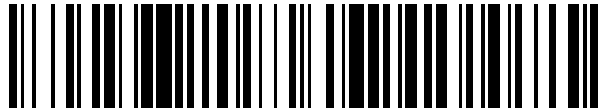


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 535 643**

51 Int. Cl.:

C12N 9/10	(2006.01)
C12N 5/04	(2006.01)
C12N 9/02	(2006.01)
C07H 21/04	(2006.01)
C12N 15/82	(2006.01)
A01H 5/00	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.04.2005 E 10011425 (5)**

97 Fecha y número de publicación de la concesión europea: **25.02.2015 EP 2295549**

54 Título: **Moléculas de ácido nucleico de nicotiana y usos de las mismas**

30 Prioridad:

29.04.2004 US 566235 P
03.09.2004 US 607357 P
03.09.2004 US 934944
17.09.2004 US 943507
15.10.2004 WO PCT/US2004/034218
15.10.2004 WO PCT/US2004/034065
25.01.2005 US 646764 P
24.03.2005 US 665097 P
24.03.2005 US 665451 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
13.05.2015

73 Titular/es:

U.S. SMOKELESS TOBACCO COMPANY LLC
(100.0%)
6603 West Broad Street
Richmond, VA 23230 , US

72 Inventor/es:

XU, DONGMEI y
NIELSEN, MARK, T.

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

Observaciones :

Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 535 643 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

Moléculas de ácido nucleico de *nicotiana* y usos de las mismas

La presente invención se refiere a enzimas citocromo p450 (en lo sucesivo, p450 y enzimas p450), en plantas de *Nicotiana*, en particular con plantas de tabaco que tienen una mutación en un gen endógeno para nicotina desmetilasa, con tabaco curado elaborado a partir de dichas plantas, y un método para reducir la expresión de nicotina desmetilasa en una planta de tabaco, como se define en las reivindicaciones.

Antecedentes

Durante la maduración o curado del tabaco se altera la expresión de diversos genes. Tales genes pueden afectar las rutas metabólicas implicadas en la formación de numerosos metabolitos secundarios, incluyendo terpenoides, polifenoles, y alcaloides que afectan a rasgos de calidad del producto final. Por ejemplo, la bioconversión de la nicotina para formar nornicotina durante la senescencia de la planta y en la fase de curado de la hoja o poscosecha, se produce en muchas especies de *Nicotiana*. La nicotina es la principal fuente de nornicotina. El alcaloide nornicotina, es un sustrato para la nitrosación mediada por microbios para formar la nitrosamina específica del tabaco (TSNA) N'-nitrosornicotina (NNN) durante el curado de la hoja o el almacenamiento y procesamiento posterior de la hoja.

Los genes expresados durante la maduración o el curado del tabaco pueden ser genes constitutivamente expresados, inducidos por etileno o relacionados con la senescencia, por ejemplo, los genes que codifican un citocromo p450. Los citocromos p450, por ejemplo, catalizan reacciones enzimáticas para una amplia gama de sustratos químicamente diferentes que incluyen al metabolismo oxidativo, peroxidativo, y reductivo de sustratos endógenos y xenobióticos. En las plantas, los p450 participan en las rutas bioquímicas que incluyen la síntesis de productos vegetales tales como fenilpropanoides, alcaloides, terpenoides, lípidos, glicósidos cianogénicos, y glucosinolatos estudiados (Chappell, Annu. Rev. Plant Physiol. Plant Mol. Biol. 46: 521 - 547, 1995). Los citocromos p450, también conocidos como proteínas hemo-tiolato p450, por lo general actúan como oxidasas terminales en las cadenas de transferencia de electrones de múltiples componentes, llamados sistemas monooxigenasa que contienen p450. Las reacciones específicas catalizadas por estos sistemas enzimáticos incluyen desmetilación, hidroxilación, epoxidación, N-oxidación, sulfoxidación, N, S, y O-desalquilaciones, desulfatación, desaminación y reducción de grupos azo, nitro, y N-óxido.

Los diversos papeles de las enzimas p450 de la planta de *Nicotiana* han sido implicados en la realización de una variedad de metabolitos vegetales tales como fenilpropanoides, alcaloides, terpenoides, lípidos, glicósidos cianogénicos, glucosinolatos, y una serie de otras entidades químicas. Algunas enzimas p450 pueden afectar la composición de metabolitos de las plantas. Por ejemplo, se ha deseado durante mucho tiempo mejorar el sabor y el aroma de ciertas plantas alterando un perfil de la planta de ácidos grasos seleccionados a través de fitomejoramiento; sin embargo, muy poco se sabe acerca de los mecanismos que intervienen en el control de los niveles de estos constituyentes de la hoja. La subregulación o sobre-regulación de las enzimas p450 asociadas con la modificación de los ácidos grasos puede facilitar la acumulación de ácidos grasos deseados que proporcionan cualidades fenotípicas más preferidas de la hoja.

La función de las enzimas p450 y su gran cantidad de funciones en componentes de las plantas aún están por descubrirse. Por ejemplo, se encontró que una clase especial de enzimas p450 cataliza la descomposición de ácidos grasos en aldehídos y β -alcoholes volátiles C6 y C9 que son los principales contribuyentes del olor "verde fresco" de frutas y verduras. El nivel de otras p450 novedosas específicas puede ser alterado para mejorar las cualidades de los constituyentes de la hoja mediante la modificación de la composición de lípidos y metabolitos de degradación relacionados en la hoja de *Nicotiana*. Varios de estos constituyentes en la hoja se ven afectados por la senescencia que estimula la maduración de las propiedades de calidad de la hoja. Sin embargo, otros informes han demostrado que las enzimas p450 juegan un papel funcional en la alteración de los ácidos grasos que están involucrados en las interacciones planta-patógeno y de resistencia a las enfermedades.

La gran multiplicidad de formas de la enzima p450, sus diferentes estructuras y funciones han hecho muy difícil su investigación sobre enzimas p450 de *Nicotiana* antes de la presente invención. Además, la clonación de enzimas p450 se ha visto obstaculizada por lo menos en parte debido a que estas proteínas localizadas en la membrana están típicamente presentes en baja cantidad y, a menudo son inestables durante la purificación. Por lo tanto, existe la necesidad para la identificación de enzimas p450 en plantas y las secuencias de ácido nucleico asociadas con esas enzimas p450. En particular, sólo se han reportado unas pocas proteínas citocromo p450 en *Nicotiana*. La invención descrita en la presente invención implica el descubrimiento de las citocromo p450 y fragmentos de citocromo p450 que corresponden a varios grupos de especies de p450 con base en su identidad de secuencia.

Además de las secuencias de p450, la presente invención abarca el descubrimiento de otras secuencias constitutivas y de etileno o inducidas por senescencia que abordan la necesidad de la regulación de las rutas

metabólicas implicadas en la formación de metabolitos secundarios que afectan la calidad de un producto de tabaco. Estas secuencias también son útiles en el desarrollo de germoplasmas de plantas que tienen características deseables para uso en programas de fitomejoramiento para desarrollar germoplasmas más deseables, y especialmente de germoplasmas que no son del tipo OGM (organismo genéticamente modificado).

5 Resumen de la invención

Los presentes inventores han identificado y caracterizado secuencias constitutivas, y de etileno e inducidas por senescencia, incluyendo un clon genómico de nicotina desmetilasa, del tabaco. También se describe aquí el uso de estas secuencias en los métodos de fitomejoramiento y en los métodos para crear una planta (por ejemplo, una planta transgénica) que tiene características deseables, tales como niveles alterados de nornicotina o N'-nitrosornicotina ("NNN") o ambos en relación con una planta de control.

En aspectos relacionados la invención presenta una planta de tabaco como se define en las reivindicaciones. En aún un aspecto relacionado adicional, la descripción presenta un cultivo de tejido de células de tabaco que pueden ser regeneradas obtenidas a partir de cualquiera de las plantas engendradas o producidas de acuerdo con los métodos descritos aquí. Tales cultivos de tejidos regeneran plantas de tabaco capaces de expresar todas las características fisiológicas y morfológicas de la planta de tabaco que tiene una variante de la expresión del gen para nicotina desmetilasa o un atributo modificado. Ejemplos de células que pueden ser regeneradas son embriones, células meristemáticas, semillas, polen, hojas, raíces, puntas de las raíces o flores o son protoplastos o callos derivados de los mismos.

En aspectos aún relacionados, la descripción presenta un método para la producción de un producto de tabaco que implica: (a) proporcionar una planta de tabaco producida de acuerdo con cualquiera de los métodos de fitomejoramiento antes mencionados; y (b) la preparación de un producto de tabaco a partir de la planta de tabaco. Los ejemplos de productos de tabaco incluyen hojas o tallos o ambos; un producto de tabaco sin humo; un tabaco húmedo o seco; tabacos para mascar; cigarrillos; cigarros; cigarritos; tabacos de pipa; o bidis.

Un aspecto adicional presenta un método para reducir la expresión o la actividad enzimática de un polipéptido del tabaco constitutivo, o inducido por etileno o inducido por senescencia en una célula vegetal. Este método implica reducir el nivel o la actividad enzimática de un polipéptido endógeno del tabaco constitutivo, o inducido por etileno o por senescencia en la célula vegetal. En una realización deseable, el polipéptido del tabaco es un p450. En otras realizaciones deseables, la célula vegetal es de una especie de *Nicotiana*.

Un aspecto adicional presenta una molécula aislada de ácido nucleico, por ejemplo, una secuencia de ADN, que contiene una secuencia de nucleótidos que codifica una nicotina desmetilasa. En realizaciones deseables, la secuencia de nucleótidos del primer aspecto es sustancialmente idéntica a una secuencia de nucleótidos que codifica una nicotina desmetilasa de tabaco, tal como la nicotina desmetilasa de tabaco que contiene una secuencia de nucleótidos que es al menos 70% idéntica a la secuencia de nucleótidos de la SEQ ID NO: 4 o la SEQ ID NO: 5, o que contiene los nucleótidos 2010 - 2949 y/o 3947 - 4562 de la SEQ ID NO: 4, o que contiene la secuencia de la SEQ ID NO: 4 o la SEQ ID NO: 5. La molécula aislada de ácido nucleico, por ejemplo, está operativamente enlazada a un promotor funcional en una célula vegetal y, deseablemente, está contenida en un vector de expresión. En otras realizaciones deseables, el vector de expresión está contenido en una célula, por ejemplo, una célula vegetal. Deseablemente, la célula vegetal, tal como una célula de una planta de tabaco, está incluida en una planta. La invención también presenta una semilla, por ejemplo, una semilla de tabaco, de una planta que contiene el vector de expresión, donde la semilla incluye una molécula aislada de ácido nucleico que hibrida bajo condiciones rigurosas con la secuencia de la SEQ ID NO: 4 operativamente enlazada a una secuencia heteróloga del promotor. Además, la invención presenta una planta derivada de una semilla germinada que contiene el vector de expresión, una hoja, ya sea verde o curada, de la planta, y un artículo de fabricación elaborado a partir de la hoja.

La secuencia de nucleótidos puede contener una secuencia que hibrida bajo condiciones rigurosas con el complemento de la secuencia de nucleótidos de la SEQ ID NO: 4 y/o la SEQ ID NO: 5, o con un fragmento de la SEQ ID NO: 4 o la SEQ ID NO: 5. Deseablemente, la secuencia de nucleótidos codifica una nicotina desmetilasa que es sustancialmente idéntica a la secuencia de aminoácidos de la SEQ ID NO: 3. La nicotina desmetilasa puede tener al menos 70% de identidad de secuencia de aminoácidos con la secuencia de aminoácidos de la nicotina desmetilasa de la SEQ ID NO: 3 o con un fragmento de una nicotina desmetilasa que tiene actividad enzimática alterada reducida en comparación con el polipéptido de longitud completa. De manera deseable, la nicotina desmetilasa incluye la secuencia de aminoácidos de la SEQ ID NO: 3.

En otro aspecto, la invención presenta una molécula aislada de ácido nucleico que contiene un promotor que hibrida bajo condiciones rigurosas con la secuencia de la SEQ ID NO: 8, o un fragmento de la misma que dirige la transcripción. Deseablemente, el promotor (i) se induce tras el tratamiento con etileno o durante la senescencia; y (ii) incluye (a) los pares de bases 1-2009 de la SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivas idénticas a 200 pares de bases consecutivos de la secuencia definida por los pares de bases 1-2009 de la SEQ ID

NO : 4, o (c) una porción de nucleótidos de 20 pares de bases idéntica en secuencia a una porción de 20 pares de bases consecutivas de la secuencia expuesta en los pares de bases 1-2009 de la SEQ ID NO: 4.

5 Un aspecto adicional de la invención presenta un promotor aislado de ácido nucleico que contiene una secuencia de nucleótidos que tiene 50% o más de identidad de secuencia con la secuencia de la SEQ ID NO: 8. Deseablemente, este promotor aislado de ácido nucleico es inducido después de tratamiento con etileno o durante la senescencia y, por ejemplo, incluye la secuencia de la SEQ ID NO: 8. Alternativamente, el promotor puede incluir un fragmento que puede ser obtenido a partir de la SEQ ID NO: 8, en donde el fragmento dirige la transcripción de un gen heterólogo o reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica). En una realización deseable la secuencia del promotor está operativamente enlazada a una secuencia heteróloga de ácido nucleico, y puede, por ejemplo estar contenido en un vector de expresión. En otras realizaciones deseables, el vector de expresión está contenido en una célula, por ejemplo, una célula vegetal. Deseablemente, la célula vegetal, tal como una célula de una planta de tabaco, está incluida en una planta. En otra realización deseable, la invención presenta una semilla, por ejemplo, una semilla de tabaco, de una planta que contiene el vector de expresión, donde la semilla incluye una molécula aislada de ácido nucleico que se hibrida bajo condiciones rigurosas con la secuencia de la SEQ ID NO: 8 operativamente enlazada a una secuencia heteróloga de ácido nucleico. Además, la invención presenta una planta derivada de una semilla germinada que contiene el promotor de este aspecto, una hoja, ya sea verde o curada, de la planta, y un artículo de fabricación elaborado a partir de la hoja.

