 de *Trifolium repens* (TRM*) alineados con los alelos de TaMYB14 de *Trifolium arvense* (TaM3, TaM4), resaltándose las diferencias en exones (gris claro) e intrones (gris oscuro).

10 La Figura 11 muestra el alineamiento de las secuencias de ADN genómico de longitud completa de los alelos de ToMYB14 de *Trifolium occidentale* (To1, To6) alineados con los alelos de TaMYB14 de *Trifolium arvense* (TaM3, TaM4), resaltándose las diferencias en exones (gris claro) e intrones (gris oscuro).

La Figura 12 muestra el alineamiento de las secuencias de ADN genómico de longitud completa de los alelos de TaMYB14 de *Trifolium arvense* (Ta*) y alelos de TafMYB14 de *Trifolium affine* (Taf*), mostrando diferencias los exones (gris claro) e intrones (gris oscuro).

15 La Figura 13 muestra el mapa según Vector NTI del constructo pHZbarSMYB, que contiene el fragmento NotI de MYB14pHANNIBAL, que contiene un segmento de ADNc de TaMYB14 de *T. arvense* en orientación de sentido (SMYB14F) y antisentido (SMYB14R) que flanquea al intrón pdk.

20 La Figura 14 muestra la reacción de PCR para la presencia de M14ApHZBAR de ADN genómico aislado de trébol blanco transformado putativamente. Líneas; A1, B1 escalera; A2-18 y B2-B15 tréboles transformados, B16 trébol blanco no transformado, B17 control plasmídico, B18 control de agua. Los cebadores fueron 35S (promotor) y PMYBR (hasta el extremo 3' del gen) que amplifican un fragmento de 1.244 pb.

La Figura 15 muestra los resultados del cribado mediante DMACA de hojas de *T. repens* de tipo salvaje (A) y transgénicas (B a D), transformadas con el constructo TaMYB14.

La Figura 16 muestra microscopía mediante aceite de tricomas (E-G), células epidérmicas (H) y célula mesófila (I-K) de folíolos transgénicos teñidos con DMACA que expresan el gen TaMyb14A.

25 La Figura 17 muestra monómeros de extracto de semillas de uva – se muestran debajo los cromatogramas SRM de los monómeros en un extracto de semillas de uva. La señal A es una suma de los iones del producto 123, 139 y 165 *m/z* del SRM de 291,3 *m/z* (catequina (C) y epicatequina (EC)). La señal B es una suma de los iones de producto 139 y 151 *m/z* del SRM de 307,3 *m/z* (galocatequina (GC) y epigalocatequina (EGC)).

30 La Figura 18 muestra dímeros y trímeros de extracto de semillas de uva. Se muestran debajo los cromatogramas SRM de los dímeros y trímeros en un extracto de semillas de uva. La señal A es una suma de los iones del producto 291, 409 y 427 *m/z* del SRM de 579,3 *m/z* (dímero PC:PC). La señal B es una suma de los iones de producto 291, 307, 427 y 443 *m/z* del SRM de 595,3 *m/z* (dímero PC:PD). La señal C es una suma de los iones de producto 291, 577 y 579 *m/z* del SRM de 867,3 *m/z* (trímero 3PC). Se proporcionan los espectros MS2 de un dímero PC:PC, un dímero PC:PD, y dos trímeros 3PC como prueba de la identificación de estos metabolitos.

35

40 La Figura 19 muestra los cromatogramas SRM de monómeros para plantas de control (trébol blanco –ve) y transgénicas (trébol blanco +ve) que expresan MYB14. La señal A es una suma de los iones del producto 123, 139 y 165 *m/z* del SRM de 291,3 *m/z* (PC; catequina y epicatequina). La señal B es una suma de los iones de producto 139 y 151 *m/z* del SRM de 307,3 *m/z* (PD; galocatequina y epigalocatequina). Las escalas del cromatograma se fijan para mostrar el aspecto de monómeros en la planta modificada. No se detectaron monómeros en la planta de control. Los espectros MS2 de epicatequina (EC) y epigalocatequina (EGC) se proporcionan a partir de la planta modificada como prueba de la identificación de estos metabolitos.

45 La Figura 20 muestra los cromatogramas SRM de monómeros para plantas de control (trébol blanco –ve) y transgénicas (trébol blanco +ve) que expresan MYB14. La señal A es una suma de los iones del producto 291, 409 y 427 *m/z* del SRM de 579,3 *m/z* (dímero PC:PC). La señal B es una suma de los iones de producto 291, 307, 427 y 443 *m/z* del SRM de 595,3 *m/z* (dímero PC:PD). La señal C es una suma de los iones de producto 307 y 443 *m/z* del SRM de 611,3 *m/z* (dímero PD:PD). Las escalas del cromatograma se fijan para mostrar el aspecto de dímeros en la planta modificada. No se detectaron dímeros en la planta de control. Los espectros MS2 de tres dímeros PD:PD (1-3) y un dímero mixto PC:PD (4) se proporcionan a partir de la planta modificada como prueba de la identificación de esos metabolitos.

50

55 La Figura 21 muestra los cromatogramas SRM de trímeros para plantas de control (trébol blanco –ve) y transgénicas (trébol blanco +ve) que expresan MYB14. La señal A es una suma de los iones del producto 291, 577 y 579 *m/z* del SRM de 867,3 *m/z* (trímero 3PC). La señal B es una suma de los iones de producto 291, 307, 427, 443, 577, 579, 593, 595 y 757 *m/z* del SRM de 883,3 *m/z* (dímero PC:PD). La señal C es una suma de los iones de producto 291, 307, 443, 593, 595, 611, 731, 757 773 *m/z* del SRM de 899,3 *m/z* (trímero 1PC:2PD). La señal D es una suma de los iones de producto 307, 443, 609, 611, 747, 773 y 789 *m/z*

del SRM de 915,3 m/z (trímero 3PD). Las escalas del cromatograma se fijan para mostrar el aspecto de trímeros en la planta modificada. No se detectaron trímeros en la planta de control. Los espectros MS2 de un trímero 3PD y un trímero mixto 1PC:2PD se proporcionan a partir de la planta modificada como prueba de la identificación de esos metabolitos.

5 La Figura 22 muestra la reacción de PCR para la presencia de M14ApHZBAR a partir de ADN genómico aislado de plántulas de tabaco putativamente transformadas. Líneas A1, escalera; A2-10 tabaco transformado, A13, 14, controles de tabaco, A15 control plasmídico. Los cebadores fueron 35S (promotor) y PMYBR (hasta el extremo 3' del gen) que amplifican un fragmento de 1.244 pb.

10 La Figura 23 muestra los resultados del cribado mediante DMACA de hojas de tabaco (*Nicotiana tabacum*) transgénico (A a G), transformadas con el constructo M14ApHZBAR.

15 La Figura 24 muestra los cromatogramas SRM para plantas de control (tipo salvaje) y modificadas (transgénicas) que expresan MYB14. La señal A es una suma de los iones del producto 123, 139 y 165 m/z del SRM de 291,3 m/z (PC; catequina y epicatequina). La señal B es una suma de los iones de producto 139 y 151 m/z del SRM de 307,3 m/z (PD; galocatequina y epigalocatequina). La señal C es una suma de los iones de producto 291, 409 y 427 m/z del SRM de 579,3 m/z (dímero PC:PC). La señal D es una suma de los iones de producto 291, 577 y 579 m/z del SRM de 867,3 m/z (trímero PC:PC:PC). Las escalas del cromatograma se fijan para mostrar el aspecto de monómeros, dímeros y trímeros en la planta modificada. No se detectaron monómeros en la planta de control. Obsérvese que no se detectaron dímeros o trímeros mixtos PC:PD o 100% PD.

20 La Figura 25 muestra los espectros MS2 de epicatequina (EC), galocatequina (GC), epigalocatequina (EGC), dímero PC:PC 1 y 2, y el trímero PC:PC:PC proporcionados de las plantas modificadas (transgénicas) que expresan MYB14, como prueba de la identificación de estos metabolitos.

25 La Figura 26 muestra la reacción de PCR para la presencia de M14pHANNIBAL en ADN genómico aislado de *T. arvense* putativamente transformado. Líneas; A1 vector de control negativo pHANNIBAL, A2 M14ApHZBAR que contiene 35S y constructo de gen genómico – control que amplifica un fragmento de 1.244 pb; A3 M14pHANNIBAL control plasmídico positivo que contiene el constructo de ARNhp, A4 pHANNIBAL que contiene el fragmento MYB en orientación antisentido en dirección 5' del terminador ocs (control negativo), A5 pHZBARSMYB control plasmídico positivo, A6 escalera, A7-18 *T. arvense* transformado, A19 ADN genómico de *T. arvense* de tipo salvaje, A20 control de agua.

30 B: B1 escalera, B2-B11 *T. arvense* transformado, B12 M14pHANNIBAL control plasmídico positivo. Los cebadores fueron 35S (promotor) y PHMYBR (hasta el extremo 3' del gen) que amplifican un fragmento de 393 pb.

35 La Figura 27 muestra los resultados del cribado mediante DMACA de callo (A) y plántulas (B a D) de *T. arvense* de tipo salvaje regenerados en medios de cultivo tisular. No se produce tinción mediante DMACA en callo y cribado mediante DMACA de plántulas de *T. arvense* transgénicas (E a L) regeneradas en medios de cultivo tisular. La tinción está enormemente disminuida en comparación con plantas de tipo salvaje.

40 La Figura 28 muestra los cuatro cromatogramas SRM de monómeros para plantas de *T. arvense* de control y genosuprimidas: La señal A es una suma de los iones del producto 123, 139 y 165 m/z del SRM de 291,3 m/z (PC; catequina y epicatequina) para una planta de control. B es una suma de los iones de producto 123, 139 y 165 m/z del SRM de 291,3 m/z (PC; catequina y epicatequina) para una planta genosuprimida. C es una suma de los iones de producto 139 y 151 m/z del SRM de 307,3 m/z (PD; galocatequina y epigalocatequina) para una planta de control. D es una suma de los iones de producto 139 y 151 m/z del SRM de 307,3 m/z (PD; galocatequina y epigalocatequina) para una planta genosuprimida. Los espectros MS2 se proporcionan de la planta de control como prueba de catequina y galocatequina en la planta de control. Las escalas de los cromatogramas para las señales A, B, C y D se han fijado para mostrar la desaparición de catequina y galocatequina en la planta genosuprimida.

45 La Figura 29 muestra los cromatogramas SRM de dímeros para plantas de *T. arvense* de control y genosuprimidas. La señal A es una suma de los iones del producto 291 y 427 m/z del SRM de 579,3 m/z (dímero PC:PC). La señal B es una suma de los iones de producto 307, 427 y 443 m/z del SRM de 595,3 m/z (dímero PC:PD). La señal C es una suma de los iones de producto 307 y 443 m/z del SRM de 611,3 m/z (dímero PD:PD). Las escalas de los cromatogramas se fijan para mostrar la desaparición de dímeros en la planta genosuprimida. Los espectros MS2 se proporcionan de la planta de control como prueba de los tres tipos de dímeros en el control.

50 La Figura 30 muestra el análisis de PCR para la presencia de pTaMyb14A de ADN genómico aislado de alfalfa putativamente transformada. Líneas L; escalera; 1-3, no transformada, 4-10 transformada, 11 tipo salvaje, 12 control de agua, 13 control plasmídico. Los cebadores fueron 35S y PMYBR (hasta el extremo 3' del gen).

55

La Figura 31 muestra el análisis de PCR para la presencia de M14ApHZBAR de ADN genómico aislado de plántulas de crucifera putativamente transformadas. Línea 8, control de crucifera; Línea 18 escalera; Línea 1-7 y 9-17 crucifera transformada. Los cebadores fueron 35S (promotor) y PMYBR (hasta el extremo 3' del gen) que amplifican un fragmento de 1.244 pb.

5 La Figura 32 muestra los resultados del cribado mediante DMACA de crucifera de tipo salvaje (*Brassica oleracea*) (A) y hojas transgénicas (B a D), transformadas con el constructo M14ApHZBARP.

10 La Figura 33 muestra los cromatogramas SRM de los iones de producto 123, 139 y 165 *m/z* del SRM de 291,3 *m/z* (catequina (C) y epicatequina (EC)) en dos controles y una crucifera transgénica que expresa MYB14. Los espectros MS2 de la epicatequina detectada en el control verde y la muestra +ve transgénica se proporcionan como prueba de identificación de estos metabolitos. No se detectó epicatequina en la muestra de control rojo.

La Figura 34 muestra un alineamiento de todas las secuencias de la proteína MYB14 de *Trifolium* identificadas por el solicitante.

La Figura 35 muestra el porcentaje de identidad entre las secuencias alineadas en la Figura 34.

15 BREVE DESCRIPCIÓN DEL LISTADO DE SECUENCIAS

SEC ID NO:	Descripción	Secuencia correspondiente
1	Polinucleótido, <i>Trifolium arvense</i> , ADNc de TaMYB14-1	Secuencia de ADNc de Ta MYB14 de gen expresado
2	Polinucleótido, <i>Trifolium arvense</i> , ADNg de TaMYB14-1	Secuencia genómica de Ta MYB14 1 del alelo 1 de <i>Trifolium arvense</i> .
3	Polinucleótido, <i>Trifolium arvense</i> , ADNg de TaMYB14-2	Secuencia genómica de Ta MYB14 2 del alelo 2 de <i>Trifolium arvense</i> .
4	Polinucleótido, <i>Trifolium affine</i> , ADNg de TafMYB14-1	Secuencia genómica de Taf MYB14 1 del alelo 1 de <i>Trifolium affine</i> .
5	Polinucleótido, <i>Trifolium affine</i> , ADNc de TafMYB14-1	Secuencia de ADNc de Taf MYB14 del gen expresado
6	Polinucleótido, <i>Trifolium affine</i> , ADNg de TafMYB14-2	Secuencia genómica de Taf MYB14 2 del alelo 2 de <i>Trifolium affine</i> .
7	Polinucleótido, <i>Trifolium occidentale</i> , ADNg de ToMYB14-1	Secuencia genómica de ToMYB14 1 del alelo 1 e <i>Trifolium occidentale</i> .
8	Polinucleótido, <i>Trifolium occidentale</i> , ADNg de ToMYB14-2	Secuencia genómica de ToMYB14 2 del alelo 2 de <i>Trifolium occidentale</i> .
9	Polinucleótido, <i>Trifolium repens</i> , ADNg de TrMYB14-1	Secuencia genómica de TrMYB14 1 del alelo 1 de <i>Trifolium repens</i> .
10	Polinucleótido, <i>Trifolium repens</i> , ADNg de TrMYB14-2	Secuencia genómica de TrMYB14 2 desde el alelo 2 de <i>Trifolium repens</i> .
11	Polinucleótido, <i>Trifolium repens</i> , ADNg de TrMYB14-3	Secuencia genómica de TrMYB14 3 del alelo 3 de <i>Trifolium repens</i> .
12	Polinucleótido, <i>Trifolium repens</i> , ADNg de TrMYB14-4	Secuencia genómica de TrMYB14 4 del alelo 4 de <i>Trifolium repens</i> .
13	Polinucleótido, <i>Trifolium arvense</i> , ADNc de TaMYB14-1	Secuencia de ADNc que representa la secuencia de ADNc de longitud completa de TaMYB14
14	Polipéptido, <i>Trifolium arvense</i> , TaMYB14-1	traducción de aminoácidos de of TaMYB14
15	Polipéptido, artificial, consenso	motivo similar al motivo de subgrupo 5 (Stracke et al., 2001) común a activadores de MYB de CT conocidos

ES 2 537 342 T3

SEC ID NO:	Descripción	Secuencia correspondiente
16	Polipéptido, artificial, consenso	motivo común a activadores de MYB de antocianinas conocidos (Motivo de subgrupo 6, Stracke et al., 2001)
17	Polipéptido, artificial, consenso	nuevo motivo de MYB de TFs de MYB14
18	Polinucleótido, artificial, cebador	Búsqueda del dominio de MYB - MYBFX
19	Polinucleótido, artificial, cebador	Búsqueda del dominio de MYB - MYBFY
20	Polinucleótido, artificial, cebador	Búsqueda del dominio de MYB - MYBFZ
21	Polinucleótido, artificial, cebador	Aislamiento de longitud completa - M14ATG
22	Polinucleótido, artificial, cebador	Aislamiento de longitud completa - M14TGA
23	Polinucleótido, artificial, cebador	Paseo génico - M14TSP1
24	Polinucleótido, artificial, cebador	Paseo génico - M14TSP2
25	Polinucleótido, artificial, cebador	Paseo génico - M14TSP3
26	Polinucleótido, artificial, cebador	Clonación en el vector- M14FATG
27	Polinucleótido, artificial, cebador	Lotus corniculatus - MYBLF
28	Polinucleótido, artificial, cebador	Lotus corniculatus - MYBLR
29	Polinucleótido, artificial, cebador	Extremo 5' UTR de MYB14 - MYB148N
30	Polinucleótido, artificial, cebador	Extremo 3' UTR de MYB14 - MYB14RR
31	Polinucleótido, artificial, cebador	Cebador para intrón 1 - 15
32	Polinucleótido, artificial, cebador	Cebador para intrón 1 - 13
33	Polinucleótido, artificial, cebador	Paseo génico - TSP4
34	Polinucleótido, artificial, cebador	Paseo génico - TSP5
35	Polinucleótido, artificial, cebador	Sitio de partida de 5' directo - MYB148F
36	Polinucleótido, artificial, cebador	Sitio de partida de 5' inverso - MYB14RR
37	Polinucleótido, artificial, cebador	Análisis de expresión/ vector de slienciamiento - MYB14F
38	Polinucleótido, artificial, cebador	Análisis de expresión/ vector de slienciamiento - MYB14R
39	Polinucleótido, artificial, cebador	Paseo génico - MYB14R2
40	Polinucleótido, artificial, cebador	Paseo génico - MYB14R3
41	Polinucleótido, artificial, cebador	Secuenciación - M13 Directo
42	Polinucleótido, artificial, cebador	Secuenciación - M13 Inverso
43	Polinucleótido, artificial, cebador	Producción de ADNc - BD SMART II™ A Oligonucleotide
44	Polinucleótido, artificial, cebador	Producción de ADNc - 3' BD SMAR™ CDS Primer II A
45	Polinucleótido, artificial, cebador	Amplificación de ARNm - 5' PCR Primer II A
46	Polipéptido, <i>Trifolium arvense</i> , TaMYB14-2	

SEC ID NO:	Descripción	Secuencia correspondiente
47	Polipéptido, <i>Trifolium affine</i> , TafMYB14-1	
48	Polipéptido, <i>Trifolium affine</i> , TafMYB14-2	
49	Polipéptido, <i>Trifolium occidentale</i> , ToMYB14-1	
50	Polinucleótido, <i>Trifolium occidentale</i> , ToMYB14-2	
51	Polipéptido, <i>Trifolium repens</i> , TrMYB14-1	
52	Polipéptido, <i>Trifolium repens</i> , TrMYB14-2	
53	Polipéptido, <i>Trifolium repens</i> , TrMYB14-3	
54	Polipéptido, <i>Trifolium repens</i> , TrMYB14-4	
55	Polinucleótido, <i>Trifolium arvense</i> , ADNc/ORF de TaMYB14-1	
56	Polinucleótido, <i>Trifolium arvense</i> , ADNc/ORF de TaMYB14-2	
57	Polinucleótido, <i>Trifolium affine</i> , ADNc/ORF de TafMYB14-1	
58	Polinucleótido, <i>Trifolium affine</i> , ADNc/ORF de TafMYB14-2	
59	Polinucleótido, <i>Trifolium occidentale</i> , ADNc/ORF de ToMYB14-1	
60	Polinucleótido, <i>Trifolium occidentale</i> , ADNc/ORF de ToMYB14-2	
61	Polinucleótido, <i>Trifolium repens</i> , ADNc/ORF de TrMYB14-1	
62	Polinucleótido, <i>Trifolium repens</i> , ADNc/ORF de TrMYB14-2	
63	Polinucleótido, <i>Trifolium repens</i> , ADNc/ORF de TrMYB14-3	
64	Polinucleótido, <i>Trifolium repens</i> , ADNc/ORF de TrMYB14-4	
65	Polinucleótido, <i>Trifolium arvense</i> , secuencia silenciamiento	
66	Polinucleótido, <i>artificial, cebador</i> , MYB F1	
67	Polinucleótido, <i>artificial, cebador</i> , MYB R	
68	Polinucleótido, <i>artificial, cebador</i> , MYB F	
69	Polinucleótido, <i>artificial, cebador</i> , MYB R1	

La invención se ilustrará ahora con referencia a los siguientes ejemplos no limitantes.

Ejemplo 1: Identificación de los genes/ácidos nucleicos/proteínas de MYB14 de la invención, y análisis de los perfiles de expresión.

5 Introducción

Usando cebadores diseñados para el dominio de MYB de especie de legumbre, el solicitante ha amplificado secuencias que codifican nuevos factores de transcripción (TFs) de MYB mediante PCR de ADNc y ADN genómico

(ADNg) aislados de un abanico de especies de *Trifolium*. Estas especies difieren en su capacidad para acumular CTs en tejido de hoja madura. Debido a que el trébol blanco no expresa genes de CT en tejido de hoja, se usó una estrategia alternativa que permitió el aislamiento del TF de MYB expresado de especies de *Trifolium* estrechamente relacionadas (*T. arvense*; *T. affine*) que sí acumulan CTs en todas las células de tejido foliar a lo largo de la vida de la hoja. Esto se logró investigando los patrones de expresión diferencial de TFs de MYB en diversos tipos de hoja de *Trifolium*; a saber, (a) en tejido de hoja de trébol blanco (*T. repens*), en el que la expresión del gen de CT está restringida a los tricomas de la hoja durante el desarrollo meristemático antes de la emergencia de la hoja; (b) en las especies estrechamente relacionadas (*T. arvense*), en las que la expresión del gen de CT se encuentra en la mayoría de las células de la hoja durante todo el tiempo de vida (excepto los pelos del tricoma); (c) en el tejido de hoja madura de trébol, en el que la biosíntesis de CTs ya ha cesado. Tal expresión temporal y espacial específica requiere la regulación diferencial por diferentes TFs de MYB específicos para la ruta de la rama de CTs. La comparación de los TFs de MYB a partir de cada tipo de hoja eliminó factores de MYB comunes que tienen funciones distintas en la biosíntesis de CTs. El análisis de los TFs de MYB aislados que quedan permitió la identificación de aquellos que son únicos para tejidos que acumulan CTs.

La secuenciación de productos de la PCR dio como resultado la identificación de TFs de MYB previamente no identificados a partir de un número de especies de *Trifolium*. La secuenciación de longitud completa de estos genes de MYB reveló un código proteico muy desigual cuando se compara con la secuencia de *AtTT2* publicada (NP_198405), incluyendo la presencia de varias supresiones e inserciones de bases en los genes de las diferentes especies de *Trifolium* (Figuras 7 y 8). La traducción de la secuencia de ADNc reveló que la proteína codificada por este TF de MYB también tiene un número sustancial de supresiones, inserciones e intercambios de aminoácidos (Figura 9). Se ha diseñado este gen *TaMYB14*. El análisis de las secuencias de ADNc de longitud completa procedentes de 2 especies diferentes de *Trifolium* reveló la presencia de tres exones y dos intrones de tamaños variables en todas las isoformas/alelos de *TaMYB14* (Figuras 10-12).

Se hicieron crecer en un invernadero semillas procedentes de un número de accesos que representan diversos genotipos de cuatro especies de *Trifolium*, respectivamente, y se determinó la presencia o ausencia de CTs en hojas usando tinción con DMACA. Se diseñaron cebadores específicos para *TaMYB14*, y los niveles de transcrito en diversos tejidos se determinaron mediante PCR. La expresión de *TaMYB14* estaba correlacionada con la acumulación de CTs en tejidos de hoja. Su expresión fue indetectable en tejidos libres de CTs. *TaMYB14* se expresó de forma muy importante en tejidos que acumulan activamente CTs, y coincidió con la expresión detectable de las dos enzimas específicamente implicadas en la biosíntesis de CTs, a saber, ANR y LAR.

La transformación y sobreexpresión de *TaMYB14* en trébol blanco (véase el Ejemplo 2) dio como resultado mayores niveles de CTs en tejidos habitualmente desprovistos de CTs. Esto muestra que la expresión de *TaMYB14* es crítica para la acumulación de CTs. La sobreexpresión de *TaMYB14* en *T. repens* por medio de transgénesis permitirá por lo tanto la acumulación de niveles significativos de CTs en tejidos foliares de diversas especies vegetales, proporcionando de ese modo los medios para mejorar la calidad de la pastura para el ganado.

Materiales y métodos

Material vegetal y análisis de niveles de taninos condensados

Se hicieron crecer en invernaderos semillas de varias variedades de cultivo de cuatro especies de legumbre que difieren en sus niveles de CT foliar. *Trifolium repens* (Huia); *T. arvense* (AZ2925; AZ4755; AZ1353); *T. affine* (AZ925), y *T. occidentale* (AZ4270). El material vegetal de diversas edades y tipos se cosechó, y el material se congeló inmediatamente en nitrógeno líquido y se trituró y se usó subsiguientemente para el aislamiento de ADN o ARN.

Tinción del material vegetal con DMACA

Los CTs se analizaron histoquímicamente usando el método de DMACA (4-dimetilamino-cinamaldehído) acidificado, esencialmente como se describe por Li et al. (1996). Este método usa el reactivo de DMACA (p-dimetilaminocinamaldehído) como una tinción histoquímica rápida que permite cribar específicamente material vegetal en busca de la acumulación muy baja de CTs. El protocolo de DMACA-HCl es muy específico para proantocianidinas. Este método se usó preferentemente con respecto al ensayo de vainillina, ya que las antocianinas interfieren seriamente con el ensayo de vainillina. Tejidos de diversas edades se muestrearon y se evaluaron.

Métodos de selección de candidatos de MYB R2R3

Se usaron dos métodos para identificar secuencias de legumbre que contienen un dominio de unión a ADN de MYB R2R3: los modelos ocultos de Markov (HMMs) y perfiles. Ambos métodos dependen de crear en primer lugar un "modelo" del dominio a partir de secuencias proteicas conocidas del dominio de unión a ADN de MYB R2R3, que entonces se usa como la base de la investigación. Los modelos de HMM y de perfil se crearon usando dominios vegetales de MYB R2R3 conocidos como se indica en la Tabla 1 a continuación. Éstos se tomaron de la Figura 2 en Miyake et. al. (2003) y de la Figura 4C en Nesi et. al. (2001; se excluyó la secuencia de MYB humana en esta figura). La distribución de especies de las secuencias usadas para construir el modelo fue la siguiente:

Fuente	Especies	Recuento de dominios
Miyake <i>et. al.</i> (2003)	<i>Lotus japonicus</i>	3
	<i>Glycine max</i>	1
Nesi <i>et. al.</i> (2001)	<i>Arabidopsis thaliana</i>	10
	<i>Zea mays</i>	3
	<i>Hordeum vulgare</i> subsp. <i>vulgare</i>	2
	<i>Oryza sativa</i>	1
	<i>Petunia x hybrida</i>	1
	<i>Picea mariana</i>	1

TABLA 1. Dominios vegetales de MYB R2R3 tomados de Miyake *et. al.* (2003) y Nesi *et. al.* (2001)

Los conjuntos de secuencias de legumbre buscados se dan en la Tabla 2 a continuación. Antes de la búsqueda, todos los conjuntos de EST y cóntigos de EST se tradujeron en seis marcos para generar secuencias proteicas adecuadas para los análisis de HMM/perfil. Las secuencias proteicas de *M. truncatula* se usaron tal cual (estas son predicciones génicas de FGENESH obtenidas de TIGR).

Se usó el programa hmmbuild de HMMER para crear un HMM a partir de los dominios de unión a ADN modelo, y se inspeccionó frente a los conjuntos de secuencias de legumbre usando el programa hmmsearch de HMMER (valor E de corte = 0,01). Se usó el programa prophecy de EMBOSS para crear un perfil a partir de los mismos dominios, y también se inspeccionó frente a las secuencias de legumbre usando el programa profit de EMBOSS (corte de puntuación = 50). En la Tabla 2 a continuación se dan los números de aciertos identificados por cada método en cada conjunto de secuencias:

Conjunto de secuencias	Número total de secuencias	Número de aciertos – Método de perfil	Número de aciertos – Método de HMM	Número de aciertos que se hacen pasar a análisis filogenético
Cóntigos de EST de trébol blanco (CS35)	17.758	18	24	17
NR de PG de trébol blanco	159.017	0	9	3
Cóntigos de EST de trébol rojo	38.099	1	2	0
Cóntigo de EST de Lotus	28.460	5	9	4
Cóntigo de EST de haba de soja	63.676	15	40	15
Proteínas predichas de <i>Medicago truncatula</i>	41.315	60	80	69
ESTs de tricoma glandular <i>Medicago sativa</i>	5.647	1	2	1
Total	353.972	100	166	109

TABLA 2: Conjuntos de secuencias de legumbre investigados

El método de HMM pareció ser más sensible que el método del perfil, identificando todos los aciertos del perfil así como muchos aciertos adicionales. Por esta razón, el método de HMM se seleccionó como el método de elección – las proteínas de los aciertos de HMM se usaron para generar los alineamientos y se hicieron pasar al análisis filogenético. Los aciertos del perfil todavía son bastante útiles: el método de perfil es más restrictivo, y por lo tanto hay una mayor probabilidad de que los candidatos del perfil representen aciertos verdaderos.

Generación de alineamientos

Las secuencias de los dominios de unión a ADN se extrajeron de los 166 candidatos de MYB R2R3 de legumbre identificados anteriormente. Los dominios proteicos se alinearon usando el programa de alineamiento hmalign de HMMER, el cual alinea los dominios usando información en el modelo de HMM original. Los alineamientos

nucleotídicos se generaron superponiendo las secuencias nucleotídicas correspondientes sobre los alineamientos proteicos, preservando de ese modo la estructura de los alineamientos al nivel proteico. Esto se realizó para obtener un alineamiento más exacto que representa mejor la estructura de los dominios.

Análisis filogenético

- 5 Se llevó a cabo un análisis filogenético sobre los dominios de unión a ADN MYB R2R3 vegetales, para ver si los nodos de árbol resultantes se podrían usar para identificar subtipos de MYB R2R3, relacionados con factores de transcripción TT2. Se extrajeron 109 dominios de unión a ADN de longitud completa a partir de los 166 candidatos de MYB R2R3 de legumbre identificados en este estudio, y se combinaron con los genes de MYB R2R3 conocidos de Nesi et. al. (2001) y Miyake et. al. (2003), dando en total 130 dominios de unión a ADN. Se generó un alineamiento proteico de estos 130 dominios usando hmalign, y las secuencias de los dominios nucleotídicos correspondientes se alinearon basándose en esto. El alineamiento nucleotídico se sometió a un análisis de probabilidad máxima para generar un árbol filogenético basado en 100 réplicas de bootstrap, usando los programas fastDNAmI y el programa consensus de Phylip para generar el árbol de consenso. Esta información se usó para diseñar tres cebadores para el dominio de MYBR2R3 de legumbre.

- 15 Aislamiento de ADN y ARN, y síntesis de ADNc

ADN genómico se aisló de tejidos vegetales recientes o congelados (100 mg) usando DNeasy[®] Plant Mini kit (Qiagen) siguiendo las instrucciones del fabricante. Las preparaciones de ADN se trataron con RNase H (Sigma) para eliminar el ARN de las muestras. El ARN total se aisló de tejidos recientes o congelados usando RNeasy[®] Plant Mini kit (Qiagen). El ARN total aislado (100 µg) se trató con DNase I libre de RNase para eliminar el ADN de las muestras durante el aislamiento, siguiendo las instrucciones del fabricante. La concentración y pureza de las muestras de ADN y de ARN se evaluaron determinando la relación de absorbancia a 260 y 280 nm usando un espectrofotómetro NanoDrop ND-100. El ARN total (1 µg) se transcribió de forma inversa en ADNc usando SMART[™] cDNA Synthesis Kit (Clontech) usando los oligonucleótidos SMART[™] CDS primer IIA y SMART II[™] A, siguiendo las instrucciones del fabricante.

- 25 Reacción en cadena de la polimerasa (PCR) y clonación TOPO de los productos de PCR

Se llevaron a cabo reacciones estándar de PCR en un Thermal Cycler (Applied Biosystems), usando como molde una cantidad de aproximadamente 5 ng de ADN o 1 µl de ADNc. Las condiciones del ciclo térmico fueron las siguientes: reacción inicial a 94°C durante 30 s, 35 ciclos a 94°C durante 30 s, 50-64°C durante 30 s (dependiendo de la T_m de los cebadores), y a 72°C durante 1-2 min. (1 min./kb), respectivamente, y una reacción final a 72°C durante 10 min.

Los productos de la PCR se separaron mediante electroforesis en gel de agarosa, y se visualizaron mediante tinción con bromuro de etidio. Las bandas de interés se cortaron, y el ADN se extrajo subsiguientemente de la sílice en gel usando el QIAquick Gel Extraction Kit (Qiagen) siguiendo las instrucciones del fabricante. Los productos de la PCR extraídos se clonaron en vectores TOPO 2.1 (Invitrogen) y se transformaron en células de *Escherichia coli* OneShot[®] mediante transformación química siguiendo las instrucciones del fabricante. Las bacterias se cultivaron subsiguientemente en placas de agar Luria-Bertani (LB; Invitrogen) precalentadas (1% de triptona, 0,5% de extracto de levadura, 1,0% de NaCl, y 1,5% de agar) que contienen 50 µg ml⁻¹ de kanamicina y 40 µl de 40 mg ml⁻¹ X-gal (5-bromo-4-cloro-3-indolil-X-D-galactopiranosido; Invitrogen), y se incubaron a 37°C toda la noche. Las colonias positivas se seleccionaron usando selección blanco-azul en combinación con selección con antibióticos. Las colonias se recogieron y se inocularon en 6 ml de caldo LB (1% de triptona, 0,5% de extracto de levadura, 1,0% de NaCl) que contiene 50 µg ml⁻¹ de kanamicina, y se incubaron a 37°C en una incubadora de agitación a 200 rpm.

Los cultivos bacterianos se extrajeron y purificaron del cultivo del caldo LB usando el Qiagen Prep Plasmid Miniprep Kit (Qiagen), siguiendo las instrucciones del fabricante.

Secuenciación del ADN

- 45 El ADN plasmídico aislado se secuenció usando el método de terminación de cadena por didesoxinucleótido (Sanger et al., 1977), usando química Big-Dye (Versión 3.1) (Applied Biosystems). Se usaron cebadores directos o inversos M13 o cebadores génicos específicos. Los productos se separaron en un analizador genético ABI Prism 3100 (Applied Biosystems), y los datos de secuencias se compararon con la información de secuencia publicada en GenBank (NCBI) usando AlignX (Invitrogen).

- 50 Resultados

Identificación y secuenciación de TaMYB14

Se aislaron ARN total y ADN genómico (ADNg) de tejido de hoja en desarrollo y madura de *T. arvense*, y el ARN total se transcribió de forma inversa en ADNc. Inicialmente, se diseñaron cebadores para la región genérica de MYB de la secuencia codificante, y se llevó a cabo la PCR. Los productos de la PCR se separaron en geles de agarosa, y se visualizaron mediante tinción con bromuro de etidio. Las bandas que oscilan en tamaño se cortaron, el ADN se

extrajo, se purificó, se clonó en vectores TOPO, y se transformó en células de *E. coli*. Se escogieron al azar doscientos transformantes del suceso de clonación, el ADN plasmídico se aisló y se secuenció subsiguientemente. Se diseñaron cebadores adicionales para secuenciar las regiones N-terminales cuando era necesario (Tabla 4).

- 5 Se identificó una matriz de MYBs parciales secuenciando el ADNc aislado; >50% fueron desconocidas, no produciendo acierto sustancial con respecto a proteínas MYB conocidas. Las restantes se identificaron como ortólogos para MYBs expresados durante estrés abiótico, respuesta a falta de agua, estímulo lumínico, estrés salino, estímulo etilénico, estímulo auxínico, estímulo con ácido abscísico, estímulo con ácido giberélico, estímulo con ácido salicílico, estímulo con ácido jasmónico, cadmio, luz, movimiento y control estomático, regulación, similar a mixta (crecimiento de células epidérmicas), disminución de ácido cafeico O-metil-transferasa, y control de meristemo.
- 10 Dos ADNc de MYB parciales codificaron una proteína que cayó dentro de los clados correctos de MYB (NO8 y NO9) cuyos miembros incluyen aquellos que se sabe que activan la biosíntesis de antocianinas o CTs. Se diseñaron cebadores para el extremo 3' del gen para aislar el extremo 5' restante, y por tanto todo el clon de ADNc. El *TaMYB14* de longitud completa contiene una región codificante de 942 pb que codifica una proteína de 314 aminoácidos. En comparación, AtTT2 codifica una proteína de 258 aminoácidos.
- 15 Resultados Blast para *TaMYB14*

La secuencia de ADNc de *TaMYB14* procedente del genotipo AZ2925 de *T. arvense* se sometió a blast frente a las bases de datos públicas. BlastN devolvió los siguientes 5 aciertos principales:

- AB300033.1 "ARNm de *Lotus japonicus* LjTT2-1 para el factor de transcripción R2R3-MYB", (valor e 3e-69)
- AB300035.1 "ARNm de *Lotus japonicus* LjTT2-3 para el factor de transcripción R2R3-MYB", (valor e 4e-62)
- 20 AB300034.1 "ARNm de *Lotus japonicus* LjTT2-2 para el factor de transcripción R2R3-MYB", (valor e 4e-59)
- AF336284.1 ARNm de *Gossypium hirsutum* GhMYB36, (valor e 1e-40)
- AB298506.1 ARNm de *Daucus carota* DcMYB3-1 para factor de transcripción, (valor e 7e-39)

A su vez, BlastX de la secuencia traducida de *TaMYB14* procedente del genotipo AZ2925 de *T. arvense* devolvió los siguientes 5 aciertos principales:

- 25 BAG12893.1 "Factor de transcripción R2R3-MYB de *Lotus japonicus* LjTT2-1", (valor e 2e-81)
- AAK19615.1AF336282_1 "*Gossypium hirsutum* GhMYB10", (valor e 3e-76);
- BAG12895.1 "Factor de transcripción R2R3-MYB de *Lotus japonicus* LjTT2-3", (valor e 8e-74);
- BAG12894.1 "Factor de transcripción R2R3-MYB de *Lotus japonicus* LjTT2-2", (valor e 2e-72);
- AAZ20431.1 "MYB11 "[*Malus x domestica*], (valor e 2e-66)

- 30 El alineamiento de ADNc de *TaMYB14* frente a *AtTT2* y otros aciertos de BLAST se muestra en la Figura 7, mostrándose en amarillo las similitudes más elevadas. La traducción del marco de lectura abierto también mostró diferencias sustanciales en la composición de aminoácidos, compartiendo una homología de 52% con respecto a TT2 de *A. thaliana* (Figura 8). Además, *TaMYB14* comparte los motivos comunes a activadores de MYB de CT conocidos (NO9).

- 35 El alineamiento de ADNc de *TaMYB14* frente a *AtTT2* y otros aciertos de BLAST se muestra en la Figura 7, resaltándose en amarillo y en azul las similitudes. La traducción del marco de lectura abierto (Figura 8) también mostró diferencias sustanciales en la composición de aminoácidos, compartiendo una homología de 52% con respecto a TT2 de *A. thaliana*, principalmente en la región del dominio de MYB.

- 40 *TaMYB14* incluye un motivo similar al motivo del subgrupo 5 (DEXWRLxxT) según Stracke et al., 2001, que es común a activadores de MYB de CT previamente conocidos.

TaMYB14 carece del motivo de subgrupo 6 (KPRPR[S/T, mostrado en SEC ID NO:16) según Stracke et al., 2001, que es común a activadores de MYB de antocianinas previamente conocidos.

Además, este alineamiento ha identificado un nuevo motivo de MYB (VI/VRTKAXR/KxSK). Este nuevo motivo (resaltado en la Figura 8) parece asociado con un número de nuevos TFs de MYB14 que regulan las rutas de CTs.

- 45 Niveles de transcrito de *TaMYB14*

La acumulación de CTs se produjo en las especies *T. arvense* y *T. affine*, en las que fue detectable a lo largo de toda la lámina foliar en la capa epidérmica abaxial y adaxial, y el pecíolo; excepto la región del peciólulo. Los CTs son detectables solamente en *T. repens* y *T. occidentale* en los tricomas foliares en la superficie epidérmica abaxial.

5 El análisis de transcritos usando cebadores específicos para *TaMYB14* reveló que este gen se expresaba solamente en tejidos que acumulan de forma activa CTs. *TaMYB14* se expresó en tejido de hoja madura e inmadura de *T. arvense*, pero no en el callo (que no sintetiza CTs). Los cebadores diseñados para *TaMYB14* también amplificaron un MYB14 en *T. repens*, que se expresó en hoja de meristemo y tricomas meristemáticos tempranos, en los que los CTs se acumulan de forma activa, pero no se detectaron en tejido de hoja madura o emergente, estolones, internodos, raíces, y pecíolos. MYB14 tampoco se detectó en tejidos maduros de *T. occidentale*, en los que los CTs solamente están presentes en los tricomas de las hojas. Los resultados del análisis se muestran en la Tabla 3 a continuación:

Especie	Librería	Resultado	Esperado	Ruta
<i>T. repens</i> Huia	hoja madura	-	-	¿CT?
<i>T. repens</i> Huia	hoja joven	-	-	
<i>T. repens</i> Huia	hoja de meristemo	+	+	
<i>T. repens</i> Huia	tricoma temprano	+	+	
<i>T. repens</i> Huia	nodos e internodos de estolón	-	-	
<i>T. repens</i> Huia	Raíces	-	-	
<i>T. repens</i> Huia	floral	- +	+	
<i>T. repens</i> Huia	pecíolos	-	-	
<i>T. occidentale</i>	planta madura	-	-	
<i>T. repens</i> Isabelle	hoja madura	-	-	Antocianina
<i>T. arvense</i>	callo	-	-	CT-ve
<i>T. arvense</i>	hoja madura	+	+	CT
<i>T. arvense</i>	hoja inmadura	+	+	

10 TABLA 3: La expresión de MYB14 también coincide con la expresión de antocianidina reductasa (ANR; BAN) y LAR, dos enzimas clave específicas para la biosíntesis de CTs en legumbres.

