



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 318 004**

(51) Int. Cl.:

C12N 15/82 (2006.01)

C12N 9/02 (2006.01)

C12N 9/90 (2006.01)

A01H 5/00 (2006.01)

C12N 15/53 (2006.01)

C12N 15/61 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **02726831 .7**

(96) Fecha de presentación : **03.05.2002**

(97) Número de publicación de la solicitud: **1392106**

(97) Fecha de publicación de la solicitud: **03.03.2004**

(54) Título: **Genes *tyrA* y usos de los mismos.**

(30) Prioridad: **09.05.2001 US 289527 P**

(45) Fecha de publicación de la mención BOPI:
01.05.2009

(45) Fecha de la publicación del folleto de la patente:
01.05.2009

(73) Titular/es: **Monsanto Technology L.L.C.**
800 North Lindbergh Boulevard
St. Louis, Missouri 63167, US

(72) Inventor/es: **Valentin, Henry, E.;**
Mitsky, Timothy, A.;
Hao, Ming;
Karunandaa, Balasulojini y
Qi, Qungang

(74) Agente: **Carpintero López, Mario**

ES 2 318 004 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

Genes *tyrA* y usos de los mismos.

5 **Campo de la invención**

La presente invención pertenece al campo de la genética y la bioquímica vegetal. Más específicamente, la invención se refiere a genes asociados con la ruta de la biosíntesis del tocoferol. La presente invención proporciona e incluye moléculas de ácido nucleico, proteínas y anticuerpos asociados con los genes de la ruta de la biosíntesis del tocoferol. La presente invención también proporciona procedimientos para utilizar tales agentes, por ejemplo, en el aislamiento de genes, el análisis de genes y la producción de plantas transgénicas. Además, la presente invención incluye plantas transgénicas modificadas para expresar proteínas asociadas con la ruta del tocoferol. Además, la presente invención incluye procedimientos para la producción de productos procedentes de la ruta de la biosíntesis del tocoferol.

15 **Antecedentes de la invención**

Los tocoferoles son un componente esencial de la dieta de los mamíferos. Las pruebas epidemiológicas indican que el complemento de tocoferol puede dar como resultado una disminución del riesgo de la enfermedad cardiovascular y el cáncer, que puede ayudar en la función inmune y está asociado con la prevención y el retraso de un número de procesos patológicos degenerativos en seres humanos (Traber y Sies, *Annu. Rev. Nutr.* 16: 321-347 (1996)). El tocoferol funciona, en parte, estabilizando la bicapa lipídica de las membranas biológicas (Skrypin y Kagan, *Biochim. Biophys. Acta* 815: 209 (1995); Kagan, N. Y. *Acad. Sci.* p 121, (1989); Gomez-Fernandez *et al.*, *Ann. N. Y. Acad. Sci.* p 109 (1989)), reduciendo los radicales libres de ácidos grasos poliinsaturados (AGPI) generados por la oxidación lipídica (Fukuzawa *et al.*, "Lipids" 17: 511-513 (1982)), y barriendo los radicales libres de oxígeno, los radicales peróxido lipídicos y las especies de oxígeno singlete (Diplock *et al.* *Ann. NY Acad. Sci.* 570: 72 (1989); Fryer, *Plant Cell Environ.* 15 (4): 381- 392 (1992)).

El α -tocopherol, habitualmente denominado vitamina E, pertenece a una clase de antioxidantes solubles en lípidos que incluye α , β , γ y δ -tocopheroles, y α , β , γ y δ -tocotrienoles. Aunque los α , β , γ y δ -tocopheroles, y los α , β , γ y δ -tocotrienoles se denominan a veces colectivamente "vitamina E", la vitamina E se define químicamente más apropiadamente como α -tocopherol. El α -tocopherol es relevante para la salud humana, en parte porque se absorbe fácilmente y es retenido por el cuerpo, y tiene un grado más elevado de bioactividad que otras especies de tocoferol (Traber y Sies, *Annu. Rev. Nutr.* 16: 321-347 (1996)). Sin embargo, otros tocoferoles tales como los β , γ y δ -tocopheroles también tienen relevantes beneficios nutricionales y para la salud.

Los tocoferoles son sintetizados fundamentalmente sólo por plantas y otros ciertos organismos fotosintéticos, incluyendo cianobacterias. Por consiguiente, los tocoferoles para la dieta de los mamíferos se obtienen casi exclusivamente de estas fuentes. Los tejidos vegetales varían considerablemente en el contenido de tocoferoles totales y en la composición de los tocoferoles, siendo el α -tocopherol la especie de tocoferol que se encuentra predominantemente en los tejidos verdes de las plantas fotosintéticas. El tejido de las hojas puede contener de 10-50 μg de tocoferoles totales por gramo de peso fresco, pero la mayoría de los cultivos de los principales alimentos básicos a nivel mundial (p. ej., el arroz, el maíz, el trigo, la patata) producen niveles de tocoferoles totales bajos a extremadamente bajos, de los cuales sólo un bajo porcentaje corresponde al α -tocopherol (Hess, "Vitamin E, α -tocopherol, In Antioxidants in Higher Plants", R. Alscher y J. Hess, Eds., CRC Press, Boca Raton. pp. 111-134 (1993)). Los cultivos de semillas oleaginosas contienen generalmente niveles mucho más elevados de tocoferoles, pero el α -tocopherol sólo está presente como un componente minoritario (Taylor y Barnes, *Chem. Ind.*, Oct.: 722-726 (1981)).

La ingestión dietética diaria recomendada de 15-30 mg de vitamina E es bastante difícil de alcanzar a partir de una dieta estadounidense media. Por ejemplo, la satisfacción de esta ingestión de vitamina E diaria recomendada supondría más de 750 gramos de hojas de espinaca en las que el α -tocopherol comprende el 60% de los tocoferoles totales o 200-400 gramos de aceite de soja. Aunque es posible aumentar la dieta con complementos, la mayoría de estos complementos contienen vitamina E fundamentalmente sintética, que tiene seis estereoisómeros, mientras que la vitamina E natural está compuesta predominantemente sólo por un único isómero. Además, los complementos tienden a ser relativamente caros, y la población en general no se siente inclinada a tomar complementos vitamínicos con regularidad. Por lo tanto, existe la necesidad en la técnica de composiciones y procedimientos que bien aumenten la producción de tocoferoles totales o aumenten el porcentaje relativo del α -tocopherol producido por las plantas.

Además de los beneficios para la salud de los tocoferoles, el aumento de los niveles de α -tocopherol en los cultivos se ha asociado con una mejor estabilidad y una estabilidad de almacenamiento más prolongada de los productos vegetales (Peterson, *Cereal-Chem.* 72 (1): 21-24 (1995); Ball, "Fat-soluble vitamin assays in food analysis A comprehensive review", Londres, Elsevier Science Publishers Ltd. (1988)). Además, se ha observado que el complemento de tocoferol de los alimentos de cerdo, de ternera y de aves aumenta significativamente la calidad de la carne y prolonga la estabilidad de almacenamiento de los productos cárnicos procesados posteriormente, retardando la oxidación lipídica tras el procesamiento que aumenta los componentes de sabor no deseados (Sante y Lacourt, *J. Sci. Food Agric.* 65 (4): 503-507 (1994); Buckley *et al.*, *J. of Animal Science* 73: 3122-3130 (1995)).

Biosíntesis del tocoferol

Los plastos de plantas superiores presentan rutas bioquímicas interconectadas que conducen a metabolitos secundarios que incluyen los tocoferoles. La ruta biosintética del tocoferol de las plantas superiores implica la condensación de ácido homogentísico y fitilpirofosfato para formar 2-metil-6 fitilplastoquinol (Fiedler *et al.*, *Planta* 155: 511-515 (1982); Soll *et al.*, *Arch. Biochem. Biophys.* 204: 544-550 (1980); Marshall *et al.*, *Phytochem.* 24: 1705-1711 (1985)). Esta ruta del tocoferol de las plantas se puede dividir en cuatro partes: 1) la síntesis del ácido homogentísico, que contribuye al anillo aromático del tocoferol; 2) la síntesis del fitilpirofosfato, que contribuye a la cadena lateral del tocoferol; 3) la ciclización, que desempeña un papel en la quiralidad y la subestructura del cromanol de la familia de la vitamina E; 4) y la metilación dependiente de la S-adenosil-metionina de un anillo aromático, que afecta a la abundancia relativa de cada una de las especies de tocoferol.

Síntesis del ácido homogentísico

El ácido homogentísico es un precursor común tanto para los tocoferoles como para las plastoquinonas. En al menos algunas bacterias, se ha publicado que la síntesis del ácido homogentísico tiene lugar mediante la conversión de corismato en prefenato y luego en *p*-hidroxifenil-piruvato mediante una prefenato deshidrogenasa bifuncional. Los ejemplos de enzimas prefenato deshidrogenasa bacterianas bifuncionales incluyen las proteínas codificadas por los genes *tyrA* de *Erwinia herbicola* y *Escherichia coli*. El producto génico *tyrA* cataliza la producción de prefenato a partir de corismato, así como la posterior deshidrogenación del prefenato para formar *p*-hidroxifenilpiruvato (p-HPP), el precursor inmediato del ácido homogentísico. Entonces se convierte el p-HPP en ácido homogentísico mediante la hidroxifenil-piruvato dioxigenasa (HPPD). Por el contrario, se cree que las plantas carecen de actividad prefenato deshidrogenasa y se cree, en general, que la síntesis del ácido homogentísico a partir de corismato tiene lugar mediante la síntesis y la conversión de un arogenato intermedio. Como las rutas implicadas en la síntesis del ácido homogentísico también son responsables de la formación de la tirosina, cualquier modificación en estas rutas también puede dar como resultado la modificación de la síntesis de la tirosina y de la síntesis de otros aminoácidos aromáticos.

Síntesis de fitilpirofosfato

Los tocoferoles forman parte de la clase de los compuestos denominados isoprenoides. Otros isoprenoides incluyen carotenoides, giberilinas, terpenos, clorofilas y ácido abscísico. Un compuesto intermedio central en la producción de los isoprenoides es el isopentenil difosfato (IPP). Se han publicado rutas citoplasmáticas y basadas en plastos para generar IPP. La ruta citoplasmática implica las enzimas acetoacetyl CoA tiolasa, HMGC CoA sintasa, HMGC CoA reductasa, mevalonato quinasa, fosfomevalonato quinasa y mevalonato pirofosfato descarboxilasa.

Recientemente, se han obtenido pruebas de la existencia de una ruta biosintética de isoprenoides basada en plastos alternativa a partir de los estudios de investigación de los grupos de Rohmer y Arigoni (Eisenreich *et al.*, *Chem. Bio.*, 5: R221-R233 (1998); Rohmer, *Prog. Drug. Res.*, 50: 135-154 (1998); Rohmer, "Comprehensive Natural Products Chemistry", Vol. 2, pp. 45-68, Barton y Nakanishi (eds.), Pergamon Press, Oxford, Inglaterra (1999)), quienes descubrieron que los patrones de marcaje de isótopos observados en los estudios acerca de ciertos terpenoides eubacterianos y vegetales podrían no explicarse en términos de la ruta del mevalonato. Arigoni y colaboradores observaron posteriormente que la 1-desoxixilulosa, o una derivado de la misma, sirve como un compuesto intermedio de la nueva ruta, actualmente denominada la ruta MEP (Rohmer *et al.*, *Biochem. J.*, 295: 517-524 (1993); Schwarz, Ph. D. thesis, Eidgenössische Technische Hochschule, Zúrich, Suiza (1994)). Estudios recientes demuestran la formación de 1-desoxixilulosa 5-fosfato (Broers, Ph. D. thesis (Eidgenössische Technische Hochschule, Zúrich, Suiza) (1994)) a partir de una molécula de cada gliceraldehído 3-fosfato (Rohmer, "Comprehensive Natural Products Chemistry", Vol. 2, pp. 45-68, Barton y Nakanishi, eds., Pergamon Press, Oxford, Inglaterra (1999)) y piruvato (Eisenreich *et al.*, *Chem. Biol.*, 5: R223-R233 (1998); Schwarz *supra*; Rohmer *et al.*, *J. Arn. Chem. Soc.*, 118: 2564-2566 (1996); y Sprenger *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 94: 12857-12862 (1997)) por una enzima codificada por el gen *dxs* (Lois *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 95: 2105-2110 (1997); y Lange *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 95: 2100-2104 (1998)). Es posible convertir además la 1-desoxixilulosa 5-fosfato en 2-C-metileritritol 4-fosfato (Arigoni *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 94: 10600-10605 (1997)) mediante una reductoisomerasa catalizada por el gen *dxr* (Bouvier *et al.*, *Plant Physio.*, 117: 1421-1431 (1998); y Rohdich *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 96: 11758-11763 (1999)).

Los genes publicados de la ruta MEP también incluyen *ygbP*, que cataliza la conversión de 2-C-metileritritol 4-fosfato de en su respectivo derivado de citidil pirofosfato e *ygbB*, que cataliza la conversión de 4-fosfocitidil-2C-metil-D-eritritol en 2C-metil-D-eritritol-3,4-ciclofosfato. Estos genes están estrechamente relacionados en el genoma de *E. coli* (Herz *et al.*, *Proc. Natl. Acad. Sci. EE.UU.*, 97 (6): 2485-2490 (2000)).

Una vez formado el IPP mediante la ruta MEP, éste es convertido en GGDP por la GGDP sintasa, y luego en fitilpirofosfato, que es el constituyente principal de la cadena lateral del tocoferol.

Combinación y ciclización

El ácido homogentísico se combina bien con fitil-pirofosfato o con solanil-pirofosfato mediante la fitil/prenil transferasa formando 2-metil-6-fetil plastoquinol o 2-metil-6-solanil plastoquinol, respectivamente. El 2-metil-6-solanil plastoquinol es un precursor de la biosíntesis de las plastoquinonas, mientras que el 2-metil-6-fetil plastoquinol es convertido finalmente en tocoferol.

Metilación del anillo aromático

La principal diferencia estructural entre cada uno de los subtipos de tocoferol es la posición de los grupos metilo alrededor del anillo fenilo. Tanto el 2-metil-6-fetil plastoquinol como el 2-metil-6-solanil plastoquinol sirven como sus-
tratos para la 2-metil-6-fetilplastoquinol/2-metil-6-solanilplastoquinol-9 metiltransferasa (Metil Transferasa 1 o MT1),
que cataliza la formación de plastoquinol-9 y γ -tocoferol, respectivamente, mediante la metilación de la posición 7. La
posterior metilación en la posición 5 del γ -tocoferol por la γ -metil-transferasa genera el α -tocoferol biológicamente
activo. La tocoferol metil transferasa 2 (TMT2) muestra una actividad similar a la MT1.

Existe la necesidad en la técnica por moléculas de ácido nucleico codificantes de enzimas implicadas en la bio-
síntesis del tocoferol, así como, las enzimas y los anticuerpos relacionados para el aumento y la modificación de la
producción del tocoferol en plantas. Existe además la necesidad de organismos transgénicos que expresen aquellas mo-
léculas de ácido nucleico implicadas en la biosíntesis del tocoferol, que sean capaces de aumentar el valor nutricional
de alimentos y fuentes de alimentación.

ID del gen	Nombre de la enzima
<i>tyrA</i>	Prefenato deshidrogenasa
<i>slr1736</i>	Fitolprenil transferasa de <i>Synechocystis</i>
<i>ATPT2</i>	Fitolprenil transferasa de <i>Arabidopsis thaliana</i>
<i>DXS</i>	1-Desoxixilulosa-5-fosfato sintasa
<i>DXR</i>	1-Desoxixilulosa-5-fosfato reductoisomerasa
<i>GGPPS</i>	Geranilgeranil pirofosfato sintasa
<i>HPPD</i>	p-hidroxifenilpiruvato dioxigenasa
<i>AANT1</i>	Transportador del adenilato
<i>slr1737</i>	Tocoferol ciclasa
<i>IDI</i>	Isopentenil difosfato isomerasa
<i>GGH</i>	Geranilgeranil reductasa
<i>MT1</i>	Metil transferasa 1
<i>tMT2</i>	Tocoferol metil transferasa 2
<i>GMT</i>	Gamma metil transferasa

Como se usan en la presente memoria, la homogentisato fitil transferasa (HPT), fitilprenil transferasa (PPT), el
shr1736 y el *ATPT2* se refiere cada uno a proteínas o genes codificantes de proteínas que tienen la misma actividad
enzimática.

Resumen de la invención

La presente invención incluye y proporciona una molécula de ácido nucleico sustancialmente purificada según lo
expuesto en la reivindicación 1.

La presente invención incluye y proporciona una planta transformada según lo expuesto en la reivindicación 7.

La presente invención incluye y proporciona un procedimiento para producir una planta que tiene niveles aumen-
tados de tocoferol según lo expuesto en la reivindicación 12.

La presente invención incluye y proporciona un procedimiento para reducir los niveles de tocoferol en una planta
que comprende: (A) transformar dicha planta con una molécula de ácido nucleico, en el que dicha molécula de ácido
nucleico comprende como componentes ligados operativamente una región promotora exógena que funciona en células
vegetales para generar la producción de una molécula de ARNm, una molécula de ácido nucleico transcrita que tiene
una cadena transcrita y una cadena no transcrita, siendo dicha cadena transcrita complementaria a una molécula de
ácido nucleico que tiene la secuencia de ácido nucleico seleccionada del grupo constituido por SEC ID N.º 1 y 3; y
estando dicha molécula de ácido nucleico transcrita ligada a una secuencia no traducida en 3' que funciona en las
células vegetales para provocar la terminación de la transcripción y la adición de ribonucleótidos poliadenilados a un
extremo 3' de la secuencia de ARNm; y (B) cultivar dicha planta transformada.

La presente invención incluye y proporciona realizaciones preferidas según lo expuesto en las reivindicaciones subordinadas.

Descripción de las secuencias de ácidos nucleicos y aminoácidos

La SEC ID N.º 1 expone una secuencia de un ácido nucleico de una molécula de ADN que codifica una prefenato deshidrogenasa bifuncional de *Erwinia herbicola*.

La SEC ID N.º 2 expone una secuencia de aminoácidos derivada de una prefenato deshidrogenasa bifuncional de *Erwinia herbicola*.

La SEC ID N.º 3 expone una secuencia de ácido nucleico de una molécula de ADN que codifica una prefenato deshidrogenasa bifuncional de *Escherichia coli*.

La SEC ID N.º 4 expone una secuencia de aminoácidos derivada de una prefenato deshidrogenasa bifuncional de *Escherichia coli*.

La SEC ID N.º 5 expone un cebador 5' usado para la amplificación de una secuencia de *tyrA* de *Erwinia herbicola*.

La SEC ID N.º 6 expone un cebador 3' usado para la amplificación de una secuencia de *tyrA* de *Erwinia herbicola*.

La SEC ID N.º 7 expone un cebador 5' usado para la amplificación de una secuencia de *tyrA* de *Escherichia coli*.

La SEC ID N.º 8 expone un cebador 3' usado para la amplificación de una secuencia de *tyrA* de *Escherichia coli*.

La SEC ID N.º 9 expone una secuencia de cebador.

La SEC ID N.º 10 expone una secuencia de cebador.

La SEC ID N.º 11 expone una secuencia de cebador.

La SEC ID N.º 12 expone una secuencia de cebador.

Breve descripción de las figuras

La figura 1 es una gráfica que compara los niveles de tocoferoles totales de las semillas de las líneas de *Arabidopsis thaliana* que albergan un constructo de expresión de *tyrA* de *Erwinia herbicola* con una secuencia dirigida a plastos, una planta de tipo natural y una planta con un vector de control.

La Figura 2 es un esquema del constructo pMON26588.

La Figura 3 es un esquema del constructo pMON26589.

La Figura 4 es un esquema del constructo pMON26591.

La Figura 5 es un esquema del constructo pMON26590.

La Figura 6 es un esquema del constructo pMON36510.

La Figura 7 es un esquema del constructo pMON36512.

La Figura 8 es un esquema del constructo pMON36511.

La Figura 9 es un esquema del constructo pMON36520.

La Figura 10 es un esquema del constructo pCGN10822.

La Figura 11 es un esquema del constructo pMON36528.

La Figura 12 es un esquema del constructo pMON69907.

La Figura 13 es un esquema del constructo pMON69909.

La figura 14 representa el contenido de tocoferoles y tocotrienoles totales de semillas de *Arabidopsis* procedentes de plantas de tipo natural y varias líneas vegetales transformadas con el vector plasmídico pMON69907.

La figura 15 representa el contenido de tocoferoles totales de semillas de *Arabidopsis* procedentes de plantas de tipo natural y varias líneas vegetales transformadas con el vector plasmídico pMON69907.

ES 2 318 004 T3

La figura 16 representa el contenido de tocoferoles y tocotrienoles totales de semillas de *Arabidopsis thaliana* procedentes de plantas de tipo natural y varias líneas vegetales transformadas por separado con los vectores plasmídicos pCGN10822, pMON36528, pMON69907 y pMON69909.

5 La figura 17 representa un análisis detallado del contenido de tocoferoles y tocotrienoles de semillas de *Arabidopsis* procedentes de líneas vegetales transformadas con el vector pMON69909 en relación con semillas vegetales de tipo natural.

Las figuras 18a y 18b muestran ejemplos de constructos para la transformación en plantas.

10

La Figura 19 es un constructo esquemático de pMON36582.

10

La figura 20 es un ejemplo de vector lanzadera (pMON36586) que alberga un casete de expresión de la homogentisato fitiltransferasa (HPT) de *Arabidopsis* como un casete de Bsp120 I/Not I. El promotor de la napina y el terminador de la napina están fusionados a los extremos 5' y 3' respectivamente para dirigir la expresión de semilla específica.

15

La figura 21 representa diversos casetes de expresión de genes mostrados en un vector lanzadera (A) o en un vector binario (B).

20

La Figura 22 es un esquema del constructo pMON36596.

La Figura 23 es un esquema del constructo pMON36597.

25

La Figura 24 es un esquema del constructo pMON77601.

La Figura 25 es un esquema del constructo pMON77602.

La Figura 26 es un esquema del constructo pMON66657.

30

La Figura 27 es un esquema del constructo pMON66659.

La Figura 28 es un esquema del constructo pMON26541.

35

La Figura 29 es un esquema del constructo pMON26543.

La Figura 30 es un esquema del constructo pMON36176.

La figura 31 es una gráfica estándar de CL/EM.

40

La figura 32 es una gráfica de CL/EM.

La figura 33 es una cromatografía de CLAR.

45

La figura 34 es una cromatografía de CLAR.

La Figura 35 es un esquema del constructo pMON69915.

La Figura 36 es un esquema del constructo pMON69919.

50

La Figura 37 es un esquema del constructo pMON10098.

La Figura 38 es un esquema del constructo pMON36520.

55

La Figura 39 es un esquema del constructo pMON43853.

La Figura 40 es un esquema del constructo pMON36525.

La Figura 41 es un esquema del constructo pMON43861.

60

La Figura 42 es un esquema del constructo pCGN7770.

La Figura 43 es un esquema del constructo pMON69911.

65

La Figura 44 es un esquema del constructo pCGN11301.

La Figura 45 es un esquema del constructo pCGN3223.

La Figura 46 es un esquema del constructo pMON36575.

La Figura 47 es un esquema del constructo pMON38207R.

5 La Figura 48 es un esquema del constructo pMON36571.

La Figura 49 es un esquema del constructo pMON36576.

10 La Figura 50 es un esquema del constructo pMON69924.

La Figura 51 es un esquema del constructo pMON69943.

La Figura 52 es un esquema del constructo pMON69929.

15 La Figura 53 es un esquema del constructo pMON69936.

La Figura 54 es un esquema del constructo pMON36592.

20 La Figura 55 es un esquema del constructo pMON69945.

La Figura 56 es un esquema del constructo pMON16602.

La Figura 57 es un esquema del constructo pMON36525.

25 La Figura 58 es un esquema del constructo pMON58171.

La Figura 59 es un esquema del constructo pMON58172.

30 La Figura 60 es un esquema del constructo pMON58170.

La Figura 61 es un esquema del constructo pMON36591.

La Figura 62 es un esquema del constructo pMON36588.

35 La Figura 63 es un esquema del constructo pMON36592.

La Figura 64 es un esquema del constructo pMON58182.

40 La Figura 65 es un esquema del constructo pMON58176.

La Figura 66 es un esquema del constructo pMON58183.

La Figura 67 es un esquema del constructo pMON58185.

45 La Figura 68 es un esquema del constructo pMON36593.

La Figura 69 es un esquema del constructo pMON36589.

50 La Figura 70 es un esquema del constructo pMON36590.

La Figura 71 es un esquema del constructo pMON67162.

La Figura 72 es un esquema del constructo pMON58178.

55 La Figura 73 es un esquema del constructo pMON58186.

La Figura 74 es un esquema del constructo pMON58188.

60 La figura 75 muestra los niveles de tocoferol y tocotrienol, así como los niveles de HGA de líneas seleccionadas.

La figura 76 muestra los niveles de tocoferol y 2M6PPQ de líneas seleccionadas.

Descripción detallada

65 Cualquiera de las moléculas de ácido nucleico reveladas en la presente memoria puede ser aumentada o sobre-expresada en una variedad de organismos, tales como plantas, que pueden dar como resultado niveles más elevados de precursores del tocoferol tales como el ácido homogentísico (HGA) y finalmente, mayores niveles de tocoferoles en tales organismos. Además, el aumento de expresión o la sobre-expresión de las proteínas expuestas en la presente

memoria también pueden dar como resultado la producción de mayores niveles de plastoquinonas. Además, la presente invención proporciona un número de agentes, por ejemplo, moléculas de ácido nucleico y proteínas asociadas con la producción de tocoferoles, y proporciona usos de tales agentes.

La presente invención incluye y proporciona constructos de ácido nucleico para la expresión de prefenato deshidrogenasas bifuncionales en organismos en los que es deseable generar una mayor producción de ácido homogentésico, plastoquinonas o tocoferoles. Tales constructos de ácido nucleico se pueden usar en organismos para los que sea deseable un mayor nivel de actividad prefenato deshidrogenasa. La invención también incluye y proporciona constructos de ácidos nucleicos para la expresión de fitil prenilttransferasas en organismos en los que sea deseable generar una mayor producción de plastoquinonas o tocoferoles, y el uso de los constructos produciendo ácidos nucleicos antisentido frente a las fitil prenilttransferasas en organismos en los que sea deseable generar una mayor producción de ácido homogentésico.

Agentes

Los agentes de la invención serán preferiblemente “biológicamente activos” con respecto a bien un atributo estructural, tal como la capacidad de un ácido nucleico para hibridarse con otra molécula de ácido nucleico, o la capacidad de una proteína para ser unida por un anticuerpo (o para competir con otra molécula por tal unión). Alternativamente, tal atributo puede ser catalítico y, por tanto, implicar la capacidad del agente de mediar una reacción o una respuesta química. Los agentes estarán preferiblemente “sustancialmente purificados”. El término “sustancialmente purificado”, como se usa en la presente memoria, se refiere a una molécula separada de sustancialmente todas las demás moléculas que se encuentran normalmente asociadas con ella en su estado nativo. Más preferiblemente, una molécula sustancialmente purificada es la especie predominante presente en una preparación. Una molécula sustancialmente purificada puede estar más del 60% libre, preferiblemente, el 75% libre, más preferiblemente, el 90% libre, y lo más preferible, el 95% libre de otras moléculas (excluyendo el disolvente) presentes en la mezcla natural. El término “sustancialmente purificado” no pretende englobar las moléculas presentes en su estado nativo.

Los agentes de la invención también pueden ser recombinantes. Como se usa en la presente memoria, el término recombinante significa cualquier agente (p.ej., ADN, péptido etc.), que es, o resulta, aunque indirectamente, de la manipulación humana de una molécula de ácido nucleico.

Se entiende que los agentes de la invención pueden estar marcados con reactivos que faciliten la detección del agente (p.ej., marcadores fluorescentes, Prober *et al.*, *Science* 238: 336-340 (1987); Albarella *et al.*, EP 144914; marcadores químicos, Sheldon *et al.*, patente estadounidense n.º 4.582.789; Albarella *et al.*, patente estadounidense n.º 4.563.417; bases modificadas, Miyoshi *et al.*, EP 119448).

Moléculas de ácido nucleico

Los agentes de la invención incluyen moléculas de ácido nucleico que codifican una prefenato deshidrogenasa bifuncional que tiene actividades tanto corismato mutasa como prefenato deshidrogenasa. En una realización preferida, la molécula de ácido nucleico comprende una secuencia de ácido nucleico que tiene las SEC ID N.º 1 ó 3.

Un fragmento de una o más de las moléculas de ácido nucleico de la invención puede ser una sonda y específicamente una sonda PCR. Una sonda PCR es una molécula de ácido nucleico capaz de iniciar una actividad polimerasa cuando está en una estructura bicatenaria con otro ácido nucleico. Existen diversos procedimientos para determinar la estructura de las sondas PCR y diversas técnicas PCR en la técnica. Se pueden usar búsquedas generadas por ordenador usando programas tales como Primer3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi), STSPipeline (www-genome.wi.mit.edu/cgi-bin/www-STSPipeline) o GeneUp (Pesole *et al.*, *BioTechniques* 25: 112-123 (1998)), por ejemplo, para identificar posibles cebadores PCR.

Otro subconjunto de moléculas de ácido nucleico de la invención incluye moléculas de ácido nucleico que codifican una proteína o un fragmento de la misma.

Las moléculas de ácido nucleico o los fragmentos de las mismas de la presente invención son capaces de hibridarse específicamente con otras moléculas de ácido nucleico en ciertas circunstancias.

Como se usa en la presente memoria, se dice que dos moléculas de ácido nucleico son capaces de hibridarse específicamente entre sí si las dos moléculas son capaces de formar una estructura de ácido nucleico antiparalela bicatenaria.

Se dice que una molécula de ácido nucleico es “complementaria” a otra molécula de ácido nucleico si presentan una complementariedad total. Como se usa en la presente memoria, se dice que las moléculas presentan una “complementariedad total” cuando cada nucleótido de una de las moléculas es complementario a un nucleótido de la otra. Se dice que dos moléculas son “mínimamente complementarias” si pueden hibridarse entre sí con una estabilidad suficiente que les permita permanecer apareadas entre sí bajo condiciones al menos “poco restrictivas” convencionales. De manera similar, se dice que las moléculas son “complementarias” si pueden hibridarse entre sí con una estabilidad suficiente que les permita permanecer apareadas entre sí bajo condiciones “muy restrictivas” convencionales. En Sambrook *et al.*, “Molecular Cloning, A Laboratory Manual”, 2ª ed., Cold Spring Harbor Press, Cold Spring Harbor,

Nueva York (1989) y Haymes *et al.*, “Nucleic Acid Hybridization, A Practical Approach”, IRL Press, Washington, DC (1985), se describen las condiciones restrictivas convencionales. Por lo tanto, se permite una complementariedad incompleta, siempre y cuando tales complementariedades no excluyan por completo la capacidad de las moléculas para formar una estructura bicatenaria. De este modo, para que una molécula de ácido nucleico pueda servir como un cebador o una sonda, sólo es necesario que sea suficientemente complementaria secuencialmente para ser capaz de formar una estructura bicatenaria estable a las concentraciones de disolvente y sales particulares empleadas.

Las condiciones restrictivas apropiadas que promueven la hibridación del ADN son, por ejemplo, 6,0 x cloruro de sodio/citrato de sodio (SSC) a aproximadamente 45°C, seguidos por un lavado de 2,0 x SSC a 20-25°C, y son conocidas por los expertos en la técnica o se pueden encontrar en “Current Protocols in Molecular Biology”, John Wiley e hijos, N. Y. (1989), 6.3.1-6.3.6. Por ejemplo, la concentración de sales en la etapa de lavado se puede seleccionar entre una condición poco restrictiva de aproximadamente 2,0 x SSC a 50°C a una condición muy restrictiva de aproximadamente 0,2 x SSC a 65°C. Además, se puede aumentar la temperatura de la etapa de lavado de condiciones poco restrictivas a la temperatura ambiente, aproximadamente 22°C, a condiciones muy restrictivas a aproximadamente 65°C. Se puede variar tanto la temperatura como la concentración salina, o se puede mantener constante bien la temperatura o la concentración salina cambiando la otra variable.

Una o más de las moléculas de ácido nucleico de la presente invención difieren en la secuencia de ácido nucleico de aquéllas para las que se proporciona una secuencia específica en la presente memoria, porque uno o más codones han sido reemplazados con un codón que codifica una sustitución conservativa del aminoácido originariamente codificado.

Moléculas proteicas y peptídicas

Una clase de agentes incluye una o más de las proteínas o los fragmentos de las mismas o las moléculas peptídicas codificadas por un agente de ácido nucleico de la invención.

Los agentes proteicos o peptídicos pueden tener prolongaciones de la secuencias de aminoácidos C-terminales o N-terminales. Una clase de prolongaciones N-terminales empleadas en una realización preferida son los péptidos de tránsito a plastos. Cuando se emplean, los péptidos de tránsito a plastos se pueden ligar operativamente a la secuencia N-terminal, permitiendo así la localización del agente peptídico o proteico en los plastos. En una realización preferida, la secuencia dirigida a plastos es una secuencia de CTPI. En otra realización, la secuencia es una secuencia de CTP2.

Como se usa en la presente memoria, el término “molécula proteína” o “peptídica” incluye cualquier molécula que comprende cinco o más aminoácidos. Se sabe en la técnica que las proteínas pueden ser sometidas a una modificación, incluyendo modificaciones post-traduccionales, tales como, pero no limitándose a, la formación de enlaces disulfuro, glicosilación, fosforilación u oligomerización. De este modo, como se usa en la presente memoria, el término “molécula proteína” o “peptídica” incluye cualquier proteína que esté modificada mediante cualquier procedimiento biológico o no biológico. Los términos “aminoácido” y “aminoácidos” se refieren a todos los L-aminoácidos naturales. Esta definición pretende incluir norleucina, norvalina, ornitina, homocisteína y homoserina.

Una o más de moléculas proteicas o fragmentos de las mismas o moléculas peptídicas se pueden producir mediante síntesis química, o más preferiblemente, mediante la expresión en un huésped bacteriano o eucariota adecuado. En Sambrook *et al.*, “Molecular Cloning: A Laboratory Manual”, 2ª Ed., Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1989) o textos similares, se describen procedimientos adecuados para la expresión.

Un “fragmento proteico” es una molécula peptídica o polipeptídica cuya secuencia de aminoácidos comprende un subconjunto de la secuencia de aminoácidos de la proteína. Una proteína o un fragmento de la misma que comprende una o más regiones peptídicas adicionales no derivadas de la proteína es una proteína de “fusión”. Tales moléculas pueden ser derivadas para que contengan hidratos de carbono u otros restos (tales como hemocianina de lapa californiana). Las moléculas proteicas o peptídicas de fusión de la invención se producen preferiblemente mediante procedimientos recombinantes.

Constructos vegetales y transformantes vegetales

Se pueden usar una o más de las moléculas de ácido nucleico de la invención en la transformación o transfección de plantas. Se puede transferir material genético exógeno a una célula vegetal y regenerar la célula vegetal en una planta completa fértil o estéril. El material genético exógeno es cualquier material genético, bien natural o de otro tipo, procedente de cualquier fuente que sea capaz de ser insertado en cualquier organismo. En una realización preferida, el material genético exógeno codifica una preferato deshidrogenasa bifuncional o fragmentos de la misma, más preferiblemente, una preferato deshidrogenasa bifuncional procedente de un organismo procariota, e incluso más preferiblemente, una preferato deshidrogenasa bifuncional procedente de *Erwinia herbicola* o *Escherichia coli*.

En una realización de la presente invención, se introduce el material genético exógeno que comprende un homólogo de *tyrA* o un fragmento del mismo en una planta con uno o más genes adicionales. En una realización, las combinaciones de genes preferidas incluyen dos o más de los siguientes genes: *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, y un constructo antisentido para la ácido homogentísico dióxigenasa (Krindl *et al.*, *Seed Sci. Res.* 1: 209: 219 (1991); Keegstra, *Cell* 56 (2):47-53 (1989); Nawrath, *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 91: 12760-12764 (1994); Xia *et al.*, *J. Gen. Microbiol* 138: 1309-1316 (1992); Cyanobase en

la web www.kazusa.or.jp/cyanobase; Lois *et al.*, *Poc. Natl. Acad. Sci. EE.UU.* 95 (5): 2105-2110 (1998); Takahashi *et al.* *Proc. Natl. Acad. Sci. EE.UU.* 95 (17), 9879-9884 (1998); Norris *et al.*, *Plant Physiol.* 117: 1317-1323 (1998); Bartley y Scolnik, *Plant Physiol.* 104: 1469-1470 (1994), Smith *et al.*, *Plant J.* 11: 83-92 (1997); WO 00/32757; WO 00/10380; Saint Guily, *et al.*, *Plant Physiol.*, 100(2): 1069-1071 (1992); Sato *et al.*, *J. DNA Res.* 7 (1): 31-63 (2000)).

5 En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *slr1736* o *ATPT2*.

En tales combinaciones, se puede dirigir uno o más de los productos génicos hacia el plasto mediante el uso de una secuencia dirigida a plastos. Alternativamente, se puede localizar uno o más de los productos génicos en el citoplasma.

10 Se pueden introducir tales genes, por ejemplo, con el homólogo de *tyrA* o un fragmento del mismo sobre un solo constructo, introducir en diferentes constructos pero en la misma transformación o introducir en plantas separadas tras lo que se realizan uno o más cruces para generar la combinación de genes deseada. En tales combinaciones, un promotor preferido es un promotor de napina y una secuencia dirigida a plastos preferida es una secuencia de CTP1.

15 Se puede transferir tal material genético bien a monocotiledóneas y dicotiledóneas incluyendo, pero no limitándose a, canola, maíz, soja, *Arabidopsis phaseolus*, cacahuete, alfalfa, trigo, arroz, avena, sorgo, centeno, tritordeo, mijo, festuca, valluco perenne, caña de azúcar, arándano, papaya, plátano, cártamo, aceite de palma, lino, melón, manzana, pepino, dendrobio, gladiolo, crisantemo, lila, algodón, eucalipto, girasol, *Brassica campestris*, *Brassica napus*, césped, remolacha, café y dioscorea (Christou, en: "Particle Bombardment for Genetic Engineering of Plants", Biotechnology

20 Intelligence Unit. Academic Press, San Diego, California (1996)), siendo la canola, el maíz, *Arabidopsis*, *Brassica campestris*, *Brassica napus*, soja, crambe, mostaza, semilla de ricino, cacahuete, sésamo, semilla de algodón, semilla de lino, cártamo, aceite de palma, lino y girasol los preferidos, y los más preferidos canola, maíz, *Arabidopsis*, *Brassica campestris*, *Brassica napus*, soja, girasol, cártamo, aceite de palma y cacahuete. En una realización más preferida, el material genético se transfiere a canola. En otra realización más preferida, el material genético se transfiere a *Brassica*

25 *napus*. En otra más realización preferida, el material genético se transfiere a soja.

La transferencia de un ácido nucleico que codifica una proteína puede dar como resultado la expresión o la sobre-expresión de la proteína en una célula transformada o una planta transgénica. Una o más de las proteínas o los fragmentos de las mismas codificadas por las moléculas de ácido nucleico de la invención pueden ser sobre-expresadas

30 en una célula transformada o una planta transformada. Tal expresión o sobre-expresión puede ser el resultado de una transferencia transitoria o estable del material genético exógeno.

En una realización preferida, la expresión o la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un

35 mayor nivel de tocotrienoles.

En una realización preferida, la expresión o la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un

40 mayor nivel de tocoferoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un mayor nivel de α -

45 tocoferoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un mayor nivel de γ -

50 tocoferoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un mayor nivel de ácido

55 homogentésico.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, en relación con una planta sin transformar con un contexto genético similar, un mayor nivel de

60 plastoquinoles o plastoquinonas.

En algunas realizaciones, los niveles de uno o más productos de la ruta de la biosíntesis del tocoferol, incluyendo uno cualquiera o más entre tocotrienoles, tocoferoles, α -tocotrienoles, γ -tocotrienoles, plastoquinoles, plastoquinonas o

65 ácido homogentésico, se aumentan en un 10% o, más preferiblemente, en un 25%, 50%, 100%, 200%, 250%, 1.000%, 2.000% o 2.500%. Los niveles de los productos se pueden aumentar en todo un organismo, tal como una planta, o en uno o más órganos o tejidos específicos del organismo. Por ejemplo, se pueden aumentar los niveles de los productos en uno o más tejidos y órganos de una planta incluyendo sin limitación: raíces, tubérculos, tallos, hojas, troncos, fruto, bayas, frutos secos, corteza, vainas, semillas y flores.

En otra realización, la sobre-expresión de una proteína de la presente invención en una planta proporciona en esa planta, o en un tejido de esa planta, en relación con una planta o el tejido de la planta sin transformar con un contexto

70 genético similar, un mayor nivel de proteína prefenato deshidrogenasa.

En otra realización preferida, la sobre-expresión de una proteína de la presente invención en una planta transformada puede proporcionar tolerancia a una variedad de estreses, p. ej., tolerancia al estrés oxidativo tal como al oxígeno o al ozono, tolerancia a la radiación UV, tolerancia al frío o tolerancia a patógenos micóticos/microbianos.

Como se usa en la presente memoria, en un aspecto preferido, la tolerancia o la resistencia a un estrés está determinada por la capacidad de una planta, cuando es desafiada por un estrés tal como el frío, de producir una planta que tenga un mayor rendimiento que una sin tal tolerancia o resistencia a ese estrés. En un aspecto particularmente preferido de la presente invención, la tolerancia o la resistencia al estrés se mide en relación con una planta con un contexto genético similar a la planta tolerante o resistente a excepción de que la planta expresa o sobre-expresa una proteína de la presente invención.

El material genético exógeno puede ser transferido a una célula huésped mediante el uso de un vector de ADN o un constructo diseñado para tal efecto. El diseño de tal vector pertenece generalmente al alcance de la técnica (véase, "Plant Molecular Biology: A Laboratory Manual", Clark (ed.), Springer, Nueva York (1997)). En algunas realizaciones de esta invención se pueden transferir una secuencia de un solo gen seleccionado del grupo constituido por *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa a la planta diana deseada. En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *slr1736* o *ATPT2*. Las plantas diana que expresan la actividad deseada desde la secuencia génica transferida se pueden someter a uno o más cruces con plantas que hayan sido transformadas con una o más secuencias de otros genes seleccionados del grupo constituido por *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr 1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa para obtener plantas que expresen dos o más de las actividades deseadas desde la secuencia génica transferida. En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *slr1736* o *ATPT2*. En otra realización, los constructos de vectores de ADN pueden ser constructos de múltiples genes que comprenden dos o más secuencias génicas seleccionadas del grupo constituido por *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa tal que la transformación con un constructo de un solo vector de ADN resulte en la expresión de las dos o más de las secuencias génicas. En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *736* o *ATPT2*.

Un constructo o un vector pueden incluir un promotor vegetal que expresa la proteína o el fragmento proteico de elección. En una realización preferida, cualquier molécula de ácido nucleico descrita en la presente memoria puede ser ligada operativamente a una región promotora que funcione en una célula vegetal para generar la producción de una molécula de ARNm. Por ejemplo, se puede usar cualquier promotor que funcione en una célula vegetal para generar la producción de una molécula de ARNm, tal como aquellos promotores descritos en la presente memoria, sin limitación. En una realización preferida, el promotor es un promotor vegetal.

Se ha descrito un número de promotores que son activos en células vegetales en la bibliografía. Estos incluyen el promotor de la nopalina sintasa (NOS) (Ebert *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 84: 5745-5749 (1987)), el promotor de la octapina sintasa (OCS) (que es portado en plásmidos de inducción tumoral de *Agrobacterium tumefaciens*), los promotores caulimovirus tales como el promotor 19S del virus del mosaico de la coliflor (CaMV) (Lawton *et al.*, *Plant Mol. Biol.* 9: 315-324 (1987)) y el promotor 35S del CaMV (Odell *et al.*, *Nature* 313: 810-812 (1985)), el promotor 35S del virus del mosaico de la escrofularia, el promotor foto-inducible de la subunidad pequeña de ribulosa-1,5-bis-fosfato carboxilasa (spRUBISCO), el promotor de Adh (Walker *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 84: 6624-6628 (1987)), el promotor de la sacarosa sintasa (Yang *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 87: 4144-4148 (1990)), el promotor del complejo de genes R (Chandler *et al.*, "The Plant Cell" 1: 1175-1183 (1989)) y el promotor del gen proteico de unión a la clorofila a/b, etc. Estos promotores se han usado para crear constructos de ADN que han sido expresados en plantas; véase, p.ej., la publicación PCT WO 84/02913. Los promotores 35S del CaMV son los preferidos para su uso en plantas. Se pueden usar promotores de los que se sepa o se haya descubierto que generan la transcripción de ADN en células vegetales en la invención.

A efectos de la expresión en tejidos fuente de la planta, tales como la hoja, la semilla, la raíz o el tallo, es preferible que los promotores utilizados tengan una expresión relativamente elevada de estos tejidos específicos. La expresión de tejido específico de una proteína de la presente invención es una realización particularmente preferida. A tal efecto, se puede hacer la selección de entre un número de promotores con expresión de tejido específico, de célula específica o aumentada. Los ejemplos de tales promotores presentados en la bibliografía incluyen el promotor GS2 de la glutamina sintetasa cloroplástica procedente del guisante (Edwards *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 87: 3459-3463 (1990)), el promotor para la fructosa-1,6-bisfosfatasa (FBPasa) cloroplástica del trigo (Lloyd *et al.*, *Mol. Gen. Genet.* 225: 209-216 (1991)), el promotor ST-LS1 fotosintético nuclear de la patata (Stockhaus *et al.*, *EMBO J.* 8: 2445-2451 (1989)), el promotor para la serina/treonina quinasa (PAL) y el promotor para la glucoamilasa (CHS) de *Arabidopsis thaliana*. También se han descrito como activos en tejidos activos fotosintéticamente el promotor para la ribulosa-1,5-bisfosfato carboxilasa (RbcS) del alerce americano oriental (*Larix laricina*), el promotor para el gen *cab*, *cab6*, del pino (Yamamoto *et al.*, *Plant Cell Physiol.* 35: 773-778 (1994)), el promotor para el gen Cab-1 del trigo (Fejes *et al.*, *Plant Mol. Biol.* 15: 921-932 (1990)), el promotor para el gen CAB-1 de la espinaca (Lubberstedt *et al.*, *Plant Physiol.* 104: 997-1006 (1994)), el promotor para el gen cab1R del arroz (Luan *et al.*, *Plant Cell.* 4: 971-981 (1992)), el promotor de la piruvato ortofosfato diquinasa (PPDK) del maíz (Matsuoka *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 90: 9586-9590 (1993)), el promotor para el gen Lhcb1*2 del tabaco (Cerdan *et al.*, *Plant Mol. Biol.* 33: 245-255 (1997)), el promotor para el simportador de sacarosa-H⁺ SUC2 de *Arabidopsis thaliana* (Truernit *et al.*, *Planta.* 196:

564-570 (1995)) y el promotor para las proteínas de la membrana tilacoidea de la espinaca (*psaD*, *psaF*, *psaE*, *PC*, *FNR*, *atpC*, *atpD*, *cab*, *rbcS*). También se pueden utilizar otros promotores para las proteínas de unión a la clorofila a/b en la invención, tales como los promotores para el gen *LhcB* y el gen *PsbP* de la mostaza blanca (*Sinapis alba*; Kretsch *et al.*, *Plant Mol. Biol.* 28: 219-229 (1995)).

A efectos de la expresión en tejidos destino de la planta, tales como el tubérculo de la planta de la patata, el fruto del tomate o la semilla del maíz, del trigo, del arroz y de la cebada, es preferible que los promotores utilizados en la invención tengan una expresión relativamente elevada en estos tejidos específicos. Se conoce un número de promotores para genes con una expresión de tubérculo específico o aumentada en tubérculos, incluyendo el promotor para la patatina de clase I (Bevan *et al.*, *EMBO J.* 8: 1899-1906 (1986); Jefferson *et al.*, *Plant Mol. Biol.* 14: 995-1006 (1990)), el promotor para los genes ADPGPP del tubérculo de la patata, tanto la subunidad grande como la pequeña, el promotor de la sacarosa sintasa (Salanoubat y Belliard, *Gene* 60: 47-56 (1987), Salanoubat y Belliard, *Gene* 84: 181-185 (1989)), el promotor para las principales proteínas de tubérculos incluyendo los complejos proteicos de 22 kd y los inhibidores de la proteasa (Hannapel, *Plant Physiol.* 101: 703-704 (1993)), el promotor para el gen de la almidón sintasa unida a gránulos (GBSS) (Visser *et al.*, *Plant Mol. Biol.* 17: 691-699 (1991)) y otros promotores para patatinas de clase I y II (Koster-Topfer *et al.*, *Mol. Gen. Genet.* 219: 390-396 (1989); Mignery *et al.*, *Gene.* 62: 27-44 (1988)).

También se pueden usar otros promotores para expresar una proteína, o un fragmento de la misma, en tejidos específicos, tales como semillas o frutos. De hecho, en una realización preferida, el promotor usado es un promotor de una semilla específica. Los ejemplos de tales promotores incluyen las regiones reguladoras de 5' de genes tales como la napina (Krindl *et al.*, *Seed Sci. Res.* 1: 209: 219 (1991)), la faseolina (Bustos, *et al.*, *Plant Cell*, 1 (9): 839-853 (1989)), el inhibidor de la tripsina de la soja (Riggs, *et al.*, *Plant Cell* 1 (6): 609-621 (1989)), ACP (Baerson, *et al.*, *Plant Mol. Biol.*, 22 (2): 255-267 (1993)), esteroil-ACP desaturasa (Slocombe, *et al.*, *Plant Physiol.* 104 (4): 167-176 (1994)), subunidad α' de la β -conglucina de la soja (7s de la soja, (Chen *et al.*, *Proc. Natl. Acad. Sci.*, 83: 8560-8564 (1986))), y la oleosina (véase, por ejemplo, Hong, *et al.*, *Plant Mol. Biol.*, 34 (3): 549-555 (1997)). Otros ejemplos incluyen el promotor de la β -conglucina (Chen *et al.*, *Dev. Genet.* 10: 112-122 (1989)). También se incluyen las zeínas, que son un grupo de proteínas de almacenamiento que se encuentran en el endosperma del maíz. Se han aislado clones genómicos para genes de zeínas (Pedersen *et al.*, *Cell* 29: 1015-1026 (1982), y Russell *et al.*, *Transgenic Res.* 6 (2): 157-168), y también se podrían usar los promotores de estos clones, incluyendo los genes de 15 kD, 16 kD, 19 kD, 22 kD y 27 kD. Otros promotores conocidos por su funcionamiento en, por ejemplo, el maíz incluyen los promotores de los siguientes genes: *waxy*, *Brittle*, *Shrunken 2*, enzimas de ramificación I y II, almidón sintasas, enzimas de desramificación, oleosinas, glutelinas y sacarosa sintasas. Un promotor particularmente preferido para la expresión del endosperma del maíz es el promotor para el gen de la glutelina del arroz, más concretamente, el promotor *Osgt-1* (Zheng *et al.*, *Mol. Cell Biol.* 13: 5829-5842 (1993)). Los ejemplos de promotores adecuados para la expresión en trigo incluyen aquellos promotores para las subunidades de ADPglucosa piro sintasa (ADPGPP), la almidón sintasa unida a gránulos y otras almidón sintasa, las enzimas de ramificación y desramificación, las proteínas abundantes en la embriogénesis, las gliadinas y las gluteninas. Los ejemplos de tales promotores en el arroz incluyen aquellos promotores para las subunidades de ADPGPP, la almidón sintasa unida a gránulos y otra almidón sintasa, las enzimas de ramificación, las enzimas de desramificación, las sacarosa sintasas y las glutelinas. Un promotor particularmente preferido es el promotor para la glutelina del arroz, el *Osgt-1*. Los ejemplos de tales promotores en la cebada incluyen aquellos promotores para las subunidades de ADPGPP, la almidón sintasa unida a gránulos y otra almidón sintasa, las enzimas de ramificación, las enzimas de desramificación, las sacarosa sintasas, las hordeínas, las globulinas embriónicas y las proteínas específicas de la aleurona. Un promotor preferido para la expresión en la semilla es un promotor de la napina.

También se pueden usar promotores de una raíz específica. Un ejemplo de tal promotor es el promotor para el gen de la chitinasa ácida (Samac *et al.*, *Plant Mol. Biol.* 25: 587-596 (1994)). La expresión en tejido de raíz también se podría realizar utilizando subdominios de una raíz específica del promotor 35S del CaMV que haya sido identificado (Lam *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 86: 7890-7894 (1989)). Hay otros promotores de células de raíz específica que incluyen aquellos descritos por Conkling *et al.* (Conkling *et al.*, *Plant Physiol* 93: 1203-1211 (1990)).

En las patentes estadounidenses n.º 5.378.619; 5.391.725; 5.428.147; 5.447.858; 5.608.144; 5.608.144; 5.614.399; 5.633.441; 5.633.435; y 4.633.436, se describen otros promotores que se pueden utilizar. Además, se puede usar un potenciador de un tejido específico (Fromm *et al.*, *The Plant Cell* 1: 977-984 (1989)).

Los constructos o los vectores también pueden incluir, con la región codificante de interés, una secuencia de ácido nucleico que actúe, en la totalidad o en una parte, para terminar la transcripción de esa región. Se ha aislado un número de tales secuencias, incluyendo la secuencia 3' de Tr7 y la secuencia 3' de NOS (Ingelbrecht *et al.*, "The Plant Cell" 1: 671-680 (1989); Bevan *et al.*, *Nucleic Acids Res.* 11: 369-385 (1983)). También se pueden proporcionar regiones de terminación de transcritos reguladores en los constructos de expresión vegetal de esta invención. Las regiones de terminación de los transcritos pueden ser proporcionados por la secuencia de ADN codificante del gen de interés o una región de terminación de la transcripción conveniente derivada de una fuente génica diferente, por ejemplo, la región de terminación de transcritos que está asociada de manera natural con la región de iniciación de transcritos. El experto en la técnica reconocerá que, en los constructos de la presente invención, se puede emplear cualquier región de terminación de transcritos conveniente que sea capaz de terminar la transcripción en una célula vegetal.

Un vector o un constructo también pueden incluir elementos reguladores. Tales ejemplos incluyen el intrón 1 de *Adh* (Callis *et al.*, *Genes and Develop.* 1: 1183-1200 (1987)), el intrón de la sacarosa sintasa (Vasil *et al.*, *Plant Physiol.*

91: 1575-1579 (1989)) y el elemento omega del VMT (Gallic *et al.*, "The Plant Cell" 1: 301-311 (1989)). Se pueden incluir éstos y otros elementos reguladores cuando sea apropiado.

Un vector o un constructo también pueden incluir marcadores seleccionables. Los marcadores seleccionables también se pueden usar para seleccionar las plantas o las células vegetales que contengan el material genético exógeno. Tales ejemplos incluyen, pero no se limitan a: un gen *neo* (Potrykus *et al.*, *Mol. Gen. Genet.* 199: 183-188 (1985)), que codifica la resistencia a la kanamicina y se puede seleccionar para usar kanamicina, RptII, G418, hpt etc.; un gen bar que codifica la resistencia a bialafos; un gen de EPSP sintasa mutante (Hinchee *et al.*, *BiolTechnology* 6: 915-922 (1988); Reynaerts *et al.*, "Selectable and Screenable Markers". En Gelvin y Schilperoort. "Plant Molecular Biology Manual", Kluwer, Dordrecht (1988); Reynaerts *et al.*, "Selectable and screenable markers". En Gelvin y Schilperoort. "Plant Molecular Biology Manual", Kluwer, Dordrecht (1988)), *aadA* (Jones *et al.*, *Mol. Gen. Genet.* (1987)), que codifica la resistencia al glifosato; un gen de la nitrilasa que confiere resistencia al bromoxinilo (Stalker *et al.*, *J. Biol. Chem.* 263: 6310-6314 (1988)); un gen de la acetolactato sintasa mutante (ALS) que confiere resistencia a la imidazolinona o a la sulfonilurea (solicitud de patente europea 154.204 (11 de sept. de 1985)), ALS (D'Halluin *et al.*, *Bio/Technology* 10: 309-314 (1992)), y un gen de DHFR resistente al metotrexato (Thillet *et al.*, *J. Biol. Chem.* 263: 12500-12508 (1988)). Un vector o un constructo también pueden incluir un péptido de tránsito.

También se puede emplear la incorporación de un péptido de tránsito cloroplástico adecuado (publicación de solicitud de patente europea número 0218571). También se pueden incorporar ampliadores de la traducción como una parte del ADN vectorial. Los constructos de ADN podrían contener una o más secuencias líder no traducidas en 5' que pueden servir para aumentar la expresión de los productos génicos a partir de los transcritos de ARNm resultantes. Tales secuencias pueden ser derivadas del promotor seleccionado para expresar el gen o pueden ser modificados específicamente para aumentar la traducción del ARNm. También se pueden obtener tales regiones de ARN víricos procedentes de genes eucariotas adecuados, o de una secuencia génica sintética. Como publicación sobre la optimización de la expresión de transgenes, véase Kozziel *et al.*, *Plant Mol. Biol.* 32: 393-405 (1996). El péptido de tránsito preferido es el CTP1. En otra realización, el péptido de tránsito es una secuencia de CTP2.

Un vector o un constructo también pueden incluir un marcador detectable. Los marcadores detectables se pueden usar para controlar la expresión. Los ejemplos de marcadores detectables incluyen: un gen de β -glucuronidasa o *uidA* (GUS) que codifica una enzima para la que se conocen diversos sustratos cromogénicos (Jefferson, *Plant Mol. Biol. Rep.* 5: 387-405 (1987); Jefferson *et al.*, *EMBO J.* 6:3901-3907 (1987)); un gen del locus R, que codifica un producto que regula la producción de pigmentos de antocianina (color rojo) en tejidos vegetales (Dellaporta *et al.*, *Stadler Symposium* 11: 263-282 (1988)); un gen de la β -lactamasa (Sutcliffe *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 75: 3737-3741 (1978)), un gen que codifica una enzima para la que se conocen diversos sustratos cromogénicos (p.ej., PADAC, una cefalosporina cromogénica); un gen de la luciferasa (Ow *et al.*, *Science* 234: 856-859 (1986)); un gen *xylE* (Zukowsky *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 80: 1101-1105 (1983)) que codifica una catecol dioxigenasa que puede convertir los catecoles cromogénicos; un gen de la α -amilasa (Ikatu *et al.*, *BiolTechnol.* 8: 241-242 (1990)); un gen de la tirosina (Katz *et al.*, *J. Gen. Microbiol.* 129: 2703-2714 (1983)) que codifica una enzima capaz de oxidar la tirosina en DOPA y dopaquinona que, a su vez, se condensa en melanina; una α -galactosidasa, que convertirá un sustrato de α -galactosa cromogénico.