20 En otro aspecto, la invención presenta un método para reducir la expresión de la nicotina desmetilasa en una planta de tabaco. Este método incluye las etapas de (i) introducir en la planta de tabaco un vector que contiene la secuencia de la SEQ ID NO: 8 o un fragmento que puede ser obtenido a partir de la SEQ ID NO: 8 operativamente enlazado a una secuencia heteróloga de ácido nucleico y (ii) expresar el vector en la planta de tabaco. En una realización deseable de este método, la expresión de la nicotina desmetilasa es silenciada. En otra realización deseable, el vector expresa ARN, tal como ARN antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi).

25 En un aspecto adicional deseable, la invención presenta una molécula aislada de ácido nucleico que contiene un intrón que hibrida bajo condiciones rigurosas con la secuencia de la SEQ ID NO: 7, o un fragmento de la misma que reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión del gen) o puede servir como un marcador molecular para identificar las secuencias de ácido nucleico de la nicotina desmetilasa. En una realización deseable, el intrón incluye (a) los pares de bases 2950 - 3946 de la SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivas idénticas a 200 pares de bases consecutivas de la secuencia definida por los pares de bases 2950 - 3946 de la SEQ ID NO: 4, o (c) una porción de nucleótidos de 20 pares de bases idéntica en secuencia a una porción de 20 pares de bases consecutivas de la secuencia expuesta en los pares de bases 2950-3946 de la SEQ ID NO: 4.

35 Otro aspecto deseable de la invención presenta un intrón aislado de ácido nucleico que incluye una secuencia de nucleótidos que tiene 50% o más de identidad de secuencia con la secuencia de la SEQ ID NO: 7, o un fragmento de la misma que reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica) o puede servir como un marcador molecular para identificar secuencias de ácido nucleico de la nicotina desmetilasa. El silenciamiento de la expresión génica, por ejemplo, implica la recombinación homóloga (por ejemplo, usando la secuencia de la SEQ ID NO: 188, o un fragmento de la misma) o una mutación que resulta en un producto génico que no tiene actividad de nicotina desmetilasa. En particular, el intrón puede incluir la secuencia de la SEQ ID NO: 7 o un fragmento que puede ser obtenido a partir de la SEQ ID NO: 7. Deseablemente, una molécula aislada de ácido nucleico que incluye un intrón está operativamente enlazada a una secuencia heteróloga de ácido nucleico y esta secuencia deseablemente está incluida en un vector de expresión. En otra realización, el vector de expresión está contenido en una célula, tal como una célula vegetal. En particular, la célula puede ser una célula de tabaco. Una planta, por ejemplo, una planta de tabaco, incluyendo una planta con una célula vegetal que contiene la secuencia de la SEQ ID NO: 7 o un fragmento que puede ser obtenido a partir de la SEQ ID NO: 7 operativamente enlazado a una secuencia heteróloga de ácido nucleico en un vector de expresión es otra realización deseable. Además, una semilla, por ejemplo, una semilla de tabaco, de una planta, en donde la semilla contiene un intrón que hibrida bajo condiciones rigurosas con la SEQ ID NO: 7 operativamente enlazada a una secuencia heteróloga de ácido nucleico es también deseable. Además, la invención presenta una planta derivada de la semilla germinada que contiene el intrón de este aspecto, una hoja, ya sea verde o curada, de la planta, y un artículo de fabricación elaborado a partir de la hoja verde o curada.

55 En aún otro aspecto, la invención presenta un método para reducir la expresión de la nicotina desmetilasa en una planta de tabaco. Este método incluye las etapas de (i) introducir en la planta de tabaco un vector que contiene la secuencia de la SEQ ID NO: 7 o un fragmento que puede ser obtenido a partir de la SEQ ID NO: 7 operativamente enlazado a una secuencia heteróloga de ácido nucleico y (ii) expresar el vector en la planta de tabaco. En una realización deseable de este método, la expresión de la nicotina desmetilasa es silenciada. En otra realización deseable, el vector expresa ARN, tal como ARN antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi).

5 En un aspecto adicional, la invención presenta una molécula aislada de ácido nucleico que contiene una región no traducida que hibrida bajo condiciones rigurosas con la secuencia de la SEQ ID NO: 9 o un fragmento de la misma que puede alterar el patrón de expresión de un gen, reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica), o puede ser utilizada como un marcador para identificar las secuencias de ácido nucleico de la nicotina desmetilasa. En una realización deseable de este aspecto, la región no traducida incluye (a) los pares de bases 4563-6347 de la SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivas idénticas a las 200 pares de bases consecutivas de la secuencia definida por los pares de bases 4563-6347 de la SEQ ID NO: 4, o (c) una porción de nucleótidos 20 pares de bases idéntica en secuencia a una porción de 20 pares de bases consecutivas de la secuencia expuesta en los pares de bases 4563-6347 de la SEQ ID NO: 4.

10 Un aspecto deseable adicional presenta una región no traducida de ácido nucleico aislado que contiene una secuencia de nucleótidos que tiene 50% o más identidad de secuencia con la secuencia de la SEQ ID NO: 9. Deseablemente, la región no traducida incluye la secuencia de la SEQ ID NO: 9 o la región no traducida incluye un fragmento que puede ser obtenido a partir de la SEQ ID NO: 9 que puede alterar el patrón de expresión de un gen, reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica), o puede ser utilizado como un marcador para identificar secuencias de ácido nucleico de nicotina desmetilasa. La región no traducida deseablemente está operativamente enlazada a una secuencia heteróloga de ácido nucleico y puede estar contenida en un vector de expresión. Además, este vector de expresión está contenido deseablemente en una célula, tal como una célula vegetal, por ejemplo, una célula de tabaco. Otra realización deseable presenta una planta, tal como una planta de tabaco, que incluye una célula vegetal que contiene un vector que incluye una secuencia aislada de ácido nucleico que tiene 50% o más identidad de secuencia con la secuencia de la SEQ ID NO: 9 y está operativamente enlazada a una secuencia heteróloga de ácido nucleico.

25 La invención también presenta una semilla, por ejemplo, una semilla de tabaco, de una planta, donde la semilla incluye una región no traducida que hibrida bajo condiciones rigurosas con la SEQ ID NO: 9 operativamente enlazada a una secuencia heteróloga de ácido nucleico. Además, la invención presenta una planta derivada de una semilla germinada que contiene la región no traducida de este aspecto de la invención, una hoja, ya sea verde o curada, de la planta, y un artículo de fabricación elaborado a partir de la hoja verde o curada.

30 Además, la invención presenta un método para reducir la expresión o alterar la actividad enzimática de la nicotina desmetilasa en una planta de tabaco. Este método incluye las etapas de (i) introducir en la planta de tabaco un vector que contiene una secuencia aislada de ácido nucleico que tiene 50% o más identidad de secuencia con la secuencia de la SEQ ID NO: 7 y está operativamente enlazada a una secuencia heteróloga de ácido nucleico y (ii) expresar el vector en la planta de tabaco. De manera deseable, la expresión de la nicotina desmetilasa es silenciada. En otras realizaciones deseables el vector expresa ARN, por ejemplo, ARN antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi).

35 Otro aspecto presenta un vector de expresión que incluye una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica una nicotina desmetilasa, donde el vector es capaz de dirigir la expresión de la nicotina desmetilasa codificada por la molécula aislada de ácido nucleico. Deseablemente, el vector incluye la secuencia de la SEQ ID NO: 4.

40 En un aspecto adicional, la divulgación presenta una planta de tabaco que tiene expresión reducida de una secuencia de ácido nucleico que codifica un polipéptido, por ejemplo, una que incluye la secuencia de la SEQ ID NO: 3, y que desmetila nicotina, donde la expresión reducida (o una reducción en la actividad enzimática) reduce el nivel de normicotina en la planta en una realización deseable, la planta de tabaco es una planta transgénica, tal como una que incluye un transgén que, cuando se expresa en la planta transgénica, silencia la expresión génica de una nicotina desmetilasa endógena de tabaco.

45 En particular, la planta transgénica incluye deseablemente uno o más de lo siguiente: un transgén que expresa una molécula antisentido de una nicotina desmetilasa de tabaco o una molécula de ARN capaz de inducir ARN de interferencia (ARNi); un transgén que, cuando se expresa en la planta transgénica, suprime conjuntamente la expresión de una nicotina desmetilasa del tabaco; un transgén que codifica un producto génico dominante negativo, por ejemplo, una forma mutada de la secuencia de aminoácidos de la SEQ ID NO: 3; una mutación puntual en un gen que codifica la secuencia de aminoácidos de la SEQ ID NO: 3; una supresión en un gen que codifica una nicotina desmetilasa del tabaco; y una inserción en un gen que codifica una nicotina desmetilasa del tabaco.

50 En otras realizaciones deseables, la expresión reducida de una secuencia de ácido nucleico que codifica un polipéptido se produce a nivel transcripcional, a nivel de la traducción, o a nivel posterior a la traducción.

55 Un aspecto adicional presenta un método para reducir la expresión de la nicotina desmetilasa de tabaco en una planta o componente de una planta. Este método implica las etapas de: (a) introducir en células vegetales un transgén que codifica nicotina desmetilasa de tabaco operativamente enlazada a un promotor funcional en las células vegetales para producir células vegetales transformadas; y (b) regenerar una planta o componente de una

planta a partir de la célula vegetal transformada, en donde la nicotina desmetilasa de tabaco es expresada en las células de la planta o componentes de la planta, reduciendo de ese modo la expresión de la nicotina desmetilasa de tabaco en una planta o componente de una planta. En realizaciones particulares de este aspecto, el transgén que codifica la nicotina desmetilasa de tabaco es constitutivamente expresado o expresado en forma inducible, por ejemplo, en una forma específica del tejido, o específica de la célula o específica de un órgano. En otra realización de este aspecto, la expresión del transgén suprime conjuntamente la expresión de nicotina desmetilasa endógena de tabaco o cualquier otro polipéptido descrito en la presente invención.

Un aspecto adicional de la invención presenta otro método para reducir la expresión de la nicotina desmetilasa del tabaco en una planta o componente de la planta tal como se define en las reivindicaciones. Este método incluye las etapas de: (a) introducir en células vegetales un transgén que codifica una molécula de ARN capaz de inducir ARN de interferencia (ARNi) operativamente enlazado a un promotor funcional en las células vegetales para producir células vegetales transformadas; y (b) regenerar una planta o componente de una planta a partir de las células vegetales transformadas, en donde la molécula antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi) de la secuencia codificante de la nicotina desmetilasa de tabaco se expresa en las células de la planta o el componente de la planta, reduciendo de este modo la expresión de la nicotina desmetilasa de tabaco en una planta o componente de una planta. Deseablemente, la molécula de ARN capaz de inducir ARN de interferencia (ARNi) de una nicotina desmetilasa de tabaco se expresa constitutivamente o se expresa de manera inducible, por ejemplo en una forma específica del tejido, específica de la célula o específica de un órgano.

Este método implica reducir el nivel de una nicotina desmetilasa endógena de tabaco, o su enzimática actividad, en la célula vegetal. Deseablemente, la célula vegetal es de una dicotiledónea, una planta solanácea, o una especie de *Nicotiana*. En este aspecto, la reducción del nivel de la nicotina desmetilasa endógena de tabaco implica expresar un transgén que codifica una molécula de ARN bicatenaria de una nicotina desmetilasa del tabaco en la célula vegetal tal como se define en las reivindicaciones. Deseablemente, el ARN bicatenario es una secuencia de ARN correspondiente a la secuencia de la SEQ ID NO: 4, o un fragmento de la misma.

En otras realizaciones deseables, la nicotina desmetilasa endógena de tabaco incluye una mutación puntual en un gen que codifica la secuencia de aminoácidos de la SEQ ID NO: 3. En otras realizaciones deseables la reducción del nivel de expresión de una nicotina desmetilasa endógena de tabaco implica una supresión en un gen que codifica una nicotina desmetilasa de tabaco o implica una inserción en un gen que codifica una nicotina desmetilasa de tabaco. La expresión reducida puede ocurrir a nivel transcripcional, a nivel de la traducción, o a nivel postraduccional.

Un aspecto adicional de la invención presenta un tabaco curado elaborado a partir de una planta de tabaco o componente de la planta como se define en las reivindicaciones, que contiene (i) niveles reducidos de nicotina desmetilasa o (ii) una nicotina desmetilasa que tiene una actividad enzimática alterada y una cantidad reducida de una nitrosamina. Deseablemente, el componente de la planta es una hoja de tabaco o tallo de tabaco. En una realización deseable, la nitrosamina es nornicotina, y el contenido de nornicotina deseablemente es inferior a 5 mg / g, 4,5 mg / g, 4,0 mg / g, 3,5 mg / g, 3,0 mg / g, más deseablemente inferior a 2,5 mg / g, 2,0 mg / g, 1,5 mg / g, 1,0 mg / g, más deseablemente inferior a 750 µg / g, 500 µg / g, 250 µg / g, 100 µg / g, incluso más deseablemente inferior a 75 µg / g, 50 µg / g, 25 µg / g, 10 µg / g, 7,0 µg / g, 5,0 µg / g, 4,0 µg / g, e incluso más deseablemente inferior a 2,0 µg / g, 1,0 µg / g, 0,5 µg / g, 0,4 µg / g, 0,2 µg / g, 0,1 µg / g, 0,05 µg / g, o 0,01 µg / g o en donde el porcentaje de alcaloides secundarios con respecto al contenido total de alcaloides allí es inferior a 90%, 70%, 50%, 30%, 10%, deseablemente inferior a 5%, 4%, 3%, 2%, 1,5%, 1%, y más deseablemente inferior a 0,75%, 0,5%, 0,25%, o 0,1%. En otra realización deseable, la nitrosamina es N'-nitrosornicotina (NNN), y el contenido de N'-NNN deseablemente es inferior a 5 mg / g, 4,5 mg / g, 4,0 mg / g, 3,5 mg / g, 3,0 mg / g, más deseablemente inferior a 2,5 mg / g, 2,0 mg / g, 1,5 mg / g, 1,0 mg / g, más deseablemente inferior a 750 µg / g, 500 µg / g, 250 µg / g, 100 µg / g, incluso más deseablemente inferior a 75 µg / g, 50 µg / g, 25 µg / g, 10 µg / g, 7,0 µg / g, 5,0 µg / g, 4,0 µg / g, e incluso más deseablemente inferior a 2,0 µg / g, 1,0 µg / g, 0,5 µg / g, 0,4 µg / g, 0,2 µg / g, 0,1 µg / g, 0,05 µg / g, o 0,01 µg / g o en donde el porcentaje de alcaloides secundarios con respecto al contenido total de alcaloides contenidos allí es inferior a 90%, 70%, 50%, 30%, 10%, deseablemente inferior a 5%, 4%, 3%, 2%, 1,5%, 1%, y más deseablemente inferior a 0,75%, 0,5%, 0,25%, o 0,1%. En realizaciones adicionales deseables de este aspecto de la invención, la planta de tabaco curada o componente de la planta es un tabaco negro, tabaco Burley, tabaco curado al humo, Virginia, tabaco curado al aire, o tabaco oriental.