15 Las Figuras 3 y 4 también mostraron la comparación de niveles de transcrito en diversos tejidos en la especie de *Trifolium*; la Figura 3 muestra niveles de transcrito de *TaMYB14* en tejidos variados de la especie de *Trifolium* y variedades de cultivo que se hacen crecer en condiciones idénticas de invernadero; Línea 1, (escalera); Línea 2, librería de ADNc de hoja madura de *T. repens* (variedad de cultivo Huia); Línea 3, librería de ADNc de raíz madura de *T. repens* (variedad de cultivo Huia); Línea 4, librería de ADNc de estolón maduro de *T. repens* (variedad de cultivo Huia); Línea 5, librería de ADNc floral madura de *T. repens* (variedad de cultivo DC111); Línea 6, ADNc de hoja emergente de *T. repens* (variedad de cultivo Huia); Línea 7, ADNc de hoja madura de *T. repens* (variedad de cultivo Isabelle rica en antocianinas); Línea 8, ADNc de hoja inmadura de *T. arvense* (variedad de cultivo AZ2925); Línea 9, ADNc de hoja madura de *T. arvense* (variedad de cultivo AZ2925); Línea 10, ADNc floral de meristemo de *T. repens* (variedad de cultivo Huia); Línea 11, ADNc de hoja de meristemo de *T. repens* (variedad de cultivo Huia); Línea 12, ADNc solamente de tricoma de meristemo de *T. repens* (variedad de cultivo Huia); Línea 13, librería de ADNc de planta (hoja, raíz y estolón) madura de *T. occidentale* (variedad de cultivo Huia); Línea 14, librería de ADNc de nodo maduro de *T. repens* (variedad de cultivo Huia); Línea 15, clon de MYB14cDNA de *T. arvense* clonado en TOPO, Línea 16, clon genómico de MYB14 de *T. arvense* clonado en TOPO, Línea 17, ADN genómico de *T. occidentale*; Línea 17, ADN genómico de *T. repens*; Línea 17, ADN genómico de *T. arvense*; Línea 20, (escalera).

30 Mientras, la Figura 4 muestra niveles de transcrito de BANYULS(A) y LAR (B) en tejidos variables de la especie de *Trifolium* y variedades de cultivo que se hacen crecer en condiciones de invernadero idénticas. Línea 1, (escalera); Línea 2, librería de ADNc de hoja madura de *T. repens* (variedad de cultivo Huia); Línea 3, librería de ADNc de raíz madura de *T. repens* (variedad de cultivo Huia); Línea 4, librería de ADNc de estolón maduro de *T. repens* (variedad de cultivo Huia); Línea 5, librería de ADNc floral madura de *T. repens* (variedad de cultivo DC111); Línea 6, ADNc de hoja emergente de *T. repens* (variedad de cultivo Huia); Línea 7, ADNc de hoja madura de *T. repens* (variedad de cultivo Isabelle rica en antocianinas); Línea 8, ADNc de hoja inmadura de *T. arvense* (variedad de cultivo AZ2925); Línea 9, ADNc de hoja madura de *T. arvense* (variedad de cultivo AZ2925); Línea 10, ADNc floral de meristemo de *T. repens* (variedad de cultivo Huia); Línea 11, ADNc de hoja de meristemo de *T. repens* (variedad de cultivo Huia); Línea 12, ADNc solamente de tricoma de meristemo de *T. repens* (variedad de cultivo Huia); Línea 13, librería de ADNc de planta (hoja, raíz y estolón) madura de *T. occidentale* (variedad de cultivo Huia); Línea 14, librería de ADNc de nodo maduro de *T. repens* (variedad de cultivo Huia); Línea 15, clon de BAN o LAR de ADNc de *T. arvense*

clonado en TOPO, Línea 16, clon genómico de BAN o LAR de *T. arvense* clonado en TOPO, Línea 17, ADN genómico de *T. occidentale*; Línea 17, ADN genómico de *T. repens*; Línea 17, ADN genómico de *T. arvense*; Línea 20, (escalera).

Identificación y secuenciación de MYB14 de ADNg de *T. arvense*, *T. affine*, *T. occidentale* y *T. repens*

5 Usando cebadores diseñados para la región de inicio y parada de *TaMYB14* (véase la Tabla 4), los inventores amplificaron homólogos de *TaMYB14* mediante PCR de ADNc y ADNg aislados de un abanico de varias especies de *Trifolium*, a saber, *T. arvense*, *T. affine*, *T. repens* y *T. occidentale*. El aislamiento de la secuencia de ADN genómico y la secuenciación de longitud completa de los productos de la PCR clonados mostró que *T. arvense* tiene dos isoformas o alelos de este gen, uno de los cuales corresponde a la secuencia de ADNc expresada, correspondiendo
10 el otro a una variante alélica/isoforma previamente no identificada de *TaMYB14*.

El alineamiento de estas isoformas o variantes alélicas reveló la presencia de varias supresiones e inserciones de bases en comparación con la secuencia de ADNc de *TaMYB14* (véase la Figura 10). La traducción de la secuencia de ADNc putativa reveló que la proteína codificada por esta isoforma o variante alélica también tiene supresiones, inserciones e intercambios de aminoácidos (véase la Figura 9). A la variante alélica se le denominó como *TaMYB14-2*.
15

Las secuencias de ADNg de longitud completa correspondientes para este gen también se aislaron a partir de otras tres especies de *Trifolium*: *T. affine*, *T. repens* y *T. occidentale*. Todos los alelos de MYB14 tuvieron tres exones y dos intrones de tamaños variables (véanse las Figuras 10-12). *T. affine* y *T. occidentale* tienen ambas un alelo, mientras que *T. repens* tiene dos alelos. Las secuencias traducidas de MYB14 de las diversas especies fueron 95% homólogas a *TaMYB14*, con cambios en la composición de aminoácidos. La mayoría de las diferencias de aminoácidos están situadas en la región única de 3' en dirección 3' del dominio de MYB.
20

Uso de cebadores	Código	Secuencia del cebador	SEC ID NO:
Búsqueda del dominio de MYB	MYBFX	GACAATGAGATAAAGAAT TACTTG	18
Búsqueda del dominio de MYB	MYBFY	AAGAGTTGTAGACTTAGM TGG	19
Búsqueda del dominio de MYB	MYBFZ	YTKGGSAACAGGTTGTC	20
Aislamiento de longitud completa	M14ATG	ATGGGGGAGAAGCCCTTGT TGTGC	21
Aislamiento de longitud completa	M14TGA	TCATTCTCCTAGTACTTCC TCACTGG	22
Paseo génico	M14TSP1	CTCTTTTTGGAAGGTTTC TCC	23
Paseo génico	M14TSP2	TTCTCCATTTTCCTTCACC ATGG	24
Paseo génico	M14TSP3	TCCAAGCACCTCTATTCA AGCC	25
Clonación en el vector	M14FATG	CTCGAGATGCAATGCTGG TTGATGGTGTGGC	26
Lotus corniculatus	MYBLF	CATTGCCCTGTAGATTCTG TAGCC	27
Lotus corniculatus	MYBLR	TGAAGATTGTTGGACACA TTGG	28
Extremo 5' UTR de MYB14	MYB148N	AGGTTGGAATACAAGACA GAC	29
Extremo 3' UTR de MYB14	MYB14RR	TCTCCTAGTACTTCCTCA CTGG	30
Cebador para intrón 1	15	ATAATCATACTAATTAACA TCAC	31

ES 2 537 342 T3

Uso de cebadores	Código	Secuencia del cebador	SEC ID NO:
Cebador para intrón 1	13	TGATAGATCATGTCATTG TG	32
Paseo génico	TSP4	GCCTTCCTTTGCACAACA AGGGC	33
Paseo génico	TSP5	GCACAACAAGGGCTTCTC CCC	34
Sitio de partida de 5' directo	MYB148F	ATGGGGAGAAGCCCTTGT TGTGC	35
Sitio de partida de 5' inverso	MYB14RR	TCTCCTAGTACTTCCTCA CTGG	36
Análisis de expresión/ vector de silenciamiento	MYB14F	CTCGAGCAATGCTGGTTG ATGGTGTGGC	37
Análisis de expresión/ vector de silenciamiento	MYB14R	TCTAGAGGACACATTTGT CTCATCAGC	38
Paseo génico	MYB14R2	TCTAGATTGAGTTTGGTC CGAACAAGG	39
Paseo génico	MYB14R3	TCTAGAAATCTTCTAGCAA ATCTGCGG	40
Secuenciación	M13 Directo	GTAAAACGACGGCCAG	41
	M13 Inverso	CAGGAAACAGCTATGAC	42
Producción de ADNc	BD SMART II™ A Oligonucleotide	AAGCAGTGGTATCAACGC AGAGTACGCGGG	43
Producción de ADNc	3' BD SMART™ CDS Primer II A	AAGCAGTGGTATCAACGC AGAGTACT(30)V N-3'	44
Amplificación del ARNm	5' PCR Primer II A	AAGCAGTGGTATCAACGC AGAGT	45

TABLA 4: Secuencias de cebadores para PCR, clonación y secuenciación de MYB14 de diversas especies de *Trifolium* (*T. arvense*; *T. repens*; *T. affine*; *T. occidentale*).

En resumen, se han identificado y aislado diez nuevas proteínas/genes MYB14, como se resume en la Tabla 5 a continuación, que también muestra la SEC ID NO: asociada con cada secuencia en el listado de secuencias:

Especie, y referencia de secuencia	SEC ID NO:			
	ADNc de longitud completa	ADNg	Proteína	ORF
<i>Trifolium arvense</i> , TaMYB14-1	1, 13	2	14	55
<i>Trifolium arvense</i> , TaMYB14-2	-	3	46	56
<i>Trifolium affine</i> , TafMYB14-1	5	4	47	57
<i>Trifolium affine</i> , TafMYB14-2	-	6	48	58
<i>Trifolium occidentale</i> , ToMYB14-1	-	7	49	59
<i>Trifolium occidentale</i> , ToMYB14-2	-	8	50	60
<i>Trifolium repens</i> , TrMYB14-1	-	9	51	61
<i>Trifolium repens</i> , TrMYB14-2	-	10	52	62

Especie, y referencia de secuencia	SEC ID NO:			
	ADNc de longitud completa	ADNg	Proteína	ORF
<i>Trifolium repens</i> , TrMYB14-3	-	11	53	63
<i>Trifolium repens</i> , TrMYB14-4	-	12	54	64

TABLA 5: Resumen de secuencias de MYB14 de la invención

En la Figura 34 se muestra un alineamiento de todas estas secuencias de MYB14. Se identificaron dos motivos de secuencia comunes a todas las secuencias proteicas de MYB14.

El primer motivo es DDEILKN (SEC ID NO:15).

- 5 El segundo motivo es $X_1VVRTX_2AX_3KCSK$ (SEC ID NO:17), en el que $X_1 = N$, $Y = O$ o H , $X_2 = K$ o R , y $X_3 = T$ o I .

La presencia de cualquiera o de ambos de estos motivos parece que es un diagnóstico para proteínas MYB14, particularmente cuando se asocia con una falta de motivo de SEC ID NO:16.

La Figura 35 muestra el porcentaje de identidad entre cada una de las proteínas MYB14 alineadas en la Figura 34.

- 10 También se ha mostrado que el patrón de expresión espacial y temporal de TaMYB14 está correlacionado consistentemente con la producción de CT en plantas *in vivo*.

Ejemplo 2: Uso de la secuencia de ácido nucleico de MYB14 de la invención para producir taninos condensados en trébol blanco (*Trifolium repens*)

Materiales y métodos

Constructos genéticos usados en el protocolo de transformación

- 15 El vector de transformación vegetal pHZBar deriva de pART27 (Gleave 1992). El casete de selección *nos-nptII-nos3'* se ha sustituido por el casete de selección CaMV35S-BAR-OCS3' con el gen *bar* (que confiere resistencia al herbicida glufosinato de amonio) expresado a partir del promotor del CaMV 35S. La clonación de los casetes de expresión en este vector binario se facilitó mediante un sitio de restricción *NotI* único y la selección de recombinantes por cribado mediante azul/blanco para β -galactosidasa. El trébol blanco se transformó usando
- 20 M14ApHZBarP, que contiene el alelo expresado de *Trifolium arvense*. Los casetes de sobreexpresión para M14ApHZBarP se clonaron en primer lugar en pART7. El constructo se transportó entonces a pHZBar como un fragmento *NotI*. Los ADN-T de los constructos genéticos, que muestran la orientación de genes clonados, se representan gráficamente en la Figura 6.

- 25 Los constructos genéticos en pHZBar se transfirieron a la cepa GV3101 de *Agrobacterium tumefaciens* como ADN plasmídico usando el método de transformación de congelación-descongelación (Ditta et al 1980). La estructura de los constructos mantenidos en *Agrobacterium* se confirmó mediante digestión de restricción de los ADN plasmídicos preparados a partir de cultivo bacteriano. Los cultivos de *Agrobacterium* se prepararon en glicerol y se transfirieron hasta -80°C para el almacenamiento a largo plazo. Los constructos genéticos mantenidos en la cepa GV3101 de *Agrobacterium* se inoculan en 25 ml de caldo MGL que contiene espectinomicina a una concentración de 100 mg/l.
- 30 Los cultivos se hacen crecer toda la noche (16 horas) en un agitador giratorio (200 rpm) a 28°C . Los cultivos bacterianos se cosechan mediante centrifugación (3000 x g, 10 minutos). El sobrenadante se elimina, y las células se resuspenden en 5 ml de disolución de MgSO_4 10 mM.

Transformación de explantes cotiledonarios

- 35 El trébol se transformó usando un método modificado de Voisey et al. (1994). Las semillas se pesaron para proporcionar aproximadamente 400-500 cotiledones (es decir, 200-250 semillas) para disección (0,06 g = 100 semillas). En un tubo de centrifugadora, las semillas se enjuagaron con etanol al 70% durante 1 minuto. Las semillas se esterilizan superficialmente en lejía (5% de cloro disponible) agitando en una mezcladora circular durante 15 minutos, seguido de cuatro lavados en agua estéril. Las semillas se empapan toda la noche a 4°C . Los cotiledones se disecan de las semillas usando un microscopio de disección. Inicialmente, se eliminan el tegumento y el endospermo.
- 40 Los cotiledones se separan del radical con el escalpelo colocando la cuchilla entre los cotiledones y cortando a lo largo del tallo que queda. Los explantes cotiledonarios se cosechan sobre un disco de filtro estéril en medio CR7.

- Para la transformación, se dispensa una alícuota de 3 μl de suspensión de *Agrobacterium* sobre cada cotiledón disecado. Las placas se cierran herméticamente y se cultivan a 25°C en un fotoperíodo de 16 horas. Después de un período de 72 horas de cocultivo, los cotiledones transformados se transfieren a placas que contienen medio CR7 suplementado con glufosinato de amonio (2,5 mg/l) y timentina (300 mg/l), y se devolvieron a la habitación de cultivo. Tras la regeneración de brotes, los explantes se transfirieron a medio CR5 suplementado con glufosinato de amonio

(2,5 mg/l) y timentina (300 mg/l). Los brotes que se regeneran se subcultivan tres semanas a medio CR5 reciente que contiene selección. A medida que se produce la formación de raíces, las plántulas se transfieren en recipientes que contienen medio CR0 que contiene selección mediante glufosinato de amonio. En esta etapa, los grandes agrupamientos de regenerantes se dividen en plántulas individuales. Las plantas enraizadas completas que crecen bajo selección se plantan entonces en macetas en tapones de turba estéril.

Metodología LCMSMS para el análisis mediante HPLC

Para extraer flavonoides para el análisis mediante HPLC, se congela tejido de hoja (0,5 g de peso reciente) en N₂ líquido, se tritura hasta un polvo fino y se extrae con ácido acético:metanol (80:20 v/v) durante 30 min. a 4°C. El desecho vegetal se peletizó en una microcentrifugadora a 13K rpm durante 10 min. El sobrenadante se eliminó y se colocó a -20°C durante 30 min. Para el análisis mediante HPLC se usó una alícuota. Se analizó una alícuota mediante HPLC usando detección tanto mediante UV-PDA como MS/MS en un sistema de espectrómetro de masas con trampa de iones Thermo LTQ. Los extractos se resolvieron en una columna de fase inversa Phenomenex Luna C18 mediante elución por gradiente con agua y acetonitrilo con 0,1% de ácido fórmico como el sistema de fase móvil. La detección de las antocianinas se realizó mediante absorción de UV a 550 nm, y los otros metabolitos se estimaron mediante detección MS1 o MS2 mediante el espectrómetro de masas.

El instrumento usado fue un espectrómetro de masas con trampa de iones lineal (Thermo LTQ) acoplado a un sistema de HPLC Thermo Finnigan Surveyor (ambos de San Jose, CA, USA) equipado con un detector de conjunto de fotodiodos (PDA) Thermo. Para la adquisición y procesamiento de datos, se usó el software Thermo Finnigan Xcalibur (versión 2.0).

Se inyectó una alícuota de muestra de 5 µl en una columna de 150 x 2,1 mm Luna C18(2) (Phenomenex, Torrance, CA) mantenida a 25°C constantemente. Los disolventes de la HPLC usados fueron: disolvente A = 0,1% de ácido fórmico en H₂O; disolvente B = 0,1% de ácido fórmico en acetonitrilo. El caudal fue 200 µl min⁻¹, y el gradiente de disolvente usado se muestra en la Tabla 6 a continuación. Los datos de PDA se recogieron a lo largo del intervalo de 220 nm-600 nm para todo el cromatograma.

Tiempo (min)	Disolvente A%	Disolvente B%
0	95	5
6	95	5
11	90	10
26	83	17
31	77	23
41	70	30
45	50	50
52	50	50
52	3	97
59	3	97
62	95	5
70	95	5

TABLA 6: Gradiente de HPLC

El espectrómetro de masas se ajustó para ionización por electropulverización en modo positivo. El voltaje de pulverización fue 4,5 kV y la temperatura del capilar 275°C, y los caudales de gas de desolvatación, gas auxiliar, y gas de barrido se ajustaron (en unidades arbitrarias/min.) a 20, 10 y 5, respectivamente. Los primeros 4 y los últimos 11 minutos del caudal procedente de la HPLC se desviaron a residuos. El MS se programó para barrer de 150-2000 m/z (barrido MS¹), y después se llevó a cabo MS³ dependiente de los datos en el ion de MS¹ más intenso. Las ventanas de aislamiento para el método de MS³ dependiente de los datos fue 2 mu (unidades de masa nominales), y la fragmentación (35% CE (energía de colisión relativa)) del ion más intenso procedente del espectro MS¹ fue seguida del aislamiento (2 mu) y fragmentación (35% CE) del ion más intenso procedente del espectro MS². El espectrómetro de masas llevó a cabo entonces secuencialmente la monitorización de la reacción seleccionada (SRM) sobre las masas en la Tabla 7 a continuación, con ventanas de aislamiento para cada SRM de 2,5 mu y fragmentación CE de 35%. Estas masas enumeradas cubren las diferentes combinaciones de masas de procianidina (catequina y/o epicatequina) y prodefinidina (galocatequina o epigalocatequina) hasta el trómero.

masa de SRM (m/z)	Intervalo de barrido de MS2 (m/z)	Compuesto diana
291,3	80-700	monómeros de PC
307,3	80-700	monómeros de PD
579,3	155-2000	dímeros de PC:PC
595,3	160-2000	dímeros de PC:PD
611,3	165-2000	dímeros de PD:PD
867,3	235-2000	trímeros de PC:PC:PC
883,3	240-2000	trímeros de PC:PC:PD
899,3	245-2000	trímeros de PC:PD:PD
915,3	250-2000	trímeros de PD:PD:PD

Tabla 7: Masas de SRM para monómeros, dímeros y trímeros.

Resultados

Análisis mediante DMACA de trébol blanco con MYB14 de ADN_G de *T. arvense*

5 Los cotiledones de trébol blanco se transformaron con el alelo de *T. arvense* correspondiente a la secuencia de ADN_C expresada, bajo el control del promotor del CaMV 35S, y se regeneraron como se describe en los métodos. Las hojas procedentes de todas las plántulas regeneradas se cribaron en busca de la producción de CTs con tinción mediante DMACA, como se describe en el Ejemplo 1. Un número de estas plantas transformadas fueron positivas para la producción de CTs, dando como resultado tinción azul cuando se tiñen con DMACA. Tal tinción se produjo en la mayoría de las células epidérmicas de los tejidos de las hojas, incluyendo las seis células centrales de tricomas de las hojas. En comparación, las plantas de trébol blanco de tipo salvaje no transformadas fueron negativas para CT, excepto los tricomas en el lado de la hoja abaxial (Figura 5). Los CTs también estuvieron presentes en algunas células de la raíz y peciolares de algunas plantas. Esto indica que la expresión constitutiva de *TaMYB14* altera el patrón temporal y espacial de la acumulación de CTs en plantas de trébol blanco.

Análisis molecular, cribado mediante DMACA, y bioquímica de trébol blanco transgénico

15 Análisis molecular de trébol blanco

El ADN extraído de plantas de trébol blanco transgénicas se estudió en busca de la integración del vector M14ApHZBAR. Se llevaron a cabo reacciones de PCR usando conjuntos de cebadores diseñados para amplificar un producto que incluye una porción del promotor 35S y la mayoría del gen *TaMYB14*. Los resultados de este análisis indicaron la integración del vector binario que contiene el gen *TaMyb14A* en el genoma del trébol blanco (Figura 14).

20 Análisis mediante DMACA de trébol blanco

Se muestran los resultados logrados a partir de la tinción con DMACA de tejidos foliares de trébol blanco (Figura 15). El colorante específico de CTs, DMACA, ha teñido fuertemente el limbo y el pecíolo de la hoja de las hojas de trébol transgénico (B, C, D, G, H), en comparación con la hoja de trébol blanco de tipo salvaje (A, E, F).

25 Además (Figura 16), las células de la capa de tricomas y las células apicales se tiñeron mucho más fuertemente (F, G) que lo observado normalmente en hojas de tipo salvaje (E). Las células oclusivas de los estomas también se tiñó fuertemente (H). Hubo una tinción evidente en el núcleo de las células epidérmicas de la misma forma que en la célula del tricoma del tallo. Las células epidérmicas se tiñeron más uniformemente de lo normal, y las células basales de la roseta también se tiñeron fuertemente (G). Se llevó a cabo el rasgado de las hojas para ayudar a establecer qué células específicas tienen tinción con DMACA (I a K). En esta ocasión, la epidermis inferior (parte más alta de la superficie exterior) se separó de la capa mesofilica. Las células epidérmicas (excepto células especializadas tales como estomas y tricomas) tuvieron poca actividad en comparación con la capa celular mesofilica. Las células mesofilicas mostraron fuerte tinción definida por toda la célula, con sublocalización definida en orgánulos específicos semejantes a vacuolas, que son obviamente múltiples por célula. Por lo tanto, hay una compartimentalización de la tinción por DMACA en las células mesofilicas.

35 Análisis mediante HPLC/LCMS de trébol blanco

El análisis bioquímico de los solicitantes del tejido transgénico transformado con M14ApHZBAR proporcionó pruebas irrefutables de que la sobreexpresión de *TaMYB14* conduce a la acumulación de monómeros, dímeros y trímeros de taninos condensados en tejido foliar en trébol blanco y tabaco. También es posible que estén presentes taninos de

cadena más largas, pero la resolución de esto está más allá del alcance de nuestro equipo.

5 Se usó extracto de semilla de uva purificado como el estándar para todas las medidas de LCMSMS HPLC, debido a que su perfil de taninos se ha caracterizado muy bien y se muestra en las Figuras 17 y 18. Este extracto permite la identificación definida de catequina (C), epicatequina (EC), galocatequina (GC) y epigalocatequina (EGC), así como la detección de dímeros PC:PC, un dímero PC:PD y dos trímeros 3PC.

Los espectros MS2 de los cuatro monómeros se proporcionan como prueba de la identificación de estos metabolitos.

10 Los flavonoides se extraen de plantas de trébol blanco transgénicas y de control de tipo salvaje, y se procesan vía HPLC/LCMS. Los resultados de estos análisis confirmaron la presencia de CT en extractos de hojas procedentes de muestras de trébol transgénico. La mayoría de los monómeros detectados fueron epicatequina y epigalocatequina, con trazas de galocatequina. Esto es consistente ya que los taninos de trébol derivan de delphinidina. No se detectaron monómeros en tejido de hoja de trébol blanco de tipo salvaje (Figura 19). También se detectaron dímeros y trímeros (Figuras 20, 21).

Ejemplo 3: Uso de la secuencia de ácido nucleico de MYB14 de la invención para producir taninos condensados en tabaco (*Nicotiana tabacum*)

15 Materiales y métodos

Construido genético usado en protocolos de transformación

El fragmento *NotI* procedente del plásmido M14ApHZBAR (Figura 6) se aisló y se clonó en pART27 (Gleave, 1992) para la transformación de tabaco. Este vector binario contiene el gen de selección *nptII* para resistencia a kanamicina bajo el control del promotor del CaMV 35S.

20 Transformación de tabaco

25 Se transformó tabaco vía el método de transformación-regeneración de discos de hojas (Horsch et al. 1985). Discos de hojas procedentes de plantas de tabaco W38 de tipo salvaje estériles se inocularon con una cepa de *Agrobacterium tumefaciens* que contiene el vector binario, y se cultivaron durante 3 días. Los discos de hojas se transfirieron entonces a medio selectivo MS que contiene 100 mg/ml de kanamicina y 300 mg/l de cefotaxima. La regeneración de los brotes se produjo a alrededor de un mes, y los explantes foliares se colocaron en medio libre de hormonas que contiene kanamicina, para la formación de raíces.

Resultados

Análisis molecular, cribado mediante DMACA, y bioquímica de tabaco transgénico

Análisis molecular de tabaco

30 El ADN extraído de plantas de tabaco transgénicas se estudió para determinar la integración del vector binario M14ApHZBAR. Las reacciones de PCR se llevaron a cabo usando conjuntos de cebadores diseñados para amplificar una porción del promotor 35S y la mayoría del gen. Los resultados de este análisis indicaron integración del vector binario, que contiene el gen TaMyb14A, en el genoma del trébol blanco (Figura 22).

Análisis mediante DMACA del tabaco

35 El análisis mediante DMACA se llevó a cabo sobre las plantas de tabaco como se describe para el trébol en el Ejemplo 1. Las plántulas de tabaco transgénicas que expresan TaMYB14A (bajo el control del promotor del virus del mosaico de la coliflor 35S) no mostraron diferencias significativas en el crecimiento en comparación con plantas de tipo salvaje. Además, se detectó CT en tejido foliar de plántulas de tabaco transgénicas derivadas de células de tabaco de tipo salvaje o de tabaco transgénico (que ya acumulan antocianina), en comparación con tabaco sin
40 transformar de tipo salvaje que no acumula CT en tejidos vegetativos. Esto indica que el gen de MYB14 de *T. arvense* es capaz de activar todos los genes de la ruta de CT en tabaco, por sí mismo. Se muestran ejemplos de la tinción mediante DMACA de hojas de tabaco transgénicas (Figura 23). El colorante específico de CT, DMACA, tiñó fuertemente el limbo de la hoja de las hojas de tabaco transgénicas (A a G) en comparación con las hojas de tipo salvaje, que están siempre desprovistas de CT.

45 Análisis mediante HPLC/LCMS de tabaco

50 El análisis mediante HPLC/LCMS se llevó a cabo para tabaco como se describe para el trébol en el Ejemplo 2. Se extrajeron flavonoides de plantas de tabaco transgénicas y de control de tipo salvaje, y se procesaron vía HPLC. Los resultados de estos análisis confirmaron la presencia de CT en extractos foliares procedentes de las muestras de tabaco transgénico. Las muestras de control de tabaco estaban desprovistas de unidades de CT. La mayoría de monómeros detectados fueron epicatequina, con pequeñas cantidades de monómeros epigalocatequina y galocatequina (Figura 24). También se detectaron dímeros y trímeros (Figura 25).

Ejemplo 4: Uso de la secuencia de ácido nucleico de MYB14 de la invención para reducir la producción de taninos condensados en *Trifolium arvense*

Materiales y métodos

Construceto genético usado en el protocolo de silenciamiento

- 5 Se usó pHANNIBAL (Helliwell y Waterhouse, 2003), un vector vegetal de ARNi de horquilla de pelo, para transformar cotiledones de *T. arvense* con un construceto que expresa porciones autocomplementarias de una secuencia homóloga a una porción del ADNc de TaMYB14. Todo el ADNc para el MYB14 (previamente aislado de una librería de hojas) se usó para amplificar un fragmento de 299 pb de longitud del ADNc procedente del extremo 3' del gen
- 10 (caatgctggttgatggtggtgctagtgattcaatgagtaacaacgaaatggaacacggttatggattttgctatttgcgatgaagaga aagaactatccgcagatttgctagaagattttaacatcgcgatgatattgcttatctgaacttttgaactctgatttctcaaatgctgca atttgcattacaatgatctattgtcacctgttcgaccaaactcaaatgttctctgatgatgagattctcaagaattggacacaatgtaact ttgctgatgagacaaatgtgtcc - SEC ID NO:65). Los cebadores se diseñaron para permitir la clonación de los fragmentos en el vector de silenciamiento pHANNIBAL (Tabla 5). El fragmento se clonó en el sitio *XhoI* en la dirección sentido delante del intrón *pdk*, o en los sitios *XbaI*, después del intrón *pdk*, en la dirección antisentido. La dirección de la clonación se determinó mediante PCR para asegurarse de que el fragmento estaba en la orientación correcta. El fragmento *NotI* procedente de MYB14pHANNIBAL que contiene el casete de ARNhp se subclonó en pHZBar (denominado pHZBARSMYB (Figura 13) y se usó en experimentos de transformación.

Cebador	Secuencia
MYB14F1	<i>TCTAGACAATGCTGGTTGATGGTGTGGC</i> (SEQ ID NO:66)
MYB14R	<i>TCTAGAGGACACATTTGTCTCATCAGC</i> (SEQ ID NO:67)
MYB14F	CTCGAGCAATGCTGGTTGATGGTGTGGC (SEQ ID NO:68)
MYB14R1	CTCGAGGGACACATTTGTCTCATCAGC (SEQ ID NO:69)

Tabla 8: Cebadores modificados para incluir un sitio de enzima de restricción *XbaI* (resaltado en cursiva) o un sitio de enzima de restricción *XhoI* (resaltado en negrita) en el extremo 5' de los cebadores, para permitir la clonación.

20 Transformación de *T. arvense*:

Se transformaron variedades de cultivo de *T. arvense* con el vector binario de silenciamiento pHZbarSMYB, esencialmente como se describe para *T. repens*, con algunas modificaciones pequeñas (Voisey et al., 1994). El nivel de glufosinato de amonio se disminuyó hasta 1,25 mg/l, y las plantas se colocaron en medio CR5 durante solamente una quincena antes de la colocación sobre medio CR0 para la regeneración de raíces.

25 Resultados

Análisis molecular, cribado mediante DMACA, y bioquímica de *Trifolium arvense* transgénico

Análisis molecular de *T. arvense*

- 30 El ADN extraído de plantas transgénicas de *T. arvense* se estudió para determinar la integración del vector binario M14pHANNIBAL. Las reacciones de PCR se llevaron a cabo usando conjuntos de cebadores diseñados para amplificar una porción del promotor 35S y el extremo 3' del fragmento génico de ADNc. Los resultados de este análisis indicaron integración del vector binario, que contiene el construceto génico de ARNhp, en el genoma (Figura 26).

Análisis mediante DMACA de *T. arvense*

- 35 El material vegetal procedente de *T. arvense* de control y algunas de las plántulas transformadas se ha teñido usando DMACA (Figura 27) como se describe en el Ejemplo 1. Las plantas transformadas se compararon con las hojas maduras de tipo salvaje también regeneradas a través de cultivo tisular, ya que el cultivo tisular afecta a la regeneración de las hojas, y el comienzo de la producción de tanino se comparó con plantas que se hicieron crecer de forma natural en el suelo derivadas de semillas. El callo de *T. arvense* de tipo salvaje no produce tanino (A), pero las células comienzan a acumular tanino en hojas que se asemejan a tejidos (B a D – color púrpura). Las plantas transgénicas tampoco producen tanino en el callo, pero el tejido foliar teñido de forma similar con DMACA mostró solamente una tinción azul de color claro (E-L), indicando que los niveles de CT se redujeron drásticamente en plantas que expresan el construceto de silenciamiento.
- 40

Análisis mediante HPLC/LCMS de *T. arvense*

- 45 Se extrajeron flavonoides de plantas de *T. arvense* transgénicas y de control de tipo salvaje, y se procesaron vía HPLC/LCMS, como se describe en el Ejemplo 2. Las plántulas de *T. arvense* de tipo salvaje (no transformadas)

tuvieron niveles detectables elevados de monómeros de CT. La mayoría de estos monómeros fueron catequina, con pequeñas cantidades de monómeros de galocatequina (Figura 28). También se detectaron dímeros (Figura 29). Por el contrario, solamente se detectaron trazas de estos compuestos en las plántulas transformadas, si es que se detectaron. Por lo tanto, el análisis mediante HPLC de plántulas de *T. arvense* silenciadas confirmó que la acumulación de CT se había reducido significativamente. Estos resultados confirman que la ausencia de CT en extractos foliares procedentes de las plantas transgénicas de *T. arvense* está asociada con la presencia del vector diseñado para silenciar la expresión de TaMYB14.

Ejemplo 5: Uso de la secuencia de ácido nucleico de MYB14 de la invención para producir taninos condensados en alfalfa (*Medicago sativa*)

10 Materiales y métodos

Transformación de alfalfa mediante bombardeo con microproyectiles

Para todos los experimentos de transformación se usó la variedad de cultivo Regen-SY (Bingham 1991). El protocolo de transformación se adaptó de Samac et al (1995). Se iniciaron cultivos de callo a partir de explantes de pecíolos, y se hicieron crecer en la oscuridad en medio Schenk y Hildebrandt media (Schenk y Hildebrandt, 1972) suplementado con ácido 2,4-diclorofenoxiacético y quinetina (SHDK). Los cultivos en desarrollo se hicieron pasar por subcultivo normal sobre medio reciente a intervalos de cuatro semanas. El callo de Regen Sy de ocho a doce semanas se transformó mediante bombardeo con microproyectiles en un aparato de sistema de suministro de partículas Bio-Rad PDS1000/He Biolistic®. Los cultivos de callo se incubaron durante un mínimo de cuatro horas en medio SHDK suplementado con una concentración 0,7 M de sorbitol y manitol para inducir la plasmólisis celular. ADN plasmídico (1 µg/µl) de p35STaMyb14A (que contiene el fragmento *NotI* procedente de M14ApHZBAR) y pCW122 (que contiene un gen *nptII* para conferir resistencia al antibiótico kanamicina; Walter et al, 1998) se precipitaron sobre partículas de tungsteno (M17, Bio-Rad) como se describe por el fabricante. Se usaron parámetros estándar (vacío de 27" de Hg, ruptura de 1100 psi, y distancia a la diana de 100 mm) para la transformación según el manual de instrucciones.

Los tejidos transformados se hicieron descansar toda la noche antes de transferirlos a medio SHDK. Después de dos días, los cultivos se transfirieron a medio SHDK que contiene selección de antibióticos (kanamicina 50 mg/l), para la selección de células transformadas. Este material se subcultivó hasta tres veces a intervalos de tres semanas antes de transferirlo a medio SH libre de hormonas o medio de Blaydes (Blaydes, 1966), y se colocó en la luz para la regeneración. Los embriones somáticos que germinan se disecaron a partir de la masa de callo y se transfirieron a un medio de Murashige y Skoog de fuerza media (Murashige y Skoog, 1962) para el desarrollo de raíces y brotes.

Objetivo

Se llevaron a cabo experimentos de transformación para introducir un plásmido que contiene el gen TaMyb14 bajo el control del promotor CaMV35S en alfalfa. El objetivo fue generar plantas que expresan TaMyb14 y cribar en busca de la acumulación de taninos condensados en tejidos foliares.

Resultados

Análisis molecular, cribado mediante DMACA, y bioquímica de alfalfa transgénica

Análisis molecular de alfalfa

El ADN extraído de la alfalfa transgénica se evaluó para determinar la integración del vector p35STaMyb14A. Se usaron conjuntos de cebadores diseñados para amplificar el producto a partir del gen *nptII* o del gen TaMyb14. Los resultados de este análisis indicaron integración de ambos constructos plasmídicos en el genoma de la alfalfa (Figura 30).

Análisis mediante DMACA de la alfalfa

Para evaluar la acumulación de taninos condensados, se puede llevar a cabo el análisis mediante DMACA para las plantas de alfalfa como se describe para el trébol en el Ejemplo 1.

Análisis mediante HPLC/LCMS de alfalfa

El análisis mediante HPLC/LCMS como se describe para trébol en el Ejemplo 2 anterior se puede usar para detectar de forma exacta la presencia de monómeros, dímeros y trímeros de taninos en alfalfa transgénica. Para llevar a cabo el análisis, se extraen flavonoides de plantas de alfalfa transgénicas y de control de tipo salvaje, como se describe para trébol. La alfalfa de tipo salvaje acumula (en el tegumento) principalmente taninos derivados de cianidina y pequeñas cantidades de taninos derivados de delphinidina (Pang et al., 2007). Las hojas de las estirpes de alfalfa transgénicas que expresan TaMYB14 se pueden evaluar para determinar la producción de monómeros de epicatequina, catequina, y epigalocatequina, y galocatequina, así como combinaciones de dímeros y trímeros de estas unidades de base.

Ejemplo 6: Uso de la secuencia de ácido nucleico de MYB14 de la invención para producir taninos condensados en col silvestre (*Brassica oleracea*)

Materiales y métodos

Transformación de estirpes de *Brassica*

- 5 Semillas de *Brassica oleracea* var. *acephala* cv. Coleor (berza forrajera roja) y Gruner (berza forrajera verde) se hicieron germinar in vitro como se describe en Christey *et al.* (1997, 2006). Explantes de hipocotilo y pecíolos cotiledonarios procedentes de plántulas de 4-5 días se cocultivaron brevemente con un cultivo de *Agrobacterium tumefaciens* que se hizo crecer toda la noche en medio LB que contiene antibióticos antes de la dilución 1:10 en medio mínimo libre de antibióticos (7,6 mM de $(\text{NH}_4)_2\text{SO}_4$, 1,7 mM de citrato de sodio, 78,7 mM de K_2HPO_4 , 0,33 M de KH_2PO_4 , 1 mM de MgSO_4 , 0,2% de sacarosa) con crecimiento mínimo durante otras 4 h. Los explantes se cultivaron en medio a base de Murashige-Skoog (MS, Murashige y Skoog, 1962) con vitaminas B5 y 2,5 mg/l de BA, y se solidificaron con agar estándar Danisco 10 gm/L. Después de 3 días de cocultivo, los explantes se transfirieron al mismo medio con la adición de 300 mg/l de Timentina (SmithKline Beecham) y 15/l de kanamicina. Los explantes se transfirieron cada 3-4 semanas a medio de selección reciente. Los brotes verdes se transfirieron a medida que aparecieron a medio a base de Linsmaier-Skoog libre de hormonas (LS, Linsmaier y Skoog, 1965) que contiene 50 mg/l de kanamicina, y se solidificaron con agar estándar Danisco 10 gm/L. Los explantes se cultivaron en cápsulas de Petri altas (9 cm de diámetro, 2 cm de altura) cerradas herméticamente con cinta quirúrgica Micropore (3M). Los brotes se cultivaron en recipientes de plástico transparentes (98 mm, 250 ml, Vertex). Todas las manipulaciones de los cultivos vegetales se realizaron a 25°C con un fotoperíodo de 16 h/día, proporcionado por luces fluorescentes blancas frías, 20 $\mu\text{E}/\text{m}^2/\text{s}$.

RESULTADOS

Análisis molecular, cribado mediante DMACA, y bioquímica de *Brassica* transgénicaAnálisis molecular de *Brassica*

- 25 El ADN extraído de plantas de brassica transgénicas se evaluaron para determinar la integración del vector binario M14ApHZBAR. Las reacciones de PCR se llevaron a cabo usando conjuntos de cebadores diseñados para amplificar una porción del promotor 35S y la mayoría del gen. Los resultados de este análisis indicaron la integración del vector binario, que contiene el gen TaMyb14A, en el genoma de brassica (mostrado en la Figura 31).

Análisis mediante DMACA de *Brassica*

- 30 El análisis mediante DMACA se llevó a cabo en las plantas de Brassica como se describe para trébol en el Ejemplo 1. Las plántulas de brassica transgénicas que expresan TaMYB14A (bajo el control del promotor del virus del mosaico de la coliflor 35S) fueron indistinguibles de las plantas de tipo salvaje. La col no transformada de tipo salvaje de cualquier variedad de cultivo que no acumule de forma natural CT en tejidos vegetativos permaneció sin teñir. Sin embargo, se detectó CT en tejido foliar de plántulas de brassica transgénicas derivadas de las variedades de cultivo que acumulan antocianinas, como se evidencia mediante la tinción positiva por DMACA. La tinción no fue tan intensa como la observada para tabaco y tréboles. Por el contrario, las plántulas transgénicas derivadas de la variedad de cultivo verde de tipo salvaje nunca se tiñeron con DMACA.

- 40 Esto indica que el gen de MYB14 de *T. arvense* es capaz de activar una porción de los genes de la ruta de CT en brassica, pero puede requerir una ruta de antocianinas activa para la producción de CTs. Los ejemplos de la tinción mediante DMACA de hojas de brassica transgénicas se muestran en los dibujos más abajo (Figura 32). El colorante específico de CT, DMACA, tiñó el limbo de la hoja de la brassica transgénica (B a D) en comparación con hojas de tipo salvaje (A), que siempre están desprovistas de CT.

Análisis mediante HPLC/LCMS de Brassica

- 45 Se extrajeron flavonoides de plantas de Brassica transgénicas y de control de tipo salvaje, y se procesaron vía HPLC como se describe para trébol en el Ejemplo 2. Los resultados de esos análisis confirmaron la presencia de CT en extractos de hojas procedentes de una muestra de brassica transgénica. La transformación de brassica se realizó tanto con brassica de color verde normal así como con una estirpe de brassica que acumula antocianina. El análisis mediante HPLC detectó epicatequina en brassica de color verde pero no monómeros de tanino en las estirpes que acumulan antocianina. La brassica transgénica que sobreexpresa TaMYB14 que acumuló CTs en la hoja derivó de una estirpe que acumula antocianina. Solamente se detectaron monómeros de epicatequina en la estirpe transgénica, como se muestra en la Figura 33.