Incluidos en el término "genes marcadores seleccionables o detectables" también están los genes que codifican un marcador secretable cuya secreción puede ser detectada como un medio de identificación o selección de células transformadas. Los ejemplos incluyen marcadores que codifican un antígeno secretable que puede ser identificado mediante la interacción con anticuerpos, o incluso enzimas secretables que se pueden detectar catalíticamente. Las proteínas secretables pertenecen a un número de clases, incluyendo proteínas difusibles pequeñas que son detectables (p.ej., mediante ELISA), pequeñas enzimas activas, que son detectables en solución extracelular (p. ej., α -amilasa, β -lactamasa, fosfinotricina transferasa) o proteínas que están insertadas o atrapadas en la pared celular (tales como proteínas que incluyen una secuencia líder tal como la que se encuentra en la unidad de expresión de la prolongación o PR-S del tabaco). Hay otros posibles genes marcadores seleccionables y/o detectables que resultarán evidentes para los expertos en la técnica.

Hay muchos procedimientos para introducir moléculas de ácido nucleico transformantes en células vegetales. Se cree que los procedimientos adecuados incluyen casi cualquier procedimiento mediante el cual se puedan introducir moléculas de ácido nucleico en una célula, tales como mediante la infección con bacterias del género *Agrobacterium* o la administración directa de moléculas de ácido nucleico tal como, por ejemplo, mediante transformación mediada por PEG, mediante electroporación o mediante aceleración de partículas revestidas de ADN, etc. (Potrykus, *Ann. Rev. Plant Physiol. Plant Mol. Biol.* 42: 205-225 (1991); Vasil, *Plant Mol. Biol.* 25: 925-937 (1994)). Por ejemplo, se ha usado la electroporación para transformar protoplastos de maíz (Fromm *et al.*, *Nature* 312: 791-793 (1986)).

Otros sistemas vectoriales adecuados para introducir ADN transformante en una célula vegetal huésped incluyen, pero no se limitan a, vectores de cromosomas artificiales binarios (BIBAC) (Hamilton *et al.*, *Gene* 200: 107-116 (1997)); y la transfección con vectores víricos de ARN (Della-Cioppa *et al.*, *Ann. N Y Acad. Sci.* (1996), 792 ("Engineering Plants for Commercial Products and Applications"), 57-61). Otros sistemas vectoriales incluyen también vectores YAC seleccionables vegetales tales como aquéllos descritos en Mullen *et al.*, "Molecular Breeding" 4: 449-457 (1988).

La tecnología para introducir el ADN en células es conocida por aquéllos expertos en la técnica. Se han descrito cuatro procedimientos generales para la administración de un gen en células:

- (1) procedimientos químicos (Graham y van der Eb, *Virology* 54: 536-539(1973));
- (2) procedimientos físicos tales como la microinyección (Capecchi, *Cell* 22: 479-488 (1980)), la electroporación (Wong y Neumann, *Biochem. Biophys. Res. Commun.* 107: 584-587 (1982); Fromm *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 82: 5824-5828 (1985); patente estadounidense n.º 5.384.253); la pistola de genes (Johnston y Tang, *Methods Cell Biol.* 43: 353-365 (1994)); y la infiltración al vacío (Bechtold *et al.*, *C. R. Acad. Sci. Paris, Life Sci.* 316:1194-1199 (1993)); (3) vectores víricos (Clapp, *Clin. Perinatol.* 20: 155-168 (1993); Lu *et al.*, *J Exp. Med.* 178: 2089-2096 (1993); Eglitis y Anderson, *Biotechniques* 6: 608-614 (1988)); y (4) mecanismos mediados por receptores (Curiel *et al.*, *Hum. Gen. Ther.* 3: 147-154 (1992), Wagner *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 89: 6099-6103 (1992)).

Los procedimientos de aceleración que se pueden usar incluyen, por ejemplo, bombardeo de microproyectiles y similares. Un ejemplo de un procedimiento para introducir moléculas de ácido nucleico transformantes en células vegetales es el bombardeo de microproyectiles. Este procedimiento ha sido revisado por Yang y Christou (eds.), "Particle Bombardment Technology for Gene Transfer", Oxford Press, Oxford, Inglaterra (1994)). Partículas no biológicas (microproyectiles) que pueden ser revestidos con ácidos nucleicos e introducidas en células mediante una fuerza de propulsión. Los ejemplos de partículas incluyen aquéllas compuestas por tungsteno, oro, platino y similares.

Una particular ventaja del bombardeo de microproyectiles es que además de ser un procedimiento eficaz para transformar reproduciblemente monocotiledóneas, no es necesario ni el aislamiento de protoplastos (Cristou *et al.*, *Plant Physiol.* 87: 671-674 (1988)) ni la susceptibilidad a la infección con *Agrobacterium*. Una realización ilustrativa de un procedimiento de introducción de ADN en células de maíz mediante aceleración es un sistema biolístico de administración de partículas α , que se puede usar para propulsar partículas revestidas de ADN a través de una pantalla, tal como una pantalla de acero inoxidable o Nytex, sobre una superficie filtrante cubierta con células de maíz cultivadas en suspensión. Gordon-Kamm *et al.* describen el procedimiento básico para revestir partículas de tungsteno con ADN (Gordon-Kamm *et al.*, *Plant Cell* 2: 603-618 (1990)). La pantalla dispersa las partículas de ácido nucleico de tungsteno de manera que no son enviadas a las células receptoras en forma de grandes agregados. Un sistema de administración de partículas adecuado para su uso con la invención es la pistola PDS-1000/He de aceleración de helio, que se encuentra disponible en los laboratorios Bio-Rad (Bio-Rad, Hercules, California) (Sanford *et al.*, *Technique* 3: 3-16 (1991)).

Para el bombardeo, se pueden concentrar sobre los filtros las células en suspensión. Los filtros que contienen las células que van a ser bombardeadas se colocan a una distancia apropiada por debajo de la placa de detención de los microproyectiles. Si se desea, también se colocan una o más pantallas entre la pistola y las células que van a ser bombardeadas.

Alternativamente, se pueden disponer embriones inmaduros u otras células diana sobre medio de cultivo sólido. Las células que van a ser bombardeadas se colocan a una distancia apropiada por debajo de la placa de detención de los microproyectiles. Si se desea, también se colocan una o más pantallas entre el dispositivo de aceleración y las células que van a ser bombardeadas. Mediante el uso de las técnicas expuestas en la presente memoria es posible obtener 1.000 o más locus de células que expresen transitoriamente un gen marcador. El número de células de un foco que expresa el producto génico exógeno 48 horas después del bombardeo habitualmente varía de una a diez, y una media de uno a tres.

En la transformación por bombardeo, es posible optimizar las condiciones de cultivo previas al bombardeo y los parámetros del bombardeo para producir los números máximos de transformantes estables. Los parámetros tanto físicos como biológicos para el bombardeo son importantes en esta tecnología. Los factores físicos son aquéllos que implican la manipulación del precipitado de ADN/microproyectiles o aquéllos que afectan al vuelo y a la velocidad de bien los macro- o micro-proyectiles. Los factores biológicos incluyen todas las etapas implicadas en la manipulación de las células antes e inmediatamente después del bombardeo, el ajuste osmótico de las células diana para ayudar a calmar el traumatismo asociado con el bombardeo y también la naturaleza del ADN transformante, tal como ADN linealizado o plásmidos super-enrollados intactos. Se cree que las manipulaciones previas al bombardeo son especialmente importantes para la correcta transformación de los embriones inmaduros.

En otra realización alternativa, se pueden transformar establemente plastos. Los procedimientos revelados para la transformación de plastos en plantas superiores incluyen la administración con una pistola de partículas de ADN que contiene un marcador seleccionable y la dirección del ADN hacia el genoma del plasto a través de recombinación de homólogos (Svab *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 87: 8526-8530 (1990); Svab y Maliga, *Proc. Natl. Acad. Sci.* (EE.UU.) 90: 913-917 (1993); Staub y Maliga, *EMBO J.* 12: 601-606 (1993); patentes estadounidenses n.º 5.451.513 y 5.545.818).

Por consiguiente, se contempla que es posible desear el ajuste de diversos aspectos de los parámetros de bombardeo en estudios a pequeña escala para optimizar las condiciones por completo. Puede que se desee en particular ajustar parámetros físicos tales como la distancia entre huecos, la distancia de vuelo, la distancia del tejido y la presión del helio. También es posible minimizar los factores de reducción de los traumatismos modificando las condiciones que

influyen en el estado fisiológico de las células receptoras y que pueden influir por tanto en la eficacia de transformación y de integración. Por ejemplo, se pueden ajustar para una transformación óptima el estado osmótico, la hidratación de los tejidos y la etapa del subcultivo o el ciclo celular de las células receptoras. La ejecución de otros ajustes rutinarios será conocida por aquéllos expertos en la técnica a la luz de la presente revelación.

La transferencia mediada por *Agrobacterium* es un sistema ampliamente aplicable para introducir genes en células vegetales, porque se puede introducir el ADN en tejidos de toda la planta, superando así la necesidad de regeneración de una planta intacta a partir de un protoplasto. El uso de vectores de la integración en plantas mediada por *Agrobacterium* para introducir ADN en células vegetales es conocido en la técnica. Véanse, por ejemplo, los procedimientos descritos por Fraley *et al.*, *BiolTechnology* 3: 629-635 (1985) y Rogers *et al.*, *Methods Enzymol.* 153: 253-277 (1987). Además, la integración del ADN de Ti es un procedimiento relativamente preciso que da como resultado pocos reajustes. La región del ADN por ser transferida está definida por las secuencias de los bordes, y el ADN intermedio se inserta habitualmente en el genoma de la planta según lo descrito (Spielmann *et al.*, *Mol. Gen. Genet.* 205: 34 (1986)).

Los vectores modernos de transformación con *Agrobacterium* son capaces de realizar la replicación en *E. coli*, así como en *Agrobacterium*, permitiendo las manipulaciones convenientes según lo descrito (Klee *et al.*, en: "Plant DNA Infectious Agents", Hohn y Schell (eds.), Springer-Verlag, Nueva York, pp. 179-203 (1985)). Además, los avances tecnológicos en vectores para la transferencia de genes mediada por *Agrobacterium* han mejorado la disposición de los genes y de los sitios de restricción en los vectores para facilitar la construcción de los vectores capaces de expresar diversos genes codificantes de polipéptidos. Los vectores descritos tienen regiones multi-ligadoras convenientes flanqueadas por un promotor y un sitio de poliadenilación para la expresión directa de genes codificantes de polipéptidos insertados y son adecuados para los presentes objetivos (Rogers *et al.*, *Methods Enzymol.* 153: 253-277 (1987)). Además, se puede usar *Agrobacterium* que contenga genes de Ti tanto armados como desarmados para las transformaciones. En aquellas cepas vegetales en las que la transformación mediada por *Agrobacterium* sea eficaz, éste es el procedimiento de elección debido a la naturaleza fácil y definida de la transferencia génica.

Una planta transgénica formada usando procedimientos de transformación con *Agrobacterium* contiene comúnmente un solo gen sobre un cromosoma. Tales plantas transgénicas se pueden denominar como heterocigóticas para el gen añadido. Es más preferible una planta transgénica que sea homocigótica para el gen estructural añadido; i.e., una planta transgénica que contenga dos genes añadidos, un gen en el mismo locus de cada cromosoma de un par de cromosomas. Se puede obtener una planta transgénica homocigótica mediante el apareamiento sexual (la autofecundación) de un segregante independiente, una planta transgénica que contenga un solo gen añadido, mediante la germinación de alguna de las semillas producidas y el análisis de las plantas resultantes producidas para el gen de interés.

Además, se entenderá que también se pueden aparear dos plantas transgénicas diferentes para producir descendencia que contenga dos genes exógenos segregantes independientemente. La autofecundación de la progenie apropiada puede producir plantas que son homocigóticas para ambos genes exógenos añadidos que codifican un polipéptido de interés. También se contemplan el retrocruzamiento hacia una planta parental y el cruzamiento lejano con una planta no transgénica, al igual que la propagación vegetativa.

Es posible realizar la transformación de protoplastos vegetales usando los procedimientos basados en la precipitación de fosfato de calcio, el tratamiento con polietilenglicol, la electroporación y combinaciones de estos tratamientos (véase, por ejemplo, Potrykus *et al.*, *Mol. Gen. Genet.* 205: 193-200 (1986); Lorz *et al.*, *Mol. Gen. Genet.* 199: 178 (1985); Fromm *et al.*, *Nature* 319: 791 (1986); Uchimiya *et al.*, *Mol. Gen. Genet.* 204: 204 (1986); Marcotte *et al.*, *Nature* 335: 454-457 (1988)).

La aplicación de estos sistemas a diferentes cepas vegetales depende de la capacidad de regeneración de esa cepa vegetal en particular a partir de protoplastos. Hay descritos procedimientos ilustrativos para la regeneración de cereales a partir de protoplastos (Fujimura *et al.*, *Plant Tissue Culture Letters* 2: 74 (1985); Toriyama *et al.*, *Theor. Appl. Genet.* 205: 34 (1986); Yamada *et al.*, *Plant Cell Rep.* 4: 85 (1986); Abdullah *et al.*, *Biotechnology* 4: 1087 (1986)).

Para transformar cepas vegetales que no pueden ser regeneradas satisfactoriamente a partir de protoplastos, se pueden utilizar otros modos de introducir ADN en células o tejidos intactos. Por ejemplo, se puede efectuar la regeneración de cereales a partir de embriones inmaduros o explantos según lo descrito (Vasil, *Biotechnology* 6: 397 (1988)). Además, se puede utilizar una tecnología de "pistola de partículas" o de microproyectiles a alta velocidad (Vasil *et al.*, *BioTechnology* 10: 667 (1992)).

Usando esta última tecnología, se transporta el ADN a través de la pared celular y se introduce en el citoplasma sobre la superficie de pequeñas partículas metálicas según lo descrito (Klein *et al.*, *Nature* 328: 70 (1987); Klein *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 85: 8502-8505 (1988); McCabe *et al.*, *BiolTechnology* 6: 923 (1988)). Las partículas metálicas penetran a través de varias cepas de las células, permitiendo así la transformación de las células en explantos de tejidos.

También se pueden usar otros procedimientos de transformación celular e incluyen, pero no se limitan a, la introducción de ADN en plantas mediante la transferencia directa de ADN en el polen (Hess *et al.*, *Intern Rev. Cytol.* 107: 367 (1987); Luo *et al.*, *Plant Mol Biol. Reporter* 6: 165 (1988)), mediante la inyección directa de ADN en los órganos reproductores de una planta (Pena *et al.*, *Nature* 325: 274 (1987)), o mediante la inyección directa de ADN en células

de embriones inmaduros seguida por la rehidratación de embriones disecados (Neuhaus *et al.*, *Theor. Appl. Genet.* 75: 30 (1987)).

La regeneración, el desarrollo y el cultivo de plantas a partir de transformantes de un solo protoplasto vegetal o a partir de diversos explantos transformados son conocidos en la técnica (Weissbach y Weissbach, en: "Methods for Plant Molecular Biology", Academic Press, San Diego, CA, (1988)). Este procedimiento de regeneración y crecimiento incluye comúnmente las etapas de selección de células transformadas, el cultivo de aquellas células individualizadas a través de las etapas habituales del desarrollo embrionario en la etapa de plántula enraizada. Los embriones y las semillas transgénicos son regenerados similarmente. Los brotes enraizados transgénicos resultantes son después plantados en un medio de crecimiento vegetal apropiado tal como tierra.

El desarrollo o la regeneración de plantas que contienen el gen foráneo exógeno que codifica una proteína de interés son conocidos en la técnica. Preferiblemente, las plantas regeneradas se auto-polinizan para proporcionar plantas transgénicas homocigóticas. Aparte de eso, el polen obtenido de las plantas regeneradas se cruza con plantas cultivadas a partir de semillas de líneas importantes agrónomicamente. Por el contrario, el polen de las plantas de estas líneas importantes se usa para polinizar plantas regeneradas. Una planta transgénica de la invención que contiene un polipéptido deseado se cultiva usando los procedimientos conocidos por los expertos en la técnica.

Hay una variedad de procedimientos para la regeneración de plantas a partir de tejido vegetal. El procedimiento particular de regeneración dependerá del tejido vegetal inicial y de la especie vegetal en particular que se vaya a regenerar.

Se han publicado procedimientos para transformar dicotiledóneas, fundamentalmente mediante el uso de *Agrobacterium tumefaciens* y obtener plantas transgénicas para el algodón (patente estadounidense n.º 5.004.863; patente estadounidense n.º 5.159.135; patente estadounidense n.º 5.518.908); la soja (patente estadounidense n.º 5.569.834; patente estadounidense n.º 5.416.011; McCabe *et al.*, *Biotechnology* 6: 923 (1988); Christou *et al.*, *Plant Physiol.* 87: 671-674 (1988)); *Brassica* (patente estadounidense n.º 5.463.174); el cacahuete (Cheng *et al.*, *Plant Cell Rep.* 15: 653-657 (1996), McKently *et al.*, *Plant Cell Rep.* 14: 699-703 (1995)); la papaya; el guisante (Grant *et al.*, *Plant Cell Rep.* 15: 254-258 (1995)); y *Arabidopsis thaliana* (Bechtold *et al.*, *C. R. Acad. Sci. Paris, Life Sci.* 316: 1194-1199 (1993)). Este último procedimiento para transformar *Arabidopsis thaliana* se denomina comúnmente "inmersión" o infiltración al vacío, o transformación germoplasmática.

También se ha descrito la transformación de monocotiledóneas mediante el uso de electroporación, bombardeo de partículas y *Agrobacterium*. La transformación y la regeneración vegetal se han realizado en el espárrago (Bytebier *et al.*, *Proc. Natl. Acad. Sci. (EE.UU.)* 84: 5354 (1987)); la cebada (Wan y Lemaux, *Plant Physiol.* 104: 37 (1994)); el maíz (Rhodes *et al.*, *Science* 240: 204 (1988); Gordon-Kamm *et al.*, *Plant Cell* 2: 603-618 (1990); Fromm *et al.*, *Bio/Technology* 8: 833 (1990); Koziel *et al.*, *Bio/Technology* 11: 194 (1993); Armstrong *et al.*, *Crop Science* 35: 550-557 (1995)); la avena (Somers *et al.*, *Bio/Technology* 10: 1589 (1992)); la pata de gallina (Horn *et al.*, *Plant Cell Rep.* 7: 469 (1988)); el arroz (Toriyama *et al.*, *Theor Appl. Genet.* 205:34 (1986); Part *et al.*, *Plant Mol. Biol.* 32: 1135-1148 (1996); Abedinia *et al.*, *Aust. J. Plant Physiol.* 24: 133-141 (1997); Zhang y Wu, *Theor. Appl. Genet.* 76: 835 (1988); Zhang *et al.*, *Plant Cell Rep.* 7: 379 (1988); Battraw y Hall, *Plant Sci.* 86: 191-202 (1992); Christou *et al.*, *Bio/Technology* 9: 957 (1991)); el centeno (De la Pena *et al.*, *Nature* 325: 274 (1987)); la caña de azúcar (Bower y Birch, *Plant J.* 2: 409 (1992)); la festuca alta (Wang *et al.*, *Bio/Technology* 10: 691 (1992)) y el trigo (Vasil *et al.*, *Bio/Technology* 10: 667 (1992); patente estadounidense n.º 5.631.152).

Se han desarrollado análisis para la expresión génica basados en la expresión transitoria de constructos de ácido nucleico clonados mediante la introducción de moléculas de ácido nucleico en células vegetales mediante el tratamiento con polietilenglicol, la electroporación o el bombardeo de partículas (Marcotte *et al.*, *Nature* 335: 454-457 (1988); Marcotte *et al.*, *Plant Cell* 1: 523-532 (1989); McCarty *et al.*, *Cell* 66: 895-905 (1991); Hattori *et al.*, *Genes Dev.* 6: 609-618 (1992); Goff *et al.*, *EMBO J.* 9:2517-25222 (1990)). Se pueden usar sistemas de expresión transitoria para diseccionar funcionalmente constructos génicos (véase, en general, Mailga *et al.*, "Methods in Plant Molecular Biology", Cold Spring Harbor Press (1995)).

Se puede introducir cualquiera de las moléculas de ácido nucleico de la invención en una célula vegetal de una manera permanente o transitoria, en combinación con otros elementos genéticos tales como vectores, promotores, potenciadores, etc. Además, se puede introducir cualquiera de las moléculas de ácido nucleico de la invención en una célula vegetal de una manera que permita la expresión o la sobre-expresión de la proteína o del fragmento de la misma codificada por la molécula de ácido nucleico.

La co-supresión es la reducción de los niveles de expresión, habitualmente a nivel del ARN, de un determinado gen endógeno o una familia de genes mediante la expresión de un constructo sentido homólogo que sea capaz de transcribir ARNm del mismo encadenamiento como el transcripto del gen endógeno (Napoli *et al.*, *Plant Cell* 2: 279-289 (1990); van der Krol *et al.*, *Plant Cell* 2: 291-299 (1990)). La co-supresión puede resultar de la transformación estable con una sola molécula de ácido nucleico de copia que sea homóloga a la secuencia de ácido nucleico encontrada con la célula (Prolls y Meyer, *Plant J.* 2: 465-475 (1992)) o con múltiples copias de una molécula de ácido nucleico que sea homóloga a una secuencia de ácido nucleico encontrada con la célula (Mittlesten *et al.*, *Mol. Gen. Genet.* 244: 325-330 (1994)). Los genes, incluso los diferentes, unidos a promotores homólogos pueden dar como resultado la cosupresión de los genes ligados (Vaucheret, *C. R. Acad. Sci. III* 316: 1471-1483 (1993); Flavell, *Proc. Natl. Acad. Sci. (EE.UU.)*

91: 3490-3496 (1994)); van Blokland *et al.*, *Plant J.* 6: 861-877 (1994); Jorgensen, *Trends Biotechnol.* 8: 340-344 (1990); Meins y Kunz, en: "Gene Inactivation and Homologous Recombination in Plants", Paszkowski (ed.), pp. 335-348, Kluwer Academic, Países Bajos (1994)).

- 5 Se entiende que es posible introducir uno o más de los ácidos nucleicos de la invención en una célula vegetal y transcribirlos usando un promotor apropiado con tal transcripción dando como resultado la cosupresión de una proteína endógena.

10 Los enfoques antisentido son un modo de evitar o reducir la función génica dirigiéndose al material genético (Mol *et al.*, *FEBS Lett.* 268: 427-430 (1990)). El objetivo del enfoque antisentido consiste en usar una secuencia complementaria al gen diana para bloquear su expresión y crear una línea celular o un organismo mutante en el que se reduzca o se inhiba selectivamente el nivel de una sola proteína seleccionada. Las técnicas antisentido tienen varias ventajas frente a otros enfoques "genéticos inversos". El sitio de la desactivación y su efecto en el desarrollo se pueden manipular mediante la elección del promotor para los genes antisentido o mediante el cronometraje de la aplicación externa o de la microinyección. El antisentido puede manipular su especificidad seleccionando bien regiones únicas del gen diana o regiones en las que comparta homología con otros genes relacionados (Hiatt *et al.*, en: Genetic Engineering, Setlow (ed.), Vol. 11, Nueva York: Plenum 49-63 (1989)).

20 Las técnicas de ARN antisentido implican la introducción de ARN que sea complementario al ARNm diana en las células, lo que resulta en dúplex de ARN:ARN específicos formados mediante el apareamiento de bases entre el sustrato antisentido y el ARNm diana (Green *et al.*, *Annu. Rev. Biochem.* 55: 569-597 (1986)). En una realización, el procedimiento implica la introducción y la expresión de una secuencia génica antisentido. Tal secuencia es una en la que parte de o todas las secuencias génicas normales están colocadas bajo un promotor en sentido opuesto, tal que la cadena "errónea" o complementaria se transcriba en un ARN antisentido no codificante que se hibrida con el ARNm diana e interfiere en su expresión (Takayama e Inouye, *Crit. Rev. Biochem. Mol. Biol.* 25: 155-184 (1990)). Los vectores antisentido se construyen mediante procedimientos estándar y se introducen en células mediante transformación, transfección, electroporación, microinyección, infección, etc. El tipo de transformación y la elección del vector determinarán si la expresión es transitoria o estable. El promotor usado para el gen antisentido puede influir en el nivel, la duración, el tejido, la especificidad o la inductibilidad de la inhibición antisentido.

30 Se entiende que la actividad de una proteína en una célula vegetal puede ser reducida o bajada mediante el cultivo de una célula vegetal transformada que contenga una molécula de ácido nucleico cuya cadena no transcrita codifique una proteína o un fragmento de la misma.

35 El silenciamiento génico post-transcripcional (PTGS) puede resultar en una inmunidad vírica o un silenciamiento génico en las plantas. El PTGS es inducido por ARNbc y es mediado por una ARN polimerasa dependiente del ARN, presente en el citoplasma, que requiere un molde de ARNbc. El ARNbc se forma mediante la hibridación de varios ARNm transgénicos complementarios o regiones complementarias del mismo transcripto. La formación de los dúplex se puede realizar usando transcriptos de un gen sentido y un gen antisentido colocados en el genoma de la planta, un solo transcripto que tenga auto-complementariedad, o transcriptos sentido y antisentido de genes unidos mediante cruzamiento. La ARN polimerasa dependiente de ARNbc forma una cadena complementaria a partir del ARNm transgénico, y las moléculas de RNasa se unen a esta cadena complementaria (ARNc). Estas moléculas de ARNc-RNasa se hibridan con el ARNm endógeno y escinden el ARN monocatenario adyacente al híbrido. Se siguen degradando los ARN monocatenarios escindidos mediante otras RNasas huésped, porque uno carecerá de un extremo 5' con casquete y el otro carecerá de la cola poli(A) (Waterhouse *et al.*, *PNAS* 95: 13959- 13964 (1998)).

Se entiende que es posible introducir uno o más de los ácidos nucleicos de la invención en una célula vegetal y transcribirlos usando un promotor apropiado con tal transcripción dando como resultado el silenciamiento génico post-traduccional de un transcripto endógeno.

50 Se han expresado anticuerpos en plantas (Hiatt *et al.*, *Nature* 342: 76-78 (1989); Conrad y Fielder, *Plant Mol. Biol.* 26: 1023-1030 (1994)). Se ha informado que la expresión citoplasmática de un Fvmc (anticuerpo Fv monocatenario) retrasa la infección por el virus del rizado moteado de la alcachofa. Las plantas transgénicas que expresan anticuerpos dirigidos frente a proteínas endógenas pueden presentar un efecto fisiológico (Philips *et al.*, *EMBO J.* 16: 4489-4496 (1997); Marion-Poll, "Trends in Plant Science" 2: 447-448 (1997)). Por ejemplo, se ha informado que los anticuerpos anti-abscísicos expresados dan como resultado una perturbación general del desarrollo de la semilla (Philips *et al.*, *EMBO J.* 16: 4489-4496 (1997)).

60 Los anticuerpos que son catalíticos también pueden ser expresados en plantas (abzimas). El principio que hay tras las abzimas es que como los anticuerpos pueden crecer frente a muchas moléculas, se puede dirigir esta capacidad de reconocimiento hacia la generación de anticuerpos que unan estados de transición para forzar el avance una reacción química (Persidas, *Nature Biotechnology* 15: 1313-1315 (1997); Baca *et al.*, *Ann. Rev. Biophys. Biomol. Struct.* 26: 461-493 (1997)). Las capacidades catalíticas de las abzimas se pueden aumentar mediante mutagénesis de sitio dirigida. Los ejemplos de abzimas se exponen, por ejemplo, en la patente estadounidense n.º 5.658.753; la patente estadounidense n.º 5.632.990; la patente estadounidense n.º 5.631.137; la patente estadounidense n.º 5.602.015; la patente estadounidense n.º 5.559.538; la patente estadounidense n.º 5.576.174; la patente estadounidense n.º 5.500.358; la patente estadounidense n.º 5.318.897; la patente estadounidense n.º 5.298.409; patente estadounidense n.º 5.258.289 y la patente estadounidense n.º 5.194.585.

Se entiende que es posible expresar cualquiera de los anticuerpos de la invención en plantas y que tal expresión puede dar como resultado un efecto fisiológico. También se entiende que cualquiera de los anticuerpos expresados puede ser catalítico.

5 Las modificaciones de los fenotipos vegetales de la presente invención pueden ser relativas a una planta que tenga un contexto genético similar que carezca del ácido nucleico introducido de interés. En un aspecto preferido, un contexto genético similar es un contexto en el que los organismos que están siendo comparados comparten el 50% o más de su material genético nuclear. En un aspecto más preferido, un contexto genético similar es un contexto en el que los organismos que están siendo comparados comparten el 75% o más, incluso, más preferiblemente, el 90% o más de su material genético nuclear. En otro aspecto incluso más preferido, un contexto genético similar es un contexto en el que los organismos que están siendo comparados son plantas, y las plantas son isogénicas excepto para cualquier material genético introducido originariamente usando técnicas de transformación de plantas.

15 La presente invención también proporciona partes de las plantas, particularmente, partes reproductoras o de almacenamiento, de la presente invención. Las partes de las plantas incluyen sin limitación semilla, endoesperma, óvulo y polen. En una realización particularmente preferida de la presente invención, la parte de la planta es una semilla. En una realización, la semilla es un constituyente de alimento para animales.

20 En otra realización, la parte de la planta es un fruto, más preferiblemente, un fruto con una estabilidad de almacenamiento mejorada. En otra realización preferida, el fruto tiene mayores niveles de un tocoferol.

La presente invención también proporciona un recipiente de más de 10.000, más preferiblemente, 20.000, e incluso más preferiblemente de 40.000 semillas en las que más del 10%, más preferiblemente, del 25%, más preferiblemente, del 50% e incluso más preferiblemente, del 75% o del 90% de las semillas son semillas derivadas de una planta de la presente invención.

30 La presente invención también proporciona un recipiente de más de 10 kg, más preferiblemente, de 25 kg, e incluso más preferiblemente de 50 kg de semillas en las que más del 10%, más preferiblemente, del 25%, más preferiblemente, del 50% e incluso más preferiblemente, del 75% o del 90% de las semillas son semillas derivadas de una planta de la presente invención.

Cualquiera de las plantas o de las partes de las mismas de la presente invención puede ser procesada para producir una preparación alimentaria, de harina, proteica u oleaginosa. Una parte de planta particularmente preferida a tal efecto es una semilla. En una realización preferida, la preparación alimentaria, de harina, proteica u oleaginosa está diseñada para animales rumiantes. Los procedimientos para producir preparaciones alimentarias, de harina, proteicas u oleaginosas son conocidos en la técnica. Véanse, por ejemplo, las patentes estadounidenses n.º 4.957.748; 5.100.679; 5.219.596; 5.936.069; 6.005.076; 6.146.669 y 6.156.227. En una realización preferida, la preparación proteica es una preparación con un alto contenido en proteínas. Tal preparación con alto contenido en proteínas tiene preferiblemente un contenido de proteínas mayor del 5% p/v, más preferiblemente, del 10% p/v e incluso más preferiblemente, del 15% p/v. En una preparación oleaginosa preferida, la preparación oleaginosa es una preparación con alto contenido en aceite derivado de una planta o de una parte de la misma de la presente invención de más del 5% p/v, más preferiblemente, del 10% p/v e incluso más preferiblemente, del 15% p/v. En una realización preferida, la preparación oleaginosa es un líquido y de un volumen mayor de 1, 5, 10 ó 50 litros. La presente invención también proporciona aceite producido a partir de plantas de la presente invención o generado mediante un procedimiento de la presente invención. Tal aceite puede ser un componente minoritario o mayoritario de cualquier producto resultante. Además, tal aceite puede ser mezclado con otros aceites. En una realización preferida, el aceite producido a partir de plantas de la presente invención o generado mediante un procedimiento de la presente invención constituye más del 0,5%, 1%, 5%, 10%, 25%, 50%, 75% o 90% en volumen o peso del componente oleaginoso de cualquier producto. En otra realización, se puede mezclar la preparación oleaginosa y puede constituir más del 10%, 25%, 35%, 50% o 75% de la mezcla en volumen. El aceite producido a partir de una planta de la presente invención se puede mezclar con uno o más disolventes orgánicos o destilados del petróleo.

55 Las plantas de la presente invención pueden ser parte de o ser generadas a partir de un programa de mejora genética. La elección del procedimiento de mejora depende del modo de reproducción de la planta, de la heredabilidad del carácter o de los caracteres que se estén mejorando y del tipo de cultivar usado comercialmente (p. ej., cultivar de híbrido F₁, cultivar de purelina, etc.). A continuación, se exponen los enfoques no restrictivos seleccionados para la mejora genética de las plantas de la presente invención. Se puede perfeccionar un programa de mejora usando una selección con marcadores de la progenie de cualquier cruzamiento. Se entiende además que se puede utilizar cualquier cultivar comercial o no comercial en un programa de mejora genética. Factores tales como, por ejemplo, el vigor de los brotes, el vigor vegetativo, la tolerancia al estrés, la resistencia a enfermedades, la ramificación, la floración, la colección de semillas, el tamaño de la semilla, la densidad de las semillas, la capacidad de permanencia y la capacidad de trillado, etc., serán los que rijan en general la elección.

65 Para caracteres altamente hereditarios, la elección de plantas individuales superiores evaluadas en una sola ubicación será eficaz, mientras que para caracteres con baja heredabilidad, la selección debería estar basada en valores medios obtenidos a partir de evaluaciones por duplicado de familias de plantas relacionadas. Los procedimientos de selección comunes incluyen la selección genealógica, la selección genealógica modificada, la selección de la masa y la

selección recurrente. En una realización preferida, se lleva a cabo un retrocruzamiento o programa de mejora genética recurrente.

La complejidad de la herencia influye en la elección del procedimiento de mejora genética. La mejora por retrocruzamiento se puede usar para transferir uno o unos cuantos genes favorables para un carácter altamente hereditario en un cultivar deseable. Este enfoque se ha usado ampliamente para la selección de cultivares resistentes a enfermedades. Se usan diversas técnicas de selección recurrente para mejorar los caracteres heredados cuantitativamente controlados por numerosos genes. El uso de la selección recurrente en cultivos auto-polinizantes depende de la facilidad de la polinización, de la frecuencia de los híbridos satisfactorios procedentes de cada polinización y del número de descendientes híbridos de cada cruce satisfactorio.

Las líneas de selección se pueden analizar y comparar con patrones apropiados en entornos representativos de la zona o las zonas diana comerciales para dos o más generaciones. Las mejores líneas son candidatos para nuevos cultivares comerciales; aquellas todavía deficientes en caracteres se pueden usar como padres para producir nuevas poblaciones para una selección posterior.

Un procedimiento para identificar una planta superior consiste en observar su rendimiento en comparación con otras plantas experimentales y con un cultivar estándar extensamente cultivado. Si hay una sola observación que no es concluyente, las observaciones por duplicado pueden proporcionar una mejor estimación de su valor genético. Un cultivador puede seleccionar y cruzar dos o más líneas parentales, siguiendo con la autofecundación y la selección repetidas, produciendo muchas nuevas combinaciones genéticas.

El desarrollo de nuevos cultivares requiere el desarrollo y la selección de variedades, el cruzamiento de estas variedades y la selección de cruces híbridos superiores. La semilla híbrida se puede producir mediante cruces manuales entre padres fértiles varones seleccionados o usando sistemas de esterilidad de varones. Los híbridos se seleccionan para ciertos caracteres de un solo gen tales como el color de la vaina, el color de la flor, el rendimiento de las semillas, el color de las pubescencias o la resistencia a herbicidas, que indican que la semilla es verdaderamente un híbrido. Hay más datos sobre las líneas parentales, así como el fenotipo del híbrido, que influyen en la decisión del cultivador sobre si continuar con el cruzamiento de un híbrido específico.

Los procedimientos de selección genealógica y de mejora por selección recurrente se pueden usar para desarrollar cultivares a partir de las poblaciones de mejora genética. Los programas de mejora combinan los caracteres deseables de dos o más cultivares o varias fuentes de características comunes en mezclas de selección a partir de las cuales se desarrollan los cultivares mediante la autofecundación y la selección de fenotipos deseados. Se pueden evaluar nuevos cultivares para determinar cuál tiene potencial comercial.

Los sistemas de selección genealógica se usan comúnmente para la mejora de cultivos auto-polinizantes. Dos padres que poseen caracteres favorables complementarios se cruzan para producir un F_1 . Se produce una población de F_2 mediante la autofecundación de uno o más F_1 . Se lleva a cabo la selección de los mejores individuos de las mejores familias. El análisis por duplicado de las familias puede comenzar en la generación F_4 para mejorar la eficacia de la selección de los caracteres con una baja heredabilidad. En una etapa avanzada de la autopolinización (i.e., F_6 y F_7), se analizan las mejores líneas o mezclas de líneas fenotípicamente similares para una posible liberación como nuevos cultivares.

La mejora por retrocruzamiento se ha usado para transferir genes para un carácter altamente heredable heredado simplemente a una línea consanguínea o un cultivar homocigótico deseable, que es el padre recurrente. El origen del carácter por ser transferido se denomina padre donante. Se espera que la planta resultante tenga los atributos del padre recurrente (p. ej., el cultivar) y el carácter deseable transferido del padre donante. Tras el cruce inicial, se seleccionan los individuos que poseen el fenotipo del padre donante y se cruzan repetidamente (retrocruzan) con el padre recurrente. Se espera que el padre resultante tenga los atributos del padre recurrente (p. ej., el cultivar) y el carácter deseable transferido del padre donante.

El procedimiento de descendencia de una sola semilla se refiere en sentido estricto a la plantación de una población segregante, la cosecha de una muestra de una semilla por planta y el uso de la muestra de una sola semilla para plantar la siguiente generación. Cuando la población se ha producido a partir de F_2 hasta el nivel deseado de autopolinización, las plantas de cuyas líneas son derivadas seguirán cada una hasta individuos F_2 diferentes. El número de plantas de una población disminuye cada generación debido a que algunas semillas no germinan o que algunas plantas no producen al menos una semilla. Por consiguiente, no todas las plantas F_2 muestreadas originariamente en la población estarán representadas por una progenie una vez completado el avance de la generación.

En un procedimiento de múltiples semillas, los cultivadores comúnmente cosechan una o más vainas de cada planta de una población y las trillan conjuntamente para agruparlas. Parte del grupo se usa para plantar la siguiente generación y la otra parte se reserva. El procedimiento ha sido denominado técnica de descendencia de una sola semilla modificada o de agrupamiento de vainas.

El procedimiento de múltiples semillas se ha usado para ahorrar mano de obra en la cosecha. Es considerablemente más rápido trillar vainas con una máquina que retirar una semilla de cada una a mano para el procedimiento de una

sola semilla. El procedimiento de múltiples semillas también permite plantar el mismo número de semillas de una población de cada generación de la autopolinización.

Se pueden encontrar descripciones de otros procedimientos de mejora genética que son comúnmente usados para diferentes caracteres y cultivos en uno de varios libros de referencia (e. g. Fehr, "Principles of Cultivar Development" Vol. 1, pp. 2-3 (1987)).

También es posible reproducir una planta transgénica de la presente invención usando la apomixis. La apomixis es un procedimiento controlado genéticamente de reproducción en plantas en el que el embrión se forma sin la unión de un huevo y un esperma. Hay tres tipos básicos de reproducción apomíctica: 1) la aposporia, en la que el embrión se desarrolla a partir de un huevo no reducido cromosómicamente en un saco embrionario derivado del núcleo, 2) diplosporia, en la que el embrión se desarrolla a partir de un huevo sin reducir en un saco embrionario derivado de la célula madre de la macrospora y 3) embrionía adventicia, en la que el embrión se desarrolla directamente a partir de una célula somática. En la mayoría de las formas de apomixis, es necesaria la pseudogamia o la fertilización de los núcleos polares para producir endoesperma para conseguir una viabilidad de las semillas. En la aposporia, se puede usar un cultivar de cría como fuente de polen para la formación de endoesperma en las semillas. El cultivar de cría no afecta a la genética del cultivar apomíctico aposporios, pues el huevo sin reducir del cultivar se desarrolla partenogenéticamente, pero hace posible la producción de endoesperma. La apomixis es importante económicamente, especialmente en plantas transgénicas, porque provoca la verdadera mejora de cualquier genotipo, independientemente de lo heterocigótico que sea. De este modo, con la reproducción apomíctica, las plantas transgénicas heterocigóticas pueden mantener su fidelidad genética a lo largo de ciclos de vida repetidos. Los procedimientos para producir plantas apomícticas son conocidos en la técnica. Véase la patente estadounidense n.º 5.811.636.

Otros organismos

Es posible introducir un ácido nucleico de la presente invención en cualquier célula u organismo tal como una célula de mamífero, un mamífero, una célula de pez, un pez, una célula de ave, un ave, una célula de alga, un alga, una célula de hongo, un hongo o una célula bacteriana. Una proteína de la presente invención se puede producir en una célula o en un organismo apropiado. El huésped y los transformantes preferidos incluyen: células micóticas, telas como *Aspergillus*, levaduras, mamíferos, particularmente, bovino y porcino, insectos, bacterias y algas. En la técnica, se conocen procedimientos para transformar tales células u organismos (EP 0 238 023; Yelton *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.), 81: 1470-1474 (1984); Malardier *et al.*, *Gene*, 78: 147-156 (1989); Becker y Guarente, en: Abelson y Simon (eds.), "Guide to Yeast Genetics and Molecular Biology", *Method Enzymol.*, Vol. 194, pp. 182-187, Academic Press, Inc., Nueva York; Ito *et al.*, *J. Bacteriology*, 153: 163 (1983) Hinnen *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.), 75: 1920 (1978); Bennett y LaSure (eds.), "More Gene Manipulations in fungi", Academic Press, CA (1991)). También se conocen procedimientos para producir proteínas de la presente invención (Kudla *et al.*, *EMBO*, 9: 1355-1364 (1990); Jarai y Buxton, *Current Genetics*, 26: 2238-2244 (1994); Verdier, *Yeast*, 6: 271-297 (1990); MacKenzie *et al.*, *Journal of Gen. Microbiol.*, 139: 2295-2307 (1993); Hartl *et al.*, *TIBS*, 19: 20-25 (1994); Bergenron *et al.*, *TIBS*, 19: 124-128 (1994); Demolder *et al.*, *J. Biotechnology*, 32: 179-189 (1994); Craig, *Science*, 260: 1902-1903 (1993); Gething y Sambrook, *Nature*, 355: 33-45 (1992); Puig y Gilbert, *J. Biol. Chem.*, 269: 7764-7771 (1994); Wang y Tsou, *FASEB Journal*, 7: 1515-1517 (1993); Robinson *et al.*, *Bio/Technology*, 1: 381-384 (1994); Enderlin y Ogrydziak, *Yeast*, 10: 67-79 (1994); Fuller *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.), 86: 1434-1438 (1989); Julius *et al.*, *Cell*, 37: 1075-1089 (1984); Julius *et al.*, *Cell* 32: 839-852 (1983).

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de tocotrienoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de tocoferoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de α -tocotrienoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de γ -tocotrienoles.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de ácido homogentísico.

En una realización preferida, la sobre-expresión de una proteína de la presente invención en una célula o un organismo proporciona en esa célula o ese organismo, en relación con una célula o un organismo con un contexto genético similar, un mayor nivel de plastoquinoles o plastoquinonas.

Anticuerpos

Un aspecto de la invención se refiere a anticuerpos, moléculas de unión antigénica monocatenarias, u otras proteínas, que se unen específicamente a una o más de las moléculas proteicas o peptídicas de la invención. En una realización, el anticuerpo se une específicamente a una proteína que tiene la secuencia de aminoácidos expuesta en las SEC ID N.º 2 y 4. En otra realización, el anticuerpo se une específicamente a una proteína de fusión que comprende una secuencia de aminoácidos seleccionada de la secuencia de aminoácidos expuesta en las SEC ID N.º 2 y 4. En otra realización, el anticuerpo se une específicamente a una proteína de fusión que comprende una secuencia de aminoácidos seleccionada de la secuencia de aminoácidos expuesta en las SEC ID N.º 2 y 4. Los anticuerpos de la invención se pueden usar para detectar cuantitativa o cualitativamente las moléculas proteicas o peptídicas de la invención, o para detectar las modificaciones post-traduccionales de las proteínas. Como se usa en la presente memoria, se dice que un anticuerpo o un péptido “se une específicamente” a una molécula proteica o peptídica de la invención si tal unión no es inhibida competitivamente por la presencia de moléculas no relacionadas.

Las moléculas de ácido nucleico que codifican toda o parte de la proteína de la invención se pueden expresar, mediante procedimientos recombinantes, para producir la proteína o los péptidos que puedan a su vez ser usados para provocar que los anticuerpos sean capaces de unirse a la proteína o el péptido expresado. Tales anticuerpos se pueden usar en inmunoanálisis para esa proteína. Tales moléculas codificantes de proteínas, o sus fragmentos, pueden ser una molécula “de fusión” (i.e., una parte de una molécula de ácido nucleico mayor), tal que, tras la expresión, se produzca una proteína de fusión. Se entiende que cualquiera de las moléculas de ácido nucleico de la invención puede ser expresada, mediante procedimientos recombinantes, para generar proteínas o péptidos codificados por estas moléculas de ácido nucleico.

Los anticuerpos que se unen específicamente a proteínas y a fragmentos de proteínas de la invención pueden ser policlonales o monoclonales, y pueden comprender inmunoglobulinas intactas o partes de unión antigénica de fragmentos de inmunoglobulinas (tales como F(ab'), F(ab')₂, o inmunoglobulinas monocatenarias producibles, por ejemplo, mediante procedimientos recombinantes. Se entiende que los profesionales están familiarizados con los materiales de información estándar que describen condiciones específicas y procedimientos para la construcción, la manipulación y el aislamiento de anticuerpos (véase, por ejemplo, Harlow y Lane, en: “Antibodies: A Laboratory Manual”, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1988)).

Según lo tratado más adelante, tales moléculas de anticuerpo o sus fragmentos se pueden usar a efectos de diagnóstico. Cuando se pretenden emplear los anticuerpos a efectos de diagnóstico, puede ser deseable derivarlos, por ejemplo, con un grupo ligando (tal como biotina) o un grupo marcador detectable (tal como un grupo fluorescente, un radioisótopo o una enzima).

La capacidad para producir anticuerpos que se unan a moléculas proteicas o peptídicas de la invención permite la identificación de compuestos miméticos derivados de aquellas moléculas. Estos compuestos miméticos pueden contener un fragmento de la proteína o del péptido, o simplemente, una región estructuralmente similar, presentando, no obstante, una capacidad para unirse específicamente a anticuerpos dirigidos frente a ese compuesto.

Usos ejemplares

Las moléculas de ácido nucleico de la invención se pueden emplear para obtener otras moléculas de ácido nucleico de la misma especie (las moléculas de ácido nucleico del maíz se pueden utilizar para obtener otras moléculas de ácido nucleico del maíz). Tales moléculas de ácido nucleico incluyen moléculas de ácido nucleico que codifican la secuencia codificante completa de una proteína y promotores y secuencias flanqueantes de tales moléculas. Además, tales moléculas de ácido nucleico incluyen moléculas de ácido nucleico que codifican otras isozimas o miembros de la familia de genes. Tales moléculas se pueden obtener fácilmente usando las moléculas de ácido nucleico anteriormente descritas o fragmentos de las mismas para rastrear ADNc o genotecas genómicas. Los procedimientos para formar tales genotecas son conocidos en la técnica.

Las moléculas de ácido nucleico de la invención también se pueden emplear para obtener homólogos de ácido nucleico. Tales homólogos incluyen las moléculas de ácido nucleico de plantas y otros organismos, incluyendo bacterias y hongos, incluyendo moléculas de ácido nucleico que codifiquen, totalmente o en parte, homólogos proteicos de otras especies vegetales o de otros organismos, secuencias de elementos genéticos, tales como promotores y elementos reguladores transcripcionales. Tales moléculas se pueden obtener fácilmente usando las moléculas de ácido nucleico anteriormente descritas o fragmentos de las mismas para rastrear ADNc o genotecas genómicas obtenidas de tales especies vegetales. Los procedimientos para formar tales genotecas son conocidos en la técnica. Tales moléculas homólogas pueden diferir en sus secuencias de nucleótidos de aquellas encontradas en una o más de las SEC ID N.º 1 y 3, y los complementos de las mismas, porque no es necesaria una complementariedad total para lograr una hibridación estable.

Se puede usar cualquiera de una variedad de procedimientos para obtener una o más de las moléculas de ácido nucleico anteriormente descritas (Zamechik *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 83: 4143-4146 (1986); Goodchild *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 85: 5507-5511 (1988); Wickstrom *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 85: 1028-1032 (1988); Holt *et al.*, *Molec. Cell. Biol.* 8: 963-973 (1988); Gerwitz *et al.*, *Science* 242: 1303-1306 (1988); Anfossi *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 86: 3379-3383 (1989); Becker *et al.*, *EMBO J.* 8: 3685-3691 (1989)).

Se pueden emplear a tal efecto sintetizadores de ácidos nucleicos automatizados. En lugar de tal síntesis, se pueden usar las moléculas de ácido nucleico reveladas para definir un par de cebadores que se puedan usar con la reacción en cadena de la polimerasa (Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.* 51: 263-273 (1986); Erlich *et al.*, Patente europea 50.424; Patente europea 84.796; Patente europea n.º 258.017; Patente europea n.º 237.362; Mullis, Patente europea n.º 201.184; Mullis *et al.*, patente estadounidense n.º 4.683.202; Erlich, patente estadounidense n.º 4.582.788; y Saiki *et al.*, patente estadounidense n.º 4.683.194) para amplificar y obtener cualquier molécula de ácido nucleico o fragmento deseados.

También se pueden obtener secuencias promotoras y otros elementos genéticos, incluyendo, pero no limitándose a, secuencias flanqueantes reguladoras transcripcionales, asociadas con una o más de las secuencias de ácidos nucleico reveladas usando la secuencia de ácido nucleico revelada proporcionada en la presente memoria. En una realización, tales secuencias se obtienen incubando moléculas de ácido nucleico de la presente invención con miembros de genotecas genómicas y recuperando los clones que se hibriden a tales moléculas de ácido nucleico de las mismas. En una segunda realización, se pueden usar procedimientos de “paseo cromosómico” o PCR inversa para obtener tales secuencias (Frohman *et al.*, *Proc. Natl. Acad. Sci. (EE.UU.)* 85: 8998-9002 (1988); Ohara *et al.*, *Proc. Natl. Acad. Sci. (EE.UU.)* 86: 5673-5677 (1989); Pang *et al.*, *Biotechniques* 22: 1046-1048 (1977); Huang *et al.*, *Methods Mol. Biol.* 69: 89-96 (1997); Huang *et al.*, *Method Mol. Biol.* 67:287-294 (1997); Benkel *et al.*, *Genet. Anal.* 13: 123-127 (1996); Hartl *et al.*, *Methods Mol. Biol.* 58: 293-301 (1996)). El término “paseo cromosómico” significa un procedimiento de ampliación de un mapa genético mediante sucesivas etapas de hibridación.

Se pueden usar las moléculas de ácido nucleico de la invención para aislar promotores de perfiles de expresión celulares mejorados, de una célula específica, titulares mejorados, de un tejido específico, evolutivamente o medioambientalmente regulados. El aislamiento y el análisis funcional de las secuencias promotoras flanqueantes en 5' de estos genes procedentes de genotecas genómicas, por ejemplo, usando procedimientos de detección genómica y técnicas de PCR darían como resultado en el aislamiento de promotores útiles y elementos reguladores transcripcionales. Estos procedimientos son conocidos por los expertos en la técnica y han sido descritos (véase, por ejemplo, Birren *et al.*, “Genome Analysis: Analyzing DNA”, 1, (1997), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.). Los promotores obtenidos utilizando las moléculas de ácido nucleico de la invención también podrían ser modificados para afectar a sus características de control. Los ejemplos de tales modificaciones incluyen, pero no se limitan a, secuencias potenciadoras. Tales elementos genéticos podrían ser usados para aumentar la expresión génica de caracteres nuevos y existentes para una mejora de cosechas.

Otro subconjunto de moléculas de ácido nucleico de la invención incluye moléculas de ácido nucleico que son marcadores. Los marcadores se pueden usar en un número de modos convencionales en el campo de la genética molecular. Tales marcadores incluyen moléculas de ácido nucleico de SEC ID N.º 1 y 3, complementos de las mismas, y fragmentos de cualquiera que pueda actuar como marcadores y otras moléculas de ácido nucleico de la presente invención que pueden actuar como marcadores.

Los marcadores genéticos de la invención incluyen marcadores “dominantes” o “codominantes”. Los “marcadores codominantes” revelan la presencia de dos o más alelos (dos por individuo diploide) en un locus. Los “marcadores dominantes” sólo revelan la presencia de un solo alelo por locus. La presencia del fenotipo del marcador dominante (p.ej., una banda de ADN) es una señal de que un alelo está en una condición bien homocigótica o heterocigótica. La ausencia del fenotipo del marcador dominante (p.ej., la ausencia de una banda de ADN) simplemente es la prueba de que hay “algún otro” alelo indefinido presente. En el caso de las poblaciones en las que los individuos son predominantemente homocigóticos y los locus son predominantemente dimórficos, los marcadores dominantes y codominantes pueden ser igualmente valiosos. A medida que las poblaciones se vuelven más heterocigóticas y multi-alelicas, los marcadores codominantes se suelen volver más informativos del genotipo que los marcadores dominantes. Las moléculas de marcadores pueden ser, por ejemplo, capaces de detectar polimorfismos tales como polimorfismos de un solo nucleótido (SNP).

Los genomas de los animales y las plantas se someten de manera natural a una mutación espontánea en el transcurso de su continua evolución (Gusella, *Ann. Rev. Biochem.* 55: 831-854 (1986)). Un “polimorfismo” es una variación o una diferencia en la secuencia del gen o de sus regiones flanqueantes que surge en alguno de los miembros de la especie. La secuencia divergente y la secuencia “original” coexisten en la población de la especie. En algunos casos, tal coexistencia está en equilibrio estable o casi-estable.

Por lo tanto, se dice que un polimorfismo es “alélico” cuando, debido a la existencia del polimorfismo, algunos miembros de una población pueden tener la secuencia original (i.e., el “alelo” original), mientras que otros miembros pueden tener la secuencia divergente (i.e., el “alelo” divergente). En el caso más simple, sólo puede existir una secuencia divergente y se dice, por tanto, que el polimorfismo es di-alelico. En otros casos, la población de la especie puede contener múltiples alelos y el polimorfismo se denomina tri-alelico, etc. Un solo gen puede tener múltiples polimorfismos no relacionados diferentes. Por ejemplo, puede tener un polimorfismo di-alelico en un sitio y un polimorfismo multi-alelico en otro sitio.

La variación que define el polimorfismo puede variar de la variación de un solo nucleótido a la inserción o la eliminación de regiones prolongadas dentro de un gen. En algunos casos, las variaciones de las secuencias de ADN están en regiones del genoma que se caracterizan por repeticiones en tándem cortas (STR) que incluyen motivos repetidos di- o trinucleotídicos en tándem de nucleótidos. Los polimorfismos caracterizados por tales repeticiones en

tándem se denominan polimorfismos de “repeticiones en tándem de un número variable” (VNTR). Las VNTR se han usado en análisis de la identidad (Weber, patente estadounidense n.º 5.075.217; Armour *et al.*, *FEBS Lett.* 307: 113-115 (1992); Jones *et al.*, *Eur. J. Haematol.* 39: 144-147 (1987); Horn *et al.*, solicitud de patente vía PCT W091/14003; Jeffreys, solicitud de patente europea 370.719; Jeffreys, patente estadounidense n.º 5.175.082; Jeffreys *et al.*, *Amer. J. Hum. Genet.* 39: 11-24 (1986); Jeffreys *et al.*, *Nature* 316: 76-79 (1985); Gray *et al.*, *Proc. R. Acad. Soc. Lond.* 243: 241-253 (1991); Moore *et al.*, *Genomics* 10: 654-660 (1991); Jeffreys *et al.*, *Anim. Genet.* 18: 1-15 (1987); Hillel *et al.*, *Anim. Genet.* 20: 145-155 (1989); Hillel *et al.*, *Genet.* 124: 783-789 (1990)).

La detección de sitios polimórficos en una muestra de ADN puede ser facilitada por el uso de procedimientos de amplificación ácidos nucleicos. Tales procedimientos aumentan específicamente la concentración de polinucleótidos que abarcan el sitio polimórfico o incluyen ese sitio y las secuencias localizadas bien distal o proximalmente al mismo. Tales moléculas amplificadas pueden ser fácilmente detectadas mediante electroforesis sobre gel u otros procedimientos.

En una realización alternativa, tales polimorfismos se pueden detectar a través del uso de una molécula de ácido nucleico marcadora que esté ligada físicamente a tal o tales polimorfismos. A tal efecto, se pueden emplear moléculas de ácido nucleico marcadoras que comprenden una secuencia de nucleótidos de un polinucleótido localizado en 1 mb del o de los polimorfismos y, más preferiblemente, en 100kb del o de los polimorfismos, y lo más preferible, en 10kb del o de los polimorfismos.

Se puede determinar la identificación de un polimorfismo en una variedad de modos. Correlacionando la presencia o la ausencia del mismo en una planta con la presencia o la ausencia de un fenotipo, es posible predecir el fenotipo de esa planta. Si un polimorfismo crea o destruye un sitio de escisión de una endonucleasa de restricción, o si da como resultado la pérdida o la inserción de ADN (p. ej., un polimorfismo de VNTR), alterará el tamaño o el perfil de los fragmentos de ADN que son generados mediante la digestión con esa endonucleasa de restricción. Como tales, los organismos que poseen una secuencia divergente se pueden distinguir de aquéllos que tienen la secuencia original mediante el análisis de los fragmentos de restricción. Los polimorfismos que se pueden identificar de esta manera se denominan “polimorfismos de la longitud de los fragmentos de restricción” (“RFLPs”) (Glassberg, solicitud de patente británica 2135774; Skolnick *et al.*, *Cytogen. Cell Genet.* 32: 58-67 (1982); Botstein *et al.*, *Am. J. Hum. Genet.* 32: 314-331 (1980); Fischer *et al.*, (solicitud vía PCT W090/13668; Uhlen, solicitud vía PCT W090/11369).