Deseablemente, la expresión de un gen endógeno para nicotina desmetilasa, o de cualquier otra secuencia de ácido nucleico descrita en la presente invención, en la planta de tabaco curada o componente de la planta es silenciada.

Otro aspecto de la invención presenta un producto de tabaco que contiene una planta de tabaco curado o componente de la planta que incluye (i) expresión reducida de una nicotina desmetilasa o cualquier otro polipéptido descrito aquí, o (ii) una nicotina desmetilasa u otro polipéptido descrito en la presente invención que tiene actividad alterada, y una cantidad reducida de una nitrosamina. Deseablemente, el producto de tabaco es tabaco sin humo, tabaco húmedo o seco, tabacos de mascar, cigarrillos, cigarrillos, cigarrillos, tabacos de pipa, o bidis. En particular, el

producto de tabaco de este aspecto puede contener tabaco negro, tabaco molido, o incluir un componente aromatizante.

La descripción también incluye un método de fabricación de un producto de tabaco, por ejemplo, un producto de tabaco sin humo, que contiene (i) expresión reducida de una nicotina desmetilasa o (ii) una nicotina desmetilasa que tiene actividad enzimática alterada (por ejemplo, reducida), y una cantidad reducida de una nitrosamina. Este método implica proporcionar una planta de tabaco curado o componente de la planta que contiene (i) un nivel reducido de nicotina desmetilasa o (ii) una nicotina desmetilasa que tiene una actividad enzimática alterada y una cantidad reducida de una nitrosamina y preparar el producto de tabaco a partir de la planta de tabaco curado o componente de la planta.

10 Definiciones

“Actividad enzimática” se entiende que incluye pero no se limita a desmetilación, hidroxilación, epoxidación, N-oxidación, sulfooxidación, N, S, y O-desalquilaciones, desulfatación, desaminación y reducción de azo, nitro, N-óxido, y otros de tales grupos químicos enzimáticamente reactivos. Actividad enzimática alterada se refiere a una disminución de la actividad enzimática (por ejemplo, de una nicotina desmetilasa de tabaco) en al menos un 10-20%, preferiblemente en al menos 25-50%, y más preferiblemente en al menos 55-95% o mayor en relación con la actividad de una enzima de control (por ejemplo, nicotina desmetilasa de tabaco de una planta de tabaco de tipo silvestre). La actividad de una enzima, tal como una nicotina desmetilasa puede analizarse utilizando métodos convencionales en la técnica, por ejemplo, usando los análisis de microsomas de levadura descritos en este documento.

El término “ácido nucleico” se refiere a un polímero de desoxirribonucleótido o ribonucleótido ya sea en forma monocatenaria o bicatenaria, o sentido o antisentido, y a menos que se limite de otra forma, abarca análogos conocidos de nucleótidos naturales que hibridan con ácidos nucleicos de una manera similar a los nucleótidos de origen natural. A menos que se indique lo contrario, una secuencia particular de ácido nucleico incluye la secuencia complementaria de la misma. Los términos “operativamente enlazado”, “en combinación operable,” y “en orden operable” se refiere al enlace funcional entre una secuencia de control de expresión del ácido nucleico (tal como un promotor, secuencia señal, o matriz de sitios de enlazamiento del factor de transcripción) y una segunda secuencia de ácido nucleico, donde la secuencia de control de la expresión afecta la transcripción y / o traducción del ácido nucleico correspondiente a la segunda secuencia. Deseablemente, una secuencia de ácido nucleico operativamente enlazada se refiere a un fragmento de un gen que está enlazado a otras secuencias del mismo gen para formar un gen de longitud completa.

El término “recombinante” cuando se usa con referencia a una célula indica que la célula replica un ácido nucleico heterólogo, expresa el ácido nucleico o expresa un péptido, péptido heterólogo, o proteína codificada por un ácido nucleico heterólogo. Las células recombinantes pueden expresar genes o fragmentos de genes, ya sea en forma sentido o antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi) que no se encuentran dentro de la forma nativa (no recombinante) de la célula. Las células recombinantes también pueden expresar genes que se encuentran en la forma nativa de la célula, pero en donde los genes se modifican y se reintroducen en la célula por medios artificiales.

Un “gen estructural” es aquella porción de un gen que comprende un segmento de ADN que codifica una proteína, polipéptido o una porción de los mismos, y excluyendo, por ejemplo, la secuencia 5' que dirige el inicio de la transcripción o la 3'UTR. El gen estructural puede codificar alternativamente un producto no traducible. El gen estructural puede ser uno que se encuentra normalmente en la célula o uno que no se encuentra normalmente en la célula o localización celular en la que se introduce, en cuyo caso se denomina un “gen heterólogo”. Un gen heterólogo puede derivarse en su totalidad o en parte de cualquier fuente conocida en la técnica, incluyendo un genoma bacteriano o ADN de episoma, eucariota, nuclear o de plásmido, ADNc, ADN viral o ADN sintetizado químicamente. Un gen estructural puede contener una o más modificaciones que podrían afectar la actividad biológica o sus características, la actividad biológica o la estructura química del producto de expresión, la tasa de expresión o la forma de control de la expresión. Tales modificaciones incluyen, pero no se limitan a, mutaciones, inserciones, supresiones y sustituciones de uno o más nucleótidos.

El gen estructural puede constituir una secuencia de codificación ininterrumpida o puede incluir uno o más intrones, delimitados por las uniones de empalme apropiadas. El gen estructural puede ser traducible o no traducible, incluyendo una molécula antisentido o de ARN capaz de inducir ARN de interferencia (ARNi). El gen estructural puede ser un compuesto de segmentos derivados de una pluralidad de fuentes y de una pluralidad de secuencias génicas (de origen natural o sintético, donde sintético se refiere a ADN que se sintetiza químicamente).

Un “exón” como se usa en la presente invención en referencia a una secuencia de ácido nucleico se entiende como una porción de la secuencia de ácido nucleico de un gen, en donde la secuencia de ácido nucleico del exón codifica al menos un aminoácido del producto del gen. Un exón es típicamente adyacente a un segmento de ADN no

codificante, tal como un intrón.

Un "intrón" como se usa en la presente invención en referencia a una secuencia de ácido nucleico se entiende como una región no codificante de un gen que está flanqueado por las regiones de codificación. Un intrón es típicamente una región no codificante de un gen que se transcribe en una molécula de ARN pero luego es escindida por empalme de ARN durante la producción del ARN mensajero u otro ARN estructural funcional.

Un "3'UTR" como se usa en la presente invención en referencia a una secuencia de ácido nucleico se entiende como una secuencia de ácido nucleico no codificante proximal a un codón de parada de un exón.

"Derivado de" se usa para significar tomado, obtenido, recibido, trazado, replicado o descendiente de una fuente (química y / o biológica). Un derivado puede ser producido por manipulación química o biológica (incluyendo, pero no limitado a, sustitución, adición, inserción, supresión, extracción, aislamiento, mutación y replicación) de la fuente original.

"Químicamente sintetizado", en relación con una secuencia de ADN, significa que se ensamblaron porciones de los nucleótidos componentes *in vitro*. La síntesis química manual de ADN puede realizarse utilizando procedimientos bien establecidos (Caruthers, Methodology of DNA and RNA Sequencing, (1983), Weissman (ed), Praeger Publishers, Nueva York, Capítulo 1.); la síntesis química automatizada puede llevarse a cabo utilizando una de una serie de máquinas comercialmente disponibles.

La alineación óptima de secuencias para comparación puede realizarse, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, Adv. Appl. Math. 2: 482 (1981), mediante el algoritmo de alineación de homología de Needleman y Wunsch, J. Mol. Biol. 48: 443 (1970), mediante la búsqueda del método de similitud de Pearson y Lipman Proc. Natl. Acad. Sci. (EE.UU.) 85: 2444 (1988), mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o mediante inspección.

La herramienta de búsqueda de alineación local básica del Centro Nacional para Información Biológica (BLAST) (Altschul et al., 1990) está disponible a partir de varias fuentes, incluyendo el Centro Nacional de Información Biológica (NCBI, Bethesda, Md.) y en la Internet, para uso en conexión con los programas de análisis de secuencias blastp, blastn, blastx, tblastn y tblastx. Se puede acceder a través de <http://www.ncbi.nlm.nih.gov/BLAST/>. Una descripción de cómo determinar la identidad de secuencia usando este programa está disponible en http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.

Los términos "identidad sustancial de aminoácidos" o "identidad sustancial de secuencias de aminoácidos" como se aplica a las secuencias de aminoácidos y como se usa aquí denotan una característica de un polipéptido, en donde el péptido comprende una secuencia que tiene al menos 70 por ciento de identidad de secuencia, preferiblemente 80 por ciento de identidad de secuencia de aminoácidos, más preferiblemente 90 por ciento de identidad de secuencia de aminoácidos, y lo más preferible al menos 99 a 100 por ciento de identidad de secuencia en comparación con la secuencia de la proteína mostrada en las figuras 10 - 17. Deseablemente, para una nicotina desmetilasa, la comparación de secuencias se compara deseablemente para una región después del motivo del citocromo p450 GXRXCX(G/A) (SEQ ID NO: 2265) para el codón de parada del péptido traducido.

Los términos "identidad sustancial de ácidos nucleicos" o "identidad sustancial de secuencia de ácidos nucleicos" tal como se aplica a secuencias de ácidos nucleicos y como se utiliza aquí denota una característica de una secuencia de polinucleótidos, en donde el polinucleótido comprende una secuencia que tiene al menos 50 por ciento, preferiblemente 60, 65, 70, o 75 por ciento de identidad de secuencia, más preferentemente 81 o 91 por ciento de identidad de secuencia de ácidos nucleicos, y más preferiblemente al menos 95, 99 o incluso 100 por ciento de identidad de secuencia en comparación con un grupo de referencia sobre la región correspondiente al primer ácido nucleico después de la región que codifica al motivo del citocromo p450 GXRXCX (G/A) (SEQ ID NO: 2265) para el codón de parada del péptido traducido.

Otra indicación de que las secuencias de nucleótidos son sustancialmente idénticas es si dos moléculas hibridan entre sí bajo condiciones rigurosas. Las condiciones rigurosas dependen de la secuencia y serán diferentes en diferentes circunstancias. Generalmente, se seleccionan las condiciones rigurosas para que sean aproximadamente de 5 °C hasta aproximadamente 20 °C, usualmente de aproximadamente 10 °C hasta aproximadamente 15 °C, inferiores al punto de fusión térmico (Tm) para la secuencia específica a una fuerza iónica y pH definidos. El Tm es la temperatura (bajo una fuerza iónica y pH definidos) a la que 50% de la secuencia objetivo hibrida con una sonda emparejada. Típicamente, condiciones rigurosas serán aquellas en las que la concentración de sal es aproximadamente 0,02 molar a pH 7 y la temperatura es al menos aproximadamente 60 °C. Por ejemplo, en un procedimiento de hibridación estándar tipo Southern, las condiciones rigurosas incluirán un lavado inicial en 6 x SSC a 42 °C seguido por uno o más lavados en 0,2 x SSC a una temperatura de al menos aproximadamente 55 °C, típicamente de aproximadamente 60 °C, y a menudo aproximadamente 65 °C.

Las secuencias de nucleótidos también son sustancialmente idénticas para los propósitos de esta invención cuando dichas secuencias de nucleótidos codifican polipéptidos y/o proteínas que son sustancialmente idénticos. Así pues, cuando una secuencia de ácido nucleico codifica esencialmente el mismo polipéptido como una segunda secuencia de ácido nucleico, las dos secuencias de ácido nucleico son sustancialmente idénticas, incluso si no se hibridan bajo condiciones rigurosas debido a la degeneración permitida por el código genético (véase, Darnell et al. (1990) Molecular Cell Biology, segunda edición Scientific American Books W.H. Freeman and Company Nueva York para una explicación de la degeneración del codón y el código genético). La pureza u homogeneidad de la proteína pueden ser indicadas por una cantidad de medios bien conocidos en la técnica, tales como electroforesis en gel de poliacrilamida de una muestra de proteína, seguido por la visualización después de la coloración. Para ciertos fines puede ser necesaria alta resolución y se pueden utilizar HPLC o un medio similar para la purificación.

Por un anticuerpo que “enlaza específicamente” o “específicamente reconoce” un polipéptido particular, tal como una nicotina desmetilasa de tabaco, se entiende un aumento de la afinidad del anticuerpo por el polipéptido en relación con una cantidad igual de cualquier otra proteína. Anticuerpos deseables son anticuerpos que se enlazan específicamente a un polipéptido que tiene una secuencia de aminoácidos mostrada en la Figura 1, las Figuras 3 y 4, o las Figuras 10 a 17. Por ejemplo, un anticuerpo que se enlaza específicamente a una nicotina desmetilasa de tabaco que contiene la secuencia de aminoácidos de la SEQ ID NO: 3 tiene deseablemente una afinidad por su antígeno que es al menos 2 veces, 5 veces, 10 veces, 30 veces, o 100 veces mayor que para una cantidad igual de cualquier otro antígeno, incluyendo antígenos relacionados. El enlazamiento de un anticuerpo a un antígeno, por ejemplo, una nicotina desmetilasa de tabaco, puede ser determinado por cualquier cantidad de métodos convencionales en la técnica, por ejemplo, análisis tipo Western, ELISA, o co-inmunoprecipitación. Los anticuerpos que enlazan específicamente un polipéptido, por ejemplo, una nicotina desmetilasa, también son útiles para purificar el polipéptido.