Ejemplo 6: Para demostrar la modificación de la población de taninos condensados por la variante de MYB14

Cualesquiera secuencias de MYB variantes, que se pueden identificar mediante métodos descritos aquí, se pueden evaluar para determinar su capacidad para alterar taninos condensados en plantas usando los métodos descritos en los Ejemplos 2 a 5.

De forma breve, las secuencias codificantes (tales como, pero sin limitarse a, aquellas de SEC ID NO: 56-64) de las secuencias variantes se pueden clonar en una expresión adecuada consistente (por ejemplo pHZBar, como se describe en el Ejemplo 2) y se puede transformar en una célula vegetal o planta. Un enfoque particularmente conveniente y relativamente simple es usar tabaco como planta de ensayo como se describe en el Ejemplo 3. El análisis mediante DMACA se puede usar como un ensayo rápido y conveniente para alteraciones en la producción de taninos condensados como se describe en el Ejemplo 1.

De esta manera, se puede confirmar rápidamente la función de variantes de MYB14 a la hora de regular la producción de taninos condensados.

También se pueden llevar a cabo análisis más detallados de los taninos condensados usando análisis mediante HPLC/LCMS como se describe en el Ejemplo 2.

Sumario de ejemplos

Los ejemplos demuestran claramente que el gen de MYB14 de la invención es útil para manipular la producción de flavonoides, específicamente taninos condensados, en un abanico de géneros de plantas, incluyendo tabaco (*Nicotiana tabacum*; familia *Solanaceae*), y en las legumbres trébol blanco (*Trifolium repens*; familia *Fabaceae*) y brassica (*Brassica oleracea*, familia *Brassicaceae*).

Se ha demostrado tanto el incremento como la disminución en la producción de taninos condensados usando los métodos y polinucleótidos de la invención.

REFERENCIAS

- Abrahams S, Lee E, Walker AR, Tanner GJ, Larkin PJ, Ashton AR (2003). The Arabidopsis TDS4 gene encodes leucoanthocyanidin dioxygenase (LDOX) and is essential for proanthocyanidin synthesis and vacuole development. *Plant Journal* 35: 624-636.
- Abrahams S, Tanner GJ, Larkin PJ, Ashton AR (2002). Identification and biochemical characterization of mutants in the proanthocyanidin pathway in Arabidopsis. *Plant Physiology* 130: 561-576.
- Aerts, RJ, Barry, TN y McNabb, WC (1999). Polyphenols and agriculture: beneficial effects of proanthocyanidins in forages. *Agric. Ecosyst. Env.* 75: 1-12.
- Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004). TT2, TT8, and TTG1 synergistically specify the expression of BANYULS and proanthocyanidin biosynthesis in Arabidopsis thaliana. *Plant J* 39: 366-380.
- Bingham, ET (1991). Registration of Alfalfa Hybrid Regen-Sy Germplasm for Tissue Culture and Transformation Research. *Crop Science* 31: 1098.
- Blaydes, DF (1966). Interaction of kinetin and various inhibitors in the growth of soybean tissue. *Physiologia Plantarum* 19:748-753.
- Blaxter, K.L., Clapperton, J.L. (1965). Prediction of the amount of methane produced by ruminants. *British Journal of Nutrition* 19: 511-522.
- Bogs J, Downey M, Harvey JS, Ashton AR, Tanner GJ, Robinson SP (2005). Proanthocyanidin synthesis and expression of genes encoding leucoanthocyanidin reductase and anthocyanidin reductase in developing grape berries and grapevine leaves. *Plant Physiology* 139: 652-663.
- Bogs J, Jaffe FW, Takos AM, Walker AR, Robinson SP (2007). The grapevine transcription factor VvMYBPA1 regulates proanthocyanidin synthesis during fruit development. *Plant Physiology* 143:1347-1361.
- Broun P. (2005). Transcriptional control of flavonoid biosynthesis: a complex network of conserved regulators involved in multiple aspects of differentiation in Arabidopsis. *Current Opinion in Plant Biology* 8:272-279.
- Burggraaf, V.T., Woodward, S.L., Woodfield, D.R., Thom, E.R., Waghorn, G.C. y Kemp, P.D. (2006). Morphology and agronomic performance of white clover with increased flowering and condensed tannin concentration. *New Zealand Journal of Agricultural Research* 49: 147-155.
- Caradus, J.R., Woodfield, D.R., Easton, H.S (2000). Improved grazing value of pasture cultivars for temperate environments. *Asian-Australasian Journal of Animal Sciences* 13 (SUPL. 1), p. 5-8.
- Christey, M.C., Sinclair, B.K., Braun, R.H. y Wyke, L. (1997). Regeneration of transgenic vegetable brassicas (*Brassica oleracea* and *B. campestris*) via Ri-mediated transformation. *Plant Cell Reports* 16: 587-593.
- Christey MC, Braun RH, Conner EL, Reader JK, White DWR, Voisey CR (2006). Cabbage white butterfly and diamond-back moth resistant Brassica oleracea plants transgenic for cry1Ba1 or cry1Ca5. *Acta Horticulturae* 706: 247-253.

- Clark, H. (2001). Ruminant Methane Emissions: A Review of the Methodology Used for National Inventory Estimations. A Report Prepared for the Ministry of Agriculture and Forestry, Nueva Zelandia.
- Choreo y Goodman, Acc. Chem. REs., (1993) 26 266-273.
- DairyInsight: Strategic Framework for Dairy Farming's Future, 2005-2015.
- 5 Damiani F, Paolocci F, Cluster PD, Arcioni S, Tanner GJ, Joseph RG, Li YG, de Majnik J, Larkin PJ (1999). The maize transcription factor Sn alters proanthocyanidin synthesis in transgenic *Lotus corniculatus* plants *Australian Journal Of Plant Physiology* 26: 159-169.
- Davies KM, Schwinn KE (2003). Transcriptional regulation of secondary metabolism. *Functional Plant Biology* 30:913-925.
- 10 de Majnik, J. Weinman, J., Djordjevic, M. Rolfe, MB. Tanner, G. Joseph, RG. Larkin PJ (2000). Anthocyanin regulatory gene expression in transgenic white clover can result in an altered pattern of pigmentation. *Australian Journal of Plant Physiology* 27:659-667.
- I, Nesi N, Perez P, Devic M, Grandjean O, Caboche M, Lepiniec L (2003). Proanthocyanidin-accumulating cells in *Arabidopsis testis*: regulation of differentiation and role in seed development. *Plant Cell* 15: 2514-2531.
- 15 Debeaujon I, Peeters AJM, Léon-Kloosterziel KM, Koornneef M (2001). The TRANSPARENT TESTA12 gene of *Arabidopsis* encodes a multidrug secondary transporter-like protein required for flavonoid sequestration in vacuoles of the seed coat endothelium. *Plant Cell* 13: 853-871.
- Ditta, G., Stanfield, S., Corbin, D., y Helsinki, S.R. (1980). Broad host range cloning system for gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proceedings of the National Academy of Sciences USA* 77: 7347-7351.
- 20 Dixon RA, Lamb CJ, Masoud S, Sewalt VJH, Paiva NL (1996). Metabolic engineering: prospects for crop improvement through the genetic manipulation of phenylpropanoid biosynthesis and defense responses - a review. *Gene* 179: 61-71.
- Dixon RA, Xie DY, Sharma SB (2005) Proanthocyanidins - a final frontier in flavonoid research? *New Phytologist* 165: 9-28.
- 25 Douglas GB, Wang Y, Waghorn GC, Barry TN, Purchas RW, Foote AG, Wilson GF (1995). Liveweight Gain And Wool Production Of Sheep Grazing *Lotus-Corniculatus* And Lucerne (*Medicago-Sativa*). *New Zealand Journal Of Agricultural Research* 38: 95-104.
- Ellison, N.W., Liston, A., Steiner, J.J., Williams, W.M., Taylor, N.L (2006). Molecular phylogenetics of the clover genus (*Trifolium-Leguminosae*) *Molecular Phylogenetics and Evolution* 39; 688-705.
- 30 Fay MF, Dale PJ (1993). Condensed Tannins in *Trifolium* species and their significance for taxonomy and plant breeding. *Genetic resources and Crop Evolution* 40:7-13.
- Freidinger, R.M., Perlow, D.S., Veber, D.F., *J. Org. Chem.* 1982, 59, 104-109.
- Gallop, M.A., Barrett, R.W., Dower, W.J., Fodor, S.P.A. y Hogan, Jr., J.C. (1997). *Nature Biotechnology*, 15 328-330.
- 35 Gleave AP (1992). A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Molecular Biology* 20: 1203-1207.
- Helliwell, C y Waterhouse, P (2003). Constructs and methods for high-throughput gene silencing in plants. *Methods* 30: 289-295.
- 40 Horsch RB, Fry JE, Hoffmann NL, Eichholtz D, Rogers SG, Fraley RT. (1985). A simple and general method for transferring genes into plants. *Science*;227:1229-1231.
- Jones, W.T., Broadhurst, R.B. y Lyttleton, J.W. (1976). The condensed tannins of pasture legume species. *Phytochemistry* 15: 1407-1409.
- Kingston-Smith AH, Thomas HM (2003). Strategies of plant breeding for improved rumen function *Annals of Applied Biology* 142:13-24.
- 45 Li, YG y Tanner G, Larkin P (1996). The DMACA-HCl Protocol and the Threshold Proanthocyanidin Content for Bloat Safety in Forage Legumes. *Journal of the Science of Food and Agriculture* 70 (1996) 98-101.
- Linsmaier, E.M. y Skoog, F. (1965). Organic growth factor requirements of tobacco tissue cultures. *Physiologia Plantarum*. 18:100-127.

- McKenna, P.B (1994). The occurrence of anthelmintic resistant sheep nematodes in the southern North Island of New Zealand. *NZ Veterinary Journal*. 42: 151-152.
- 5 McMahon LR, McAllister TA, Berg BP, Majak W, Acharya SN, Popp JD, Coulman BE, Wang Y, Cheng KJ (2000). A review of the effects of forage condensed tannins on ruminal fermentation and bloat in grazing cattle. *Canadian Journal of Plant Science* 80: 469-485.
- Marten, G. C., Ehle, F.R. y Ristau, E. A. (1987). Performance and photosensitization of cattle related to forage quality of four legumes. *Crop Science* 27: 138-145.
- Mehrtens F, Kranz H, Bednarek P, Weisshaar B (2005). The Arabidopsis transcription factor MYB12 is a flavonol-specific regulator of phenylpropanoid biosynthesis. *Physiologia Plantarum*. 138: 1083-1096.
- 10 Miyake K, Ito T, Senda M, Ishikawa R, Harada T, Niizeki M, Akada S (2003). Isolation of a subfamily of genes for R2R3-MYB transcription factors showing up-regulated expression under nitrogen nutrient-limited conditions. *Plant Molecular Biology* 53: 237-245.
- Molan. A.L. Waghorn, G.C., McNabb, W.C. (2001). Effect of condensed tannins on egg hatching and larval development of *Trichostrongylus colubriformis* in vitro. *The Veterinary Record* 150: 65-69.
- 15 Murashige T y Skoog F (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiologia Plantarum* 15(3): 473-497.
- Nagai, U., Sato, K. *Tetrahedron Lett.* 1985, 26, 647-650.
- Nesi N, Debeaujon I, Jond C, Pelletier G, Caboche M, Lepiniec L (2000). The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of DFR and BAN genes in Arabidopsis siliques. *Plant Cell* 12: 1863-1878.
- 20 Nesi N, Debeaujon I, Jond C, Stewart AJ, Jenkins GI, Caboche M, Lepiniec L (2002). The TRANSPARENT TESTA16 locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell* 14: 2463-2479.
- Nesi N, Jond C, Debeaujon I, Caboche M, Lepiniec L (2001). The Arabidopsis TT2 gene encodes an R2R3 MYB domain protein that acts as a key determinant for proanthocyanidin accumulation in developing seed. *Plant Cell* 13: 2099-2114.
- 25 Niezen, J.H., Waghorn, T.S., Charleston, W.A.G. y Waghorn, G.C. (1995). Growth and gastrointestinal nematode parasitism in lambs grazing either lucerne (*Medicago sativa*) or sulla (*Hedysarum coronarium*) which contains condensed tannins. *J. Agric. Sci. (Cam)* 125, p. 281-289.
- 30 Niezen, J.H., Waghorn, T.S., Waghorn, G.C. y Charleston, W.A.G. (1993) Internal parasites and lamb production - a role for plants containing condensed tannins?. *Proc. NZL. Soc. Anim. Prod.* 53, pp. 235-238.
- Olson *et al.*, (1993) *J. Med. Chem.*, 36 3039-3049.
- Pang Y, Peel GJ, Wright E, Wang Z, Dixon RA (2007). Early steps in proanthocyanidin biosynthesis in the model legume *Medicago truncatula*. *Plant Physiology* 145(3):601-615.
- 35 Pfeiffer J, Kuhnel C, Brandt J, Duy D, Punyasiri PAN, Forkmann G, Fischer TC (2006). Biosynthesis of flavan 3-ols by leucoanthocyanidin 4-reductases and anthocyanidin reductases in leaves of grape (*Vitis vinifera* L.), apple (*Malus x domestica* Borkh.) and other crops. *Plant Physiology and Biochemistry* 44: 323-334.
- Puchala, R., Min, B. R., Goetsch, A. L. y Sahlu, T. (2005). The effect of a condensed tannin-containing forage on methane emission by goats. *Journal of Animal Science* 83:182-186.
- 40 Ray H, Yu M, Auser P, Blahut-Beatty L, McKersie B, Bowley S, Westcott N, Coulman B, Lloyd A, Gruber MY (2003). Expression of Anthocyanins and Proanthocyanidins after Transformation of Alfalfa with Maize Lc. *Plant Physiology*, 132: 1448-1463.
- Robbins MP, Paolocci F, Hughes JW, Turchetti V, Allison G, Arcioni S, Morris P, Damiani F (2003). Sn, a maize bHLH gene, modulates anthocyanin and condensed tannin pathways in *Lotus corniculatus*. *Journal of Experimental Botany* 54:381: 239-248., DOI: 10.1093/jxb/erg022
- 45 Rumbaugh, M.D. (1985). Breeding bloat-safe cultivars of bloat-causing legumes. En: Barnes, R.F., Ball, P.R., Bringham, R.W., Martin, G.C., Minson, D.J. (Eds.), *Forage Legumes for Energy-Efficient Animal Production*. USDA, Washington. Proc. Bilateral Workshop, Palmerston North, NZ, Abril de 1984, p. 238-245.
- Samac, DA (1995). Strain specificity in transformation of alfalfa by *Agrobacterium tumefaciens*. *Plant Cell, Tissue and Organ Culture* 43: 271-277.
- 50

- Sanger F, Nicklen S, Coulson AR (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences USA* 74: 5463-5467.
- Schenk, RU y Hildebrandt, AC (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Canadian Journal of Botany* 50: 199-204.
- 5 Sharma, S.B. y Dixon, R.A. (2005). Metabolic engineering of proanthocyanidins by ectopic expression of transcription factors in *Arabidopsis thaliana*. *Plant Journal* 44:62- 75.
- Debeaujon Smythe, M.L., von Itzstein, M., *J. Am. Chem. Soc.* 1994, 116, 2725-2733.
- Stracke R, Werber M, Weisshaar B (2001). The R2R3-MYB gene family in *Arabidopsis thaliana*. *Current Opinion in Plant Biology* 4: 447-456.
- 10 Sykes. A.R y Coop. R.L (2001). Interaction between nutrition and gastrointestinal parasitism in sheep New Zealand *Veterinary Journal.* 49: 222-226.
- Tanner GJ, Francki KT, Abrahams S, Watson JM, Larkin PJ, Ashton AR (2003). Proanthocyanidin biosynthesis in plants - Purification of legume leucoanthocyanidin reductase and molecular cloning of its ADNc. *Journal of Biological Chemistry* 278:31647-31656.
- 15 Tanner GJ, Moore AE, Larkin PJ (1994). Proanthocyanidins Inhibit Hydrolysis Of Leaf Proteins By Rumen Microflora In-Vitro *British Journal Of Nutrition* 71: 947-958.
- Voisey, C.R.; White, D.W.R.; Dudas, B.; Appleby, R.D.; Ealing, P.M.; Scott, A.G. (1994). *Agrobacterium*-mediated transformation of white clover using direct shoot organogenesis. *Plant Cell Reports* 13: 309-314.
- 20 Waghorn, G.C., Douglas, G.B., Niezen, J.H., McNabb, W.C y Foote, A.G (1998). Forages with condensed tannins-their management and nutritive value for ruminants. *Proceedings of the New Zealand Grasslands Association* 60: 89-98.
- Walker AR, Davison PA, Bolognesi-Winfield AC, James CM, Srinivasan N, Blundel TL, Esch JJ, Marks MD, Gray JC (1999). The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *Plant Cell* 11: 1337-1349.
- 25 Walter C, Grace LJ, Wagner A, White DWR, Walden AR, Donaldson SS, Hinton H, Gardner RC, Smith DR (1998). Stable transformation and regeneration of transgenic plants of *Pinus radiata* D.Don. *Plant Cell Reports* 17: 460-469.
- Wei YL, Li JN, Lu J, Tang ZL, Pu DC, Chai YR (2007). Molecular cloning of *Brassica napus* TRANSPARENT TESTA 2 gene family encoding potential MYB regulatory proteins of proanthocyanidin biosynthesis. *Molecular Biology Reports* 34:105-120.
- 30 Winkel-Shirley B (2001). Flavonoid biosynthesis: a colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiology* 126: 485-493.
- Winkel-Shirley, B. (2002). A mutational approach to dissection of flavonoid biosynthesis in *Arabidopsis*. En *Recent Advances in Phytochemistry: Proceedings of the Annual Meeting of the Phytochemical Society of North America*, Vol. 36, J.T. Romeo, ed (New York: Elsevier), p. 95-110.
- 35 Woodfield, D., McNabb, W., Kennedy, L., Cousins, G. y Caradus, J. (1998). Floral and foliar content in white clover. *Proceedings of the 15th Trifolium Conference*, P.19.
- Woodward, S. L., Waghorn, G. C., Ulyatt, M. J. y Lassey. K. R. (2001). Early indications that feeding *Lotus* will reduce methane emission from ruminants. *Proceedings New Zealand Society of Animal Production* 61:23-26.
- 40 Xie DY, Sharma SB, Dixon RA (2004). Anthocyanidin reductases from *Medicago truncatula* and *Arabidopsis thaliana*. *Archives Of Biochemistry and Biophysics* 422: 91-102.
- Xie DY, Sharma SB, Paiva NL, Paiva NL, Ferreira D, Dixon RA (2003). Role of anthocyanidin reductase, encoded by BANYULS in plant flavonoid biosynthesis. *Science* 299: 396-399.
- Xie DY, Sharma SB, Wright E, Wang ZY, Dixon RA (2006). Metabolic engineering of proanthocyanidins through co-expression of anthocyanidin reductase and the PAP1 MYB transcription factor. *Plant Journal* 45: 895-907.
- 45 Yoshida, K, Iwasaka, R, Kaneko T, Sato s, Tabata, S. Sakuta M (2008). Functional differentiation of *Lotus japonicus* TT2s, R2R3 MYB transcription factors comprising a multigene family. *Plant Cell Physiology* 49:157-169.

LISTADO DE SECUENCIAS

<110> GRASSLANZ TECHNOLOGY LIMITED

ES 2 537 342 T3

<120> NUEVOS GENES IMPLICADOS EN BIOSÍNTESIS

<130> 609676 HCF

<160> 69

<170> PatentIn version 3.3

5 <210> 1

<211> 1243

<212> ADN

<213> *Trifolium arvense*

<400> 1

```

gaattcgccc ttaagcagtg gtatcaacgc agagtacgcg ggggaagtta ttttaatttta      60
tctacatcaa acacttcaag aggttggaaat acaagacaga ctaattaaga ataacatcaa      120
tggggagaag cccttgttgt gcaaaggaag gcttgaatag aggtgcttgg acaactcaag      180
aagacaaaat cctcactgaa tacattaagc tccatggtga aggaaaatgg agaaaccttc      240
caaaaagagc agatttaaaa agatgtggaa aaagttgtag acttagatgg ttgaattatc      300
taagaccaga tattaagcga ggtaatatat ccccggatga agaagaactt attatccgac      360
ttcaciaaact actcggaaac agatggtctc taatagccgg aagacttcca gggcgaacag      420
acaatgaaat aaagaactac tggaacacaa atttaggaaa aaaggttaag gatcttaatc      480
aacaaaacac caacaattct tctcctaacta aactttctgc tcaacaaaa aatgcaaaga      540
tcaaacagaa acagatcaat cctaagccaa tgaagccaaa ctcaaagtgt gtccgtacaa      600
aagctaccaa gtgttctaag gtattgttca taaactcact ccccaactca ccaatgcatg      660
atitgcagaa caaagctgag gcagagacaa caacaaagcc atcaatgctg gttgatggtg      720
tggctagtga ttcaatgagt aacaacgaaa tggaacacgg ttatggattt ttgtcatttt      780
gcgatgaaga gaaagaacta tccgcagatt tgctagaaga ttttaacatc gcggatgata      840
tttgcttata tgaacttttg aactctgatt tctcaaatgc gtgcaatttc gattacaatg      900
atctattgtc accttgttcg gaccaaactc aaatgttctc tgatgatgag atttctcaaga      960
attggacaca atgtaacttt gctgatgaga caaatgtgtc caacaacctt cattcttttg     1020
cttcctttct tgaatccagt gaggaagtac taggagaatg ataataaaaa ttcattttcc     1080
aataaaatta actactctag gttttttttt ttttttttta atttcaattt catgttaggg     1140
tggtttaata aataaatata ttctatggtt taatattgca aaaaaaaaaa aaaaaaaaaa     1200
aaaaagtact ctgcttgat accactgctt aagggcgaat tcc                               1243

```

10

<210> 2

<211> 1290

<212> ADN

<213> *Trifolium arvense*

15 <400> 2

```

gaattcgccc ttaggttggga atacaagaca gactaattaa gaataacatc aatggggaga      60

```

ES 2 537 342 T3

agcccttggt gtgcaaagga aggcttgaat agaggtgctt ggacaactca agaagacaaa 120
atcctcactg aatacattaa gctccatggt gaaggaaaat ggagaaacct tccaaaaaga 180
gcaggttcat tcattctagt atcttgcaat tatagatcaa tcactttcat acttttgttt 240
gcttataaat tttcttgcat tttttcttca attttccatg tgaaatgcaa attactagta 300
cattattatg gatatgtttt tgcaaatatg tgtatgccat gcaggtttaa aaagatgcgg 360
aaaaagttgt agacttagat ggttgaatta tctaagacca gatattaagc gaggtaatat 420
atcctcggat gaagaagaac ttatcatcag acttcacaaa ctactcggaa acaggtaaaa 480
gtaccgacat aatcactaac ttattaacat ttatctataa tttgtttttt ttgacaatta 540
gtactactaa tttattttta taatgtgtgc taatttgctt tgcctttaat ttgtggtaga 600
tgggtctctaa tagccggaag acttccagga cgaacagaca atgaaataaa gaactactgg 660
aacacaaatt taggaaaaaa ggttaaggat cttaatcaac aaaacaccaa caattcttct 720
cctactaac tctctgctca accaaaaaat gcaaagatca aacagaaaca gatcaatcct 780
aagccaatga agccaaactc aatgtttgtc cgtacaaaag ctaccaagtg ttctaaggta 840
ttgttcataa actcactccc caactcacca atgcatgatt tgcagaacaa agctgaggca 900
gagacaacaa caaagccatc aatgctggtt gatggtgtgg ctagtgattc aatgagtaac 960
aacgaaatgg aacacggtta tggatttttg tcattttgcy atgaagagaa agaactatcc 1020
gcagatttgc tagaagattt taacatcgcg gatgatattt gcttatctga acttttgaac 1080
tctgatttct caaatgctg caatttcgat tacaatgatc tattgtcacc ttgttcggac 1140
caaaactcaa tgttctctga tgatgagatt ctcaagaatt ggacacaatg taactttgct 1200
gatgagacaa atgtgtccaa caaccttcat tcttttgctt cctttcttga atccagttag 1260
gaagtactag gagaatgaaa gggcgaattc 1290

<210> 3

<211> 1296

<212> ADN

5 <213> *Trifolium arvense*

<400> 3

gaattcgccc ttaggttggga atacaagaca gactaattaa gaataacatc aatggggaga 60
agcccttggt gtgcaaagga aggcttgaat agaggtgctt ggacaactca agaagacaaa 120
atcctcactg aatacattaa gctccatggt gaaggaaaat ggagaaacct tccaaaaaga 180
gcaggttcat tcattctgta tcttacaatt atagattaac cactttcata cttttgttt 240
cttataaatt ttcttgattt ttttcttcca ttttcatga gaaatgcaaa ttactagtac 300
attattatgg acatgttttg gcaaatatgt ttatgccatg cagattttaaa aagatgtgga 360
aaaagttgta gacttagatg gttgaattat ctaagaccag atattaagcg aggtaatata 420
tccccggatg aagaagaact tattatccga cttcacaaac tactcggaaa caggtaaagt 480
cctaacataa tcactaactt attaacgttt gtctataatt tgtttttttt gaccattagt 540

ES 2 537 342 T3

actactaatt taattttaca atgtgtgcta atttgcttgt ctttaatttg tggtagatgg 600
tctctaatag ccggaagact tccagggcga acagacaatg aaataaagaa ctactggaac 660
acaaatttag gaaaaaagg taaaggatcct gatcaacaaa acaccaacaa ttcttctcct 720
actaaactct ctgctcaacc aaaaaatgca gagatcaaac agaaacagat caatcctaag 780
ccaaactcat atgtgtgccg tacaaaagct accaagtgtt ctaaggtatt gttcataaac 840
tcacccccca actcaccacc aatgcatgat ttgcagagca aagctgaggc agagacaaca 900
acaacaacaa agccatcaat gccatcaatg ctggttgatg gtgtggctag tgattcaatg 960
agtaacaacg aatggaatg cggtaatgga tttttgtcat tttgcatga agagaaagaa 1020
ctatccgcag atttgctaga agattttaac atcgcggatg atatttgctt atctgaattt 1080
ctaaacttcg atttctcaaa tgcgtgcgat atcgattaca atgatctatt gtcgccttgt 1140
tcggaccaa ctcaaatggt ccctgatgat gagattctaa agaattggac acaatgtaac 1200
tttgctgatg agacaaatgt gtccaacaac cttcagtctt ctgcttccct tcttgaatcc 1260
agtgaggaag tactaggaga atgaaagggc gaattc 1296

<210> 4

<211> 1239

<212> ADN

5 <213> *Trifolium affine*

<400> 4

gaattcgccc ttatggggag aagcccttgt tgtgcgaagg aaggcttgaa tagaggtgct 60
tggacaactc aagaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa 120
tggagaaacc ttccaaaaag agcaggttca ttcattctgt atcttacaat tatagattaa 180
ccactttcat acttttgttt gcttataaat tttcttgtat tttttcttcc atttttcatg 240
agaaatgcaa attactagta cattattatg gacatgtttt tgcaaatatg tttatgccat 300
gcaggtttaa aaagatgtgg aaaaagtgtg agacttagat ggttgaatta tctaagacta 360
gatattaagc gaggtaatat atcctcggat gaagaagaac ttatcatccg acttcacaaa 420
ttactcggaa acaggtaaag tcctaacata atcactaact tattaacggt tgtctataac 480
ttgttttttt gacaattagt actactaatt taattttata atgtgtgcta atttgcttgt 540
ctttaatttg tggtagatgg tctctaatag ccggaagact tccaggacga acagacaatg 600
aaataaagaa ctactggaac acaaatttag gaaaaaagg taaaggatcct aatcaagaaa 660
acaccaacaa ttcttctcct actaaacttt ctgctcaact aaaaaatgca aagatcaaac 720
agaaacagat caatcctaag ccaatggagc caaacacaaa tgttgtccgt acaaaagcta 780
ccaagtgttc taaggcattg ttcataaact cccccccaa ctaccacca atgcatgatt 840
tgcagaacaa agctgaggca gagacaacaa caaagtcac aatgccatca atgctggttg 900
atggcgtggc tagtgattca atgagtaaca acgaaatgga atacggatgat ggatttgttt 960
cattttgcga tgacgataaa gaactatccg cagatttgct agaagatttt aacatctcgg 1020
atgatatttg cttatccgaa tttctaaact tcgatttctc aaatgcgtgc aatttcgatt 1080
acaacgatct attgtcgctt tgttcggacc aaacacaaat gttctctgat gatgagattc 1140
tcaagaattc gacaccatgt aactttgctg ctgagacaaa ttatgtgtcc aacaaccaat 1200
ccagtgagga agtactagga gaatgaaagg gcgaattct 1239

ES 2 537 342 T3

<210> 5

<211> 933

<212> ADN

<213> *Trifolium affine*

5 <400> 5

```

atggggagaa gcccttgttg tgcgaaggaa ggcttgaata gaggtgcttg gacaactcaa      60
gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt     120
ccaaaaagag caggtttaaa aagatgtgga aaaagttgta gacttagatg gttgaattat     180
ctaagactag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatccga     240
cttcacaaat tactcggaaa cagatggtct ctaatagccg gaagacttcc aggacgaaca     300
gacaatgaaa taaagaacta ctggaacaca aatntagga   aaaaggtaa  ggatcttaat     360
caagaaaaca ccaacaattc ttctcctact aaactttctg ctcaactaaa aaatgcaaag     420
atcaaacaga aacagatcaa tcctaagcca atggagcca  actcaaatgt  tgtccgtaca     480
aaagctacca agtgttctaa ggcatgttc  ataaactcac cccccaactc accaccaatg     540
catgatttgc agaacaaagc tgaggcagag acaacaaca  agtcatcaat  gccatcaatg     600
ctggttgatg gcgtggctag tgattcaatg agtaacaacg aaatggaata  cggtgatgga     660
tttgtttcat tttgcatga cgataaagaa ctatccgag  atttgctaga  agattttaac     720
atctcggatg atatttgctt atccgaattt ctaaacttcg atttctcaa  tgcgtgcaat     780
ttcgattaca acgatctatt gtcgccttgt tcggaccaa  cacaaatgtt  ctctgatgat     840
gagattctca agaattcgac accatgtaac tttgctgctg agacaaatta  tgtgtccaac     900
aaccaatcca gtgaggaagt actaggagaa  tga                                     933

```

<210> 6

<211> 1238

<212> ADN

10 <213> *Trifolium affine*

<400> 6

```

gaattcggcc ttatggggag aagcccttgt tgtgcaaagg aaggcttgaa tagagggtgct     60
tggacaactc aagaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa     120
tggagaaacc ttccaaaag  agcaggttca ttcatctgt  atcttacaat  tatagattaa     180
ccactttcat acttttgttt tcttataaat tttcttgtat tttttcttcc atttttcatg     240
agaaatgcaa attactagta cattattatg gacatgtttt tgcaaata  tttatgccat     300
gcaggtttaa aaagatgtgg aaaaagttgt agacttagat ggttgaatta  tctaagacca     360
gatattaagc gaggtaatat atcctcggat gaagaagaac ttatcatccg acttcacaaa     420
ctactcggaa acaggtaaag tcataacata atcattaatt tattaacggt  tatctataat     480

```

ES 2 537 342 T3

ttgttttttt gacaattatc actacaaatt taattttata atgtgcgcta atttgcttgt 540
 ctttaatttg tggtagatgg tctctaatag ccggaagact tccagggcga acagacaatg 600
 aaataaagaa ctactggaac acaaatttag gaaaaaagg taaggatctt aatcaagaaa 660
 acaccaacaa ttcttctcct actaaacttt ctgctcaact aaaaaatgca aagatcaaac 720
 agaaacagat caatcctaag ccaatggagc caaactcaa tgttgtccgt acaaaagcta 780
 ccaagtgttc taaggcattg ttcataaact cccccccaa ctcaccacca atgcatgatt 840
 tgcagaacaa agctgaggca gagacaacaa caaagtcac aatgccatca atgctgggtg 900
 atggcgtggc tagtgattca atgagtaaca acgaaatgga atacggtgat ggatttgttt 960
 cttttgcca tgacgataaa gaactatccg cagatttgct agaagatttt aacatctcgg 1020
 atgatatttg cttatccgaa tttctaaact tcgatttctc aaatgctgac aatttcgatt 1080
 acaacgatct attgtcgcct tgttcggacc aaacacaaat gttctctggt gatgagattc 1140
 tcaagaattc gacacaatgt aactttgctg ctgagacaaa ttatgtgtcc aacaaccaat 1200
 ccagtgagga agtactagga gaatgaaagg gcaattc 1238

<210> 7

<211> 1252

<212> ADN

5 <213> *Trifolium occidentale*

<400> 7

gaattcggcc ttatggggag aagcccttgt tgtgcaaagg aaggcttgaa tagagggtgct 60
 tggacaactc aagaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa 120
 tggagaaacc ttccaaaaag agcaggttca ttcattctag tatcttgcaa ttatagatca 180
 atcactttca tacttttgtt tgcttataaa ttttcttgca ttttttctc aattttccat 240
 gtgaaatgca aattactagt acattattat ggatatgttt ttgcaaata gtgtatgcca 300
 tgcgaggttt aaaaagatgc ggaaaaagt gttagactag atggttgaat tatctaagac 360
 cagatattaa gcgaggaat atatcctcgg atgaagaaga acttatcatc agacttcaca 420
 aactactcgg aaacaggtaa aagtaccgac ataactacta acttattaac atttatctat 480
 aatttgtttt ttttgacaat tagtactact aatttaattt tataatgtgt gctaatttgc 540
 tttgccttta atttgtggta gatggctct aatagccgga agacttccag gacgaacaga 600
 caatgaaata aagaactact ggaacacaaa tttaggaaaa aaggtttaagg atcttaatca 660
 acaaacacc aacaagtctt ctctactaa actctctgct caaccaaaa atgcaaagat 720
 caaacagaaa cagatcaatc ctaagccaat gaagccaaac tcaaagtgtg tccgtacaag 780
 agctaccaag tgttctaagg tattgttcat aaactcactc cccaactcac caatgcatga 840
 tttgcagaac aaagctgagg cagagacaac acaaagcca tcaatgctgg ttgatgggtg 900
 ggctagtgat tcaatgagta acaacgaaat ggaacacggg tatggatttt tgtcattttg 960
 cgatgaagag aaagaactat ccgcagattt gctagaagat tttaacatcg cggatgatat 1020
 ttgcttatct gaacttttga actctgattt ctcaaatgcg tgcaatttcg attacaatga 1080
 tctattgtcm ccttgttcgg accaaactca aatgttctct gatgatgaga ttctcaagaa 1140
 ttggacacaa tgtaactttg ctgatgagac aaatgtgtcc aacaacctt attcttttgc 1200
 ttcctttctt gaatccagtg aggaagtact aggagaatga aagggcgaat tc 1252

ES 2 537 342 T3

<210> 8

<211> 1164

<212> ADN

<213> *Triforium occidentale*

5 <400> 8

```

gaattcgccc ttatggggag aagcccttgt tgtgcaaagg aaggtttgaa tagaggtgct      60
tggacagctc atgaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa    120
tggagaaacc ttccaaaag agcaggttca ttcattctgt atcttactat ttatagatca    180
ataatcactt tcatgtattt ttttccctc cattttccat tagaaatgca aattaatagt    240
acattattat ggacatgttt ttccaggttt aaaaagatgt ggaaaaagtt gtagacttag    300
atggttgaat tatcttagac cagatattaa gagaggtaat atatcgtccg atgaagaaga    360
acttatcatt agacttcaca aactacttgg aaaccggtaa agtatcgaca taatcactaa    420
cttactaaca tttgtttata atgtgtacta attgcgattc ctttgatttg tggtagatgg    480
tctctaatag ccggaagact tccagggcga acagacaatg aaataaaaaa ttactggaac    540
acgaatttag gaaaaaagggt taaggatctt aatcaacaaa acaccaacaa ttcttctcct    600
actaaacctt ctgctcaacc aaaaaatgca aagatcaaac agaaacaaca gatcaataat    660
cctaagccaa tgaagccaaa ctcgaatggt gtccgtacaa aagctacca atgttctaag    720
gtattgttca taaactcacc accaatgcat aatttgcaga acaaagctga ggcagagaca    780
aaaacaaga catcaatggt ggttaatggt gtagctagtg attcaatgag taacaacgaa    840
atggaacgag gtaatggatt tttgtcattt cgcgatgaag agaaagaact atccgctgat    900
ttgctagatg attttaacat cgcggatgac atttgcttat ccgaatttct aaactccgat    960
ttctcaaatg cgtgcaattt cgattacaat gatctattgt caccttgttc ggatcaaac    1020
caaatgttct ctgatgatga gattctcaag aattggacac aatgtaactt tgctgatgag    1080
acaaatgtgt ccaacaacct tcattctttt gcttcctttc tcgaatccag tgaggaagta    1140
ctaggagaat gaaagggcga attc                                     1164

```

<210> 9

<211> 1205

<212> ADN

10 <213> *Trifolium repens*

<400> 9

```

gaattcgccc ttatggggag aagcccttgt tgtgcaaaag aaggcttgaa tagaggtgct      60
tggacagctc atgaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa    120
tggagaaacc ttccaaaag agcaggttca ttcattctgt atcttactat tatagatcaa    180

```

ES 2 537 342 T3

taatcacttt cacacttttt ttttactta taaattttca tgtatttttt cttccatttt 240
 ccattagaaa tgcaaattaa tagtacatta ttatggacat gttttttcaa aaatgtgtat 300
 tccatgcagg tttaaaaaga tgtggaaaaa gttgtagact aagggtggtg aattatctta 360
 gaccggatat taagagagg t aatataatcgt cggatgaaga agaacttatc attagacttc 420
 acaaactact cggaaaccgg taaagtatcg acataatcac tgacttacta acatttgttt 480
 ataatgtgtg ctaattgctc ttcctttgat ttgtggtaga tggctctctaa tagccggaag 540
 acttccaggg cgaacagaca atgaaataaa gaactactgg aacacaaatt taggaaaaaa 600
 agttaaggat cttaatcaac aaaacaccaa caattcttct cctactaaac cttctgctca 660
 accaaaaaat gcaaatatca aacagaaaca acagatcaat cctaagccaa tgaagccaaa 720
 ctgcaatggt gtccgtacaa aagctaccaa atgttctaag gtattgttca taaactcacc 780
 accaatgcat aatttgcaga acaaagctga ggcagagaca aaaacaaagc cattaatgct 840
 ggttaatggt gtagctagtg attcaatgag taacaacgaa atggaacgcg gtaatggatt 900
 tttgtcattt tgcgacgaag agaaagaact atccgcagat ttgctagatg attttaacat 960
 cgcggtgat atttgcattat ctgaatttct aaactccgat ttctcaaatg cgtgcaattt 1020
 cgattacaat gatctattgt cgccttgttc ggatcaaaact caaatgttct ctgatgatga 1080
 gattctcaag aattggacac aatgtaactt tgctgatgag acaaatgtgt ccaacaacct 1140
 taattctttt gcttcttttc tcgaatccag tgaggaagta ctaggagaat gaaagggcga 1200
 attct 1205

<210> 10

<211> 1202

<212> ADN

5 <213> *Trifolium repens*

<400> 10

gaattcgccc ttatggggag aagcccttgt tgtgcaaaaag aaggcttgaa tagaggtgct 60
 tggacagctc atgaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa 120
 tggagaaacc ttccaaaag agcaggttca ttcattctgt atcttactat tatagatcaa 180
 taatcacttt cacacttttt tttacttata aattttcatg tttttttct tccattttcc 240
 attagaaatg caaattaata gtacattatt atggacatgt tttttcaaaa atgtgtattc 300
 catgcagggt taaaaagatg tggaaaaagt tgtagactaa ggtggttgaa ttatcttaga 360
 ccggatatta agagaggtaa tatatcgtcg gatgaagaag aacttatcat tagacttcac 420
 aaactactcg gaaaccggta aagtatcgac ataactacta acttactaac atttgtttat 480
 aatgtgtgct aattgctctt cctttgattt gtggtagatg gtctctaata gccggaagac 540
 ttccagggcg aacagacaat gaaataaaga actactggaa cacaaattta ggaaaaaaag 600
 ttaaggatct taatcaacaa aacaccaaca attcttctcc tactaaacct tctgctcaac 660
 caaaaaatgc aaatatcaaa cagaacaac agatcaatcc taagccaatg aagccaaact 720
 cgaatgttgt ccgtacaaaa gctaccaaat gttctaagggt attgttcata aactcaccac 780

ES 2 537 342 T3

caatgcataa tttgcagaac aaagctgagg cagayacaaa aacaaaylca ttaatgctgg 840
 ttaatggtgt agctagtgat tcaatgagta acaacgaaat ggaacgcggt aatggatttt 900
 tgtcattttg cgacgaagag aaagaactat ccgcagattt gctagatgat tttaacatcg 960
 cggatgatat ttgcttacct gaatttctaa actccgattt ctcaaatgcg tgcaatttcg 1020
 attacaatga tctattgtcg ccttgctcgg atcaaaactca aatgttctct gatgatgaga 1080
 ttctcaagaa ttggacacaa tgtaactttg ctgatgagac aaatgtgtcc aacaacctta 1140
 attcttttgc ttcttttctc gaatccagtg aggaagtact aggagaatga aagggcgaat 1200
 tc 1202

<210> 11

<211> 1203

<212> ADN

5 <213> *Trifolium repens*

<400> 11

gaattcgccc ttatggggag aagcccttgt tgtgcaaaag aaggcttgaa tagagggtgct 60
 tggacagctc atgaagacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa 120
 tggagaaacc ttccaaaaag agcaggttca ttcattctgt atcttactat tatagatcaa 180
 tagtcacttt cacacttttt ttttacttat aaattttcat gtattttttc ttccattttc 240
 cattagaaat gcaaattaat agtacattat tatggacatg ttttttcaaa aatgtgtatt 300
 ccatgcaggt ttaaaaagat gtggaaaaag ttgtagacta aggtggttga attatcttag 360
 accggatatt aagagaggta atatatcgtc ggatgaagaa gaacttatca ttagacttca 420
 caaactactc ggaaaccggt aaagtatcga cataatcact aacttactaa catttgttta 480
 taatgtgtgc taattgtctt tcctttgatt tgtggtagat ggtctctaata agccggaaga 540
 cttccagggc gaacagacaa tgaaataaag aactactgga acacaaattt aggaaaaaaa 600
 gttaaggatc ttaatcaaca aaacaccaac aattcttctc ctactaaacc ttctgctcaa 660
 ccaaaaaatg caaatatcaa acagaaacaa cagatcaatc ctaagccaat gaagccaaac 720
 tcgaatggtg tccgtacaaa agctacaaa tgttctaagg tattgttcat aaactcacca 780
 ccaatgcata atttgagaa caaagctgag gcagagacaa agacaaagcc attaatgctg 840
 gttaatggtg tagctagtga ttcaatgagt aacaacgaaa tggaacgcgg taatggattt 900
 ttgtcatttt gcgacgaaga gaaagaacta tccgcagatt tgctagatga ttttaacatc 960
 gcggatgata tttgcttatc tgaatttcta aactccgatt tctcaaatgc gtgcaatttc 1020
 gattacaatg atctattgtc gccttgctcg gatcaaaactc aaatgttctc tgatgatgag 1080
 attctcaaga attggacaca atgtaacttt gctgatgaga caaatgtgtc caacaacctt 1140
 cattcttttg cttcctttct cgaatccagt gaggaagtac taggagaatg aaagggcgaa 1200
 ttc 1203