Los polimorfismos también se pueden identificar mediante análisis de polimorfismos de configuración monocatenaria (SSCP) (Elles, “Methods in a Molecular Medicine: Molecular Diagnosis of Genetic Diseases”, Humana Press (1996)); Orita *et al.*, *Genomics* 5: 874-879 (1989)). Se ha descrito un número de protocolos para el SSCP incluyendo, pero no limitándose a, Lee *et al.*, *Anal. Biochem.* 205: 289-293 (1992); Suzuki *et al.*, *Anal. Biochem.* 192: 82-84 (1991); Lo *et al.*, “Nucleic Acids Research” 20: 1005-1009 (1992); Sarkar *et al.*, *Genomics* 13: 441-443 (1992). Se entiende que se pueden utilizar uno o más de los ácidos nucleicos de la invención como marcadores o sondas para detectar los polimorfismos mediante el análisis de SSCP.

También se pueden encontrar polimorfismos usando una técnica de análisis de la huella del ADN denominada polimorfismo de la longitud de fragmentos amplificados (AFLP), que se basa en la amplificación por PCR selectiva de fragmentos de restricción procedentes de un digesto total de ADN genómico para perfilar ese ADN (Vos *et al.*, *Nucleic Acids Res.* 23: 4407-4414 (1995)). Este procedimiento permite la co-amplificación específica de elevados números de fragmentos de restricción, que pueden ser visualizados por PCR sin conocer la secuencia de ácido nucleico. Se entiende que se pueden utilizar uno o más de los ácidos nucleicos de la invención como marcadores o sondas para detectar los polimorfismos mediante el análisis de polimorfismos AFLP o para la huella del ARN.

También se pueden encontrar polimorfismos usando ADN polimórfico amplificado aleatoriamente (RAPD) (Williams *et al.*, *Nucl. Acids Res.* 18: 6531-6535 (1990)) y secuencias polimórficas amplificadas escindibles (CAPS) (Lyamichev *et al.*, *Science* 260: 778-783 (1993)). Se entiende que se pueden utilizar una o más de las moléculas de ácido nucleico de la invención como marcadores o sondas para detectar los polimorfismos mediante el análisis de RAPD o CAPS.

Los polimorfismos de de un solo nucleótido (SNP) tienen lugar generalmente a una mayor frecuencia que otros marcadores polimórficos y están espaciados con una mayor uniformidad por un genoma que otras formas descritas de polimorfismo. La mayor frecuencia y uniformidad de los SNP significa que hay una mayor probabilidad de que tal polimorfismo sea encontrado cerca o en un locus genético de interés que la que habría en el caso de otros polimorfismos. Los SNP se localizan en regiones codificantes de proteínas y en regiones no codificantes de un genoma. Algunos de estos SNP pueden dar como resultado una expresión de proteínas defectuosa o divergente (p. ej., como resultado de mutaciones o un corte y empalme defectuoso). El análisis (el genotipado) de SNP caracterizados puede requerir sólo un análisis más o menos en lugar de una medición larga y pesada, lo que facilita la automatización.

Los SNP se pueden caracterizar usando cualquiera de una variedad de procedimientos. Tales procedimientos incluyen la secuenciación directa o indirecta del sitio, el uso de enzimas de restricción (Botstein *et al.*, *Am. J. Hum. Genet.* 32: 314-331 (1980); Konieczny y Ausubel, *Plant J.* 4: 403-410 (1993)), análisis enzimáticos y químicos de apareamientos erróneos (Myers *et al.*, *Nature* 313: 495-498 (1985)), PCR de alelo específico (Newton *et al.*, *Nucl. Acids Res.* 17: 2503-2516 (1989); Wu *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 86: 2757-2760 (1989)), la reacción en cadena de la ligasa (Barany, *Proc. Natl. Acad. Sci. EE.UU.* 88: 189-193 (1991)), el análisis de polimorfismos de con-

figuración monocatenaria (Labruno *et al.*, *Am. J. Hum. Genet.* 48: 1115-1120 (1991)), la prolongación de cebadores de una sola base (Kuppuswamy *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 88: 1143-1147 (1991), Goelet US 6.004.744; Goelet 5.888.819), el análisis de unión de oligonucleótidos basados en ELISA en fase sólida (Nikiforov *et al.*, *Nucl. Acids Res.* 22: 4167-4175 (1994), la huella dideoxi (Sarkar *et al.*, *Genomics* 13: 441-443 (1992)), el análisis de detención de la fluorescencia de oligonucleótidos (Livak *et al.*, *PCR Methods Appl.* 4: 357-362 (1995a)), el análisis de TaqMan® de hibridación de alelo específico de la nucleasa de 5' (Livak *et al.*, *Nature Genet.* 9: 341-342 (1995)), el análisis de incorporación de terminadores con colorantes dirigido a un molde (TDI) (Chen y Kwok, *Nucl. Acids Res.* 25: 347-353 (1997)), el análisis molecular beacon de alelo específico (Tyagi *et al.*, *Nature Biotech.* 16: 49-53 (1998)), el análisis PinPoint (Haff y Smirnov, *Genome Res.* 7: 378-388 (1997)), el análisis dCAPS (Neff *et al.*, *Plant J.* 14: 387-392 (1998)), la pirosecuenciación (Ronaghi *et al.*, *Analytical Biochemistry* 267: 65-71 (1999); Ronaghi *et al.* solicitud vía PCT WO 98/13523; Nyren *et al.* solicitud vía PCT WO 98/28440; <http://www.pyrosequencing.com>), usando una espectrometría de masas, p.ej., el sistema Masscode™ (Howbert *et al.* WO 99/05319; Howber *et al.* WO 97/27331; <http://www.rapigene.com>; Becker *et al.* solicitud vía PCT WO 98/26095; Becker *et al.* solicitud vía PCT; WO 98/12355; Becker *et al.* solicitud vía PCT WO 97/33000; Monforte *et al.* US 5.965.363), escisión invasiva de sondas de oligonucleótido (Lyamichev *et al.* *Nature Biotechnology* 17: 292-296; <http://www.twt.com>), y usando alineamientos de oligonucleótidos de alta densidad (Hacia *et al.* *Nature Genetics* 22: 164-167; <http://www.affymetrix.com>).

Los polimorfismos también se pueden detectar usando oligonucleótidos de alelo específico (ASO) que se pueden usar, por ejemplo, en combinación con una tecnología basada en la hibridación incluyendo hibridaciones de transferencia southern, northern y de puntos, hibridaciones de transferencia de puntos inversa e hibridaciones realizadas sobre microalineamientos y la tecnología relacionada.

Las condiciones restrictivas de la hibridación para la detección de polimorfismos son muy dependientes de una variedad de factores, incluyendo la longitud del oligonucleótido de alelo específico, la composición secuencial, el grado de complementariedad (i.e., la presencia o la ausencia de apareamientos erróneos de bases), la concentración de sales y otros factores tales como la formamida y la temperatura. Estos factores son importantes tanto durante la propia hibridación como durante los posteriores lavados realizados para eliminar el polinucleótido diana que no se haya hibridado específicamente. En la práctica, las condiciones del lavado más restrictivo final son las más fundamentales. Además, la cantidad de polinucleótido diana que es capaz de hibridarse al oligonucleótido de alelo específico también está gobernada por factores tales como la concentración de tanto el ASO como el polinucleótido diana, la presencia y la concentración de factores que actúan para "amarrar" las moléculas de agua, para concentrar eficazmente los reactivos (p.ej., PEG, dextrano, sulfato de dextrano, etc.) bien si los ácidos nucleicos están inmovilizados o en disolución, y la duración de la hibridación y las etapas de lavado.

Las hibridaciones se realizan preferiblemente por debajo de la temperatura de fusión (T_f) del ASO. Cuanto más se acerca la hibridación y/o la etapa de lavado a la T_f , más restrictivas serán las condiciones. La T_f de un oligonucleótido puede ser determinada, por ejemplo, según la siguiente fórmula: $T_f = 81,5 + 16,6 \times (\log_{10} [\text{Na}^+]) + 0,41 \times (\%G+C) - 675/n$; en la que $[\text{Na}^+]$ es la concentración molar de sales de Na^+ o de cualquier otro catión adecuado y n = número de bases del oligonucleótido. Hay otras fórmulas para determinar la T_f disponibles y conocidas por aquellos expertos habituales en la técnica.

Preferiblemente, se ajustan las condiciones restrictivas para permitir que un determinado ASO se hibride diferencialmente con un polinucleótido diana del alelo correcto y un polinucleótido diana del alelo incorrecto. Preferiblemente, habrá al menos un diferencial del doble entre la señal producida por el ASO en hibridación con un polinucleótido diana del alelo correcto y el nivel de la señal producida mediante la hibridación cruzada del ASO con un polinucleótido diana del alelo incorrecto (p. ej., un ASO específico para un alelo mutante en hibridación cruzada con un alelo de tipo natural). En realizaciones más preferidas de la presente invención, hay al menos un diferencial cinco veces mayor de la señal. En realizaciones muy preferidas de la presente invención, hay al menos un diferencial de la señal de un orden de magnitud entre el ASO en hibridación con un polinucleótido diana del alelo correcto y el nivel de la señal producida mediante la hibridación cruzada del ASO con un polinucleótido diana del alelo incorrecto.

Aunque en la presente memoria se describen ciertos procedimientos para la detección de polimorfismos, se pueden utilizar otras metodologías de detección. Por ejemplo, hay más metodologías conocidas y se exponen en Birren *et al.*, *Genome Analysis*, 4: 135-186, "A Laboratory Manual. Mapping Genomes", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1999); Maliga *et al.*, "Methods in Plant Molecular Biology. A Laboratory Course Manual", Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1995); Paterson, "Biotechnology Intelligence Unit: Genome Mapping in Plants", R. G. Landes Co., Georgetown, TX, y Academic Press, San Diego, CA (1996); "The Maize Handbook", Freeling y Walbot, eds., Springer-Verlag, Nueva York, NY (1994); "Methods in Molecular Medicine: Molecular Diagnosis of Genetic Diseases", Elles, ed., Humana Press, Totowa, NJ (1996); Clark, ed., "Plant Molecular Biology: A Laboratory Manual", Clark, ed., Springer-Verlag, Berlín, Alemania (1997).

Los requisitos para una selección la ayuda de marcadores en un programa de mejora genética vegetal son: (1) el o los marcadores deberían cosegregarse o estar estrechamente unidos con el carácter deseado; (2) debería haber un procedimiento eficaz de rastreo de grandes poblaciones para el o los marcadores moleculares; y (3) la técnica de rastreo debería tener una elevada reproductibilidad en laboratorios y, preferiblemente, ser fácil de usar y económica.

El enlace genético de moléculas marcadoras se puede establecer mediante un modelo de mapeado génico tal como, sin limitación, el modelo de marcadores flanqueantes descrito por Lander y Botstein, *Genetics* 121: 185-199 (1989)

y el mapeado de intervalos, basado en procedimientos de probabilidad máxima descritos por Lander y Botstein, *Genetics* 121: 185-199 (1989) y puesto en marcha en el paquete informático MAPMAKER/QTL (Lincoln y Lander, "Mapping Genes Controlling Quantitative Traits Using MAPMAKER/QTL", Whitehead Institute for Biomedical Research, Massachusetts, (1990). Otros programas incluyen el Qgene, Versión 2.23 (1996), Departamento de biometría y mejora genética vegetal, 266 Emerson Hall, Cornell University, Ithaca, NY. El uso del programa Qgene es un enfoque particularmente preferido.

Se calcula una estimación de la probabilidad máxima (MLE) para la presencia de un marcador junto con una MLE suponiendo un efecto nulo del QTL para evitar falsos positivos. Entonces se calcula una lógica de razón de posibilidades (LOD) como: $LOD = \log_{10} (MLE \text{ para la presencia de un QTL} / MLE \text{ dada no unida a un QTL})$.

La puntuación de la LOD indica esencialmente cuánto mayor es la probabilidad de que los datos hayan surgido suponiendo la presencia de un QTL que en su ausencia. El valor umbral de LOD para evitar un falso positivo con una determinada confianza, es decir del 95%, depende del número de marcadores y de la longitud del genoma. En Lander y Botstein, *Genetics* 121: 185-199 (1989), se exponen gráficas que indican umbrales de LOD descritos además por Arus y Moreno-González, "Plant Breeding", Hayward *et al.*, (eds.) Chapman & Hall, London, pp. 314-331 (1993).

En una realización preferida de la presente invención el marcador de ácido nucleico presenta una puntuación de LOD de más de 2,0, más preferiblemente, de 2,5, incluso más preferiblemente de más de 3,0 ó 4,0 con el carácter o el fenotipo de interés. En una realización preferida, el carácter de interés es el de niveles o composiciones de tocoferol modificados.

Se pueden usar otros modelos. Se han descrito muchas modificaciones y enfoques alternativos al mapeado de intervalos, incluyendo el uso de procedimientos no paramétricos (Kruglyak y Lander, *Genetics* 139: 1421-1428 (1995)). También se pueden usar procedimientos de regresión múltiple, en los que el carácter es sometido a una regresión sobre un gran número de marcadores (Jansen, "Biometrics in Plant Breeding", van Oijen y Jansen (eds.), "Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding", Países Bajos, pp. 116-124 (1994); Weber y Wricke, "Advances in Plant Breeding", Blackwell, Berlín, 16 (1994)). Jansen y Stam, *Genetics* 136: 1447-1455 (1994), y Zeng, *Genetics* 136: 1457-1468 (1994) han descrito procedimientos que combinan el mapeado de intervalos con el análisis de regresión, mediante los cuales se somete el fenotipo a una regresión sobre un solo QTL putativo en un determinado intervalo de marcadores y al mismo tiempo sobre un número de marcadores que sirven como "cofactores". Generalmente, el uso de cofactores reduce el margen de error y el error de muestreo de las posiciones de QTL estimadas (Utz y Melchinger, "Biometrics in Plant Breeding", van Oijen y Jansen (eds.), "Proceedings of the Ninth Meeting of the Eucarpia Section Biometrics in Plant Breeding", Países Bajos, pp. 195-204 (1994), mediante lo que se mejora la eficacia del mapeado de QTL (Zeng, *Genetics* 136: 1457-1468 (1994)). Se pueden ampliar estos modelos a experimentos multi-ambientales para analizar interacciones entre genotipos y el medio ambiente (Jansen *et al.*, *Theo. Appl. Genet.* 91: 33-37 (1995)).

Se entiende que se pueden utilizar una o más de las moléculas de ácido nucleico de la invención como marcadores moleculares. También se entiende que se pueden utilizar una o más de las moléculas proteicas de la invención como marcadores moleculares.

En una realización preferida, el polimorfismo está presente y es rastreado en una población de mapeado, p. ej., una colección de plantas capaces de ser usadas con marcadores tales como marcadores polimórficos para mapear la posición genética de caracteres. La elección de la población de mapeado apropiada depende habitualmente del tipo de sistemas marcadores empleados (Tanksley *et al.*, J. P. Gustafson y R. Appels (eds.), Plenum Press, Nueva York, pp: 157-173 (1988)). Se debe tener en cuenta el origen de los padres (adaptados frente a exóticos) usados en la población de mapeado. El apareamiento de cromosomas y las velocidades de recombinación pueden ser gravemente perturbadas (inhibidas) en cruces de amplio espectro (adaptados x exóticos) y generalmente producen distancias de enlace enormemente reducidas. Los cruces de amplio espectro proporcionarán habitualmente poblaciones segregantes con un número relativamente elevado de polimorfismos en comparación con la progenie en un cruce restringido (adaptado x adaptado).

Una población F_2 es la primera generación de la autofecundación (auto-polinización) una vez producida la semilla híbrida. Habitualmente, se autofecunda una sola planta F_1 para generar una población segregante para todos los genes en un patrón de Mendel (1: 2: 1). Se obtiene una información genética máxima de una población F_2 completamente clasificada usando un sistema de marcadores codominantes (Mather, "Measurement of Linkage in Heredity": Methuen y Co., (1938)). En el caso de los marcadores dominantes, se requieren análisis de la progenie (p. ej., F_3 , BCF_2) para identificar los heterocigotos, con el fin de clasificar la población. Sin embargo, este procedimiento suele ser prohibitivo debido al coste y al tiempo implicado en el análisis de la progenie. El análisis de la progenie de individuos F_2 se usa habitualmente en la construcción de mapas en los que los fenotipos no reflejan de manera sistemática el genotipo (p. ej., la resistencia a una enfermedad) o en los que la expresión del carácter está controlada por un QTL. Los datos de separación procedentes de las poblaciones del análisis de la progenie (p.ej., F_3 o BCF_2) se pueden usar en la construcción de mapas. Entonces se puede aplicar la selección con marcadores a la progenie del cruce en base a las asociaciones en el mapa de marcador-carácter (F_2 , F_3), en las que los grupos de enlace no hayan sido disociados completamente mediante recombinaciones (i.e., desequilibrio máximo).

Se pueden usar las líneas consanguíneas recombinantes (RIL) (líneas relacionadas genéticamente; habitualmente $>F_5$, desarrolladas a partir de líneas de F_2 de autofecundación continua hacia la homocigosidad) como una población de mapeado. La información obtenida de los marcadores dominantes se puede maximizar mediante el uso de RIL, porque todos los locus son homocigóticos o casi lo son. En condiciones de fuerte unión (i.e., una recombinación de aproximadamente $<10\%$) los marcadores dominantes y codominantes evaluados en las poblaciones de RIL proporcionan más información por individuo que cualquier tipo de marcador de las poblaciones de retrocruzamiento (Reiter. *Proc. Natl. Acad. Sci. (EE.UU.)* 89: 1477-1481 (1992)). Sin embargo, a medida que crece la distancia entre los marcadores (i.e., los locus se vuelven más independientes), disminuye la información de las poblaciones de RIL de manera espectacular en comparación con los marcadores codominantes.

Las poblaciones de un retrocruzamiento (p.ej., las generadas a partir del cruce entre una variedad satisfactoria (padre recurrente) y otra variedad (padre donante) que porta un carácter que no está presente en el primero) se pueden utilizar como una población de mapeado. Se puede hacer una serie de retrocruzamientos con el padre recurrente para recuperar la mayoría de los caracteres deseables. Por lo tanto, se crea una población constituida por individuos casi como el padre recurrente, pero portando cada individuo cantidades variables o mosaico de regiones genómicas procedentes del padre donante. Las poblaciones de retrocruzamiento pueden ser útiles para mapear marcadores dominantes si todos los locus del padre recurrente son homocigóticos, y el donante y el padre recurrente tienen alelos de marcadores polimórficos contrastados (Reiter *et al.*, *Proc. Natl. Acad. Sci. (EE.UU.)* 89: 1477-1481 (1992)). La información obtenida a partir de poblaciones de retrocruzamiento que usan marcadores bien codominantes o dominantes es menor que la obtenida a partir de poblaciones F_2 , porque se muestrea uno en lugar de dos gametos recombinantes por planta. Sin embargo, las poblaciones de retrocruzamiento son más informativas (a una saturación de marcadores baja) en comparación con las RIL, pues la distancia entre los locus enlazados aumenta en las poblaciones de RIL (i.e., una recombinación del aproximadamente $0,15\%$). El aumento de la recombinación puede ser beneficioso para la resolución de los enlaces fuertes, pero puede no ser deseable en la construcción de mapas con una saturación baja de marcadores.

Se pueden usar líneas isogénicas cercanas (NIL) (creadas mediante muchos retrocruzamientos para producir una colección de individuos que es casi idéntica en composición genética a excepción del carácter o la región genómica interrogante) como una población de mapeado. En el mapeado con NIL, sólo se espera el mapeado de una parte de los locus polimórficos hasta una región seleccionada.

El análisis de segregantes agrupados (BSA) es un procedimiento desarrollado para la rápida identificación del enlace entre marcadores y caracteres de interés (Michelmore *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 88: 9828-9832 (1991)). En los BSA, se representan dos muestras de ADN agrupadas procedentes de una población segregante originaria de un solo cruce. Estos grupos contienen individuos que son idénticos para un determinado carácter (resistentes o susceptibles a una determinada enfermedad) o una región genómica pero arbitraria en regiones no alineadas (i.e., heterocigótico). Las regiones no unidas a la región diana no diferirán entre las muestras agrupadas de muchos individuos en el BSA.

En un aspecto de la presente invención, se usan una o más moléculas nucleicas de la presente invención para determinar el nivel (i.e., la concentración de ARNm de una muestra, etc.) en una planta (preferiblemente canola, maíz, *Brassica campestris*, *Brassica napus*, soja, crambe, mostaza, aceite de ricino, cacahuete, sésamo, semilla de algodón, semilla de lino, cártamo, aceite de palma, lino o girasol) o un patrón (i.e., la cinética de la expresión, la velocidad de descomposición, el perfil de estabilidad, etc.) de la expresión de una proteína codificada en parte o por completo por una o más de las moléculas de ácido nucleico de la presente invención (colectivamente, la "respuesta a la expresión" de una célula o de un tejido).

Como se usa en la presente memoria, la respuesta a la expresión manifestada por una célula o por un tejido se dice que está "modificada" si difiere de la respuesta a la expresión de células o de tejidos de plantas que no presentan el fenotipo. Para determinar si una respuesta a la expresión está modificada, se compara la respuesta a la expresión manifestada por la célula o el tejido de la planta que presenta el fenotipo con la de una muestra celular o tisular similar de una planta que no presenta el fenotipo. Como se comprenderá, no es necesario volver a determinar la respuesta a la expresión de la muestra celular o tisular de las plantas que no presentan el fenotipo cada vez que se hace tal comparación; más bien, se puede comparar la respuesta a la expresión de una determinada planta con los valores obtenidos previamente de plantas normales. Como se usa en la presente memoria, el fenotipo del organismo es cualquiera de entre una o más características de un organismo (p. ej., resistencia a una enfermedad, tolerancia a una plaga, tolerancia ambiental tal como la tolerancia a un estrés abiótico, la esterilidad masculina, la mejora de la calidad o el rendimiento, etc.). Un cambio de genotipo o fenotipo puede ser transitorio o permanente. También como se usa en la presente memoria, una muestra tisular es cualquier muestra que comprenda más de una célula. En un aspecto preferido, una muestra tisular comprende células que comparten una característica común (p. ej., derivadas de una raíz, semilla, flor, hoja, tallo o polen, etc.).

En un aspecto de la presente invención, se puede realizar una evaluación para determinar si una determinada molécula de ARNm está presente. Se utilizan una o más de las moléculas de ácido nucleico de la presente invención para detectar la presencia o la cantidad de la especie de ARNm. Entonces se incuban tales moléculas con extractos celulares o tisulares de una planta en condiciones suficientes para permitir la hibridación del ácido nucleico. La detección de las moléculas híbridas de sonda-ARNm bicatenarias indica la presencia del ARNm; la cantidad de tal híbrido formado es proporcional a la cantidad de ARNm. De este modo, se pueden usar tales sondas para determinar el nivel y el grado de producción del ARNm en células o tejidos de una planta. Tal hibridación de ácidos nucleicos se puede realizar en

condiciones cuantitativas (proporcionando así un valor numérico de la cantidad del ARNm presente). Alternativamente, se puede realizar el análisis como un análisis cualitativo que indica bien que el ARNm está presente o que su nivel supera un valor predefinido fijado por el usuario.

- 5 Se puede usar un número de procedimientos para comparar la respuesta a la expresión entre dos o más muestras de células o tejido. Estos procedimientos incluyen análisis de hibridación, tales como transferencia northern, análisis de protección de RNasa e hibridación *in situ*. Alternativamente, los procedimientos incluyen análisis de tipo PCR. En un procedimiento preferido, la respuesta a la expresión se compara hibridando los ácidos nucleicos de dos o más muestras con un alineamiento de ácidos nucleicos. El alineamiento contiene una pluralidad de secuencias sospechosas
10 conocidas o sospechosas de estar presentes en las células o el tejido de las muestras.

- Una ventaja de una hibridación *in situ* frente a técnicas más convencionales para la detección de ácidos nucleicos es que permite a un investigador determinar la población espacial exacta (Angerer *et al.*, *Dev. Biol.* 101: 477-484 (1984); Angerer *et al.*, *Dev. Biol.* 112: 157-166 (1985); Dixon *et al.*, *EMBO J.* 10: 1317-1324 (1991)). Se puede usar
15 la hibridación *in situ* para medir el nivel de estado constante de la acumulación de ARN (Hardin *et al.*, *J. Mol. Biol.* 202: 417-431 (1989)). Se ha creado un número de protocolos para la hibridación *in situ*, cada uno con condiciones para la preparación tisular, la hibridación y el lavado (Meyerowitz, *Plant Mol. Biol. Rep.* 5: 242-250 (1987); Cox y Goldberg, en: "Plant Molecular Biology: A Practical Approach", Shaw (ed.), pp. 1-35, IRL Press, Oxford (1988); Raikhel *et al.*, "In situ RNA hybridization in plant tissues", en: "Plant Molecular Biology Manual", vol. B9: 1-32,
20 Kluwer Academic Publisher, Dordrecht, Bélgica (1989)).

- La hibridación *in situ* también permite la localización de proteínas dentro de un tejido o una célula (Wilkinson, "In Situ Hybridization", Oxford University Press, Oxford (1992); Langdale, "In Situ Hybridization" en: "The Maize Handbook, Freeling y Walbot" (eds.), pp. 165-179, Springer-Verlag, Nueva York (1994)). Se entiende que es posible
25 utilizar una o más moléculas de la invención, preferiblemente una o más moléculas de ácido nucleico, o fragmentos de las mismas, de la invención, o uno o más anticuerpos de la invención para detectar el nivel o el patrón de una proteína o un ARNm de la misma mediante hibridación *in situ*.

- La hibridación *in situ* fluorescente permite la localización de una determinada secuencia de ADN a lo largo de
30 un cromosoma, que sea útil, entre otros usos, para el mapeado de genes, el seguimiento de cromosomas en líneas híbridas o la detección de cromosomas con translocaciones, transversiones o eliminaciones. La hibridación *in situ* se ha usado para identificar cromosomas de varias especies vegetales (Griffor *et al.*, *Plant Mol. Biol.* 17: 101-109 (1991); Gustafson *et al.*, *Proc. Natl. Acad. Sci.* (EE.UU.) 87: 1899-1902 (1990); Mukai y Gill, *Genome* 34: 448-452 (1991); Schwarzscher y Heslop-Harrison, *Genome* 34: 317-323 (1991); Wang *et al.*, *Jpn. J. Genet.* 66: 313-316 (1991); Parra y Windle, *Nature Genetics* 5: 17-21 (1993)). Se entiende que se pueden usar las moléculas de ácido nucleico de la
35 invención como sondas o marcadores para localizar secuencias a lo largo de un cromosoma.

- Otro procedimiento para localizar la expresión de una molécula es la impresión de tejidos. La impresión de tejidos proporciona un modo de detectar, al mismo tiempo sobre la misma membrana, muchas partes de tejido procedentes de
40 diferentes plantas o diferentes etapas evolutivas (Yomo y Taylor, *Planta* 112:35-43 (1973); Harris y Chrispeels, *Plant Physiol.* 56: 292-299 (1975); Cassab y Varner, *J. Cell. Biol.* 105: 2581-2588 (1987); Spruce *et al.*, *Phytochemistry* 26: 2901-2903 (1987); Barres *et al.*, *Neuron* 5: 527-544 (1990); Reid y Pont-Lezica, "Tissue Printing: Tools for the Study of Anatomy, Histochemistry and Gene Expression", Academic Press, Nueva York, Nueva York (1992); Reid *et al.*, *Plant Physiol.* 93: 160-165 (1990); Ye *et al.*, *Plant J.* 1: 175-183 (1991)).

- Cualquier experto en la técnica puede recurrir a textos de referencia generales para obtener descripciones detalladas de técnicas conocidas tratadas en la presente memoria o técnicas equivalentes. Estos textos incluyen "Current Protocols in Molecular Biology" de Ausubel, *et al.*, eds., John Wiley e hijos, N.Y. (1989), y los suplementos de septiembre
50 (1998), "Molecular Cloning, A Laboratory Manual", Sambrook *et al.*, 2ª Ed., Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1989), "Genome Analysis: A Laboratory Manual 1: Analyzing DNA", Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1997); "Genome Analysis: A Laboratory Manual 2: Detecting Genes", Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1998); "Genome Analysis: A Laboratory Manual 3: Cloning Systems", Birren *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1999); "Genome Analysis: A Laboratory Manual 4: Mapping Genomes", Birren *et al.*, Cold Spring Harbor Press, Cold Spring
55 Harbor, Nueva York (1999); "Plant Molecular Biology: A Laboratory Manual", Clark, Springer-Verlag, Berlin, (1997), "Methods in Plant Molecular Biology", Maliga *et al.*, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York (1995). Estos textos también pueden ser consultados, por supuesto, para elaborar o usar un aspecto de la invención. Se entiende que cualquiera de los agentes de la invención puede ser sustancialmente purificado y/o ser biológicamente activo y/o recombinante.

- Habiendo ya descrito de manera general la invención, se comprenderá más fácilmente la misma en referencia a los
60 siguientes ejemplos que se proporcionan a modo de ilustración y no pretenden ser restrictivos de la presente invención a no ser que se especifique lo contrario.

Ejemplo 1

Clonación de tyrA de Erwinia herbicola

Se obtienen los vectores pJX1, pJX181 y pJX184 (Zhao y Jensen de "Molecular Evolution" 36 (2): 107-20 (1993)). Se amplifica el gen *tyrA* por PCR usando cebadores tyrA5' (ACT GCC ATG GTG GCT GAA CTG ACC G (SEC ID N.º 5)) y tyrA3' (ACT GGA ATT CTT ATT ATG GGC GGC TGT CAT TG (SEC ID N.º 6)) y ADN plasmídico procedente de los vectores pJX1, pJX181 y pJX184 como ADN molde. Se llevan a cabo reacciones PCR usando el equipo de PCR de alta fidelidad Expand™ de Boehringer Mannheim en un volumen total de 50 µl según el protocolo del fabricante. Se amplifica el gen *tyrA* mediante 30 ciclos de PCR bajo las siguientes condiciones: incubación de 10 min a 95°C, seguida por 30 ciclos de 1 min a 95°C, apareamiento de 1 min a 56°C y prolongación de 1,5 min a 72°C. Estas reacciones fueron seguidas por una incubación a 72°C durante 5 minutos. La forma del producto de la PCR pJX184 se selecciona para la clonación de los genes y se digiere con NcoI y EcoRI. Se une el fragmento de restricción purificado sobre gel en el pSE280 digerido por NcoI/EcoRI purificado sobre gel (Invitrogen Co., Carlsbad, CA) resultando en la formación de pMON26588 (Figura 2). El inserto de *tyrA* en pMON26588 es verificado mediante secuenciación de ADN.

Ejemplo 2

Clonación de tyrA de Escherichia coli

Se amplifica el gen *tyrA* de *E. coli* por PCR usando cebadores tyrAecoli5' (ACT GCC ATG GTT GCT GAA TTG ACC G (SEC ID N.º 7)) y tyrAecoli3' (ACT GGA ATT CTT ATT ACT GGC GAT TG (SEC ID N.º 8)) y ADN genómico total de DH5a de *E. coli* como ADN molde. Se aísla el ADN genómico total de *E. coli* usando el equipo Qiaamp Tissue de Qiagen (Qiagen Inc. Valencia, CA). Se llevan a cabo reacciones PCR usando el equipo de PCR de alta fidelidad Expand™ de Boehringer Mannheim en un volumen total de 50 µl según el protocolo del fabricante. Se amplifica el gen *tyrA* mediante 30 ciclos de PCR bajo las siguientes condiciones: incubación de 10 min a 95°C, seguida por 30 ciclos de 1 min a 95°C, apareamiento de 1 min a 56°C y prolongación de 1,5 min a 72°C. Estas reacciones fueron seguidas por una incubación a 72°C durante 5 minutos. Se digiere el producto de la PCR con NcoI y EcoRI. Se une el fragmento de restricción purificado sobre gel en pSE280 digerido con NcoI/EcoRI y purificado sobre gel (Invitrogen Co., Carlsbad, CA) resultando en la formación de pMON26589 (Figura 3). El inserto de *tyrA* en pMON26589 es verificado por secuenciación de ADN.

Ejemplo 3

Expresión de prefenato deshidrogenasa bifuncional en E. coli

Se transforman los vectores pMON26588 y pMON26589 en DH5a de *E. coli* y se cultivan las células en un cultivo de LB de 15 ml hasta una densidad óptica a 600 nm de aproximadamente 0,6, y se inducen añadiendo IPTG hasta una concentración final de 0,66 µM. Tras una incubación durante 2 a 3 horas, se cosechan las células. Se vuelven a suspender los sedimentos celulares en 0,5 ml de Tris/HCl 25 mM, pH 8,2 y se perturba el desarrollo de las células mediante un tratamiento de ultrasonidos. Se sedimentan las membranas y los restos celulares mediante centrifugación a 100.000 xg durante tres horas. Se usa el sobrenadante en análisis enzimáticos como un extracto celular crudo. Se mide la actividad prefenato deshidrogenasa en un volumen final de 1,5 ml que contiene EDTA 1 mM, DTE 1 mM, NAD 1 mM y prefenato (sal de Ba) 1 mM en Tris/HCl 25 mM, pH 8,2. Se determina la actividad específica de la prefenato deshidrogenasa controlando la conversión de NAD⁺ en NADH según lo descrito en "Methods in Enzymology" Vol. 17 (Parte A) páginas 564-574 (1970). En la siguiente tabla 1 se muestran los resultados.

TABLA 1

Denominación del vector	Gen	Actividad específica
		µmol/mg x min
Control de tipo natural		2 x 10 ⁻⁵
pMON26588	<i>TyrA de Erwinia herbicola</i>	5,75
pMON26589	<i>TyrA de Escherichia coli</i>	3,44

Ejemplo 4

Colocación del gen tyrA bajo el control del promotor T7

- 5 Se escinden los genes *tyrA* de *E. coli* y *E. herbicola* como fragmentos de NcoI/EcoRI desde pMON26589 y pMON26588, se purifican sobre gel y se clonan en pMON26541 digerido por NcoI/EcoRI y purificado sobre gel (Figura 28), dando como resultado la formación de pMON26591 y pMON26590, respectivamente (figuras 4 y 5). Estos vectores colocan al gen *tyrA* bajo el control del promotor T7.

10

Ejemplo 5

Preparación de un vector de expresión vegetal con tyrA

- 15 Se selecciona el gen *tyrA* de *E. herbicola* para la expresión vegetal. Se escinde el gen de pMON26590 mediante el digesto de restricción NcoI/EcoRI, se purifica sobre gel y se liga en pMON26541 digerido con NcoI/EcoRI y purificado sobre gel, dando como resultado la formación del vector lanzadera pMON36510 (figura 6). Estas uniones fusionan el gen *tyrA* bacteriano con el CTP1, que es el péptido diana cloroplástico de la subunidad pequeña de la ribulosa bifosfato carboxilasa de *Arabidopsis*, y se coloca bajo el control del promotor e35S.

20

- Para colocar el gen *tyrA* bajo el control del promotor de napina, se digiere pMON36510 con EcoRI, se rellenan los extremos usando el fragmento de Klenow (Maniatis) y se digiere el vector purificado sobre gel con Bgl II. Se purifica sobre gel el fragmento más pequeño codificante del gen *tyrA* fusionado con CTP1 y se une para una unión en pCGN3223 digerido y purificado sobre gel (figura 45). Para realizar esta unión, se digiere pCGN3224 con PstI, se rellenan los extremos con fragmento de Klenow (Maniatis) y se digiere posteriormente el vector con Bgl II y se purifica sobre gel. La unión del vector purificado y la fusión de CTP1::*tyrA* purificada da como resultado la formación de pMON36512 (figura 7).

25

- Para transferir el gen *tyrA* de *E. herbicola* a un vector binario de *Arabidopsis*, se digiere pMON36510 con HindIII y Sac I, y se fusiona el fragmento purificado sobre gel que porta el promotor e35S con CTP1, y se une *tyrA* en pMON26543 digerido con HindIII-SacI y purificado sobre gel (figura 29), lo que resulta en la formación de pMON36511 (figura 8). Este vector contiene *tyrA* bajo el control del promotor e35S. El vector de expresión binario pNapina se obtiene uniendo el fragmento de NotI purificado sobre gel que alberga el casete de expresión de pNapina::CTP1::*tyrA*::3' de napina en pMON36176 digerido con NotI (figura 30), lo que da como resultado la formación de pMON36520 (figura 9).

35

Ejemplo 6

- 40 Se prepara la transformación de *Arabidopsis* con pMON36520 y pMON36511 de *Agrobacterium* que han sido transformados con pMON36520 y pMON36511 como se explica a continuación. Se extienden 100 µl de un cultivo de una noche sobre una placa de LB en agar con antibióticos. Se coloca la placa al revés en una cámara a 30°C durante una noche. Se retiran las placas una vez que las colonias han crecido (24-48 horas).

- 45 Se inicia un cultivo a pequeña escala colocando 10 ml de medio de LB líquido en un tubo de 50 ml. Se añaden 10 µl de kanamicina (50 µg/µl), 10 µl de espectinomicina (75-100 µg/µl) y 10 µl de cloranfenicol (25 µg/µl). Se añade *Agrobacterium* de una placa y se agita el tubo y se coloca en un agitador a 30°C durante una noche.

- Tras el cultivo de una noche de 10 ml de cultivo, se retira el mismo a un matraz de 500 ml. Se colocan 200 ml de LB líquido en un matraz, se añaden 200 µl de kanamicina (50 µg/µl), 200 µl de espectinomicina (75-100 µg/µl) y 200 µl de cloranfenicol (25 µg/µl), y después se añade el cultivo de una noche de 10 ml entero. Se coloca el matraz de 500 ml en un agitador a 30°C y se cultiva durante una noche.

50

- Se coloca el cultivo de 200 ml entero en un tubo de centrifugación y se centrifuga durante 25 minutos a 3.750 rpm y a 19°C. Tras la centrifugación, se vierte el líquido y se vuelven a suspender los sedimentos en 25 ml de solución de sacarosa al 5% (Silwet al 0,05%).

55

- Se colocan 900 µl de solución de sacarosa y 100 µl del cultivo bacteriano de 25 ml en una cubeta y se agita la cubeta tapada con una envoltura transparente. Se toma una lectura de la DO del blanco con 1 ml de solución de sacarosa y luego se toman lecturas de todas las soluciones bacterianas. Se registra la DO (a una longitud de onda de 600) de cada cultivo. Entonces se realizan los siguientes cálculos: $C_1 V_1 = C_2 V_2$; $C_1 V_1 = (0,8) (200 \text{ ml})$; $C_1 V_1 = 160$; $V_1 = 160/C_1$; y $V_1 = X \text{ ml}/10$ para determinar la $DO_{600} = 0,8$ de un cultivo de *Agrobacterium*.

60

- Se empapan las plantas durante al menos 30 minutos en agua antes de sumergirlas. Se vierte la solución bacteriana en un recipiente de plástico poco profundo y se sumergen las partes aéreas de la planta (flores y semillas prematuras, rosetas) en la solución durante 3-5 segundos con una agitación suave. Se colocan las plantas sumergidas de lado en una placa negra revestida con un pañal y se cubren con un domo durante una noche (16-24 horas) para mantener una humedad elevada. Se retira la tapa y se reanudan las condiciones de crecimiento normales de la planta durante 4 semanas.

65

ES 2 318 004 T3

Tras la transformación y el tratamiento a humedad elevada, se mantienen las plantas a 22°C, a una HR del 60% y un fotoperíodo de 16 horas durante 4 semanas. A los 5-7 días de la transformación, las plantas han dado fruto. Se realiza semanalmente una fertilización con un fertilizador de 20-20-20 semanas. Tras 4 semanas de crecimiento, se colocan las plantas en un invernadero y se detiene el riego para potenciar el secado de la planta destinado a cosechar las semillas. Las plantas están listas para la cosecha de las semillas tras 1-1,5 semanas de secado.

Las semillas se cosechan cortando la base de la planta por debajo de los conos, colocando la planta sobre un tamiz de semillas y una hoja blanca de papel, y pasando las semillas por el orificio del cono para recoger las semillas limpias mediante tamizado.

Se esterilizan las semillas conectando una manguera de secado al vacío en una campana de evacuación de humos/banco de flujo. Se colocan 100 ml de blanqueador en un vaso de precipitados de 250 ml y se añaden 3 ml de HCl concentrado al blanqueador. Se coloca el vaso de precipitados en el desecador y las semillas de los tubos de semillas en un soporte de tubos del desecador. Se pone una tapa sobre el desecador y se inicia el vacío. Se deja toda la noche el desecador, pero no más de 16 horas.

Una vez esterilizadas, se colocan las semillas sobre medio de selección (preparado añadiendo 10 g (2 g/l) de Phytagel; 10,75 g (2,15 g/l) de sales basales MS (M-5524 de Sigma); 50 g (10 g/l) de sacarosa y 6 ml (1,2 ml/l) de solución de kanamicina (950 mg/ml); 5 ml (1 ml/l) de solución de cefotaxima (250 mg/ml) y 5 ml (1 ml/l) de solución de carbenicilina (250 mg/ml) hasta un volumen total de 5 litros a un pH de 5,7). Se golpean ligeramente los tubos de semillas sobre una placa para distribuir las semillas escasamente. Se envuelven las placas con una película transparente y se colocan en un refrigerador a 4°C durante 1-2 días de tratamiento frío. Tras este tratamiento de frío, se colocan las placas en una cámara a 28°C para la germinación.

Las plántulas seleccionadas son verdes y tienen un desarrollo de hojas secundarias. Se mueven las plántulas seleccionadas a la tierra una vez desarrolladas las hojas secundarias.

Se plantan las plántulas en tierra y se tapan con un domo durante 5 días para mantener una humedad elevada. Se transplantan las plántulas a un invernadero una vez que las silicuas inferiores comienzan a amarillear.

Se cultivan las semillas de las plántulas seleccionadas en macetas de 6,35 cm con tierra (1/2 Metro-200; 1/2 PGX Mix). Se amontona la tierra y se cubre la maceta con una malla. Se sujeta la malla a la maceta con una tira de caucho. Se siembran las semillas y se tapan con un domo de germinación.

Se cultivan las plántulas en un fotoperíodo de 12 h a una humedad relativa del 70% a 22°C. Se suministra agua cada dos días según lo necesario y se aplica un fertilizador 20-20-20 de Peter desde abajo cada dos semanas.

Ejemplo 7

Se cultivan las plantas de semillas transformadas del ejemplo 6 que representan 20 transformaciones independientes y se recogen las semillas para producir semillas T₂. Se cultivan las semillas T₂ y se analizan en cuanto a los niveles de tocoferol. En la figura 1, se expresan los niveles de tocoferol como los nanogramos de tocoferol total por cada miligramo de semilla. Se determinan los niveles de tocoferol añadiendo de 10 a 15 mg de semillas de *Arabidopsis* a un microtubo de 2 ml. Se añaden al tubo una masa de 1 g de micropelotas de 0,5 mm (Biospecifics Technologies Corp., Lynbrook, NY) y 500 µl de pirogalol al 1% (Sigma Chem, St. Louis, MO) en etanol que contiene 5 µg/ml de tocol. Se agita la muestra dos veces durante 45 segundos en un FastPrep (Bio101/Savant) a una velocidad de 6,5. Se filtra el extracto (acrodisco de 0,2 µm de PTFE Gelman, filtros de jeringa de 13 mm, Pall Gelman Laboratory Inc, Ann Arbor, MI) en un tubo automuestreador. Se realiza una CLAR sobre una columna de CLAR de sílice Zorbax, 4,6 mm x 250 mm (5 µm) con una detección de la fluorescencia usando una CLAR de Hewlett Packard (Agilent Technologies, Palo Alto CA). Se realiza la excitación de las muestras a 290 nm y se controla la emisión a 336 nm. Se separan los tocoferoles con un gradiente de hexano metil-t-butil éter usando un volumen de inyección de 20 µl, un caudal de 1,5 ml/min y un tiempo de ejecución de 12 min (40°C). Se calcula la concentración y la composición de los tocoferoles en base a curvas estándar para los α, β, δ y γ-tocoferoles y α, β, δ y γ-tocotrienoles usando un programa Chemstation (Agilent Technologies, Palo Alto CA).

Ejemplo 8

Plantas transformadas con *tyrA* y otros genes de la biosíntesis del tocoferol

Se transforman plantas de canola, *Brassica napus*, *Arabidopsis* y soja con una variedad de constructos de ADN usando un enfoque de bombardeo de partículas esencialmente según lo expuesto en Christou, "In Particle Bombardment for the Genetic Engineering of Plants", Biotechnology Intelligence Unit, Academic Press, San Diego, California (1996) y usando una transformación mediada por *Agrobacterium*. Se producen dos conjuntos de constructos de ADN. El primer conjunto de constructos es el de los "constructos de un solo gen". Cada uno de los siguientes genes se inserta en un constructo de ADN vegetal independiente bajo el control de un promotor de napina (Krindl *et al.*, *Seed Sci. Res.* 1:209: 219 (1991)) y los productos de los genes se pueden dirigir hacia el plasto mediante un péptido dirigido a plastos

codificado tal como el CTP1 (Keegstra, *Cell* 56 (2): 247-53 (1989); Nawrath, *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 91: 12760-12764 (1994)) o el CTP2: un gen *tyrA* de *E. herbicola* (Xia *et al.*, *J. Gen. Microbiol.* 138: 1309-1316 (1992)), un gen *slr1736* (en la Cyanobase de la web www.kazusa.or.jp/cyanobase), un gen *ATPT2* vegetal (Smith *et al.*, *Plant J.* 11: 83-92 (1997)), un gen *dxs* (Lois *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 95 (5): 2105-2110 (1998)), un gen *dxr* (Takahashi *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 95 (17), 9879-9884 (1998)), un gen *HPPD* de *Arabidopsis thaliana* (Norris *et al.*, *Plant Physiol.* 117: 1317-1323 (1998)), un gen *GGH* (Keller *et al.*, *Eur. J. Biochem.* 251: 413-417 (1998)), un gen *GGPPS* de *Arabidopsis thaliana* (Bartley y Scolnik, *Plant Physiol.* 104: 1469-1470 (1994)), un gen *AANT1* (Saint Guily, *et al.*, *Plant Physiol.*, 100 (2): 1069-1071 (1992)), se usó un gen *MT1* (la secuencia de *MT1* de *Synechocystis* (Número de identificador general del NCBI 1653572) en la búsqueda de blastos frente a EST de la cepa PCC 7120 de *Anabaena sp.* (Kaneko 2001). Se descubrió una secuencia con una homología sustancial con *MT1* de *Synechocystis* en una búsqueda de blastos frente a EST de la cepa PCC 7120 de *Anabaena sp.* (Kaneko *et al.*, *DNA Research* 8 (5): 205-213 (2001)), un gen *TMT2* (según lo revelado en la solicitud estadounidense US 2003/0150015), un gen *GMT* (según lo revelado en los documentos U.S. 2003/0154513; WO 00/32757, WO 00/10380) y un gen *slr1737* (en la cyanobase de la web www.kazusa.or.jp/cyanobase) y un constructo antisentido para la ácido homogentísico dioxigenasa (Sato *et al.*, *J. DNA Res.* 7 (1): 31-63 (2000)). Se transforma cada constructo en al menos una planta de canola, *Brassica napus*, *Arabidopsis* y soja. Se seleccionan las plantas que expresan cada uno de estos genes para participar en más cruces. Se analiza también la composición y el nivel de tocoferol de cada planta usando el procedimiento expuesto en el ejemplo 7. Se llevan a cabo los cruces para cada especie para generar plantas transgénicas que tengan uno o más de los de la siguiente combinación de genes introducidos: *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa. En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *slr1736* o *ATPT2*.

Se analiza también la composición y el nivel de tocoferol de cada planta generada mediante los cruces (incluyendo todos los cruces intermedios) usando el procedimiento expuesto en el ejemplo 7. La progenie de los transformantes de estos constructos será cruzada entre sí para apilar los genes adicionales y alcanzar el nivel deseado de tocoferol.

Se genera el segundo conjunto de constructos de ADN y se denomina “constructos de múltiples genes”. Los constructos de múltiples genes contienen múltiples genes cada uno bajo el control de un promotor de napina (Krindl *et al.*, *Seed Sci. Res.* 1: 209: 219 (1991)) y los productos de cada uno de los genes son dirigidos al plasto por un péptido dirigido a plastos codificado. El constructo de múltiples genes puede tener dos o más de los siguientes genes: *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa. En una combinación preferida, el constructo o los constructos de ácido nucleico codifican, además de *tyrA*, *HPPD* y bien *slr1736* o *ATPT2*.

Se transforma entonces cada constructo en al menos una planta de canola, *Brassica napus*, *Arabidopsis* y soja. Se analiza también la composición y el nivel de tocoferol de cada planta usando el procedimiento expuesto en el ejemplo 7. La progenie de los transformantes de estos constructos será cruzada entre sí para apilar los genes adicionales y alcanzar el nivel deseado de tocoferol.

Ejemplo 9

Plantas de *Arabidopsis* transformadas con *tyrA*, *ATPT2* y otros genes de la biosíntesis del tocoferol

Se transforman plantas de *Arabidopsis* de tipo natural y líneas vegetales de *Arabidopsis* con el vector plasmídico pMON69907 (figura 12), se cultivan y se recogen las semillas según lo descrito en los ejemplos anteriores, y se analiza la semilla en cuanto al contenido de tocoferoles y tocotrienoles según lo descrito anteriormente. El plásmido pMON69907 codifica una pifenato deshidrogenasa bifuncional (*tyrA*) y una fitil preniltransferasa (*ATPT2*). La figura 14 representa el contenido de tocoferoles y tocotrienoles totales de semillas de *Arabidopsis* procedentes de plantas de tipo natural y de varias líneas vegetales transformadas con el vector plasmídico pMON69907. La figura 15 representa el contenido de tocoferoles totales de semillas de *Arabidopsis* procedentes de una planta de tipo natural y de varias líneas vegetales transformadas con el vector plasmídico pMON69907. La figura 31 muestra los patrones de CL/EM para el tocoferol y el tocotrienol. La figura 32 muestra los resultados de la LC/EM para las líneas seleccionadas, mostrando la presencia de tocotrienoles. La figura 33 muestra un cromatograma de CLAR/FLD de extracto de semillas de control, que muestra la falta de presencia de tocotrienoles. La figura 34 muestra un cromatograma de CLAR/FLD de extracto de semillas de control que muestra la presencia de tocotrienoles en las líneas seleccionadas.

Ejemplo 10

Plantas de *Arabidopsis* transformadas con *tyrA* y otros genes de la biosíntesis del tocoferol

Se preparan los constructos de expresión pCGN10822, pMON36528, pMON69907 y pMON69909 mostrados en las figuras 10-13, respectivamente.

Se transforman plantas de *Arabidopsis* con los vectores indicados usando las técnicas de transformación descritas en el ejemplo 8. Se aíslan los transformantes y se cultivan en líneas individuales mediante auto-polinización y se recoge la semilla de cada línea. Se analiza la composición de los tocoferoles y tocotrienoles de las semillas de cada línea usando el procedimiento expuesto en el ejemplo 7. La figura 16 muestra los niveles de tocoferoles y tocotrienoles totales para las líneas vegetales que albergan los constructos descritos o un control. En la figura 17, se muestra un análisis de las semillas T₂ de las líneas vegetales derivadas de la transformación con el vector pMON69909 en relación con el tipo natural. Las líneas vegetales transformadas con pMON69909 demuestran un aumento sustancial de los tocoferoles totales y los tocotrienoles totales, con aumentos mayores en el delta tocoferol, alfa tocotrienol, delta tocotrienol y gamma tocotrienol. Alguna semilla de las plantas que albergan el vector pMON69909 muestran una coloración oscura como resultado de la acumulación de ácido homogentísico, lo que queda confirmado por el análisis de CL/EM (véanse las figuras 31 y 32).

La expresión heteróloga de *tyrA* en las semillas de plantas de *Arabidopsis* transgénicas produce un aumento de 1,6 veces de los niveles de tocoferol de las semillas en comparación con la líneas control. Otra enzima clave esencial para la biosíntesis del tocoferol es la HPT, que está implicada en la condensación del fitil pirofosfato (PPP) y el homogentisato (HGA) para producir 2-metil-6-ftilplastoquinol (2M6PPQ), un precursor para la síntesis de cuatro isoformas diferentes de tocoferoles. La sobre-expresión de HPT_{Arabidopsis} (ATPT2) y HPT_{Synechocystis} (slr1736) independientemente en semillas de *A. thaliana* transgénicas resulta en un aumento de 1,6 veces de los tocoferoles de las semillas. Se observa que un transportador del adenilato putativo procedente de *A. thaliana* (AANT1) expresado como un solo gen aumenta los niveles de tocoferol en las semillas hasta 1,4 veces en *A. thaliana*. Para analizar si una combinación de estos genes daría como resultado un efecto sinérgico sobre la biosíntesis del tocoferol, se analizaron diversas combinaciones en *A. thaliana*.

Se analizan, en cuanto al contenido y la composición de los tocoferoles de las semillas, las semillas de *Arabidopsis* de T₂ que albergan los constructos de los dos genes *ATPT2* y *tyrA* (pMON69907), los constructos de los tres genes *ATPT2 tyrA* y *HPPD* (pMON69909), los constructos de los tres genes *ATPT2, tyrA* y *GGPPS* (pMON69915 (figura 35)) y los constructos de los tres genes *ATPT2, tyrA* y *AANT1* (pMON69919 (figura 36)). El contenido de los tocoferoles y tocotrienoles totales de las semillas aumenta hasta aproximadamente 2,4-veces en las líneas transformadas con pMON69907 (vector de dos genes) y hasta 5 veces en las líneas que portan el vector de tres genes (pMON69909) (véase la figura 16 y 17). El HPPD expresado como un solo gen en *A. thaliana* da como resultado un aumento apenas detectable de los niveles de tocoferol. La combinación de *HPPD* con *ATPT2* no resulta en un mayor aumento de los niveles de tocoferol en comparación con las líneas que albergan sólo *ATPT2* (no se muestran los datos). Por el contrario, cuando se combina *HPPD* con *tyrA* y *ATPT2*, los niveles de tocoferoles y tocotrienoles se duplican en comparación con la combinación de *tyrA* y *ATPT2*. Las semillas que albergan el constructo de tres genes pMON69909 aparecen mucho más oscuras que las semillas control. Además, se sabe que las plantas dicotiledóneas de tipo natural no acumulan tocotrienoles. Sin embargo, las semillas de *A. thaliana* transgénica que albergan los cuatro constructos acumulan niveles sustanciales de tocotrienoles (confirmados por CLAR y para las muestras seleccionadas por CL/EM (véanse las figuras 31, 32, 33 y 34). El contenido de tocoferoles y tocotrienoles de las semillas que albergan el constructo de expresión de tres genes pMON69909 constan del 60% de tocotrienoles y el 40% de tocoferoles. Cuando se eleva la disponibilidad del HGA endógeno mediante la sobre-expresión de las enzimas biosintéticas de HGA (*tyrA* y *HPPD*) junto con la HPT, la HPT utilizaría geranilgeranil pirofosfato (*GGPP*) y HGA para producir tocotrienoles en lugar de tocoferoles en condiciones limitadas por la disponibilidad del nivel endógeno de la geranilgeranil reductasa (*GGH*). La *GGH* funciona sobre la hidrogenación del *GGPP* en *PPP*, un sustrato para la *HPT* en la síntesis del tocoferol. Por consiguiente, la mayor acumulación de tocotrienoles observada en los constructos analizados puede ser superada por la sobre-expresión de la *GGH* en combinación con *tyrA*, *HPPD* y *HPT*.

Ejemplo 11

Plantas transformadas con *tyrA* y otros genes de la biosíntesis del tocoferol

Se transforman plantas con los constructos de ADN mostrados en las tablas 2 y 3 que se presentan a continuación, empleando las técnicas descritas en el ejemplo 8. Los constructos contienen uno o más de los genes bajo el control de un promotor de napina (Krindl *et al.*, *Seed Sci. Res.* 1: 209: 219 (1991)), el promotor 7Sa' (Chen *et al.*, *PNAS* 83 (22): 8560-8564 (1998)) o el promotor Arc5 (Goossens *et al.*, *Plant Physiol.* 120: 1095-1104 (1999)). Los productos de los genes se pueden dirigir al plasto mediante un péptido dirigido a plastos codificado tal como el CTP1 (Keegstra, *Cell* 56(2): 247-53 (1989); Nawrath, *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 91: 12760-12764 (1994)) o CTP2. Se usa uno o más de los siguientes genes: un gen *tyrA* de *E. herbicola* (Xia *et al.*, *J. Gen. Microbiol.* 138: 1309-1316 (1992)), un gen slr1736 (en la Cyanobase de web www.kazusa.org.jp/cyanobase), un gen *ATPT2* (Smith *et al.*, *Plant J.* 11: 83-92 (1997)), un gen *dxs* de *E. coli* (Lois *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 95 (5): 2105-2110 (1998)), un gen *dxr* (Takahashi *et al.*, *Proc. Natl. Acad. Sci. EE.UU.* 95 (17), 9879-9884 (1998)), un gen *HPPD* (Norris *et al.*, *Plant Physiol.* 117: 1317-1323 (1998)), un gen *GGH* (Keller *et al.*, *Eur. R Biochem.* 251: 413- 417 (1998)), un gen *GGPPS* de *Arabidopsis thaliana* (Bartley y Scolnik, *Plant Physiol.* 104: 1469-1470 (1994)), un gen *AANT1* (Saint Guily, *et al.*, *Plant Physiol.*, 100 (2): 1069-1071 (1992)), un gen *MT1* (como el anterior para el ejemplo 8), un gen *TMT2* (como el anterior para el ejemplo 8), un gen *GMT* (como el anterior para el ejemplo 8, y los documentos WO 00/32757, WO 00/10380), un gen slr1737 (de la Cyanobase de la web www.kazusa.org.jp/cyanobase), y un constructo antisentido para la ácido homogentísico dioxigenasa (denominada HGD_{AS}) (Sato *et al.*, *J. DNA Res.* 7 (1) 31-63 (2000)). Se transforma cada constructo en al menos una planta de canola, *Brassica napus*, *Arabidopsis* y soja. Se analiza también

la composición y el nivel de tocoferol de cada planta usando el procedimiento expuesto en el ejemplo 7. Los ejemplos de las plantas transformadas con *tyrA* y otros genes de la biosíntesis del tocoferol incluyen plantas de *Arabidopsis* transformadas con los constructos expuestos en la tabla 2 y plantas de soja transformadas con los constructos de la tabla 3.