Tal como se utiliza aquí, el término “vector” se usa en referencia con moléculas de ácido nucleico que transfieren un segmento(s) de ADN en una célula. Un vector puede actuar para replicar el ADN y puede reproducirse independientemente en una célula huésped. El término “vehículo” a veces se utiliza como sinónimo de “vector”. El término “vector de expresión” tal como se utiliza aquí se refiere a una molécula de ADN recombinante que contiene una secuencia de codificación deseada y secuencias apropiadas de ácido nucleico necesarias para la expresión de la secuencia de codificación operativamente enlazada en un organismo huésped particular. Las secuencias necesarias de ácido nucleico para la expresión en procariotas incluyen generalmente un promotor, un operador (opcional) y un sitio de enlazamiento al ribosoma, a menudo junto con otras secuencias. Deseablemente, el promotor incluye la secuencia de la SEQ ID NO: 8, o un fragmento de la misma que dirige la transcripción. También son deseables secuencias promotoras que tienen al menos 50%, 60%, 75%, 80%, 90%, 95%, o incluso 99% de identidad de secuencia con la secuencia de la SEQ ID NO: 8 y que dirige la transcripción. Se sabe que las células eucariotas utilizan promotores, potenciadores y señales de terminación y de poliadenilación, tal como la secuencia 3'UTR de la SEQ ID NO: 9. En algunos casos, se ha observado que los vectores de expresión de una planta requieren la presencia de intrones derivados de la planta, tales como el intrón que tiene la secuencia de la SEQ ID NO: 7, para tener una expresión estable. Como tal, la secuencia de la SEQ ID NO: 7, o cualquier otro intrón que tenga una unión apropiada de empalme de ARN se pueden usar como se describe adicionalmente en la presente invención. Vectores deseables incluyen secuencias de ácido nucleico mostradas en las Figuras 1, 3 a 7, o 10 a 17.

Para el propósito de la regeneración de plantas completas modificadas genéticamente con raíces, se puede insertar un ácido nucleico en células vegetales, por ejemplo, por cualquier técnica, tal como la inoculación *in vivo* o por cualquiera de las técnicas conocidas de cultivo de tejido *in vitro* para producir células vegetales transformadas que puedan ser regeneradas en plantas completas. Así, por ejemplo, la inserción en células vegetales puede ser por inoculación *in vitro* mediante *A. tumefaciens* patógeno o no patógeno. También se pueden emplear otras técnicas de cultivo de tejidos.

“Tejido de la planta”, “componente de la planta” o “célula vegetal” incluye tejidos de plantas diferenciados y no diferenciados, incluyendo, pero no limitado a, raíces, brotes, hojas, polen, semillas, tejido tumoral y diversas formas de células en cultivo, tales como células individuales, protoplastos, embriones y tejido de callo. El tejido de la planta puede ser *in planta* o en un órgano, tejido o cultivo celular.

“Célula vegetal” como se usa en este documento incluye células vegetales *in planta* y células vegetales y protoplastos en cultivo. “ADNc” o “ADN complementario” generalmente se refiere a una molécula de ADN monocatenario con una secuencia de nucleótidos que es complementaria a una molécula de ARN no procesado que contiene un intrón, o un ARNm procesado que carecen de intrones. El ADNc se forma por la acción de la enzima transcriptasa inversa en una plantilla de ARN.

“Tabaco” como se usa en este documento incluye tabaco curado al humo, Virginia, Burley, oscuro, Oriental, y otros tipos de plantas dentro del género *Nicotiana*. La semilla del género *Nicotiana* está fácilmente disponible comercialmente en forma de *Nicotiana tabacum*.

Los "artículos de fabricación" o "productos del tabaco" incluyen productos tales como tabaco húmedo y seco, tabacos de mascar, cigarrillos, cigarras, cigarritos, tabacos de pipa, bidis, y los productos similares derivados del tabaco.

5 Por "silenciamiento génico" se entiende una disminución en el nivel de expresión génica (por ejemplo, la expresión de un gen que codifica una nicotina desmetilasa de tabaco) en al menos un 30-50%, preferiblemente en al menos el 50-80%, y más preferiblemente en al menos 80-95% o más con respecto al nivel en una planta de control (por ejemplo, una planta de tabaco de tipo silvestre). La reducción de dichos niveles de expresión puede llevarse a cabo mediante el empleo de métodos convencionales que son conocidos en la técnica, incluyendo, sin limitación, ARN de interferencia, triple hebra de interferencia, ribozimas, recombinación homóloga, silenciamiento génico inducido por virus, tecnologías antisentido y de cosupresión, expresión de un producto del gen negativo dominante, o a través de la generación de genes mutados usando técnicas de mutagénesis estándar, tales como las descritas en la presente invención. Los niveles de un polipéptido o transcripto de nicotina desmetilasa de tabaco, o ambos, se controlan de acuerdo con cualquier técnica estándar, incluyendo, pero sin limitarse a, transferencia tipo Northern, protección de RNasa, o inmunotransferencia.

15 Por un "fragmento" o "porción" de una secuencia de aminoácidos se entiende al menos por ejemplo, 20, 15, 30, 50, 75, 100, 250, 300, 400, o 500 aminoácidos contiguos de cualquiera de las secuencias de aminoácidos mostradas en las Figuras 1, 3, 4, y 10 - 17. Los ejemplos de fragmentos deseables son los aminoácidos 1 a 313 de la secuencia de la SEQ ID NO: 3 y los aminoácidos 314 a 517 de la secuencia de la SEQ ID NO: 3, así como la secuencia de las SEQ ID NOS: 2 y 63. Además, con respecto a un fragmento o porción de una secuencia de ácido nucleico, los fragmentos deseables incluyen al menos 100, 250, 500, 750, 1000, o 1500 ácidos nucleicos contiguos de cualquiera de las secuencias de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17. Los ejemplos de fragmentos deseables son los ácidos nucleicos 1 - 2009, 2010 - 2949, 2950 - 3946, 3947 - 4562, 4563 - 6347, y 4731 - 6347 de la secuencia de la SEQ ID NO: 4.

25 Por un "polipéptido sustancialmente puro" se entiende un polipéptido que ha sido separado de la mayoría de componentes que lo acompañan de forma natural; sin embargo, también se considera que otras proteínas que se encuentran en la fracción microsomal asociadas con una preparación que tiene una actividad enzimática de al menos 8,3 pKat / mg de proteína que son un polipéptido sustancialmente puro. Típicamente, el polipéptido es sustancialmente puro cuando está al menos 60%, en peso, libre de las proteínas y moléculas orgánicas de origen natural con las que se asocia naturalmente. Preferiblemente, la preparación es al menos 75%, más preferiblemente al menos 90%, y lo más preferible al menos 99%, en peso, el polipéptido deseable. Un polipéptido sustancialmente puro puede ser obtenido, por ejemplo, por extracción a partir de una fuente natural (por ejemplo, una célula de una planta de tabaco); por expresión de un ácido nucleico recombinante que codifica al polipéptido; o por síntesis química de la proteína. La pureza se puede medir por cualquier método apropiado, por ejemplo, cromatografía en columna, electroforesis en gel de poliacrilamida, o mediante análisis por HPLC.

35 Por "molécula aislada de ácido nucleico" se entiende una secuencia de ácido nucleico libre de las secuencias de ácido nucleico que flanquean de forma natural la secuencia de la molécula de ácido nucleico en el genoma de un organismo.

40 Por "célula transformada" se entiende una célula en la que (o en un antepasado de la misma) se ha introducido, por medio de técnicas de ADN recombinante, una molécula de ADN, por ejemplo, una molécula de ADN que codifica una nicotina desmetilasa de tabaco o cualquiera de las secuencias de ácido nucleico descritas en este documento (por ejemplo, las secuencias de ácido nucleico mostradas en las Figuras 1, 3 a 7, y 10 a 17).

45 Por un "nicotina desmetilasa de tabaco" o "nicotina desmetilasa" como se usa aquí, se entiende un polipéptido que es sustancialmente idéntico a la secuencia de la SEQ ID NO: 3. De manera deseable, una nicotina desmetilasa de tabaco es capaz de convertir nicotina ($C_{10}H_{14}N_2$, también conocida como 3-(1-metil-2-pirrolidinil)piridina) en nornicotina ($C_9H_{12}N_2$). La actividad de una nicotina desmetilasa de tabaco puede ser analizada mediante métodos convencionales en la técnica, tales como midiendo la desmetilación de nicotina radioactiva mediante microsomas expresados en levadura, tal como se describe en la presente invención.

Según lo previsto en la presente invención, los términos "citocromo p450" y "p450" se utilizan indistintamente.

50 Otras características y ventajas de la invención serán evidentes a partir de la siguiente descripción detallada, los dibujos, y las reivindicaciones.

Breve descripción de los dibujos

La Figura 1 es la secuencia de ácido nucleico de D35-BG11 (SEQ ID NO: 1) y su producto de traducción (SEQ ID NO: 2).

La Figura 2 es un diagrama esquemático de la estructura genómica del gen para la nicotina desmetilasa de tabaco.

La Figura 3 es la secuencia genómica de ácido nucleico para la nicotina desmetilasa de tabaco (SEQ ID NO: 4) y su producto de traducción (SEQ ID NO: 3)

5 La Figura 4 es la secuencia de ácido nucleico de la región codificante del gen para nicotina desmetilasa de tabaco (SEQ ID NO: 5), y su producto de traducción (SEQ ID NO: 6).

La Figura 5 es la secuencia de ácido nucleico de un intrón (SEQ ID NO: 7) presente en la secuencia genómica para la nicotina desmetilasa de tabaco.

La Figura 6 es la secuencia de ácido nucleico del promotor del gen para desmetilasa de tabaco (SEQ ID NO: 8).

10 La Figura 7 es la secuencia de ácido nucleico de la 3'UTR del gen para nicotina desmetilasa de tabaco (SEQ ID NO: 9).

Figura 8 es una imagen de electroforesis que muestra productos de PCR de líneas de tabaco con el conjunto de cebadores de longitud completa ("FL") Geno.

15 La Figura 9 es una imagen de electroforesis que muestra productos de PCR de líneas de tabaco con conjuntos de cebadores (1), (2), (3) y (4) como se indica en el Ejemplo 17. Los tamaños aproximados de las bandas son de 3.500 nucleótidos (nt) para la FL, 2600 nt para (1), 1400 nt para (2), 600 nt para (3) y 1400 nt para (4).

La Figura 10 es la secuencia de ácido nucleico de D35-BG11 (SEQ ID NO: 62), y la secuencia de aminoácidos de D35-BG1 (SEQ ID NO: 63), que es la traducción de D35-BG11 (SEQ ID NO: 62).

La Figura 11 es la secuencia de ácido nucleico de D120-AH4 (SEQ ID NO: 188), y la secuencia de aminoácidos de D120-AH4 (SEQ ID NO: 189), que es la traducción de D120-AH4 (SEQ ID NO: 188).

20 La Figura 12 es la secuencia de ácido nucleico de D121-AA8 (SEQ ID NO: 190), y la secuencia de aminoácidos de D121-AA8 (SEQ ID NO: 191), que es la traducción de D121-AA8 (SEQ ID NO: 190).

La Figura 13 es la secuencia de ácido nucleico de D122-AF10 (SEQ ID NO: 192), y la secuencia de aminoácidos de D122-AF10 (SEQ ID NO: 193), que es la traducción de D122-AF10 (SEQ ID NO: 192).

25 La Figura 14 es la secuencia de ácido nucleico de D208-AC8 (SEQ ID NO: 226), y la secuencia de aminoácidos de D208-AC8 (SEQ ID NO: 227), que es la traducción de D208-AC8 (SEQ ID NO: 226).

La Figura 15 es la secuencia de ácido nucleico de D103-AH3 (SEQ ID NO: 230), y la secuencia de aminoácidos de D103-AH3 (SEQ ID NO: 231), que es la traducción de D103-AH3 (SEQ ID NO: 230).

La Figura 16 es la secuencia de ácido nucleico de D208-AD9 (SEQ ID NO: 232), y la secuencia de aminoácidos de D208-AD9 (SEQ ID NO: 233), que es la traducción de D208-AD9 (SEQ ID NO: 232).

30 La Figura 17 es la secuencia de ácido nucleico de D235-AB1 (SEQ ID NO: 254), y la secuencia de aminoácidos de D235-AB1 (SEQ ID NO: 255), que es la traducción de D235-AB1 (SEQ ID NO: 254).

35 La Figura 18 es un conjunto de alineaciones de secuencias de aminoácidos que consisten de la alineación de secuencia de D208-AD9 (SEQ ID NO: 2196), D120-AH4 (SEQ ID NO: 2197), D121-AA8 (SEQ ID NO: 2198), D122 - AF10 (SEQ ID NO: 2199), D103-AH3 (SEQ ID NO: 2200), D208-AC8 (SEQ ID NO: 2201), y D235-AB1 (SEQ ID NO: 2202); la alineación de secuencia de D244-AD4 (SEQ ID NO: 2203), D244-AB6 (SEQ ID NO: 2204), D285-AA8 (SEQ ID NO: 2205), D285-AB9 (SEQ ID NO: 2206), y D268-AE2 (SEQ ID NO: 2207); y la alineación de secuencia de D-100A AC3 (SEQ ID NO: 2208) y D100A-BE2 (SEQ ID NO: 2209).

40 La Figura 19 es un diagrama que muestra la clonación de fragmentos de ADNc del citocromo p450 por PCR. Los cebadores utilizados para la clonación están listados: DM (SEQ ID NO: 2255), DM4 (SEQ ID NO: 2256), DM12 (SEQ ID NO: 2257), DM13 (SEQ ID NO: 2258), DM17 (SEQ ID NO: 2259), OLIGO d(T) (SEQ ID NO: 2260), T7 (SEQ ID NO: 2261), y SP6 (SEQ ID NO: 2262).

La Figura 20 es una tabla que enumera las secuencias del conjunto de sondas (SEQ ID NO: 385 - (SEQ ID NO: 445) de clones en el microarreglo GeneChip®.