<210> 12

<211> 1206

10 <212> ADN

<213> *Trifolium repens*

<400> 12

ES 2 537 342 T3

gaattcgccc ttatggggag aagcccttgt tgtgcaaaag aaggcttgaa tagaggtgct 60
 tggacagctc atgaggacaa aatcctcact gaatacatta agctccatgg tgaaggaaaa 120
 tggagaaacc ttccaaaaag agcaggttca ttcattctgt atcttactat tatagatcaa 180
 taatcacttt cacacttttt ttttttact tataaatttt catgtatfff ttcttccatt 240
 ttccattaga aatgcaaatt aatagtcacat tattatggac atgttttttc aaaaatgtgt 300
 attccatgca ggtttaaaaa gatgtggaaa aagttgtaga ctaagggtgt tgaattatct 360
 tagaccggat attaagagag gtaatatatc gtcggatgaa gaagaactta tcattagact 420
 tcacaaacta ctcggaaacc ggtaaagtat cgacataatc actaacttac taacatttgt 480
 ttataatgtg tgctaattgc tcttcctttg atttgtggta gatggtctct aatagccgga 540
 agacttccag ggcgaacaga caatgaaata aagaactact ggaacacaaa tttaggaaaa 600
 aaagttaagg atcttaatca acaaaacacc aacaattctt ctcctactaa accttctgct 660
 caacaaaaaa atgcaaatat caaacagaaa caacagatca atcctaagcc aatgaagcca 720
 aactcgaatg ttgtccgtac aaaagctacc aatgttcta aggtattgtt cataaactca 780
 ccaccaatgc ataatttgca gaacaaagct gaggcagaga caaaaacaaa gccattaatg 840
 ctggttaatg gtgtagctag tgattcaatg agtaacaacg aatggaacg cggtaatgga 900
 tttttgtcat tttgcgacga agagaaagaa ctatccgag atttgctaga tgattttaac 960
 atcgcggatg atatttgctt atctgaattt ctaaactccg atttctcaaa tgcgtgcaat 1020
 ttcgattaca atgatctatt gtcgccttgt tcggatcaaa ctcaaagtgt ctctgatgat 1080
 gagattctca agaattggac acaatgtaac tttgctgatg agacaaatgt gtccaacaac 1140
 cttaatcttt ttgcttcttt tctcgaatcc agtgaggag tactaggaga atgaaagggc 1200
 gaattc 1206

<210> 13

<211> 1243

<212> ADN

5 <213> *Trifolium arvense*

<400> 13

gaattcgccc ttaagcagtg gtatcaacgc agagtacgcg ggggaagtta ttttaatttta 60
 tctacatcaa acacttcaag aggttggat acaagacaga ctaattaaga ataacatcaa 120
 tggggagaag cccttgttgt gcaaaggaag gcttgaatag aggtgcttgg acaactcaag 180
 aagacaaaat cctcactgaa tacattaagc tccatggtga aggaaaatgg agaaacctc 240
 caaaaagagc agatttaaaa agatgtggaa aaagttgtag acttagatgg ttgaattatc 300
 taagaccaga tattaagcga ggtaatatat ccccggatga agaagaactt attatccgac 360
 ttcacaaact actcggaaac agatggtctc taatagccgg aagacttcca gggcgaacag 420

ES 2 537 342 T3

acaatgaaat aaagaactac tggaacacaa atttaggaaa aaaggttaag gatcttaatc . 480
 aacaaaacac caacaattct tctcctacta aactttctgc tcaaccaaaa aatgcaaaga 540
 tcaaacagaa acagatcaat cctaagccaa tgaagccaaa ctcaaatggt gtccgtacaa 600
 aagctaccaa gtgttctaag gtattgttca taaactcact ccccaactca ccaatgcatg 660
 atttgcagaa caaagctgag gcagagacaa caacaagcc atcaatgctg gttgatggtg 720
 tggctagtga ttcaatgagt aacaacgaaa tggaacacgg ttatggattt ttgtcatttt 780
 gcgatgaaga gaaagaacta tccgcagatt tgctagaaga ttttaacatc gcggatgata 840
 tttgcttata tgaacttttg aactctgatt tctcaaatgc gtgcaatttc gattacaatg 900
 atctattgtc acctgttgc gaccaaaactc aaatgttctc tgatgatgag attctcaaga 960
 attggacaca atgtaacttt gctgatgaga caaatgtgtc caacaacctt cattcttttg 1020
 cttcctttct tgaatccagt gaggaagtac taggagaatg ataataaaaa ttcattttcc 1080
 aataaaatta actactctag gttttttttt ttttttttta atttcaattt catgttaggg 1140
 tggtttaata aataaatata ttctatggtt taatattgca aaaaaaaaaa aaaaaaaaaa 1200
 aaaaagtact ctgcgttgat accactgctt aagggcgaat tcc 1243

<210> 14

<211> 313

<212> PRT

5 <213> *Trifolium arvense*

<400> 14

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

ES 2 537 342 T3

Pro Thr Lys Leu Ser Ala Gln Pro Lys Asn Ala Lys Ile Lys Gln Lys
 130 135 140

Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg Thr
 145 150 155 160

Lys Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Leu Pro Asn
 165 170 175

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

Lys Pro Ser Met Leu Val Asp Gly Val Ala Ser Asp Ser Met Ser Asn
 195 200 205

Asn Glu Met Glu His Gly Tyr Gly Phe Leu Ser Phe Cys Asp Glu Glu
 210 215 220

Lys Glu Leu Ser Ala Asp Leu Leu Glu Asp Phe Asn Ile Ala Asp Asp
 225 230 240

Ile Cys Leu Ser Glu Leu Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn
 245 250 255

Phe Asp Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met
 260 265 270

Phe Ser Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala
 275 280 285

Asp Glu Thr Asn Val Ser Asn Asn Leu His Ser Phe Ala Ser Phe Leu
 290 295 300

Glu ser ser Glu Glu Val Leu Gly Glu
 305 310

<210> 15

<211> 7

<212> PRT

5 <213> Artificial

<220>

<223> Motivo

<400> 15

Asp Asp Glu Ile Leu Lys Asn
 1 5

10 <210> 16

<211> 7

<212> PRT

<213> Artificial

<220>

15 <223> Motivo

<400> 16

Lys Pro Arg Pro Arg Ser Thr
 1 5

<210> 17
 <211> 12
 <212> PRT

5 <213> Artificial
 <220>
 <223> Motivo
 <220>
 <221> CARACTERÍSTICA DIVERSA

10 <222> (1)..(1)
 <223> Xaa puede ser Asn, Tyr o His
 <220>
 <221> CARACTERÍSTICA DIVERSA
 <222> (6)..(6)

15 <223> Xaa puede ser Lys o Arg
 <220>
 <221> CARACTERÍSTICA DIVERSA
 <222> (8)..(8)
 <223> Xaa puede ser Ile o Thr

20 <400> 17
 xaa val val Arg Thr Xaa Ala Xaa Lys Cys Ser Lys
 1 5 10

<210> 18
 <211> 24
 <212> ADN

25 <213> Artificial
 <220>
 <223> Secuencia de cebador
 <400> 18
 gacaatgaga taaagaatta ctg 24

30 <210> 19
 <211> 21
 <212> ADN
 <213> Artificial
 <220>

35 <223> Secuencia de cebador
 <400> 19

aagagttgta gacttagmtg g 21

<210> 20

<211> 17

<212> ADN

5 <213> Artificial

<220>

<223> Secuencia de cebador

<400> 20

ytkggsaaca ggtgtc 17

10 <210> 21

<211> 23

<212> ADN

<213> Artificial

<220>

15 <223> Secuencia de cebador

<400> 21

atggggagaa gccctgttg tgc 23

<210> 22

<211> 26

20 <212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

<400> 22

25 tcattctcct agtactcct cactgg 26

<210> 23

<211> 21

<212> ADN

<213> Artificial

30 <220>

<223> Secuencia de cebador

<400> 23

ctcttttg aagtttctc c 21

<210> 24

35 <211> 23

<212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

<400> 24

ttctccattt tcctcacca tgg 23

5 <210> 25

<211> 22

<212> ADN

<213> Artificial

<220>

10 <223> Secuencia de cebador

<400> 25

tccaagcacc tctattcaag cc 22

<210> 26

<211> 31

15 <212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

<400> 26

20 ctcgagatgc aatgctgggt gatggtgtgg c 31

<210> 27

<211> 23

<212> ADN

<213> Artificial

25 <220>

<223> Secuencia de cebador

<400> 27

cattgcctgt agattctgta gcc 23

<210> 28

30 <211> 22

<212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

35 <400> 28

tgaagattgt tggacacatt gg 22

<210> 29

<211> 21
<212> ADN
<213> Artificial
<220>
5 <223> Secuencia de cebador
<400> 29
aggttgaat acaagacaga c 21
<210> 30
<211> 22
10 <212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
<400> 30
15 tctcctagta cttcctcact gg 22
<210> 31
<211> 23
<212> ADN
<213> Artificial
20 <220>
<223> Secuencia de cebador
<400> 31
ataatcatac taattaacat cac 23
<210> 32
25 <211> 20
<212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
30 <400> 32
tgatagatca tgcattgtg 20
<210> 33
<211> 23
<212> ADN
35 <213> Artificial
<220>
<223> Secuencia de cebador

<400> 33
gccttccttt gcacaacaag ggc 23
<210> 34
<211> 21
5 <212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
<400> 34
10 gcacaacaag ggcttctccc c 21
<210> 35
<211> 23
<212> ADN
<213> Artificial
15 <220>
<223> Secuencia de cebador
<400> 35
atggggagaa gccctgttg tgc 23
<210> 36
20 <211> 22
<212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
25 <400> 36
tctcctagta cttcctcact gg 22
<210> 37
<211> 28
<212> ADN
30 <213> Artificial
<220>
<223> Secuencia de cebador
<400> 37
ctcgagcaat gctggtgat ggtgtggc 28
35 <210> 38
<211> 27
<212> ADN

<213> Artificial
<220>
<223> Secuencia de cebador
<400> 38
5 tctagaggac acattgtct catcagc 27
<210> 39
<211> 27
<212> ADN
<213> Artificial
10 <220>
<223> Secuencia de cebador
<400> 39
tctagattga gtttggccg aacaagg 27
<210> 40
15 <211> 27
<212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
20 <400> 40
tctagaaatc ttctagcaaa tctgcgg 27
<210> 41
<211> 16
<212> ADN
25 <213> Artificial
<220>
<223> Secuencia de cebador
<400> 41
gtaaaacgac ggccag 16
30 <210> 42
<211> 17
<212> ADN
<213> Artificial
<220>
35 <223> Secuencia de cebador
<400> 42
caggaaacag ctatgac 17

<210> 43
<211> 30
<212> ADN
<213> Artificial
5 <220>
<223> Secuencia de cebador
<400> 43
aagcagtggg atcaacgcag agtacgcggg 30
<210> 44
10 <211> 28
<212> ADN
<213> Artificial
<220>
<223> Secuencia de cebador
15 <220>
<221> CARACTERÍSTICA DIVERSA
<222> (28)..(28)
<223> n es a, c, g, o t
<400> 44
20 aagcagtggg atcaacgcag agtactvn 28
<210> 45
<211> 23
<212> ADN
<213> Artificial
25 <220>
<223> Secuencia de cebador
<400> 45
aagcagtggg atcaacgcag agt 23
<210> 46
30 <211> 316
<212> PRT
<213> *Trifolium arvense*
<400> 46

ES 2 537 342 T3

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
1 5 10 15

Trp Thr Thr Gln Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
50 55 60

Ile Lys Arg Gly Asn Ile Ser Pro Asp Glu Glu Glu Leu Ile Ile Arg
65 70 75 80

Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
85 90 95

Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
100 105 110

Gly Lys Lys Val Lys Asp Leu Asp Gln Gln Asn Thr Asn Asn Ser Ser
115 120 125

Pro Thr Lys Leu Ser Ala Gln Pro Lys Asn Ala Glu Ile Lys Gln Lys
130 135 140

Gln Ile Asn Pro Lys Pro Asn Ser Tyr Val Val Arg Thr Lys Ala Thr
145 150 155 160

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

Met His Asp Leu Gln Ser Lys Ala Glu Ala Glu Thr Thr Thr Thr Thr
180 185 190

Lys Pro Ser Met Pro Ser Met Leu Val Asp Gly Val Ala Ser Asp Ser
195 200 205

Met Ser Asn Asn Glu Met Glu Cys Gly Asn Gly Phe Leu Ser Phe Cys
210 215 220

Asp Glu Glu Lys Glu Leu Ser Ala Asp Leu Leu Glu Asp Phe Asn Ile
225 230 235 240

Ala Asp Asp Ile Cys Leu Ser Glu Phe Leu Asn Phe Asp Phe Ser Asn
245 250 255

Ala Cys Asp Ile Asp Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln
260 265 270

Thr Gln Met Phe Pro Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys
275 280 285

Asn Phe Ala Asp Glu Thr Asn Val Ser Asn Asn Leu Gln Ser Ser Ala
290 295 300

Ser Phe Leu Glu Ser Ser Glu Glu Val Leu Gly Glu
305 310 315

ES 2 537 342 T3

<210> 47

<211> 310

<212> PRT

<213> *Trifolium affine*

5 <400> 47

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
1 5 10 15

ES 2 537 342 T3

Trp Thr Thr Gln Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
 20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
 35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Leu Asp
 50 55 60

Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
 65 70 75 80

Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
 85 90 95

Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
 100 105 110

Gly Lys Lys Val Lys Asp Leu Asn Gln Glu Asn Thr Asn Asn Ser Ser
 115 120 125

Pro Thr Lys Leu Ser Ala Gln Leu Lys Asn Ala Lys Ile Lys Gln Lys
 130 135 140

Gln Ile Asn Pro Lys Pro Met Glu Pro Asn Ser Asn Val Val Arg Thr
 145 150 155 160

Lys Ala Thr Lys Cys Ser Lys Ala Leu Phe Ile Asn Ser Pro Pro Asn
 165 170 175

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

Thr Lys Ser Ser Met Pro Ser Met Leu Val Asp Gly Val Ala Ser Asp
 195 200 205

Ser Met Ser Asn Asn Glu Met Glu Tyr Gly Asp Gly Phe Val Ser Phe
 210 215 220

Cys Asp Asp Asp Lys Glu Leu Ser Ala Asp Leu Leu Glu Asp Phe Asn
 225 230 235 240

Ile Ser Asp Asp Ile Cys Leu Ser Glu Phe Leu Asn Phe Asp Phe Ser
 245 250 255

Asn Ala Cys Asn Phe Asp Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp
 260 265 270

Gln Thr Gln Met Phe Ser Asp Asp Glu Ile Leu Lys Asn Ser Thr Pro
 275 280 285

Cys Asn Phe Ala Ala Glu Thr Asn Tyr Val Ser Asn Asn Gln Ser Ser
 290 295 300

Glu Glu Val Leu Gly Glu
 305 310

<210> 48

<211> 296

5 <212> PRT

ES 2 537 342 T3

<213> *Trifolium affine*

<400> 48

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
 1 5 10 15
 Trp Thr Thr Gln Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
 20 25 30
 Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
 35 40 45
 Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
 50 55 60
 Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
 65 70 75 80
 Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
 85 90 95
 Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
 100 105 110
 Gly Lys Lys Val Lys Asp Leu Asn Gln Glu Asn Thr Asn Asn Ser Ser
 115 120 125
 Pro Thr Lys Leu Ser Ala Gln Leu Lys Asn Ala Lys Ile Lys Gln Lys
 130 135 140
 Gln Ile Asn Pro Lys Pro Met Glu Pro Asn Ser Asn Val Val Arg Thr
 145 150 155 160
 Lys Ala Thr Lys Cys Ser Lys Ala Leu Phe Ile Asn Ser Pro Pro Asn
 165 170 175
 Ser Pro Pro Met His Asp Leu Gln Asn Lys Ala Glu Ala Glu Thr Thr
 180 185 190
 Thr Lys Ser Ser Met Pro Ser Met Leu Val Asp Gly Val Ala Ser Asp
 195 200 205
 Ser Met Ser Asn Asn Glu Met Glu Tyr Gly Asp Gly Phe Val Ser Phe
 210 215 220
 Cys Asp Asp Asp Lys Glu Leu Ser Ala Asp Leu Leu Glu Asp Phe Asn
 225 230 235 240
 Ile Ser Asp Asp Ile Cys Leu Ser Glu Phe Leu Asn Phe Asp Phe Ser
 245 250 255
 Asn Ala Cys Asn Phe Asp Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp
 260 265 270
 Gln Thr Gln Met Phe Ser Asp Asp Glu Ile Leu Lys Asn Ser Thr Gln
 275 280 285
 Cys Asn Phe Ala Ala Glu Thr Asn
 290 295

5 <210> 49

ES 2 537 342 T3

<211> 313

<212> PRT

<213> *Trifolium occidentale*

<400> 49

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
 1 5 10 15
 Trp Thr Thr Gln Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
 20 25 30
 Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
 35 40 45
 Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
 50 55 60
 Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
 65 70 75 80
 Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
 85 90 95
 Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
 100 105 110
 Gly Lys Lys Val Lys Asp Leu Asn Gln Gln Asn Thr Asn Lys Ser Ser
 115 120 125
 Pro Thr Lys Leu Ser Ala Gln Pro Lys Asn Ala Lys Ile Lys Gln Lys
 130 135 140

5

ES 2 537 342 T3

Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg Thr
145 150 155 160

Arg Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Leu Pro Asn
165 170 175

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

Lys Pro Ser Met Leu Val Asp Gly Val Ala Ser Asp Ser Met Ser Asn
195 200 205

Asn Glu Met Glu His Gly Tyr Gly Phe Leu Ser Phe Cys Asp Glu Glu
210 215 220

Lys Glu Leu Ser Ala Asp Leu Leu Glu Asp Phe Asn Ile Ala Asp Asp
225 230 235 240

Ile Cys Leu Ser Glu Leu Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn
245 250 255

Phe Asp Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met
260 265 270

Phe Ser Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala
275 280 285

Asp Glu Thr Asn Val Ser Asn Asn Leu His Ser Phe Ala Ser Phe Leu
290 295 300

Glu Ser Ser Glu Glu Val Leu Gly Glu
305 310

<210> 50

<211> 312

<212> PRT

5 <213> *Trifolium occidentale*

<400> 50

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
1 5 10 15

Trp Thr Ala His Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
50 55 60

Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg

ES 2 537 342 T3

65					70					75					80
Leu	His	Lys	Leu	Leu	Gly	Asn	Arg	Trp	Ser	Leu	Ile	Ala	Gly	Arg	Leu
				85					90					95	
Pro	Gly	Arg	Thr	Asp	Asn	Glu	Ile	Lys	Asn	Tyr	Trp	Asn	Thr	Asn	Leu
			100					105					110		
Gly	Lys	Lys	Val	Lys	Asp	Leu	Asn	Gln	Gln	Asn	Thr	Asn	Asn	Ser	Ser
		115					120					125			
Pro	Thr	Lys	Pro	Ser	Ala	Gln	Pro	Lys	Asn	Ala	Lys	Ile	Lys	Gln	Lys
	130					135					140				
Gln	Gln	Ile	Asn	Asn	Pro	Lys	Pro	Met	Lys	Pro	Asn	Ser	Asn	Val	Val
145					150					155					160
Arg	Thr	Lys	Ala	Thr	Lys	Cys	Ser	Lys	Val	Leu	Phe	Ile	Asn	Ser	Pro
				165					170					175	
Pro	Met	His	Asn	Leu	Gln	Asn	Lys	Ala	Glu	Ala	Glu	Thr	Lys	Thr	Lys
			180					185					190		
Thr	Ser	Met	Leu	Val	Asn	Gly	Val	Ala	Ser	Asp	Ser	Met	Ser	Asn	Asn
		195					200					205			
Glu	Met	Glu	Arg	Gly	Asn	Gly	Phe	Leu	Ser	Phe	Arg	Asp	Glu	Glu	Lys
	210					215					220				
Glu	Leu	Ser	Ala	Asp	Leu	Leu	Asp	Asp	Phe	Asn	Ile	Ala	Asp	Asp	Ile
225					230					235					240
Cys	Leu	Ser	Glu	Phe	Leu	Asn	Ser	Asp	Phe	Ser	Asn	Ala	Cys	Asn	Phe
				245					250					255	
Asp	Tyr	Asn	Asp	Leu	Leu	Ser	Pro	Cys	Ser	Asp	Gln	Thr	Gln	Met	Phe
			260					265					270		
Ser	Asp	Asp	Glu	Ile	Leu	Lys	Asn	Trp	Thr	Gln	Cys	Asn	Phe	Ala	Asp
		275					280					285			
Glu	Thr	Asn	Val	Ser	Asn	Asn	Leu	His	Ser	Phe	Ala	Ser	Phe	Leu	Glu
	290					295					300				
Ser	Ser	Glu	Glu	Val	Leu	Gly	Glu								
305					310										
<210>	51														
<211>	311														
<212>	PRT														
5	<213>	<i>Trifolium repens</i>													
<400>	51														

ES 2 537 342 T3

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
1 5 10 15

Trp Thr Ala His Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
50 55 60

Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
65 70 75 80

Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
85 90 95

Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
100 105 110

Gly Lys Lys Val Lys Asp Leu Asn Gln Gln Asn Thr Asn Asn Ser Ser
115 120 125

Pro Thr Lys Pro Ser Ala Gln Pro Lys Asn Ala Asn Ile Lys Gln Lys
130 135 140

Gln Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg
145 150 155 160

Thr Lys Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Pro Pro
165 170 175

Met His Asn Leu Gln Asn Lys Ala Glu Ala Glu Thr Lys Thr Lys Pro
180 185 190

Leu Met Leu Val Asn Gly Val Ala Ser Asp Ser Met Ser Asn Asn Glu
195 200 205

Met Glu Arg Gly Asn Gly Phe Leu Ser Phe Cys Asp Glu Glu Lys Glu
210 215 220

Leu Ser Ala Asp Leu Leu Asp Asp Phe Asn Ile Ala Asp Asp Ile Cys
225 230 235 240

Leu Ser Glu Phe Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn Phe Asp
245 250 255

Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met Phe Ser
260 265 270

Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala Asp Glu
275 280 285

Thr Asn Val Ser Asn Asn Leu Asn Ser Phe Ala Ser Phe Leu Glu Ser
290 295 300

Ser Glu Glu Val Leu Gly Glu
305 310

<210> 52

ES 2 537 342 T3

<211> 311

<212> PRT

<213> *Trifolium repens*

<400> 52

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
1 5 10 15

Trp Thr Ala His Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
50 55 60

Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
65 70 75 80

Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
85 90 95

Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
100 105 110

Gly Lys Lys Val Lys Asp Leu Asn Gln Gln Asn Thr Asn Asn Ser Ser
115 120 125

Pro Thr Lys Pro Ser Ala Gln Pro Lys Asn Ala Asn Ile Lys Gln Lys
130 135 140

Gln Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg
145 150 155 160

Thr Lys Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Pro Pro
165 170 175

Met His Asn Leu Gln Asn Lys Ala Glu Ala Glu Thr Lys Thr Lys Pro
180 185 190

5

ES 2 537 342 T3

Leu Met Leu Val Asn Gly Val Ala Ser Asp Ser Met Ser Asn Asn Glu
 195 200 205

Met Glu Arg Gly Asn Gly Phe Leu Ser Phe Cys Asp Glu Glu Lys Glu
 210 215 220

Leu Ser Ala Asp Leu Leu Asp Asp Phe Asn Ile Ala Asp Asp Ile Cys
 225 230 235 240

Leu Pro Glu Phe Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn Phe Asp
 245 250 255

Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met Phe Ser
 260 265 270

Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala Asp Glu
 275 280 285

Thr Asn Val Ser Asn Asn Leu Asn Ser Phe Ala Ser Phe Leu Glu Ser
 290 295 300

Ser Glu Glu Val Leu Gly Glu
 305 310

<210> 53

<211> 311

<212> PRT

5 <213> *Trifolium repens*

<400> 53

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
 1 5 10 15

Trp Thr Ala His Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
 20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg
 35 40 45

Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
 50 55 60

Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
 65 70 75 80

Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
 85 90 95

Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
 100 105 110

ES 2 537 342 T3

Gly Lys Lys Val Lys Asp Leu Asn Gln Gln Asn Thr Asn Asn Ser Ser
 115 120 125

Pro Thr Lys Pro Ser Ala Gln Pro Lys Asn Ala Asn Ile Lys Gln Lys
 130 135 140

Gln Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg
 145 150 155 160

Thr Lys Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Pro Pro
 165 170 175

Met His Asn Leu Gln Asn Lys Ala Glu Ala Glu Thr Lys Thr Lys Pro
 180 185 190

Leu Met Leu Val Asn Gly Val Ala Ser Asp Ser Met Ser Asn Asn Glu
 195 200 205

Met Glu Arg Gly Asn Gly Phe Leu Ser Phe Cys Asp Glu Glu Lys Glu
 210 215 220

Leu Ser Ala Asp Leu Leu Asp Asp Phe Asn Ile Ala Asp Asp Ile Cys
 225 230 235 240

Leu Ser Glu Phe Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn Phe Asp
 245 250 255

Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met Phe Ser
 260 265 270

Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala Asp Glu
 275 280 285

Thr Asn Val Ser Asn Asn Leu His Ser Phe Ala Ser Phe Leu Glu Ser
 290 295 300

Ser Glu Glu Val Leu Gly Glu
 305 310

<210> 54

<211> 311

<212> PRT

5 <213> *Trifolium repens*

<400> 54

Met Gly Arg Ser Pro Cys Cys Ala Lys Glu Gly Leu Asn Arg Gly Ala
 1 5 10 15

Trp Thr Ala His Glu Asp Lys Ile Leu Thr Glu Tyr Ile Lys Leu His
 20 25 30

Gly Glu Gly Lys Trp Arg Asn Leu Pro Lys Arg Ala Gly Leu Lys Arg

ES 2 537 342 T3

35 40 45
 Cys Gly Lys Ser Cys Arg Leu Arg Trp Leu Asn Tyr Leu Arg Pro Asp
 50 55 60
 Ile Lys Arg Gly Asn Ile Ser Ser Asp Glu Glu Glu Leu Ile Ile Arg
 65 70 75
 Leu His Lys Leu Leu Gly Asn Arg Trp Ser Leu Ile Ala Gly Arg Leu
 85 90 95
 Pro Gly Arg Thr Asp Asn Glu Ile Lys Asn Tyr Trp Asn Thr Asn Leu
 100 105 110
 Gly Lys Lys Val Lys Asp Leu Asn Gln Gln Asn Thr Asn Asn Ser Ser
 115 120 125
 Pro Thr Lys Pro Ser Ala Gln Pro Lys Asn Ala Asn Ile Lys Gln Lys
 130 135 140
 Gln Gln Ile Asn Pro Lys Pro Met Lys Pro Asn Ser Asn Val Val Arg
 145 150 155 160
 Thr Lys Ala Thr Lys Cys Ser Lys Val Leu Phe Ile Asn Ser Pro Pro
 165 170 175
 Met His Asn Leu Gln Asn Lys Ala Glu Ala Glu Thr Lys Thr Lys Pro
 180 185 190
 Leu Met Leu Val Asn Gly Val Ala Ser Asp Ser Met Ser Asn Asn Glu
 195 200
 Met Glu Arg Gly Asn Gly Phe Leu Ser Phe Cys Asp Glu Glu Lys Glu
 210 215 220
 Leu Ser Ala Asp Leu Leu Asp Asp Phe Asn Ile Ala Asp Asp Ile Cys
 225 230 235 240
 Leu Ser Glu Phe Leu Asn Ser Asp Phe Ser Asn Ala Cys Asn Phe Asp
 245 250 255
 Tyr Asn Asp Leu Leu Ser Pro Cys Ser Asp Gln Thr Gln Met Phe Ser
 260 265 270
 Asp Asp Glu Ile Leu Lys Asn Trp Thr Gln Cys Asn Phe Ala Asp Glu
 275 280 285
 Thr Asn Val Ser Asn Asn Leu His Ser Phe Ala Ser Phe Leu Glu Ser
 290 295 300
 Ser Glu Glu Val Leu Gly Glu
 305 310

<210> 55

<211> 942

5 <212> ADN

<213> *Trifolium arvense*

<400> 55

ES 2 537 342 T3

atggggagaa gcccttggtg tgcaaaggaa ggcttgaata gaggtgcttg gacaactcaa 60
 gaagacaaaa tcctcactga atacattaag ctccatgggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag cagggtttaa aagatgcgga aaaagttgta gacttagatg gttgaattat 180
 ctaagaccag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatcaga 240
 cttcacaaac tactcggaaa cagatggtct ctaatagccg gaagacttcc aggacgaaca 300
 gacaatgaaa taaagaacta ctggaacaca aatntagga aaaaaggtaa ggatcttaat 360
 caacaaaaca ccaacaattc ttctcctact aaactctctg ctcaaccaa aaatgcaaag 420
 atcaaacaga aacagatcaa tcctaagcca atgaagccaa actcaaagt tgtccgtaca 480
 aaagctacca agtgttctaa ggtattgttc ataaactcac tccccaaact accaatgcat 540
 gatttgacaga acaaagctga ggagagaca acaacaaagc catcaatgct ggttgatggg 600
 gtggctagtg attcaatgag taacaacgaa atggaacacg gttatggatt tttgtcattt 660
 tgcgatgaag agaaagaact atccgcagat ttgctagaag attttaacat cgcgatgat 720
 atttgcttat ctgaactttt gaactctgat ttctcaaatg cgtgcaattt cgattacaat 780
 gatctattgt caccttggtc ggaccaaact caaatgttct ctgatgatga gattctcaag 840
 aattggacac aatgtaactt tgctgatgag acaaatgtgt ccaacaacct tcattctttt 900
 gcttcctttc ttgaatccag tgaggaagta ctaggagaat ga 942

<210> 56

<211> 933

<212> ADN

5 <213> *Trifolium arvense*

<400> 56

atggggagaa gcccttggtg tgcaaaggaa ggcttgaata gaggtgcttg gacaactcaa 60
 gaagacaaaa tcctcactga atacattaag ctccatgggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag cagggtttaa aagatgtgga aaaagttgta gacttagatg gttgaattat 180
 ctaagaccag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatccga 240
 cttcacaaac tactcggaaa cagatggtct ctaatagccg gaagacttcc agggcgaaca 300
 gacaatgaaa taaagaacta ctggaacaca aatntagga aaaaaggtaa ggatcttaat 360
 caagaaaaca ccaacaattc ttctcctact aaactttctg ctcaactaaa aaatgcaaag 420
 atcaaacaga aacagatcaa tcctaagcca atggagccaa actcaaagt tgtccgtaca 480
 aaagctacca agtgttctaa ggcattgttc ataaactcac cccccaaact accaccaatg 540
 catgatttgc agaacaaagc tgaggcagag acaacaaca agtcatcaat gccatcaatg 600
 ctggttgatg gcggtgctag tgattcaatg agtaacaacg aaatggaata cggtgatgga 660
 tttgtttcat tttgcgatga cgataaagaa ctatccgcag atttgctaga agattttaac 720
 atctcggatg atatttgctt atccgaattt ctaaacttctg atttctcaa tgcgtgcaat 780
 ttcgattaca acgatctatt gtcgccttgt tcggacaaa cacaaatgtt ctctggtgat 840
 gagattctca agaattcgac acaatgtaac tttgctgctg agacaaatta tgtgtccaac 900
 aaccaatcca gtgaggaagt actaggagaa tga 933

<210> 57

10 <211> 933

<212> ADN

<213> *Trifolium affine*

<400> 57

atggggagaa gcccttggtg tgcaaggaa ggcttgaata gaggtgcttg gacaactcaa 60
 gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag caggtttaaa aagatgtgga aaaagttgta gacttagatg gttgaattat 180
 ctaagactag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatccga 240
 cttcacaat tactcggaaa cagatggtct ctaatagccg gaagacttcc aggacgaaca 300
 gacaatgaaa taaagaacta ctggaacaca aathtagga aaaaggtaa ggatcttaat 360
 caagaaaaca ccaacaattc ttctcctact aaactttctg ctcaactaaa aaatgcaaag 420
 atcaaacaga aacagatcaa tcctaagcca atggagccaa actcaaatgt tgtccgtaca 480
 aaagctacca agtgttctaa ggcattgttc ataaactcac cccccaactc accaccaatg 540
 catgatttgc agaacaaagc tgaggcagag acaacaacaa agtcatcaat gccatcaatg 600
 ctggttgatg gcggtggctag tgattcaatg agtaacaacg aaatggaata cggatgatgga 660
 tttgtttcat tttgcatga cgataaagaa ctatccgcag atttgctaga agattttaac 720
 atctcggatg atatttgctt atccgaattt ctaaacttcg atttctcaa tgcgtgcaat 780
 ttcgattaca acgatctatt gtcgccttgt tcggaccaa cacaaatgtt ctctgatgat 840
 gagattctca agaattcgac accatgtaac tttgctgctg agacaaatta tgtgtccaac 900
 aaccaatcca gtgaggaagt actaggagaa tga 933

5 <210> 58

<211> 891

<212> ADN

<213> *Trifolium affine*

<400> 58

atggggagaa gcccttggtg tgcaaggaa ggcttgaata gaggtgcttg gacaactcaa 60
 gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag caggtttaaa aagatgtgga aaaagttgta gacttagatg gttgaattat 180
 ctaagaccag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatccga 240
 10 cttcacaac tactcggaaa cagatggtct ctaatagccg gaagacttcc agggcgaaca 300
 gacaatgaaa taaagaacta ctggaacaca aathtagga aaaaggtaa ggatcttaat 360
 caagaaaaca ccaacaattc ttctcctact aaactttctg ctcaactaaa aaatgcaaag 420
 atcaaacaga aacagatcaa tcctaagcca atggagccaa actcaaatgt tgtccgtaca 480
 aaagctacca agtgttctaa ggcattgttc ataaactcac cccccaactc accaccaatg 540
 catgatttgc agaacaaagc tgaggcagag acaacaacaa agtcatcaat gccatcaatg 600
 ctggttgatg gcggtggctag tgattcaatg agtaacaacg aaatggaata cggatgatgga 660
 tttgtttcat tttgcatga cgataaagaa ctatccgcag atttgctaga agattttaac 720
 atctcggatg atatttgctt atccgaattt ctaaacttcg atttctcaa tgcgtgcaat 780
 ttcgattaca acgatctatt gtcgccttgt tcggaccaa cacaaatgtt ctctgatgat 840
 gagattctca agaattcgac acaatgtaac tttgctgctg agacaaatta a 891

ES 2 537 342 T3

<210> 59

<211> 942

<212> ADN

<213> *Trifolium occidentale*

5 <400> 59

```

atggggagaa gcccttggtg tgcaaaggaa ggcttgaata gaggtgcttg gacaactcaa      60
gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt      120
ccaaaaagag caggtttaaa aagatgcgga aaaagtgtga gacttagatg gttgaattat      180
ctaagaccag atattaagcg aggtaatata tcctcggatg aagaagaact tatcatcaga      240
cttcacaaac tactcggaaa cagatggtct ctaatagccg gaagacttcc aggacgaaca      300
gacaatgaaa taaagaacta ctggaacaca aatttaggaa aaaaggtaa ggatcttaat      360
caacaaaaca ccaacaagtc ttctcctact aaactctctg ctcaaccaa aaatgcaaag      420
atcaaacaga aacagatcaa tcctaagcca atgaagcaa actcaaagt tgtccgtaca      480
agagctacca agtgttctaa ggtattgttc ataaactcac tccccactc accaatgcat      540
gatttgcaga acaaagctga ggcagagaca acaacaaagc catcaatgct ggttgatggt      600
gtggctagtg attcaatgag taacaacgaa atggaacacg gttatggatt tttgtcattt      660
tgcgatgaag agaaagaact atccgcagat ttgctagaag attttaacat cgcggatgat      720
atttgcttat ctgaactttt gaactctgat ttctcaaag cgtgcaattt cgattacaat      780
gatctattgt cmccttggtc ggaccaaact caaatgttct ctgatgatga gattctcaag      840
aattggacac aatgtaactt tgctgatgag acaaatgtgt ccaacaacct tcattctttt      900
gcttcctttc ttgaatccag tgaggaagta ctaggagaat ga                          942

```

<210> 60

<211> 939

<212> ADN

10 <213> *Trifolium occidentale*

<400> 60

```

atggggagaa gcccttggtg tgcaaaggaa ggttgaata gaggtgcttg gacagctcat      60

```

ES 2 537 342 T3

gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag cagggtttaa aagatgtgga aaaagttgta gacttagatg gttgaattat 180
 cttagaccag atattaagag aggtaatata tcgtccgatg aagaagaact tadcattaga 240
 cttcacaac tacttggaaa ccgatggtct ctaatagccg gaagacttcc agggcgaaca 300
 gacaatgaaa taaaaaatta ctggaacacg aathtagga aaaaggtaa ggatcttaat 360
 caacaaaaca ccaacaattc ttctcctact aaaccttctg ctcaaccaa aatgcaaag 420
 atcaaacaga aacaacagat caataatct aagccaatga agccaaactc gaatgtgtc 480
 cgtacaaaag ctaccaatg ttctaaggta ttgttcataa actcaccacc aatgcataat 540
 ttgcagaaca aagctgaggc agagacaaaa acaagacat caatgttgt taatgtgtgta 600
 gctagtgatt caatgagtaa caacgaaatg gaacgaggta atggatTTTT gtcatttcgc 660
 gatgaagaga aagaactatc cgctgattg ctagatgatt ttaacatcgc ggatgacatt 720
 tgcttatccg aatttctaaa ctccgatttc tcaaatgctg gcaatttcga ttacaatgat 780
 ctattgtcac cttgttcgga tcaactcaa atgttctctg atgatgagat tctcaagaat 840
 tggacacaat gtaactttgc tgatgagaca aatgtgtcca acaaccttca ttcttttgct 900
 tcctttctcg aatccagtga ggaagtacta ggagaatga 939

<210> 61

<211> 936

<212> ADN

5 <213> *Trifolium repens*

<400> 61

atggggagaa gcccttggtg tgcaaaagaa ggcttgaata gaggtgcttg gacagctcat 60
 gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt 120
 ccaaaaagag cagggtttaa aagatgtgga aaaagttgta gactaaggatg gttgaattat 180
 cttagaccgg atattaagag aggtaatata tcgtccgatg aagaagaact tadcattaga 240
 cttcacaac tactcggaaa ccgatggtct ctaatagccg gaagacttcc agggcgaaca 300
 gacaatgaaa taagaacta ctggaacaca aathtagga aaaaagttaa ggatcttaat 360
 caacaaaaca ccaacaattc ttctcctact aaaccttctg ctcaaccaa aatgcaaat 420
 atcaaacaga aacaacagat caatcctaag ccaatgaagc caaactcga tggtgtccgt 480
 acaaaagcta ccaaatgttc taaggattg ttcataaact caccaccaat gcataatttg 540
 cagaacaaaag ctgaggcaga gacaaaaaca aagccattaa tgctggttaa tgggtgtagct 600
 agtgattcaa tgagtaaca cgaaatgga cgcggtaatg gatttttgtc attttgcgac 660
 gaagagaaag aactatccgc agatttgcta gatgatttta acatcgcgga tgatatttgc 720
 ttatctgaat ttctaaactc cgatttctca aatgcgtgca atttcgatta caatgatcta 780
 ttgtcgcctt gttcggatca aactcaaatg ttctctgatg atgagattct caagaattgg 840
 acacaatgta actttgctga tgagacaaat gtgtccaaca accttaattc ttttgcttct 900
 tttctcgaat ccagtgagga agtactagga gaatga 936

<210> 62

<211> 936

10 <212> ADN

ES 2 537 342 T3

<213> *Trifolium repens*

<400> 62

atggggagaa gcccttggtg tgcaaaagaa ggcttgaata gaggtgcttg gacagctcat	60
gaagacaaaa tcctcactga atacattaag ctccatgggtg aaggaaaatg gagaaacctt	120
ccaaaaagag caggtttaaa aagatgtgga aaaagttgta gactaagggtg gttgaattat	180
cttagaccgg atattaagag aggtaatata tcgtcggatg aagaagaact tadcattaga	240
cttcacaaac tactcggaaa ccgatggtct ctaatagccg gaagacttcc agggcgaaca	300
gacaatgaaa taaagaacta ctggaacaca aatthagga aaaaagttaa ggatcttaat	360
caacaaaaca ccaacaattc ttctcctact aaaccttctg ctcaaccaa aaatgcaaat	420
atcaaacaga aacaacagat caatcctaag ccaatgaagc caaactcgaa tgttgctcgt	480
acaaaagcta ccaaatgttc taaggatttg ttcataaact caccaccaat gcataatttg	540
cagaacaaag ctgaggcaga gacaaaaaca aagccattaa tgctgggtaa tgggtgtagct	600
agtgattcaa tgagtaacaa cgaaatggaa cgcggtaatg gatttttgtc attttgcgac	660
gaagagaaaag aactatccgc agatttgcta gatgatttta acatcgcgga tgatatttgc	720
ttacctgaat ttctaaactc cgatttctca aatgcgtgca atttcgatta caatgatcta	780
ttgtcgcctt gttcggatca aactcaaatg ttctctgatg atgagattct caagaattgg	840
acacaatgta actttgctga tgagacaaat gtgtccaaca accttaattc ttttgcttct	900
tttctcgaat ccagtgagga agtactagga gaatga	936