Se pueden someter las plantas con las características deseadas a más cruces para generar plantas transgénicas que tengan uno o más genes de la siguiente combinación de genes introducidos: *tyrA*, *slr1736*, *ATPT2*, *dxs*, *dxr*, *GGH*, *GGPPS*, *HPPD*, *MT1*, *TMT2*, *GMT*, *AANT1*, *slr1737*, y un constructo antisentido para la ácido homogentísico dioxigenasa. Alternativamente, se pueden cruzar las plantas para apilar múltiples copias de uno o más de los genes anteriormente mencionados en una planta transgénica.

TABLA 2

Se preparan y analizan las siguientes combinaciones de genes en <i>Arabidopsis thaliana</i>	
1	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::CTP2::HPPD _{A. thaliana} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina
2	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::AANT1 _{A. thaliana} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina
3	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::GGPPS _{A. thaliana} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina
4	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::CTP1::DXS _{E. coli} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina/pNapina::GGH::3' de napina
5	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::CTP1::DXR _{E. coli} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina/pNapina::GGH::3' de napina
6	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::HGD _{AS} ::3' de napina
7	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::HGD _{AS} ::3' de napina/pNapina::ATPT2 _{A. thaliana}
8	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::HGD _{AS} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina/pNapina::GGH::3' de napina
9	pNapina::CTP1::TyrA _{E. herbicola} ::3' de napina/pNapina::HGD _{AS} ::3' de napina/pNapina::ATPT2 _{A. thaliana} ::3' de napina/pNapina::CTP1::DXS _{E. coli} ::3' de napina/pNapina::GGH::3' de napina

TABLA 3

Se preparan y analizan las siguientes combinaciones de genes en <i>Glycine max</i>	
1	p7S::CTP2::HPPD::3' de E9/p7Sa::CTP1::TyrA::3' de E9
2	pArc5::ATPT2::3' de Arc/p7Sa::CTP1::TyrA::3' de E9/pNapina::GGH::3' de Napina
3	pArc5::ATPT2::3' de Arc/p7S::CTP1::TyrA::3' de E9/pNapina::GGH::3' de Napina/pNapina::CTP1::DXS::3' de Napina

Ejemplo 12

Construcción de vectores binarios vegetales que albergan combinaciones de tyrA con otros genes de la síntesis del tocoferol

Los componentes de cada casete de expresión génica incluyen un promotor, en este ejemplo, el promotor de napina, un terminador, un péptido dirigido a plastos (que puede ser un péptido dirigido a plastos nativo o un péptido diana cloroplástico fusionado al terminal N) y un gen de interés, según lo mostrado en las figuras 18a y 18b. Los casetes de expresión se pueden orientar de cabeza a cola, de cabeza a cabeza o la orientación puede variar.

La clonación se realiza usando casetes de expresión flanqueados por los sitios de restricción de Bsp120 I y Not I. Se construye un vector lanzadera (pMON36582 (figura 19)) apareando los cebadores SV MCS 1A y SV MCS 1B:

SV MCS 1A Xma I Bsp120 I Eag Xba I EcoRI
GATCT CCCGGG AA GGCCC CGGCCG TCTAGA GAATTC

Not I Asc I Age Iv
GCGGCCGC GGCGCGCC ACCGGT (SEC ID N.º 9)

SV MCS 1B Xma I Bsp120 I Eag Xba I EcoRI
TCGA ACCGGT GGCGCGCC GCGGCCGC GAATTC TCTAGA

Not I Asc I Age I
CGGCCG GGGCCC TT CCCGGG A (SEC ID N.º 10)

y uniéndolos en pSP72 digerido por Bgl II y Xho I, y purificado sobre gel (Promega, www.promega.com). El vector resultante se denomina pMON36582 (figura 19). El vector es confirmado mediante secuenciación de ADN.

Todos los casetes de expresión génica están preparados para ser flanqueados por los sitios de restricción de Not I. Se aíslan estos casetes mediante la digestión de los vectores anteriores con Not I, seguida por la purificación sobre gel de los casetes de expresión. Se digiere pMON36582 con Eag I, que corta dos veces en este vector, una vez en el sitio de Not I y una vez 19 pb secuencia arriba del sitio de Not I. Ambos salientes son compatibles con Not I. Los casetes de expresión de Not I son ligados en el pMON36582 digerido con Eag I purificado sobre gel, resultando en un vector con un solo sitio de Not I. El casete de expresión está por tanto disponible como un casete de Bsp120 I/Not I. En la figura 20, se muestra pMON36586 como un ejemplo de un casete de expresión para la homogentisato fitiltransferasa de *Arabidopsis* disponible como un casete de Bsp120 I/Not. Este vector se obtiene según lo descrito anteriormente.

La preparación de los casetes de expresión se realiza en un vector lanzadera, tal como pMON36586. Los casetes de expresión génica son liberados desde otros vectores lanzadera mediante digestos de Bsp120/Not I y se unen en un vector lanzadera tal como pMON36586, que ha sido digerido con Not I. El vector resultante alberga un casete de expresión génica adicional y un solo sitio de Not I. Este procedimiento se puede repetir según lo requerido. Una vez realizada la preparación del gen, se pueden liberar los casetes de expresión combinados mediante el digesto de Bsp120 I/Not I (pMON10098 (figura 37)). Entonces se purifica el fragmento resultante que porta los casetes de expresión y se une en un solo sitio de Not I de un vector binario. Alternativamente, se puede realizar la preparación de los casetes de expresión génica directamente en un vector binario (figura 21). Un vector binario se define por la presencia de las secuencias del borde izquierdo y derecho que son necesarias para la transferencia de ADN de *Agrobacterium* a las células vegetales. Todos los reactivos químicos y las enzimas para esta manipulación son de grado molecular. Estos reactivos y estas enzimas se utilizan según las instrucciones del suministrador. Se usan técnicas de clonación molecular estándar.

Se representan varios ejemplos de constructos binarios vegetales, sus componentes y los mapas plasmídicos. Los ejemplos representados que contienen las combinaciones de *tyrA* con otros genes de interés para el diseño de la ruta del tocoferol se enumeran en las figuras 18a y 18b. Los componentes de estos constructos también se proporcionan en la tabla 4. Los mapas vectoriales mostrados en las Figuras 22-27 representan diversos constructos.

TABLA 4

Lista de vectores binarios por ser transformados en Arabidopsis thaliana para diseñar la biosíntesis del tocoferol

pMon n.º	Combinación de genes	Elementos genéticos
36596	HPPD _{At} /tyrA	5' de napina y 3' de napina; CTP1 y 2
36597	HPPD _{At} /tyrA/GGH _{Syn}	5' de napina y 3' de napina; CTP1 y 2;
77601	HPPD _{At} /tyrA/GGH _{At} /ATPT2	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
77602	HPPD _{At} /tyrA/GGH _{Syn} /ATPT2	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
66657	HPPD _{At} /tyrA/GGH _{Syn} /slr1736	5' de napina y 3' de napina; CTP1 y 2; CTP

ES 2 318 004 T3

		nativos
5	66659 HPPD _{At} /tyrA/GGH _{At} /ATPT2/TMT2	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
10	HPPD _{At} /tyrA/GGH _{Syn} /ATPT2/MT1	5' de napina y 3' de napina; cTP1 y 2; CTP nativos
15	HPPD _{At} /tyrA/ATP2/TMT2	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
20	HPPD _{At} /tyrA/GGH _{At} /ATPT2/MT1/DxS	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
25	HPPD _{At} /tyrA/GGH _{At} /ATPT2/DxS _{E.coli}	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
30	HPPD _{At} /tyrA/GGH _{At} /slr1736/MT1/GGPPS _{At}	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
35	HPPD _{At} /tyrA/ATPT2/GGH _{At} /DxR _{At}	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
40	HPPD _{At} /tyrA/GGH _{At} /slr1736/Cyclasa _{Syn} /MT 1	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
45	HPPD _{At} /tyrA/GGH _{Syn} /slr1736/Cyclasa _{Syn}	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
50	HPPD _{At} /tyrA/slrl736/Cyclasa _{Syn}	5' de napina y 3' de napina; cTP1 y 2; CTP nativos
55		
60		
65		

	HPPD _{At} /tyrA/GGH _{At} /ATPT2/GMT _{At}	5' de napina y 3' de napina; CTP1 y 2; CTP nativos
--	---	--

Ejemplo 13

Construcción de vectores codificantes de múltiples enzimas

Este ejemplo expone el uso de una preferato deshidrogenasa (tyrA) tal como tyrA de *Erwinia herbicola* en combinación con otras enzimas clave en la ruta biosintética del tocoferol para aumentar la producción de tocoferol en semillas de plantas transgénicas tales como semillas de *Arabidopsis thaliana*. Las enzimas combinadas con tyrA incluyen ATP2, *p*-hidroxifenilpiruvato dioxigenasa (HPPD_{Arabidopsis}) y geranilgeranilpifosfato sintasa (GGPPS_{Arabidopsis}) de *Arabidopsis thaliana*. Además, se analizó tyrA también en combinación con ATP2 y un transportador de adenilato putativo (AANT1_{Arabidopsis}) de *Arabidopsis thaliana*.

La construcción de un vector de dos genes que alberga los casetes de expresión de tyrA y ATP2 de una semilla específica se realiza según lo siguiente. Se somete ADN plasmídico purificado de pMON36520 (figura 38) a un digesto de KpnI parcial y se une con un fragmento de Kpn I purificado sobre gel de 4,2 kpb aislado de pMON43853 (figura 39). El inserto de 4,2 kb de pMON43853 contiene el casete de expresión del gen PPT (pNapina::ATPT2::3' de napina). Se usa el vector binario vegetal resultante pMON69907 (figura 12) para la transformación de *Arabidopsis thaliana* para analizar el efecto combinatorio de la expresión de tyrA y ATP2 de una semilla específica.

Para aumentar más la biosíntesis del tocoferol, se expresa HPPD_{Arabidopsis} además de tyrA y ATP2 en semilla de *Arabidopsis thaliana*. Esto se realizó añadiendo un casete de expresión de semilla específica para HPPD_{Arabidopsis} a pMON69907 dando como resultado la formación de pMON69909. El vector binario pMON69909 se construyó digiriendo parcialmente pMON69907 con KpnI. El pMON69907 con un solo corte de KpnI se purifica sobre gel y se une con un inserto de KpnI/KpnI de 4,6 kb aislado de pMON36525 (figura 40). El inserto KpnI/KpnI de 4,6 kb de pMON36525 contiene el casete de expresión del gen HPPD, pNapina::CTP2::HPPD_{Arabidopsis}::3' de napina para dirigir la expresión dirigida a plastos de semilla específica de HPPD. El CTP2 es una señal diana cloroplástica procedente del gen 5-enolpiruvilsikimato-3-fosfato sintasa (EPSPS) de *Arabidopsis*.

Se construye el vector binario pMON69915 (figura 35) para analizar efecto de las combinaciones de los tres genes, tyrA, ATP2 y GGPPS_{Arabidopsis} sobre la producción de tocoferoles de las semillas. Se digirió parcialmente el vector pMON69907 con KpnI. El pMON69907 con un solo corte de KpnI se purifica sobre gel y se une con un fragmento de KpnI/KpnI de 4,3 kb purificado sobre gel de pMON43861 (figura 41) para crear pMON69915. El fragmento de KpnI de pMON43861 contiene el casete de expresión génica para la geranilgeranildifosfato sintasa de *Arabidopsis* procedente de *Arabidopsis thaliana* (pNapina::GGPPS_{Arabidopsis}::3' de napina). Se identifica el ADNc de GGPPS como un clon de EST, buscando en una base de datos de EST con la información secuencial disponible en la bibliografía (Okada *et al.*, *Plant Physiol.* 122: 1045-1056 (2000)). Se digiere el clon de EST con NcoI y se redondean los extremos rellenando el saliente de 5' con el fragmento de Klenow. Posteriormente, se digiere el clon con BamHI y se escinde el fragmento de ADNc. Se une el fragmento de ADNc romo/BamHI purificado sobre gel con vector pCGN7770 digerido con BglII/SalI (con los extremos redondeados con SalI) (figura 42) para crear pMON43861.

Se construye el vector binario vegetal pMON69919 (figura 36) para analizar la expresión combinada de tyrA, ATP2 y AANT1_{Arabidopsis} sobre los niveles de tocoferol de las semillas. Para generar este vector, se digiere parcialmente pMON69907 con KpnI. El pMON69907 con un solo corte de KpnI se purifica sobre gel y se une con un fragmento de KpnI/KpnI purificado sobre gel de 4,2 kb de pMON69911 (figura 43). El fragmento de 4,2 kb contiene un casete de expresión de semilla específica para el transportador de adenilato de *Arabidopsis* AANT1 (pNapina::AANT1_{Arabidopsis}::3' de napina). Se genera pMON69911 escindiendo el fragmento de AANT1 de pCGN11301 (figura 44) con SalI y PstI (se redondea el sitio de PstI retirando el saliente de 3' con Klenow) y luego se une a pCGN7770 digerido con SalI/XhoI (con extremos redondeados con XhoI).

Se identifican varios clones de longitud completa en las bases de datos de EST usando la secuencia parcial publicada de AANT1 (Saint-Guily *et al.*, *Plant Physiol.* 100 (2): 1069-1071 (1992)). La región codificante de AANT1 es amplificada por PCR usando los cebadores, AANT1F 5'-GGATCCGCGGCCGACCA**TGGTTGATCAAGTT**CAGCA (SEC ID N.º 11) y AANT1R 5'-GAGCTCCTGCAGGAAGCT**TTTAGGCACCTCCTGATCCGT**-3' (SEC ID N.º 12). El sitio de *NotI* (subrayado) se coloca secuencia arriba del codón de iniciación (cursivas) en el cebador AANT1F, mientras que el sitio de *Sse8387I* (subrayado) fue colocado secuencia abajo del codón de terminación (cursivas) en AANT1R. Primero se clonan los productos de la PCR en pCR2.1 y se verifican los insertos mediante la secuenciación de ambas cadenas. Posteriormente, se insertan los fragmentos de *NotI*/*Sse8387I* en los sitios de *NotI*/*Sse8387I* del casete de expresión de la napina en pCGN9979 en una orientación sentido con respecto al promotor de napina para generar pCGN11301. Los constructos de expresión vegetales se usan para la transformación de *Arabidopsis thaliana* mediante una transformación mediada por *Agrobacterium* según lo descrito anteriormente.

Ejemplo 14

Expresión de vectores codificantes de múltiples enzimas en plantas

- 5 Se transforman plantas de *Arabidopsis thaliana* con vectores del ejemplo 13 usando la técnica de transformación ofrecida en el ejemplo 8. Los resultados para pMON69909 se proporcionan en las Figuras 14, 15, 16, 17 y 31-34. En la siguiente tabla se ofrecen más resultados, así como en las figuras 75 y 76, que muestran los niveles de tocoferol, tocotrienol, homogentisato y 2-metilfitilplastoquinol para las plantas transformadas que tienen pMON69909.

10

15

20

25

30

35

40

45

50

55

Constructo	Elementos genéticos	Aumento de tocoferol/tocotrienol	% de tocoferol	% de tocotrienol
pMON69907	pNapina::HPT _{Arabidopsis} ::3' de napina/pNapina::CTP1::tyrA _{Erwinia herbicola} ::3' de napina	2,4 veces mayor	91	9
pMON69909	pNapina::HPT _{Arabidopsis} ::3' de napina/pNapina::CTP 1::tyrA _{Erwinia herbicola} ::3' de napina/ pNapina::CTP2::HPPD _{Arabidopsis} ::3' de napina	5 veces mayor	38	62
pMON69915	pNapina: HPT _{Arabidopsis} : 3' de napina/pNapina::CTP1::tyrA _{E. herbicola} ::3' de napina/ pNapina::GGPPS _{Arabidopsis} ::3' de napina	2,9 veces mayor	86	14
pMON69919	pNapina::HPT _{Arabidopsis} ::3' de napina/pNapina::CTP 1::tyrA _{E. herbicola} ::3' de napina /pNapina::AANT1 _{Arabidopsis} ::3' de napina	3 veces mayor	89	11

Ejemplo 15

60 *Expresión de constructos en soja*

- Este ejemplo describe el procedimiento implicado en la preparación de vectores binarios vegetales para analizar tyrA solo y en combinación con otras enzimas clave en la ruta biosintética del tocoferol para aumentar la producción del tocoferol en semillas de *Glycine max* transgénicas. La siguiente tabla describe los vectores binarios vegetales preparados para la transformación de *G. max* con su respectivo gen de los casetes de expresión de interés para la expresión de semilla específica de los transgenes.

Lista de los constructos transformados en *G. max*

Número de constructo	Elementos genéticos
pMON36575	p7Sa':::CTP1::tyrA _{E. herbicola} ::3' de E9
pMON69924	p7Sa':::CTP2::HPPD _{Arabidopsis} ::3' de E9/p7Sa':::CTP1::tyrA _{E. herbicola} ::3' de E9
pMON69943	p7Sa':::CTP2::HPPD _{Arabidopsis} ::3' de E9/p7Sa':::CTP1::tyrA _{E. herbicola} ::3' de E9/pArcelina-5::CTP1::slr1736::3' de arcelina
pMON69945	p7Sa':::CTP2::HPPD _{Arabidopsis} ::3' de E9/p7Sa':::CTP1::tyrA _{E. herbicola} ::3' de E9/pArcelina-5::CTP1::slr1736::3' de arcelina/ pNapina::GGH _{Arabidopsis} ::3' de napina

Se prepara el pMON36575 (Figura 46) uniendo el fragmento de NotI purificado sobre gel de 3 kb procedente de pMON38207R (Figura 47) en el sitio de NotI de pMON36571 (Figura 48) que contiene el casete de expresión p7Sa':::CTP1::tyrA_{E. herbicola}::3' de E9. El CTP1 codifica la secuencia señal diana cloroplástica de la subunidad pequeña de RUBISCO de *Arabidopsis*. El fragmento de NotI de 3 kb contiene el casete marcador seleccionable, pFMV::CTP2::CP4syn::3' de E9. El CTP2 codifica la secuencia señal diana cloroplástica procedente del gen 5-enolpiruvilsikimato-3-fosfato sintasa (EPSPS) de *Arabidopsis*. El CP4syn es un gen sintético de EPSPS. El vector pMON36575 se digiere además con HindIII para liberar el fragmento de 3 kb que contiene el casete de expresión p7Sa':::CTP1::tyrA_{E. herbicola}::3' de E9. Se redondean los extremos del fragmento rellenando los salientes de 5' con fragmento de Klenow, se purifica sobre gel y se une por el sitio de PmeI de pMON36576 (figura 49) que porte el casete de expresión de p7Sa':::CTP2::HPPD_{Arabidopsis}::3' de E9 para generar el pMON69924 (figura 50).

Se prepara el vector binario vegetal pMON69943 (figura 51) digiriendo pMON69929 (figura 52), que contiene el casete de expresión p7Sa':::CTP2::HPPD_{Arabidopsis}::3' de E9, con NotI y uniéndolo con el fragmento purificado sobre gel de 7,3 kb generado mediante la digestión de pMON69936 (figura 53) con Bsp120I y NotI. Este fragmento contiene los casetes de expresión de p7Sa':::CTP1::tyrA_{E. herbicola}::3' de E9 y pArcelina-5::CTP1::slr1736::3' de Arcelina. Se digiere más el vector pMON69943 con NotI y se une con un fragmento purificado sobre gel de Bsp120I/NotI de 4,5 kb procedente de pMON36592 (figura 54) para generar pMON69945 (figura 55). El fragmento de pMON36592 contiene el casete de expresión de pNapina::GGH_{Arabidopsis}::3' de napina.

Se transforma *Glucine max* con los vectores descritos según el procedimiento expuesto en el documento WO 00/61771 A3, páginas 99-100.

Ejemplo 16

Combinaciones de múltiples genes expresadas en canola

Todos los casetes de expresión génica usados para la expresión en canola se preparan como casetes de Not I que contenían el promotor de napina, un gen de interés y el terminador de napina. Los genes de interés están fundidos N-terminalmente con el péptido diana cloroplástico, a no ser que haya presente un péptido diana cloroplástico natural. Todas las combinaciones de genes se preparan en un solo vector de múltiples genes.

Para facilitar la construcción de los vectores de múltiples genes, se aíslan los casetes de expresión de Not I mediante la digestión con Not I desde pMON16602 (figura 56), pMON36525 (figura 57), pMON36520 (figura 38) y se clonan en pMON36582 digerido por Eag I y purificado sobre gel (figura 19), dando como resultado la formación de pMON58171 (figura 58) (casete de expresión slr1736), pMON58172 (figura 59) (casete de expresión de HPPD_{Arabidopsis})

ES 2 318 004 T3

y pMON58170 (figura 60) (casete de expresión de $\text{tyrA}_{\text{E herbicola}}$). Todos los casetes de expresión resultantes están flanqueados por Bsp120 I y Not I.

Se obtiene un casete de expresión dirigido por napina para la GGH de *Arabidopsis* mediante el aislamiento y la purificación sobre gel de un fragmento de Not I/Hind III de 3.191 pb procedente de pMON36591 (figura 61) y un fragmento de Not I/Hind III de 5.612 pb procedente de pMON36588 (figura 62). Se unen estos dos fragmentos purificados, dando como resultado la formación de pMON36592 (figura 63). Se digiere el vector pMON36592 con Bsp120I y Not I, se purifica sobre gel el casete de expresión de GGH y se une en pMON36582 purificado sobre gel y digerido con Eag I (figura 19), dando como resultado la formación de pMON58182 (figura 64).

Se obtienen vectores de múltiples genes combinando estos cuatro genes mediante la digestión de los vectores pMON58171, pMON58172, pMON58170 y pMON58182 con Bsp120I y Not I, seguida por la purificación sobre gel de los fragmentos más largos de cada constructo. Estos fragmentos contienen los casetes de expresión de slr1736, HPPD, tyrA y GGH, respectivamente. Se unen los casete de expresión de tyrA procedentes de pMON58170 en pMON58171 digerido con Not I y tratado con fosfatasa alcalina, dando como resultado la formación del vector de dos genes pMON58176 (figura 65) que contiene los casetes de expresión génica para tyrA y slr1736, respectivamente. Se vuelve a digerir este vector con Not I, se trata con fosfatasa alcalina y se une con el casete de expresión de HPPD procedente de pMON58172. El vector de tres genes resultante, pMON58183, (figura 66) contiene los casetes de expresión de HPPD, tyrA y slr1736. Se digiere también el vector pMON58183 con Bsp120 I, se trata con fosfatasa alcalina y se une con el casete de expresión de GGH purificado sobre gel (véase la purificación anterior), dando como resultado la formación de pMON58185 (figura 67).

Se prepara el vector lanzadera pMON36593 (figura 68) (que contiene los casetes de expresión de tyrA y HPPD) uniendo un casete de expresión de tyrA purificado sobre gel digerido con Bsp120I/Not I procedente de pMON36589 (figura 69) en pMON36590 digerido con NotI y tratado con fosfatasa alcalina (figura 70).

Se escinden los casetes de expresión de genes combinados mediante el digesto Bsp120 I/Not I procedente de pMON36593 (HPPD/tyrA), pMON58183 (HPPD/tyrA/sl1736) y pMON58185 (HPPD, tyrA, slr1736, GGH). Se purifican sobre gel estos casetes de expresión de genes combinados y se unen en pMON67162 digerido con Not I y tratado con fosfatasa alcalina (figura 71), dando como resultado la formación de los vectores binarios pMON58178 (figura 72), pMON58186 (figura 73) y pMON58188 (figura 74), respectivamente. Los tres últimos vectores binarios se usan para la transformación en canola.

REIVINDICACIONES

1. Una molécula de ácido nucleico sustancialmente purificada, molécula que comprende como componentes ligados operativamente: (A) una región promotora que funciona en una célula vegetal para generar la producción de una molécula de ARNm; (B) una molécula de ácido nucleico heteróloga a (A), siendo la molécula de ácido nucleico codificante de una enzima con actividades corismato mutasa y prefenato deshidrogenasa, y que comprende además (i) dos o más casetes de expresión, cada uno de los cuales expresa un miembro seleccionado entre fitilprenil transferasa procedente de *Synechocystis* (*slr1736*), fitilprenil transferasa procedente de *Arabidopsis thaliana* (*ATPT2*), 1-desoxixilulosa-5-fosfato sintasa (*dxs*), 1-desoxixilulosa-5-fosfato reductoisomerasa (*dxr*), geranilgeranil reductasa (*GGH*), geranilgeranil pirofosfato sintasa (*GGPPS*), *p*-hidroxifenilpiruvato dioxigenasa (*HPPD*), metil transferasa 1 (*MT1*), tocoferol metil transferasa 2 (*TMT2*), gamma metil transferasa (*GMT*), transportador de adenilato (*AANT1*), tocoferol ciclasa (*slr 1737*) y un constructo antisentido para la ácido homogentísico dioxigenasa, en la que la enzima codificada por la molécula de ácido nucleico heteróloga presenta la secuencia de aminoácidos de SEC ID N.º 2 o SEC ID N.º 4.
2. La molécula de ácido nucleico de la reivindicación 1, molécula de ácido nucleico que comprende la secuencia de nucleótidos de SEC ID N.º 1 o SEC ID N.º 3.
3. La molécula de ácido nucleico de la reivindicación 1, molécula de ácido nucleico que comprende además una secuencia no traducida en 3' que funciona en dicha célula vegetal para provocar la terminación de la transcripción y la adición de ribonucleótidos poliadenilados en el extremo 3' de la molécula de ARNm.
4. La molécula de ácido nucleico de la reivindicación 1, en la que la molécula de ácido nucleico heteróloga comprende además un casete de expresión que expresa fitil prenilttransferasa, en concreto, un casete de expresión que expresa hidroxifenilpiruvato deshidrogenada.
5. La molécula de ácido nucleico de la reivindicación 1, en la que dichos dos o más casetes de expresión expresan *HPPD* y bien *slr1736* o la *ATPT2*.
6. La molécula de ácido nucleico de la reivindicación 1, en la que la molécula de ácido nucleico heteróloga es una molécula de ácido nucleico bicatenaria.
7. Una planta transformada que tiene la molécula de ácido nucleico de la reivindicación 1 o de la reivindicación 3.
8. La planta transformada de la reivindicación 7, en la que dicha molécula de ácido nucleico comprende además un casete de expresión que expresa fitil prenilttransferasa, en concreto, un casete de expresión que expresa hidroxifenilpiruvato deshidrogenada.
9. La planta transformada de la reivindicación 7, planta que se selecciona entre canola, maíz, *Arabidopsis*, *Brassica campestris*, *Brassica napus*, soja, crambe, mostaza, semilla de ricino, cacahuete, sésamo, semilla de algodón, semilla de lino, cártamo, aceite de palma, lino y girasol, en concreto, siendo dicha planta soja, canola o *Brassica napus*.
10. La planta transformada de la reivindicación 7, planta que presenta niveles aumentados de tocoferol, tocotrienol, α -tocopherol, α -tocotrienol, γ -tocopherol y/o γ -tocotrienol en comparación con una planta con un contexto genético similar pero que carece de dicha molécula de ácido nucleico de la reivindicación 3, en concreto, en la que dichos niveles están aumentados en al menos aproximadamente el 25%, en al menos aproximadamente el 250% o en al menos aproximadamente el 2.500% en comparación con la planta con un contexto genético similar pero que carece de dicha molécula de ácido nucleico de la reivindicación 3.
11. La planta transformada de la reivindicación 7, en la que la molécula de ácido nucleico de la reivindicación 3 comprende además una secuencia dirigida a plastos, en la que dicha secuencia dirigida a plastos está ligada operativamente a la molécula de ácido nucleico de la reivindicación 3 para hacer que un transcripto de dicha molécula de ácido nucleico heteróloga codifique además una secuencia peptídica dirigida a plastos ligada operativamente a dicha secuencia de aminoácidos.
12. Un procedimiento de producción de una planta que tiene niveles aumentados de tocoferol, procedimiento que comprende las etapas de:
 - (A) transformar dicha planta con la molécula de ácido nucleico de la reivindicación 1; y
 - (B) cultivar dicha planta.
13. El procedimiento de la reivindicación 12, en el que dicha molécula de ácido nucleico comprende además un casete de expresión que expresa fitil prenilttransferasa, en concreto, un casete de expresión que expresa hidroxifenilpiruvato deshidrogenada.
14. El procedimiento de la reivindicación 12, en el que dichos dos o más casetes de expresión expresan *HPPD* y bien *slr1736* o *ATPT2*.

ES 2 318 004 T3

15. El procedimiento de la reivindicación 12, en el que dicha molécula de ácido nucleico está ligada a una secuencia no traducida en 3' que funciona en la planta para provocar la terminación de la transcripción y la adición de ribonucleótidos poliadenilados en el extremo 3' de una molécula de ARNm, y en el que la expresión de dicha molécula de ácido nucleico resulta en la sobre-expresión de dicha proteína.

16. El procedimiento de la reivindicación 12, en el que dicha planta se selecciona entre canola, maíz, *Arabidopsis*, *Brassica campestris*, *Brassica napus*, soja, crambe, mostaza, semilla de ricino, cacahuete, sésamo, semilla de algodón, semilla de lino, cártamo, aceite de palma, lino y girasol, en concreto, siendo dicha planta soja, canola o *Brassica napus*.

17. El procedimiento de la reivindicación 12, en el que dicha planta presenta niveles aumentados de tocoferol, α -tocoferol, γ -tocoferol y/o tocotrienol en comparación con una planta con un contexto genético similar pero que carece de dicha molécula de ácido nucleico.

18. Un procedimiento para reducir los niveles de tocoferol en una planta, procedimiento que comprende las etapas de:

(A) transformar dicha planta con la molécula de ácido nucleico de la reivindicación 3, teniendo la molécula de ácido nucleico una cadena transcrita y una cadena no transcrita, en la que dicha cadena transcrita es complementaria a la molécula de ácido nucleico que tiene la secuencia de ácido nucleico de SEC ID N.º 1 o SEC ID N.º 3; y

(B) cultivar dicha planta transformada.