Descripción detallada

Tradicionalmente, numerosas etapas estaban involucradas en el desarrollo de cualquier germoplasma vegetal novedoso deseable. El fitomejoramiento de las plantas comienza con el análisis y definición de los problemas y debilidades del germoplasma presente, el establecimiento de los objetivos del programa, y la definición de objetivos específicos de fitomejoramiento. La siguiente etapa es la selección del germoplasma que posee los rasgos para cumplir los objetivos del programa. El objetivo es combinar en una sola variedad una combinación mejorada de rasgos deseables del germoplasma progenitor. Los rasgos deseables incluyen, por ejemplo, un mayor rendimiento de semilla, resistencia a enfermedades e insectos, tolerancia a la sequía y el calor, y mejores cualidades agronómicas. Sin embargo, estos procesos, que conducen a la etapa final de comercialización y distribución, pueden tomar de seis a doce años desde el momento en que se hace el primer cruzamiento. Por lo tanto, el desarrollo de nuevas variedades es un proceso que consume tiempo y requiere una planificación anticipada precisa, el uso eficiente de los recursos, y un mínimo de cambios de dirección.

La mejora de las variedades vegetales a través de la transformación genética se ha vuelto cada vez más importante para el fitomejoramiento moderno. Los genes de interés comercial potencial, tales como los genes que confieren rasgos específicos deseados en la planta de resistencia a enfermedades, resistencia a los insectos, o mejora de la calidad, se pueden incorporar en especies de cultivo a través de diversas tecnologías de transferencia de genes. La capacidad para manipular la expresión génica proporciona un medio de producir nuevas características en las plantas transformadas. En algunas situaciones pueden ser deseables niveles altos o mayores de expresión génica. Por ejemplo, es deseable aumentar la producción de una proteína que por sí misma maximiza la resistencia a las enfermedades, el rendimiento, sabor, o cualquier otro atributo comercialmente deseable de una planta. Del mismo modo, la regulación de la expresión de genes endógenos, por ejemplo, mediante el silenciamiento de genes puede dar lugar a plantas o productos vegetales más valiosos.

Durante la maduración o curado del tabaco, la activación, sobrerregulación, o subregulación de cualquiera de los genes identificados como inducidos por etileno o relacionados con la senescencia (por ejemplo, aquellos que tienen la secuencia de las SEQ ID NOS: 4, 188, y 226) pueden afectar aquellas rutas metabólicas implicadas en la formación de numerosos metabolitos secundarios, incluyendo terpenoides, polifenoles, alcaloides, etc. que afectan los rasgos de calidad del producto final (por ejemplo, resistencia a enfermedades, resistencia a insectos, la mejora de la calidad, aroma modificado, sabor modificado, y similares). Afectados del mismo modo por los genes identificados en la presente invención pueden ser las rutas metabólicas asociadas con el rasgo y el tipo de materia seca acumulada durante la senescencia o la repartición de la materia seca dentro de la planta durante la senescencia. Los cambios en la velocidad y el tipo de acumulación de almidón, la formación de lignina, la deposición de celulosa, y la translocación de azúcar se pueden demostrar. El control de los genes identificados en la presente invención también puede afectar a aquellas rutas metabólicas implicadas en la determinación de las tasas de senescencia, la uniformidad de la senescencia dentro de una hoja y entre las hojas de una sola planta, y la inducción de la senescencia por medios artificiales o naturales. Los agentes que inducen la senescencia o las actividades que estimulan o activan los genes identificados en la presente invención incluyen, por ejemplo compuestos químicos tales como peróxidos diluidos, pesticidas, herbicidas, reguladores del crecimiento, tratamientos térmicos, lesiones, o gases como el ozono y concentraciones elevadas de dióxido de carbono.

Identificación de tabaco constitutivamente expresado, o secuencias inducidas por etileno o por senescencia

Se extrajo el ARN a partir de tejido de *Nicotiana* de líneas de *Nicotiana* convertidoras y no convertidoras. El ARN extraído fue luego utilizado para crear ADNc. Se generaron luego secuencias de ácido nucleico, utilizando dos estrategias.

En la primera estrategia, se extrajo el ARN enriquecido en poli A de tejido de la planta y se elaboró ADNc por PCR de transcripción inversa. A continuación, se utilizó ADNc monocatenario para crear poblaciones específicas de p450 por PCR usando cebadores degenerados más un cebador inverso oligo d(T). El diseño del cebador se basó en los motivos altamente conservados de otras secuencias de genes para el citocromo p450 de la planta. Ejemplos de cebadores degenerados específicos se exponen en la Figura 1 de las publicaciones de las solicitudes de patente de los Estados Unidos Nos. 2004/0103449 A1, 2004/0111759 A1 y 2004/0117869 A1. Se analizó adicionalmente la secuencia de fragmentos de plásmidos que contenían insertos de tamaño apropiado. El tamaño de estos insertos iba típicamente desde aproximadamente 300 hasta aproximadamente 800 nucleótidos, dependiendo de qué cebadores se utilizaron.

En una segunda estrategia, se construyó inicialmente una biblioteca de ADNc. El ADNc en los plásmidos se utilizó para crear poblaciones específicas de p450 por PCR usando cebadores degenerados más el cebador T7 en el plásmido como cebador inverso. Como en la primera estrategia, se analizó adicionalmente la secuencia de fragmentos de plásmidos que contenían insertos de tamaño apropiado.

Se pueden usar líneas de plantas de *Nicotiana* que se sabe que producen altos niveles de nornicotina (convertidor) y

líneas de plantas que tienen bajos niveles de norcicotina como materiales de partida. Las hojas pueden ser removidas luego de las plantas y tratadas con etileno para activar las actividades enzimáticas de p450 definidas en la presente invención. El ARN total se extrae usando técnicas conocidas en la técnica. Se pueden generar luego fragmentos de ADNc usando PCR (RT-PCR) con el cebador oligo d(T) (SEQ ID NO: 2260) como se describe en la Figura 19. La biblioteca de ADNc puede luego ser construida como se describe más detalladamente en los ejemplos en la presente invención.

La región conservada de las enzimas tipo p450 se utilizó como plantilla para cebadores degenerados, ejemplos de los cuales se muestran en la Figura 19. Usando cebadores degenerados, se amplificaron bandas específicas de p450 mediante PCR. Se identificaron bandas indicativas para enzimas tipo p450 mediante secuenciación del ADN. Se caracterizaron fragmentos de PCR mediante la búsqueda con BLAST, alineación u otras herramientas para identificar a los candidatos adecuados.

La información de la secuencia de fragmentos identificados fue utilizada para desarrollar cebadores para la PCR. Estos cebadores en combinación con cebadores de plásmidos en la biblioteca de ADNc se utilizaron para clonar genes para p450 de longitud completa. Se llevó a cabo un análisis tipo Southern inverso a gran escala para examinar la expresión diferencial de todos los clones de fragmentos obtenidos y, en algunos casos, los clones de longitud completa. Estos ensayos tipo Southern inverso a gran escala pueden llevarse a cabo utilizando ADNc totales marcados de diferentes tejidos como sonda para hibridarse con fragmentos de ADN clonados con el fin de seleccionar todos los insertos clonados. Se usaron también ensayos de transferencia tipo Northern radioactivos (P^{32}) y no radiactivos para caracterizar fragmentos de p450 clonados y clones de longitud completa.

Una vez que se obtienen células vegetales que expresan el nivel deseado de enzima p450, se pueden regenerar tejidos vegetales y plantas enteras a partir de las mismas usando métodos y técnicas bien conocidas en la técnica. Las plantas regeneradas se reproducen luego por medios convencionales y los genes introducidos pueden transferirse a otras cepas y variedades cultivadas mediante técnicas convencionales de reproducción de plantas.

Los genes relacionados con la senescencia o inducidos por etileno, por ejemplo, aquellos identificados en las SEQ ID NOS: 4, 188, y 226 puede codificar enzimas que son importantes determinantes de los parámetros de calidad de la hoja de tabaco para una variedad de productos de tabaco. Los productos de tabaco incluyen tabaco húmedo o seco, tabacos de mascar, cigarrillos, cigarrillos, cigarrillos, tabacos para pipa, bidis y productos para fumar similares. Los parámetros de calidad de la hoja pueden incluir: atributos visuales como el color, uniformidad de la superficie, textura, o abigarramiento; características estructurales o físicas como las ejemplificadas por la relación lámina a tallo, oleaginosidad, potencial de llenado de cigarrillos, densidad aparente, retención de humedad, y flexibilidad; rasgos químicos o bioquímicos relacionados con el sabor, el aroma, la capacidad de fermentación, tasas de quemado, temperaturas de quemado, absorción y liberación del sabor artificial; y generación de los componentes del humo incluyendo alquitrán o material en partículas, alcaloides y otros atributos similares. Las reacciones enzimáticas resultantes de estos genes inducidos por el etileno o relacionados con senescencia también pueden producir metabolitos secundarios que influyen sobre las interacciones con patógenos o insectos que afectan el rendimiento y la calidad de la hoja de tabaco. Por ejemplo, Wagner, et al. (Nature Biotechnology, 19: 371 - 374, 2001) demostró que la supresión de un gen para la hidroxilasa p450 aumenta en gran medida la acumulación de cembratien-ol, un metabolito secundario que influye sobre la resistencia a los ácidos.

Generación de anticuerpos

Los anticuerpos específicos de péptidos se elaboraron mediante la derivación de su secuencia de aminoácidos y seleccionando regiones peptídicas que eran antigénicas y únicas con respecto a otros clones. Se elaboraron anticuerpos de conejo con péptidos sintéticos conjugados con una proteína portadora. Se realizaron análisis de transferencia tipo Western u otros métodos inmunológicos en tejido vegetal utilizando estos anticuerpos. Además, se elaboraron anticuerpos específicos de péptidos para varios clones de longitud completa mediante la derivación de su secuencia de aminoácidos y seleccionando regiones peptídicas que eran potencialmente antigénicas y eran únicas con respecto a otros clones. Se elaboraron anticuerpos de conejo con péptidos sintéticos conjugados con una proteína portadora. Los análisis de transferencia tipo Western se realizaron utilizando estos anticuerpos.

Subregulación de la expresión génica y alteración de la actividad enzimática

Las plantas que tienen menor expresión de un polipéptido se generan de acuerdo a los métodos convencionales de silenciamiento génico. (Para una revisión, véase Arndt y Rank, Genome 40: 785 - 797, 1997; Turner y Schuch, Journal of Chemical Technology and Biotechnology 75: 869 - 882, 2000; y Klink y Wolniak, Journal of Plant Growth Regulation 19 (4): 371 - 384, 2000). En particular, se pueden usar secuencias de ácido nucleico para la nicotina desmetilasa de tabaco (por ejemplo, las SEQ ID NOS: 4, 5, 7, 8, y 9, o fragmentos de las mismas, tales como la secuencia de las SEQ ID NOS: 1 y 62), así como secuencias de ácido nucleico sustancialmente idénticas (por ejemplo, la secuencia de la SEQ ID NO: 188) para alterar fenotipos de tabaco o metabolitos de tabaco, por ejemplo, norcicotina en cualquier especie de *Nicotiana*. La disminución de la expresión de un gen para nicotina desmetilasa

de tabaco puede conseguirse usando, por ejemplo, el ARN de interferencia (ARNi) (Smith y otros, Nature 407: 319 - 320, 2000; Fire et al., Nature 391: 306 - 311, 1998; Waterhouse et al., PNAS 95: 13959 - 13964, 1998; Stalberg et al., Plant Molecular Biology 23: 671 - 683, 1993; Brignetti et al., EMBO J. 17: 6739 - 6746, 1998; Allen et al., Nature Biotechnology 22: 1559 - 1566, 2004); silenciamiento génico inducido por virus ("VIGS") (Baulcombe, Current Opinions in Plant Biology, 2: 109 - 113, 1999; Cogoni y Macino, Genes Dev 10: 638 - 643, 2000; Ngelbrecht et al., PNAS 91: 10502 -10506, 1994); silenciamiento del gen objetivo mediante la transferencia de un gen endógeno de la planta en la orientación sentido (Jorgensen et al., Plant Mol Biol. 31: 957 - 973, 1996); expresión de un gen antisentido; recombinación homóloga (Ohl et al., Homologous Recombination and Gene Silencing in Plants. Kluwer, Dordrecht, Países Bajos, 1994); Sistemas Cre / lox (Qin et al., PNAS 91: 1706 - 1710, 1994; Koshinsky et al., The Plant Journal 23: 715 - 722, 2000; Chou, et al., Plant and Animal Genome VII Conference Abstracts. San Diego, CA, 17 - 21 de enero de 1999); atrapamiento de genes y etiquetado de ADN-T (Burns et al., Genes Dev. 8: 1087 - 1105, 1994; Spradling, et al., PNAS 92: 10824 - 10830, 1995; Skarnes et al., Bio/Technology 8, 827 - 831, 1990; Sundaresan, et al., Genes Dev. 9: 1797 - 1810, 1995); y cualquiera de los otros sistemas posibles de silenciamiento de genes que están disponibles en las áreas de la ciencia que resultan en la subregulación de la expresión de un polipéptido de tabaco o en una reducción de su actividad enzimática. Tal como se establece adicionalmente en este documento, cualquiera de las secuencias de ácido nucleico proporcionadas en este documento puede ser subregulada o sobrerregulada utilizando las técnicas descritas en la presente invención y otras tecnologías encontradas en la técnica. Se describen ejemplos de métodos con más detalle a continuación.

ARN de interferencia

ARN de interferencia ("ARNi") es un proceso generalmente aplicable para inducir silenciamiento génico postraduccional potente y específico en muchos organismos, incluyendo plantas (véase, por ejemplo, Boshier et al., Nat. Cell Biol. 2: E31 - 36, 2000; y Tavemarakis et al., Nat. Genetics 24: 180 - 183, 2000). El ARNi implica la introducción de ARN con carácter parcial o completamente bicatenario en la célula o en el entorno extracelular. La inhibición es específica porque se escoge esa secuencia de nucleótidos de una porción del gen objetivo (por ejemplo, una nicotina desmetilasa de tabaco) para producir RNA inhibidor. La porción elegida generalmente abarca los exones del gen objetivo, pero la porción elegida también puede incluir regiones no traducidas (UTR), así como intrones (por ejemplo, la secuencia de la SEQ ID NO: 7, o una secuencia de ácido nucleico a partir de un gen de una planta deseada, tal como cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17).