<210> 63

5 <211> 936

<212> ADN

<213> *Trifolium repens*

<400> 63

atggggagaa gcccttggtg tgcaaaagaa ggcttgaata gaggtgcttg gacagctcat	60
gaagacaaaa tcctcactga atacattaag ctccatgggtg aaggaaaatg gagaaacctt	120
ccaaaaagag caggtttaaa aagatgtgga aaaagttgta gactaagggtg gttgaattat	180
cttagaccgg atattaagag aggtaatata tcgtcggatg aagaagaact tadcattaga	240
cttcacaaac tactcggaaa ccgatggtct ctaatagccg gaagacttcc agggcgaaca	300
gacaatgaaa taaagaacta ctggaacaca aatthagga aaaaagttaa ggatcttaat	360
caacaaaaca ccaacaattc ttctcctact aaaccttctg ctcaaccaa aaatgcaaat	420
atcaaacaga aacaacagat caatcctaag ccaatgaagc caaactcgaa tgttgctcgt	480
acaaaagcta ccaaatgttc taaggatttg ttcataaact caccaccaat gcataatttg	540
cagaacaaag ctgaggcaga gacaaagaca aagccattaa tgctgggtaa tgggtgtagct	600
agtgattcaa tgagtaacaa cgaaatggaa cgcggtaatg gatttttgtc attttgcgac	660
gaagagaaaag aactatccgc agatttgcta gatgatttta acatcgcgga tgatatttgc	720
ttatctgaat ttctaaactc cgatttctca aatgcgtgca atttcgatta caatgatcta	780
ttgtcgcctt gttcggatca aactcaaatg ttctctgatg atgagattct caagaattgg	840
acacaatgta actttgctga tgagacaaat gtgtccaaca accttcattc ttttgcttcc	900
tttctcgaat ccagtgagga agtactagga gaatga	936

10

ES 2 537 342 T3

<210> 64

<211> 936

<212> ADN

<213> *Trifolium repens*

5 <400> 64

```

atggggagaa gcccttggtg tgcaaaagaa ggcttgaata gaggtgcttg gacagctcat    60
gaagacaaaa tcctcactga atacattaag ctccatggtg aaggaaaatg gagaaacctt    120
ccaaaaagag caggtttaa aagatgtgga aaaagttgta gactaagggtg gttgaattat    180
cttagaccgg atattaagag aggtaatata tcgtcggatg aagaagaact taccattaga    240
cttcacaaac tactcggaaa ccgatggtct ctaatagccg gaagacttcc agggcgaaca    300
gacaatgaaa taaagaacta ctggaacaca aatttaggaa aaaaagttaa ggatcttaat    360
caacaaaaca ccaacaattc ttctcctact aaaccttctg ctcaaccaa aaatgcaaat    420
atcaaacaga aacaacagat caatcctaag ccaatgaagc caaactcgaa tgttgtccgt    480
acaaaagcta ccaaatgttc taaggtattg ttcataaact caccaccaat gcataatttg    540
cagaacaaag ctgaggcaga gacaaagaca aagccattaa tgctgggtta tgggtgtagct    600
agtgattcaa tgagtaacaa cgaaatggaa cgcggtaatg gatttttgtc attttgcgac    660
gaagagaaaag aactatccgc agatttgcta gatgatttta acatcgcgga tgatatttgc    720
ttatctgaat ttctaaactc cgatttctca aatgcgtgca atttcgatta caatgatcta    780
ttgtcgctt gttcggatca aactcaaatg ttctctgatg atgagattct caagaattgg    840
acacaatgta actttgctga tgagacaaat gtgtccaaca accttcattc ttttgcttcc    900
tttctcgaat ccagtgagga agtactagga gaatga                                936

```

<210> 65

<211> 299

<212> ADN

10 <213> *Trifolium arvense*

<400> 65

```

caatgctggt tgatggtgtg gctagtgatt caatgagtaa caacgaaatg gaacacggtt    60
atggattttt gtcattttgc gatgaagaga aagaactatc cgcagatttg ctagaagatt    120
ttaacatcgc ggatgatatt tgcttatctg aacttttgaa ctctgatttc tcaaatgctg    180
gcaatttcga ttacaatgat ctattgtcac cttgttcgga ccaaactcaa atgttctctg    240
atgatgagat tctcaagaat tggacacaat gtaactttgc tgatgagaca aatgtgtcc    299

```

<210> 66

<211> 28

15 <212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

<400> 66

tctagacaat gctggttgat ggtgtggc 28

<210> 67

<211> 27

<212> ADN

5 <213> Artificial

<220>

<223> Secuencia de cebador

<400> 67

tctagaggac acattgtct catcagc 27

10 <210> 68

<211> 28

<212> ADN

<213> Artificial

<220>

15 <223> Secuencia de cebador

<400> 68

ctcgagcaat gctggttgat ggtgtggc 28

<210> 69

<211> 27

20 <212> ADN

<213> Artificial

<220>

<223> Secuencia de cebador

<400> 69

25 ctcgagggac acattgtct catcagc 27

REIVINDICACIONES

1. Una molécula de ácido nucleico aislada que codifica al menos uno de:
- a) un polipéptido MYB14 que comprende una secuencia con al menos 70% de identidad con SEC ID NO: 14, en el que el % de identidad se calcula a lo largo de toda la longitud de SEC ID NO: 14; y
- 5 b) un fragmento funcional de SEC ID NO: 14;
- en la que el polipéptido MYB14 en a) y el fragmento funcional en b) son capaces de incrementar la producción de taninos condensados en plantas.
2. La molécula de ácido nucleico aislada de la reivindicación 1, en la que el polipéptido MYB14 comprende la secuencia de SEC ID NO: 14.
- 10 3. La molécula de ácido nucleico aislada de una cualquiera de las reivindicaciones 1 a 2, en la que el polipéptido MYB14, o fragmento funcional, comprende una secuencia de aminoácidos de SEC ID NO: 17.
4. La molécula de ácido nucleico aislada de la reivindicación 1, en la que el ácido nucleico tiene una secuencia seleccionada del grupo que consiste en:
- a) SEC ID NO: 1, 2 ó 55;
- 15 b) un complemento de la secuencia (o secuencias) en a);
- c) una secuencia con al menos 70% de identidad con una secuencia en a) o b) a lo largo de toda la longitud de la secuencia;
 - d) un fragmento funcional de una secuencia en a) o b);
- en la que la secuencia de c) y el fragmento funcional de d) codifican un polipéptido que es capaz de incrementar la producción de taninos condensados en plantas.
- 20 5. Un polipéptido MYB14 aislado, que comprende al menos uno de:
- a) una secuencia con al menos 70% de identidad con SEC ID NO: 14, en el que el % de identidad se calcula a lo largo de toda la longitud de SEC ID NO: 14; y
 - b) un fragmento funcional de SEC ID NO: 14;
- 25 en el que la secuencia en a) y el fragmento funcional en b) son capaces de incrementar la producción de taninos condensados en plantas.
6. Un constructo que incluye una secuencia nucleotídica como se describe en una cualquiera de las reivindicaciones 1 a 5.
7. El constructo de la reivindicación 6, que incluye:
- 30 - al menos un promotor; y
- la molécula de ácido nucleico;
- en el que el promotor está enlazado operativamente a la molécula de ácido nucleico para controlar la expresión de la molécula de ácido nucleico.
- 35 8. Una célula hospedante, célula vegetal o planta, que se ha alterado a partir del tipo salvaje al transformarla para incluir una molécula de ácido nucleico como se describe en una cualquiera de las reivindicaciones 1 a 4, o un constructo de la reivindicación 6 ó 7.
9. La semilla de una planta de la reivindicación 8, que se altera a partir de una semilla de tipo salvaje, transformándola para incluir una molécula de ácido nucleico como se describe en una cualquiera de las reivindicaciones 1 a 4, o un constructo de la reivindicación 6 ó 7.
- 40 10. Un método para producir una composición que incluye un tanino condensado, en el que el método incluye la etapa de obtener el tanino condensado a partir de una planta de la reivindicación 9, o una parte de la misma.
11. Un método para producir una planta alterada o célula vegetal usando una molécula de ácido nucleico como se describe en una cualquiera de las reivindicaciones 1 a 5 para alterar la planta o célula vegetal, en el que la célula vegetal o planta se altera en al menos una de:
- 45 (a) producción incrementada de al menos un tanino condensado, o monómero del mismo, y

(b) la producción incrementada de tanino condensado se selecciona de catequina, epicatequina, epigallocatequina y galocatequina.

12. Una parte, semilla, fruto, material cosechado, propágulo o progenie de una planta de la reivindicación 9, que se modifica genéticamente para comprender al menos una molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 5, o un constructo de la reivindicación 7 u 8.

5

Figura 1

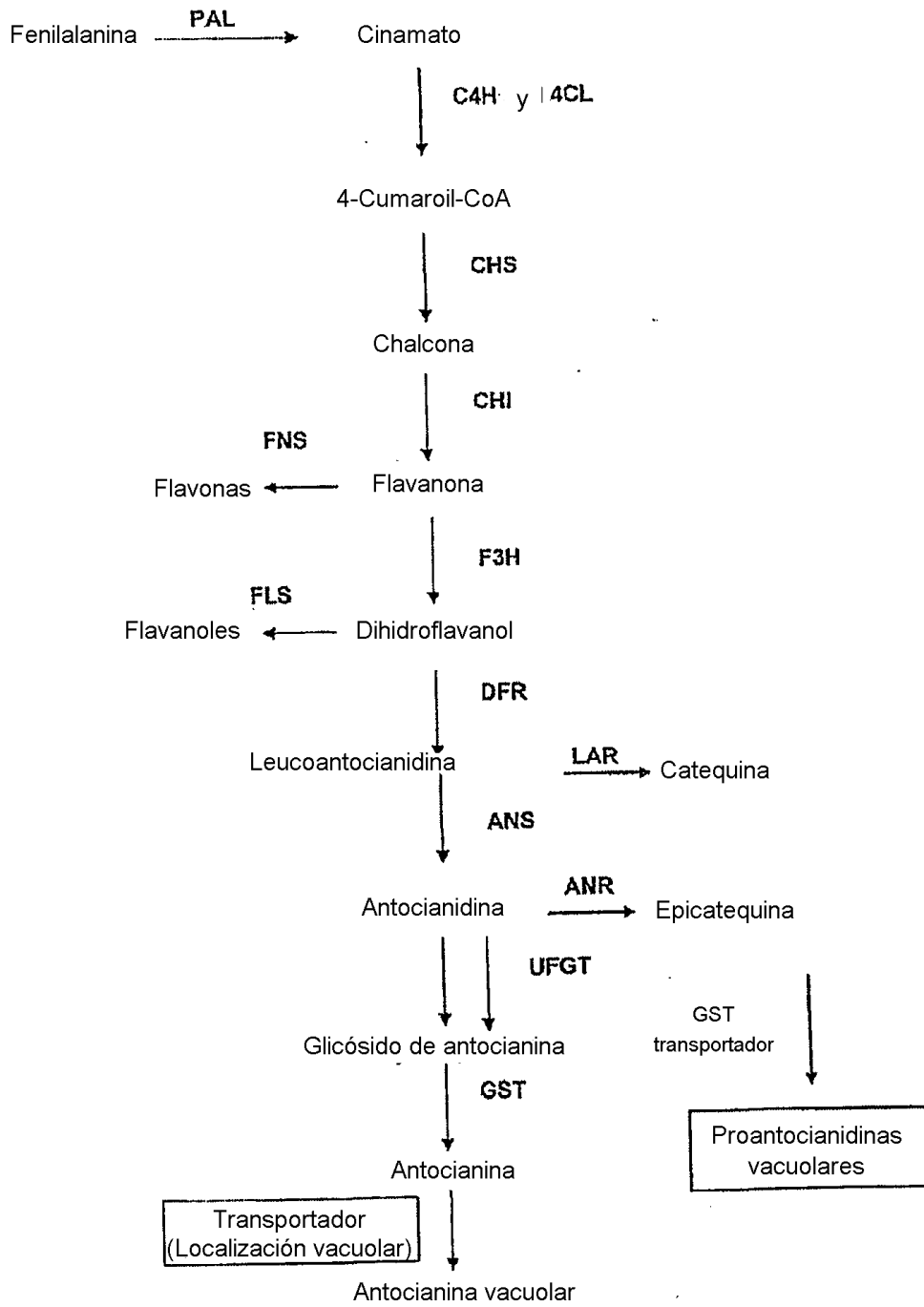


Figura 2A

gaaltcgccttaagcagtggtatcaacgcagagtacgcgggggaagttatattatctacatcaaacacttcaagagg
ttgaatacaagacagactaattaagaataacatcaatgggggagaagccctgtgtgcaaaggaaggcttgaatagaggf
gcttggacaactcaagaagacaaaatcctcactgaatacattaagctccatggtgaaggaaaatggagaaaccttccaaa
aagagcagatttaaaaagatgtggaaaaagtgtgacttagatggttgaattatctaagaccagatattaagcgaggtaata
tatccccggatgaagaagaacttattatccgacttcacaaactactcggaaacagatggtctc taatagccggaagacttcc
agggcgaacagacaatgaaataaagaacttggaaacacaaattaggaaaaaaggtaaggatcttaatcaacaaaaca
ccaacaattcttctc factaaactttctgctcaacaaaaaatgcaagatcaaacagaaacagatcaatcctaagccaat
gaagccaaactcaaatgtgttcgtaacaaaagctaccaagtgttctaaggattgttcataaactcactcccaactcaccia
tgcattgattgcagaacaaagctgaggcagagacaacaacaaagccatcaatgctggttgatggtggtgctagtgattcaa
tgagtaacaacgaaatggaacacgggtatggattttgtcattttgcgatgaagagaagaactatccgcagatttctagaa
gatttaacatcgcggatgataattgcttactgaacttttgaactctgatttctcaaatgctgcaatttcgattacaatgatctatt
gtcaccttgttcggacaaaactcaaatgttctctgatgatgagattcfaagaattggacacaatgtaactttgctgatgagac
aaatgtgtccaacaaccttatttcttcttcttctgaatccagtgaggaagtactaggagaatgataaaaaattcatttt
ccaataaaaataactactctaggfttttttttttttttaatttcaatttcatgtaggggtggttaataataaataatattctatggtta
atattgcaaaaaaaaaaaaaaaaaaaaaaaaaaagactctgcgttgataccactgcttaagggcgaattcc

Figura 2B

MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCGKSCRLRWLNYL
RPDIKRGNISSDEEELIIRLHKLLGNRWSLIAGR L PGR TDNEIKNYWNTNLGKKVKDLNQQNTN
NSSPTKLSAQPKNAKIKQKQINPKPMKPNSNVVVRTKATKCSKVLFINSLPNSPMHDLQNKAEA
ETTTKPSMLVDGVASDSMSNMEMEHGYGFLSFCDEEKELSADLLED FNIADDICLSELLNSDF
SNACNFDYNDLLSPCSDQTQMFSDDEILKNWTQCNFADET N VSNLH SFASFLESSEEV LGE*

Figura 3

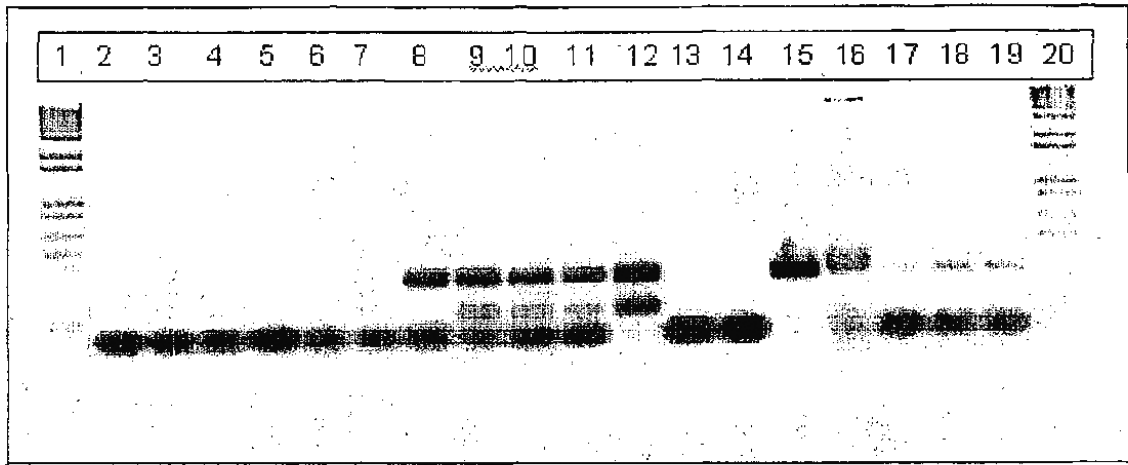


Figura 4

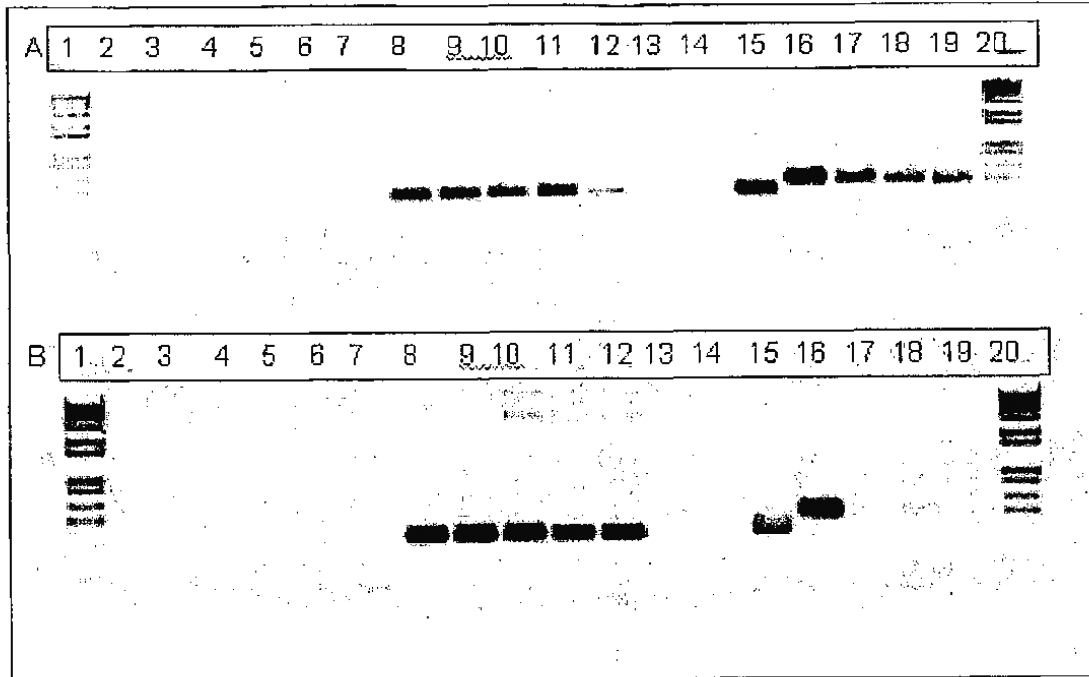


Figura 5

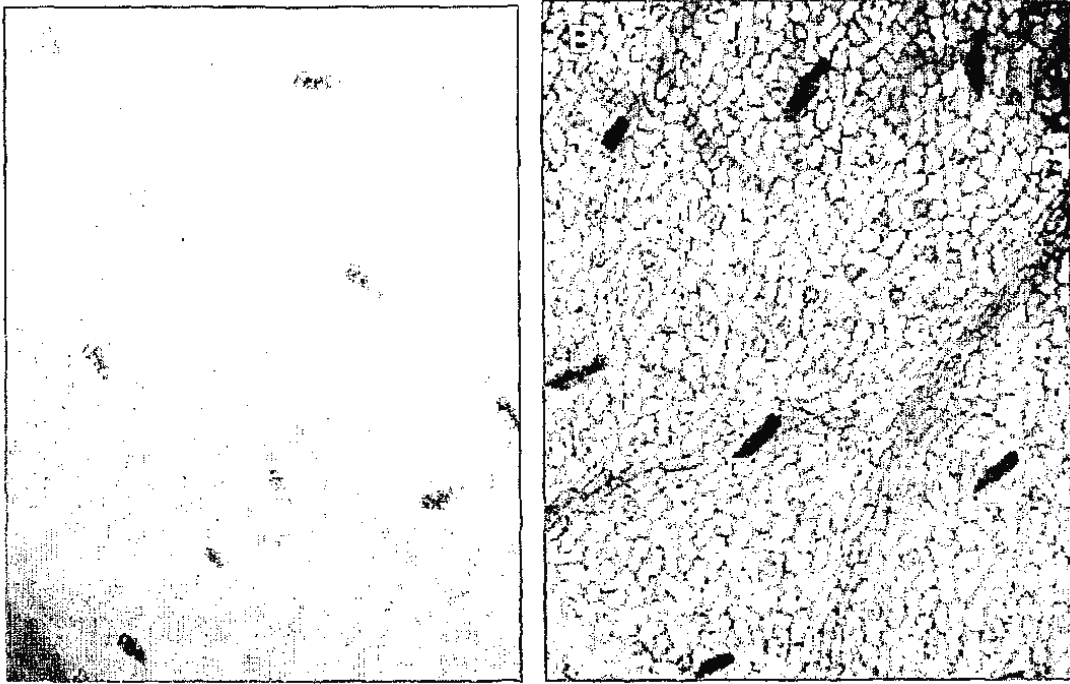
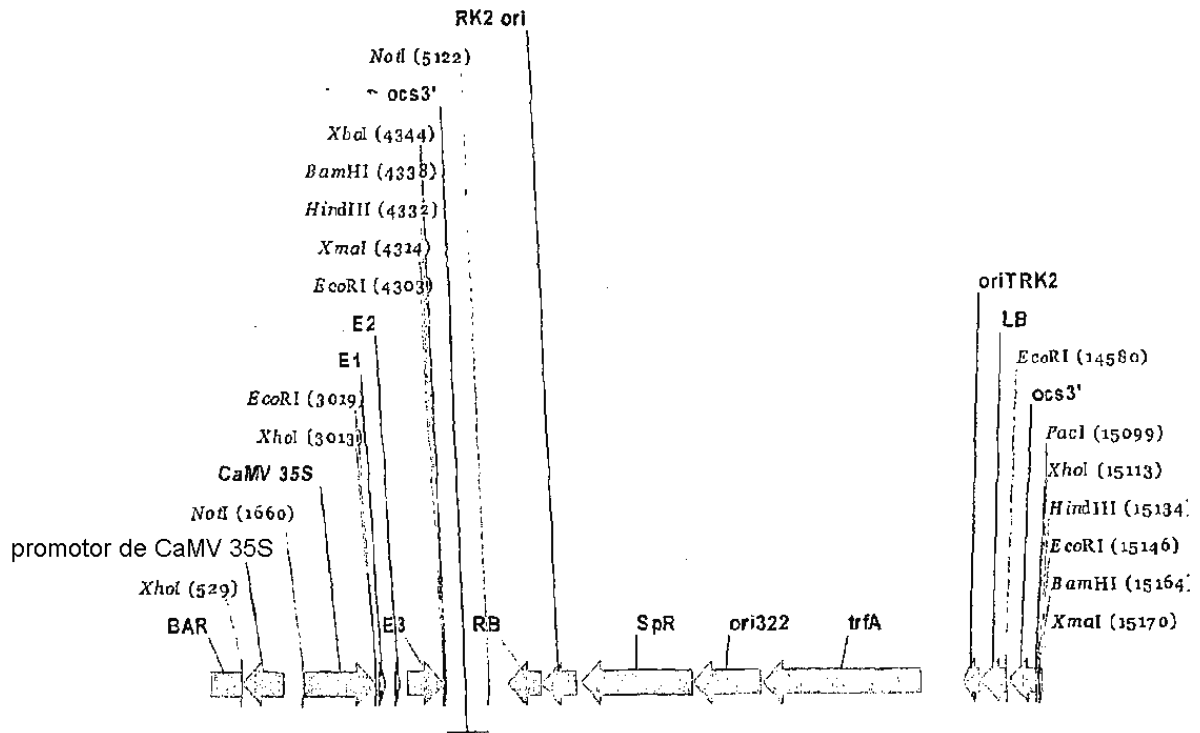


Figura 6



M14ApHZBarP

15212 pb

Figura 7

		101		150
LjTT2a	(1)	-----ATGGGAAAGCCCTTGTGTGTTCA-----		A
MYB14TaF	(101)	CTAATTAAGAATAACATCAATGGGGAGAAGCCCTTGTGTGTTCA-----		A
MYB92Gmax	(1)	-----GCAAAAATGGGAAAGGCTCCTTGTGTGTTCC-----		A
DcMYB3	(1)	-----GAAGAATGGGAAAGGAGCCCTTGTGTGTTCA-----		A
GHMYB10	(1)	-----ATGGGAAAGGAGTCCCTTGTGTGTTCT-----		A
BnTT2-3	(43)	CACAACAACAAGAGAGATGATGAGAAGAGAGAAAAGTAGTAAGGTGAAGA		
GHMYB36	(1)	-----ATGGGAAAGTCCAATTGTGTGTTCC-----		A
		151		200
LjTT2a	(26)	AGCAGGTTTGAACCGAGGTGCCTGGACAGCACAGGAAGACCAAATCCTC		
MYB14TaF	(145)	AGGAAGSCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTC		
MYB92Gmax	(33)	AAGTGGGGTTGCACAAAGGTCCATGGACTCCTAAAAGAAGATGCATTTGCTT		
DcMYB3	(31)	AAGTGGGGCTGAACAAAGGAGCCTGGACCACCTGCTGAGGACAAAATTCCTC		
GHMYB10	(26)	AGGAAAGCCCTTAACAGAGGAGCTTGGACTGCTCTTGAAGACAAAATTCCTT		
BnTT2-3	(93)	AAGAGGAGTTAAACAGAGGGGCTTGGACCGATCAAGAAGACAAAGTCCCTT		
GHMYB36	(26)	AGGAAGGACTCAACAAAGGAGCTTGGACTGCTTTAGAAGATAAAATACCTT		
		201		250
LjTT2a	(76)	CGAGACTATGTTTCACTCCATGGCCAAGSAAATGGAGGACCTTCCCTCA		
MYB14TaF	(195)	ACTGAATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAA		
MYB92Gmax	(83)	ACCAAGTATATCCAAGCTCATGGAGAAGGCCAATGGAAMTCACTACCCAA		
DcMYB3	(81)	ACTGATTCATTCATCTTCACTGCTGAGGTTGATGGAGAAACCTTCCCAA		
GHMYB10	(76)	AAAGATTATATCAAGTACACGGTGAAGGTCGTTGGAGAAATCTCCCAA		
BnTT2-3	(143)	AAAGACTATATCATGTTCCACGGCGAAGGAAAATGGAGCACACTCCAAA		
GHMYB36	(76)	GCATCATATATTCATGTTTCACTGCTGAAGGCAAAATGGAGAAACCTCCCAA		
		251		300
LjTT2a	(126)	AAGTGCAGGTTTGAACCGTGTGGCAAAGCTGTAGACTTAGATGGTTGA		
MYB14TaF	(245)	AAGAGCAGATTTAAAGAGATGTGGAAAAGTGTAGACTTAGATGGTTGA		
MYB92Gmax	(133)	AAAAGCAGGGCTTCTTAGATGTGGAAAAGTGTAGATTAGATGGATGA		
DcMYB3	(131)	AAGAGCAGGTTTGAAGAGATGCCGAAAGAGTGCAGGCTGAGATGGTTGA		
GHMYB10	(126)	AAGAGCTGGTCTTAAGAGATGTGGAAAAGTGTAGGCTTCGGTGGTTGA		
BnTT2-3	(193)	CCAAGCTGGTCTCAAGAGGTGTGGCAAAGCTGCAGACTTCGGTGGAGA		
GHMYB36	(126)	GAGAGCTGGTGTGAAGAGATGTGGCAAAGTGCAGACTTAGATGGCTGA		
		301		350
LjTT2a	(176)	ATTATCTAAGACCAGATATCAAAGAGGCAATATATCCAGAGATGAAGAA		
MYB14TaF	(295)	ATTATCTAAGACCAGATATTAAGCGAGTAATATATCCCCGGATGAAGAA		
MYB92Gmax	(183)	ACTATCTGAGACCAGACATAAAGAGAGGGAACATAGCACCAGAGAAGAT		
DcMYB3	(181)	ATTATTTGAGACCCGATATCAAGAGAGGCAACATTTCTGATGATGAAGAA		
GHMYB10	(176)	ATTATTTGAGACCTGATATTAAGAGGTAACATATCACCTGACGAGGAA		
BnTT2-3	(243)	ACTACTTGAGACCAGGCATAAAGCGCGAAACATCTCATCTGATGAAGAA		
GHMYB36	(176)	ATTATCTTAGACCAGATATTAAGAGGCAACATCTCTCATGATGAAGAA		
		351		400
LjTT2a	(226)	GAGCTTATCATCCGACTTCACAAGCTCCTAGGAAACAGATGGTCTCTAAT		
MYB14TaF	(345)	GAACTTATFATCCGACTTCACAAACTACTCGSAAACAGATGGTCTCTAAT		
MYB92Gmax	(233)	GATCTTATAATCAGAAATGCATTCACCTTTGGGAAACAGATGGTCCCTCAT		
DcMYB3	(231)	GACCTCATCATTCGTCTTCACAAGCTTCTCGGTAATAGGTGGTCTTTAAT		
GHMYB10	(226)	GAGCTTATCATCAAATCCACAAGCTTCTGGGAAACAGATGGTCTTTGAT		
BnTT2-3	(293)	GAACTTATAATCCGCTCCATAATCTCCTTGGAAACAGATGGTGGTTGAT		
GHMYB36	(226)	GAACTCATTATAAGACTCCATAATCTTCTTGGCAACAGATGGTCTTTAAT		
		401		450
LjTT2a	(276)	AGCTGGAAGGCTTCCAGGAAGAACAGACAAATGAGATAAAGAATCTACTGGA		
MYB14TaF	(395)	AGCCGGAAGACTTCCAGGGCGAACAGACAATGAAATAAAGAATCTACTGGA		
MYB92Gmax	(283)	AGCAGGAAGGTTACCAGGGAGAACAGACAATGAAATAAAGAATCTACTGGA		
DcMYB3	(281)	AGCTGGAAGGCTCCCTGGCCGAACAGACAATGAAATAAAGAATCTACTGGA		
GHMYB10	(276)	AGCTGGAAGGCTTCCAGGCGAACAGACAATGAAATAAAGAATCTACTGGA		
BnTT2-3	(343)	AGCTGGGAGGCTTCCAGGGCGAACAGACAATGAAATAAAGAATCTACTGGA		
GHMYB36	(276)	AGCTGGAAGGCTACCCGGCGAACAGACAATGAAATAAAGAATCTACTGGA		

ES 2 537 342 T3

		451		500
LjTT2a	(326)	ACACCAATCTATGTAAGAGCTTCAAGATGGTG---	TTGATGTTGGTGA	
MYB14TaF	(445)	ACACAAATTTAGGAAAAGCTTAAGGA---TC---	TTAATCAACAAA	
MYB92Gmax	(333)	ACACCCATCTAAGCAAAAAGCTGAAAAT-----	TCAGGACAGAA	
DcMYB3	(331)	ACACGCATTTAGGAAAAGGCTCATGATAATCACACTTCATCTGCAGCT		
GHMYB10	(326)	ACACCAACTTAAGTAAAAGAGTTCCGA-----	TCCTCAA	
BnTT2-3	(393)	ACTCAAACCTCCGCAAAAGACTTCCAAAT-----	CTCAAACCA	
GHMYB36	(326)	ACACTACTTTAGGTAAGAGACTAAAGC-----	TCAGCATCCATT	
		501		550
LjTT2a	(372)	CT---CCAAAACCCCATCTTCACAAGAAAAGAACAAATCACCATGATCAGAA		
MYB14TaF	(488)	CA---CCAACAATTCTTCTCCT-----ACTAAACTTTCTGCTCAACCAAAA		
MYB92Gmax	(374)	GA---CACAGACACACAAAATGTAGAGAATCCTCAAGAAGAGGCTGCA		
DcMYB3	(381)	GCTCCAAGACCCCGACTAAC---AATGCAACAAAGAAGACGAAGAA		
GHMYB10	(363)	GT---GACCCGCCCTCCTCG-----AAAAACCCGAGCGGCTCGA		
BnTT2-3	(433)	CCACAGAAAAGTCGAAAACATTCCAA---CAACAACACATGAATAAAGT		
GHMYB36	(367)	GTGCTAAAACGATACCAACCGAGT-CTAGGCTCAATGAACCCCTCGAAA		
		551		600
LjTT2a	(420)	A---GCAAAGCCTCAATCTGTACTCCCTCAGTATTCTCCTCATCACAGC		
MYB14TaF	(531)	AAT-GCAAAGA-TCAAACAGAAACAGA-TCAATCCTAAGCCAAATGA-AGC		
MYB92Gmax	(422)	A---GTGATGGTGGCAACAACAACAAAAGAGAAGAAGAAGAAGAAGACCG		
DcMYB3	(428)	ACACAAGAAGAAGCGCGAGAATCTGAGCCAATTAACCCGGAATCAAGG		
GHMYB10	(403)	CGAGCAACTGCTGGTAAATGGCAATACC---AATGGTAATGGTAGTGGTACT		
BnTT2-3	(481)	ATGTCTTATACGTCCAAAGCGATPAG---GATCCCAAAGGCTCTGACATT		
GHMYB36	(415)	A---GTTCAC-TAAAATCGAAGTATTTCGAACATAAAGCTATTAGGTGT		
		601		650
LjTT2a	(467)	CTAAAACCAAT--AATGTGATTTCGTACAAAGGCATCGAAGTCTCCAAAG		
MYB14TaF	(577)	C---AAACTCA--AATGTETCCCTACAAAAGCTACCAAGTGTTCAGG		
MYB92Gmax	(469)	TGGCAAAAAGA--ACAAGCAGAAGCAAAAGGCAAGAAA---TGATG		
DcMYB3	(478)	C---CAATGCA--TCCGATGTTAGGGCCAAAGCCGCTCTGG-----ACGA		
GHMYB10	(451)	T---CCTCGACA--CACTGGTGGGACAAAGGCGACAAAGTCTCCAAAG		
BnTT2-3	(529)	TCAGAAATCAGAGTAGTATTGGTAGTACCA-GTCTTCTTACTGT---GAAGG		
GHMYB36	(460)	A---GCAGCAAGGTGATGCTCCATPACAACCCTGCAACT---CATCA		
		651		700
LjTT2a	(515)	TGCTGCTCCGGGATCCTCTTCTCC--CTTGCCG-CCAATGCAACCCAG		
MYB14TaF	(622)	TATTGTTCATAAA--CTCA-CTCC--CCAACCTCA-CCAATGCAT-----		
MYB92Gmax	(513)	AGCCGC-CAAAAGACACAAGTTTAC--CTACCAAACCAATTAGAGTGAAG		
DcMYB3	(518)	GGCTGATCAATCAACTCATACTAGTACTAGTACCATGGAGCCATTTGGTTC		
GHMYB10	(497)	TTTT---CATAAACCTCATCAC---ACACA--CAAACAG---A		
BnTT2-3	(576)	AAAACGTGATTGA---TCATCAAGCTGGTTCTCCTTCTGTTGTTGG---G		
GHMYB36	(504)	ACATGGTCAACATCACTGTACAAA-----TAATAATGAAGAAATGGCTGG		
		701		750
LjTT2a	(562)	AGCGACGATTT-CATCGCAAATTATTAGAAGAGCAGAAGGAGAGCCAT		
MYB14TaF	(660)	-----GATTTGCAGAACAAAGCTGAG-GCAGAGACAACAACAAGCCAT		
MYB92Gmax	(560)	GCAATGTATTTACAAAGAACGGATAGTAACACCTTCACCTTTGATTCCAA		
DcMYB3	(568)	AACAAGCATTACAAAATAAGACTACT---GATCAATCTTCGGATCTGG		
GHMYB10	(532)	-----CACCAAAGCCTTCC-----TCAACTT-GTTCAAA		
BnTT2-3	(619)	--A---CAT---CTTAAATTCGATTTT-----GATAAAATTCAGTCTCAG		
GHMYB36	(549)	---TGCTATTGCAACAATTGAAGCTC---ACAATGGAATTCAAATGCTC		
		751		800
LjTT2a	(611)	-----TGCTTTC-TGCTGTGGCCAAATGATTTTACTA-----GTGGCG		
MYB14TaF	(703)	CAATGCTGGTTGA-TGGTGTGGCTAGTGAATCAATGA-----GTAPCA		
MYB92Gmax	(610)	TTTCAGCTAGTGGG-TCAACAAGCCAAAGAGAAGGAGGAAGCCCGTGCACA		
DcMYB3	(613)	TCCCTGGCGTGACTCCACCGACATGTGCTTAAACGGATTTT---CTTAAT		
GHMYB10	(561)	T---CATGGGGA-TCACCGGAACCTAAAACAATGAATGAGTGTATT		
BnTT2-3	(656)	TATCTCTCTCTGATTTAATGGGCTTTGATGTTTGGGT---TGTGGAA		
GHMYB36	(592)	GAGTCATTTGACACTGATGGCGGCTCAAATTTGTTGAGCTT-CGAGATC		
		801		850
LjTT2a	(647)	ACGAAG----ACGGG-GTCTTTCATTTGATCCTTGTGGA--A--ATGAG		
MYB14TaF	(745)	ACGAAATGGAACACG-GTTATGGATTTTGTCAATTTTGGC--ATGAAGAG		
MYB92Gmax	(659)	AAAGAATCAAACGTG-GTTAGTGAAGTTGGTAATGTGGGAGAAGAAAGT		
DcMYB3	(661)	ATGATTTCTCAGTTTGTAAACACTGATATTAATCACCAGGATTCACAC		
GHMYB10	(606)	ACCGAT---AATGTC-----AGAATCCGAGATGAAGGG-----ACGAC		
BnTT2-3	(702)	ACGTAA-----TGCTCTTGTTCATCTGACGAGGTGCTA-----G-GAG		

ES 2 537 342 T3

GHMYB36 (640) AATGAAC----TGTT-----GAAATCACACGATGGTGGAGAAT-TTGAG
851 900
LjTT2a (688) AAGGAAC--CTCCACGGATTGCTCTTGGATTGGACATTGG---TGAA
MYB14TaF (792) AAAGAAC--ATCCGCAGATTGCTAGAGATTTAACATCGCGGATGAT
MYB92Gmax (708) ATGGTTTGGCTTCCTCAGTGAGGACCATGACTTAGTCAACGCTCTCAGAT
DcMYB3 (711) ATGGAGAGCGGCTCGCTTGTTCGTCTGGAGAGCCATATAATGCAGAT
GHMYB10 (642) C-GATCATATATCATCGGATTTTACATTTGACTTCAACATGGG---AGAG
BnTT2-3 (741) ATTATGTT---TCGGCTGAT--GCTTCTGTCTGGGTAACTTT---GAT
GHMYB36 (679) GAGAATCTATGCAGCAGCATTTCCTTGGGTGAGGCAATGCTTAAAG-G
901 950
LjTT2a (733) ATTGCTTG--CCTGAAT-TTATCAACTCA-GATT--TT----TCATATG
MYB14TaF (840) ATTGCTTA--TCIGAAC-TTTTGAACCTT-GATT--TC---TCAAATG
MYB92Gmax (758) ATTGAATGC--CACTCTT-ATTTCCACACA-GATCA-TG---GCCAACCT
DcMYB3 (761) ACTGGAGGAGTCTCGGAA-TGCAGAGGAAC-CATGCTG---GTTTCTA
GHMYB10 (688) TTTGTTTA--TCGGATC-TTTTGAATTC-GATTCTCGGATGTAACG
BnTT2-3 (782) CTTAATAGACCTTTCAT-CTTGTCTTCAAGAAGATTG---TCTCTGGG
GHMYB36 (728) ATTGGTCTACATGTCATTGTCTTGATGACA--ATGG-TG---CCACTGA
951 1000
LjTT2a (773) TGTGTACTTCAGCTACAA-----CAC-TCATGAGGATCTAATG
MYB14TaF (880) CGTGCAATTTGATTACAAATGATCTATTGTCACTTGTTCGGACCAAACT
MYB92Gmax (799) ACAGCAATTTGATGAGAAAT----ATTCCAGCTCTGGAACATGGACC-
DcMYB3 (805) ACTCTAATCTTATTTTACCTC--ATTATCAGAGTGTATTAGTGGGTGAT
GHMYB10 (734) AGCTTAATACAGCAATGTTTGGATTTCCTCACCTTCACCGGATCAGCCT
BnTT2-3 (828) ACTTAAATTTGATAGACCTT-----ATCGTAAA--TCATCATATATACG
GHMYB36 (771) TTTGGAATCATGGCCCTTTTGC---TTGACAC---TGATGAATGGCCA
1001 1050
LjTT2a (811) C---TTTTTCCGAGAACACACTTGTCCAG---GCACAGAA-GTACCCTC
MYB14TaF (930) CAATGTTCTCTGATGATGAGATTTCTCAAGAATTGGACACAATGTAACCT
MYB92Gmax (843) ---ATGGCCAATTCGAACTCAATTCATTG-----CAGAACTTTATT
DcMYB3 (852) TGGTTGGCCTAATATGTGAAAACGGGAAGTG-----TACATTTACTGT
GHMYB10 (784) CCTATGGATTTCTCCGACGAAA-TGCTAAA-----AGAGTGGACGGC
BnTT2-3 (870) TC-TACCTCTGTACGAAACAAAGTATATATTTATATTTCTGTTGAACGCT
GHMYB36 (814) TGA-----
1051 1100
LjTT2a (853) GG-TGATGAAACAATTTGGTAAATAA-----TTGTTT-----
MYB14TaF (980) TGCTGATGAGACAAATGTTCCAAACAACCTTCATTCCTTTGCTTCTTTC
MYB92Gmax (883) AG---ATTAAAAGAATATCAACAAGAT-----TTGTTCACTTCA---TG
DcMYB3 (897) TGTTTCATTTACTTAACTTCCCGGAAATAAAGATGCATGTATCATAGTTC
GHMYB10 (825) CGCCGCTTCACTCACTGCTGTCAACCAAGTGGCGCTTCCAATC---TC
BnTT2-3 (919) TCTAATTACAAGTAAATCT-----
GHMYB36 (817) -----
1101 1150
LjTT2a (886) -----AATGAGGAGA---AGGAT---AATGGTTGCTAA-----
MYB14TaF (1030) TTGAATCCAGTGAGGAAGTACTNGSAG---AATGATAATAAAAATTCATT
MYB92Gmax (922) AAGATCACATTGCTTACAT---ATAAA---CTTTGTTGATAGATCATATG
DcMYB3 (947) AA---ATAATGACTACTTCTGATGTGTGAATGTTGTGTAATAAAAAA
GHMYB10 (871) CAGT---CCTTGCCCTCCATTATTGAA---AATGGAATTGAATGA-----
BnTT2-3 (939) -----
GHMYB36 (817) -----
1151 1200
LjTT2a (913) -----
MYB14TaF (1077) TTCCAATAAAAATTAACCTACTCTAGGTTTTTTTTTTTTTTTAAATTC
MYB92Gmax (966) TAAATATATCTGTAATGATCTCTGAGTTATGAGATCTTTTTTGTCTTTA
DcMYB3 (993) AAAAAAA-----
GHMYB10 (910) -----
BnTT2-3 (939) -----
GHMYB36 (817) -----
1201 1250
LjTT2a (913) -----
MYB14TaF (1127) ATTTTCATGTTAGGGTGGTTTAATAAATAAATATACTCTATGGTTAATAT
MYB92Gmax (1016) ATAAATATPCGCCATCTAACTCAAAAAAAAAAAAA-----
DcMYB3 (1001) -----
GHMYB10 (910) -----
BnTT2-3 (939) -----
GHMYB36 (817) -----