19. Una célula que comprende la molécula de ácido nucleico de cualquiera de las reivindicaciones 1, 4 ó 5.

20. La célula de la reivindicación 19, célula que es una célula bacteriana o una célula de alga verde azulada.

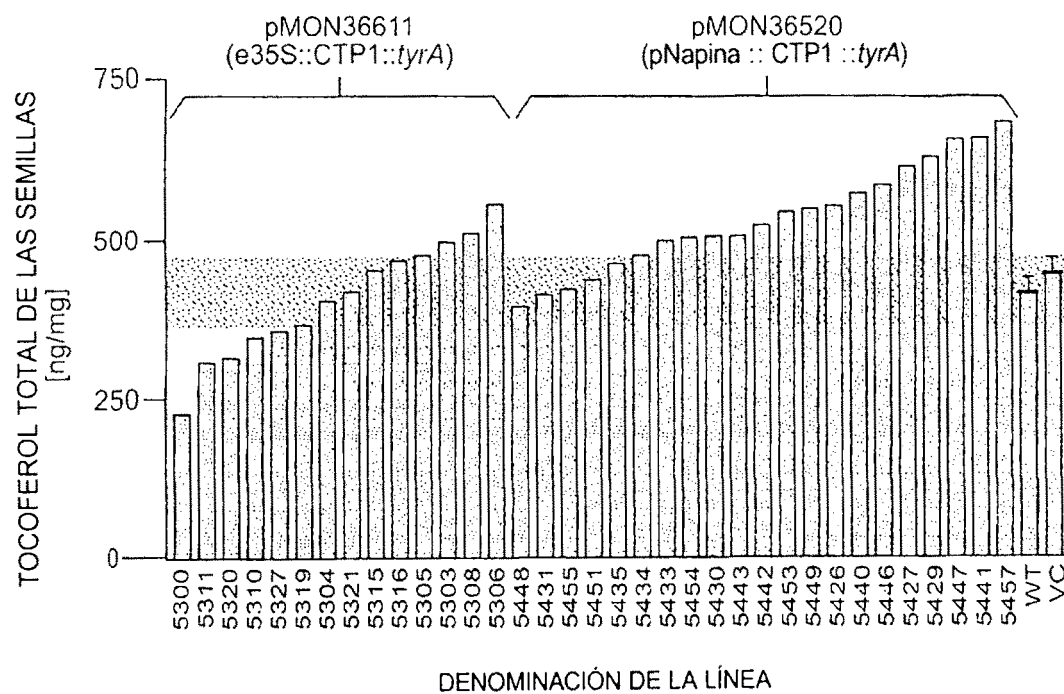


FIG. 1

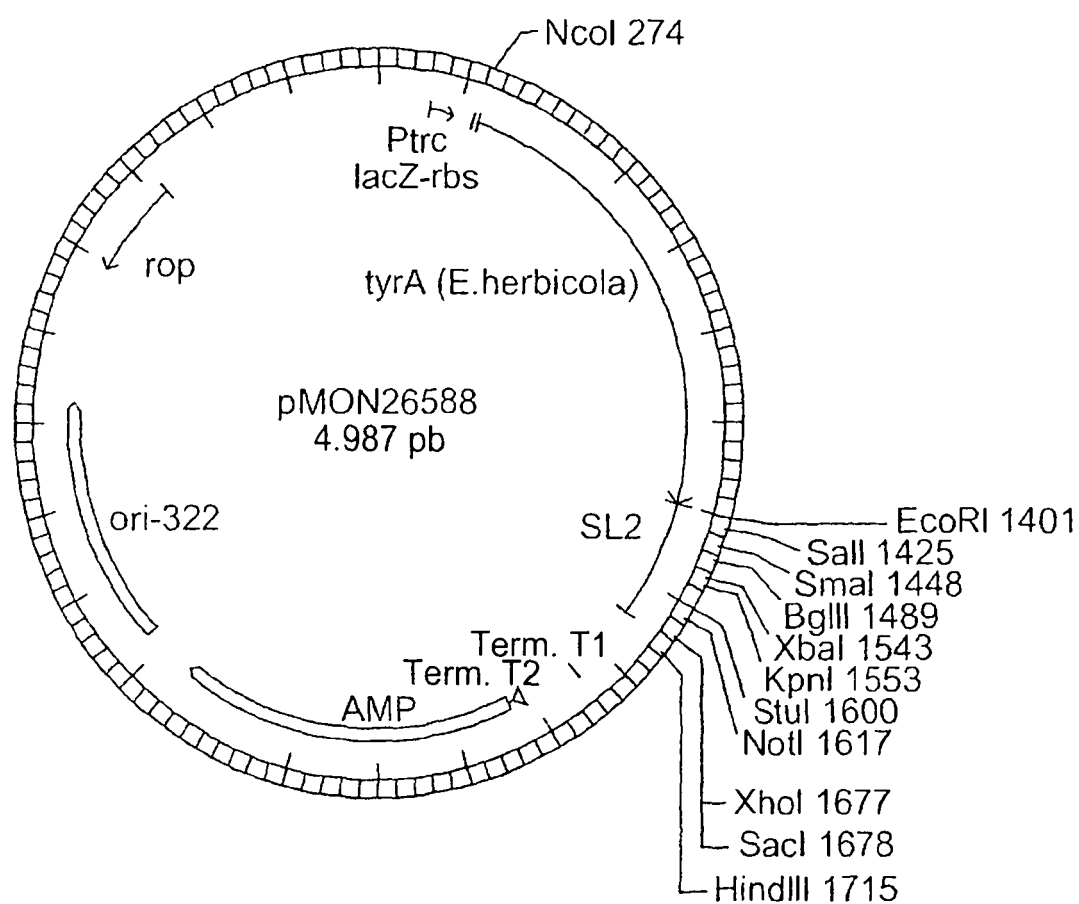


FIG. 2

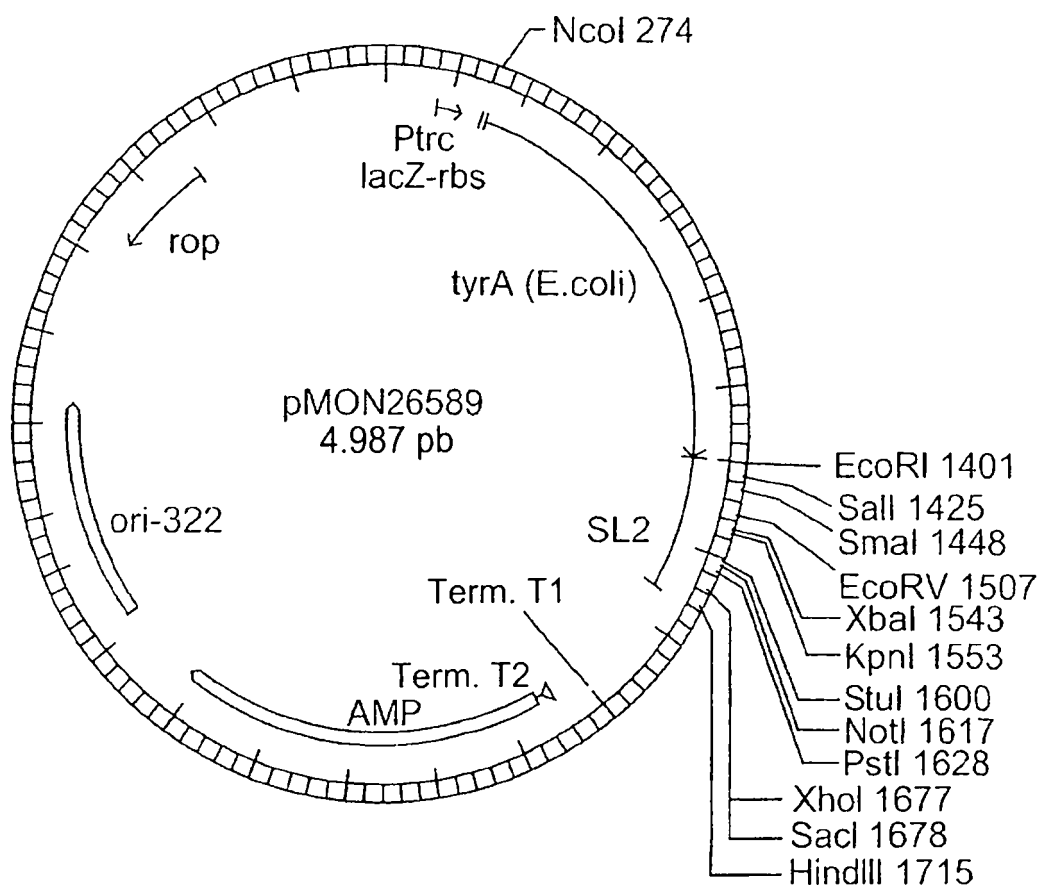


FIG. 3

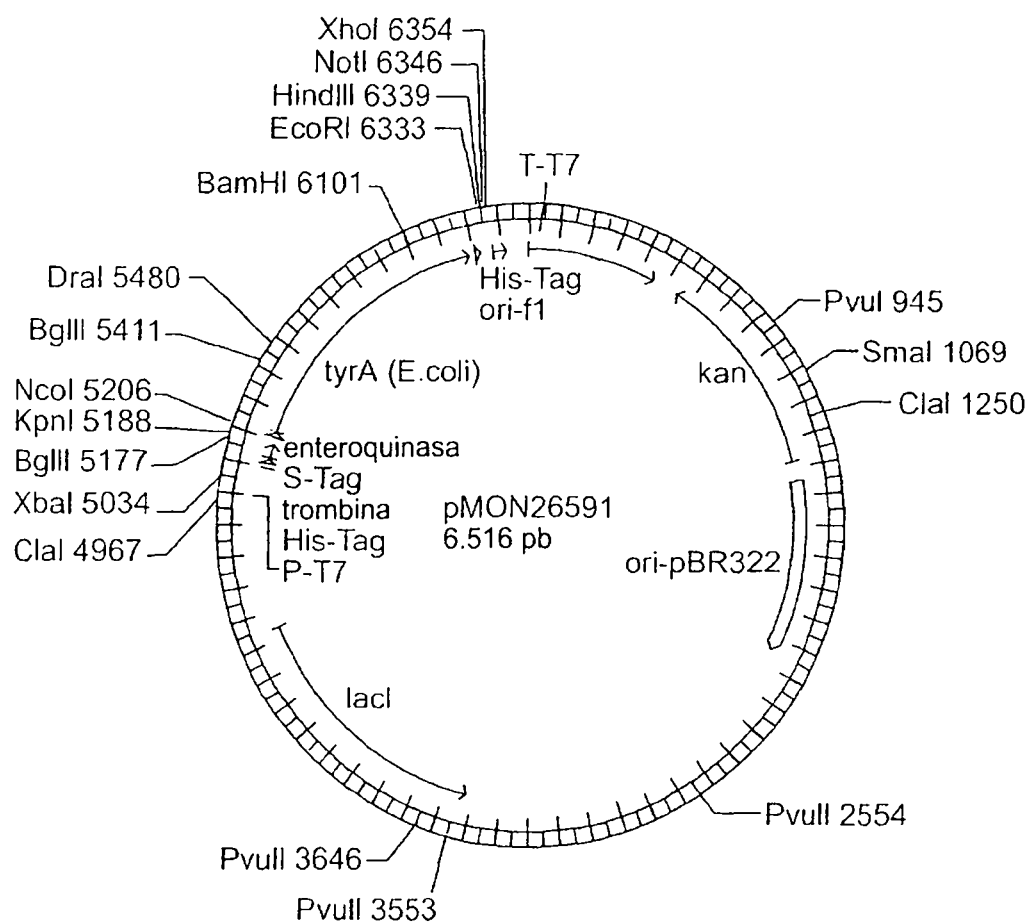


FIG. 4

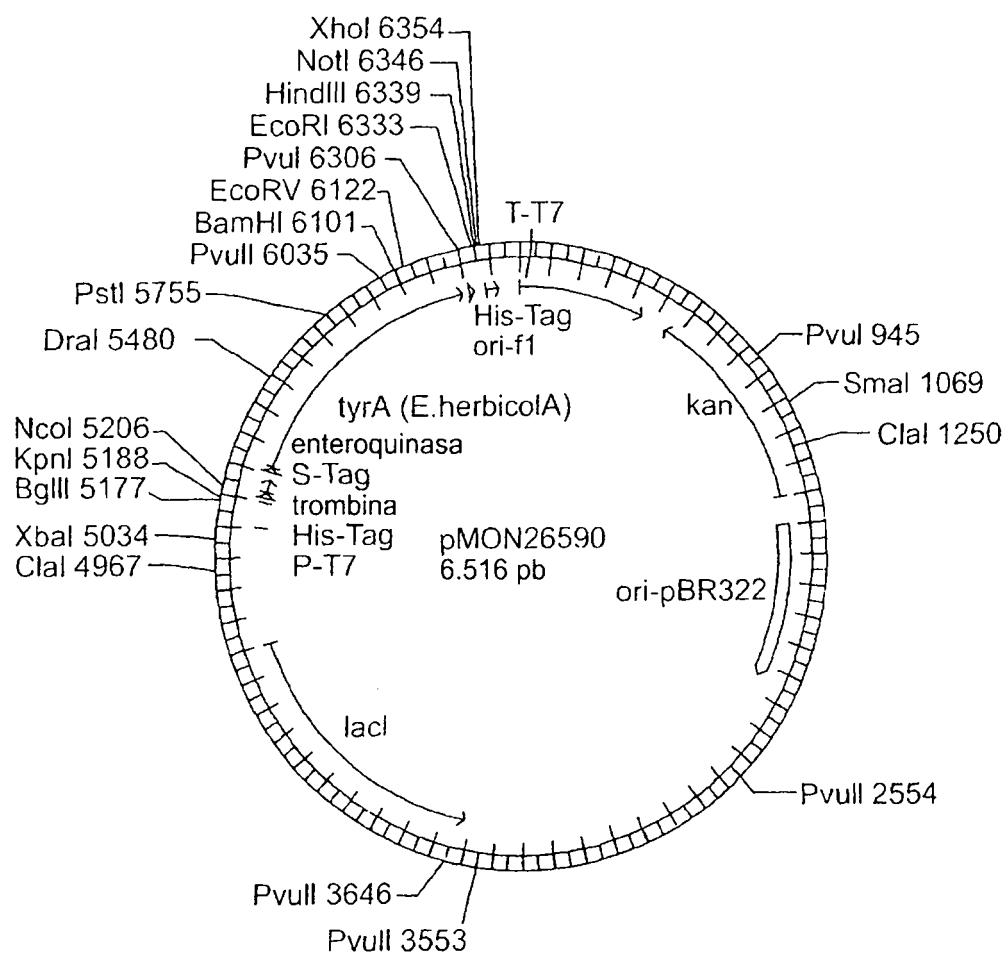


FIG. 5

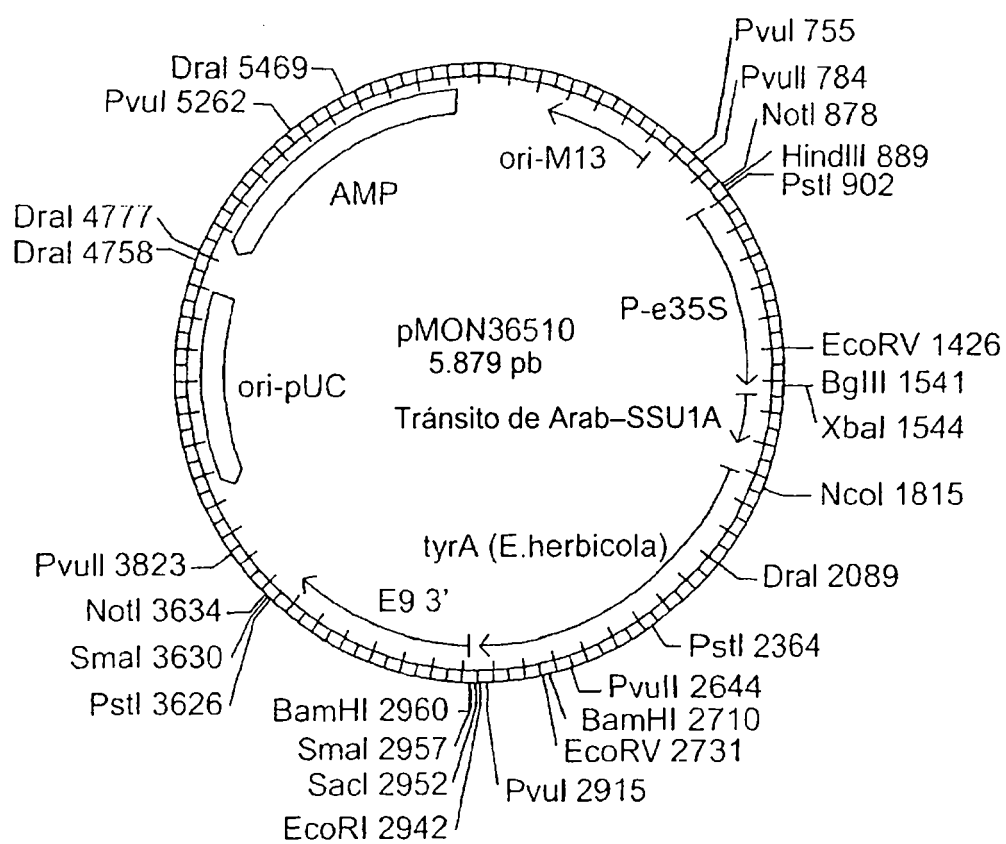


FIG. 6

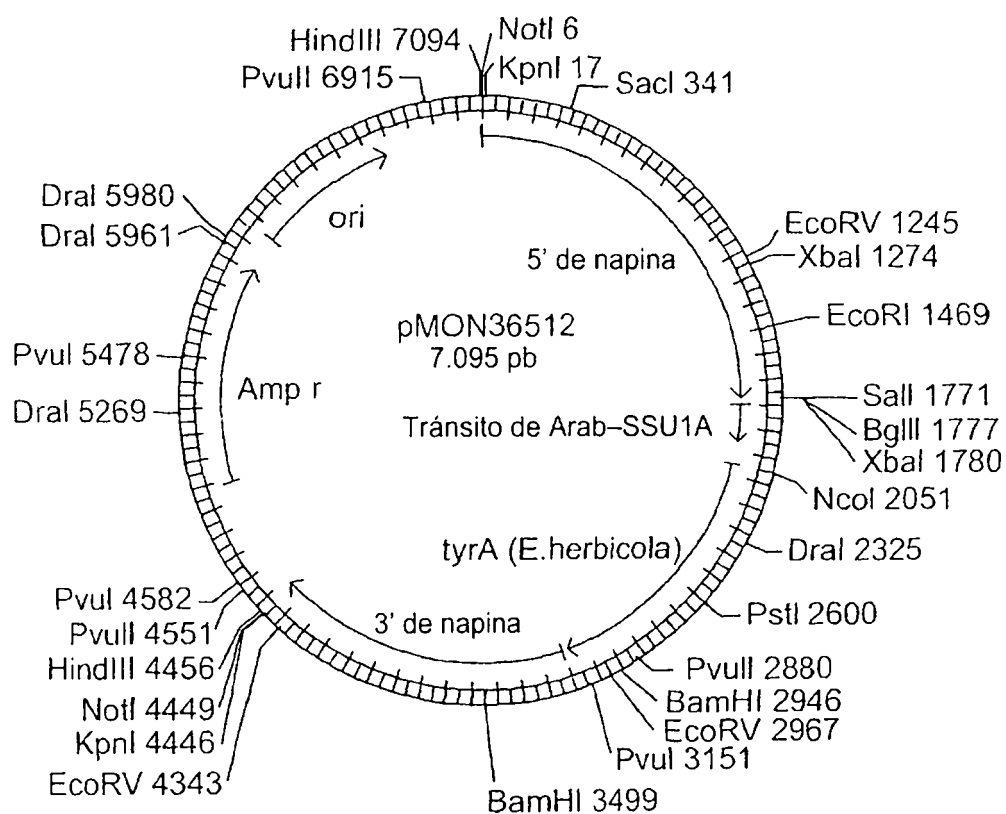


FIG. 7

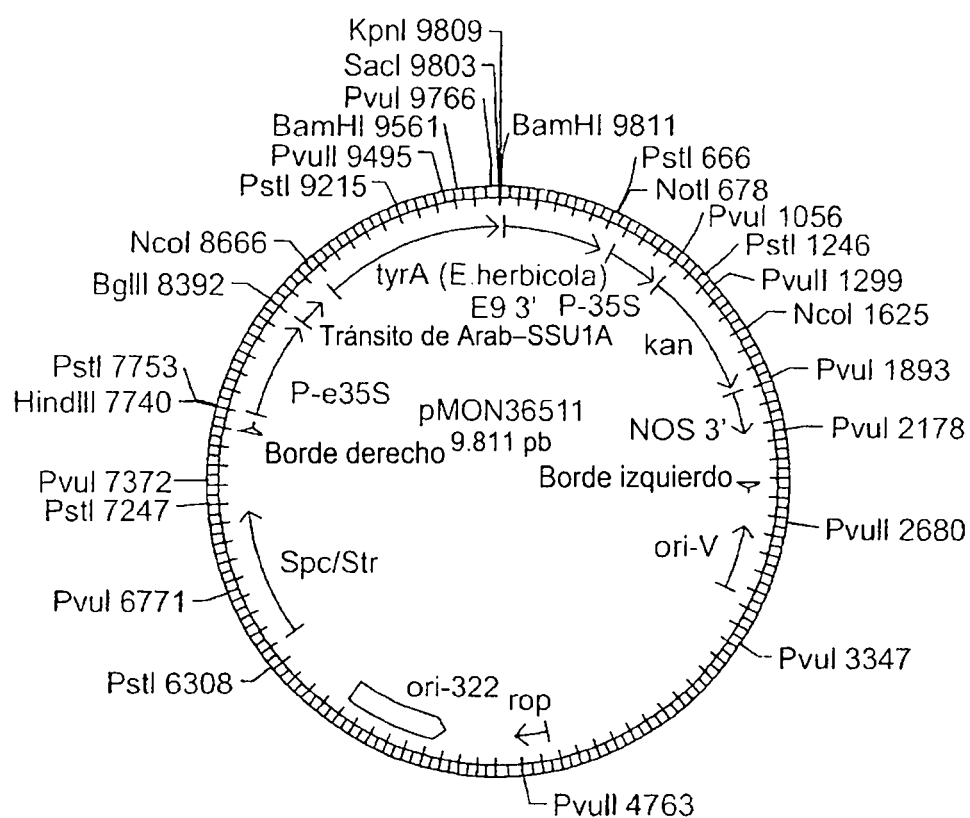


FIG. 8

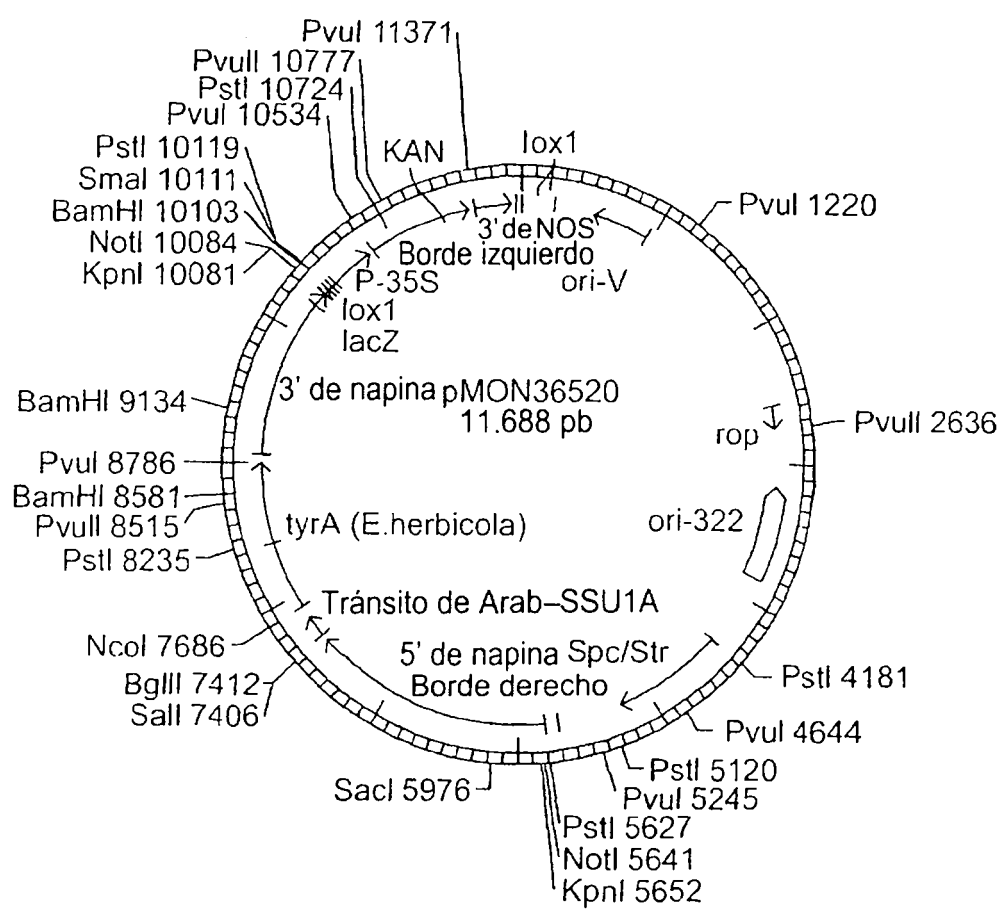


FIG. 9

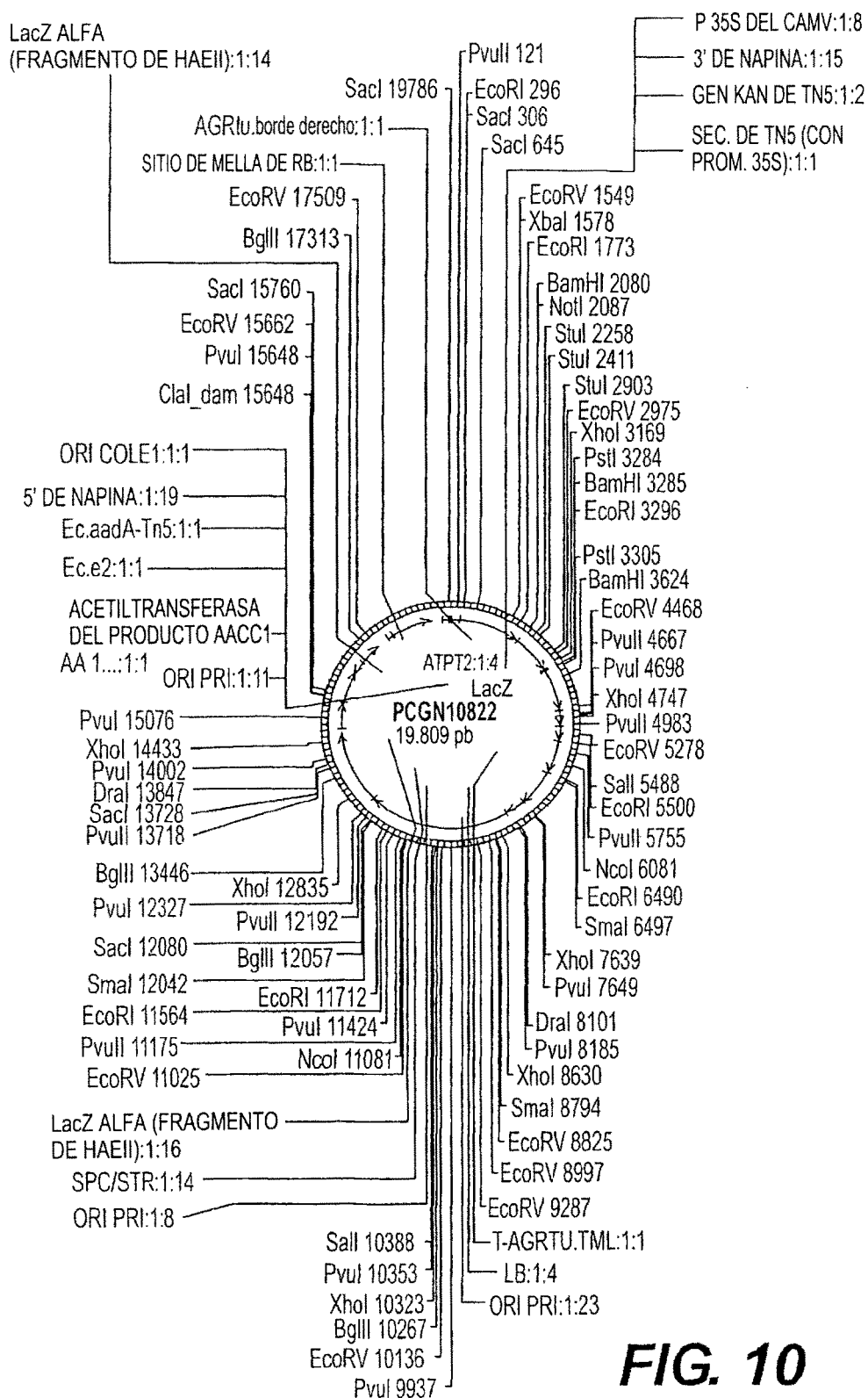


FIG. 10

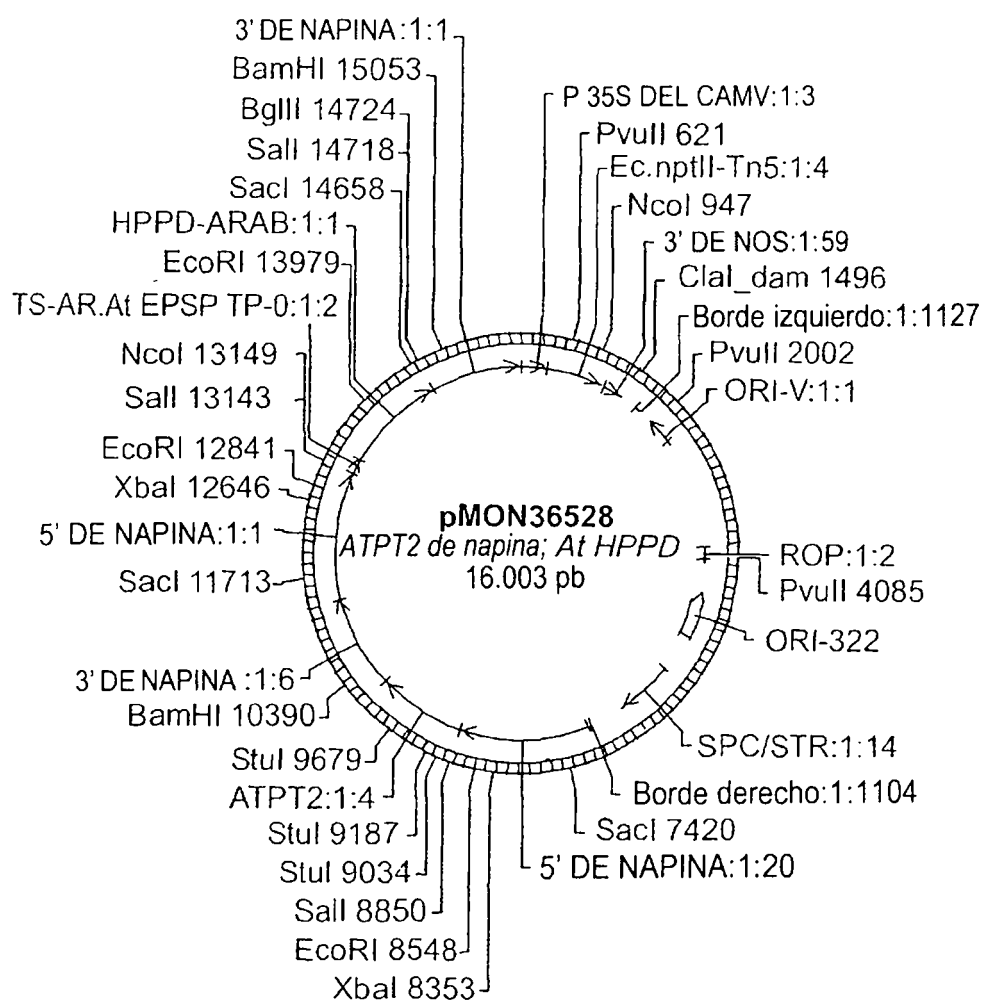


FIG. 11

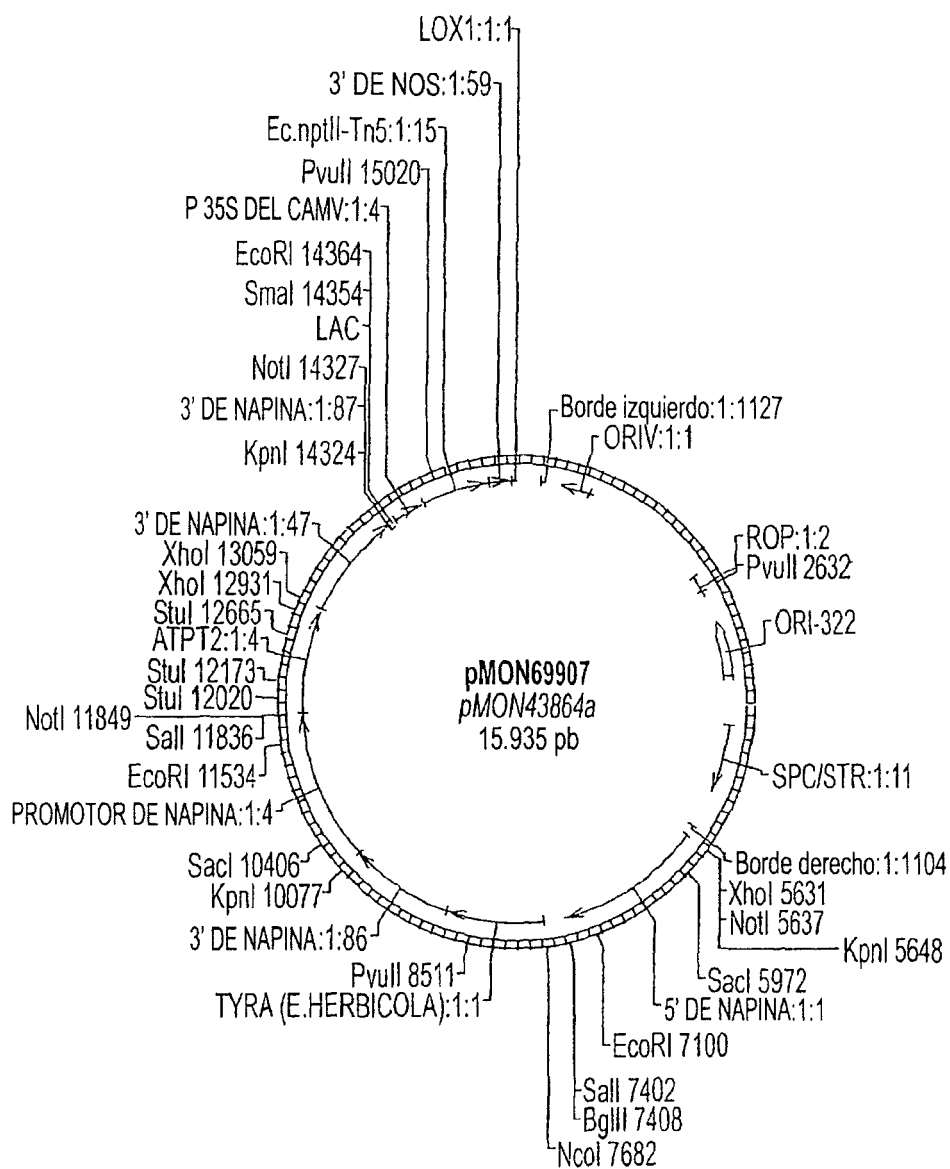


FIG. 12

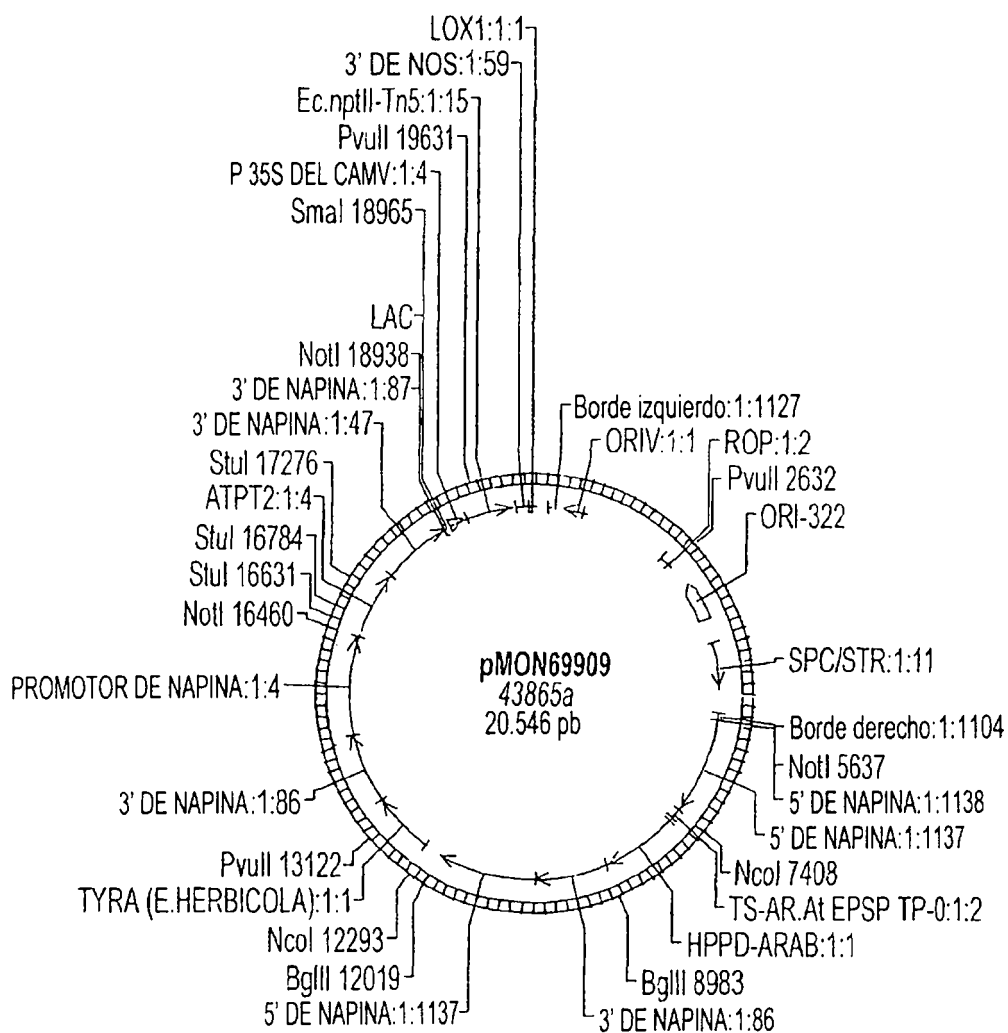


FIG. 13

pMON 69907 [pnapina::ATPT2::3' de napina::pnapina::CTP1::tyrA::3' de napina]

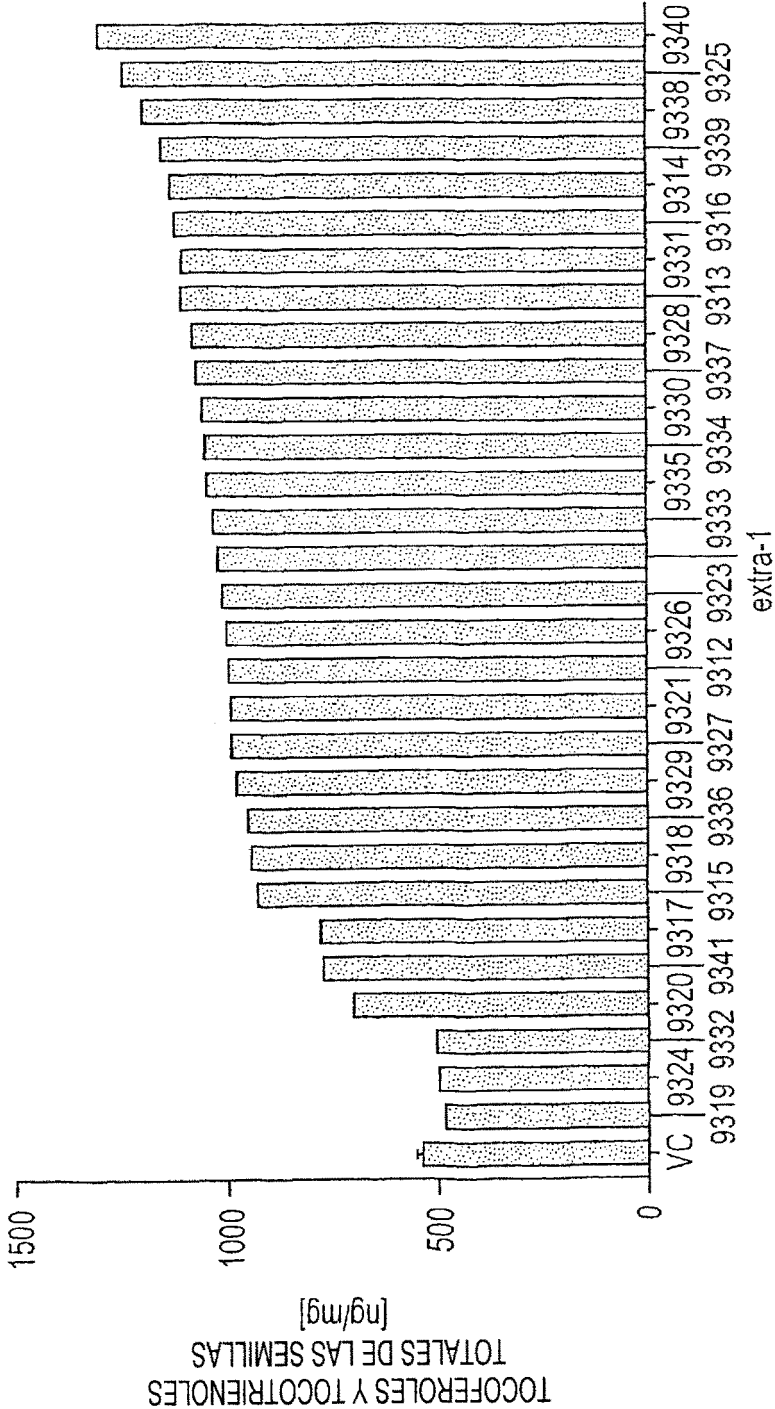


FIG. 14 DENOMINACIÓN DE LA LÍNEA

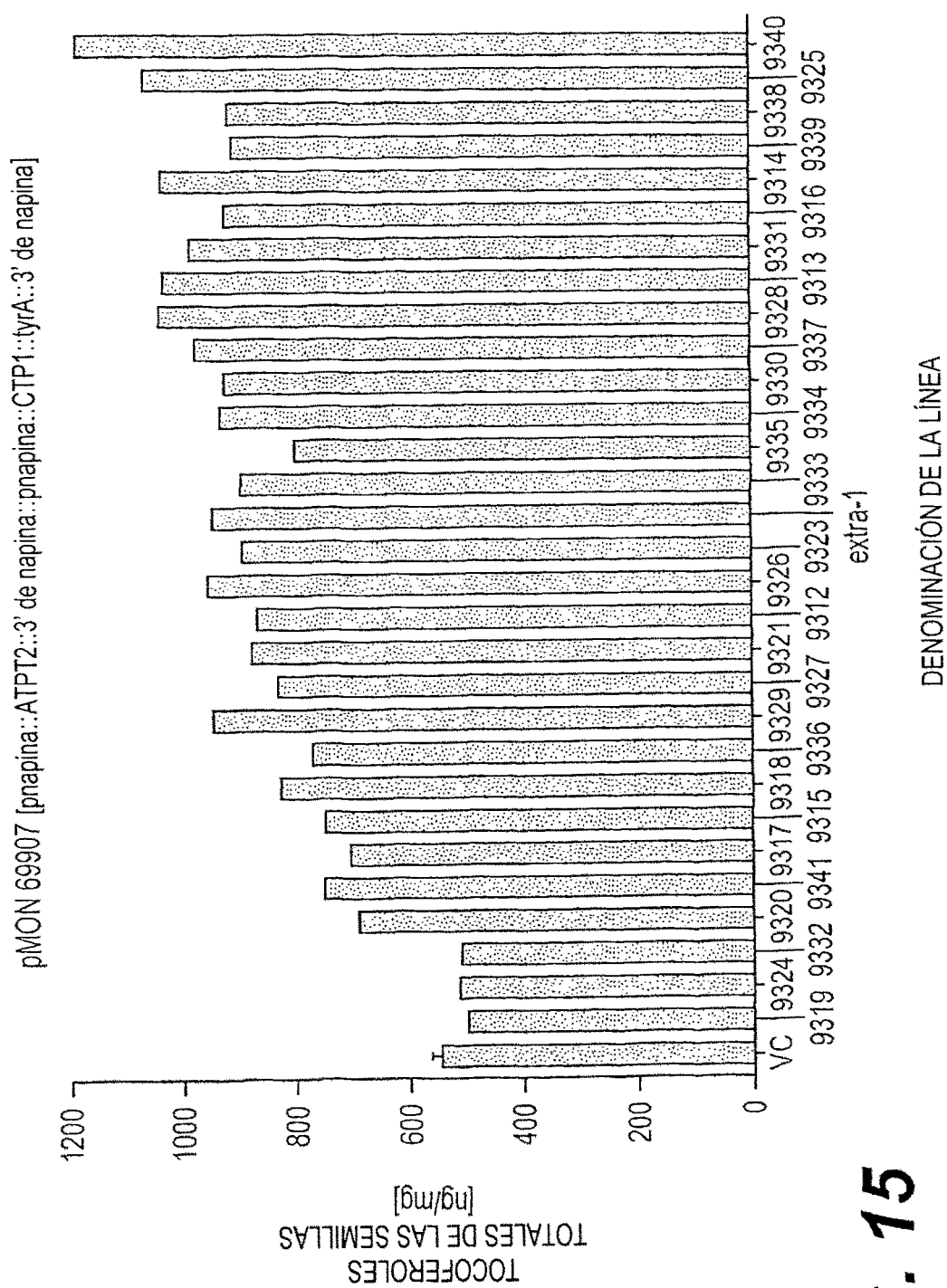


FIG. 15

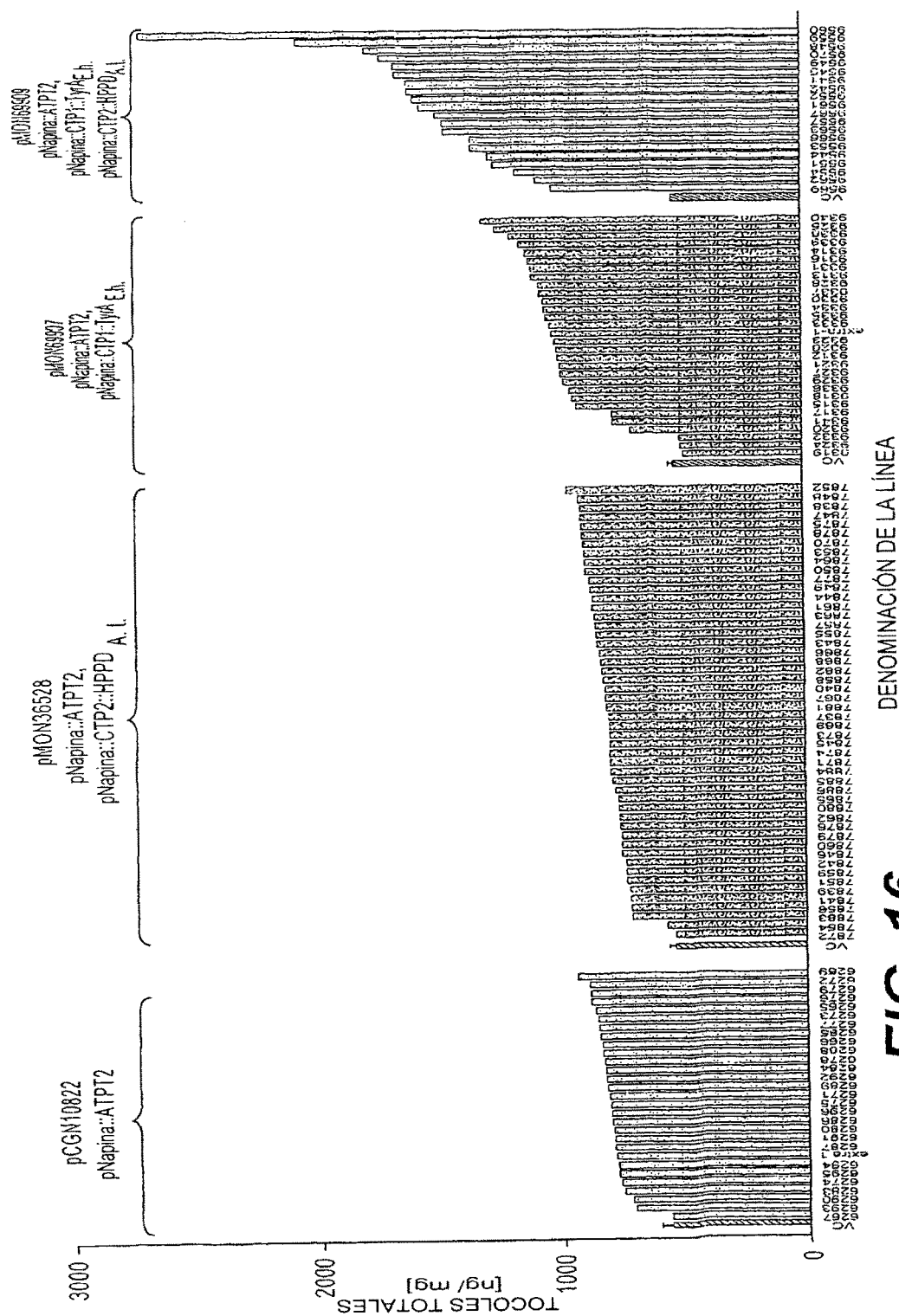


FIG. 16

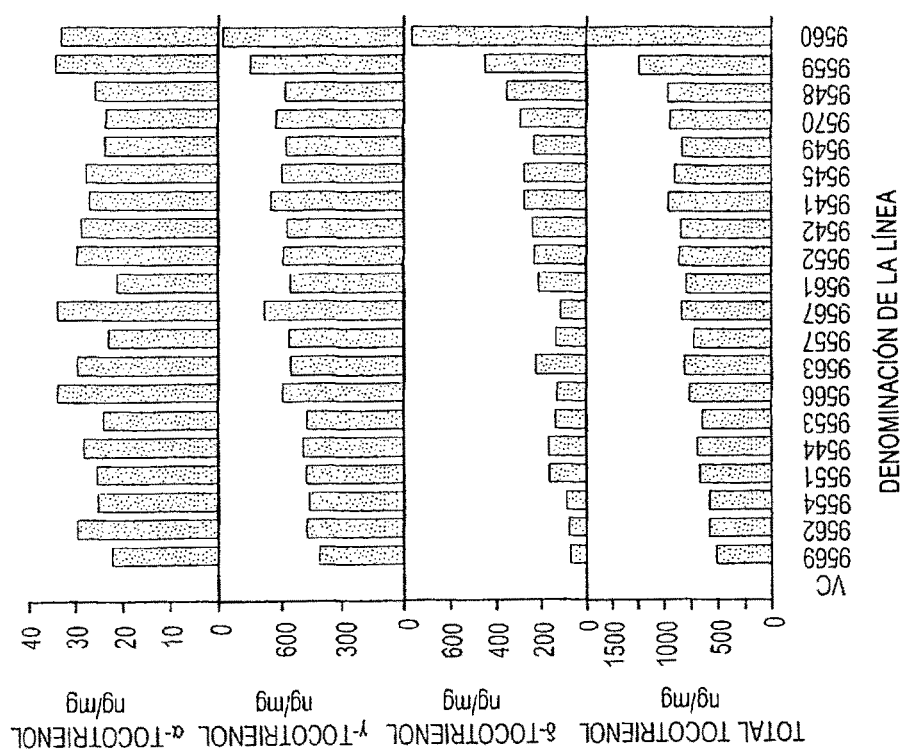
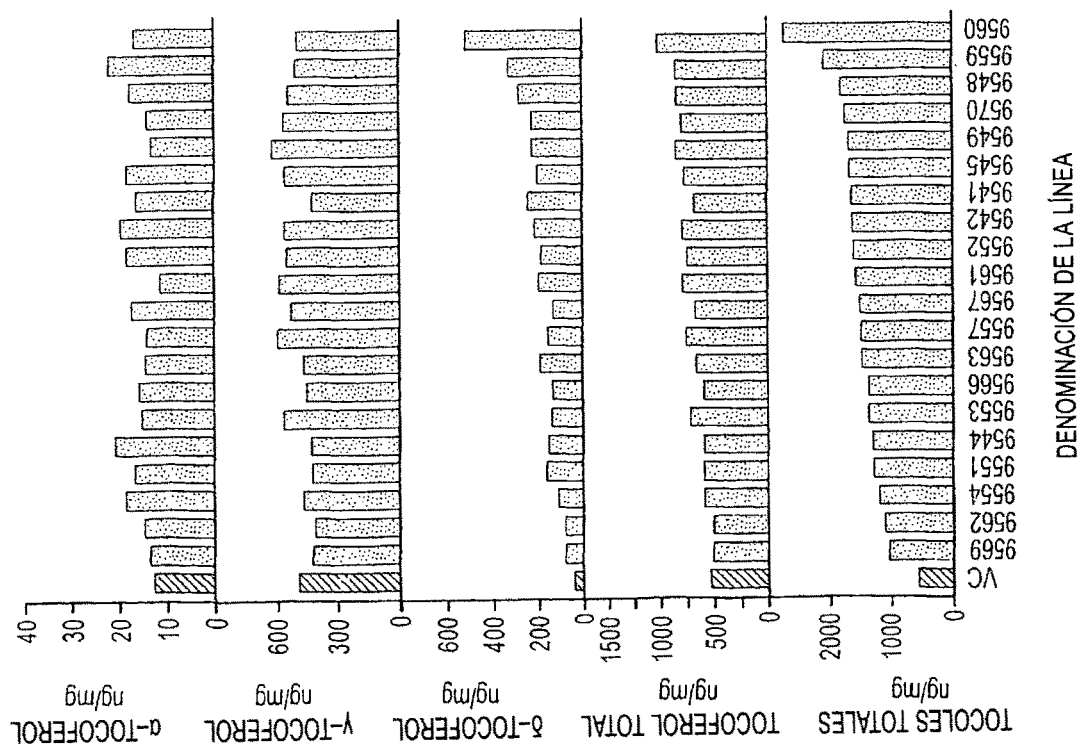


FIG. 17



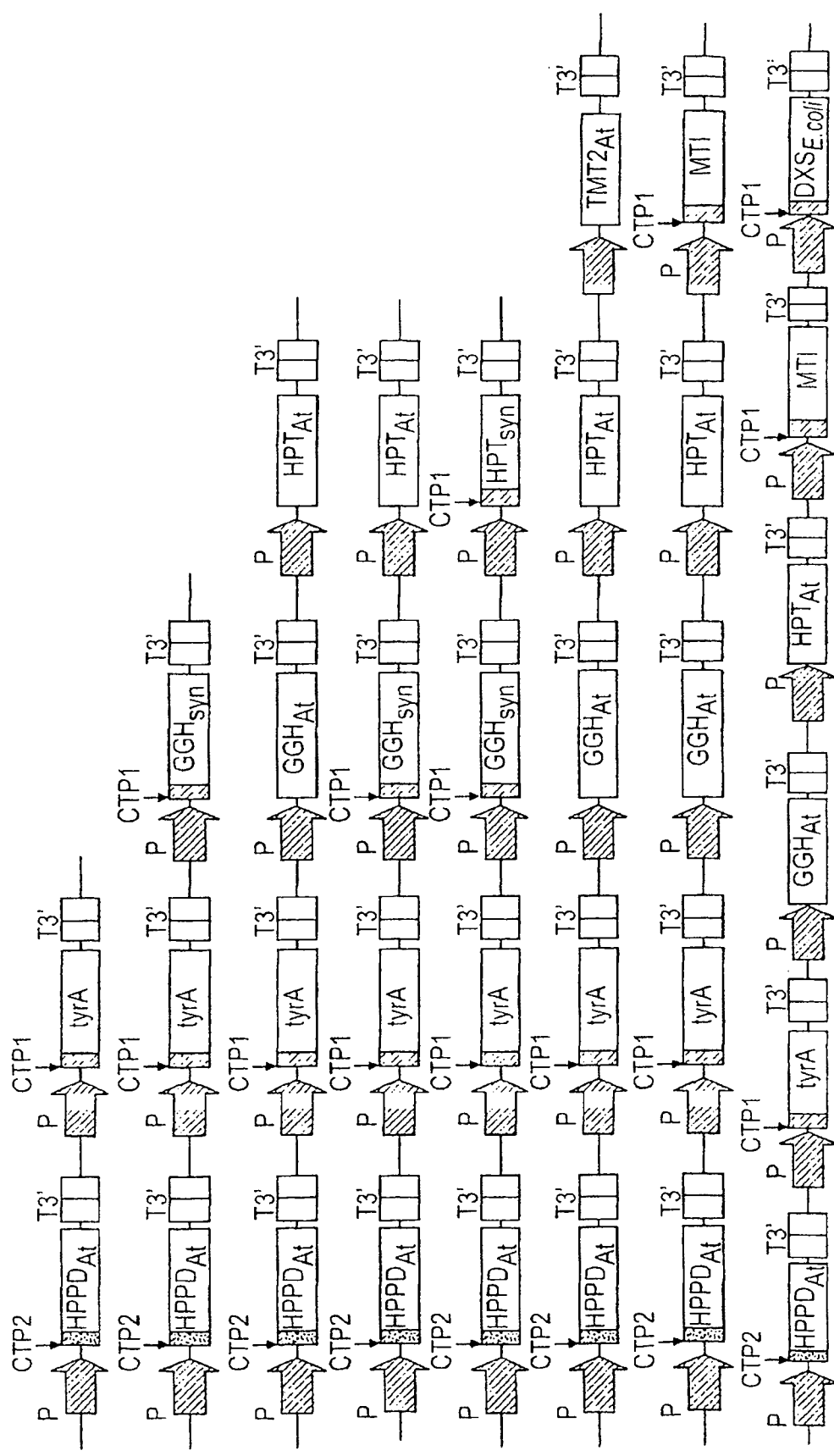
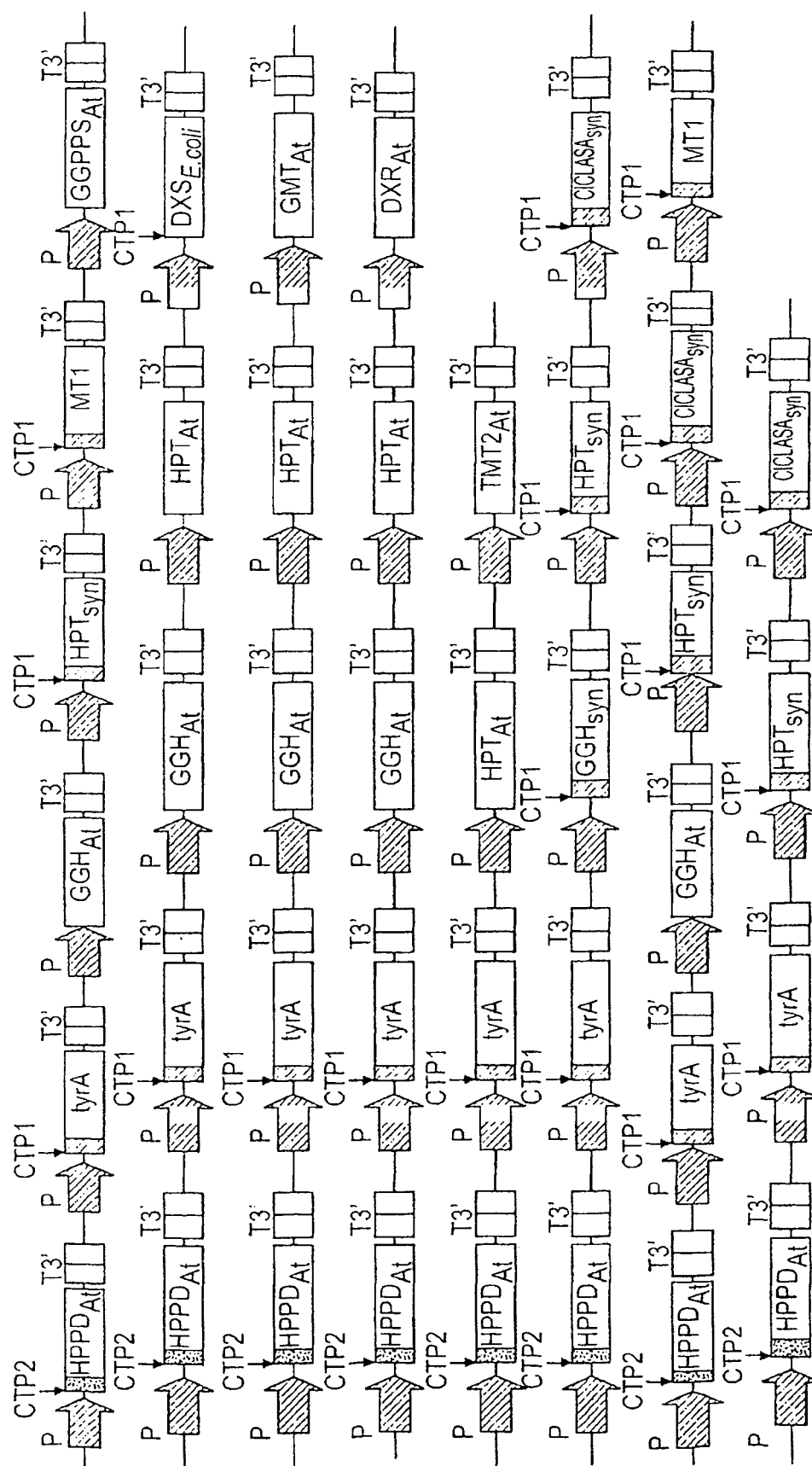