Por ejemplo, para la construcción de vectores de transformación que producen ARN capaces de formación de dúplex, dos secuencias de ácido nucleico, una en la orientación sentido y el otro en la orientación antisentido, pueden ser operativamente enlazadas, y colocadas bajo el control de un promotor viral fuerte, tal como 35S del CaMV o el promotor aislado del virus del estriado marrón de la yuca (CBSV). Sin embargo, el uso del promotor endógeno, tal como el promotor de la nicotina desmetilasa que tiene la secuencia de la SEQ ID NO: 8, o un fragmento de la misma que dirige la transcripción, también puede ser deseable. La longitud de las secuencias de ácido nucleico de la nicotina desmetilasa de tabaco incluidas en una construcción de este tipo es deseable de al menos 25 nucleótidos, pero puede abarcar una secuencia que incluye al gen para la nicotina desmetilasa de tabaco de longitud completa.

Los constructos que producen ARN capaces de formación de dúplex pueden introducirse en el genoma de una planta, tal como una planta de tabaco, mediante transformación mediada por *Agrobacterium* (Chuang et al., Proc. Natl. Acad. Sci. EE.UU. 97: 4985 - 4990, 2000), causando interferencia genética específica y heredable en una nicotina desmetilasa de tabaco. El ARN bicatenario también puede ser introducido directamente en la célula (es decir, intracelularmente) o introducido extracelularmente, por ejemplo, mediante el baño de una semilla, planta de semillero, o de la planta en una solución que contiene el ARN bicatenario.

Dependiendo de la dosis de material de ARN bicatenario suministrado, el ARNi puede proporcionar pérdida parcial o completa de la función para el gen objetivo. Se pueden obtener una reducción o pérdida de la expresión génica en al menos 99% de las células objetivo. En general, dosis inferiores del material inyectado y mayores tiempos después de la administración de ARNi dan como resultado la inhibición en una fracción más pequeña de células.

El ARN usado en el ARNi puede comprender una o más hebras de ribonucleótido polimerizado; puede incluir modificaciones ya sea en la cadena principal de fosfato-azúcar o en el nucleósido. La estructura bicatenaria puede estar formada por una sola hebra de ARN auto-complementaria o por dos hebras de ARN complementarias y la formación del dúplex de ARN puede ser iniciada dentro o fuera de la célula. El ARN puede ser introducido en una cantidad que permite el suministro de al menos una copia por célula. Sin embargo, dosis más altas (por ejemplo, al menos 5, 10, 100, 500 o 1000 copias por célula) de material bicatenario pueden producir una inhibición más eficaz. La inhibición es específica cuando las secuencias de nucleótidos correspondientes a la región dúplex del RNA están dirigidos para inhibición genética. Se prefiere ARN que contiene una secuencia de nucleótidos idéntica a una porción del gen objetivo para la inhibición. Las secuencias de ARN con inserciones, supresiones, y mutaciones puntuales individuales en relación con la secuencia objetivo también pueden ser eficaces para la inhibición. Por lo tanto, la identidad de la secuencia puede ser optimizada por algoritmos de alineación conocidos en la técnica y el cálculo de

la diferencia porcentual entre las secuencias de nucleótidos. Alternativamente, la región dúplex del RNA puede ser definida funcionalmente como una secuencia de nucleótidos que es capaz de hibridarse con una porción del transcrito del gen objetivo.

- 5 Además, el ARN utilizado para el ARNi puede ser sintetizado *in vivo* o *in vitro*. Por ejemplo, la ARN polimerasa endógena en la célula puede mediar la transcripción *in vivo*, o se puede usar ARN polimerasa clonada para la transcripción *in vivo* o *in vitro*. Para la transcripción a partir de un transgén *in vivo* o un constructo de expresión, se puede usar una región reguladora para transcribir la hebra de ARN (o hebras).

Triple hebra de interferencia

- 10 La expresión del gen endógeno para nicotina desmetilasa de tabaco o la expresión de un fragmento de ácido nucleico a partir de un gen de una planta deseada, tal como cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17, también puede ser subregulada por direccionamiento de secuencias de desoxirribonucleótido complementarias a la región reguladora de un gen de tabaco (por ejemplo, regiones promotoras o potenciadoras) para formar estructuras de triple hélice que impiden la transcripción del gen de tabaco en las células objetivo. (Véase, en general, Helene, *Anticancer Drug Des.* 6: 569 - 584, 1991; Helene et al., *Ann. N.Y. Acad. Sci.* 660: 27 - 36, 1992; y Maher, *Bioassays* 14: 807 - 815, 1992.)

- 20 Las moléculas de ácido nucleico usadas en la formación de la triple hélice para la inhibición de la transcripción son preferiblemente monocatenarias y compuestas por desoxirribonucleótidos. La composición base de estos oligonucleótidos debe promover la formación de la triple hélice a través de las reglas de apareamiento de bases de Hoogsteen, que generalmente requieren que estén presentes tramos considerables ya sea de purinas o de pirimidinas en una hebra de un dúplex. Las secuencias de nucleótidos pueden basarse en pirimidina, lo que se traducirá en las tripletas TAT y CGC a través de las tres hebras asociadas de la triple hélice resultante. Las moléculas ricas en pirimidinas proporcionan complementariedad de bases a una región rica en purina de una sola hebra del dúplex en una orientación paralela a esa hebra. Además, se pueden escoger las moléculas de ácido nucleico que sean ricas en purina, por ejemplo, que contengan un tramo de residuos de G. Estas moléculas formarán una triple hélice con un dúplex de ADN que es rico en pares GC, en donde la mayoría de los residuos de purina están ubicados en una sola hebra del dúplex objetivo, lo que resulta en tripletes CGC a través de las tres hebras en el tríplex.

- 25 Alternativamente, las secuencias potenciales que pueden ser seleccionadas para la formación de la triple hélice pueden incrementarse mediante la creación de una molécula de ácido nucleico en "zigzag". Las moléculas en zigzag se sintetizan en una forma alternante 5'-3', 3'-5', de tal manera que se aparean primero las bases con la primera cadena de un dúplex y luego con la otra, eliminando la necesidad de que esté presente un tramo considerable ya sea de purinas o de pirimidinas en una hebra de un dúplex.

Ribozimas

- 35 Las ribozimas son moléculas de ARN que actúan como enzimas y pueden ser modificadas genéticamente para escindir otras moléculas de ARN. Una ribozima puede ser diseñada para emparejarse específicamente con prácticamente cualquier ARN objetivo y escindir la cadena principal de fosfodiéster en una ubicación específica, inactivando por lo tanto funcionalmente el ARN objetivo. La ribozima misma no se consume en este proceso y puede actuar catalíticamente para escindir múltiples copias de moléculas objetivo de ARNm. Por consiguiente, las ribozimas también se pueden utilizar como un medio para subregular la expresión de una nicotina desmetilasa de tabaco. El diseño y uso de ribozimas específicas del ARN objetivo se describen en Haseloff et al. (*Nature* 334: 585 - 591, 1988). Preferiblemente, la ribozima incluye al menos aproximadamente 20 nucleótidos continuos complementarios a la secuencia objetivo (por ejemplo, una nicotina desmetilasa de tabaco o un fragmento de ácido nucleico de un gen de una planta deseada, tal como cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17) en cada lado del sitio activo de la ribozima.

- 45 Además, las secuencias de ribozimas también se pueden incluir dentro de un RNA antisentido para conferir actividad de escisión al ARN en el ARN antisentido y, de ese modo, aumentar la eficacia del constructo antisentido.

Recombinación homóloga

- 50 La tecnología de reemplazo de genes es otro método deseable para la subregulación de la expresión de un gen dado. La tecnología de reemplazo de genes se basa en la recombinación homóloga (véase, Schnable et al., *Curr. Opinions Plant Biol.* 1: 123 - 129, 1998). La secuencia de ácido nucleico de la enzima de interés tal como una nicotina desmetilasa de tabaco o un polipéptido codificado por cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17 pueden ser manipulados por mutagénesis (por ejemplo, inserciones, supresiones, duplicaciones o reemplazos) para disminuir la función enzimática. La secuencia alterada puede ser luego introducida en el genoma para reemplazar la existente, por ejemplo, del gen de tipo silvestre, a través de recombinación

homóloga (Puchta et al., Proc. Natl. Acad. Sci. EE.UU. 93: 5055 - 5060, 1996; y Kempin et al., Nature 389: 802 - 803, 1997). Alternativamente, un gen endógeno para nicotina desmetilasa de tabaco puede ser reemplazado con un gen que no tenga actividad de desmetilasa, por ejemplo, la secuencia de la SEQ ID NO: 188.

Cosupresión

5 Un método deseable adicional de silenciar la expresión génica es la cosupresión (también conocida como la
supresión sentido). Esta técnica, que implica la introducción de un ácido nucleico, por ejemplo, un fragmento de
ácido nucleico de un gen de una planta deseada, tal como cualquier secuencia de ácido nucleico mostrada en las
Figuras 1, 3 a 7, y 10 - 17, configurada en la orientación sentido, que se ha demostrado que bloquea eficazmente la
transcripción de genes objetivo (véase, por ejemplo, Napoli et al., Plant Cell, 2: 279 - 289, 1990 y Jorgensen et al.,
10 patente de los Estados Unidos No. 5.034.323).

Generalmente, la supresión sentido implica la transcripción de la secuencia introducida. Sin embargo, también
puede ocurrir cosupresión cuando la secuencia introducida no contiene una secuencia codificante en sí misma, sino
únicamente secuencias de intrón o no traducidas u otras de tales secuencias sustancialmente idénticas a las
secuencias presentes en el transcrito primario del gen endógeno que va a ser reprimido. La secuencia introducida en
15 general será sustancialmente idéntica a la del gen endógeno blanco de la represión. Tal identidad es típicamente
aproximadamente mayor al 50%, pero se prefieren identidades superiores (por ejemplo, del 80% o incluso del 95%)
porque dan lugar a una represión más eficaz. El efecto de la cosupresión también se puede aplicar a otras proteínas
dentro de una familia similar de genes que exhiben homología u homología sustancial. Los segmentos de un gen de
una planta se pueden utilizar directamente, por ejemplo, para inhibir la expresión de genes homólogos en diferentes
20 especies de plantas.

En la supresión sentido, la secuencia introducida, que requiere menos de la identidad absoluta, no necesita ser de
longitud completa, ya sea en relación con el producto de transcripción primaria o con el ARNm totalmente
procesado. Un mayor grado de identidad de secuencia en una secuencia más corta que la longitud completa
compensa una secuencia más larga de identidad menor. Además, la secuencia introducida no necesita tener el
mismo patrón de intrón o de exón, y la identidad de segmentos no codificantes puede ser igualmente eficaz. Se
25 prefieren las secuencias de al menos 50 pares de bases, siendo más preferible con secuencias introducidas de
mayor longitud (véase, por ejemplo, aquellos métodos descritos por Jorgensen et al., patente de los Estados Unidos
No. 5.034.323).

Supresión antisentido

30 En la tecnología antisentido, un segmento de ácido nucleico del gen de la planta deseada, tal como cualquier
secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17, se clona y se enlaza operativamente a una
región de control de la expresión de tal manera que se sintetiza la cadena antisentido de ARN. El constructo se
transforma a continuación en plantas y se produce la cadena antisentido de ARN. En células vegetales, se ha
demostrado que el ARN antisentido inhibe la expresión génica.

35 El segmento de ácido nucleico que se introduce en la supresión antisentido es generalmente sustancialmente
idéntica a al menos una porción del gen o genes endógenos que van a ser reprimidos, pero no tiene por qué ser
idéntico. Las secuencias de ácido nucleico de la nicotina desmetilasa de tabaco descritas en este documento
pueden incluirse en vectores diseñados de tal manera que el efecto inhibitorio se aplica a otras proteínas dentro de
una familia de genes que exhiben homología u homología sustancial con el gen objetivo. Los segmentos de un gen
40 de una planta se pueden utilizar, por ejemplo, directamente para inhibir la expresión de genes homólogos en
diferentes variedades de tabaco.

La secuencia introducida tampoco necesita ser de longitud completa con respecto ya sea al producto de
transcripción primaria o al ARNm totalmente procesado. Generalmente, se puede utilizar mayor homología para
compensar el uso de una secuencia más corta. Además, la secuencia introducida no necesita tener el mismo patrón
45 de intrón o de exón, y la homología de segmentos no codificantes será igualmente eficaz. En general, tal secuencia
antisentido usualmente tendrá al menos 15 pares de bases, preferiblemente aproximadamente 15-200 pares de
bases, y más preferiblemente 200 - 2000 pares de bases de longitud o mayor. La secuencia antisentido puede ser
complementaria a toda o a una porción del gen que va a ser suprimido, y, como apreciarán los expertos en la
técnica, el sitio o sitios particulares a los que se une la secuencia antisentido así como la longitud de la secuencia
antisentido variarán, dependiendo del grado de inhibición deseado y de la unicidad de la secuencia antisentido. Un
50 constructo transcripcional que expresa una secuencia de nucleótidos antisentido reguladora negativa de la planta
incluye, en la dirección de la transcripción, un promotor, codificando la secuencia que codifica el ARN antisentido en
la cadena sentido, y una región de terminación de la transcripción. Las secuencias antisentido pueden ser
construidas y expresadas como se describe, por ejemplo, en van der Krol et al. (Gene 72: 45 - 50, 1988); Rodermel
et al. (Cell 55: 673 - 681, 1988); Mol et al. (FEBS Lett. 268: 427 - 430, 1990); Weigel y Nilsson (Nature 377: 495 -
55 500, 1995); Cheung et al., (Cell 82: 383 - 393, 1995); y Shewmaker et al. (patente de los Estados Unidos No.

5.107.065).

Negativos dominantes

5 Las plantas transgénicas que expresan un transgén que codifica un producto génico dominante negativo de un producto génico de tabaco se pueden ensayar en entornos artificiales o en el campo para demostrar que el transgén que confiere subregula un producto génico de tabaco en la planta transgénica. Los transgenes negativos dominantes se construyen de acuerdo con métodos conocidos en la técnica. Típicamente, un gen dominante negativo codifica un polipéptido regulador negativo mutante de un producto génico de tabaco que, cuando se sobreexpresa, interrumpe la actividad de la enzima de tipo silvestre.