ES 2 537 342 T3

		1251	1300
LjTT2a	(913)	-----	
MYB14TaF	(1177)	TGCAAAAAAAAAAAAAAAAAAAAAAAAAAGTACTCTGCGTTGATAACCACT	
MYB92Gmax	(1050)	-----	
DcMYB3	(1001)	-----	
GHMYB10	(910)	-----	
BnTT2-3	(939)	-----	
GHMYB36	(817)	-----	
		1301	1317
LjTT2a	(913)	-----	
MYB14TaF	(1227)	GCTTAAGGGCGAATTCC	
GHMYB36	(817)	-----	

Figura 8

		1	50
At TT2	(1)	-MGRATTSVRREELNRGAWTDHEDKILRDYITTHGEGKWSTLPNQAGLK	
BnTT2-1	(1)	MMRARESSKVKKEELNRGAWTDQEDKILKDYIMFHGEGKWSTLPNQAGLK	
Zm P1	(1)	-MGR--ACCAKEGVKRGAWTAKEDDTLAAYVKAHGEGKWRNLPKQAGLR	
MYB10Gh	(1)	-MGRS--PCCSKEGLNRGAWTALEDKILKDYIKVHGEGRWRNLPKRAGLK	
MYB14FTa	(1)	-MGRS--PCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLK	
VvMYBPA1	(1)	-MGR--PCCSKVGLHRGSWTAREDTLLTKYIQAKGEGHWRSLPKKAGLL	
LjTT2a	(1)	-MGRS--PCCSKQGLNRGAWTAQEDQILRDYVHLHGQGWNRNLPQSAGLK	
MYB185Gmax	(1)	-MGR--PCCSKVGLHRGFWTPREDAILLTKYIQTHGEGQWRSLPKRAGLL	
MYB11Malus	(1)	-MGRSP-CCSKDEGLNRGAWTAMEDKVLTEYIGNHGEGKWRNLPKRAGLK	
		51	100
At TT2	(50)	RCGKSCRLRWKNYLRLPGIKRGNISSDEEELIIRLHNLLGNRWSLIAGRLP	
BnTT2-1	(51)	RCGKSCRLRWKNYLRLPGIKRGNISSDEEELIIRLHNLLGNRWSLIAGRLP	
Zm P1	(48)	RCGKSCRLRWLNLYLRPNIKRGNISYDEEDLIVRLHKLLGNRWSLIAGRLP	
MYB10Gh	(48)	RCGKSCRLRWLNLYLRPDIKRNISPOEELIIRLHKLLGNRWSLIAGRLP	
MYB14FTa	(48)	RCGKSCRLRWLNLYLRPDIKRNISDEEELIIRLHKLLGNRWSLIAGRLP	
VvMYBPA1	(48)	HCGKSCRLRWLNLYLRPDIKRNITPDKDDLIIIRLKSLLGNRWSLIAGRLP	
LjTT2a	(48)	RCGKSCRLRWLNLYLRPDIKRNISRDEEELIIRLHKLLGNRWSLIAGRLP	
MYB185Gmax	(48)	RCGKSCRLRWLNLYLRPDIKRNITPEEDDLIVRMHSLGNRWSLIAGRLP	
MYB11Malus	(49)	RCGKSCRLRWLNLYLRPDIKRNITRDEEELIIRLHKLLGNRWSLIAGRLP	
		101	150
At TT2	(100)	GRTDNEIKMFWNSNLRKRLPKTQTK-----OPKRIKHS	
BnTT2-1	(101)	GRTDNEIKMFWNSNLRKRLPKSQTN-----QOKSRKHS	
Zm P1	(98)	GRTDNEIKNYWNTNLGRRAGAAGAS-----RVVFAPDT	
MYB10Gh	(98)	GRTDNEIKNYWNTNLSKRVS DRQKSP-----AAPSCKPEAARRGTA	
MYB14FTa	(98)	GRTDNEIKNYWNTNLGKVKVDLNQONTNNSPTKLSAQPKNAKIKQKQIN	
VvMYBPA1	(98)	GRTDNSIKNYWNTNLSKRLRSQGTDPNTH-----KKMTEPPEPKRRKNT	
LjTT2a	(98)	GRTDNEIKNYWNTNLSKRVQDGVVGVGSKTPSSQEKNNHDKAKFQSVZ	
MYB185Gmax	(98)	GRTDNEIKNYWNTNLSKRLRNOGTDP-----KTHDKLTEAPEKKGK	
MYB11Malus	(99)	GRTDNEIKNYWNTNLSKRVQVEGRSC-----SDGNRRPTQEKPKPT	
		151	200
At TT2	(133)	TNN----ENNVCVIRTKAIRCSKTLFSDLSLQK-KSSTSPUPLKEQEM	
BnTT2-1	(134)	NNNN----MNVCVIRPKAIRPKALTFQNSS---IGSTSLTVKENVI	
Zm P1	(131)	GSHA----TPAASGSRMTGGQKGAAPRADLGS---PGSAAVWAPKAAR	
MYB10Gh	(139)	GNGN----TNGSGSGSSTHVVRTRATRCSKVFIN-PHHHTONRHPKPS	
MYB14FTa	(148)	PKPM----KPNVVRTKATKCSKVLFINSLPNSP-MHDLQNAEAETTT	
VvMYBPA1	(142)	RTRT----NNGGSKRVKISKDLNSNHKVLHPKP-VRVTSLISMSRNS	
LjTT2a	(148)	PSVFSSQPKNNVIRTKASKCSKVLRLDPLLPCPPMQTQSDDFIAKLE	
MYB185Gmax	(139)	KKKNKQKNENKSGSEKTLVYLPKPIRVKALSSCIPTDSTLTLNNSATA	
MYB11Malus	(140)	LSPKPTNISCTKVRTKASRCKVLPHEKQKFGYSTEQVVNAAPTLDQ	
VI/VRTKaxR/KxSK (Nuevo motivo asociado con TFs de MYB que regulan las rutas de CTs)			
		201	250
At TT2	(177)	DQGGSS-----LMGDLEFDLDRHSEFHFPDL-MDFDGLDCGNVMSLVS	
BnTT2-1	(177)	DHQAGSPS----LLGDLKIDFDKIQSEYLFSDL-MDFDGLGCGNVMSLVS	
Zm P1	(174)	CTGGLFFHRTPHAGETETPTPMMAGGGGGGEA-RSSDDCSSAASVSPLV	
MYB10Gh	(184)	TCSNHGDHREPKTMLNELLPIMSSENEGTTDH-ISSDFTFDENMGFCLE	
MYB14FTa	(193)	KPSMLVDGVASDSMSNNEMEHGYGFLSFCDEEKELSADLLEDENIADDIC	
VvMYBPA1	(187)	FESNTVSGSGSGSSGGNGESLPWPSFRDIRDDKVIGVDGVDFEFGDQGGQ	
LjTT2a	(198)	EAEGEPLLSAVANDFTSGDEQVLSFDPCGNEKELESTDLLLDLQIGEICL	
MYB185Gmax	(189)	STSEEK-----VQS-PEAEVKEVNMVWVGDDADNNGGIEIFFGEDHDLV	
MYB11Malus	(190)	AVNNPM-----VGIDPLPMSFLDDENNNSCFELVDFKMDENFL	
		251	300

ES 2 537 342 T3

At TT2 (220) SNEILGELVP----AQGNLDLNRPF~~TSCHHRGDE~~EDWLRDFTC-----
 BnTT2-1 (222) SDEVLGDYVSTDTSCLG~~NLDLNRPF~~TSCLQ----EDCLWDFNC-----
 Zm P1 (223) GSSQHDP~~CFSGDG~~-DGDWMD~~DV~~RALASFLS--DEEWLRCHTATQ~~LV~~---
 MYB10Gh (233) SDL~~LN~~SDFCDV~~NELNYS~~NGFDSS~~PSP~~QPPM~~DFS~~DEMLKEWTAA~~ASTHCC~~
 MYB14FTa (243) LSE~~LL~~NSDFS-NACNFDY~~N~~DL~~L~~SPCSD~~Q~~TMFS~~DD~~EILKNW~~T~~QCNFADET
 VvMYBPA1 (237) DLVASSDPESQSKMPPT~~DN~~SLDK~~L~~VEEY~~L~~QL~~L~~EREDTQV~~OLD~~STAESLLI
 LjTT2a (248) PEF~~LN~~SDFS~~Y~~VCDFS~~YN~~THE~~D~~LML~~F~~SENTLVQA~~K~~YLG~~DE~~TNL~~V~~NNCFNE
 MYB185Gmax (232) NNTASYEECYSDVHT~~DD~~HGTLE~~K~~LMEY~~L~~QL~~L~~NVEEK~~P~~DEL~~D~~STAQSLLV
 MYB11Malus (230) SDE~~LN~~VDFS~~V~~LYNNEGAGKAAAAAT~~T~~EDTSNKL~~H~~GPDL~~R~~SSKAP~~I~~IESEL
 DEXWRLxxT (Motivo de subgrupo 5; Stracke et al., 2001)

		301	323
At TT2	(259)	-----	-----
BnTT2-1	(261)	-----	-----
Zm P1	(267)	-----	-----
MYB10Gh	(283)	HQSAASNLC SL PPFIENGIE---	
MYB14FTa	(292)	NVSN N LHSEFAS F LESSEEV L GE-	
VvMYBPA1	(287)	-----	-----
LjTT2a	(298)	EKDNGC-----	-----
MYB185Gmax	(282)	-----	-----
MYB11Malus	(280)	DCWLVDN-----	-----

Figura 9

		1		50
MYB14FTa	(1)	MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
TaMYB14-2S	(1)	MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
TrMYB14f	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
TrMYB14d	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
ToMYB14	(1)	MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
Taf11cDNA	(1)	MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
Consenso	(1)	MGRSPCCAKEGLNRGAWTTQEDKILTEYIKLHGEGKWRNLPKRAGLKR	CG	
		51		100
MYB14FTa	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
TaMYB14-2S	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
TrMYB14f	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
TrMYB14d	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
ToMYB14	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
Taf11cDNA	(51)	KSCRLRWLNLYLRDIDKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
Consenso	(51)	KSCRLRWLNLYLRPDIKRGNISDDEEELIIRLHKLLGNRWSLIAGRLPGRT		
		101		150
MYB14FTa	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKLSAQPKNAKIKQKQ-INPK		
TaMYB14-2S	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKLSAQPKNAEIKQKQ----I		
TrMYB14f	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKPSAQPKNANIKQKQINPK		
TrMYB14d	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKPSAQPKNANIKQKQINPK		
ToMYB14	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKLSAQPKNAKIKQKQ-INPK		
Taf11cDNA	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKLSAQLKNAKIKQKQ-INPK		
Consenso	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSSPKLSAQPKNAKIKQKQ-INPK		
		151		200
MYB14FTa	(150)	PMKPNSNVVRTKATKCSKVLFINSLEPNSP-MHDLQNKAEAETTQK-----		
TaMYB14-2S	(147)	NPKPNSYVVRTKATKCSKVLEINSPEPNSPMMIDLQSKAEAETTQTKPSM		
TrMYB14f	(151)	PMKPNSNVVRTKATKCSKVLFINSPP-----MHNLQNKAEAETTKK-----		
TrMYB14d	(151)	PMKPNSNVVRTKATKCSKVLFINSPEP-----MHNLQNKAEAETTKK-----		
ToMYB14	(150)	PMKPNSNVVRTKATKCSKVLFINSLEPNSP-MHDLQNKAEAETTQK-----		
Taf11cDNA	(150)	PMPEPNSNVVRTKATKCSKALFINSPEPNSPMMHDLQNKAEAETTQKS--SM		
Consenso	(151)	PMKPNSNVVRTKATKCSKVLFINSPEPNSP-MHDLQNKAEAETTQK-----		
		201		250
MYB14FTa	(194)	PSMLVDGVASDSMSNNEHEGYGFLSFCDEEKELSADLLEDFNIADDICL		
TaMYB14-2S	(197)	PSMLVDGVASDSMSNNEHECGNGFLSFCDEEKELSADLLEDFNIADDICL		
TrMYB14f	(192)	PLMLVNGVASDSMSNNEHEMERNNGFLSFCDEEKELSADLLEDFNIADDICL		
TrMYB14d	(192)	PLMLVNGVASDSMSNNEHEMERNNGFLSFCDEEKELSADLLEDFNIADDICL		
ToMYB14	(194)	PSMLVDGVASDSMSNNEHEGYGFLSFCDEEKELSADLLEDFNIADDICL		
Taf11cDNA	(198)	PSMLVDGVASDSMSNNEHEYGDGFVSFCDDKELSADLLEDFNIADDICL		
Consenso	(201)	PSMLVDGVASDSMSNNEHEGNGFLSFCDEEKELSADLLEDFNIADDICL		
		251		300
MYB14FTa	(244)	SELLNSDFSNA CNFDYNDLLSPCS DQTQMFSDDEILKNWTQCNFADETN-		
TaMYB14-2S	(247)	SEFLNFDFSNA CDYNDLLSPCS DQTQMFDPDEILKNWTQCNFADETN-		
TrMYB14f	(242)	SEFLNSDFSNA CNFDCNDLLSPCS DQTQMFSDDEILKNWTQCNFADETN-		
TrMYB14d	(242)	SEFLNSDFSNA CNFDYNDLLSPCS DQTQMFSDDEILKNWTQCNFADETN-		
ToMYB14	(244)	SELLNSDFSNA CNFDYNDLLSPCS DQTQMFSDDEILKNWTQCNFADETN-		
Taf11cDNA	(248)	SEFLNFDFSNA CNFDYNDLLSPCS DQTQMFSDDEILKNSTPCNFAAETNY		
Consenso	(251)	SEFLNSDFSNA CNFDYNDLLSPCS DQTQMFSDDEILKNWTQCNFADETNV		

		301		321
MYB14FTa	(294)	VSNNLHSFASFLESSEEV	LGEE-	
TaMYB14-2S	(297)	VSNNLQSSASFLESSEEV	LGEE-	
TrMYB14f	(292)	VSNNLNSFASFLESSEEV	LGEE-	
TrMYB14d	(292)	VSNNLHSFASFLESSEEV	LGEE-	
ToMYB14	(294)	VSNNLHSFASFLESSEEV	LGEE-	
Taf11cDNA	(298)	VSNN-QS-----S	EEVLGE-	
Consenso	(301)	VSNNLNSFASFLESSEEV	LGEE-	

ES 2 537 342 T3

Figura 10

		1	50
TRM4	(1)	-----	-----
TRM6	(1)	-----	-----
TRM3	(1)	-----	-----
TRM1	(1)	-----	-----
TRM5	(1)	-----	-----
TRM14	(1)	-----	-----
MYB14TaF	(1)	GAATTCGCCCTTAAGCAGTGGTATCAACGCAGAGTACGCGGGGAAGTTA	
TaM3	(1)	-----	-----
TaM4	(1)	-----	-----
		51	100
TRM4	(1)	-----	-----
TRM6	(1)	-----	-----
TRM3	(1)	-----	-----
TRM1	(1)	-----	-----
TRM5	(1)	-----	-----
TRM14	(1)	-----	-----
MYB14TaF	(51)	TTTAATTTTATCTACATCAAACACTTCAAGAGGTTGGAATACAAGACAGA	
TaM3	(1)	-----GAATTCGCCCTTAGGTTGGAATACAAGACAGA	
TaM4	(1)	-----GAATTCGCCCTTAGGTTGGAATACAAGACAGA	
		101	150
TRM4	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
TRM6	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
TRM3	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
TRM1	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
TRM5	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
TRM14	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAAAGAA	
MYB14TaF	(101)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAAGGAA	
TaM3	(33)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAAGGAA	
TaM4	(33)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAAGGAA	
		151	200
TRM4	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
TRM6	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
TRM3	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
TRM1	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
TRM5	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
TRM14	(43)	GGCTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
MYB14TaF	(150)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
TaM3	(82)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
TaM4	(82)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
		201	250
TRM4	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TRM6	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TRM3	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TRM1	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TRM5	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TRM14	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
MYB14TaF	(200)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TaM3	(132)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
TaM4	(132)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAAGAG	
		251	300
TRM4	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	
TRM6	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	
TRM3	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	
TRM1	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	
TRM5	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	
TRM14	(143)	CAGGTTCAATTCATTCT-GTATCTTACTATTATAGATCAATAATCACTTTC	

ES 2 537 342 T3

MYB14TaF (250) CAG-----
TaM3 (182) CAGGTTCAATTCATTCTAGTATCTTGCAATATAGATCAAT---CACTTTC
TaM4 (182) CAGGTTCAATTCATTCT-GTATCTTACAATATAGATTAAC---CACTTTC

301 350
TRM4 (192) ACACTTTTTTTTTA---CTTATAAATTTTCATGTATTTTTCTTCCATTT
TRM6 (192) ACACTTTTTTTTTA---CTTATAAATTTTCATGTATTTTTCTTCCATTT
TRM3 (192) ACACTTTTTTTTTA---CTTATAAATTTTCATGTATTTTTCTTCCATTT
TRM1 (192) ACACTTTTTTTTTT--ACTTATAAATTTTCATGTATTTTTCTTCCATTT
TRM5 (192) ACACTTTTTTTTTTTACTTAIAAATTTTCATGTATTTTTCTTCCATTT
TRM14 (192) ACACTTTTTTTTTTT-ACTTATAAATTTTCATGTATTTTTCTTCCATTT
MYB14TaF (253) -----
TaM3 (229) ATACTTTTGTGTTG---CTTATAAATTTCTTGCATTTTTCTTCAATTT
TaM4 (228) ATACTTTTGTGTTG---CTTATAAATTTCTTGTATTTTTCTTCCATTT

351 400
TRM4 (239) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
TRM6 (239) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
TRM3 (238) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
TRM1 (240) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
TRM5 (242) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
TRM14 (241) TCCATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTTTTCA
MYB14TaF (253) -----
TaM3 (275) TCCATGTGAAATGCAAATTACTAGTACATTATTATGGATATGTTTTTGCA
TaM4 (274) TCCATGAGAAATGCAAATTACTAGTACATTATTATGGACATGTTTTTGCA

401 450
TRM4 (289) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
TRM6 (289) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
TRM3 (288) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
TRM1 (290) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
TRM5 (292) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
TRM14 (291) AAAATGTGTATTCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGAC
MYB14TaF (253) -----ATTAAAAAGATGTGGAAAAAGTTGTAGAC
TaM3 (325) AATATGTGTATGCCATGCAGGTTTAAAAAGATGCGGAAAAAGTTGTAGAC
TaM4 (324) AATATGTTTATGCCATGCAGATTAAAAAGATGTGGAAAAAGTTGTAGAC

451 500
TRM4 (339) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
TRM6 (339) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
TRM3 (338) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
TRM1 (340) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
TRM5 (342) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
TRM14 (341) TAAGGTGGTTGAATTATCTTAGACCGGATATTAAGAGAGGTAATATATCC
MYB14TaF (283) TTAGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCC
TaM3 (375) TTAGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCC
TaM4 (374) TTAGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCC

501 550
TRM4 (389) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
TRM6 (389) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
TRM3 (388) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
TRM1 (390) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
TRM5 (392) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
TRM14 (391) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACCG
MYB14TaF (333) CCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACAG
TaM3 (425) TCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACAG
TaM4 (424) CCGGATGAAGAAGAACTTATCATTAGACTTCACAACTACTCGGAAACAG

551 600
TRM4 (439) GTAAA-GTATCGACATAATCACTAACTTACTAACATT-----TG
TRM6 (439) GTAAA-GTATCGACATAATCACTAACTTACTAACATT-----TG
TRM3 (438) GTAAA-GTATCGACATAATCACTAACTTACTAACATT-----TG
TRM1 (440) GTAAA-GTATCGACATAATCACTGACTTACTAACATT-----TG

ES 2 537 342 T3

```

TRM5 (442) GTAAA-GTATCGACATAATCACTAÄCTTACTAACATT-----TG
TRM14 (441) GTAAA-GTATCGACATAATCACTAÄCTTACTAACATT-----TG
MYB14TaF (383)
TaM3 (475) GTAAAAGTACCGACATAATCACTAÄCTTATTAAACÄTTTATCTATAAÄTTTG
TaM4 (474) GTAAA-GTCCCTAACATAAÄCTTACTAACÄTTATTAAACGTTTGTCTATAAÄTTTG

601 650
TRM4 (477) T-----TTATAATGTGTGCTAAT
TRM6 (477) T-----TTATAATGTGTGCTAAT
TRM3 (476) T-----TTATAATGTGTGCTAAT
TRM1 (478) T-----TTATAATGTGTGCTAAT
TRM5 (480) T-----TTATAATGTGTGCTAAT
TRM14 (479) T-----TTATAATGTGTGCTAAT
MYB14TaF (383)
TaM3 (525) TTTTTTTTGACAÄTTAGTACTACTAÄTTAÄTTTATAÄTTGTGTGCTAÄT
TaM4 (523) TTTTTTTTGACCÄTTAGTACTACTAÄTTAÄTTTACAÄTTGTGTGCTAÄT

651 700
TRM4 (495) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TRM6 (495) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TRM3 (494) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TRM1 (496) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TRM5 (498) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TRM14 (497) T-GCTCTTCCTTTGÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
MYB14TaF (383) -----ATGGTCTCTAÄTAGCCGGAAGÄCTT
TaM3 (575) TTGCTTTTGCCTTTÄÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT
TaM4 (573) TTGCT-TGTCTTTÄÄTTTGTGGTAGÄTGGTCTCTAÄTAGCCGGAAGÄCTT

701 750
TRM4 (544) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TRM6 (544) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TRM3 (543) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TRM1 (545) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TRM5 (547) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TRM14 (546) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
MYB14TaF (408) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TaM3 (625) CCAGGACGÄACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG
TaM4 (622) CCAGGGCGAACAGÄCAÄTGAÄATAÄGAÄCTACTGGÄACACÄÄÄTTTAGG

751 800
TRM4 (594) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TRM6 (594) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TRM3 (593) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TRM1 (595) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TRM5 (597) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TRM14 (596) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
MYB14TaF (458) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TaM3 (675) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA
TaM4 (672) AÄÄÄÄÄÄGTTAÄGGÄTCTTÄÄTCAÄCAÄÄÄÄCÄÄCAÄÄTCTTCTCCTA

801 850
TRM4 (644) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
TRM6 (644) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
TRM3 (643) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
TRM1 (645) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
TRM5 (647) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
TRM14 (646) CTÄÄÄCCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄTATCAÄÄCAGÄÄÄÄCÄÄCÄG
MYB14TaF (508) CTÄÄÄCTTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄGATCAÄÄCAGÄÄÄÄCÄÄCÄG
TaM3 (725) CTÄÄÄCTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄGATCAÄÄCAGÄÄÄÄCÄÄCÄG
TaM4 (722) CTÄÄÄCTCTGCTCAÄCCÄÄÄÄÄÄÄTGCÄÄÄGATCAÄÄCAGÄÄÄÄCÄÄCÄG

851 900
TRM4 (694) ATCAÄTCTTÄÄGCCÄÄTGAÄGCCÄÄÄÄCTCGÄÄTGTGTCCGTÄÄÄÄÄGCT
TRM6 (694) ATCAÄTCTTÄÄGCCÄÄTGAÄGCCÄÄÄÄCTCGÄÄTGTGTCCGTÄÄÄÄÄGCT

```

ES 2 537 342 T3

TRM3	(693)	ATCAATCCTAAGCCAATGAAGCCAACTCGAATGTTGTCCGTACAAAAGC	
TRM1	(695)	ATCAATCCTAAGCCAATGAAGCCAACTCGAATGTTGTCCGTACAAAAGC	
TRM5	(697)	ATCAATCCTAAGCCAATGAAGCCAACTCGAATGTTGTCCGTACAAAAGC	
TRM14	(696)	ATCAATCCTAAGCCAATGAAGCCAACTCGAATGTTGTCCGTACAAAAGC	
MYB14TaF	(555)	ATCAATCCTAAGCCAATGAAGCCAACTCAATGTTGTCCGTACAAAAGC	
TaM3	(772)	ATCAATCCTAAGCCAATGAAGCCAACTCAATGTTGTCCGTACAAAAGC	
TaM4	(769)	ATCAATCCTAAGCCAA-----ACTCATATGTTGTCCGTACAAAAGC	
		901	950
TRM4	(744)	TACCAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
TRM6	(744)	TACCAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
TRM3	(743)	TACCAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
TRM1	(745)	TACCAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
TRM5	(747)	TACCAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
TRM14	(746)	TACCAATTGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA-	
MYB14TaF	(605)	TACCAAGTGTCTAAGGTATTGTTTCATAAACTCACTCCCAACTCACCA-	
TaM3	(822)	TACCAAGTGTCTAAGGTATTGTTTCATAAACTCACTCCCAACTCACCA-	
TaM4	(810)	TACCAAGTGTCTAAGGTATTGTTTCATAAACTCACCCCAACTCACCA-	
		951	1000
TRM4	(784)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
TRM6	(784)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
TRM3	(783)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
TRM1	(785)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
TRM5	(787)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
TRM14	(786)	--ATGCATAATTTGCAGAACAAAGCTGAGGCAGAGACAAAACAAA----	
MYB14TaF	(654)	--ATGCATGATTTGCAGAACAAAGCTGAGGCAGAGACAACAACAAA----	
TaM3	(871)	--ATGCATGATTTGCAGAACAAAGCTGAGGCAGAGACAACAACAAA----	
TaM4	(860)	CAATGCATGATTTGCAGAGCAAAGCTGAGGCAGAGACAACAACAACA	
		1001	1050
TRM4	(828)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
TRM6	(828)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
TRM3	(827)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
TRM1	(829)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
TRM5	(831)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
TRM14	(830)	-----GCCATTAATGCTGGTTAATGGTGTAGCTAGTGATTCAAT	
MYB14TaF	(698)	-----GCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAAT	
TaM3	(915)	-----GCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAAT	
TaM4	(910)	AAGCCATCAATGCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAAT	
		1051	1100
TRM4	(867)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
TRM6	(867)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
TRM3	(866)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
TRM1	(868)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
TRM5	(870)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
TRM14	(869)	GAGTAACAACGAAATGGAACGCGGTAATGGATTTTGTGATTTTGCGACG	
MYB14TaF	(737)	GAGTAACAACGAAATGGAACACGGTTATGGATTTTGTGATTTTGCGATG	
TaM3	(954)	GAGTAACAACGAAATGGAACACGGTTATGGATTTTGTGATTTTGCGATG	
TaM4	(960)	GAGTAACAACGAAATGGAATGCGGTAATGGATTTTGTGATTTTGCGATG	
		1101	1150
TRM4	(917)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
TRM6	(917)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
TRM3	(916)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
TRM1	(918)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
TRM5	(920)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
TRM14	(919)	AAGAGAAAGAACTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGAT	
MYB14TaF	(787)	AAGAGAAAGAACTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGAT	
TaM3	(1004)	AAGAGAAAGAACTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGAT	
TaM4	(1010)	AAGAGAAAGAACTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGAT	

ES 2 537 342 T3

		1151		1200
TRM4	(967)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
TRM6	(967)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
TRM3	(966)	GATATTTGCTTACCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
TRM1	(968)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
TRM5	(970)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
TRM14	(969)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
MYB14TaF	(837)	GATATTTGCTTATCTGAACTTTTGAACCTCTGATTTCTCAAATGCCGTGCAA		
TaM3	(1054)	GATATTTGCTTATCTGAACTTTTGAACCTCTGATTTCTCAAATGCCGTGCAA		
TaM4	(1060)	GATATTTGCTTATCTGAATTTCTAAACTCCGATTTCTCAAATGCCGTGCAA		
		1201		1250
TRM4	(1017)	TTTCGATTACAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
TRM6	(1017)	TTTCGATTGCAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
TRM3	(1016)	TTTCGATTACAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
TRM1	(1018)	TTTCGATTACAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
TRM5	(1020)	TTTCGATTACAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
TRM14	(1019)	TTTCGATTACAATGATCTATTGTGCGCCTTGTTCCGGATCAAACCTCAAATGT		
MYB14TaF	(887)	TTTCGATTACAATGATCTATTGTGACCTTGTTCCGGACCAAACCTCAAATGT		
TaM3	(1104)	TTTCGATTACAATGATCTATTGTGACCTTGTTCCGGACCAAACCTCAAATGT		
TaM4	(1110)	TATCGATTACAATGATCTATTGTGCGCCTTGTTCCGGACCAAACCTCAAATGT		
		1251		1300
TRM4	(1067)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TRM6	(1067)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TRM3	(1066)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TRM1	(1068)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TRM5	(1070)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TRM14	(1069)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
MYB14TaF	(937)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TaM3	(1154)	TCTCTGATGATGAGATTCTCAAGAATTGGACACAATGTAACCTTTGCTGAT		
TaM4	(1160)	TCCCTGATGATGAGATTCTAAGAATTGGACACAATGTAACCTTTGCTGAT		
		1301		1350
TRM4	(1117)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TRM6	(1117)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TRM3	(1116)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TRM1	(1118)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TRM5	(1120)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TRM14	(1119)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
MYB14TaF	(987)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TaM3	(1204)	GAGACAAATGTGTCCAACAACCTTCAATTCCTTTTGCTTCCTTTCTCGAATC		
TaM4	(1210)	GAGACAAATGTGTCCAACAACCTTCAGTCTTCTGCTTCCTTTCTCGAATC		
		1351		1400
TRM4	(1167)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TRM6	(1167)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TRM3	(1166)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TRM1	(1168)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TRM5	(1170)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TRM14	(1169)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
MYB14TaF	(1037)	CAGTGAGGAAGTACTAGGAGAATGATAATAAAAATTCATTTTCCAATPAAA		
TaM3	(1254)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
TaM4	(1260)	CAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----		
		1401		1450
TRM4	(1204)	-----		
TRM6	(1204)	-----		
TRM3	(1203)	-----		
TRM1	(1206)	-----		
TRM5	(1207)	-----		
TRM14	(1206)	-----		
MYB14TaF	(1087)	ATTAACCTACTCTAGGTTTTTTTTTTTTTTTTTAATTTCAATTTCAATGTT		
TaM3	(1291)	-----		

ES 2 537 342 T3

```

TaM4 (1297) -----
                1451                                     1500
    TRM4 (1204) -----
    TRM6 (1204) -----
    TRM3 (1203) -----
    TRM1 (1206) -----
    TRM5 (1207) -----
    TRM14 (1206) -----
    MYB14TaF (1137) AGGGTGGTTTAATAAAATAAATATATTTCTATGGTTTAATATTGCAAAAAAA
    TaM3 (1291) -----
    TaM4 (1297) -----

                1501                                     1550
    TRM4 (1204) -----
    TRM6 (1204) -----
    TRM3 (1203) -----
    TRM1 (1206) -----
    TRM5 (1207) -----
    TRM14 (1206) -----
    MYB14TaF (1187) AAAAAAAAAAAAAAAAAAAGTACTCTGCGTTGATACCACTGCTTAAGGGC
    TaM3 (1291) -----
    TaM4 (1297) -----

                1551
    TRM4 (1204) -----
    TRM6 (1204) -----
    TRM3 (1203) -----
    TRM1 (1206) -----
    TRM5 (1207) -----
    TRM14 (1206) -----
    MYB14TaF (1237) GAATTCC
    TaM3 (1291) -----
    TaM4 (1297) -----
  
```

ES 2 537 342 T3

Figura 11

		1	50
MYB14TaF	(1)	GAATTCGCCCTTAAGCAGTGGTATCAACGCAGAGTACGCCGGGGGAAGTTA	
TaM3	(1)	-----	
TaM4	(1)	-----	
To1	(1)	-----	
To6	(1)	-----	
		51	100
MYB14TaF	(51)	TTTAATTTTATCTACATCAAACACTTCAAGAGGTTGGAAATACAAGACAGA	
TaM3	(1)	-----GAATTCGCCCTTAGGTTGGAAATACAAGACAGA	
TaM4	(1)	-----GAATTCGCCCTTAGGTTGGAAATACAAGACAGA	
To1	(1)	-----	
To6	(1)	-----	
		101	150
MYB14TaF	(101)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAGGAA	
TaM3	(33)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAGGAA	
TaM4	(33)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTGTGCAAAGGAA	
To1	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAGGAA	
To6	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTGTGCAAAGGAA	
		151	200
MYB14TaF	(150)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
TaM3	(82)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
TaM4	(82)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
To1	(43)	GGTTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
To6	(43)	GGTTTGAATAGAGGTGCTTGGACAGCTCATGAAGACAAAATCCTCACTGA	
		201	250
MYB14TaF	(200)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAACCTTCCAAAAGAG	
TaM3	(132)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAACCTTCCAAAAGAG	
TaM4	(132)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAACCTTCCAAAAGAG	
To1	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAACCTTCCAAAAGAG	
To6	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAACCTTCCAAAAGAG	
		251	300
MYB14TaF	(250)	CAG-----	
TaM3	(182)	CAGGTTTCATTTCATTCTAGTATCTTGCAATT-ATAGATCAAT---CACTTT	
TaM4	(182)	CAGGTTTCATTTCATTCT-GTATCTTACAATT-ATAGATTAAC---CACTTT	
To1	(143)	CAGGTTTCATTTCATTCTG-ATATCTTACTATTTATAGATCAATAATCACTTT	
To6	(143)	CAGGTTTCATTTCATTCT-GTATCTTACTATTTATAGATCAATAATCACTTT	
		301	350
MYB14TaF	(253)	-----	
TaM3	(228)	CATACTTTTGTTTG-CTTATAAATTTCTTGCATTTTTCTTCAATTTTC	
TaM4	(227)	CATACTTTTGTTTG-CTTATAAATTTCTTGTATTTTTCTTCCATTTTT	
To1	(192)	CAT-----GTATTT-----TTTTTCCTTCCATTTTT	
To6	(192)	CAT-----GTATTT-----TTTTTCCTTCCATTTTT	
		351	400
MYB14TaF	(253)	-----	
TaM3	(277)	CATGTGAAATGCAAATTAAGTACTAGTACATTATTATGGATATGTTTTGCAAA	
TaM4	(276)	CATGAGAAATGCAAATTAAGTACTAGTACATTATTATGGACATGTTTTGGCAAA	
To1	(218)	CATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTT-----	
To6	(218)	CATTAGAAATGCAAATTAATAGTACATTATTATGGACATGTTTT-----	
		401	450
MYB14TaF	(253)	-----ATTTAAAAGATGTGGAAAAGTTGTAGACTT	
TaM3	(327)	TATGTGTATGCCATGCAGGTTTAAAAGATGCGGAAAAGTTGTAGACTT	
TaM4	(326)	TATGTTTTATGCCATGCAGATTTAAAAGATGTGGAAAAGTTGTAGACTT	
To1	(262)	-----TCCAGGTTTAAAAGATGTGGAAAAGTTGTAGACTT	
To6	(262)	-----TCCAGGTTTAAAAGATGTGGAAAAGTTGTAGACTT	

ES 2 537 342 T3

		451	500
MYB14TaF	(285)	AGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATATCCCC	
TaM3	(377)	AGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATATCTC	
TaM4	(376)	AGATGGTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATATCCCC	
To1	(299)	AGATGGTTGAATTATCTTAGACCAGATATTAAGAGAGGTAATATATATCGTC	
To6	(299)	AGATGGTTGAATTATCTTAGACCAGATATTAAGAGAGGTAATATATATCGTC	
		501	550
MYB14TaF	(335)	GGATGAAGAAGAACTTATATCCGACTTCACAACTACTCGGAAACAG--	
TaM3	(427)	GGATGAAGAAGAACTTATCATCAGACTTCACAACTACTCGGAAACAGGT	
TaM4	(426)	GGATGAAGAAGAACTTATATCCGACTTCACAACTACTCGGAAACAGGT	
To1	(349)	CGATGAAGAAGAACTTATCATTAGACTTCACAACTACTTGGAAACCGGT	
To6	(349)	CGATGAAGAAGAACTTATCATTAGACTTCACAACTACTTGGAAACCGGT	
		551	600
MYB14TaF	(383)	-----	
TaM3	(477)	AAAAGTACCGACATAATCACTAACTTATTAACATTTATCTATAAATTTGTT	
TaM4	(476)	AAA-GTCCCTAACATAATCACTAACTTATTAACGTTTGTCTATAAATTTGTT	
To1	(399)	AAA-GTATCGACATAATCACTAACTTACTAACATT-----TGT-	
To6	(399)	AAA-GTATCGACATAATCACTAACTTACTAACATT-----TGT-	
		601	650
MYB14TaF	(383)	-----	
TaM3	(527)	TTTTTTGACAATTAGTACTACTAATTTAATTTTATAATGTGTGCTAATTT	
TaM4	(525)	TTTTTTGACCATTAGTACTACTAATTTAATTTTACAATGTGTGCTAATTT	
To1	(436)	-----TTATAATGTGTACTAATTT-	
To6	(436)	-----TTATAATGTGTACTAATTT-	
		651	700
MYB14TaF	(383)	-----ATGGTCTCTAATAGCCGGAAGACTTCC	
TaM3	(577)	GCTTTGCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCC	
TaM4	(575)	GCT-TGCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCC	
To1	(454)	GCGATTCTTTGATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCC	
To6	(454)	GCGATTCTTTGATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCC	
		701	750
MYB14TaF	(410)	AGGGCGAACAGACAATGAAATAAAGAACTACTGGAACACAAATTTAGGAA	
TaM3	(627)	AGGACGAAACAGACAATGAAATAAAGAACTACTGGAACACAAATTTAGGAA	
TaM4	(624)	AGGGCGAACAGACAATGAAATAAAGAACTACTGGAACACAAATTTAGGAA	
To1	(504)	AGGGCGAACAGACAATGAAATAAAGAATTTACTGGAACACGAAATTTAGGAA	
To6	(504)	AGGGCGAACAGACAATGAAATAAAGAATTTACTGGAACACGAAATTTAGGAA	
		751	800
MYB14TaF	(460)	AAAAGGTTAAGGATCTTAAATCAACAAAACACCAACAAATCTTCTCCTACT	
TaM3	(677)	AAAAGGTTAAGGATCTTAAATCAACAAAACACCAACAAATCTTCTCCTACT	
TaM4	(674)	AAAAGGTTAAGGATCTTGAATCAACAAAACACCAACAAATCTTCTCCTACT	
To1	(554)	AAAAGGTTAAGGATCTTAAATCAACAAAACACCAACAAATCTTCTCCTACT	
To6	(554)	AAAAGGTTAAGGATCTTAAATCAACAAAACACCAACAAATCTTCTCCTACT	
		801	850
MYB14TaF	(310)	AAACTTCTGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACA---GAT	
TaM3	(727)	AAACTCTGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACA---GAT	
TaM4	(724)	AAACTCTGTCTCAACCAAAAAATGCAAGATCAAACAGAAACA---GAT	
To1	(604)	AAACTTCTGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACAACAGAT	
To6	(604)	AAACTTCTGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACAACAGAT	
		851	900
MYB14TaF	(557)	CAAT---CCTAAGCCAATGAAGCCAAACTCAAATGTTGTCCGTACAAAAG	
TaM3	(774)	CAAT---CCTAAGCCAATGAAGCCAAACTCAAATGTTGTCCGTACAAAAG	
TaM4	(771)	CAAT---CCTAAGCCAA-----ACTCATATGTTGTCCGTACAAAAG	
To1	(654)	CAATAATCCTAAGCCAATGAAGCCAAACTCGAATGTTGTCCGTACAAAAG	
To6	(654)	CAATAATCCTAAGCCAATGAAGCCAAACTCGAATGTTGTCCGTACAAAAG	
		901	950
MYB14TaF	(604)	CTACCAAGTGTCTAAGGTATTGTTTCATAAACTCACTCCCAACTCACCA	