FIG. 18A

**FIG. 18B**

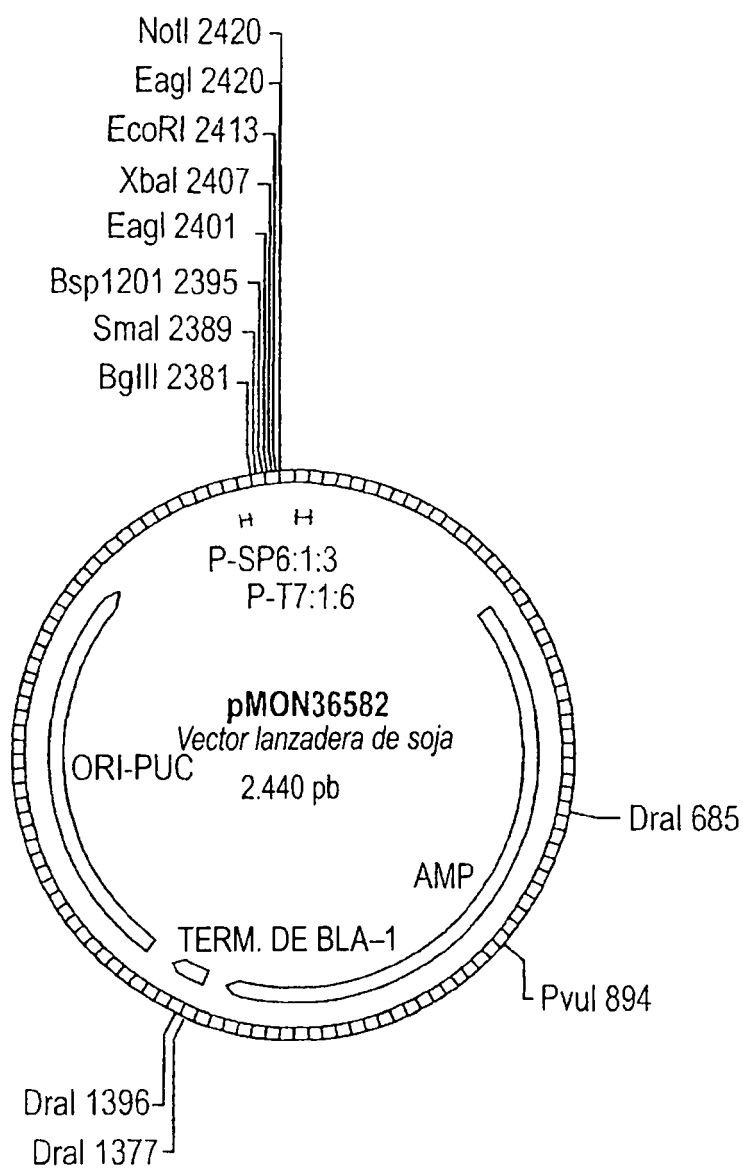


FIG. 19

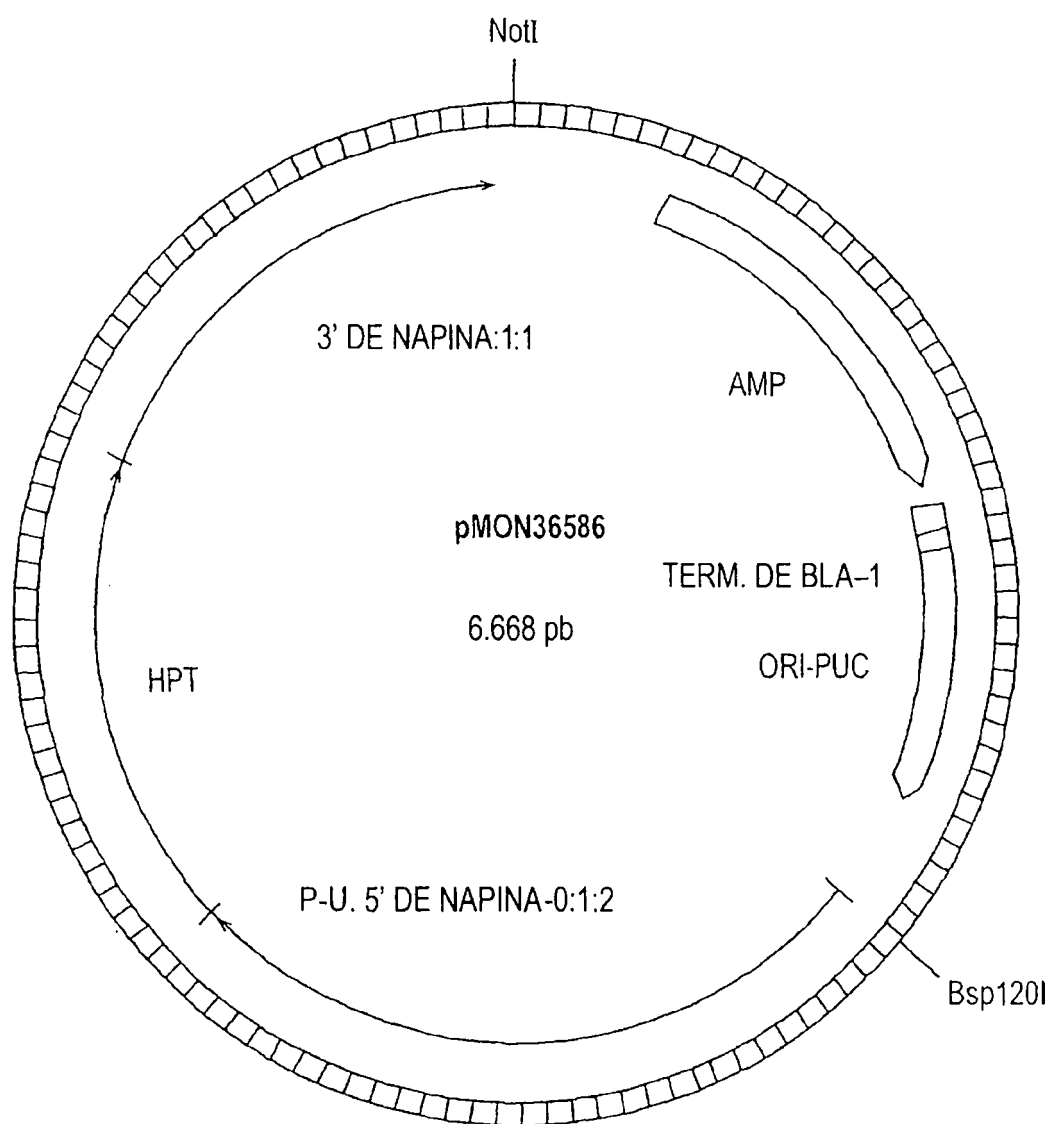


FIG. 20

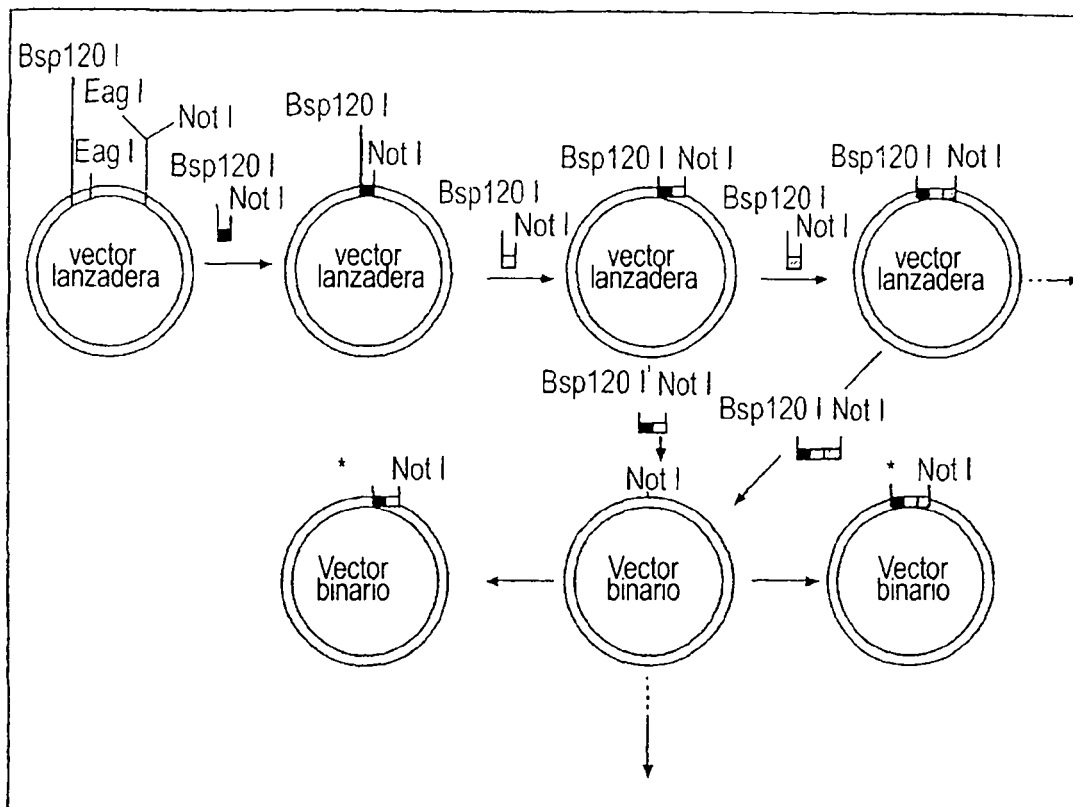


FIG. 21A

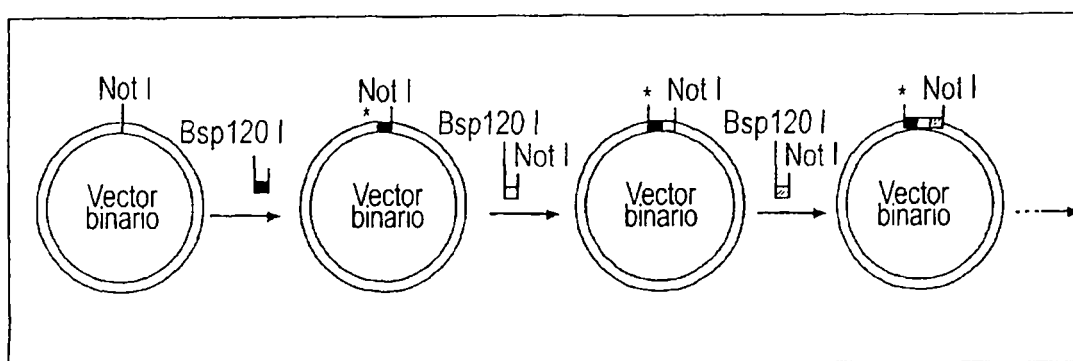


FIG. 21B

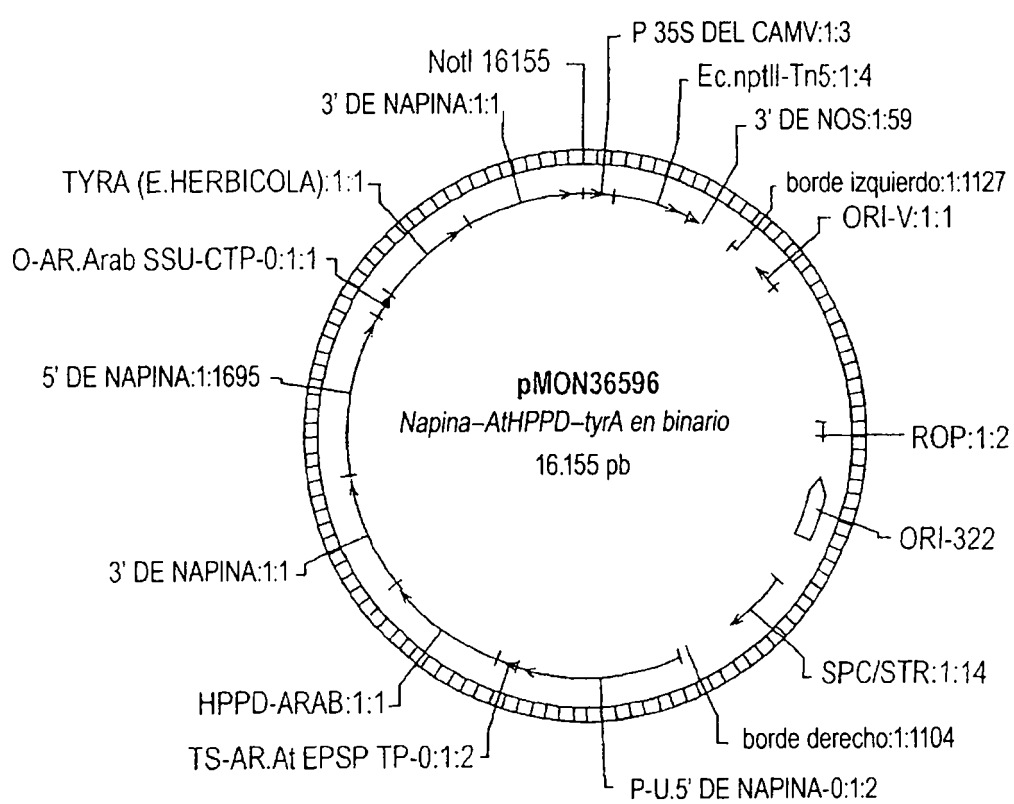


FIG. 22

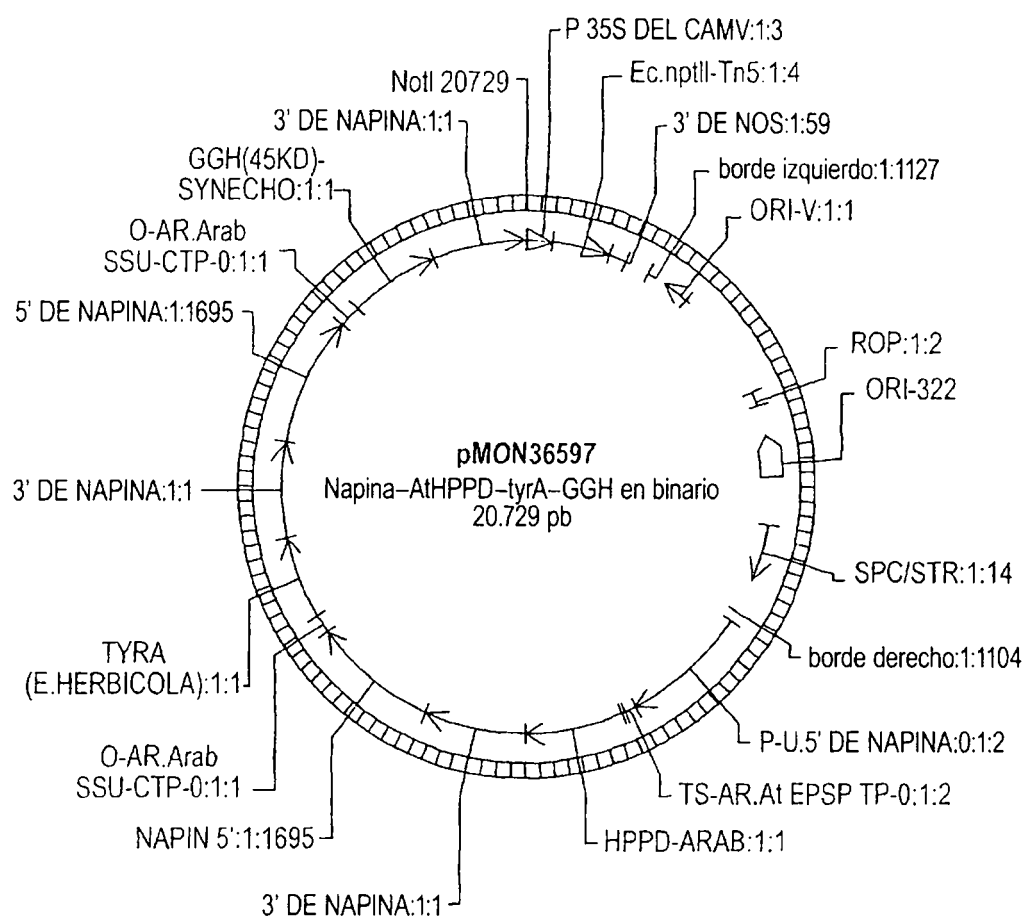


FIG. 23

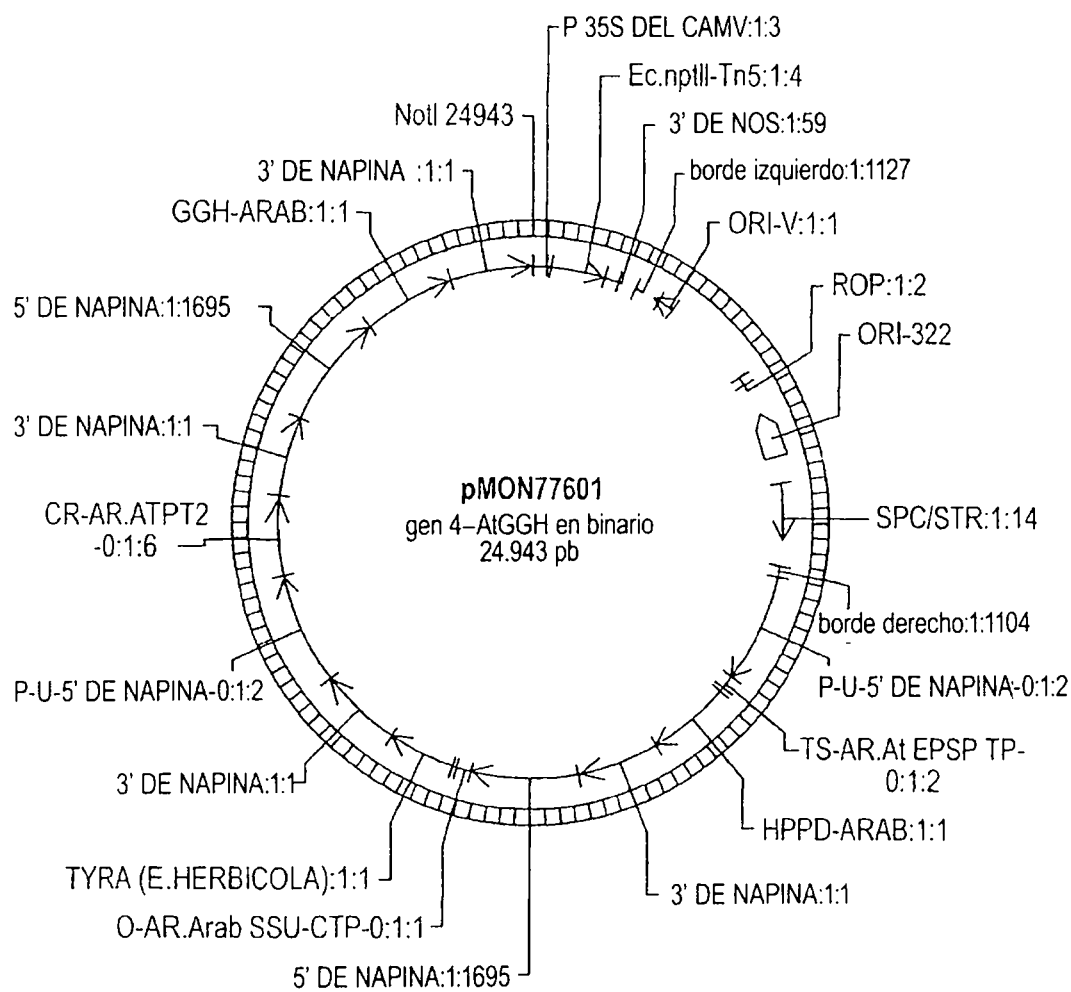


FIG. 24

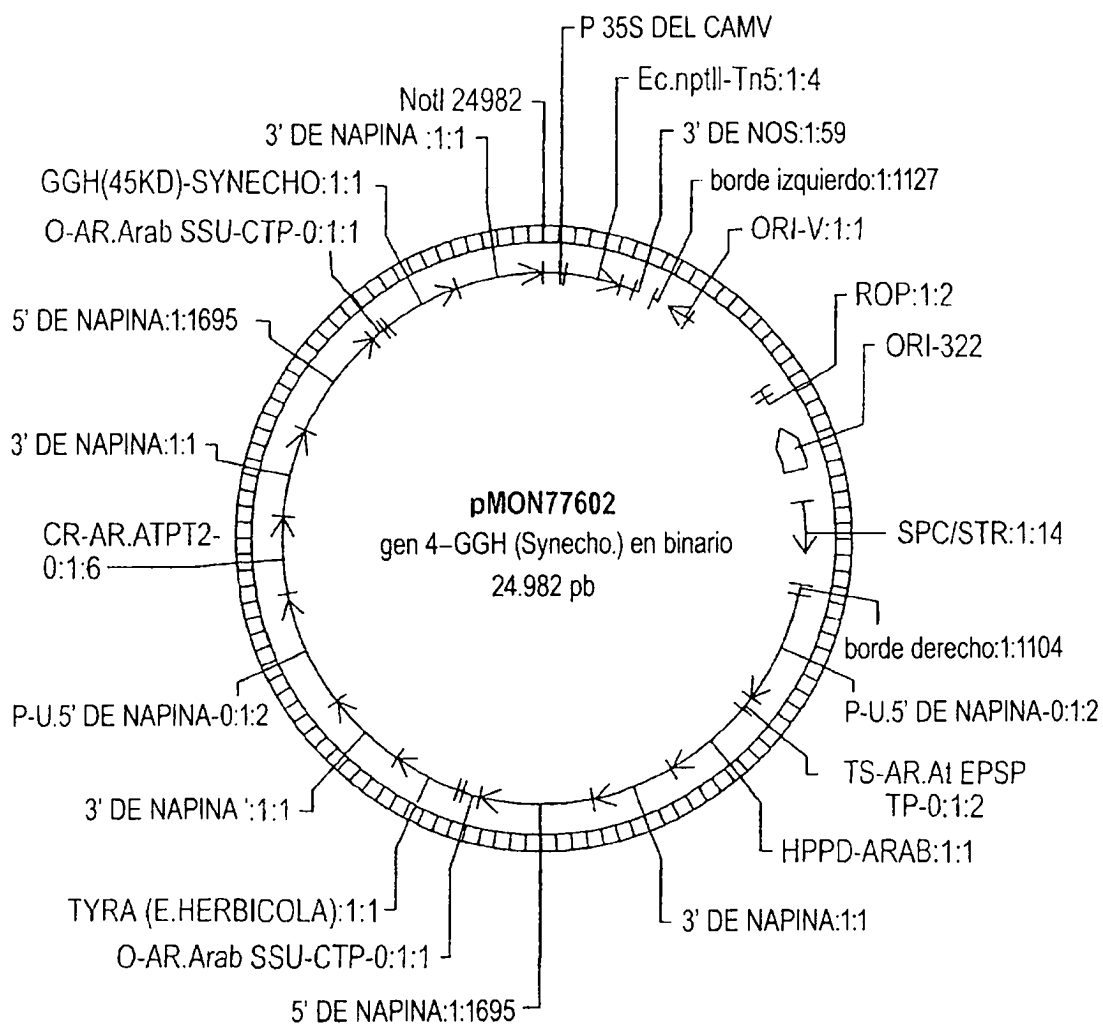


FIG. 25

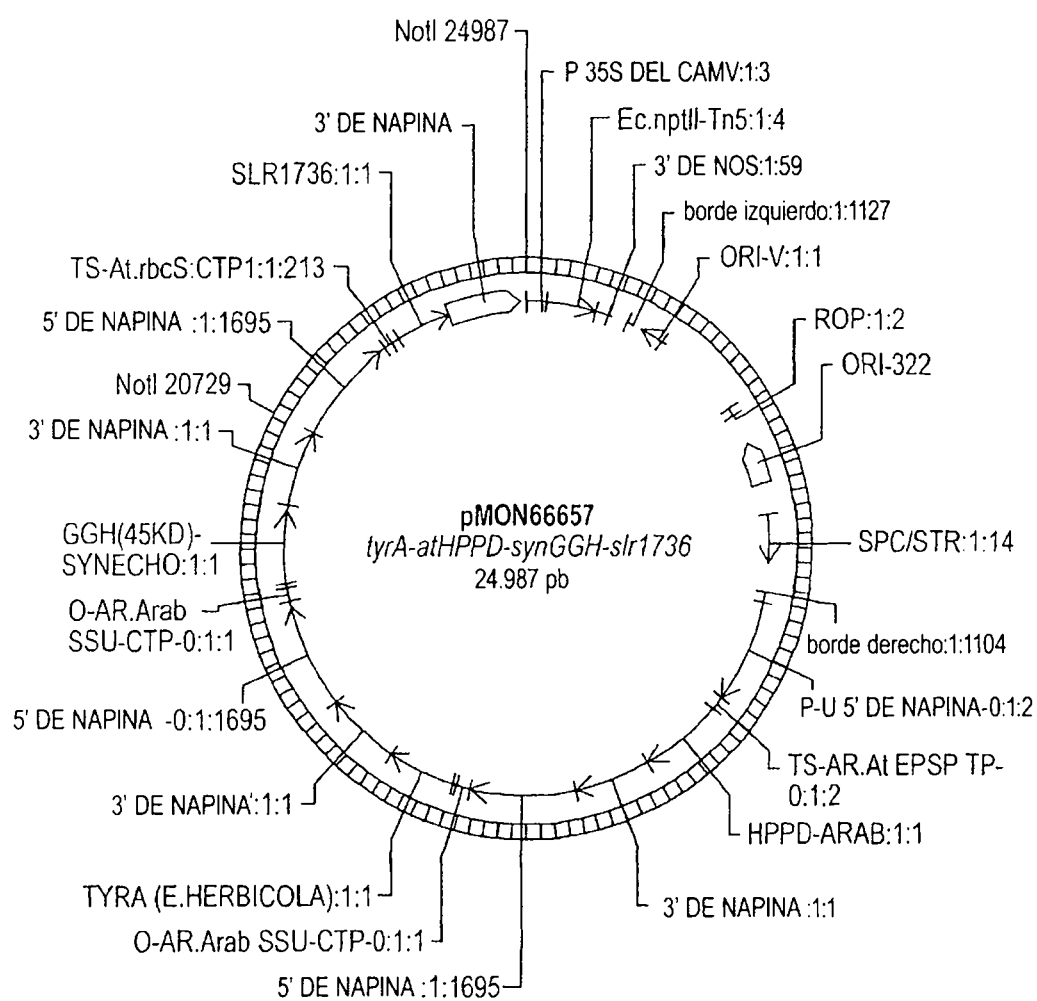


FIG. 26

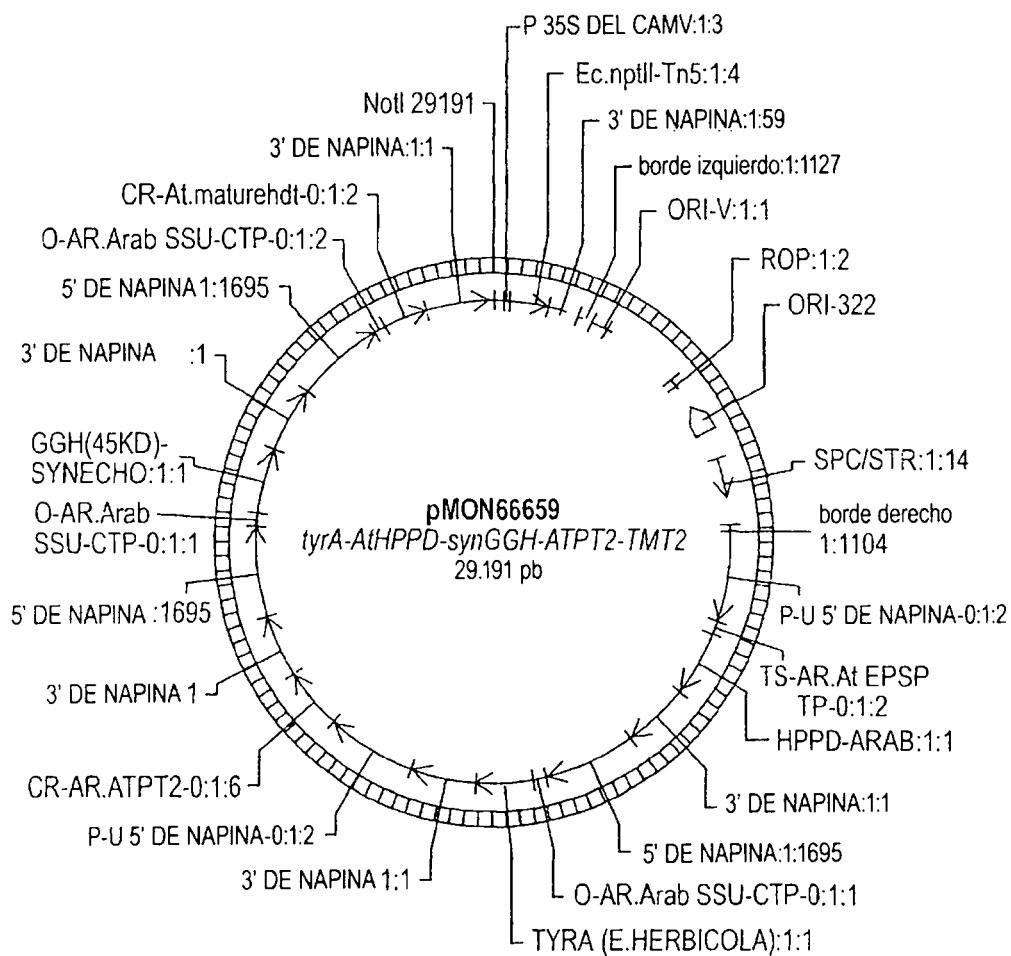


FIG. 27

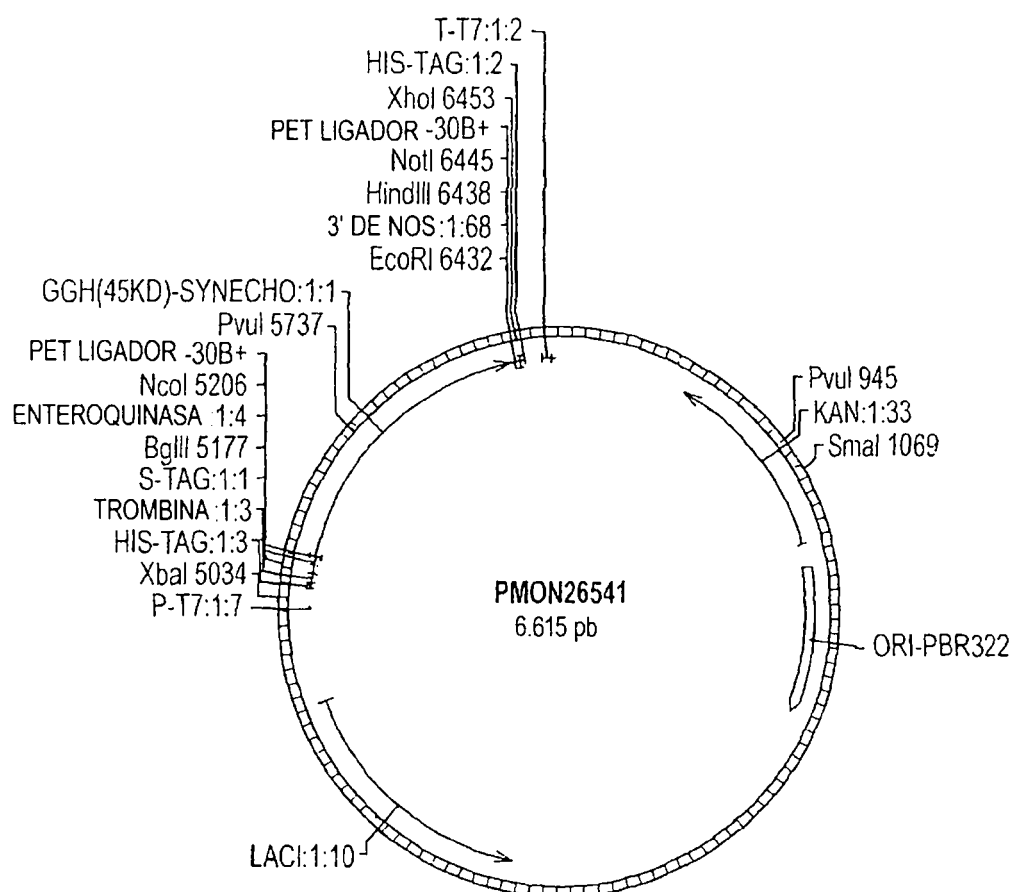


FIG. 28

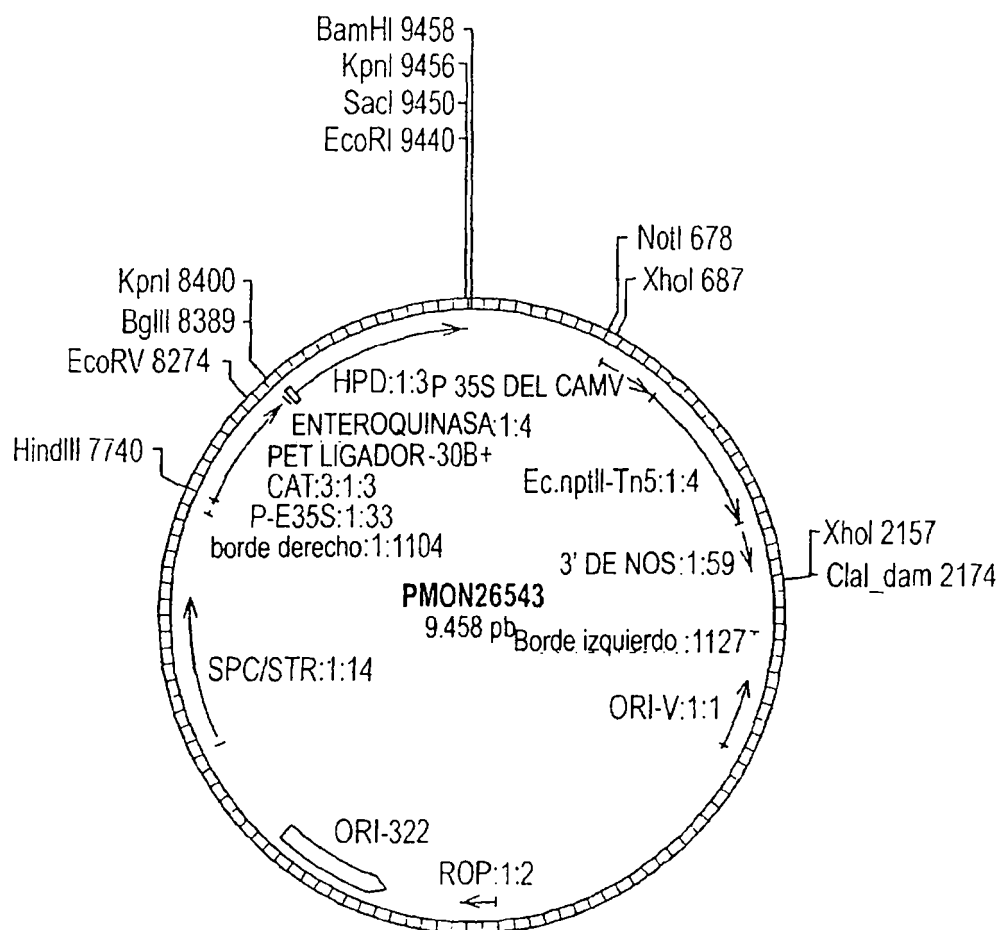


FIG. 29

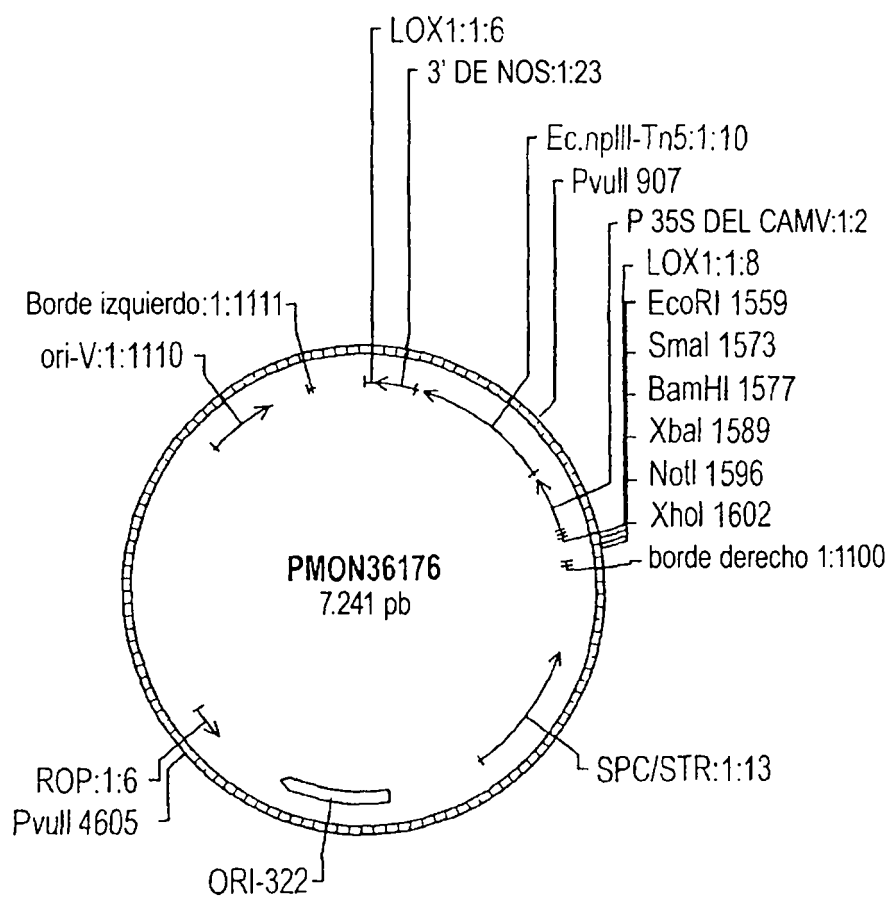


FIG. 30

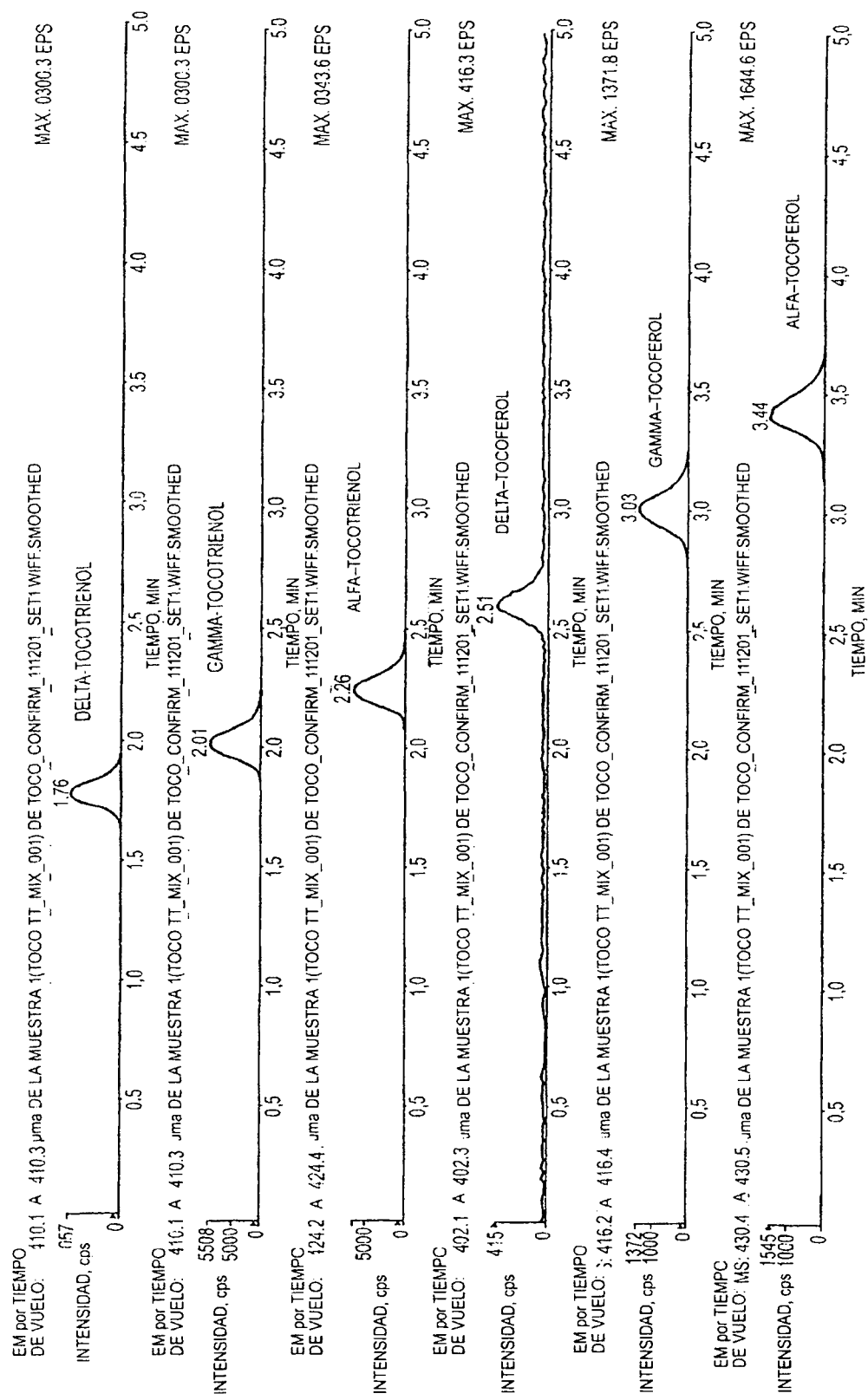
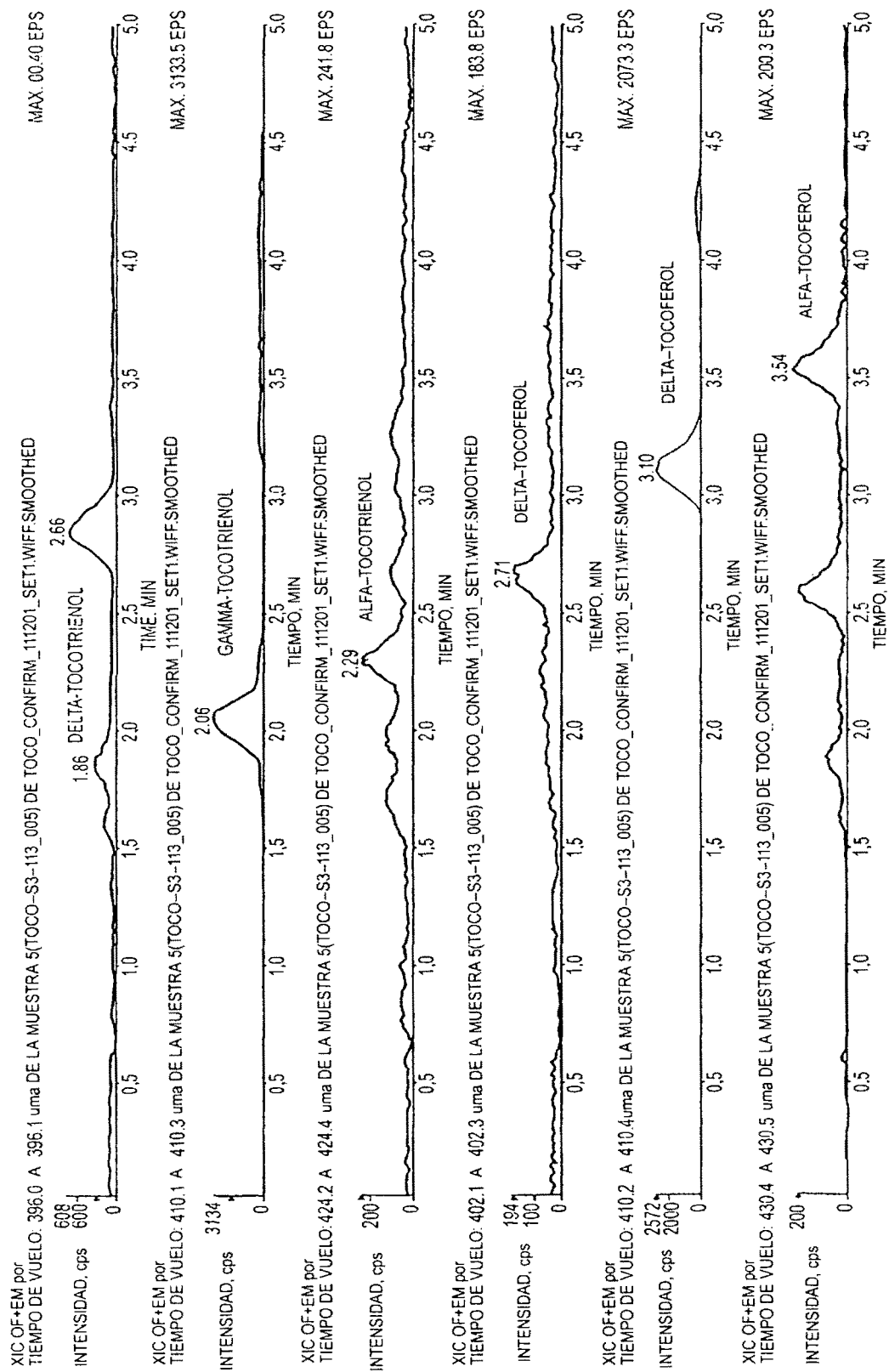


FIG. 31

**FIG. 32**

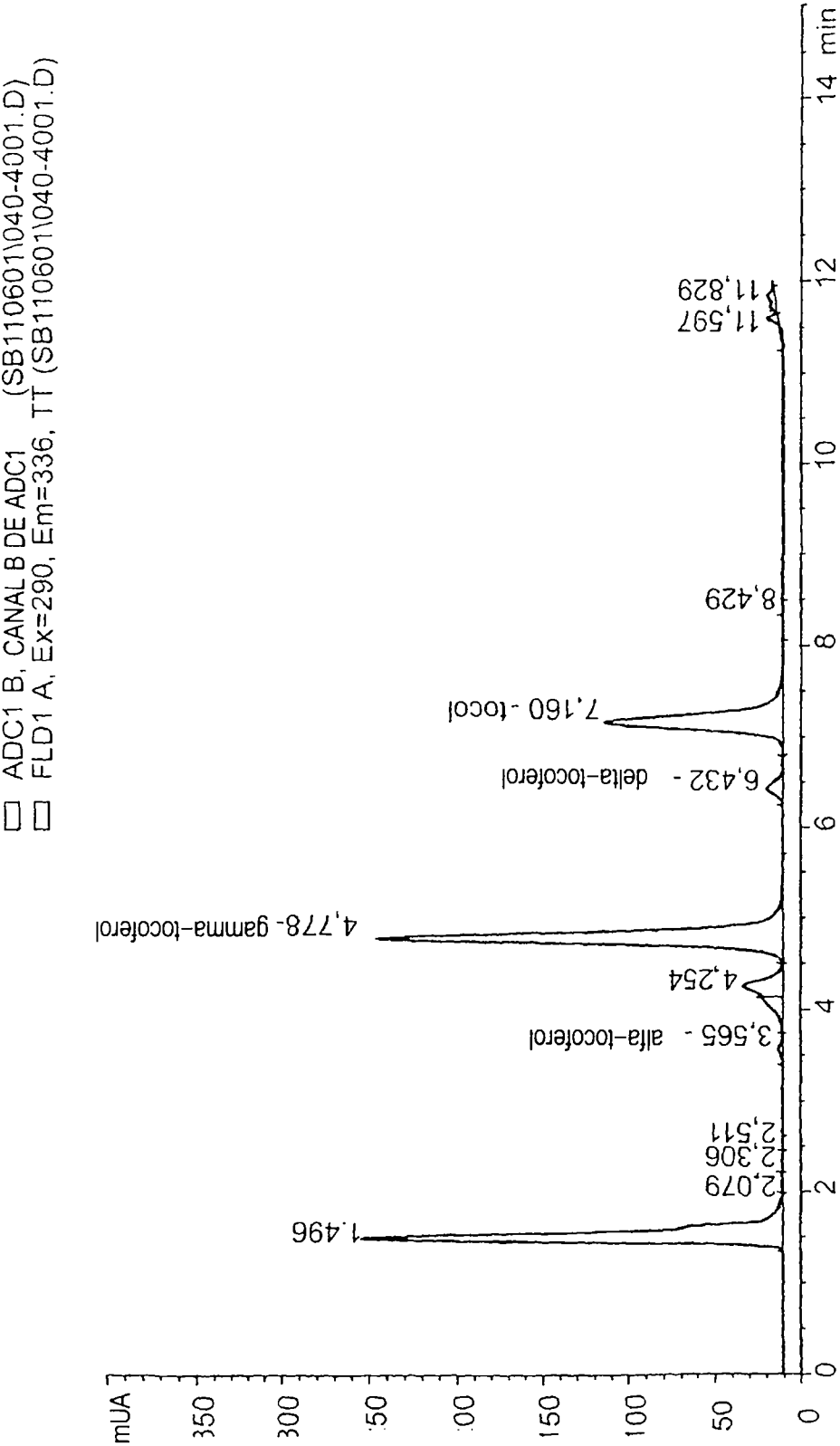


FIG. 33

☐ ADC1 B, CANAL B DE ADC1 (SB110601\035-3501.D)
☐ FLD1 A, Ex=290, Em=336, TT (SB110601\035-3501.D)