Mutantes

10 Las plantas que tienen menor expresión o actividad enzimática de un producto génico de tabaco también pueden ser generadas utilizando metodologías estándar de mutagénesis. Tales métodos de mutagénesis incluyen, sin limitación, el tratamiento de las semillas con metilsulfato de etilo (Hildering y Verkerk, en, *The use of induced mutations in plant breeding*. Pergamon press, páginas 317 - 320, 1965) o la irradiación UV, rayos X, y la irradiación de neutrones rápidos (véase, por ejemplo, Verkerk, *Neth. J. Agric. Sci.* 19: 197 - 203, 1971; y Poehlman, *Breeding Field Crops*, Van Nostrand Reinhold, Nueva York (3.sup.rd ed), 1987), el uso de transposones (Fedoroff et al., 1984; patente de los Estados Unidos No. 4.732.856 y la patente de los Estados Unidos No. 5.013.658), así como metodologías de inserción de T-ADN (Hoekema et al., 1983; patente de los Estados Unidos No. 5.149.645). Los tipos de mutaciones que pueden estar presentes en un gen del tabaco incluyen, por ejemplo, mutaciones puntuales, supresiones, inserciones, duplicaciones, e inversiones. Tales mutaciones deseablemente están presentes en la región codificante de un gen de tabaco; sin embargo, también pueden ser deseables las mutaciones en la región promotora, y el intrón, o una región no traducida de un gen de tabaco.

25 Por ejemplo, se puede usar mutagénesis por inserción de T-ADN para generar mutaciones por inserción en un gen del tabaco para subregular la expresión del gen. Teóricamente, se requieren cerca de 100.000 inserciones independientes de T-ADN para un 95% de probabilidad de lograr una inserción en cualquier gen dado (McKinnet, *Plant J.* 8: 613 - 622, 1995; y Forsthoefel et al., *Aust. J. Plant Physiol.* 19: 353 - 366, 1992). Las líneas etiquetadas con T-ADN de plantas se pueden cribar usando el análisis de la reacción en cadena de la polimerasa (PCR). Por ejemplo, se puede diseñar un cebador para un extremo del T-ADN y se puede diseñar otro cebador para el gen de interés y ambos cebadores se puede utilizar en el análisis de la PCR. Si no se obtiene un producto de la PCR, entonces no hay inserción en el gen de interés. En contraste, si se obtiene un producto de la PCR, entonces hay una inserción en el gen de interés.

35 La expresión de un producto génico mutado de tabaco puede ser evaluada de acuerdo con procedimientos estándar (por ejemplo, los descritos en la presente invención) y, opcionalmente, puede ser comparada con la expresión de la enzima no mutada. Cuando se compara con plantas no mutadas, las plantas mutadas que tienen menor expresión de un gen que codifica un producto génico de tabaco son realizaciones deseables de la presente invención. Una planta que tiene una mutación en cualquiera de las secuencias de ácido nucleico descritas en la presente invención puede utilizarse en un programa de fitomejoramiento como se describe aquí.

Promotores vegetales

40 Un promotor deseable es un promotor de caulimovirus, por ejemplo, un promotor del virus del mosaico de la coliflor (CaMV), o el promotor del virus del mosaico de la vena de la yuca (CsVMV). Estos promotores confieren altos niveles de expresión en la mayoría de los tejidos vegetales, y la actividad de estos promotores no depende de proteínas codificadas en forma viral. El CaMV es una fuente tanto para los promotores 35S como 19S. Los ejemplos de constructos de expresión de plantas que utilizan estos promotores son conocidos en la técnica. En la mayoría de los tejidos de las plantas transgénicas, el promotor 35S del CaMV es un promotor fuerte. El promotor del CaMV es también muy activo en monocotiledóneas. Además, la actividad de este promotor se puede aumentar adicionalmente (es decir, entre 2 - 10 veces) por duplicación del promotor 35S del CaMV.

Otros promotores vegetales útiles incluyen, sin limitación, al promotor de la nopalina sintasa (NOS), al promotor de octopina sintasa, al promotor del virus del mosaico de escrofularia (FMV), al promotor de actina del arroz, y al sistema promotor de ubiquitina.

50 Los ejemplos de promotores de monocotiledóneas incluyen, sin limitación, al promotor del virus del moteado amarillo de la Commelina, al promotor del virus de la caña de azúcar, al promotor del virus baciliforme del tungro del arroz, al elemento del virus del rayado del maíz, y al promotor del virus de enanismo del trigo.

Para ciertas aplicaciones, puede ser deseable producir un producto génico de tabaco, tal como un producto génico

mutante dominante negativo, en un tejido apropiado, a un nivel apropiado, o en un tiempo de desarrollo apropiado. Para este propósito, hay una variedad de promotores génicos, cada uno con sus propias características distintivas incorporadas en sus secuencias reguladoras, que se muestran que son regulados en respuesta a señales inducibles tales como el medio ambiente, hormonas, y/o señales de desarrollo. Estos incluyen, sin limitación, promotores génicos que son responsables de la expresión génica regulada por el calor, la expresión génica regulada por la luz (por ejemplo, el *rbcS-3A* del guisante; el promotor *rbcS* de maíz, el de un gen de la proteína que enlaza clorofila a/b que se encuentra en el guisante; o el promotor *Arabssu*), la expresión de genes regulados por hormonas (por ejemplo, las secuencias sensibles al ácido abscísico (ABA) del gen *Em* del trigo; los promotores HVA1 y HVA22 inducibles por ABA, y *rd29A* de la cebada y *Arabidopsis*; y la expresión génica inducida por lesiones (por ejemplo, de *wun1*), la expresión génica específica de un órgano (por ejemplo, del gen de la proteína de almacenamiento específica del tubérculo; el gen de la zeína de 23 kDa del maíz descrito; o el gen de la β -faseolina del frijol francés), o promotores inducibles por patógenos (por ejemplo, los promotores *PR-1*, *prp-1*, o β -1,3 glucanasa, el promotor *wiria* del trigo inducible por hongos, y los promotores inducibles por nemátodos, *TobRB7-5A* y *Hmg-1*, del tabaco y de perejil, respectivamente).

15 Vectores de expresión de plantas

Típicamente, los vectores de expresión de plantas incluyen (1) un gen clonado de una planta bajo el control transcripcional de secuencias reguladoras 5' y 3' y (2) un marcador seleccionable dominante. Tales vectores de expresión de plantas también pueden contener, si se desea, una región reguladora del promotor (por ejemplo, uno que confiere expresión inducible o constitutiva, inducida por patógenos o por lesiones, regulada por el desarrollo o por el medio ambiente o, específica de tejido o de la célula), un sitio de partida de inicio de la transcripción, un sitio de enlazamiento al ribosoma, una señal de procesamiento de ARN, un sitio de terminación de la transcripción, y/o una señal de poliadenilación.

Vectores de expresión de la planta también pueden incluir opcionalmente señales de procesamiento del ARN, por ejemplo, intrones, que se ha demostrado que son importantes para la síntesis y acumulación eficiente de ARN. La ubicación de las secuencias de empalme de ARN puede influir drásticamente en el nivel de expresión del transgén en las plantas. En vista de este hecho, se puede colocar un intrón secuencia arriba o secuencia abajo de una secuencia de codificación de la nicotina desmetilasa del tabaco en el transgén para alterar los niveles de expresión del gen.

Además de las secuencias reguladoras de control 5' mencionados anteriormente, los vectores de expresión también pueden incluir regiones reguladoras de control que están generalmente presentes en las regiones 3' de los genes de las plantas. Por ejemplo, la región terminadora 3' se puede incluir en el vector de expresión para aumentar la estabilidad del ARNm. Una de tales regiones terminadoras se puede derivar de la región terminadora PI-II de la patata. Además, otros terminadores comúnmente utilizados se derivan de las señales de la octopina o nopalina sintasa.

El vector de expresión de la planta también contiene típicamente un gen marcador seleccionable dominante utilizado para identificar aquellas células que se han transformado. Genes seleccionables útiles para los sistemas vegetales incluyen el gen de la aminoglucósido fosfotransferasa del transposón Tn5 (Aph II), genes que codifican genes de resistencia a antibióticos, por ejemplo, aquellos que codifican la resistencia a la higromicina, kanamicina, bleomicina, neomicina, G418, estreptomycinina o espectinomycinina. Los genes requeridos para la fotosíntesis también pueden ser utilizados como marcadores seleccionables en cepas fotosintéticas deficientes. Por último, los genes que codifican la resistencia a herbicidas pueden utilizarse como marcadores seleccionables; los genes útiles de resistencia a herbicidas incluyen el gen *bar* que codifica la enzima fosfinotricina acetiltransferasa y que confiere resistencia al herbicida de amplio espectro Basta® (Bayer CropScience Deutschland GmbH, Langenfeld, Alemania). Otros marcadores seleccionables incluyen genes que proporcionan resistencia a otros herbicidas tales como el glifosato y similares, y herbicidas de imidazolinonas, sulfonilureas, triazolopirimidina, tales como clorosulfurona, bromoxinilo, dalapon, y similares. Además, los genes que codifican la dihidrofolato reductasa se pueden usar en combinación con moléculas tales como metotrexato.

El uso eficiente de marcadores seleccionables se ve facilitada por una determinación de la susceptibilidad de una célula de la planta a un agente seleccionable particular y una determinación de la concentración de este agente que mata efectivamente la mayoría, si no todas, las células transformadas. Algunas concentraciones útiles de antibióticos para la transformación del tabaco incluyen, por ejemplo, 20 - 100 μ g/ml (kanamicina), 20 - 50 μ g/ml (higromicina), o 5 - 10 μ g/ml (bleomicina). Una estrategia útil para la selección de transformantes para resistencia a herbicidas es descrita, por ejemplo, por Vasil (Cell Culture and Somatic Cell Genetics of Plants, Vol I, II, III Laboratory Procedures and their Applications, Academic Press, Nueva York, 1984).

Además de un marcador seleccionable, puede ser deseable usar un gen reportero. En algunos casos, se puede usar un gen reportero sin un marcador seleccionable. Los genes reporteros son genes que típicamente no están presentes o se expresan en el organismo o tejido receptor. El gen reportero típicamente codifica para una proteína

que proporciona algún cambio fenotípico o propiedad enzimática. Ejemplos de tales genes se proporcionan en Weising et al. (Ann. Rev. Genetics 22: 421, 1988). Los genes reporteros preferidos incluyen, sin limitación al gen de la glucuronidasa (GUS) y genes GFP.

5 Tras la construcción del vector de expresión vegetal, varios métodos convencionales están disponibles para la introducción del vector en un huésped vegetal, generando de ese modo una planta transgénica. Estos métodos incluyen (1) transformación mediada por *Agrobacterium* (*A. tumefaciens* o *A. rhizogenes*) (véase, por ejemplo, Lichtenstein y Fuller en: Genetic Engineering, vol 6, PWJ Rigby, ed, Londres, Academic Press, 1987; y Lichtenstein, CP, y Draper, J., en: DNA Cloning, vol II, D.M. Glover, ed, Oxford, IRI Press, 1985; las patentes de los Estados Unidos Nos. 4.693.976, 4.762.785, 4.940.838, 5.004.863, 5.104.310, 5.149.645, 5.159.135, 5.177.010, 5.231.019, 10 5.463.174, 5.469.976, y 5.464.763; y las solicitudes de patente europea Nos. 0131624, 0159418, 0120516, 0176112, 0116718, 0290799, 0292435, 0320500, y 0627752, y las solicitudes de patente europea Nos. 0267159 y 0604622), (2) el sistema de suministro de partículas (véase, por ejemplo, las patentes de los Estados Unidos Nos. 4.945.050 y 5.141.131), (3) protocolos de microinyección, (4) procedimientos con polietilenglicol (PEG), (5) absorción de ADN mediada por liposomas, (6) protocolos de electroporación (véase, por ejemplo, el documento WO 87/06614 y las patentes de los Estados Unidos Nos. 5.384.253, 5.472.869, 5.641.664, 5.679.558, 5.712.135, 6.002.070, y 15 6.074.877, (7) el método de agitación tipo vórtice, u (8) la así llamada metodología de bigotes (véase, por ejemplo, Coffee et al., patentes de los Estados Unidos Nos. 5.302.523 y 5.464.765). El tipo de tejido vegetal que puede ser transformado con un vector de expresión incluye tejido embrionario, tejido de callo tipo I y II, hipocótilos, meristemo, y similares.

20 Una vez introducido en el tejido vegetal, la expresión del gen estructural puede analizarse por cualquier medio conocido en la técnica, y puede medirse la expresión como ARNm transcrito, proteína sintetizada, o la cantidad de silenciamiento de genes que se produce según lo determinado por la supervisión de metabolitos a través de análisis químico de alcaloides secundarios en el tabaco (como se describe en la presente invención; véase también la patente de los Estados Unidos No. 5.583.021). Se conocen técnicas para el cultivo *in vitro* de tejido vegetal, y en un número de casos, para la regeneración en plantas completas (véase, por ejemplo, las patentes de los Estados Unidos Nos. 5.595.733 y 5.766.900). Los procedimientos para transferir el complejo de expresión introducido en variedades de cultivo comercialmente útiles son conocidos por los expertos en la técnica.

30 Una vez que se obtienen células vegetales que expresan el nivel deseado de un producto génico deseable, se pueden regenerar tejidos de las plantas y plantas enteras a partir de las mismas usando métodos y técnicas bien conocidas en la técnica. Las plantas regeneradas se reproducen luego por medios convencionales y los genes introducidos pueden transferirse a otras cepas y variedades cultivadas mediante técnicas convencionales de fitomejoramiento de plantas.

35 Las plantas de tabaco transgénicas pueden incorporar un ácido nucleico de cualquier porción del gen genómico en diferentes orientaciones o bien para subregulación, por ejemplo, orientación antisentido, o sobreexpresión, por ejemplo, orientación sentido. La sobreexpresión de la secuencia de ácido nucleico que codifica la totalidad o una parte funcional de una secuencia de aminoácidos de un gen de tabaco de longitud completa es deseable para aumentar la expresión del producto génico dentro de las líneas de *Nicotiana*.

Determinación de los niveles de transcripción o traducción de un gen de tabaco

40 La expresión génica se puede medir, por ejemplo, mediante análisis estándar de transferencia tipo Northern (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001), y Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, (1989)) usando un gen o fragmento de gen de tabaco como sonda de hibridación. La determinación de los niveles de expresión de ARN también puede ser ayudada por PCR de transcripción inversa (rtPCR), incluyendo rtPCR cuantitativa (véase, por ejemplo, Kawasaki et al., en PCR Technology: Principles and Applications of DNA Amplification (H.A. Erlich, Ed) Stockton Press (1989); 45 Wang et al., en PCR Protocols: A Guide to Methods and Applications (M.A. Innis, et al., Eds.) Academic Press (1990); y Freeman et al., Biotechniques 26: 112 -122 y 124 - 125, 1999). Técnicas adicionales bien conocidas para la determinación de la expresión de un gen de tabaco incluyen hibridación *in situ*, y la hibridación fluorescente *in situ* (véase, por ejemplo, Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001)). Las técnicas estándar anteriores también son útiles para comparar el nivel de expresión entre las plantas, 50 por ejemplo, entre una planta que tiene una mutación en un gen de tabaco y una planta de control.