ES 2 537 342 T3

TaM3	(821)	CTACCAAGTGTCTAAGGTATTGTTTCATAAACTCACTCCCAACTCACCA	
TaM4	(809)	CTACCAAGTGTCTAAGGTATTGTTTCATAAACTCACCCCCCAACTCACCA	
To1	(704)	CTACCAAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA	
To6	(704)	CTACCAAAATGTTCTAAGGTATTGTTTCATAAACTCAC-----CACCA	
		951	1000
MYB14TaF	(654)	---ATGCATGATTTGCAGAACAAGCTGAGGCAGAGACAACAACAAA---	
TaM3	(871)	---ATGCATGATTTGCAGAACAAGCTGAGGCAGAGACAACAACAAA---	
TaM4	(859)	CCAATGCATGATTTGCAGAACAAGCTGAGGCAGAGACAACAACAACAA	
To1	(745)	---ATGCATAATTTGCAGAACAAGCTGAGGCAGAGACAACAACAAA---	
To6	(745)	---ATGCATAATTTGCAGAACAAGCTGAGGCAGAGACAACAACAAA---	
		1001	1050
MYB14TaF	(698)	-----GCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAA	
TaM3	(915)	-----GCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAA	
TaM4	(909)	AAAGCCATCAATGCCATCAATGCTGGTTGATGGTGTGGCTAGTGATTCAA	
To1	(789)	-----GACATCAATGTTGGTTAATGGTGTAGCTAGTGATTCAA	
To6	(789)	-----GACATCAATGTTGGTTAATGGTGTAGCTAGTGATTCAA	
		1051	1100
MYB14TaF	(736)	TGAGTAACAACGAAATGGAACACGGTTATGGATTTTTGTCAATTTGCGAT	
TaM3	(953)	TGAGTAACAACGAAATGGAACACGGTTATGGATTTTTGTCAATTTGCGAT	
TaM4	(959)	TGAGTAACAACGAAATGGAATGCGGTAATGGATTTTTGTCAATTTGCGAT	
To1	(827)	TGAGTAACAACGAAATGGAACGAGGTAATGGATTTTTGTCAATTTGCGGAT	
To6	(827)	TGAGTAACAACGAAATGGAACGGGTAATGGATTTTTGTCAATTTGCGGAT	
		1101	1150
MYB14TaF	(786)	GAAGAGAAAGAAGTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGA	
TaM3	(1003)	GAAGAGAAAGAAGTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGA	
TaM4	(1009)	GAAGAGAAAGAAGTATCCGCAGATTTGCTAGAGATTTTAACATCGCGGA	
To1	(877)	GAAGAGAAAGAAGTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGA	
To6	(877)	GAAGAGAAAGAAGTATCCGCAGATTTGCTAGATGATTTTAACATCGCGGA	
		1151	1200
MYB14TaF	(836)	TGATATTTGCTTATCTGAACCTTTTGAACCTCTGATTTCTCAAATGCGTGCA	
TaM3	(1053)	TGATATTTGCTTATCTGAACCTTTTGAACCTCTGATTTCTCAAATGCGTGCA	
TaM4	(1059)	TGATATTTGCTTATCTGAATTTCTAAACTTCTGATTTCTCAAATGCGTGCA	
To1	(927)	TGACATTTGCTTATCCGAATTTCTAAACTCCGATTTCTCAAATGCGTGCA	
To6	(927)	TGACATTTGCTTATCCGAATTTCTAAACTCCGATTTCTCAAATGCGTGCA	
		1201	1250
MYB14TaF	(886)	ATTTTCGATTACAATGATCTATTGTCACCTTGTTCGGACCAAACTCAAATG	
TaM3	(1103)	ATTTTCGATTACAATGATCTATTGTCACCTTGTTCGGACCAAACTCAAATG	
TaM4	(1109)	ATTTTCGATTACAATGATCTATTGTCACCTTGTTCGGACCAAACTCAAATG	
To1	(977)	ATTTTCGATTACAATGATCTATTGTCACCTTGTTCGGATCAAACCTCAAATG	
To6	(977)	ATTTTCGATTACAATGATCTATTGTCACCTTGTTCGGATCAAACCTCAAATG	
		1251	1300
MYB14TaF	(936)	TTCTCTGATGATGAGATTTCTCAAGAATTGGACACAATGTAACCTTTGCTGA	
TaM3	(1153)	TTCTCTGATGATGAGATTTCTCAAGAATTGGACACAATGTAACCTTTGCTGA	
TaM4	(1159)	TTCCCTGATGATGAGATTTCTAAAGAATTGGACACAATGTAACCTTTGCTGA	
To1	(1027)	TTCTCTGATGATGAGATTTCTCAAGAATTGGACACAATGTAACCTTTGCTGA	
To6	(1027)	TTCTCTGATGATGAGATTTCTCAAGAATTGGACACAATGTAACCTTTGCTGA	
		1301	1350
MYB14TaF	(986)	TGAGACAAATGTGTCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAAT	
TaM3	(1203)	TGAGACAAATGTGTCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAAT	
TaM4	(1209)	TGAGACAAATGTGTCCAACAACCTTCAGTCTTCTGCTTCCTTTCTTGAAT	
To1	(1077)	TGAGACAAATGTGTCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAAT	
To6	(1077)	TGAGACAAATGTGTCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAAT	
		1351	1400

ES 2 537 342 T3

```

MYB14TaF (1036) CCAGTGAGGAAGTACTAGGAGAATGATAATAAAAATTCATTTTCCAATAA
TaM3 (1253) CCAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----
TaM4 (1259) CCAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----
To1 (1127) CCAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----
To6 (1127) CCAGTGAGGAAGTACTAGGAGAATGAAAGGGCGAATTC-----

1401 1450
MYB14TaF (1086) AATTAACTACTCTAGGTTTTTTTTTTTTTTTTTTTAAATTCATTCATGT
TaM3 (1291) -----
TaM4 (1297) -----
To1 (1165) -----
To6 (1165) -----

1451 1500
MYB14TaF (1136) TAGGGTGGTTTAAATAAATAAAATATATTCTATGGTTTAAATATGCAAAAAA
TaM3 (1291) -----
TaM4 (1297) -----
To1 (1165) -----
To6 (1165) -----

1501 1550
MYB14TaF (1186) AAAAAAAAAAAAAAAAAAAGTACTCTGCGTTGATACCACTGCCTTAAGGG
TaM3 (1291) -----
TaM4 (1297) -----
To1 (1165) -----
To6 (1165) -----

1551
MYB14TaF (1236) CGAATTC
TaM3 (1291) -----
TaM4 (1297) -----
To1 (1165) -----
To6 (1165) -----

```

Figura 12

		1	50
MYB14TaFF	(1)	GAATTCGCCCTTAAGCAGTGGTATCAACGCAGAGTACGCGGGGAAGTTA	
TaM3	(1)	-----	
Taf11	(1)	-----	
Taf2 r#2	(1)	-----	
Taf3	(1)	-----	
Taf7	(1)	-----	
Taf4	(1)	-----	
Taf10	(1)	-----	
		51	100
MYB14TaFF	(51)	TTTAATTTTATCTACATCAAACACTTCAAGAGGTTGGAATACAAGACAGA	
TaM3	(1)	-----GAATTCGCCCTTAGGTTGGAATACAAGACAGA	
Taf11	(1)	-----	
Taf2 r#2	(1)	-----	
Taf3	(1)	-----	
Taf7	(1)	-----	
Taf4	(1)	-----	
Taf10	(1)	-----	
		101	150
MYB14TaFF	(101)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
TaM3	(33)	CTAATTAAGAATAACATCA-ATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf11	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf2 r#2	(1)	-----GGAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf3	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf7	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf4	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
Taf10	(1)	-----GAATTCGCCCTTATGGGGAGAAGCCCTTGTTGTGCAAAGGAA	
		151	200
MYB14TaFF	(150)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
TaM3	(82)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf11	(43)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf2 r#2	(44)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf3	(43)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf7	(43)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf4	(43)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
Taf10	(43)	GGCTTGAATAGAGGTGCTTGGACAACCTCAAGAAGACAAAATCCTCACTGA	
		201	250
MYB14TaFF	(200)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
TaM3	(132)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf11	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf2 r#2	(94)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf3	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf7	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf4	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
Taf10	(93)	ATACATTAAGCTCCATGGTGAAGGAAAATGGAGAAACCTTCCAAAAGAG	
		251	300
MYB14TaFF	(250)	CAG-----	
TaM3	(182)	CAGGTTCAATTCATTCTAGTATCTTGCAATTATAGATCAATCACTTTTCATA	
Taf11	(143)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	
Taf2 r#2	(144)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	
Taf3	(143)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	
Taf7	(143)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	
Taf4	(143)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	
Taf10	(143)	CAGGTTCAATTCATTCT-GTATCTTACAATTATAGATTAACCACTTTTCATA	

ES 2 537 342 T3

		301	350
MYB14TaFF	(253)	-----	
TaM3	(232)	CTTTTGTTTGCTTATAAAATTTCTTGTGATTTTTCTTCAATTTTCCATGT	
Taf11	(192)	CTTTTGTTTGCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
Taf2 r#2	(193)	CTTTTGTTTCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
Taf3	(192)	CTTTTGTTTCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
Taf7	(192)	CTTTTGTTTCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
Taf4	(192)	CTTTTGTTTCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
Taf10	(192)	CTTTTGTTTCTTATAAAATTTCTTGTATTTTTCTTCCATTTTTCATGA	
		351	400
MYB14TaFF	(253)	-----	
TaM3	(282)	GAAATGCAAATTACTAGTACATTTATTAAGGATATGTTTTTGCAAATATGT	
Taf11	(242)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
Taf2 r#2	(243)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
Taf3	(242)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
Taf7	(242)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
Taf4	(242)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
Taf10	(242)	GAAATGCAAATTACTAGTACATTTATTAAGGACATGTTTTTGCAAATATGT	
		401	450
MYB14TaFF	(253)	-----GTTTAAAAAGATGCGGAAAAAGTTGTAGACTTAGATG	
TaM3	(332)	GTATGCCATGCAGGTTTAAAAAGATGCGGAAAAAGTTGTAGACTTAGATG	
Taf11	(292)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
Taf2 r#2	(293)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
Taf3	(292)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
Taf7	(292)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
Taf4	(292)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
Taf10	(292)	TTATGCCATGCAGGTTTAAAAAGATGTGGAAAAAGTTGTAGACTTAGATG	
		451	500
MYB14TaFF	(290)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
TaM3	(382)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf11	(342)	GTTGAATTATCTAAGACTAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf2 r#2	(343)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf3	(342)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf7	(342)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf4	(342)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
Taf10	(342)	GTTGAATTATCTAAGACCAGATATTAAGCGAGGTAATATATCCTCGGATG	
		501	550
MYB14TaFF	(340)	AAGAAGAACTTATCATCAGACTTCACAAACTACTCGGAAACAG-----	
TaM3	(432)	AAGAAGAACTTATCATCAGACTTCACAAACTACTCGGAAACAGGTAAAAG	
Taf11	(392)	AAGAAGAACTTATCATCCGACTTCACAAATTAATCTCGGAAACAGGTAAA-G	
Taf2 r#2	(393)	AAGAAGAACTTATCATCCGACTTCACAAACTACTCGGAAACAGGTAAA-G	
Taf3	(392)	AAGAAGAACTTATCATCCGACTTCACAAACTACTCGGAAACAGGTAAA-G	
Taf7	(392)	AAGAAGAACTTATCATCCGACTTCACAAACTACTCGGAAACAGGTAAA-G	
Taf4	(392)	AAGAAGAACTTATCATCCGACTTCACAAACTACTCGGAAACAGGTAAA-G	
Taf10	(392)	AAGAAGAACTTATCATCCGACTTCACAAACTACTCGGAAACAGGTAAA-G	
		551	600
MYB14TaFF	(383)	-----	
TaM3	(482)	TACCGACATAATCACTAACTTATTAACATTTATCTATAAATTTGTTTTTTT	
Taf11	(441)	TCCTAACATAATCACTAACTTATTAACGTTTGTCTATAAATTTGTTTTTTT	
Taf2 r#2	(442)	TCATAACATAATCATTAATTTATTAACGGTTATCTATAAATTTGTTTTTTT	
Taf3	(441)	TCATAACATGATCATTAATTTATTAACGGTTATCTATAAATTTGTTTTTTT	
Taf7	(441)	TCATAACATAATCATTAATTTATTAACGGTTATCTATAAATTTGTTTTTTT	
Taf4	(441)	TCATAACATAATCATTAATTTATTAACGGTTATCTATAAATTTGTTTTTTT	
Taf10	(441)	TCATAACATAATCATTAATTTATTAACGGTTATCTATAAATTTGTTTTTTT	
		601	650
MYB14TaFF	(383)	-----	
TaM3	(532)	TGACAATTAGTACTACTAATTTAATTTTATAATGTGTGCTAATTTGCTTT	
Taf11	(491)	-GACAATTAGTACTACTAATTTAATTTTATAATGTGTGCTAATTTGCTTT-	

ES 2 537 342 T3

Taf2 r#2 (492) -GACAATTATTACTACAAAATTTAATTTTATAATGTGTGCTAAATTTGCTT-
Taf3 (491) -GACAATTATCACTACAAAATTTAATTTTATAATGTGCGCTAATTTGCTT-
Taf7 (491) -GACAATTATCACTACAAAATTTAATTTTATAATGTGCGCTAATTTGCTT-
Taf4 (491) -GACAATTATCACTACAAAATTTAATTTTATAATGTGCGCTAATTTGCTT-
Taf10 (491) -GACAATTATCACTACAAAATTTAATTTTATAATGTGCGCTAATTTGCTT-

651 700

MYB14TaFF (383) -----ATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Tam3 (582) GCTTTTAAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf11 (539) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf2 r#2 (540) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf3 (539) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf7 (539) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf4 (539) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC
Taf10 (539) GTCTTTAATTTGTGGTAGATGGTCTCTAATAGCCGGAAGACTTCCAGGAC

701 750

MYB14TaFF (415) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Tam3 (632) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf11 (589) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf2 r#2 (590) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf3 (589) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf7 (589) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf4 (589) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG
Taf10 (589) GAACAGACAATGAAATAAAGAAGACTACTGGAACACAAATTTAGGAAAAAAG

751 800

MYB14TaFF (465) GTTAAGGATCTTAATCAACAAAACACCAACAATTCCTCTCTACTAAACT
Tam3 (682) GTTAAGGATCTTAATCAACAAAACACCAACAAGTCTTCTCTACTAAACT
Taf11 (639) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT
Taf2 r#2 (640) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT
Taf3 (639) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT
Taf7 (639) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT
Taf4 (639) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT
Taf10 (639) GTTAAGGATCTTAATCAAGAAAACACCAACAATTCCTCTCTACTAAACT

801 850

MYB14TaFF (515) CTCGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Tam3 (732) CTCGTCTCAACCAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf11 (689) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf2 r#2 (690) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf3 (689) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf7 (689) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf4 (689) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA
Taf10 (689) TTCTGCTCAACTAAAAAATGCAAAGATCAAACAGAAACAGATCAATCCTA

851 900

MYB14TaFF (565) AGCCAATGAAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Tam3 (782) AGCCAATGAAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf11 (739) AGCCAATGGAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf2 r#2 (740) AGCCAATGAAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf3 (739) AGCCAATGGAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf7 (739) AGCCAATGGAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf4 (739) AGCCAATGGAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT
Taf10 (739) AGCCAATGGAGCCAAACTCAAATGTTGTCCGTACAAAAGCTACCAAGTGT

901 950

MYB14TaFF (615) TCTAAGGTATTGTTTCATAAACTCACTCCCCAACTACCA---ATGCATGA
Tam3 (832) TCTAAGGTATTGTTTCATAAACTCACTCCCCAACTACCA---ATGCATGA
Taf11 (789) TCTAAGGCATTGTTTCATAAACTCACCCCAACTCACCACCAATGCATGA
Taf2 r#2 (790) TCTAAGGTATTGTTTCATAAACTCACCCCAACTCACCACCAATGCATGA
Taf3 (789) TCTAAGGCATTGTTTCATAAACTCACCCCAACTCACCACCAATGCATGA
Taf7 (789) TCTAAGGCATTGTTTCATAAACTCACCCCAACTCACCACCAATGCATGA
Taf4 (789) TCTAAGGCATTGTTTCATAAACTCACCCCAACTCACCACCAATGCATGA

ES 2 537 342 T3

Taf10 (789) TCTAAGGCATTGTTTCATAAACTCACCCCCCAACTCACCACCAATGCATGA

951 1000

MYB14TaFF (662) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGCCATCAATGC---

TaM3 (879) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGCCATCAATGC---

Taf11 (839) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGTCATCAATGCCAT

Taf2 r#2 (840) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGCCATCAATGCCAT

Taf3 (839) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGTCATCAATGCCAT

Taf7 (839) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGTCATCAATGCCAT

Taf4 (839) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGTCATCAATGCCAT

Taf10 (839) TTTGCAGAACAAGCTGAGGCAGAGACAACAACAAGTCATCAATGCCAT

1001 1050

MYB14TaFF (709) -----TGGTTGATGGTGTGGCTAGTGATTCAATGAGTAACACGAAATG

TaM3 (926) -----TGGTTGATGGTGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf11 (889) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf2 r#2 (890) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf3 (889) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf7 (889) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf4 (889) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

Taf10 (889) CAATGCTGGTTGATGGCGTGGCTAGTGATTCAATGAGTAACACGAAATG

1051 1100

MYB14TaFF (753) GAACACGGTTATGGATTTTGTTCATTTTCCGATGAAGAGAAAGAACTATC

TaM3 (970) GAACACGGTTATGGATTTTGTTCATTTTCCGATGAAGAGAAAGAACTATC

Taf11 (939) GAATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

Taf2 r#2 (940) GGATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

Taf3 (939) GAATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

Taf7 (939) GAATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

Taf4 (939) GAATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

Taf10 (939) GAATACGGTGATGGATTTGTTTCATTTTCCGATGACGATAAAGAACTATC

1101 1150

MYB14TaFF (803) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

TaM3 (1020) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf11 (989) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf2 r#2 (990) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf3 (989) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf7 (989) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf4 (989) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

Taf10 (989) CGCAGATTTGCTAGAAGATTTTAAACATCCGGATGATATTTGCTTATCTG

1151 1200

MYB14TaFF (853) AACTTTTGAACCTCTGATTTCTCAAATGCGTGCAATTTTCGATTACAATGAT

TaM3 (1070) AACTTTTGAACCTCTGATTTCTCAAATGCGTGCAATTTTCGATTACAATGAT

Taf11 (1039) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

Taf2 r#2 (1040) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

Taf3 (1039) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

Taf7 (1039) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

Taf4 (1039) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

Taf10 (1039) AATTTCTAAACTTCGATTTCTCAAATGCGTGCAATTTTCGATTACAACGAT

1201 1250

MYB14TaFF (903) CTATTGTCACCTTGTTCGGACCAAACCTCAAATGTTCTCTGATGATGAGAT

TaM3 (1120) CTATTGTCACCTTGTTCGGACCAAACCTCAAATGTTCTCTGATGATGAGAT

Taf11 (1089) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

Taf2 r#2 (1090) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

Taf3 (1089) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

Taf7 (1089) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

Taf4 (1089) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

Taf10 (1089) CTATTGTCACCTTGTTCGGACCAAACCAAAATGTTCTCTGATGATGAGAT

ES 2 537 342 T3

		1251	1300
MYB14TaFF	(953)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGATGAGACAAAT----	GTG
TaM3	(1170)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGATGAGACAAAT----	GTG
Taf11	(1139)	TCTCAAGAATTGGACACCATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
Taf2 r#2	(1140)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
Taf3	(1139)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
Taf7	(1139)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
Taf4	(1139)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
Taf10	(1139)	TCTCAAGAATTGGACACAATGTAACCTTTGCTGCTGAGACAAATTAATGTG	
		1301	1350
MYB14TaFF	(999)	TCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAATCCAGTGAGGAAGT	
TaM3	(1216)	TCCAACAACCTTCATTCTTTTGCTTCCTTTCTTGAATCCAGTGAGGAAGT	
Taf11	(1189)	TCCAACAACC-----AATCCAGTGAGGAAGT	
Taf2 r#2	(1190)	TCCAACAACC-----AATCCAGTGAGGAAGT	
Taf3	(1189)	TCCAACAACC-----AATCCAGTGAGGAAGT	
Taf7	(1189)	TCCAACAACC-----AATCCAGTGAGGAAGT	
Taf4	(1189)	TCCAACAACC-----AATCCAGTGAGGAAGT	
Taf10	(1189)	TCCAACAACC-----AATCCAGTGAGGAAGT	
		1351	1400
MYB14TaFF	(1049)	ACTAGGAGAATGATAATAAAAATTCATTTTCCAATAAAAATTAECTACTCT	
TaM3	(1266)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf11	(1215)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf2 r#2	(1216)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf3	(1215)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf7	(1215)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf4	(1215)	ACTAGGAGAATGAAAGGGCGAATTC-----	
Taf10	(1215)	ACTAGGAGAATGAAAGGGCGAATTC-----	
		1401	1450
MYB14TaFF	(1099)	AGGTTTTTTTTTTTTTTTTTTAATTCAATTCATGTTAGGGTGGTTTAA	
TaM3	(1291)	-----	
Taf11	(1241)	-----	
Taf2 r#2	(1241)	-----	
Taf3	(1240)	-----	
Taf7	(1240)	-----	
Taf4	(1240)	-----	
Taf10	(1240)	-----	
		1451	1500
MYB14TaFF	(1149)	TAAATAAATATATCTATGGTTTAAATATTGCAAAAAAAAAAAAAAAAAAAAA	
TaM3	(1291)	-----	
Taf11	(1241)	-----	
Taf2 r#2	(1241)	-----	
Taf3	(1240)	-----	
Taf7	(1240)	-----	
Taf4	(1240)	-----	
Taf10	(1240)	-----	
		1501	1545
MYB14TaFF	(1199)	AAAAAAGTACTCTGCGTTGATACCACTGCTTAAGGGCGAATTC	
TaM3	(1291)	-----	
Taf11	(1241)	-----	
Taf2 r#2	(1241)	-----	
Taf3	(1240)	-----	
Taf7	(1240)	-----	
Taf4	(1240)	-----	
Taf10	(1240)	-----	

Figura 13

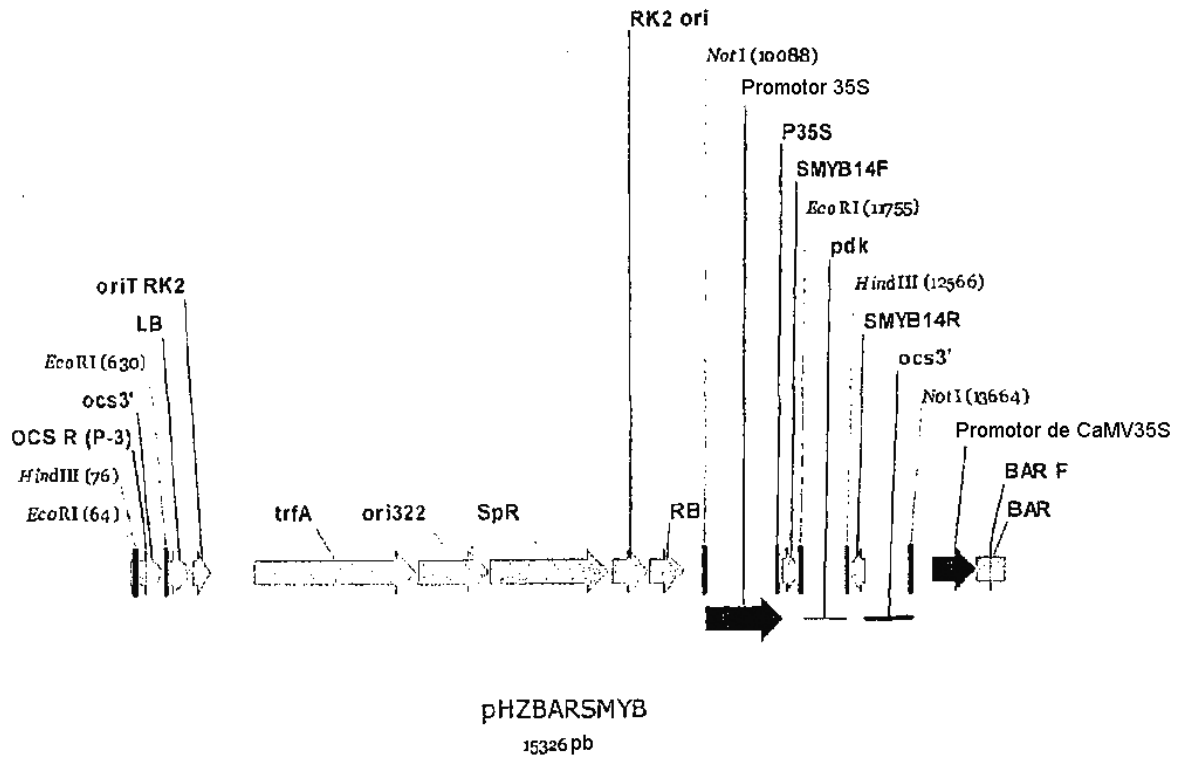


Figura 14

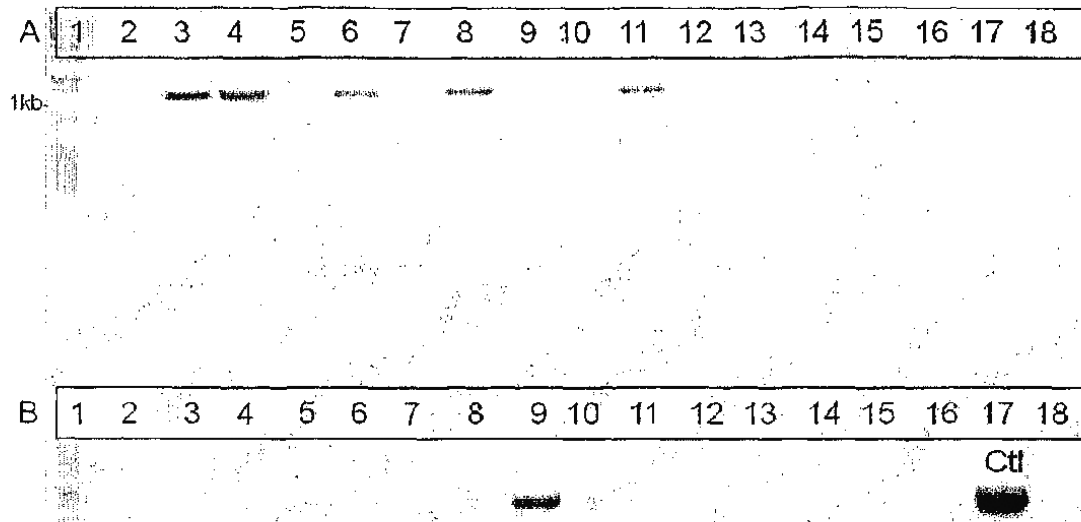
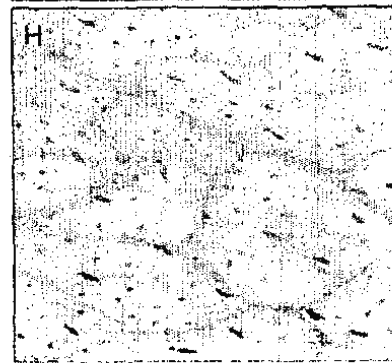
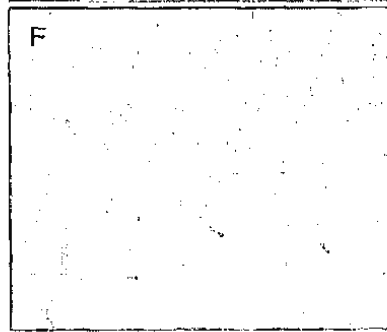
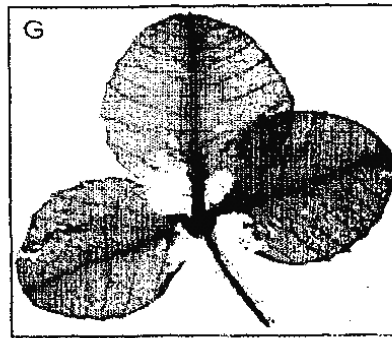
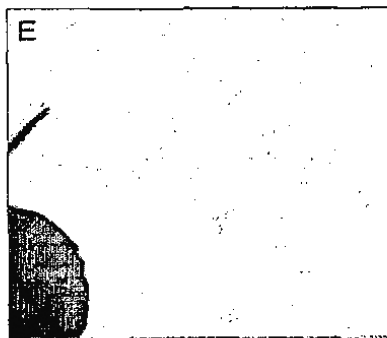
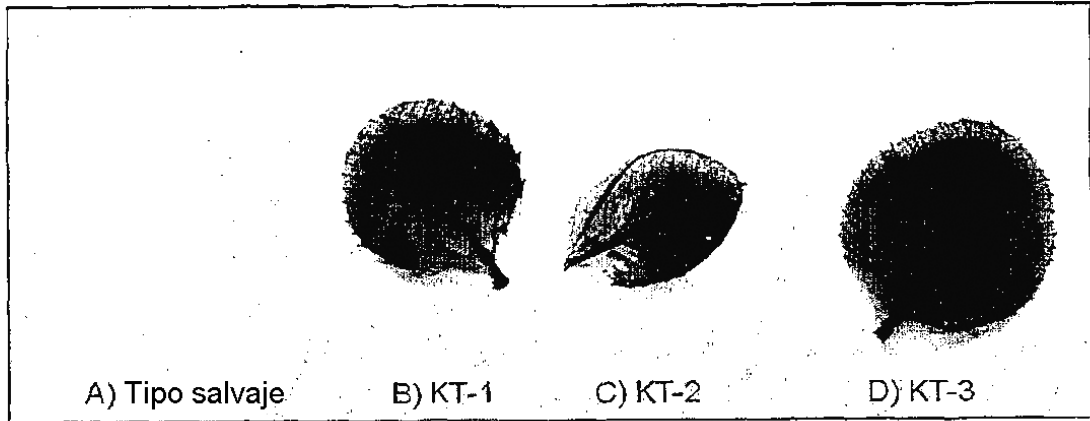


Figura 15



Tipo salvaje

D48 transgénico

Figura 16

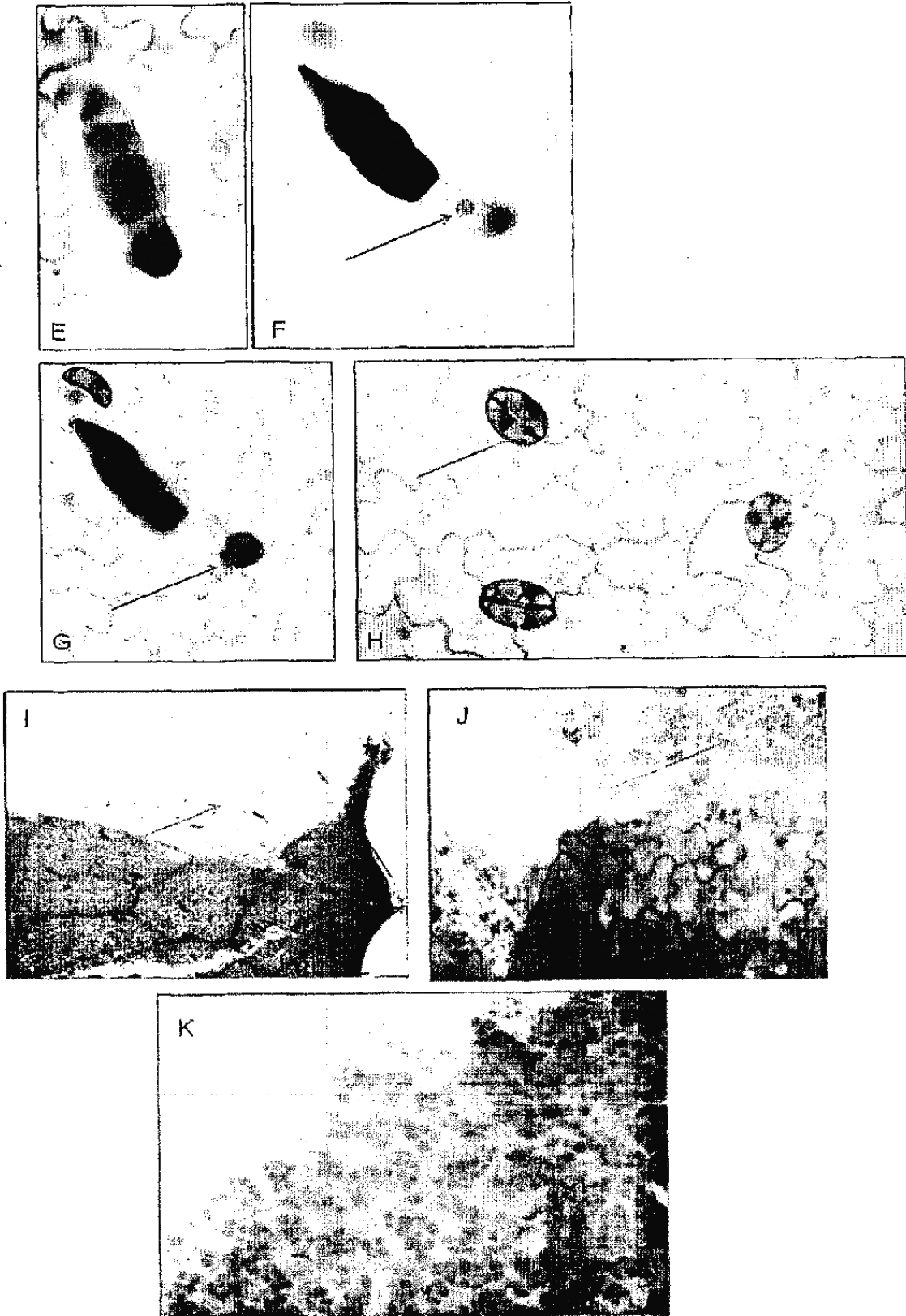


Figura 17

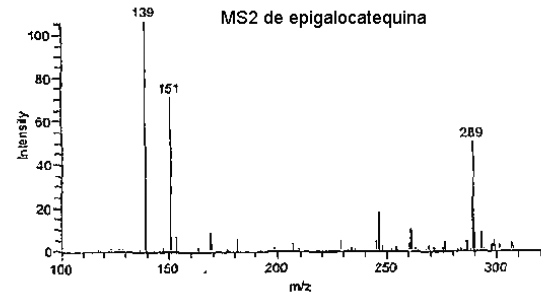
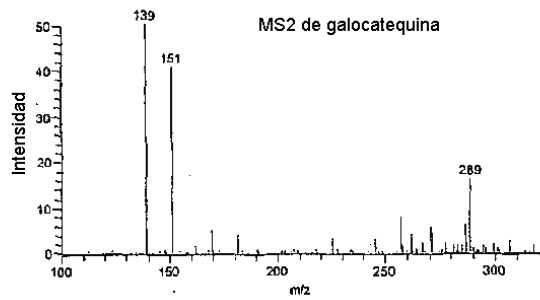
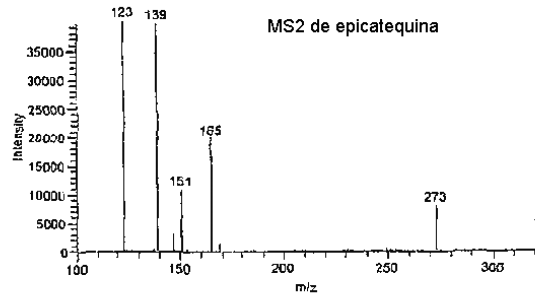
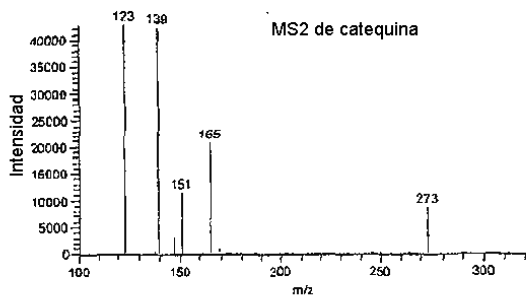
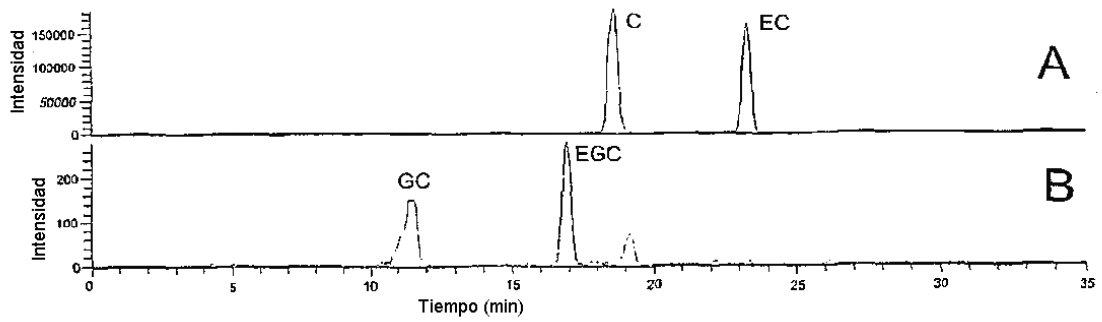