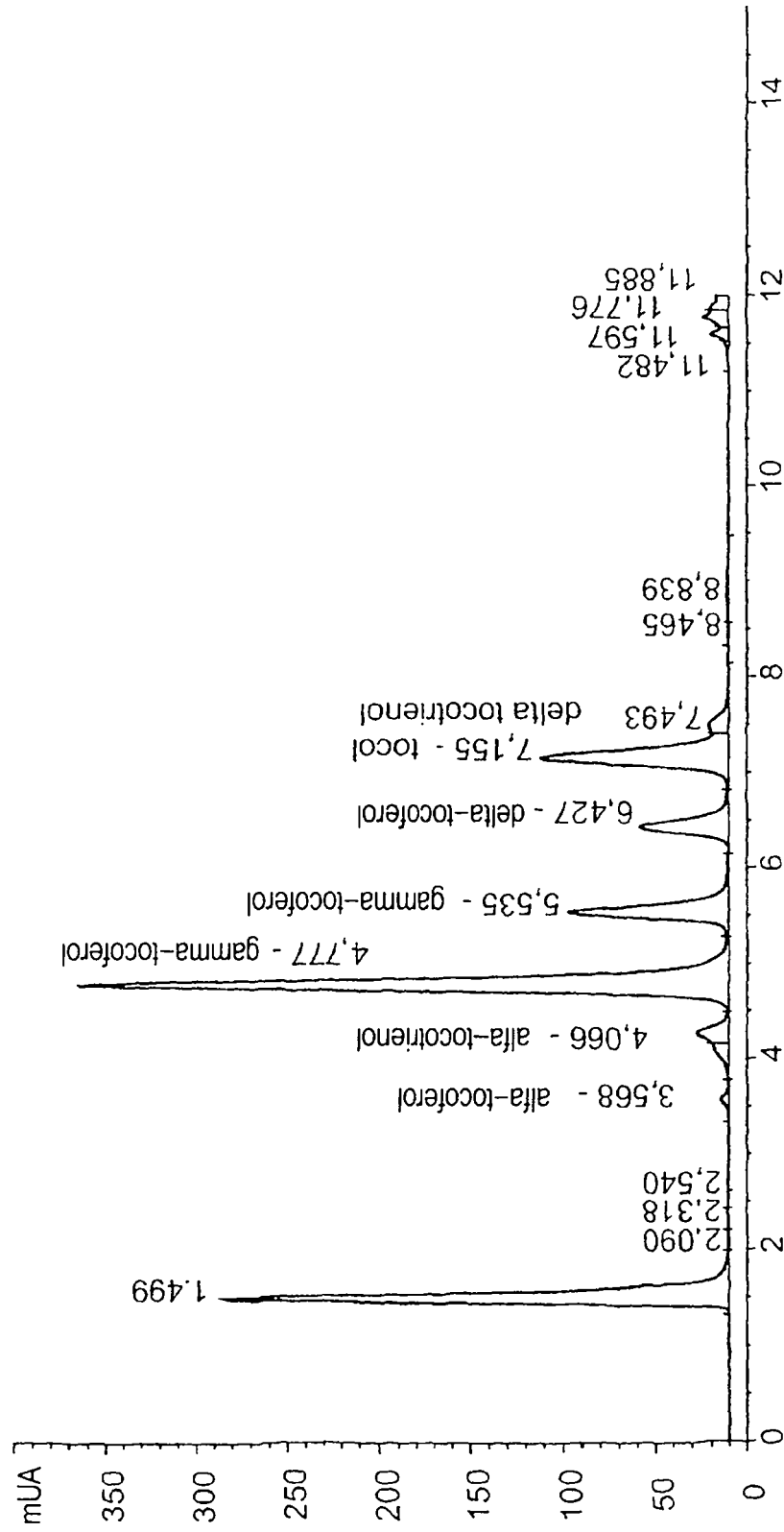


FIG. 34

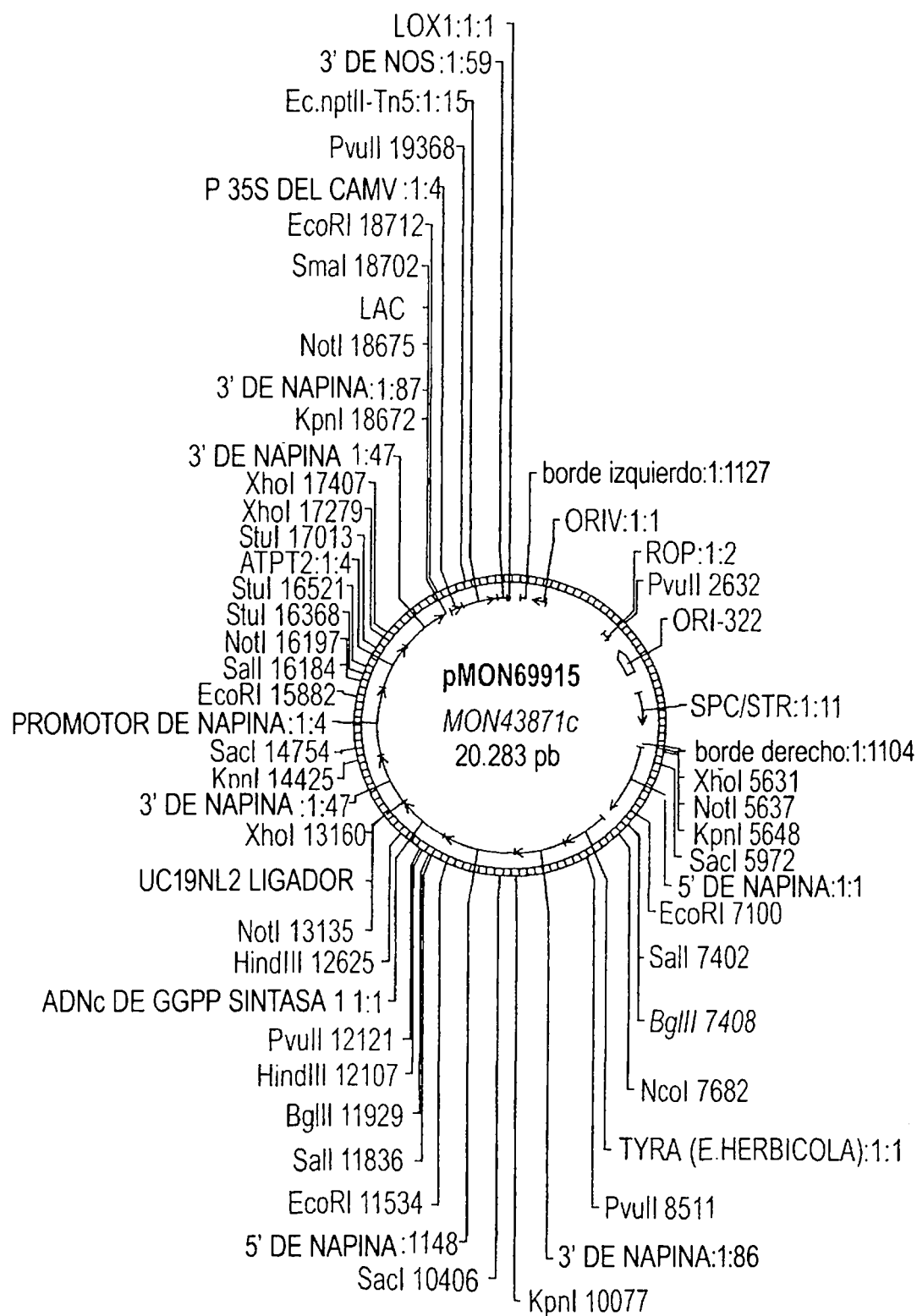


FIG. 35

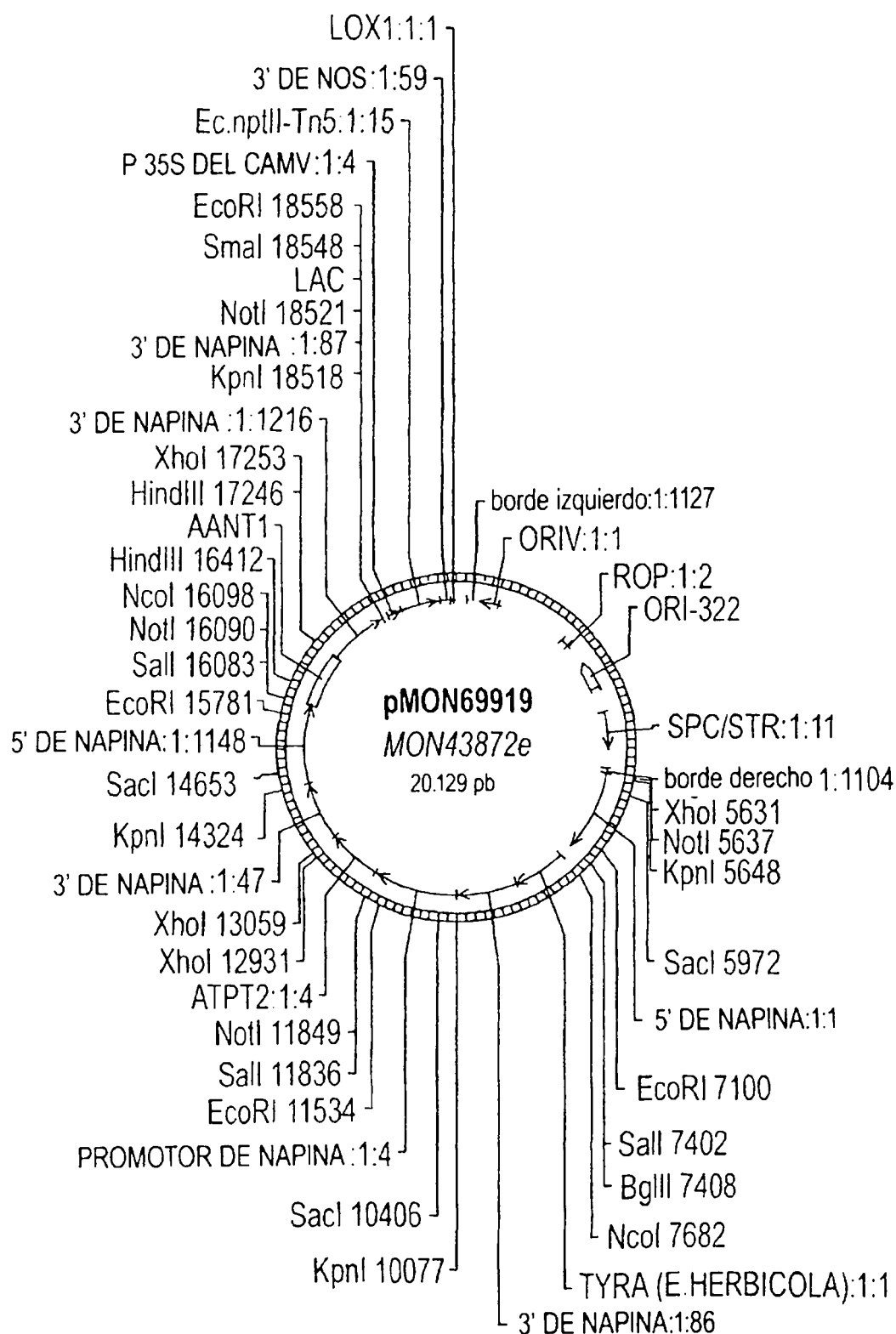


FIG. 36

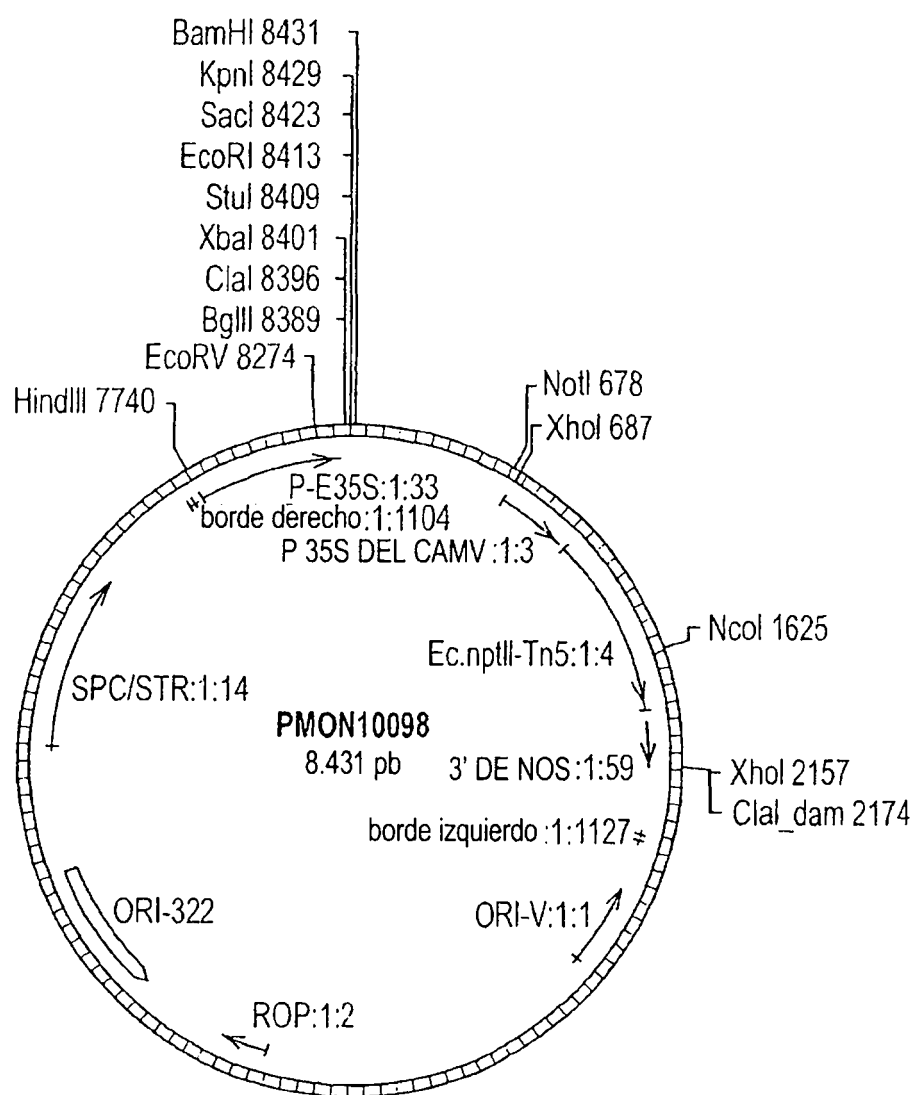


FIG. 37

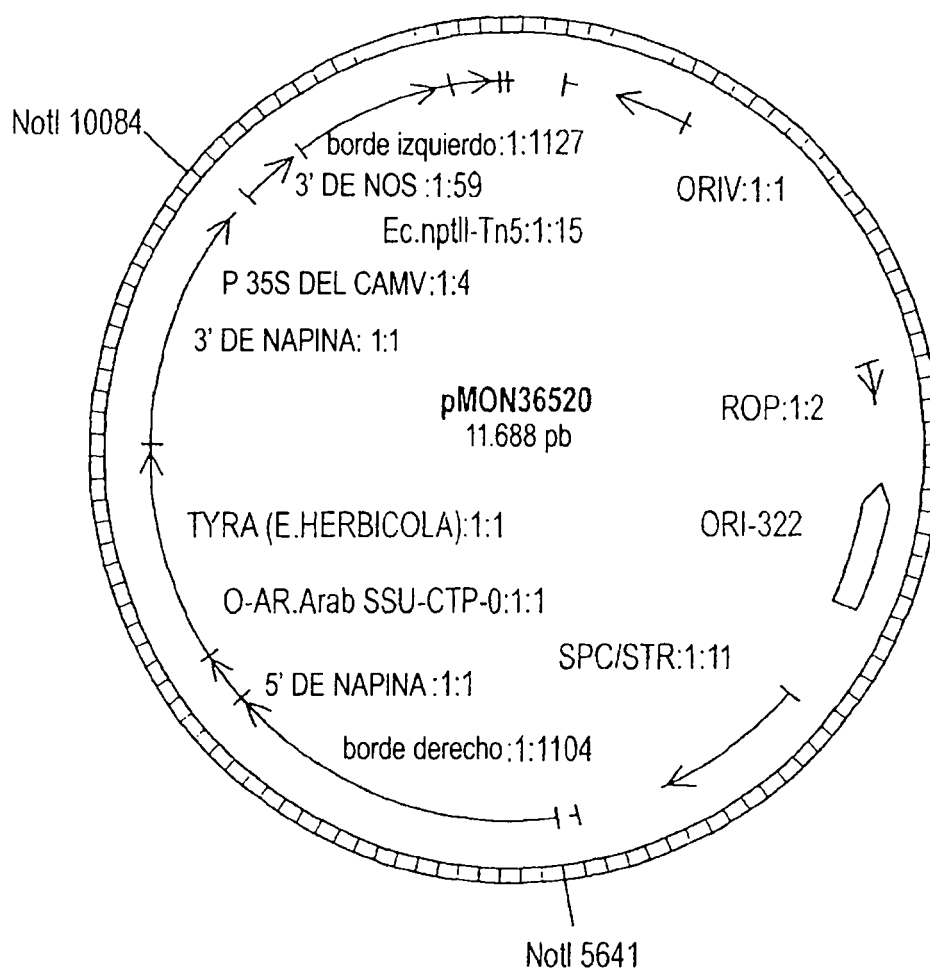


FIG. 38

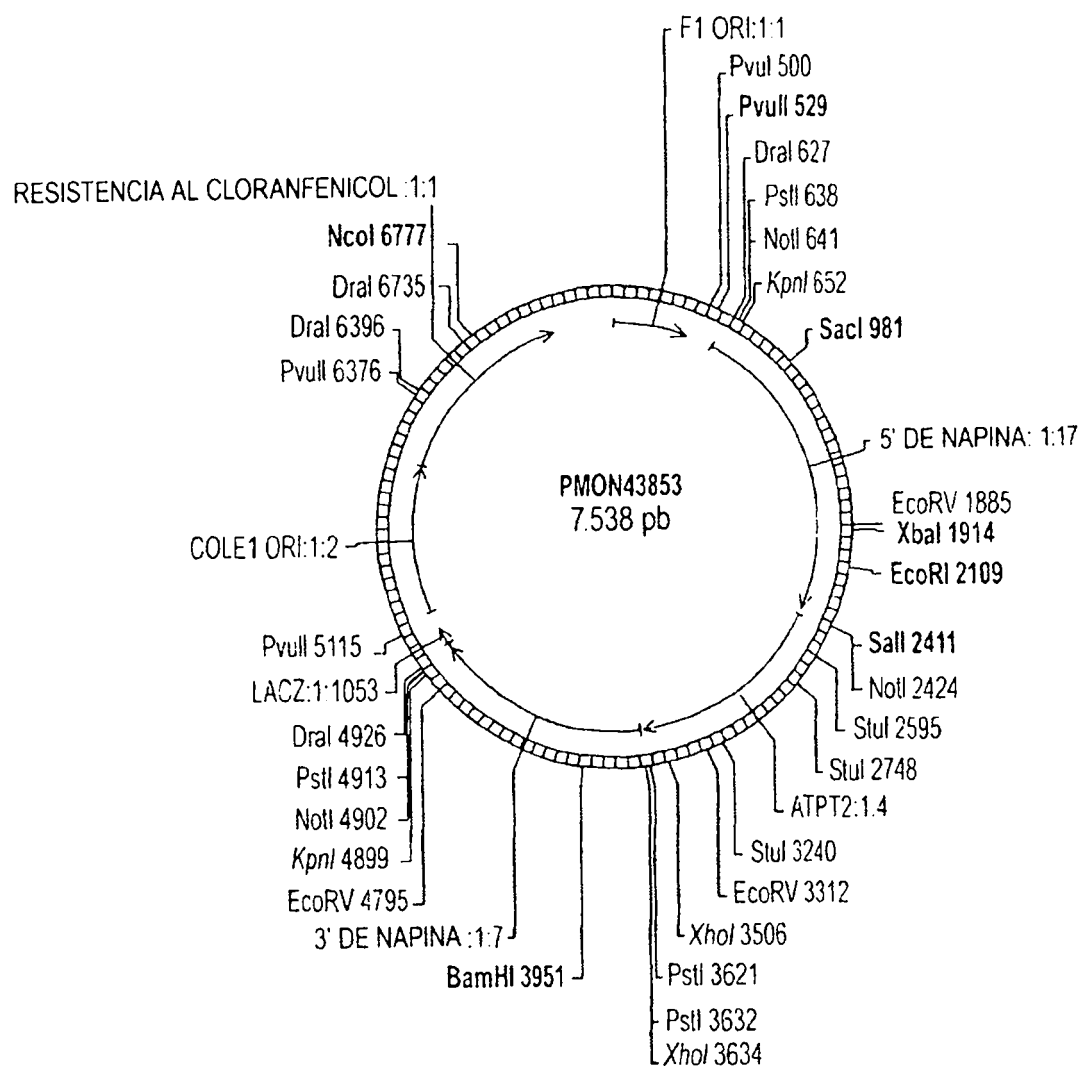


FIG. 39

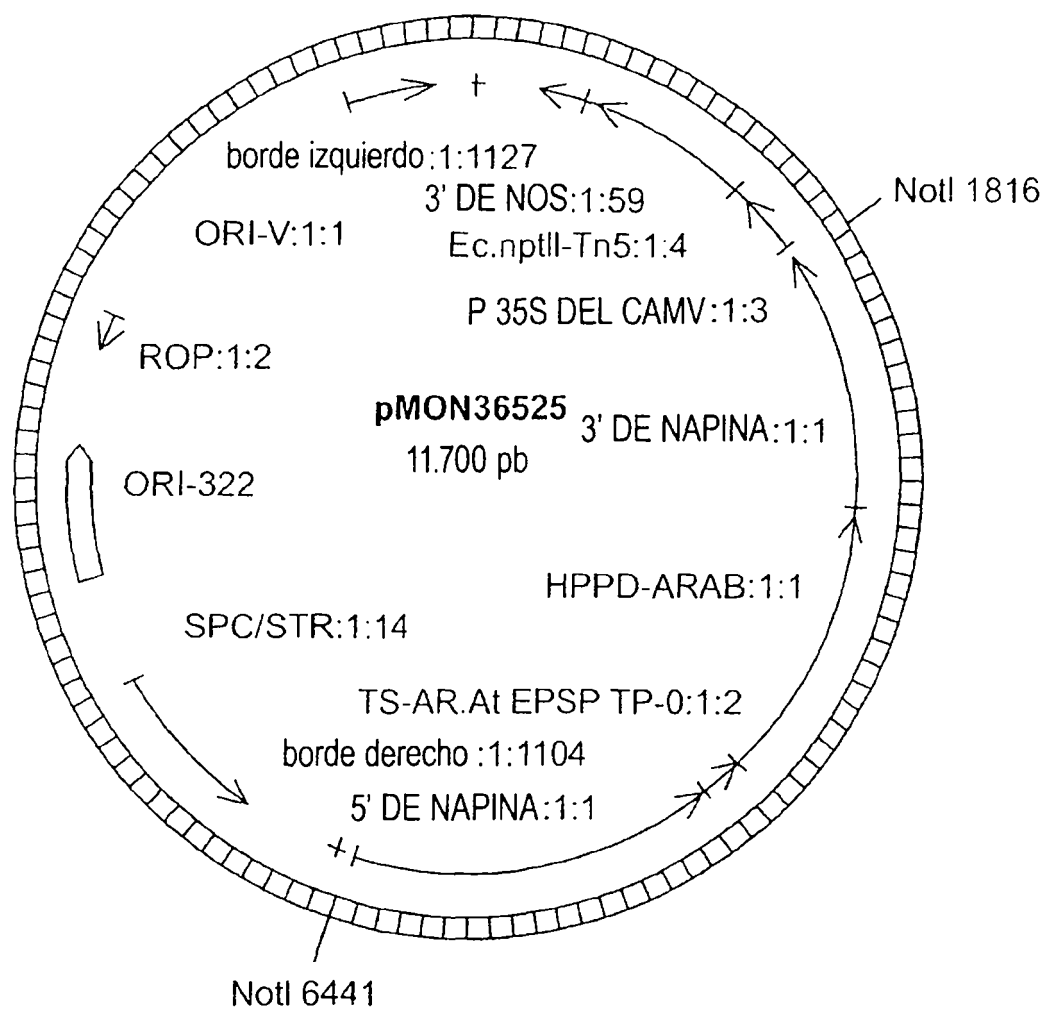


FIG. 40

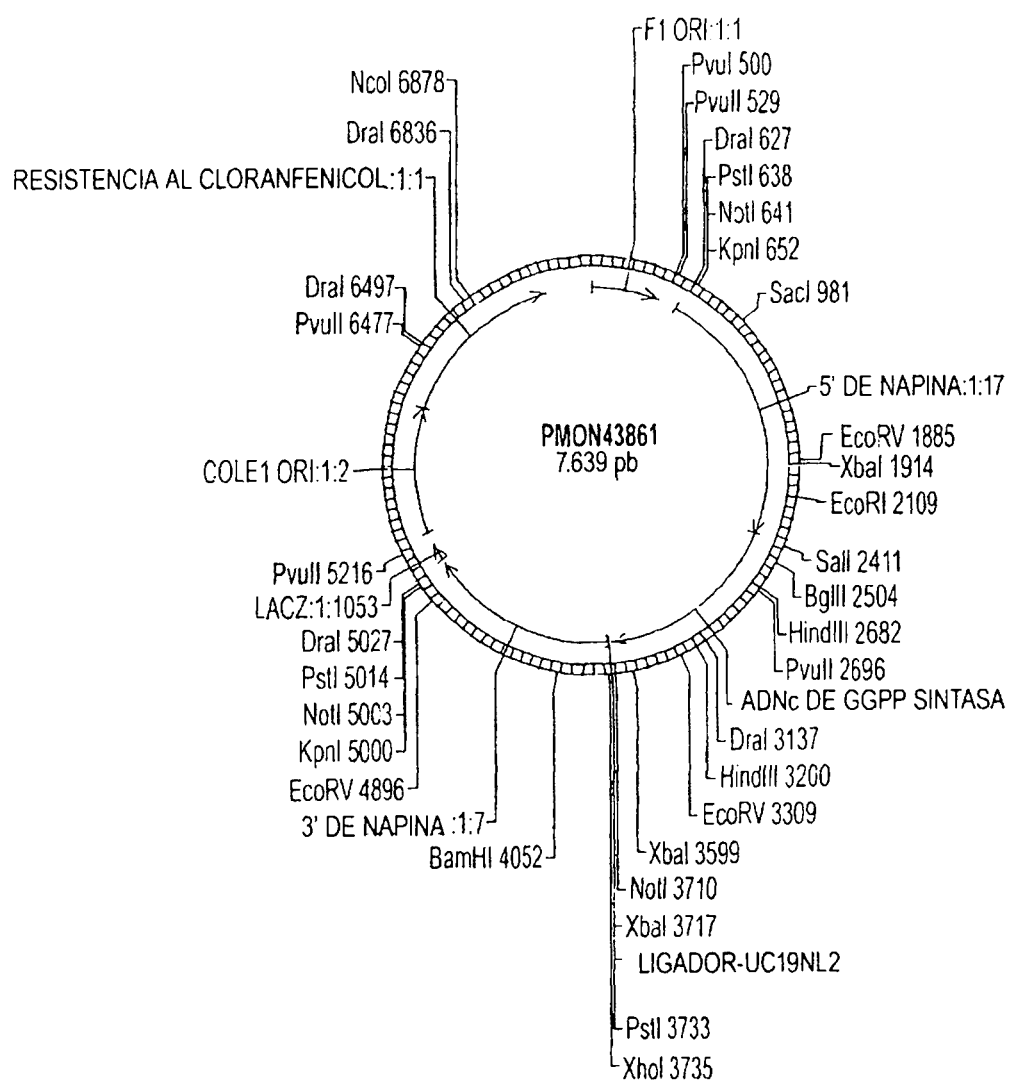


FIG. 41

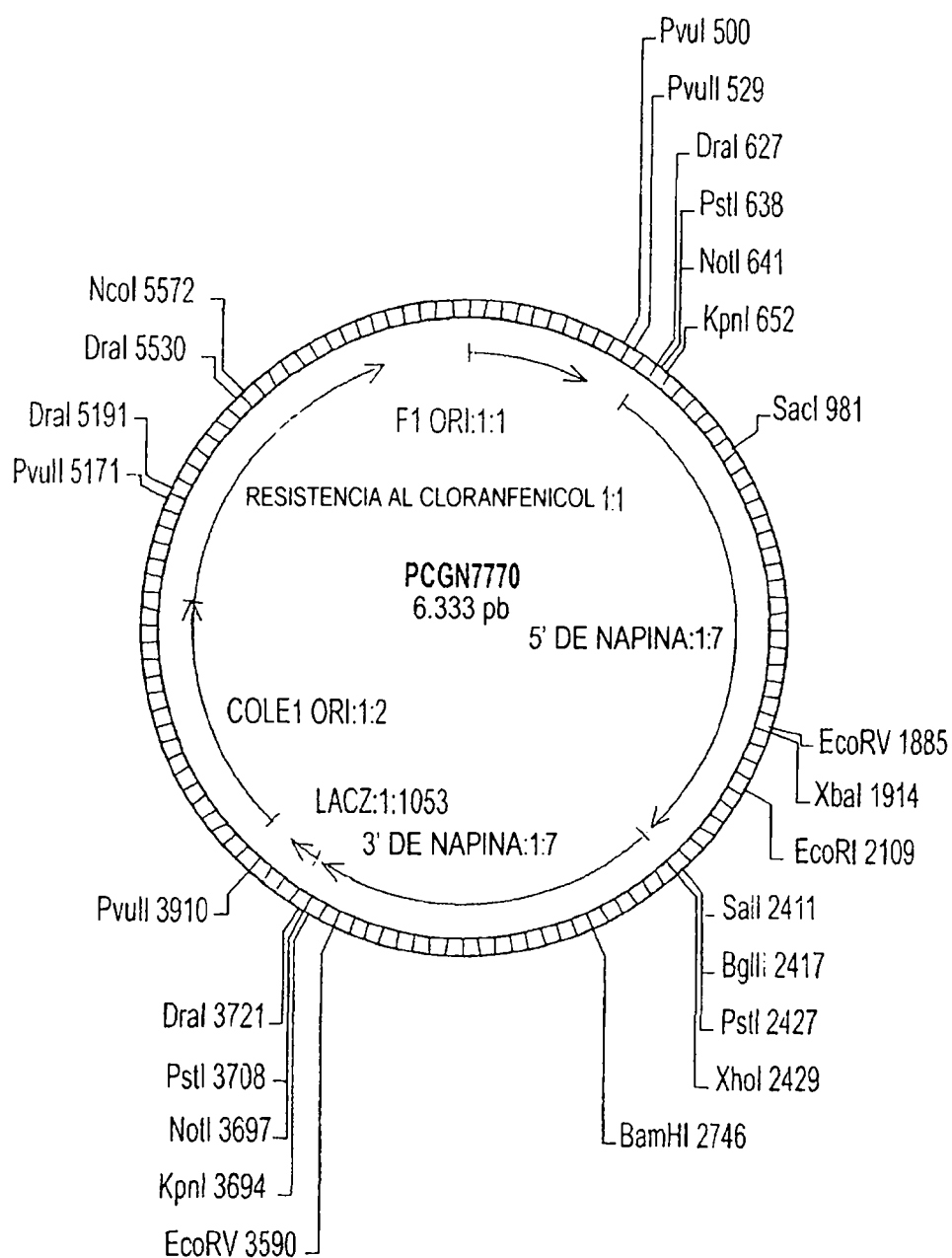


FIG. 42

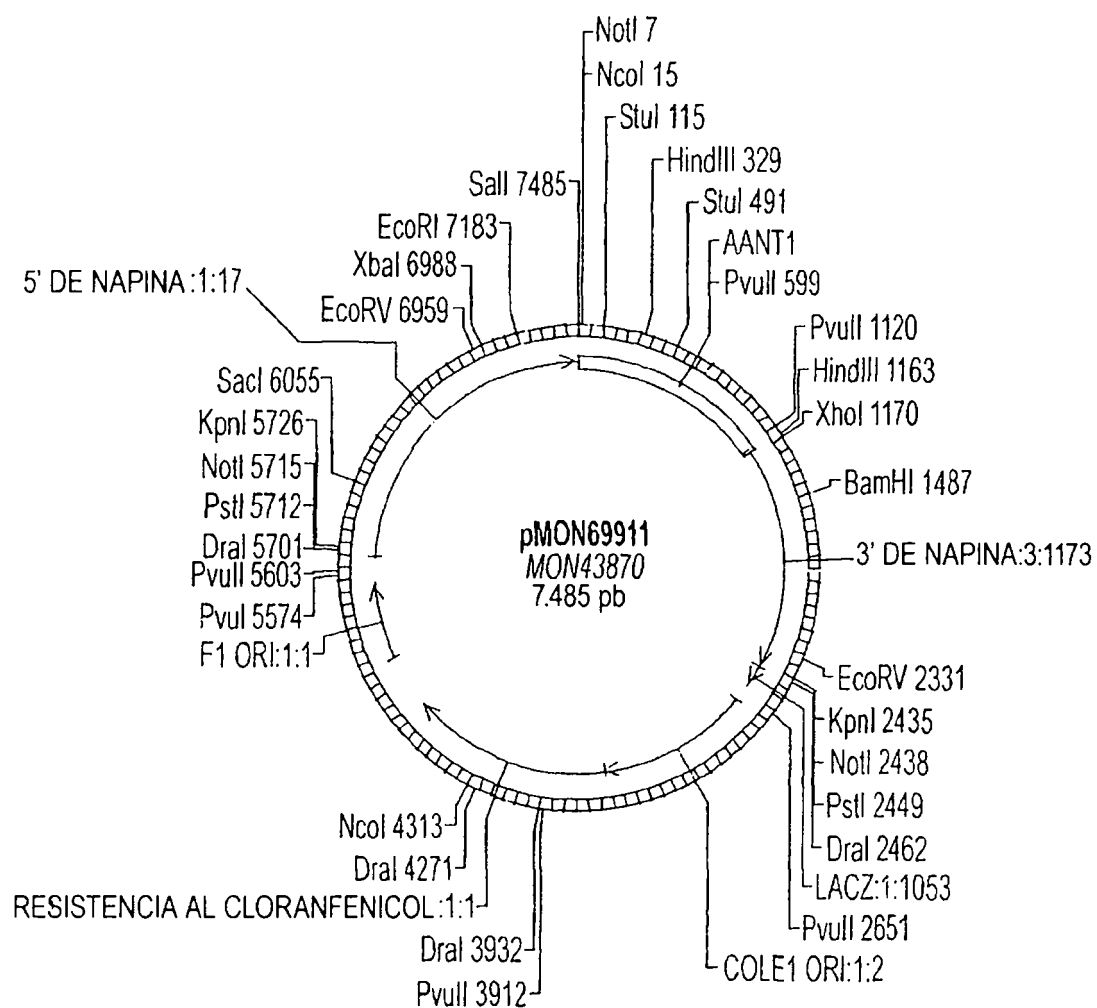


FIG. 43

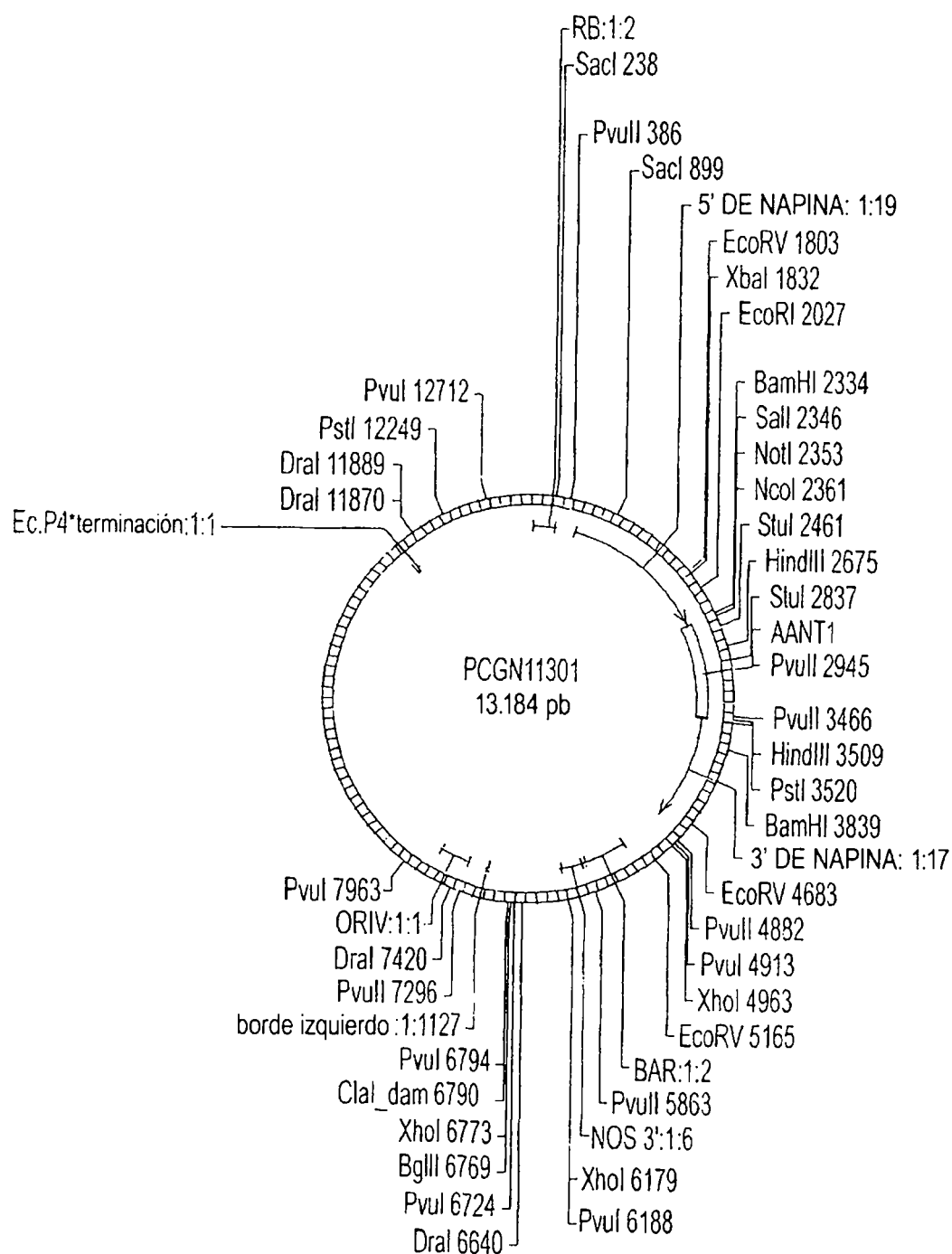


FIG. 44

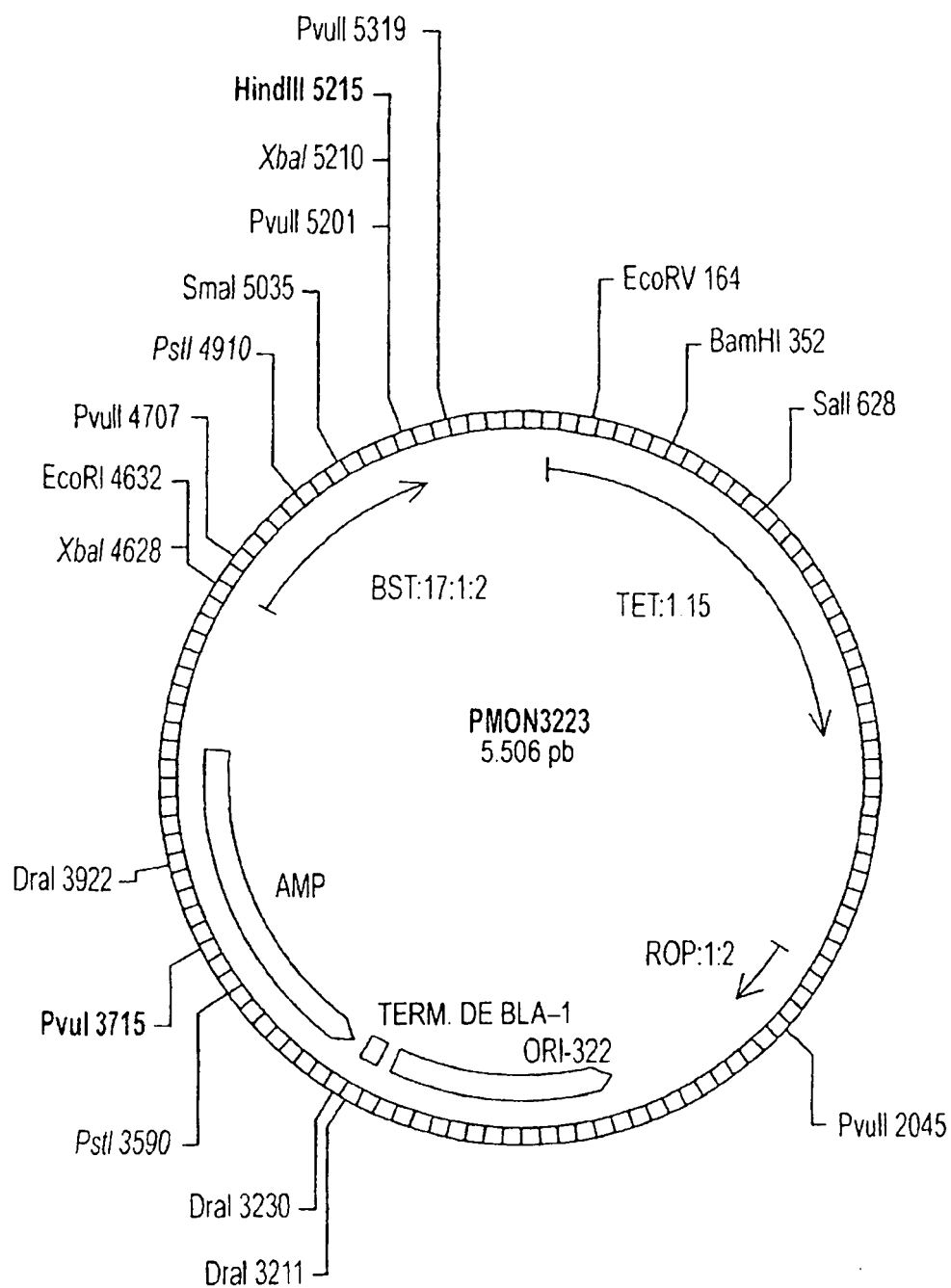


FIG. 45

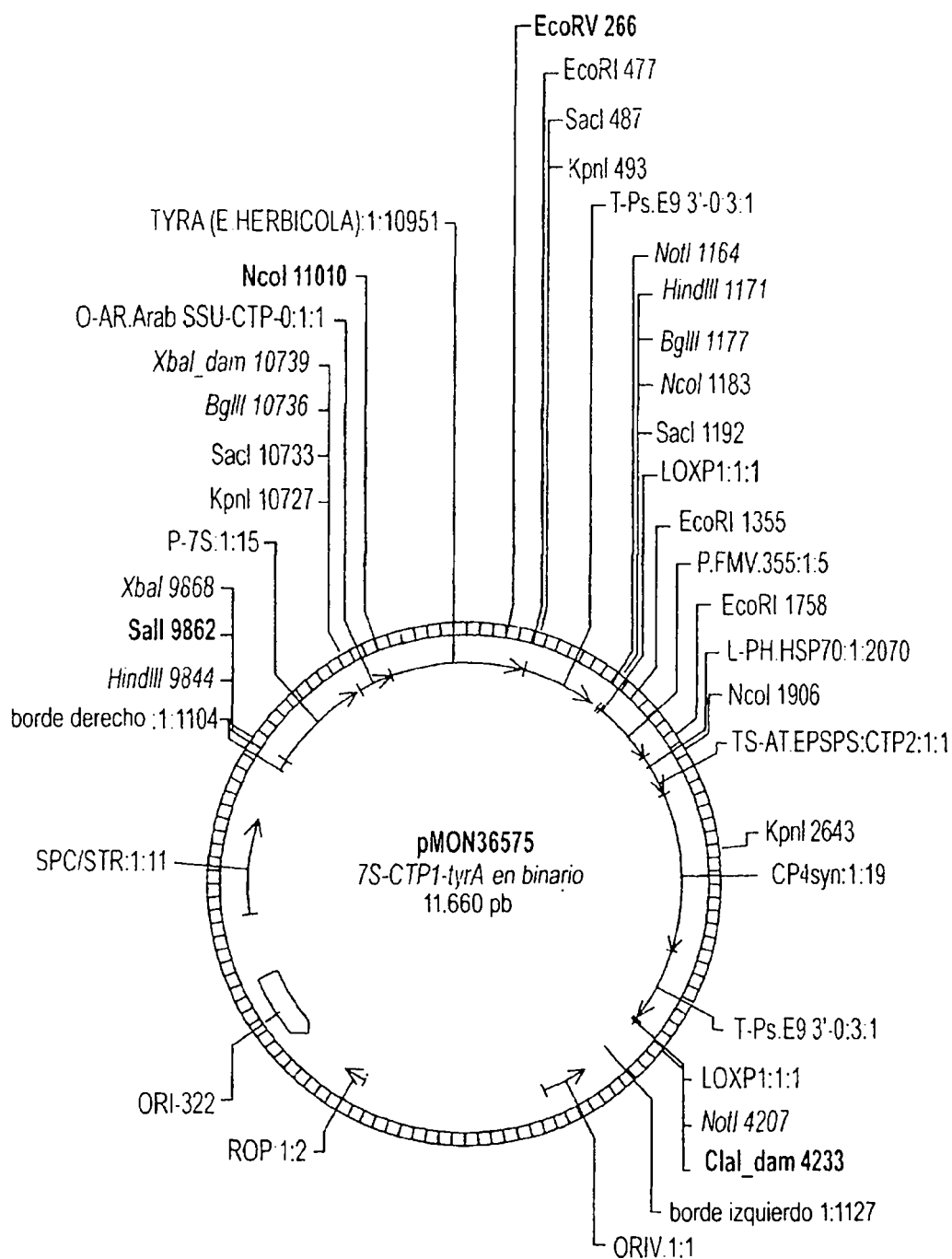


FIG. 46

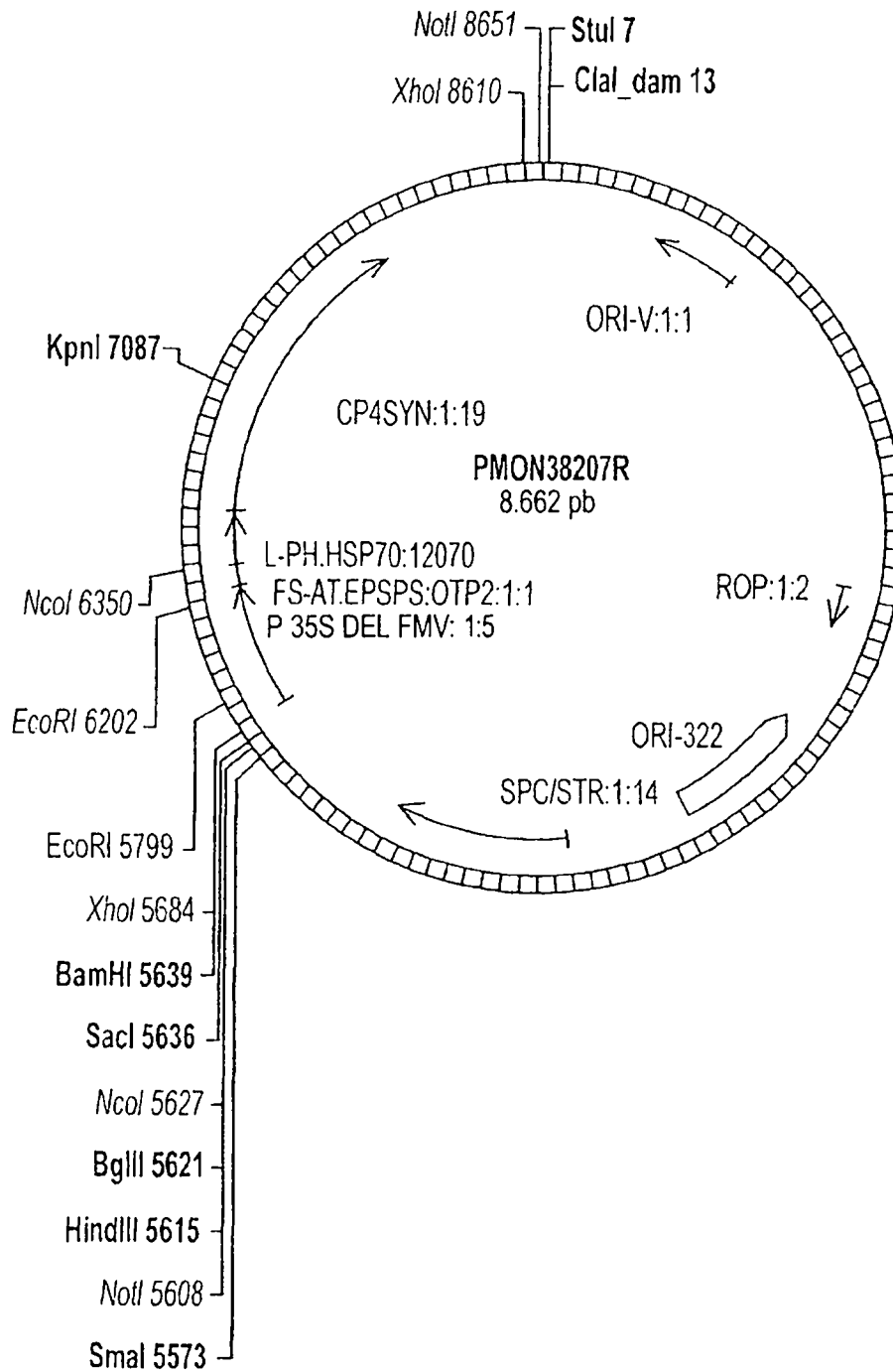


FIG. 47

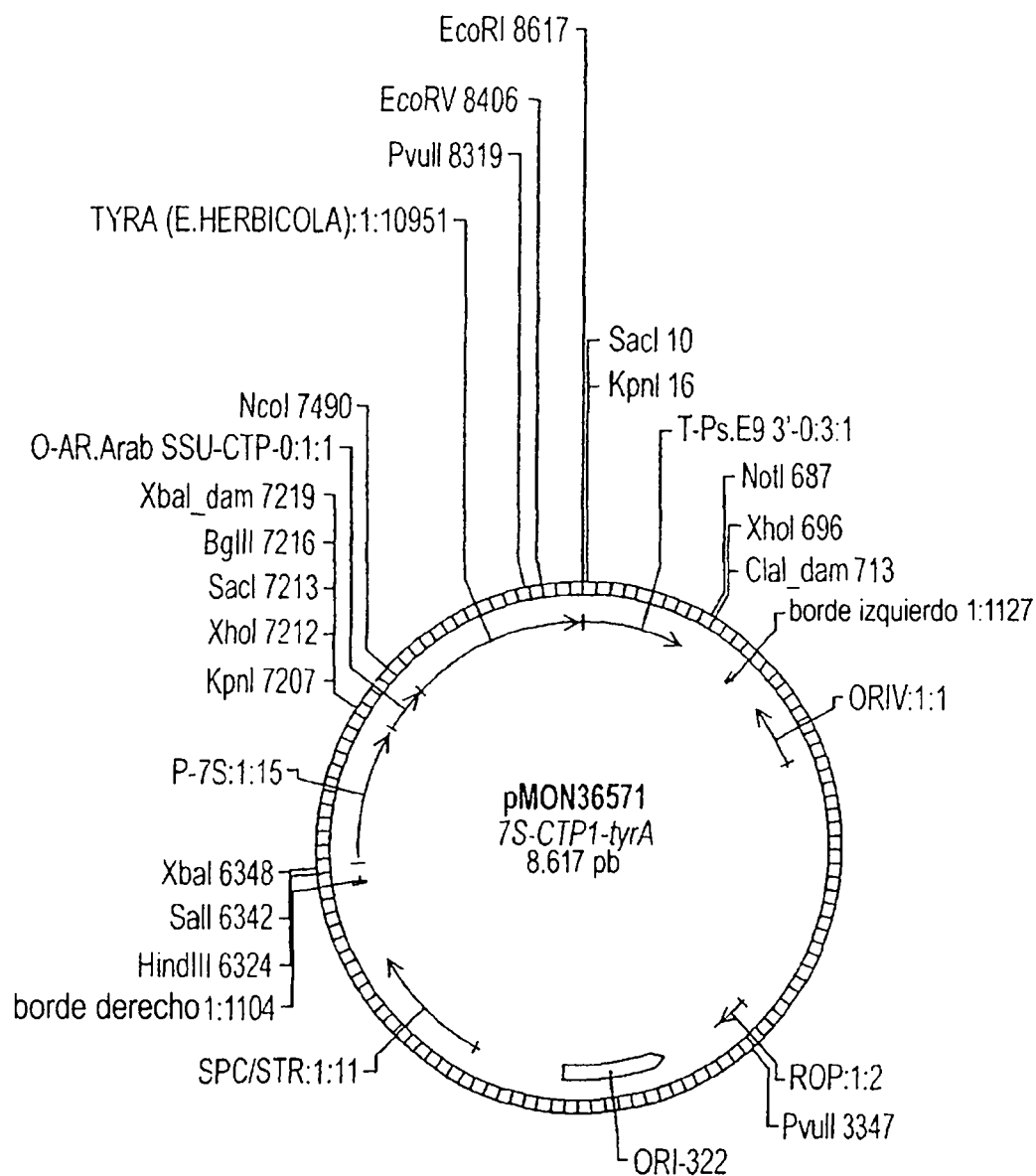


FIG. 48

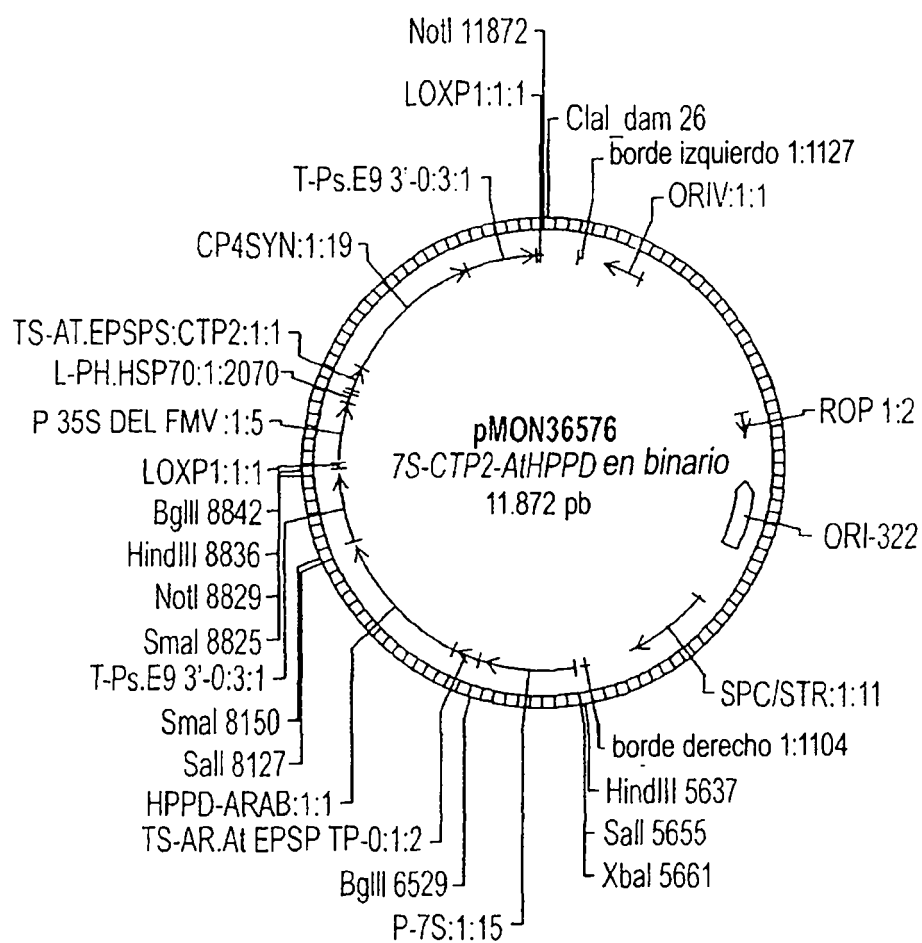


FIG. 49

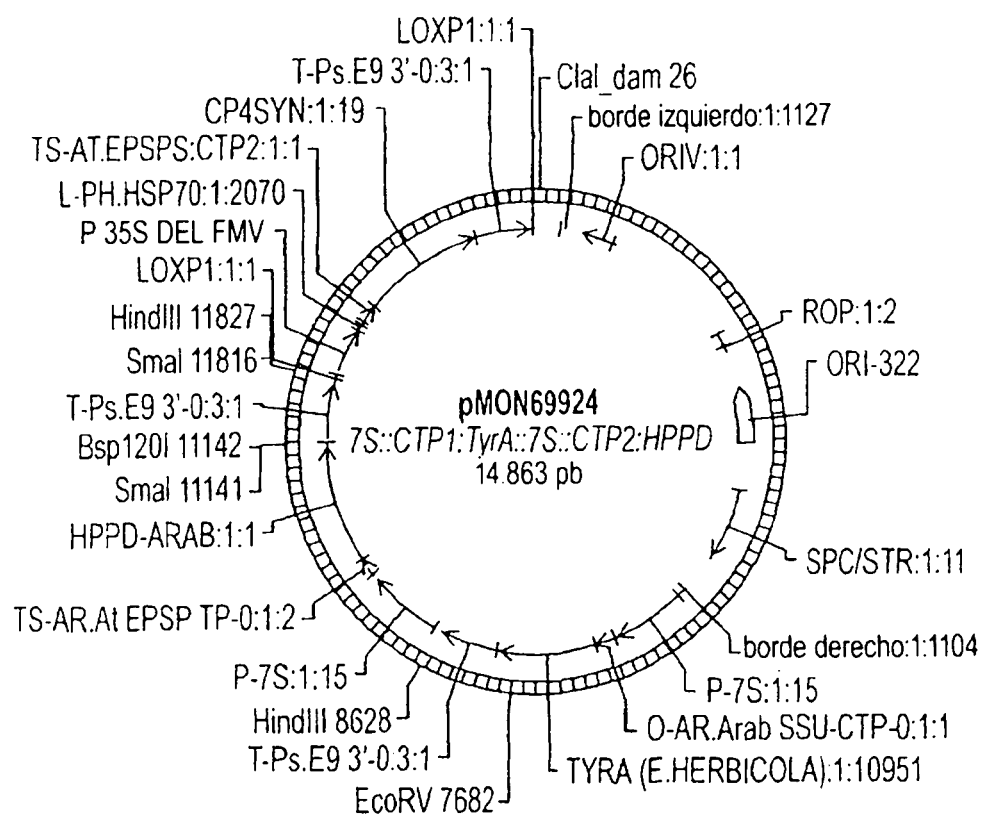


FIG. 50

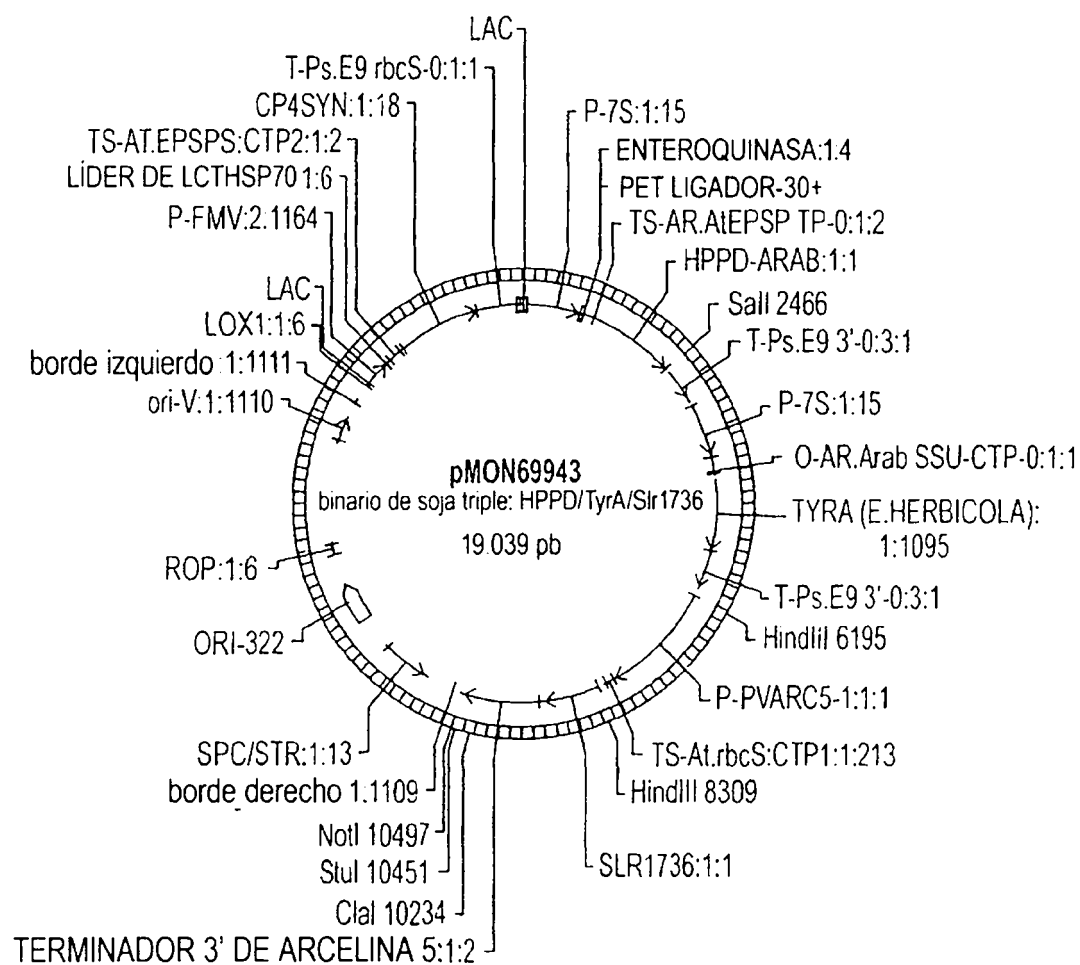


FIG. 51

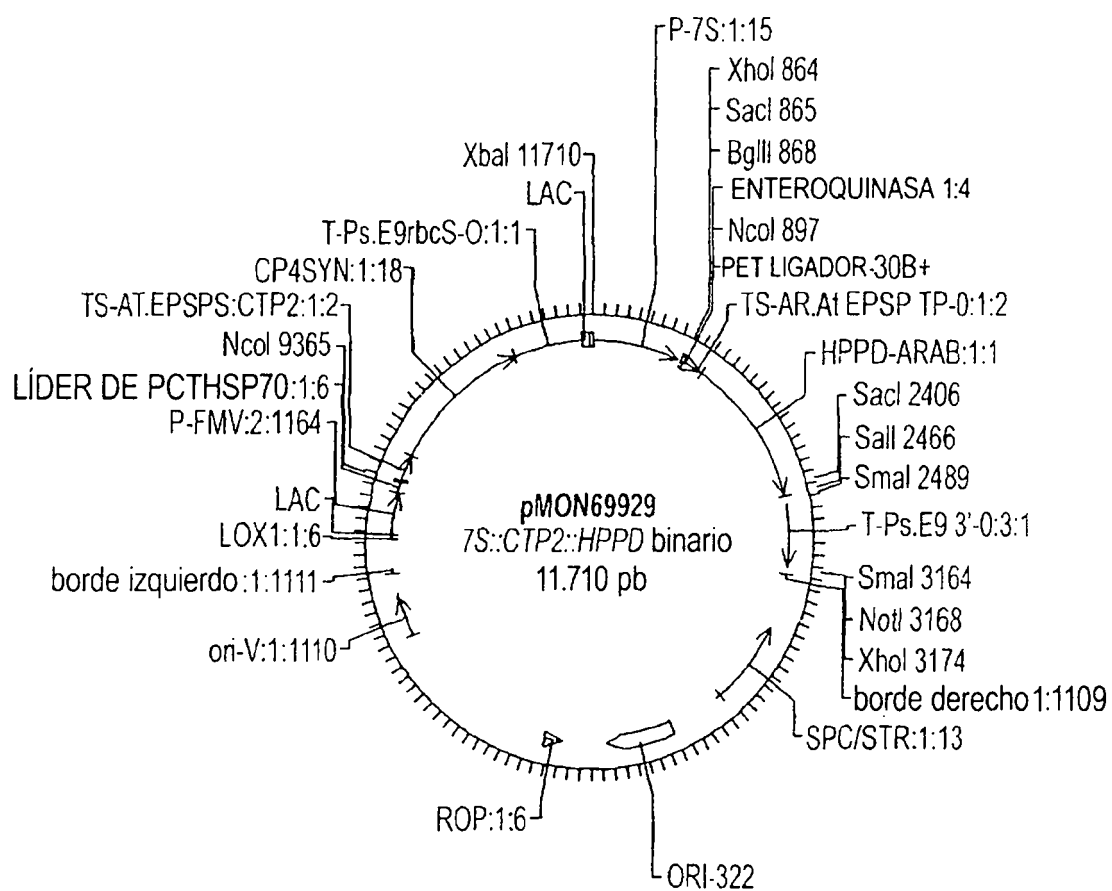


FIG. 52

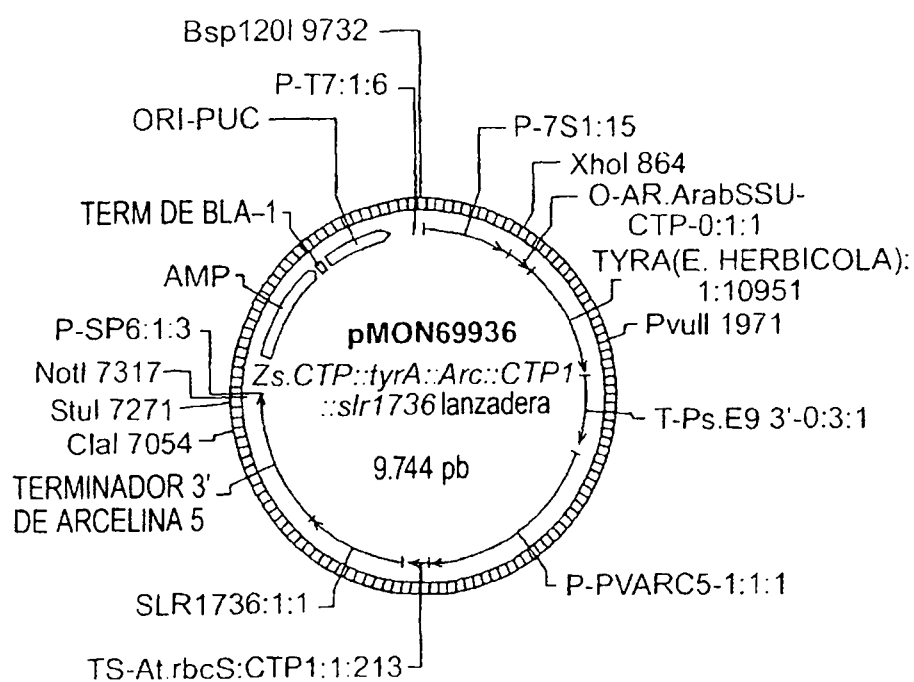


FIG. 53

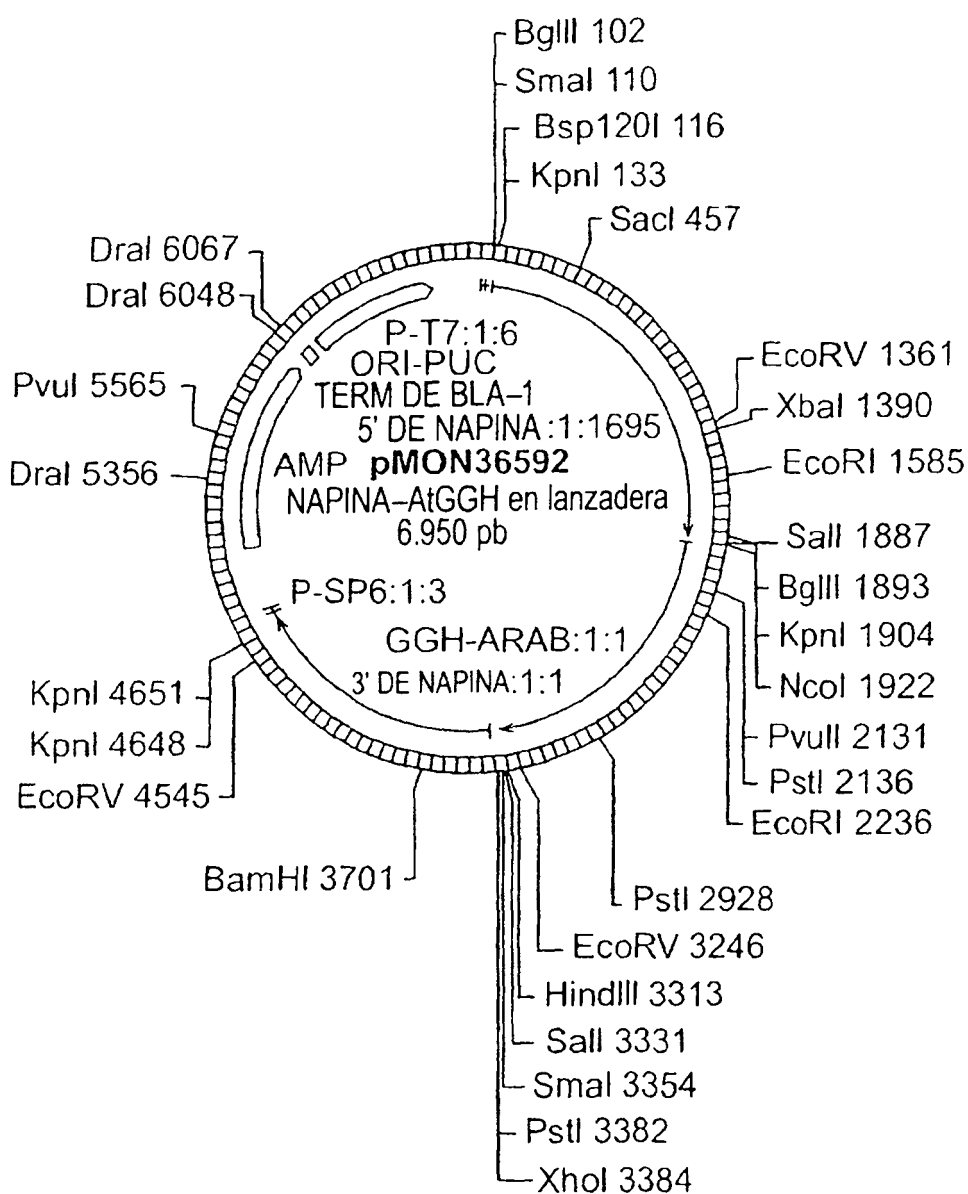


FIG. 54

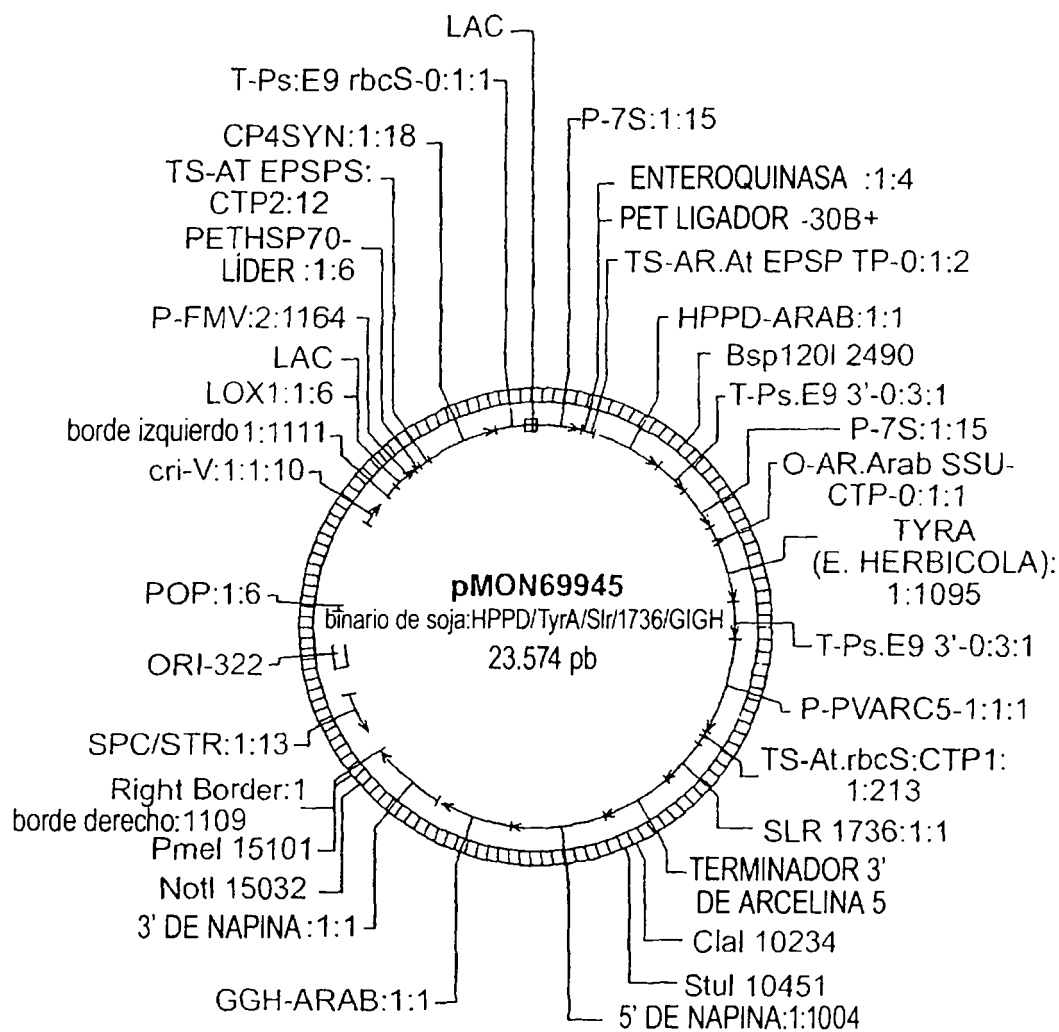


FIG. 55

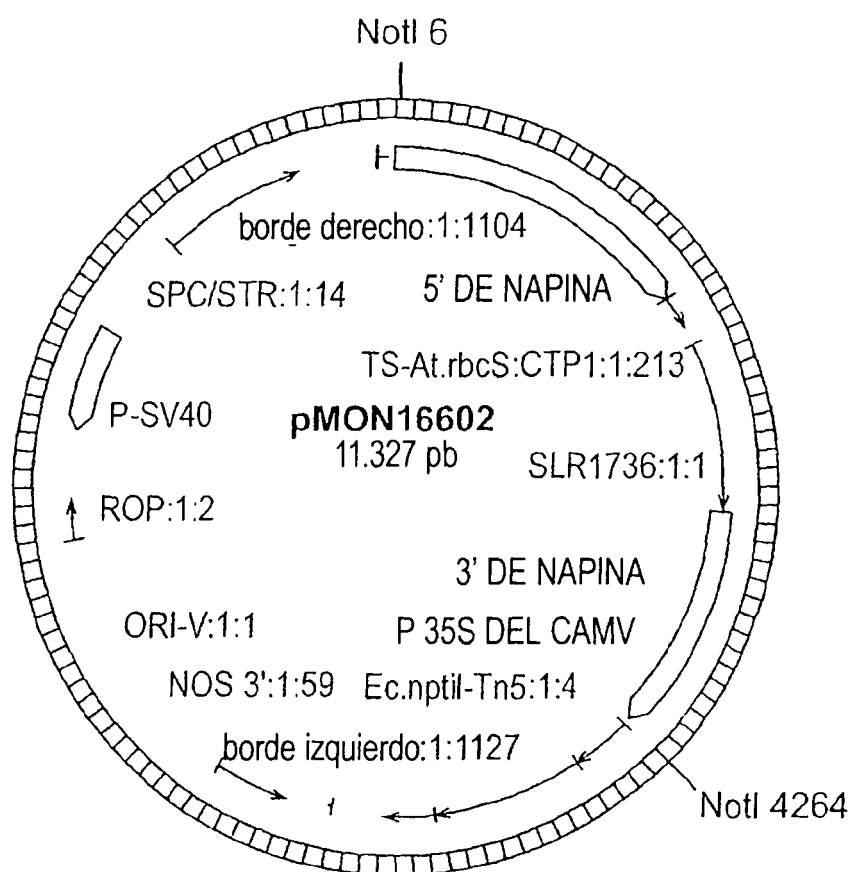


FIG. 56

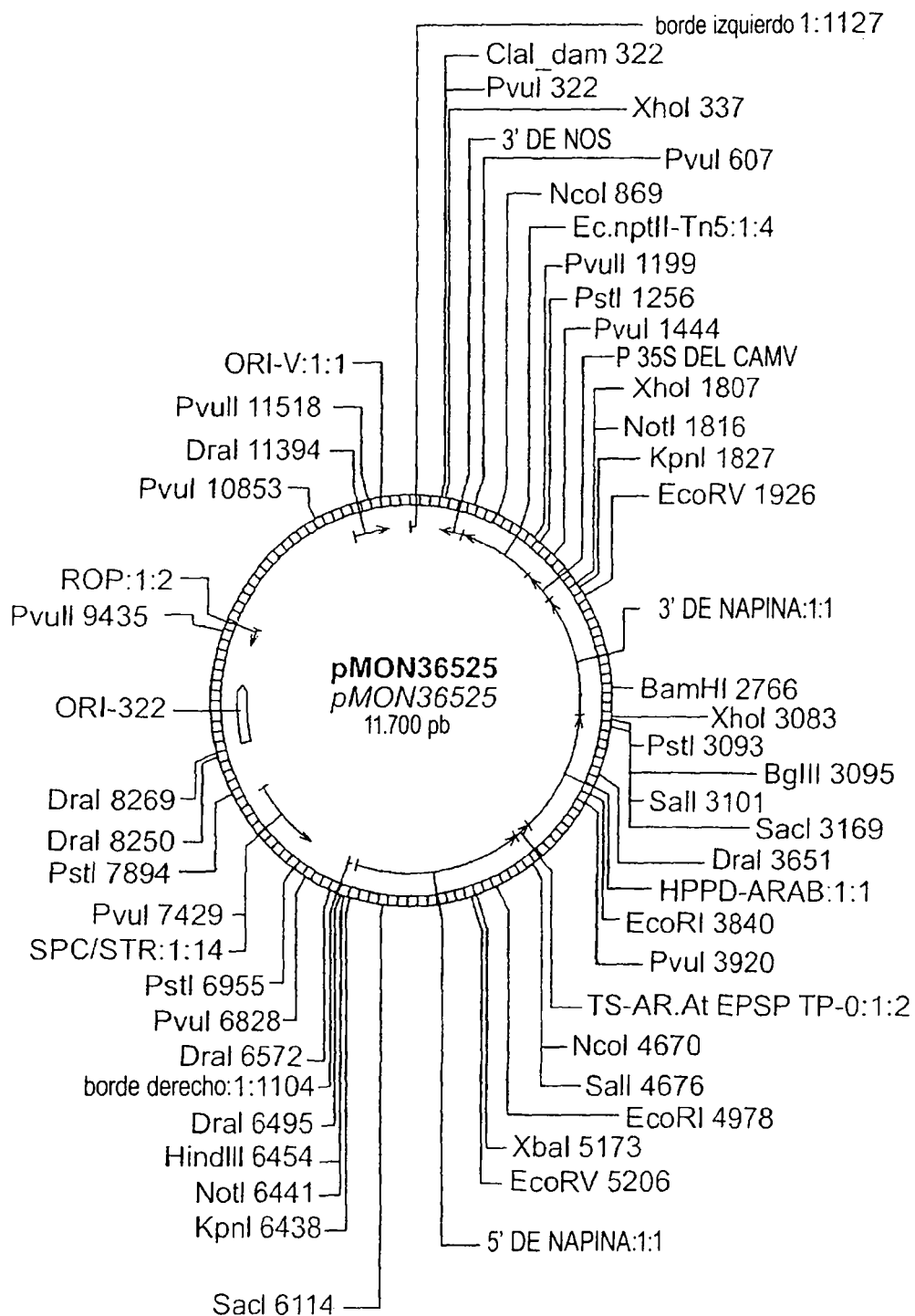


FIG. 57

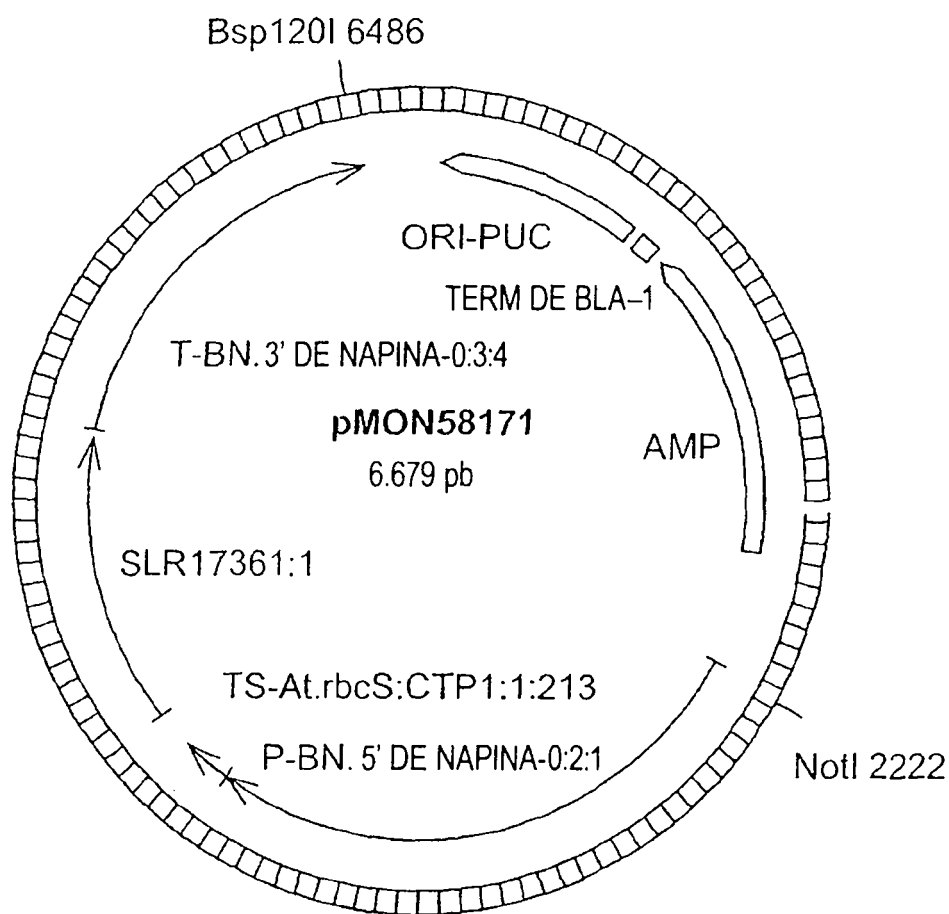


FIG. 58

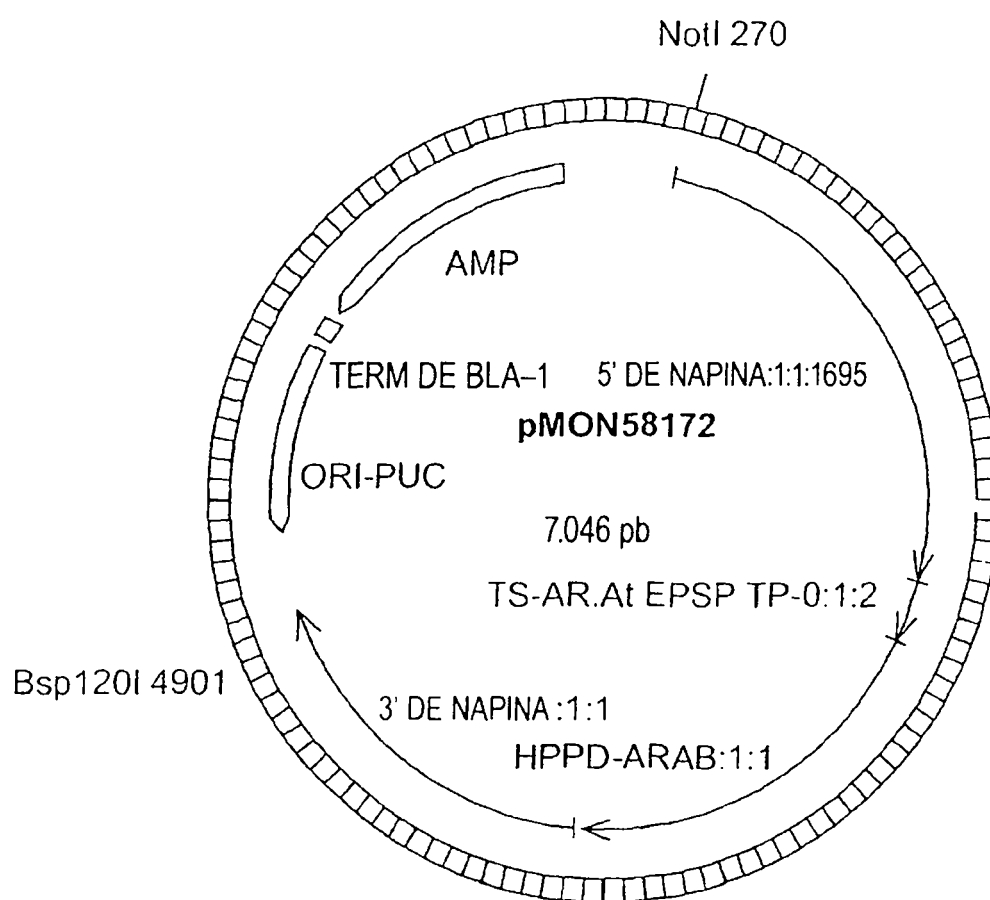


FIG. 59

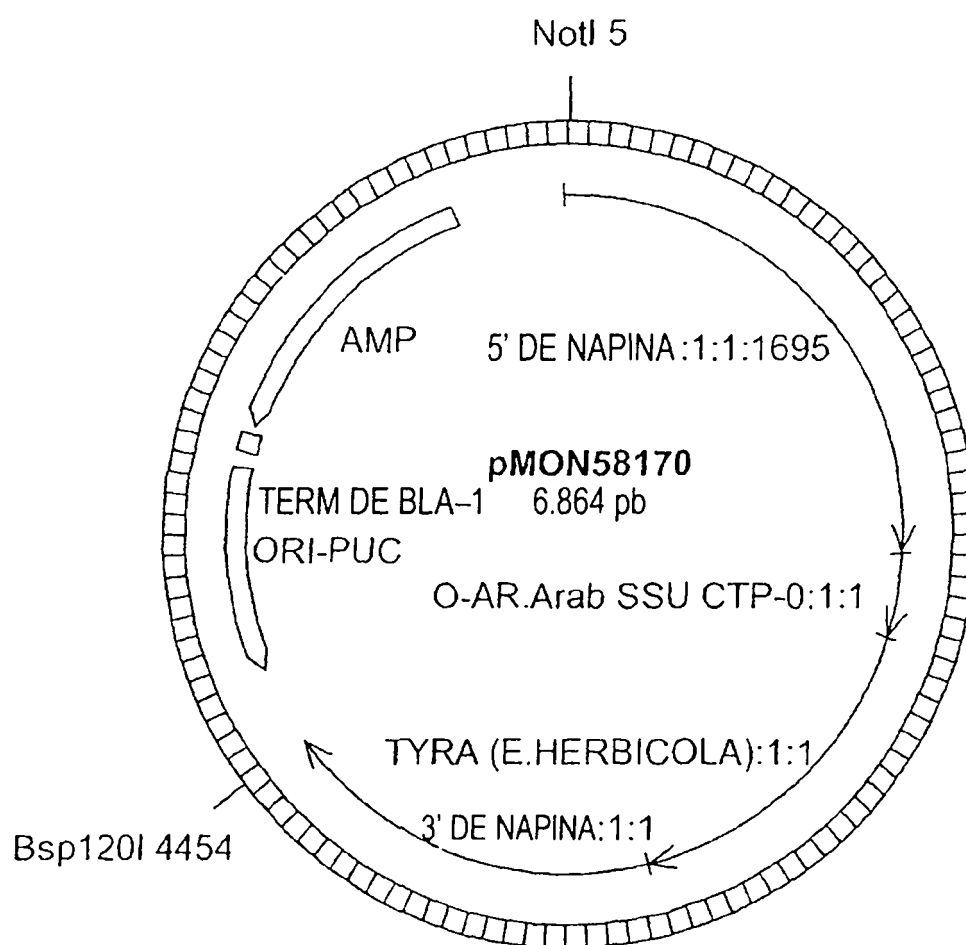


FIG. 60

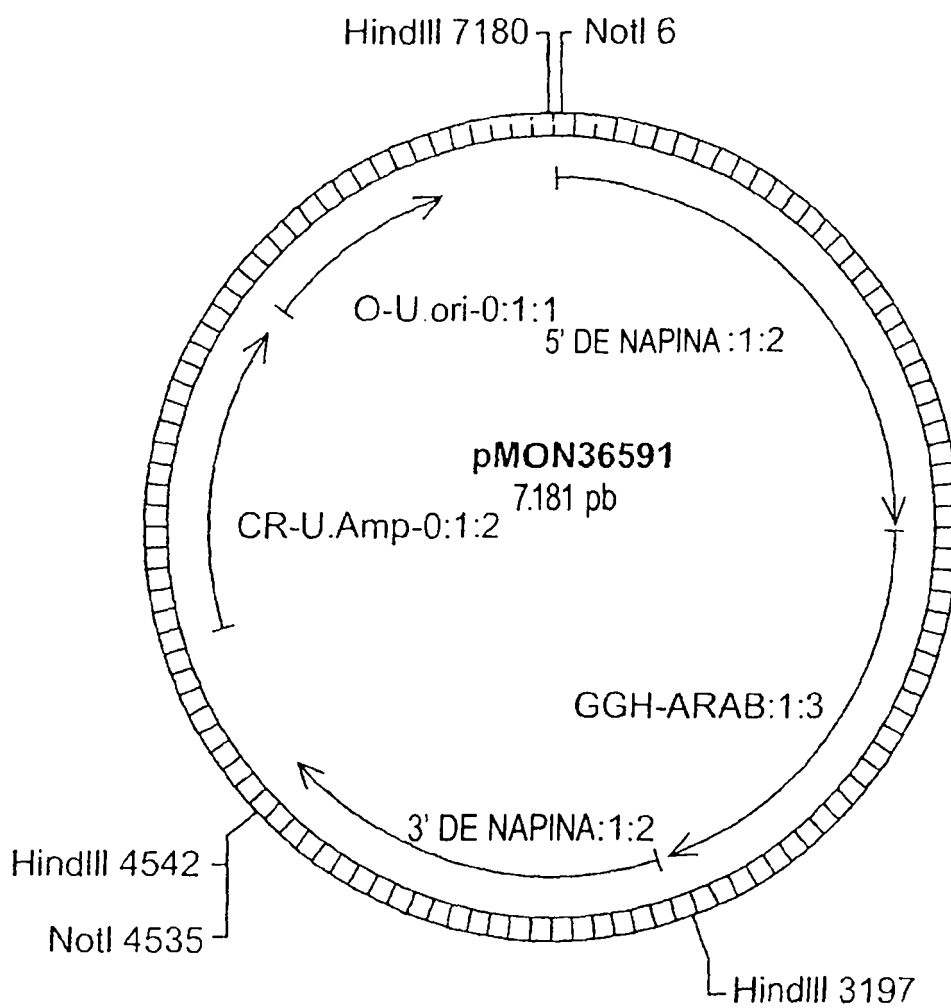


FIG. 61

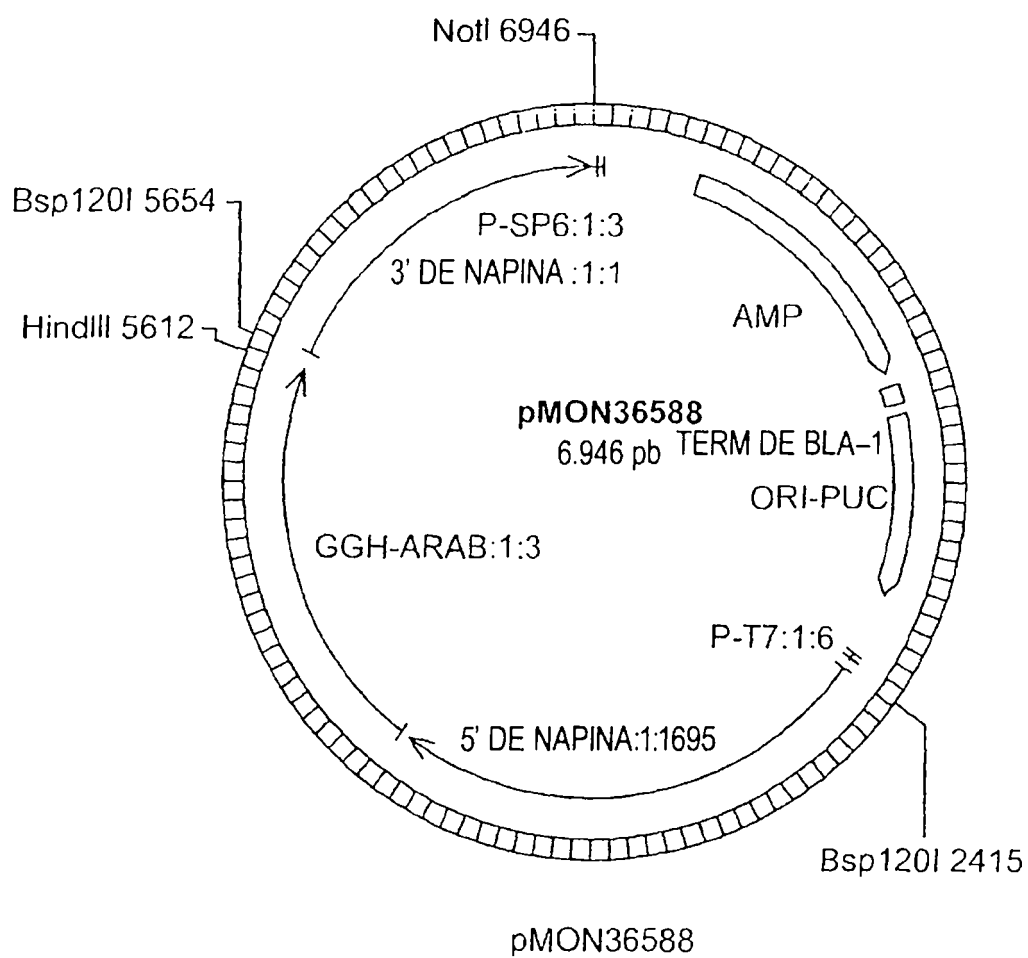


FIG. 62

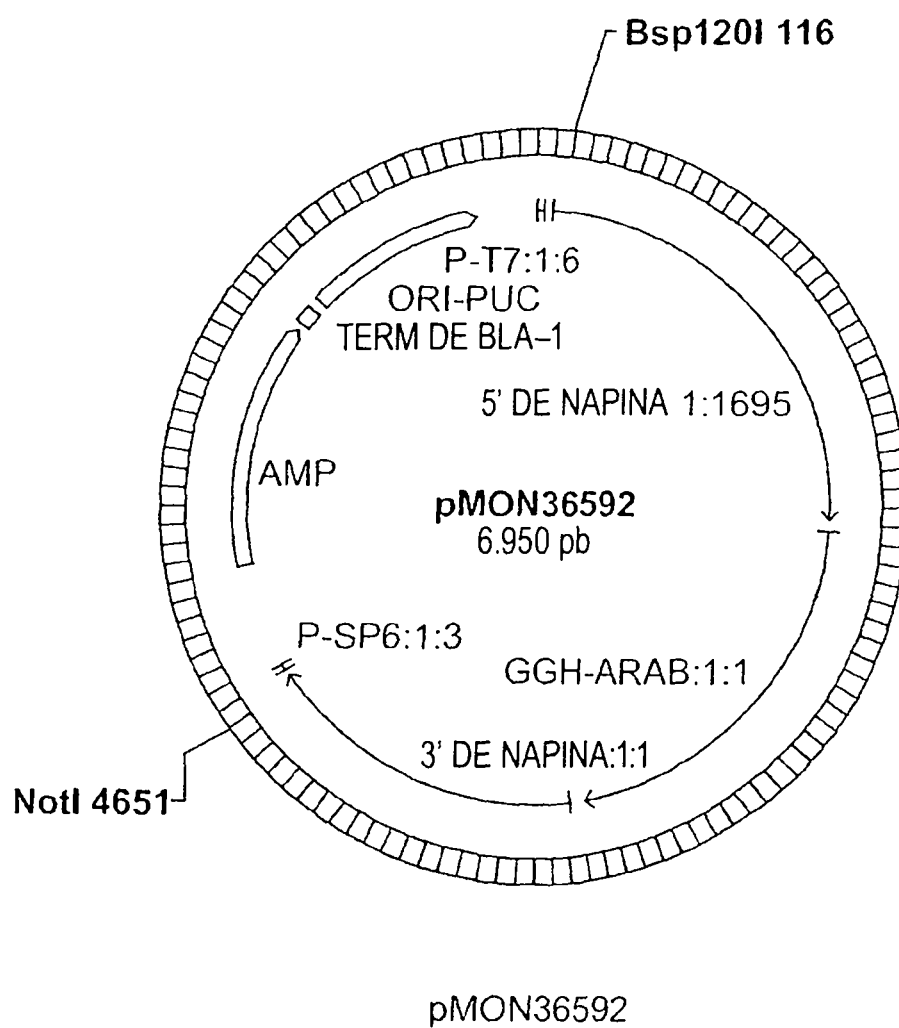


FIG. 63

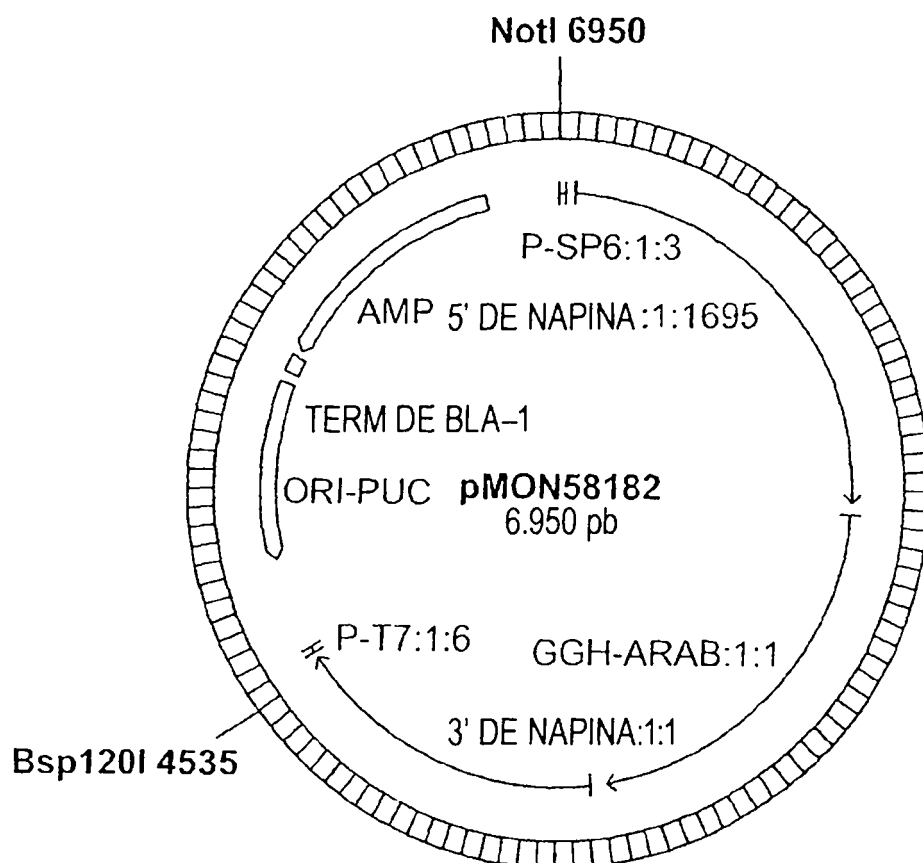


FIG. 64

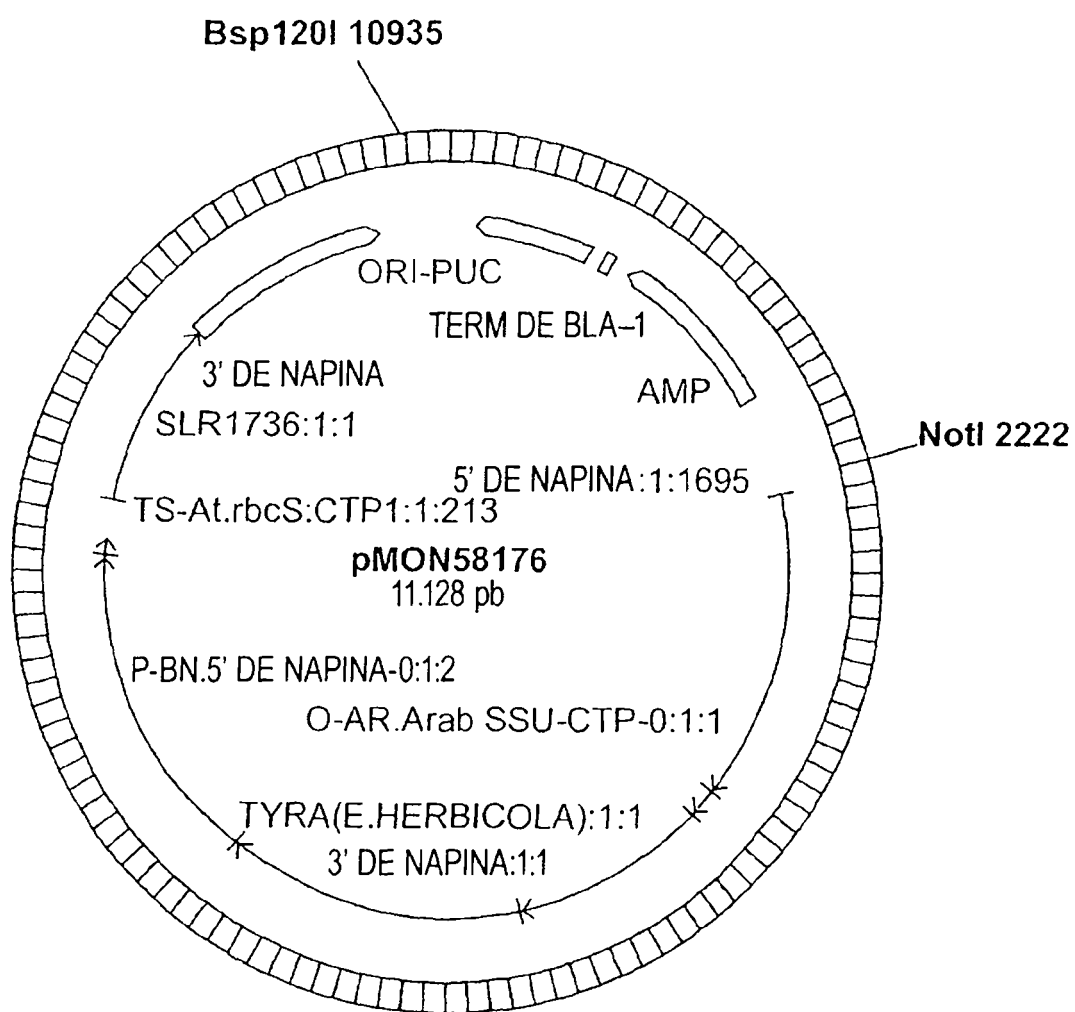


FIG. 65

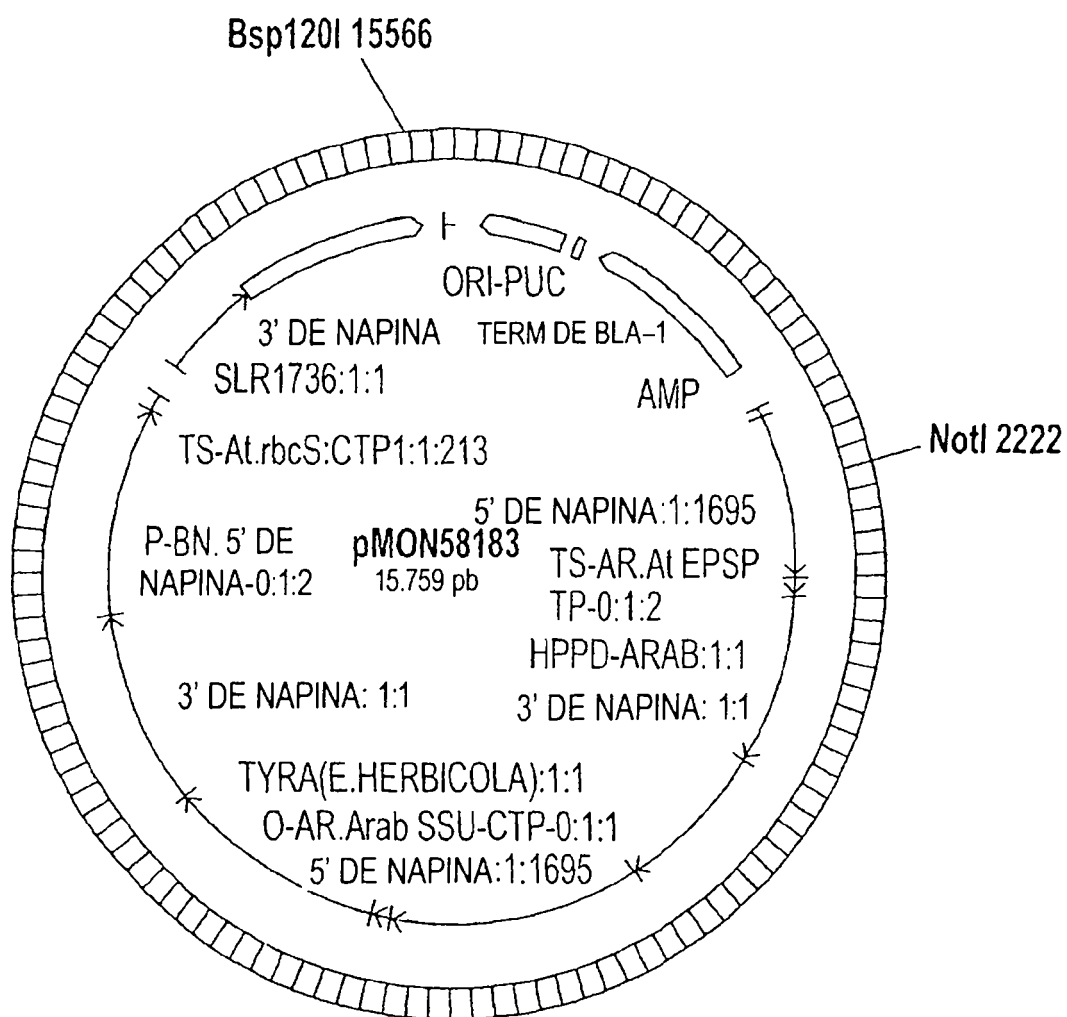


FIG. 66

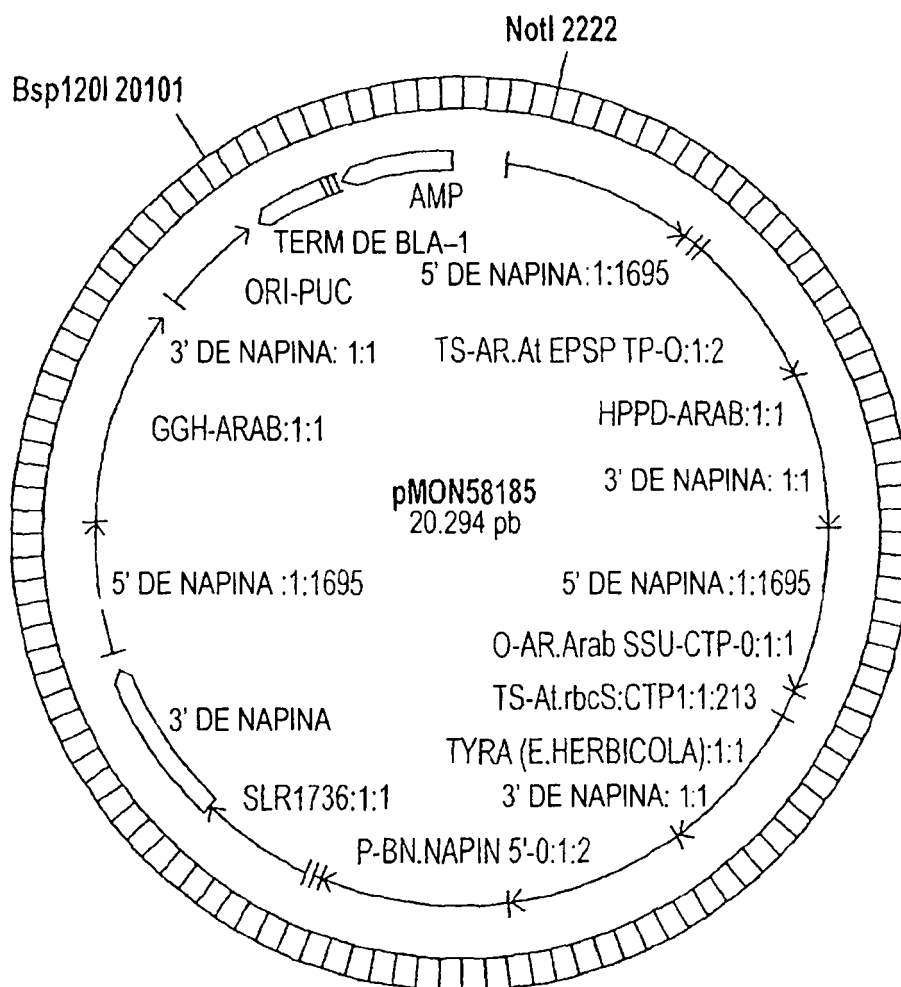


FIG. 67

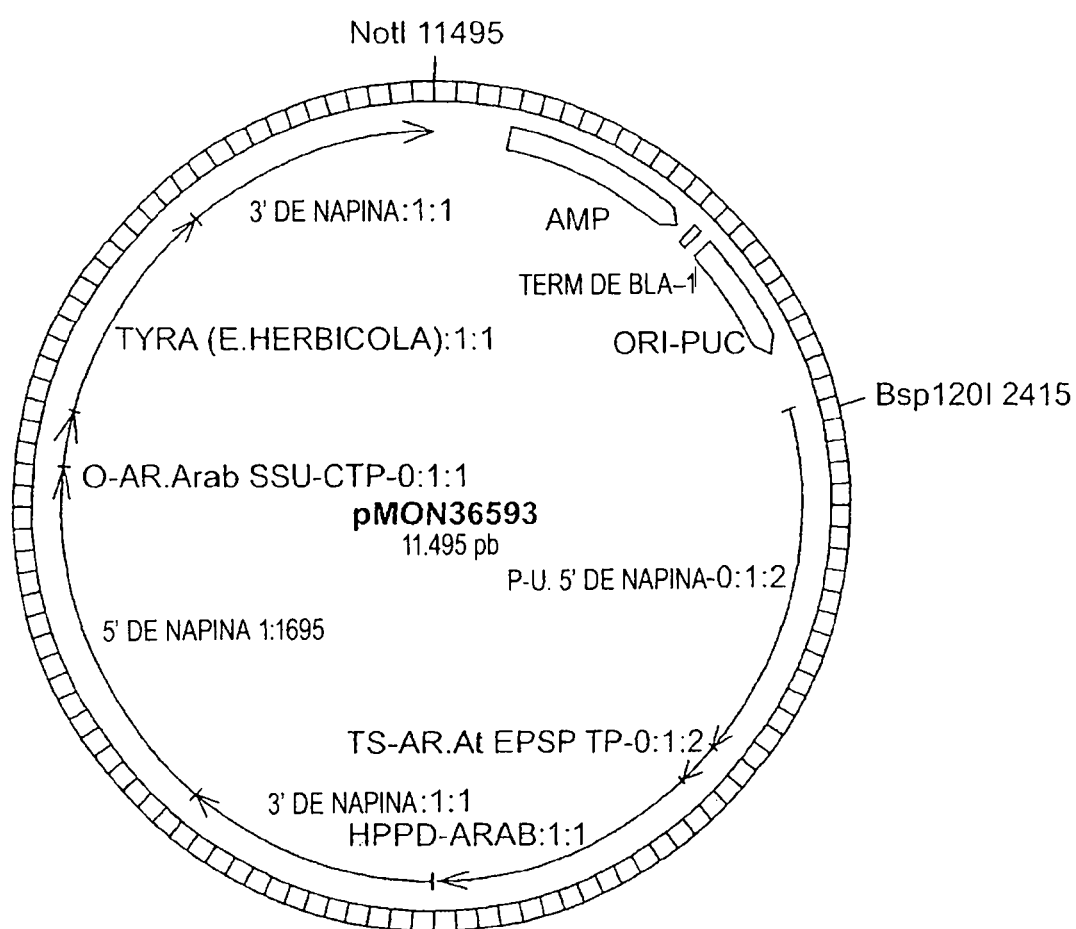


FIG. 68

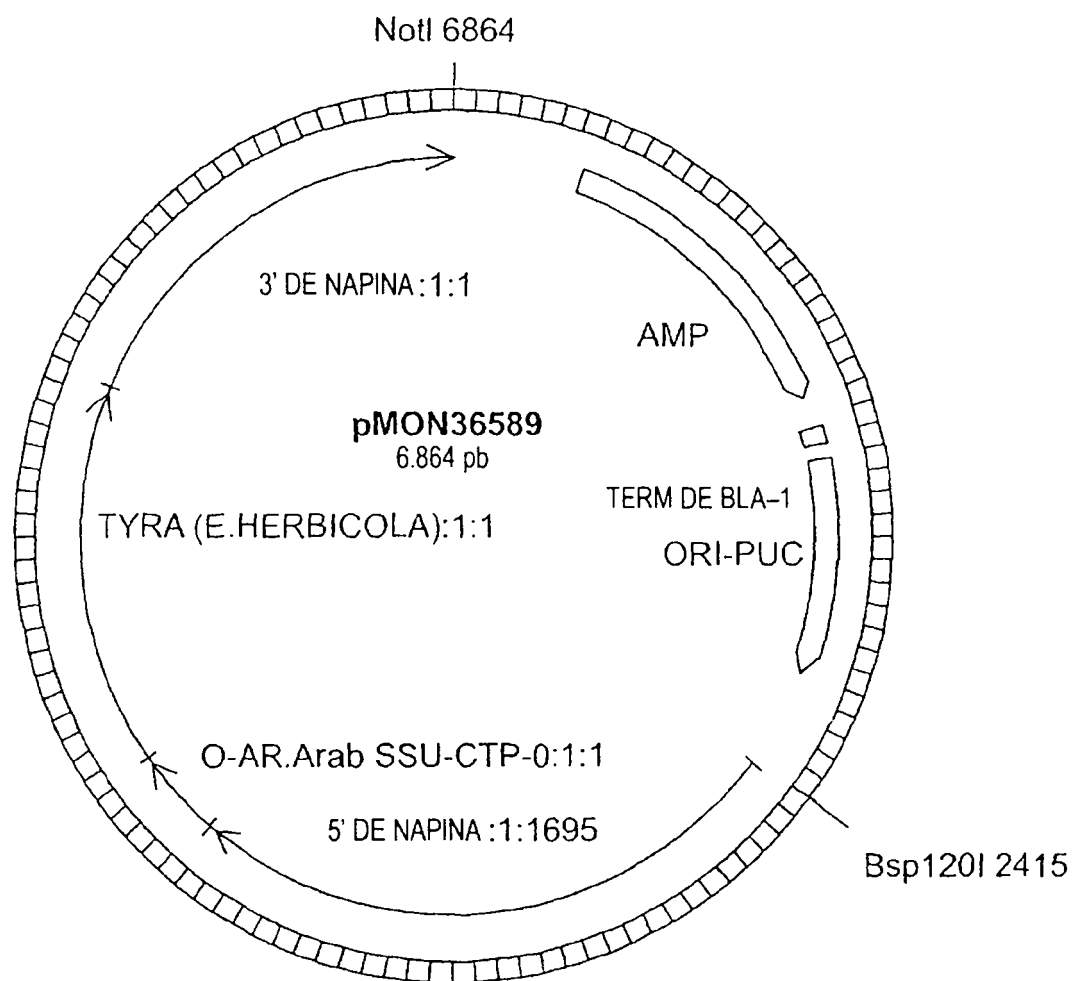


FIG. 69

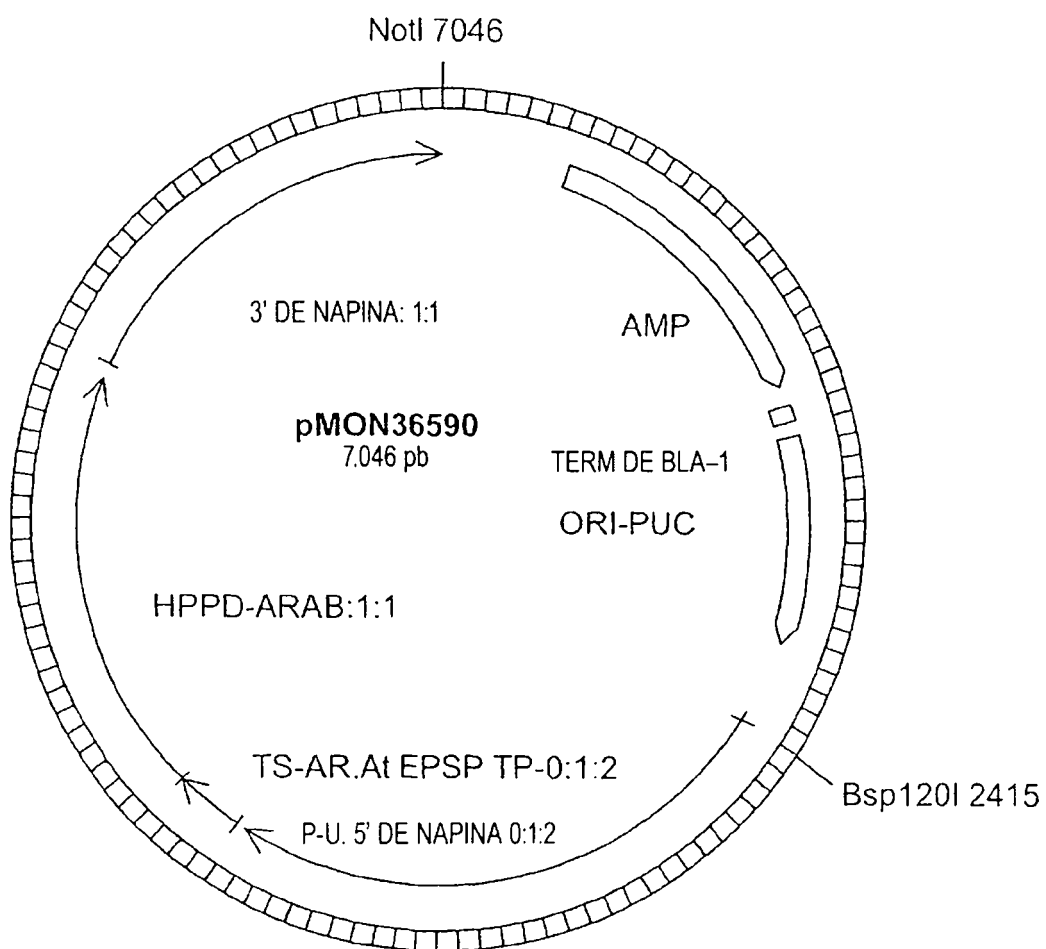


FIG. 70

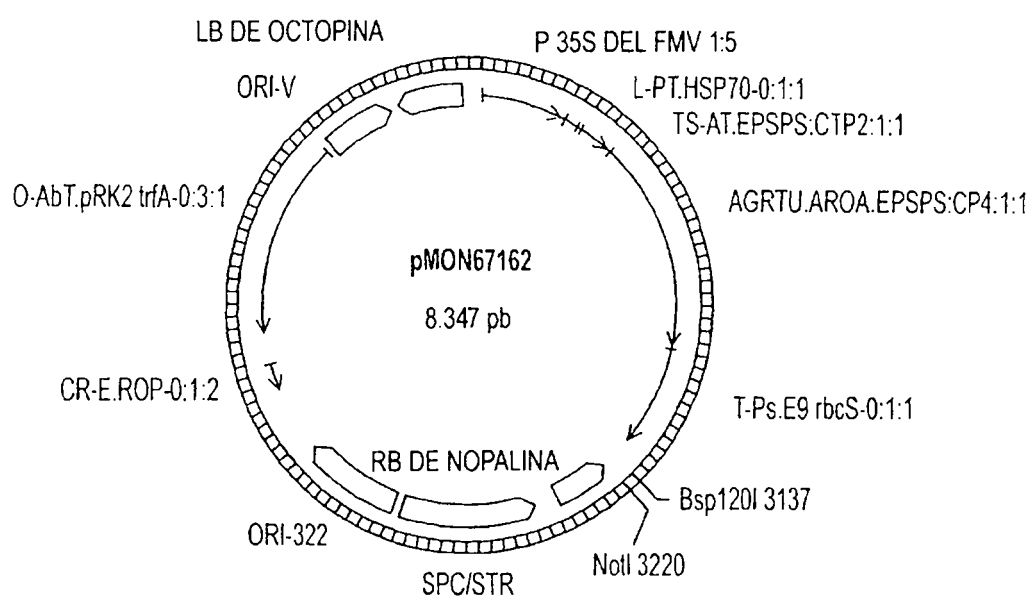


FIG. 71

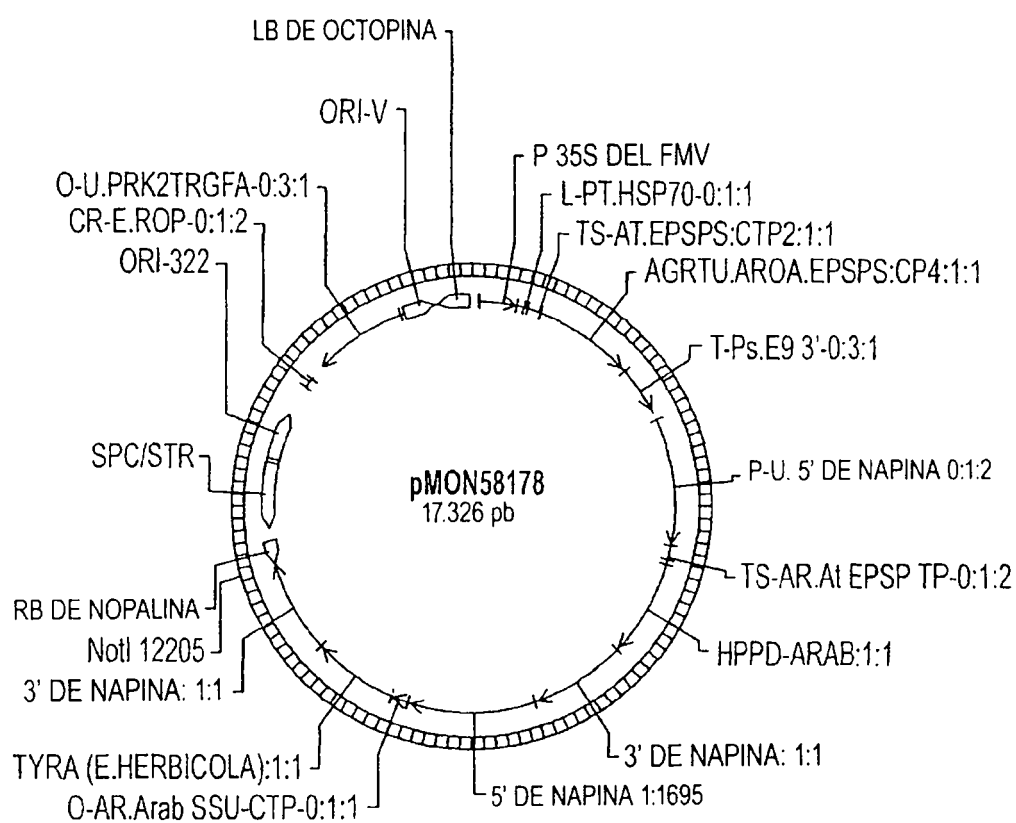


FIG. 72

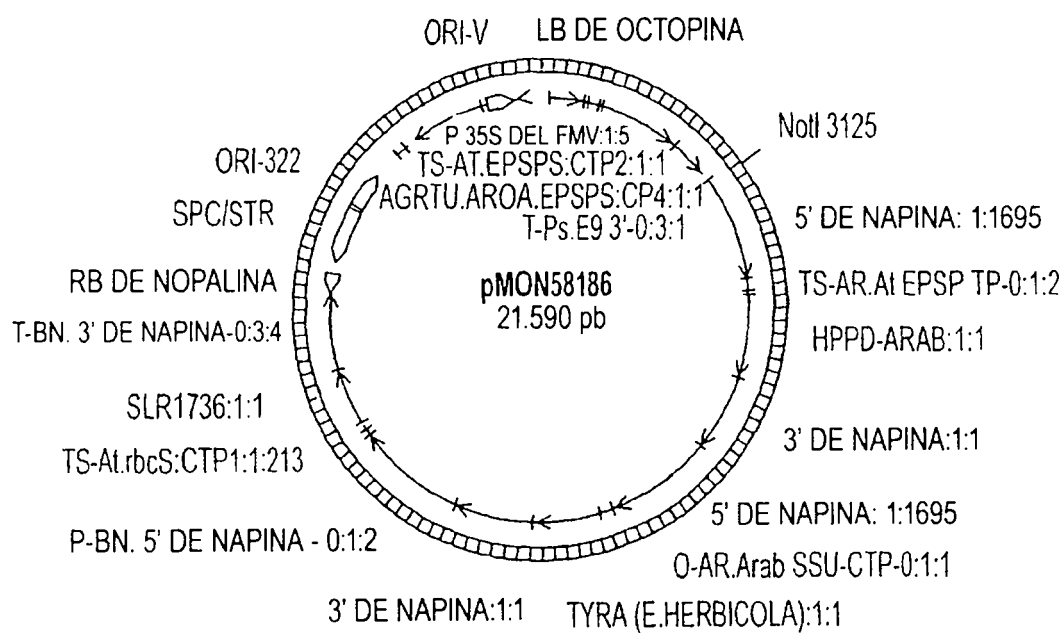


FIG. 73

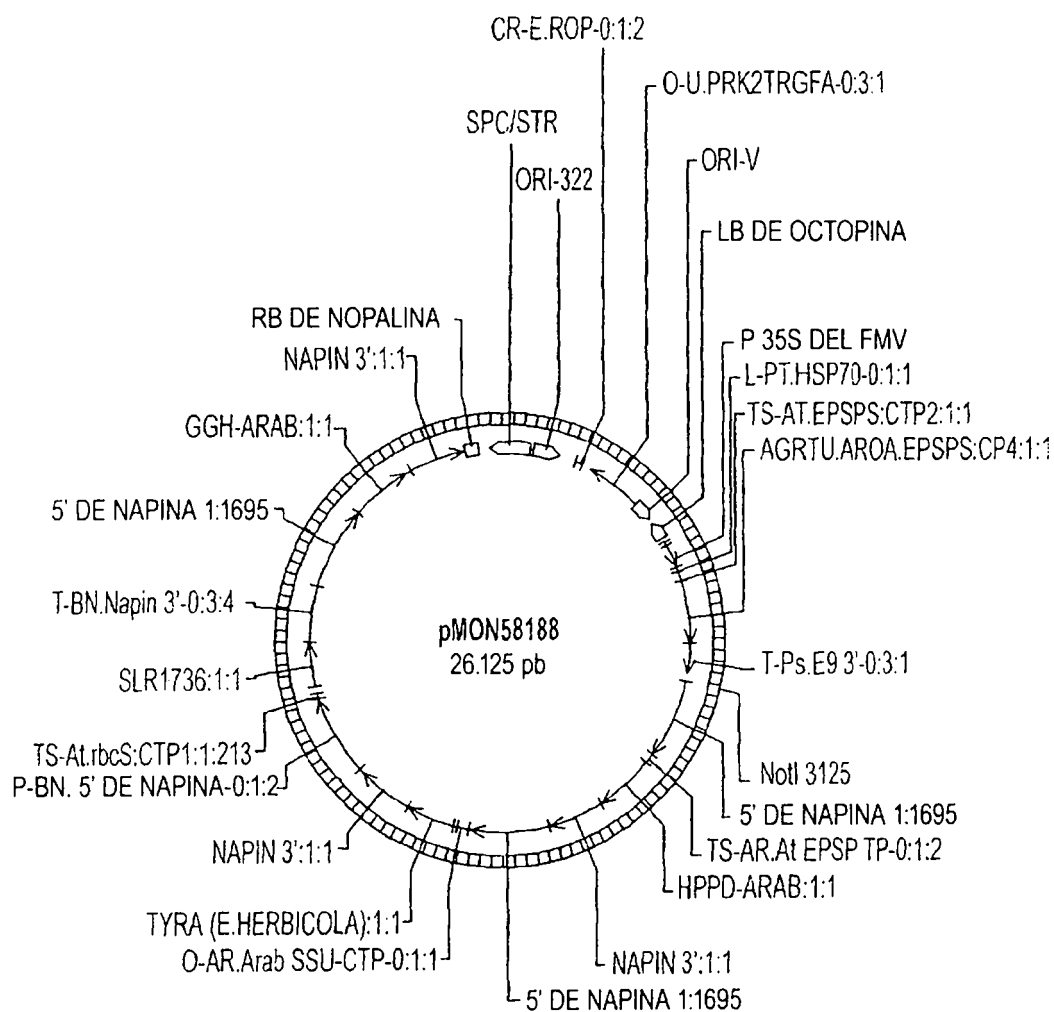


FIG. 74

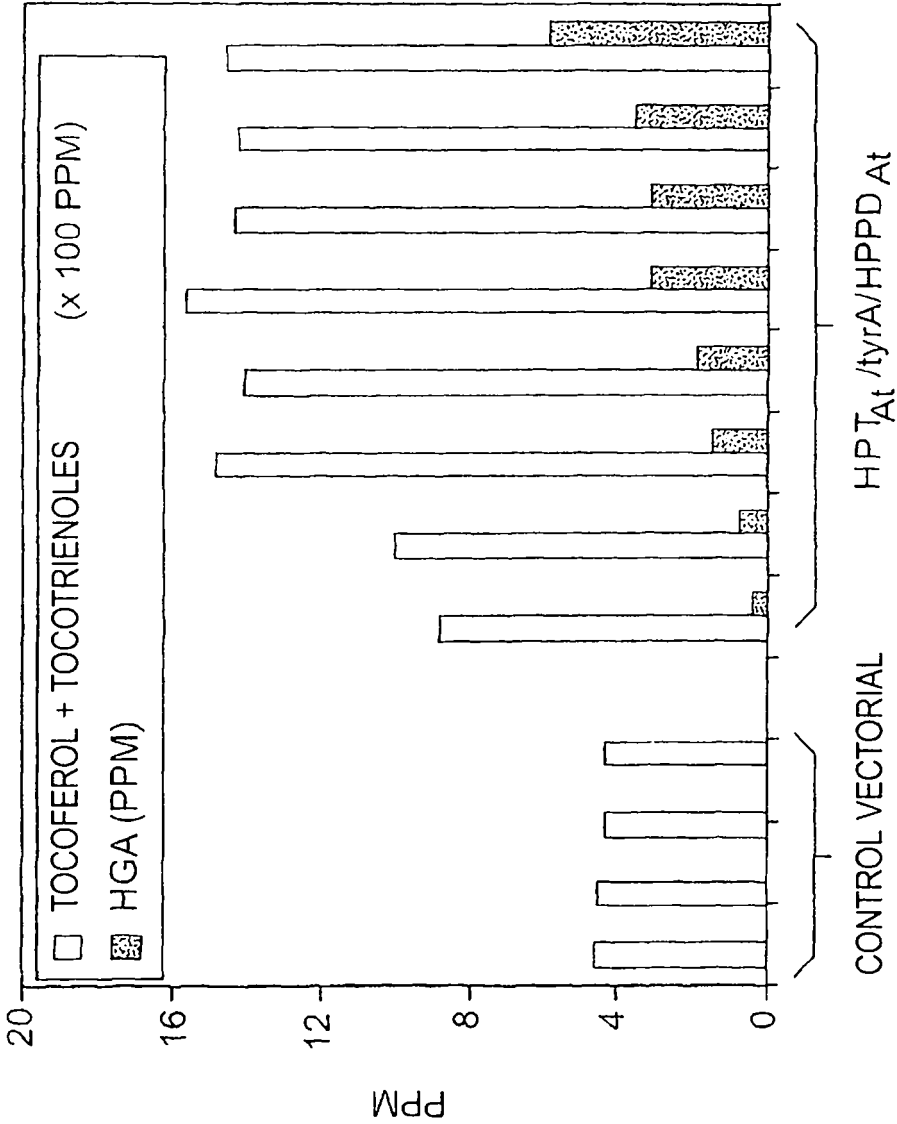


FIG. 75

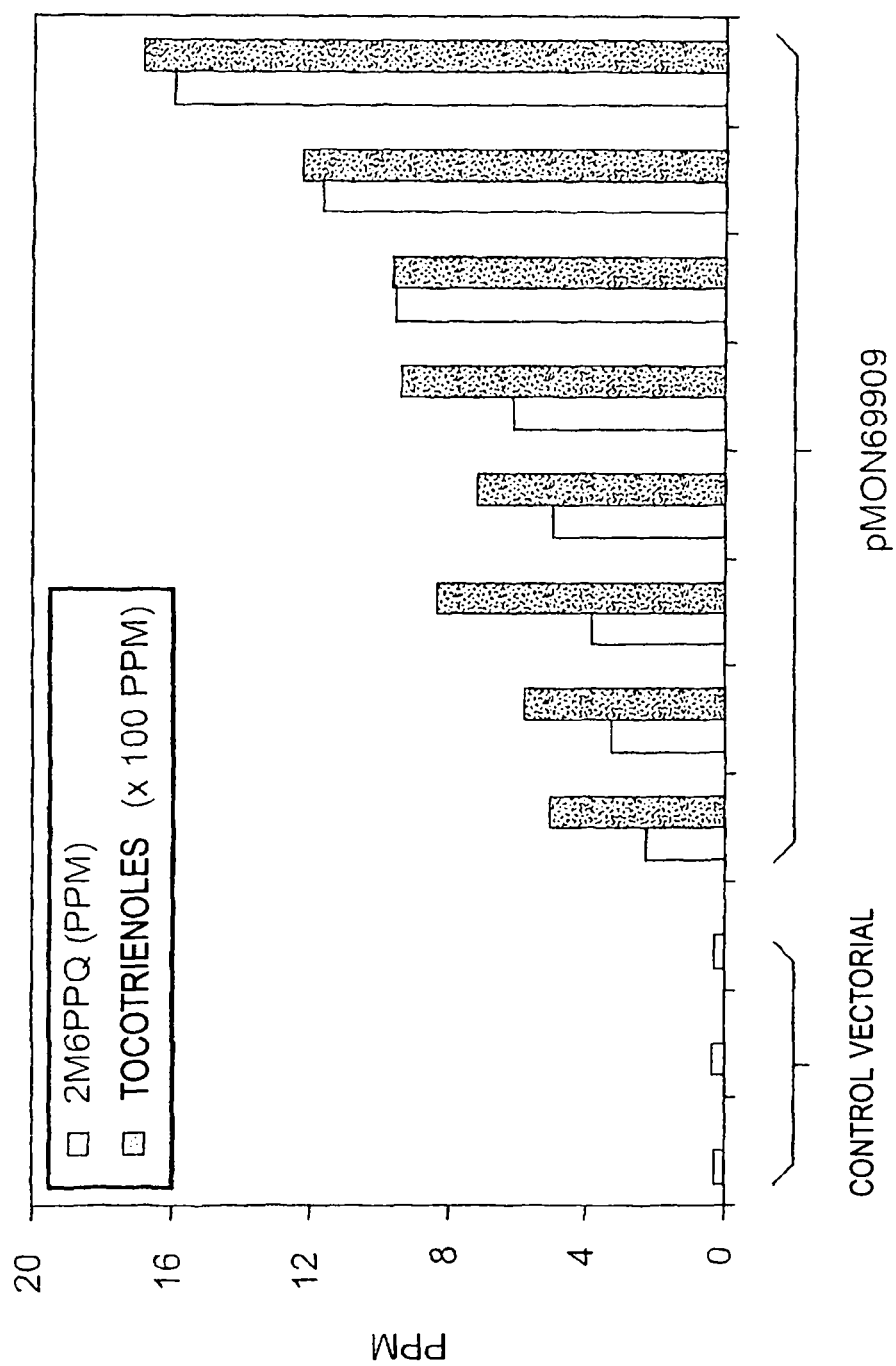


FIG. 76

ES 2 318 004 T3

LISTA DE SECUENCIAS

<110> Valentin, Henry E.
 Mitsky, Timothy A
 5
 <120> Genes TyrA y usos de los mismos
 <130> 16515.149
 10
 <140> Por ser asignada
 <141> 03-05-2002
 15 <150> US 60/289.527
 <151> 09-05-2001
 <160> 12
 20 <170> PatentIn versión 3.0
 <210> 1
 25 <211> 1122
 <212> ADN
 <213> *Erwinia herbicola*
 30 <400> 1
 atgggtggctg aactgaccgc gttacgcgat caaattgaca gtgtagataa agcgcctgctg 60
 gatctgctgg ctaagcgact ggaactggtg gccgaggtag gtgaggtgaa gagccggttac 120
 35 ggccctgccta tctatgtgcc tgagcgtgag gcgtcgatgc tggcttcgcg tcgcaaagag 180
 gccgaagcgc tcggcggtacc accggatctg attgaggatg tgctgcgtcg cgtgatgcgg 240
 gaatcctata ccagcgagaa tgataaaggc tttaaaaccc tctgtcctga actgcgcccc 300
 gtgggtgattg tcgggtggtaa gggccagatg ggccggctgt ttgaaaaaat gctcgggcta 360
 tcaggctaca cgggttaaaac gctggataaa gaggactggc ctcaggctga gactctgctc 420
 45 agcgatgcgc gaatggtgat cattagcgtg ccgattcacc tgaccgagca ggtgattgcc 480
 caactgccac cactgccgga agattgtatt ctggtcgatac tggcgctcagt caaaaaccgg 540
 cctctgcagg caatgctggc tgcccataac gggcctgtac tgggtctgca tccgatgttt 600
 50 ggcccggaaca gcggcagcct ggcaaaacag gtgggtggtct ggtgtgatgg aagacaaccg 660
 gaagcgtatc agtggttcct ggagcagatt cagggtctggg gtgcgcgtct gcacgcgtatc 720
 55 agcgcctgtt agcatgacca gaacatggca ttcattcagg cgctgcgtca ctttgctacc 780
 ttcgcttatg gtctgcattt agccgaagag aacgtcaatc tggatcagct gctggcgctc 840
 tcgtcgccca ttaccggct tgaactggcg atgggtggggc ggttggttcgc tcaggatccg 900
 60 caactctatg cggatatcat catgtcttca gagagtaatc tggcgctgat aaaacgctat 960
 taccagcggg ttgggtgaagc gattgcgctg ctggagcagg gcgacaagca ggcgtttatc 1020
 65 gccagcttta accgggttga acagtggttt ggcgatcacg caaaacgctt cctggtcgaa 1080
 agccgaagcc tgttgcgata ggccaatgac agccgcccac aa 1122

ES 2 318 004 T3

<210> 2

<211> 373

<212> PRT

5 <213> *Erwinia herbicola*

<400> 2

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

ES 2 318 004 T3

	Asn	Leu	Asp	Gln	Leu	Leu	Ala	Leu	Ser	Ser	Pro	Ile	Tyr	Arg	Leu	Glu	
5	Leu	Ala	Met	Val	Gly	Arg	Leu	Phe	Ala	Gln	Asp	Pro	Gln	Leu	Tyr	Ala	
10	Asp	Ile	Ile	Met	Ser	Ser	Glu	Ser	Asn	Leu	Ala	Leu	Ile	Lys	Arg	Tyr	
	Tyr	Gln	Arg	Phe	Gly	Glu	Ala	Ile	Ala	Leu	Leu	Glu	Gln	Gly	Asp	Lys	
15	Gln	Ala	Phe	Ile	Ala	Ser	Phe	Asn	Arg	Val	Glu	Gln	Trp	Phe	Gly	Asp	
	His	Ala	Lys	Arg	Phe	Leu	Val	Glu	Ser	Arg	Ser	Leu	Leu	Arg	Ser	Ala	
20																	
	Asn	Asp	Ser	Arg	Pro												
25	<210>	3															
	<211>	1122															
	<212>	ADN															
	<213>	<i>Escherichia coli</i>															
30	<400>	3															
35	atggttgctg	aattgaccgc	attacgcgat	caaattgatg	aagtcgataa	agcgcctgctg										60	
	aattttattag	cgaagcgtct	ggaactgggt	gctgaagtgg	gcgaggtgaa	aagccgccttt										120	
	ggactgccta	tttatgttcc	ggagcgagag	gcatctatgt	tggcctcgcg	tcgtgcagag										180	
40	gcggaagctc	tgggtgtacc	gccagatctg	attgaggatg	ttttgcgtcg	ggtgatgcgt										240	
	gaatcttact	ccagtgaaaa	cgacaaagga	tttaaaacgc	tttgtcctgc	gttacgcccg										300	
45	gtagttatcg	ttggcggcgg	cggtcagatg	ggacgtctgt	tcgagaagat	gctgacactc										360	
	tcgggttatc	aggtgcggat	tctggagcaa	catgactggg	atcgagcggc	tgatattgtt										420	
	gccgatgccg	gaatggtgat	tgttagtgtg	ccaatccacg	ttactgagca	agttattggc										480	
50																	
	aaattaccgc	ctttaccgaa	agattgtatt	ctggttgatc	tggcatcagt	gaaaaatgga										540	
	ccattacagg	ccatgctggc	ggcgcacgat	ggcccgggtac	tgggggttaca	cccaatgttc										600	
55	ggtccggaca	gcggtagcct	ggcaaagcaa	gttgtggtct	ggtgtgatgg	acgtaaaccg										660	
	gaagcatacc	aatggtttct	ggagcaaatt	caggtctggg	gcgctcgggt	gcatcgtatt										720	
	agcgcctgctg	agcacgatca	gaatatggcg	tttattcagg	cactgcgcca	ctttgctact										780	
60	tttgcttacg	ggctgcacct	ggcagaagaa	aatgttcagc	ttgagcaact	tctggcgctc										840	
	tcttcgccga	tttaccgcct	tgagctggcg	atggtcgggc	gactgttcgc	tcaggatccg										900	
	cagctttatg	ccgacattat	tatgtcgtca	gagcgtaatc	tggcggttaat	caaacgttac										960	
65	tataagcggt	tcggcgaggc	gattgagttg	ctggagcagg	gcgataagca	ggcgtttatt										1020	

ES 2 318 004 T3

gacagtttcc gcaaggtgga gcactgggtc ggcgattacg cacagcgttt tcagagtga 1080

agccgcgtgt tattgctca ggcgaatgac aatcgccagt aa 1122

5

<210> 4

<211> 373

10 <212> PRT

<213> *Escherichia coli*

<400> 4

15

Met Val Ala Glu Leu Thr Ala Leu Arg Asp Gln Ile Asp Glu Val Asp
1 5 10 15

20

Lys Ala Leu Leu Asn Leu Leu Ala Lys Arg Leu Glu Leu Val Ala Glu
20 25 30

Val Gly Glu Val Lys Ser Arg Phe Gly Leu Pro Ile Tyr Val Pro Glu
35 40 45

25

Arg Glu Ala Ser Met Leu Ala Ser Arg Arg Ala Glu Ala Glu Ala Leu
50 55 60

30

Gly Val Pro Pro Asp Leu Ile Glu Asp Val Leu Arg Arg Val Met Arg
65 70 75 80

Glu Ser Tyr Ser Ser Glu Asn Asp Lys Gly Phe Lys Thr Leu Cys Pro
85 90 95

35

Ala Leu Arg Pro Val Val Ile Val Gly Gly Gly Gly Gln Met Gly Arg
100 105 110

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

40

Glu Gln His Asp Trp Asp Arg Ala Ala Asp Ile Val Ala Asp Ala Gly
130 135 140

Met Val Ile Val Ser Val Pro Ile His Val Thr Glu Gln Val Ile Gly
145 150 155 160

45

Lys Leu Pro Pro Leu Pro Lys Asp Cys Ile Leu Val Asp Leu Ala Ser
165 170 175

Val Lys Asn Gly Pro Leu Gln Ala Met Leu Ala Ala His Asp Gly Pro
180 185 190

50

Val Leu Gly Leu His Pro Met Phe Gly Pro Asp Ser Gly Ser Leu Ala
195 200 205

Lys Gln Val Val Val Trp Cys Asp Gly Arg Lys Pro Glu Ala Tyr Gln
210 215 220

55

Trp Phe Leu Glu Gln Ile Gln Val Trp Gly Ala Arg Leu His Arg Ile
225 230 235 240

60

Ser Ala Val Glu His Asp Gln Asn Met Ala Phe Ile Gln Ala Leu Arg
245 250 255

His Phe Ala Thr Phe Ala Tyr Gly Leu His Leu Ala Glu Glu Asn Val

65

ES 2 318 004 T3

	260	265	270
5	Gln Leu Glu Gln Leu Leu Ala Leu Ser Ser Pro Ile Tyr Arg Leu Glu 275 280 285		
	Leu Ala Met Val Gly Arg Leu Phe Ala Gln Asp Pro Gln Leu Tyr Ala 290 295 300		
10	Asp Ile Ile Met Ser Ser Glu Arg Asn Leu Ala Leu Ile Lys Arg Tyr 305 310 315 320		
	Tyr Lys Arg Phe Gly Glu Ala Ile Glu Leu Leu Glu Gln Gly Asp Lys 325 330 335		
15	Gln Ala Phe Ile Asp Ser Phe Arg Lys Val Glu His Trp Phe Gly Asp 340 345 350		
20	Tyr Ala Gln Arg Phe Gln Ser Glu Ser Arg Val Leu Leu Arg Gln Ala 355 360 365		
	Asn Asp Asn Arg Gln 370		

25
 <210> 5
 <211> 25
 <212> ADN
 30 <213> secuencia artificial

<220>
 <223> cebador sintético
 35
 <400> 5

40 actgccatgg tggctgaact gaccg 25

<210> 6
 <211> 32
 45 <212> ADN
 <213> secuencia artificial

<220>
 50 <223> cebador sintético
 <400> 6

55 actggaattc ttattatggg cggctgtcat tg 32

<210> 7
 60 <211> 25
 <212> ADN
 <213> secuencia artificial

65 <220>
 <223> cebador sintético

ES 2 318 004 T3

<400> 7

5 actgccatgg ttgctgaatt gaccg 25

<210> 8

<211> 26

10 <212> ADN

<213> secuencia artificial

<220>

15 <223> cebador sintético

<400> 8

20 actggaattc ttattactgg cgattg 26

<210> 9

<211> 59

25 <212> ADN

<213> secuencia artificial

<230>

30 <223> cebador sintético

<400> 9

35 gatctcccgg gaagggcccc ggccgtctag agaattcgcg gccgcggcgc gccaccggt 59

<210> 10

40 <211> 59

<212> ADN

<213> secuencia artificial

<220>

45 <223> cebador sintético

<400> 10

50 tcgaaccggt ggcgcgccgc ggccgcgaat tctctagacg gccggggccc ttcccggga 59

<210> 11

<211> 37

<212> ADN

60 <213> secuencia artificial

<220>

<223> cebador sintético

65

ES 2 318 004 T3

<400> 11

ggatccgcg cgcacccatg gttgatcaag ttcagca 37

5

<210> 12

<211> 39

10 <212> ADN

<213> secuencia artificial

<230>

15 <223> cebador sintético

<400> 12

20 gagctcctgc aggaagcttt taggcacctc ctgatccgt 39

25

30

35

40

45

50

55

60

65