Figura 18

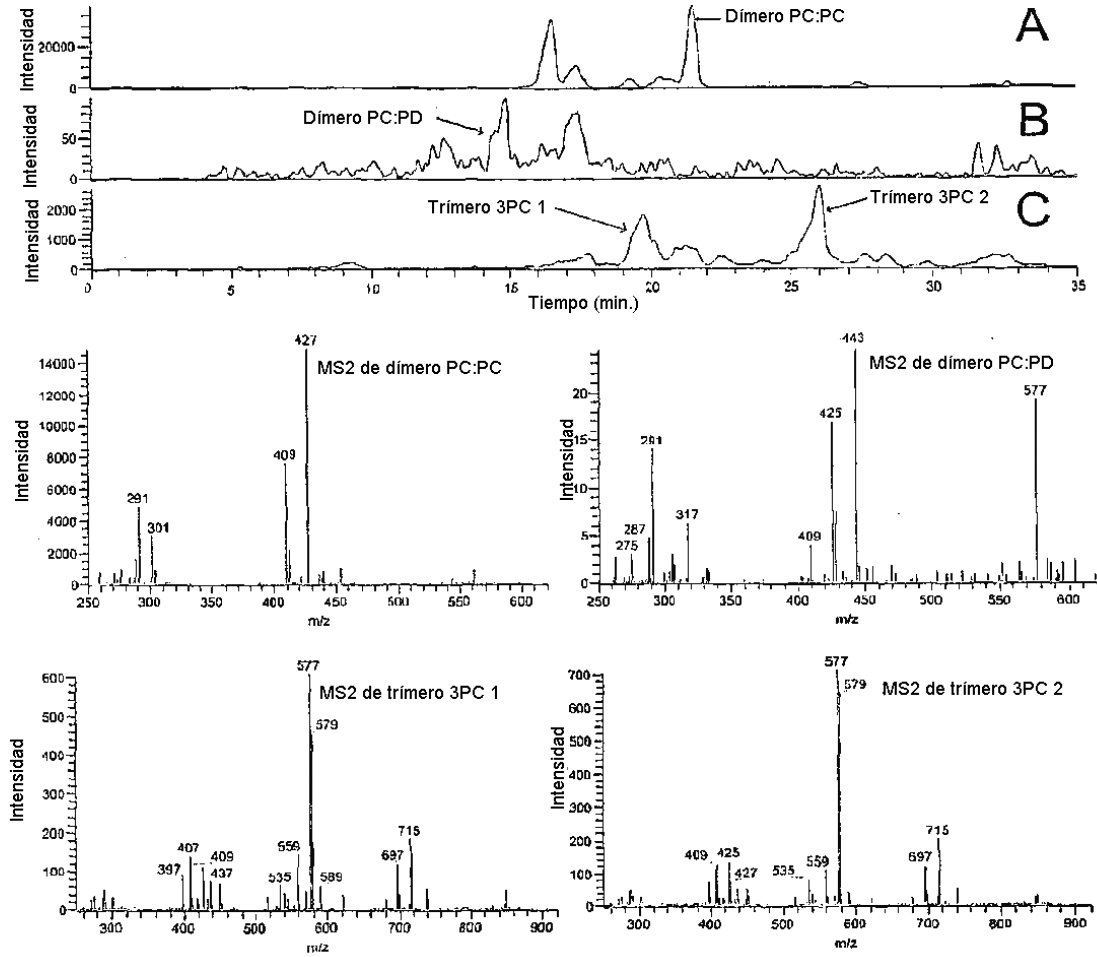


Figura 19

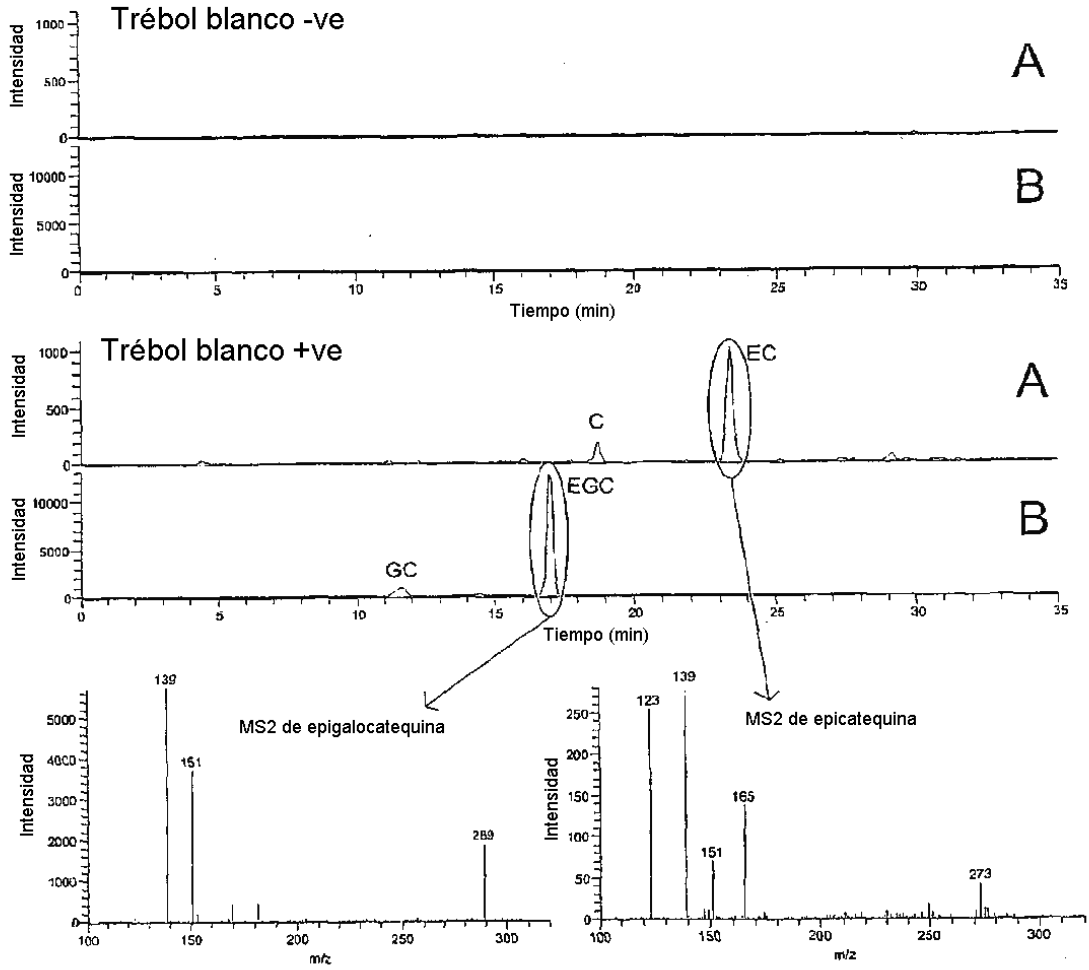


Figura 20

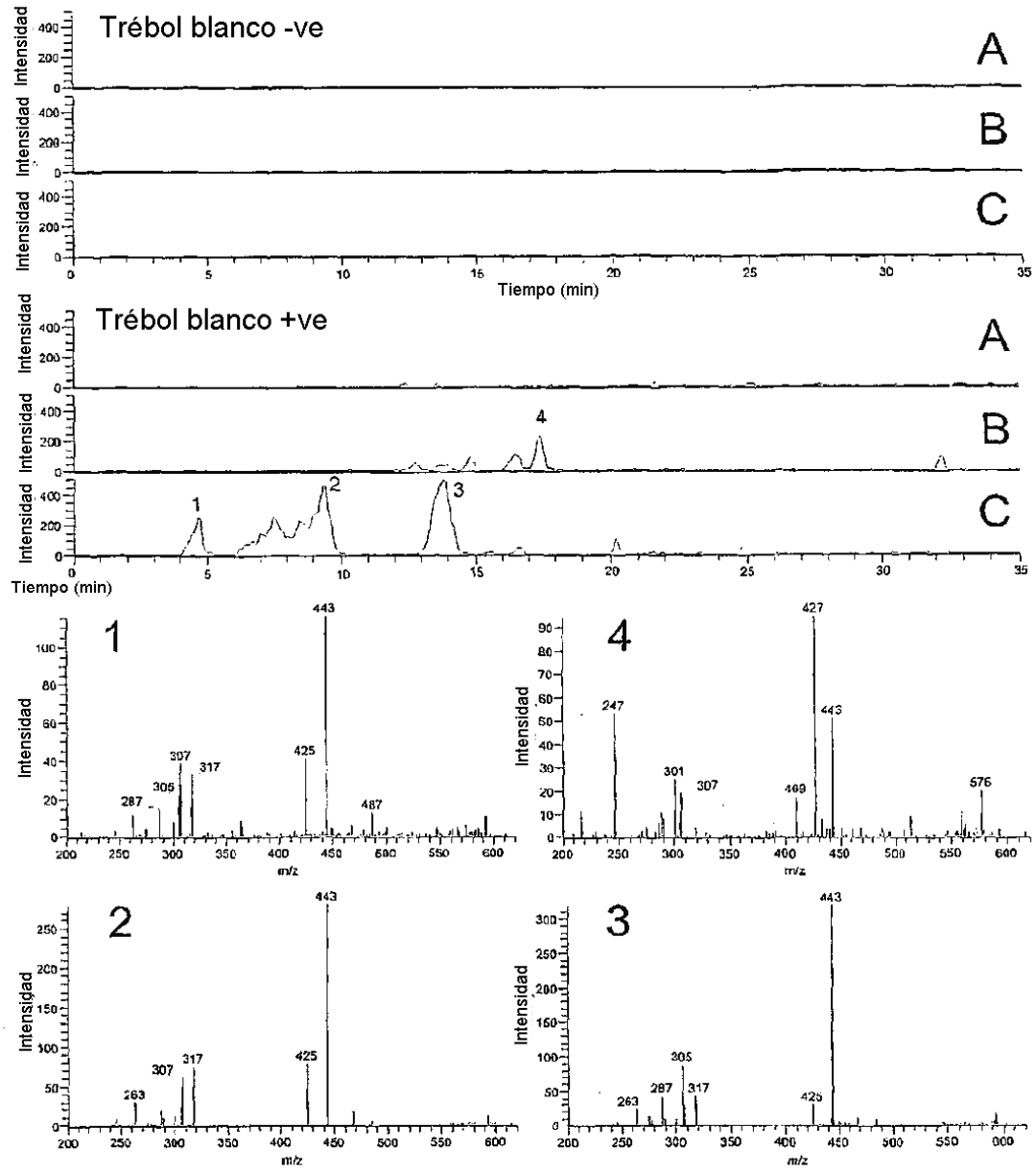


Figura 21

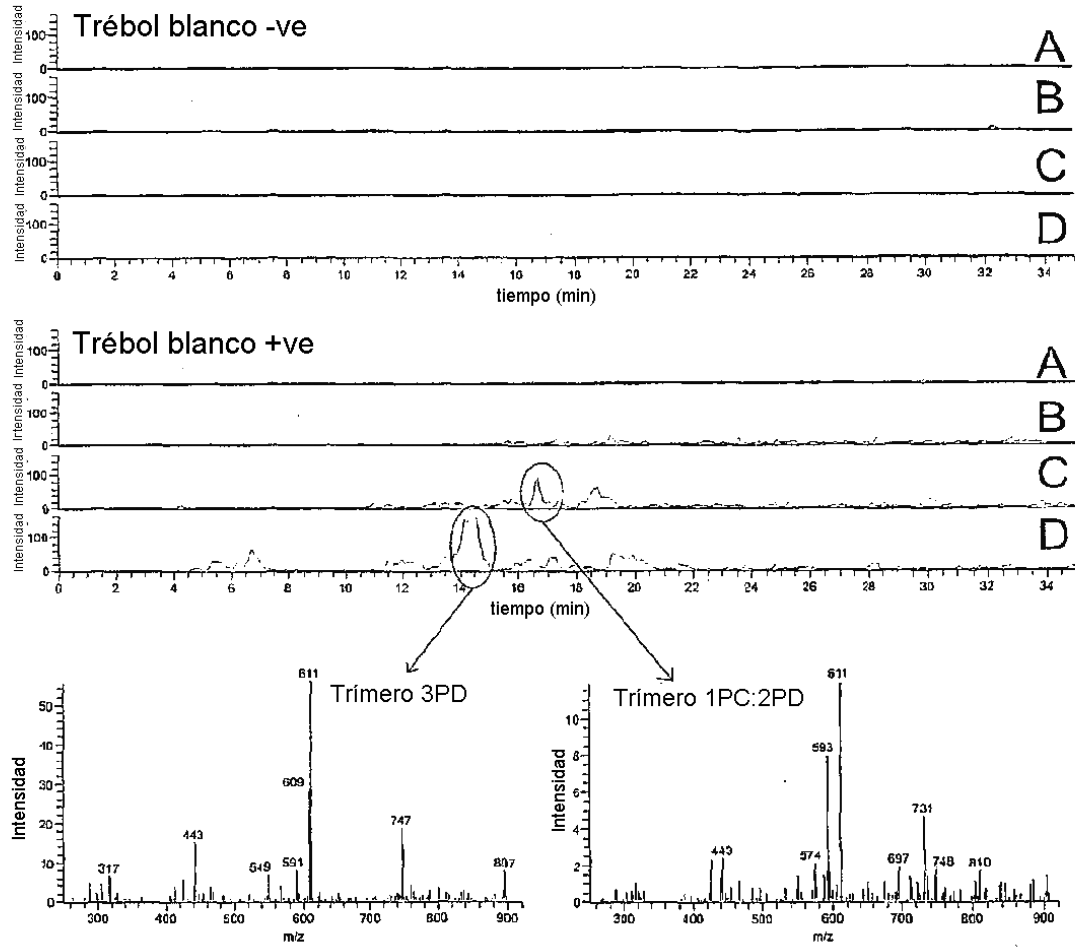


Figura 22

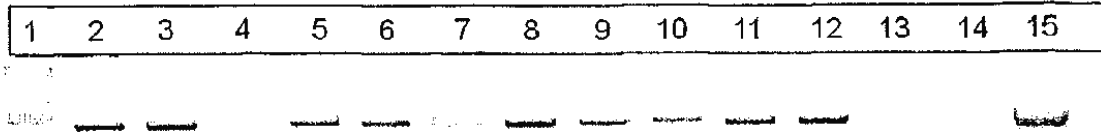


Figura 23



Figura 24

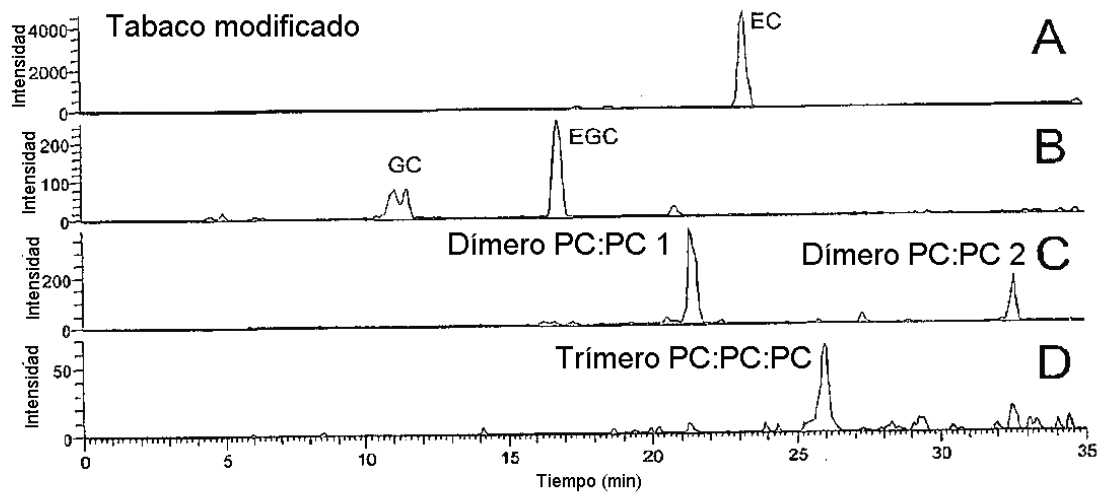
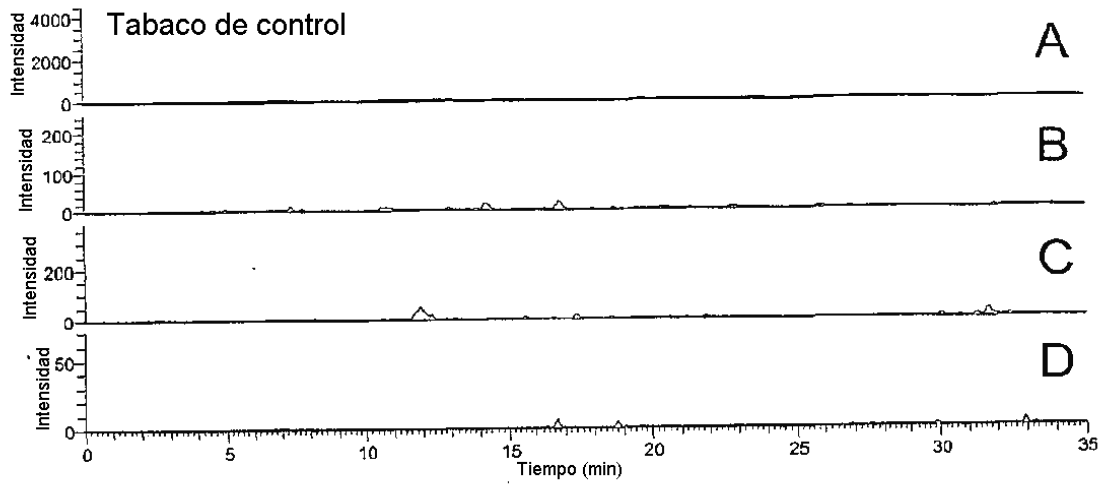


Figura 25

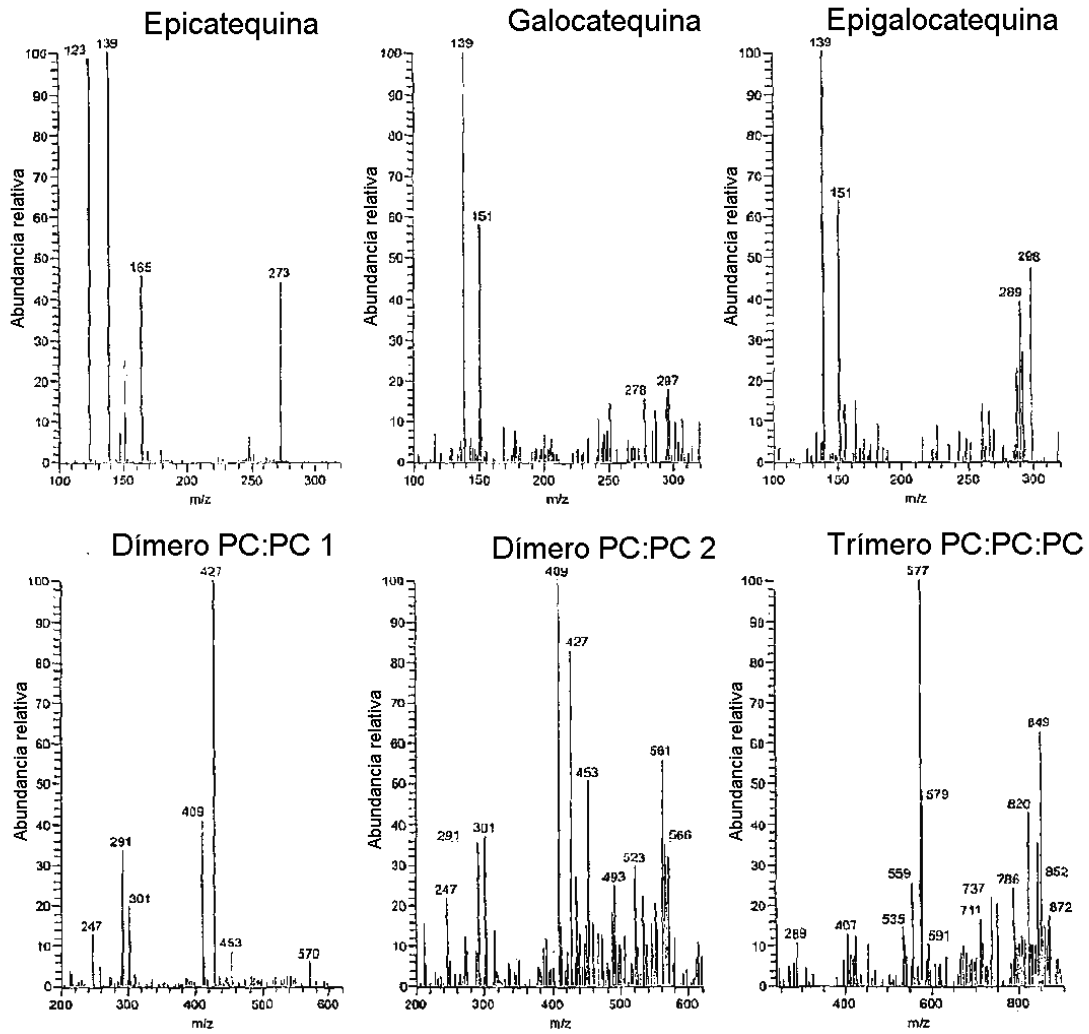


Figura 26

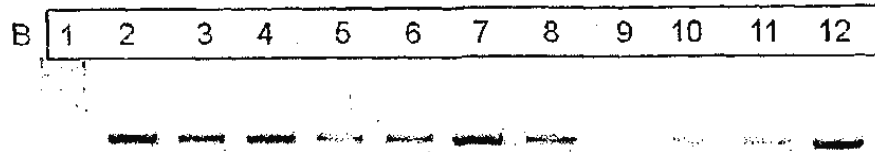
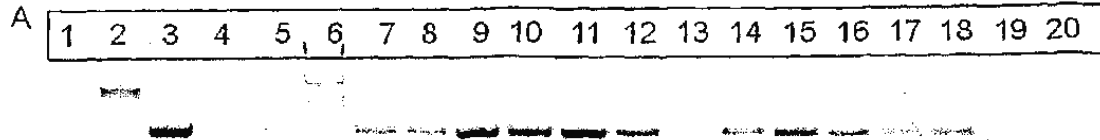


Figura 27

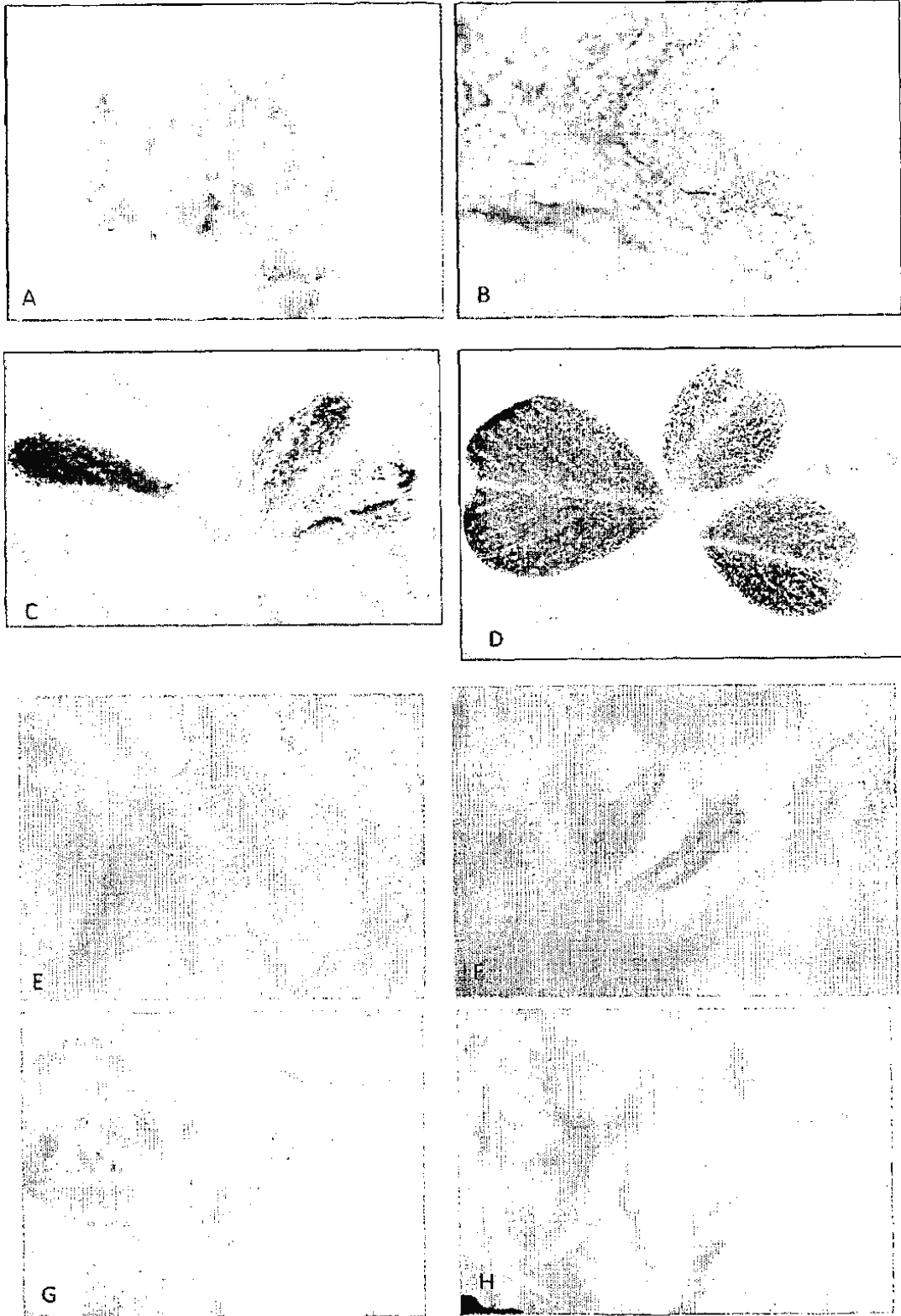


Figura 27a

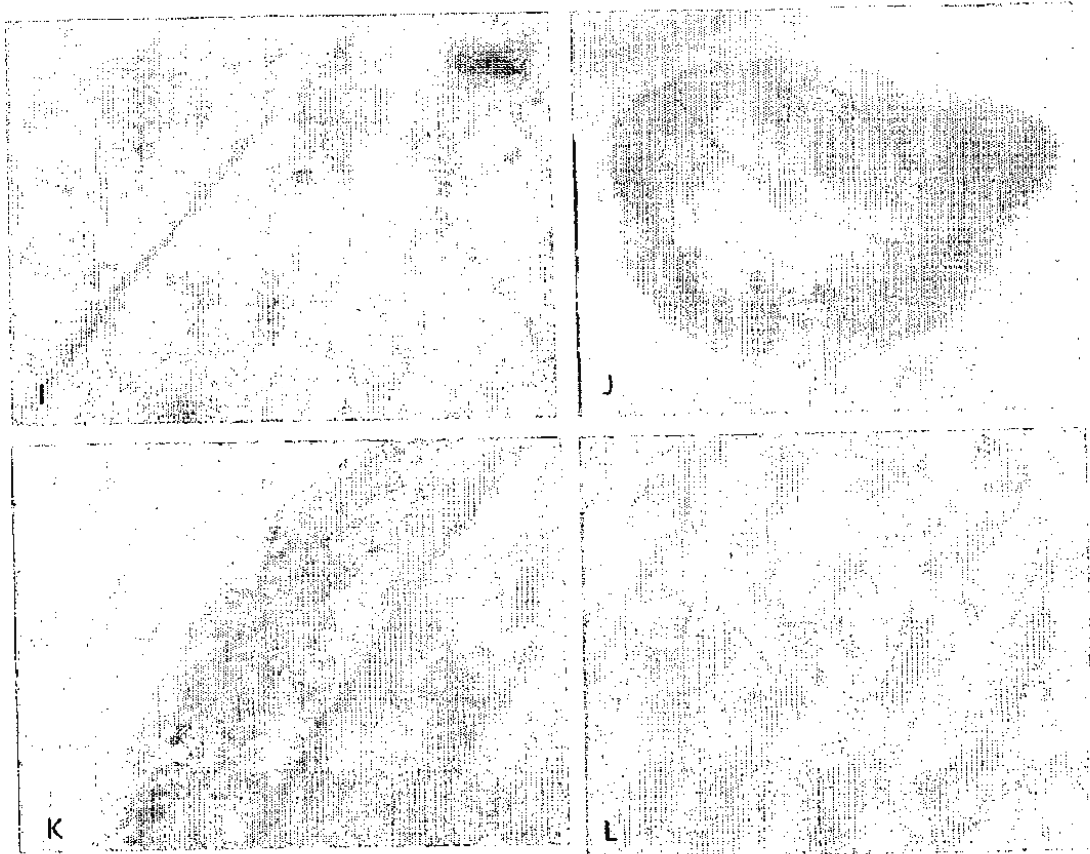


Figura 28

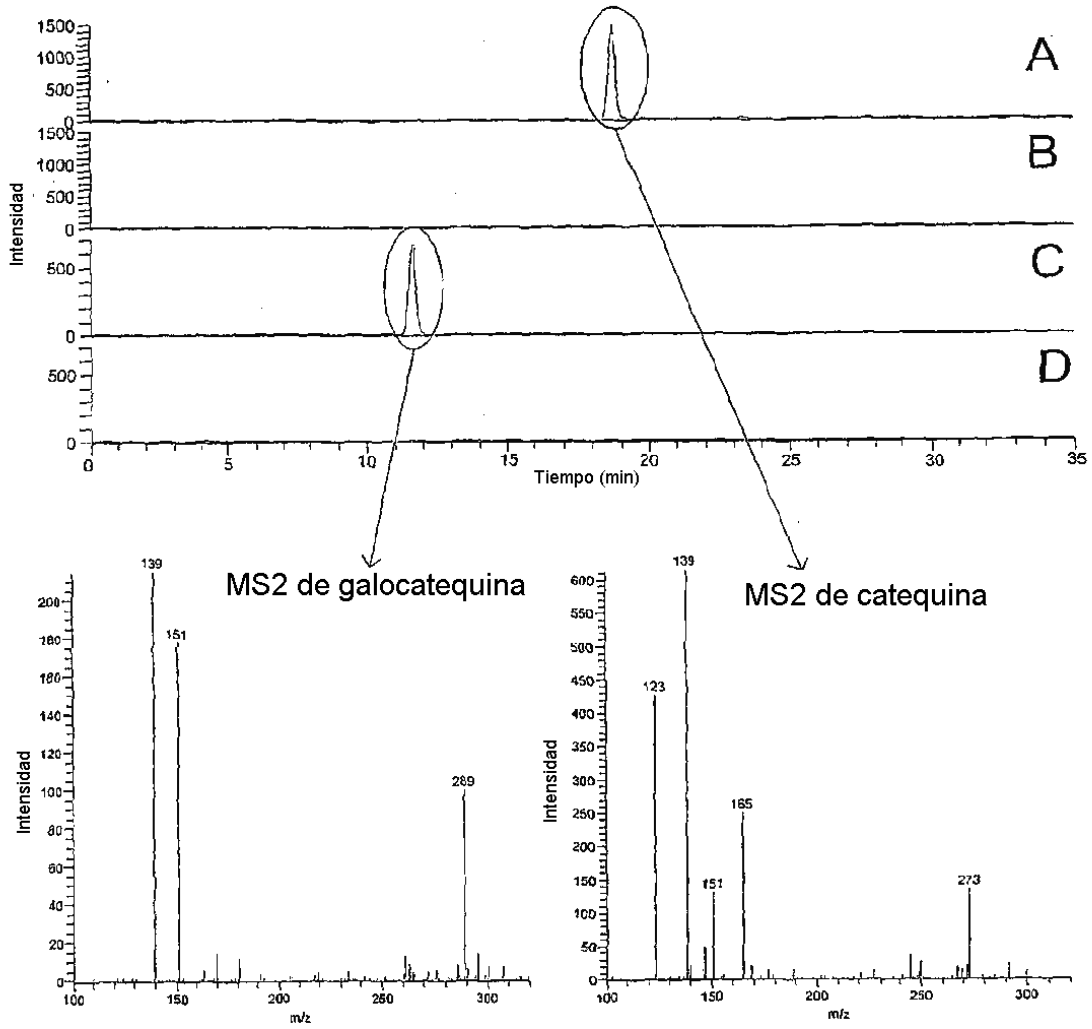


Figura 29

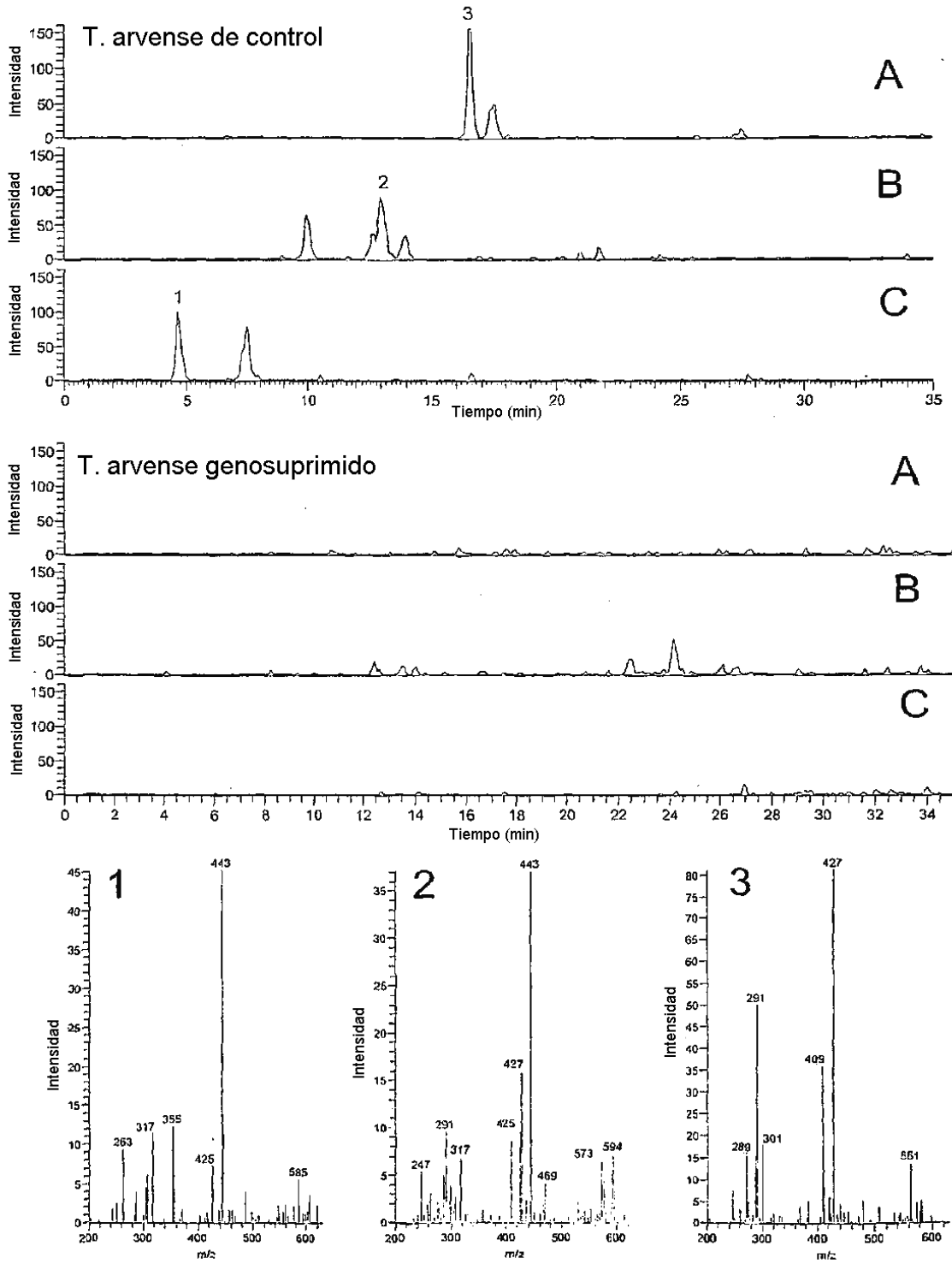


Figura 30

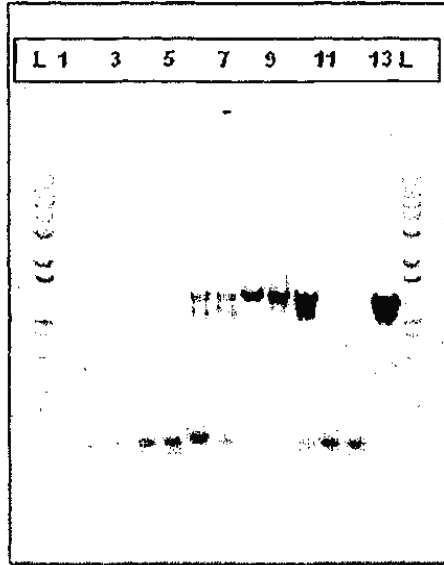


Figura 31

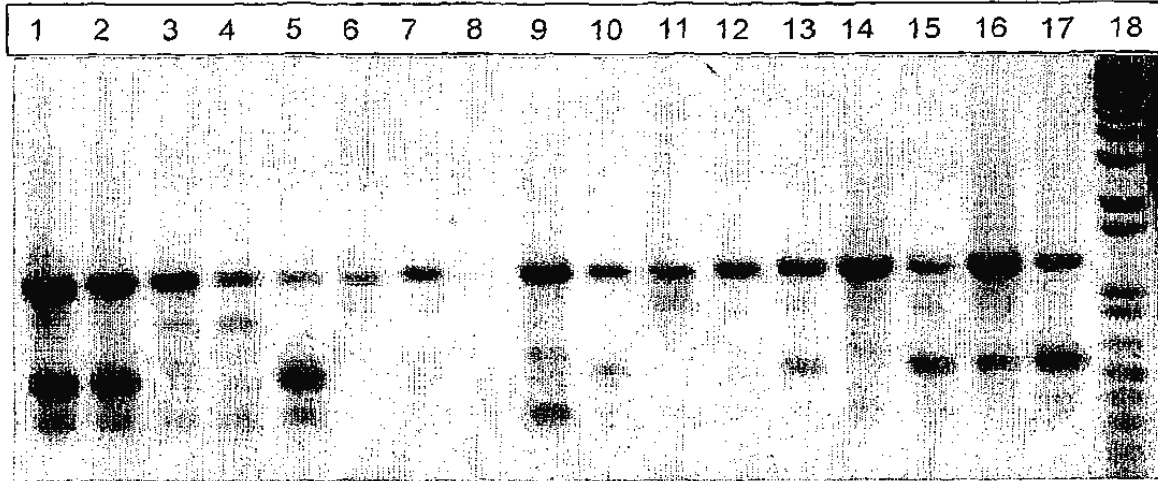


Figura 32



Figura 33

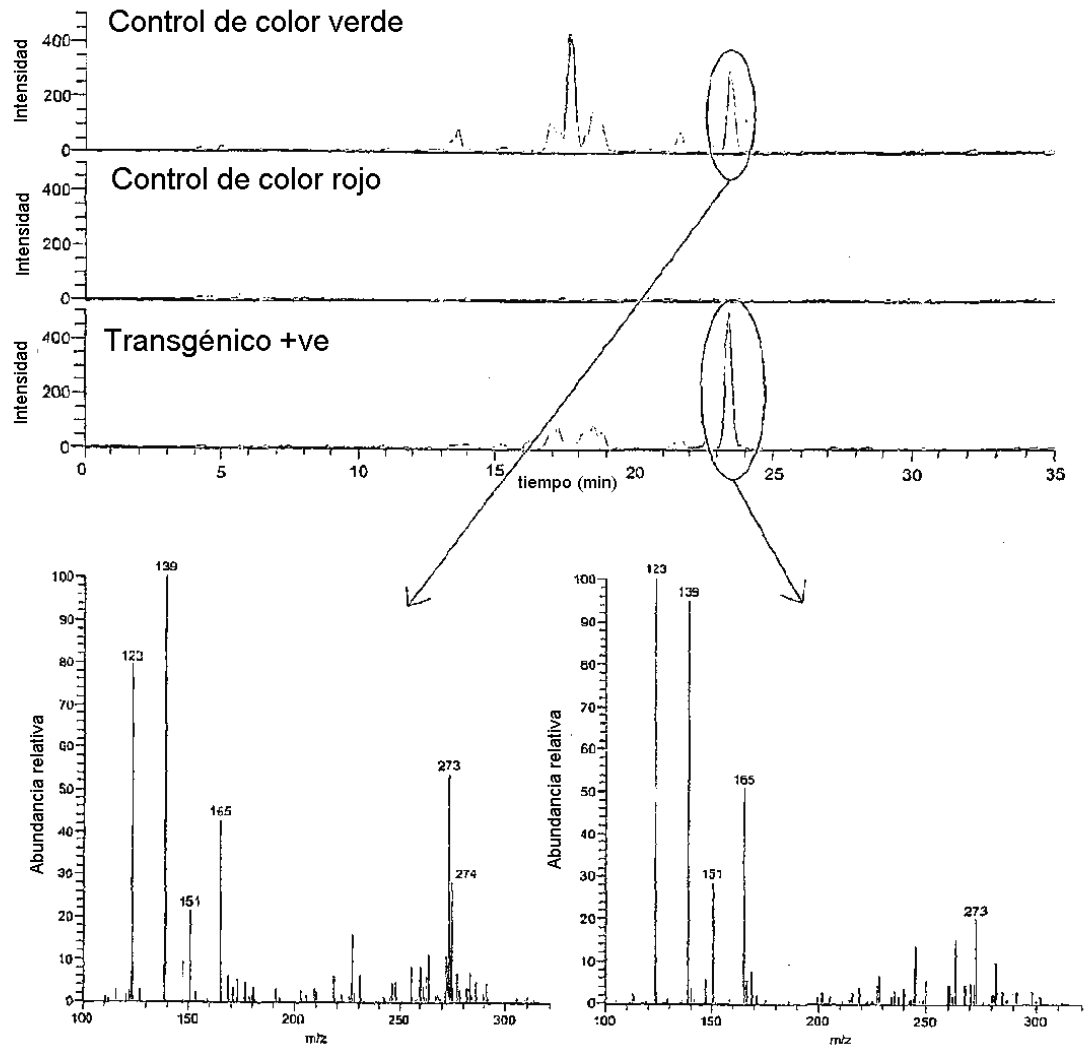


Figura 34

		1	50
TaMYB14-1	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TaMYB14-2	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TafMYB14-1	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TafMYB14-2	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
ToMYB14-1	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
ToMYB14-2	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TrMYB14-1	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TrMYB14-2	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TrMYB14-3	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
TrMYB14-4	(1)	MGRSPCCAKEGLNRGAWTAHEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
Consenso	(1)	MGRSPCCAKEGLNRGAWTQEDKILTEYIKLHGEGKWRNLPKRAGLKRCG	
		51	100
TaMYB14-1	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TaMYB14-2	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TafMYB14-1	(51)	KSCRLRWLNLYLRDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TafMYB14-2	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
ToMYB14-1	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
ToMYB14-2	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TrMYB14-1	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TrMYB14-2	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TrMYB14-3	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
TrMYB14-4	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
Consenso	(51)	KSCRLRWLNLYLRPDIKRGNISDEEELIIRLHKLLGNRWSLIAGRLPGRT	
		101	150
TaMYB14-1	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSPTKLSAQPKNAKIKQKQI--NP	
TaMYB14-2	(101)	DNEIKNYWNTNLGKKVKDLNQONTNNSPTKLSAQPKNAEIKQKQI--NP	
TafMYB14-1	(101)	DNEIKNYWNTNLGKKVKDLNQENTNNSPTKLSAQLKNAKIKQKQI--NP	
TafMYB14-2	(101)	DNEIKNYWNTNLGKKVKDLNQENTNNSPTKLSAQLKNAKIKQKQI--NP	

ToMYB14-1 (101) DNEIKNYWNTNLGKKVKDLNQONTNKSSPTKLSAQPKNAKIKOKQOI--NP
 ToMYB14-2 (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNAKIKOKQOINNP
 TrMYB14-1 (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNANIKOKQOI--NP
 TrMYB14-2 (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNANIKOKQOI--NP
 TrMYB14-3 (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNANIKOKQOI--NP
 TrMYB14-4 (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNANIKOKQOI--NP
 Consenso (101) DNEIKNYWNTNLGKKVKDLNQONTNNSPTKPSAQPKNAKIKOKQOI NP
 151 200

TaMYB14-1 (149) KPMKPN SNVVRTKATKCSKVLFINSLPNSP-MHDLQNKAEAEITTT-----
 TaMYB14-2 (149) K---PNSYVVRTKATKCSKVLFINSPNSPPMHDLQSKAEAEITTTTTPS
 TafMYB14-1 (149) KPEPNSNVVRTKATKCSKALFINSPNSPPMHDLQNKAEAEITTT--KSS
 TafMYB14-2 (149) KPEPNSNVVRTKATKCSKALFINSPNSPPMHDLQNKAEAEITTT--KSS
 ToMYB14-1 (149) KPMKPN SNVVRTKATKCSKVLFINSLPNSP-MHDLQNKAEAEITTT-----
 ToMYB14-2 (151) KPMKPN SNVVRTKATKCSKVLFINSP-----MHNLQNKAEAEETKT-----
 TrMYB14-1 (150) KPMKPN SNVVRTKATKCSKVLFINSP-----MHNLQNKAEAEETKT-----
 TrMYB14-2 (150) KPMKPN SNVVRTKATKCSKVLFINSP-----MHNLQNKAEAEETKT-----
 TrMYB14-3 (150) KPMKPN SNVVRTKATKCSKVLFINSP-----MHNLQNKAEAEETKT-----
 TrMYB14-4 (150) KPMKPN SNVVRTKATKCSKVLFINSP-----MHNLQNKAEAEETKT-----
 Consenso (151) KPMKPN SNVVRTKATKCSKVLFINSPNSP MHNLQNKAEAEITTT
 201 250

TaMYB14-1 (193) KPSMLVDGVASDSMSNNEMEHGYGFLSFCDEEKELSADLLEDFNIAADDIC
 TaMYB14-2 (196) MPSMLVDGVASDSMSNNEMECGNGFLSFCDEEKELSADLLEDFNIAADDIC
 TafMYB14-1 (197) MPSMLVDGVASDSMSNNEMEYGDGFVVSFCDDDKELSADLLEDFNISDDIC
 TafMYB14-2 (197) MPSMLVDGVASDSMSNNEMEYGDGFVVSFCDDDKELSADLLEDFNISDDIC
 ToMYB14-1 (193) KPSMLVDGVASDSMSNNEMEHGYGFLSFCDEEKELSADLLEDFNIAADDIC
 ToMYB14-2 (192) KTSMLVNGVASDSMSNNEMERGNGFLSFRDEEKELSADLLEDFNIAADDIC
 TrMYB14-1 (191) KPLMLVNGVASDSMSNNEMERGNGFLSFCDEEKELSADLLEDFNIAADDIC
 TrMYB14-2 (191) KPLMLVNGVASDSMSNNEMERGNGFLSFCDEEKELSADLLEDFNIAADDIC
 TrMYB14-3 (191) KPLMLVNGVASDSMSNNEMERGNGFLSFCDEEKELSADLLEDFNIAADDIC

ES 2 537 342 T3

TrMYB14-4 (191) KPLMLVNGVASDSMSNNEMERGNGLSFCDEEKELSADLLDDFNIADDIC
 Consenso (201) KPSMLVNGVASDSMSNNEMERGNGLSFCDEEKELSADLLDDFNIADDIC
 251 300

TaMYB14-1 (243) LSELLNSDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TaMYB14-2 (246) LSEFLNDFSNACDIDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TafMYB14-1 (247) LSEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNSTPCNFAAET
 TafMYB14-2 (247) LSEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNSTQCNFAAET
 ToMYB14-1 (243) LSELLNSDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 ToMYB14-2 (242) LSEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TrMYB14-1 (241) LSEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TrMYB14-2 (241) LPEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TrMYB14-3 (241) LSEFLNSDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 TrMYB14-4 (241) LSEFLNDFSNACNFDYNDLLSPCSDQTMFSDDEILKNWTQCNFADET
 Consenso (251) LSEFLNDFSNACNFDYNDLLSPCSDQTMFS[REDACTED]WTQCNFADET
 301 321

TaMYB14-1 (293) VSNNLHSFASFLESSEEVIGE
 TaMYB14-2 (296) VSNNLQSSASFLESSEEVIGE
 TafMYB14-1 (297) YVSNNQ-----SSEEVIGE
 TafMYB14-2 (297) -----
 ToMYB14-1 (293) VSNNLHSFASFLESSEEVIGE
 ToMYB14-2 (292) VSNNLHSFA-----
 TrMYB14-1 (291) VSNNLNSFASFLESSEEVIGE
 TrMYB14-2 (291) VSNNLNSFASFLESSEEVIGE
 TrMYB14-3 (291) VSNNLHSFASFLESSEEVIGE
 TrMYB14-4 (291) VSNNLHSFASFLESSEEVIGE
 Consenso (301) VSNNLHSFASFLESSEEVIGE

Figura 35

	TaMYB14-1	TaMYB14-2	TafMYB14-1	TafMYB14-2	ToMYB14-1	ToMYB14-2	TrMYB14-1	TrMYB14-2	TrMYB14-3	TrMYB14-4
TaMYB14-1			92	94	99	95	95	95	95	95
TaMYB14-2	95		92	93	94	92	93	92	93	93
TafMYB14-1				99	92	89	90	89	90	90
TafMYB14-2					94	92	92	91	92	92
ToMYB14-1						95	94	94	95	95
ToMYB14-2							98	98	99	99
TrMYB14-1								100	100	100
TrMYB14-2									99	99
TrMYB14-3										100
TrMYB14-4										