Si se desea, la expresión de un gen de tabaco (por ejemplo, una secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17, o una fragmento de la misma) puede medirse al nivel de la producción de proteína utilizando el mismo enfoque general y técnicas de análisis convencionales de proteína incluyendo ensayos de Bradford, ensayos espectrofotométricos, y técnicas de detección inmunológicas, tales como transferencia tipo Western o inmunoprecipitación con un anticuerpo específico para el polipéptido deseable (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001), y Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY, (1989)).

La actividad de cualquier polipéptido descrito en este documento puede ser analizada utilizando métodos convencionales en la técnica. Por ejemplo, la actividad de una p450 se analiza normalmente utilizando ensayos basados en fluorescencia (véase, por ejemplo, Donato et al. Drug Metab Dispos. 32: 699 - 706, 2004). En particular, la actividad de una nicotina desmetilasa puede ser analizada como se describe en la presente invención usando ensayos de microsomas de levadura.

Usos

La regulación del gen endógeno correspondiente a cualquiera de las secuencias descritas en este documento, por ejemplo, el silenciamiento génico puede dar lugar a plantas o productos vegetales más valiosos. En particular, las secuencias identificadas en este documento como las inducidas por etileno o las relacionados con la senescencia (por ejemplo, aquellas que tienen la secuencia de las SEQ ID NOS: 4, 188, y 226, o una secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10 - 17, o una fragmento de la misma) pueden ser utilizadas para afectar las rutas metabólicas implicadas en la formación de numerosos metabolitos secundarios, incluyendo terpenoides, polifenoles, alcaloides, etc., que afectan los rasgos de calidad del producto final. Del mismo modo, los genes identificados en la presente invención pueden ser utilizados para regular las rutas metabólicas asociadas con la tasa y el tipo de la materia seca acumulada durante la senescencia o la repartición de la materia seca dentro de la planta durante la senescencia. La regulación de los genes identificados en este documento también puede usarse para afectar a las rutas metabólicas implicadas en la determinación de las tasas de senescencia, la uniformidad de la senescencia dentro de una hoja y entre las hojas de una sola planta, y la inducción de la senescencia por agentes o actividades que estimulan o activan los genes identificados en la presente invención, y, de ese modo, controlar la calidad de un producto o artículo de fabricación que incluye una hoja u otro componente de la planta.

La región promotora de un gen descrito en este documento puede ser utilizada para dirigir la expresión de cualquier producto génico deseable para mejorar la calidad de la cosecha o mejorar rasgos específicos. La regulación de una secuencia del promotor también puede ser utilizada para subregular genes endógenos de tabaco, incluyendo genes implicados en la biosíntesis de alcaloides y/o en otras rutas. Deseablemente, se utiliza una región promotora del gen de tabaco u otra región reguladora de la transcripción para alterar propiedades químicas tales como el contenido de nicotina y los niveles de nitrosamina en una planta. Además, los motivos del promotor, que pueden identificarse fácilmente en una secuencia del promotor usando métodos convencionales en la técnica, pueden utilizarse para identificar los factores que se asocian con, o que regulan la expresión de un producto génico de tabaco, por ejemplo, un p450.

Además, cualquiera de las secuencias de la presente invención (por ejemplo, las secuencias de ácido nucleico mostradas en las Figura 1, en las Figuras 3 a 7, y 10 - 17, o fragmentos de las mismas) pueden usarse en métodos que reducen la expresión génica o alteran la actividad enzimática de un producto génico, tal como un p450, utilizando técnicas convencionales descritas en la presente invención. Tales técnicas incluyen, sin limitación, ARN de interferencia, triple hebra de interferencia, ribozimas, recombinación homóloga, silenciamiento génico inducido por virus, tecnologías antisentido y de cosupresión, la expresión de un producto génico dominante negativo, y la generación de genes mutados usando técnicas de mutagénesis estándar. Por ejemplo, la reducción de la expresión de p450 o la alteración de la actividad enzimática de p450 se puede utilizar para alterar los ácidos grasos que están implicados en las interacciones planta-patógeno y la resistencia a enfermedades o pueden utilizarse para alterar el perfil de una planta de ácidos grasos seleccionados y por lo tanto alterar el sabor o el aroma de la planta o el componente de la planta.

Además, utilizando métodos convencionales, se puede utilizar cualquier porción de un gen del tabaco, incluyendo el promotor, la secuencia codificadora, un intrón, o una 3'UTR, o un fragmento de los mismos, como un marcador genético para aislar genes relacionados, promotores o regiones reguladoras, para la selección de los genes relacionados en otras especies de tabaco o de *Nicotiana*, o para determinar si una planta tiene una mutación en un gen endógeno correspondiente. También se puede usar una porción de un gen del tabaco para controlar el flujo de genes a través de un esfuerzo de fitomejoramiento para rastrear la introgresión o la pérdida de un gen particular.

Por ejemplo, *Nicotiana tabacum* es un alotetraploide, como lo son varias de las otras especies de *Nicotiana*, y los marcadores genéticos podrían ser utilizados para identificar genes homólogos o genes relacionados en el genoma progenitor diferente del genoma en el que reside el gen original. Un marcador para el gen relacionado también podría usarse para seleccionar germoplasma de tabaco existente, poblaciones segregantes o sintéticas creadas por hibridaciones, poblaciones creadas a partir de tratamientos mutagénicos o de diferentes métodos de cultivo de tejidos. Como tales, las secuencias de ácido nucleico descritas en la presente invención (por ejemplo, las secuencias de ácido nucleico mostradas en la Figura 1, las Figuras 3 a 7, y 10 - 17, o fragmentos de las mismas) pueden utilizarse para identificar o afectar a genes implicados en la enfermedad o resistencia a los insectos, las propiedades de sabor y aroma, tolerancia a los herbicidas, los factores de calidad relacionados con constituyentes indeseables, o que aumenten el rendimiento de las hojas, o afectan a los componentes de las hojas de las plantas, tales como ligninas, celulosa, etc., relacionados con rasgos estructurales o el contenido de fibra.

Productos

Los productos del tabaco que tienen una cantidad reducida de contenido de nitrosamina se fabrican utilizando cualquiera de los materiales de la planta de tabaco descritos en este documento de acuerdo con métodos convencionales conocidos en la técnica. En una realización, los productos de tabaco se fabrican usando material de la planta de tabaco obtenido de una planta de tabaco curado. La planta de tabaco curado puede contener o haber sido fitomejorada para tener una actividad reducida de nicotina desmetilasa. Por ejemplo, la planta de tabaco curado puede ser una planta de tabaco resultante de un cruce que incluye una planta de tabaco identificada por tener una variante de expresión de la nicotina desmetilasa. Deseablemente, el producto de tabaco tiene una cantidad reducida de nornicotina o NNN de menos de aproximadamente 5 mg / g, 4,5 mg / g, 4,0 mg / g, 3,5 mg / g, 3,0 mg / g, 2,5 mg / g, 2,0 mg / g, 1,5 mg / g, 1,0 mg / g, 750 µg / g, 500 µg / g, 250 µg / g, 100 µg / g, 75 µg / g, 50 µg / g, 25 µg / g, 10 µg / g, 7,0 µg / g, 5,0 µg / g, 4,0 µg / g, 2,0 µg / g, 1,0 µg / g, 0,5 µg / g, 0,4 µg / g, 0,2 µg / g, 0,1 µg / g, 0,05 µg / g, o 0,01 µg / g o en donde el porcentaje de alcaloides secundarios con respecto al contenido total de alcaloides contenido allí es inferior al 90%, 70%, 50%, 30%, 10%, 5%, 4%, 3%, 2%, 1,5%, 1%, 0,75%, 0,5%, 0,25%, o 0,1%. La frase "una cantidad reducida" se refiere a una cantidad de nornicotina o NNN o ambas en una planta de tabaco o componente de la planta de tabaco o un producto de tabaco que es menor que el que se encontraría en una planta de tabaco o componente de la planta de tipo silvestre o producto de tabaco de la misma variedad de tabaco, procesado de la misma manera, que no se volvió transgénico para nornicotina o NNN reducida. En un ejemplo, una planta de tabaco de tipo silvestre de la misma variedad que ha sido procesada de la misma manera se utiliza como control para medir si se ha obtenido una reducción de nornicotina o NNN o ambas por los métodos descritos en la presente invención. En otro ejemplo, se evalúan plantas que tienen una cantidad reducida de contenido de nitrosamina utilizando métodos convencionales, por ejemplo, mediante el control de la presencia o ausencia de un gen o producto génico, por ejemplo, una nicotina desmetilasa, o una mutación particular en un gen. En aún otro ejemplo, el contenido de nitrosamina de las plantas resultantes de un programa de fitomejoramiento se compara con el contenido de nitrosamina de la línea destinataria o de la línea del donante, o ambos, usada para reproducir la planta que tiene la cantidad reducida de nitrosamina. Otros controles adecuados conocidos en la técnica también se utilizan según sea necesario. Los niveles de nornicotina y NNN o de ambos se miden de acuerdo con métodos bien conocidos en la técnica del tabaco.

Los siguientes ejemplos ilustran métodos para llevar a cabo la invención y debe entenderse que son ilustrativos, pero no limitantes, del alcance de la invención que se define en las reivindicaciones adjuntas.

30 Ejemplo 1

Desarrollo del tejido de la planta y tratamiento con etileno

Crecimiento de la planta

Las plantas fueron sembradas en macetas y se desarrollaron en un invernadero durante 4 semanas. Se trasplantaron las plántulas de 4 semanas de edad a macetas individuales y crecieron en el invernadero durante 2 meses. Se regaron las plantas 2 veces al día con agua que contenía 150 ppm de fertilizante de NPK durante el crecimiento. Las hojas verdes expandidas fueron separadas de las plantas para hacer el tratamiento con etileno descrito a continuación.

Línea celular 78379

La línea de tabaco 78379, que es una línea de tabaco Burley liberada por la Universidad de Kentucky se utilizó como fuente de material vegetal. Se cultivaron cien plantas como estándar en la técnica de cultivo de tabaco, se las trasplantó, y etiquetó con un número distintivo (1-100). La fertilización y el manejo de campo se realizaron como se recomienda.

Tres cuartas partes de las 100 plantas convirtieron entre el 20 y el 100% de la nicotina en nornicotina. Una cuarta parte de las 100 plantas convirtió menos del 5% de la nicotina en nornicotina. La planta número 87 tuvo la menor conversión (2%), mientras que la planta número 21 tuvo un 100% de conversión. Las plantas que convirtieron menos del 3% se clasificaron como no convertidoras. Se realizó la autopolinización de la semilla de la planta número 87 y de la planta número 21, así como también se cruzaron (21 x 87 y 87 x 21) para estudiar las diferencias genéticas y fenotípicas. Las plantas producto de la autofecundación de la planta 21 eran convertidoras, y el 99% de las autofecundadas a partir de planta 87 eran no convertidores. El otro 1% de las plantas de la número 87 mostraron una conversión baja (5-15%). Las plantas provenientes de cruzamientos recíprocos eran todas convertidoras.

Línea celular 4407

La línea de Nicotiana 4407, que es una línea Burley, se utilizó como fuente de material vegetal. Se seleccionaron las

plantas en forma uniforme y representativa (100) y se etiquetaron. De las 100 plantas, 97 eran no convertidoras y tres eran convertidoras. La planta número 56 tuvo la menor cantidad de conversión (1,2%) y la planta número 58 tuvo el más alto nivel de conversión (96%). Las semillas autopolinizadas y las semillas cruzadas se hicieron con estas dos plantas.

- 5 Las plantas de la planta 58 autofecundada fueron segregadas con una relación 3:1 de convertidoras con respecto a las no convertidoras. Las plantas 58 - 33 y 58 - 25 fueron identificadas como líneas de plantas homocigotas convertidoras y no convertidoras, respectivamente. La conversión estable de 58 - 33 se confirmó por análisis de su progenie.

Línea celular PBLB01

- 10 PBLB01 es una línea Burley desarrollada por ProfiGen, Inc. y fue utilizada como fuente de material vegetal. La planta convertidora fue seleccionada de las semillas básicas de PBLB01.

Procedimientos de tratamiento con etileno

- 15 Las hojas verdes se separaron de las plantas cultivadas en invernadero de 2 - 3 meses y se rociaron con una solución de etileno al 0,3% (marca de la preparación Ethephon (Rhône-Poulenc)). Cada hoja rociada fue colgada en un estante de curado equipado con un humidificador y cubierto con plástico. Durante el tratamiento, se rociaron periódicamente las hojas de muestra con la solución de etileno. Aproximadamente 24 - 48 horas después del tratamiento con etileno, se recogieron las hojas para la extracción del RNA. Se tomó otra submuestra para el análisis de los constituyentes metabólicos para determinar la concentración de los metabolitos de las hojas y los constituyentes más específicos de interés, tales como una variedad de alcaloides.

- 20 Como un ejemplo, el análisis de los alcaloides se puede realizar de la siguiente manera. Las muestras (0,1 g) se agitaron a 150 rpm con 0,5 ml de NaOH 2 N, y una solución de extracción de 5 ml que contenía quinolina como estándar interno y metil-t-butil éter. Las muestras se analizaron en un cromatógrafo de gases HP 6890 equipado con un detector FID. Se usó una temperatura de 250 °C para el detector y el inyector. Se usó una columna HP (30 m - 0,32 mm - 1 mm) que consistía de sílice fundida entrecruzada con 5% de fenol y 95% de metil silicona en un gradiente de temperatura de 110 - 185 °C a razón de 10 °C por minuto. La columna se operó a 100 °C con una velocidad de flujo de 1,7 cm³ · min⁻¹ con una relación de división de 40: 1 con un volumen de inyección de 2:1 utilizando helio como el gas portador.
- 25

Ejemplo 2

Aislamiento del ARN

- 30 Para las extracciones de ARN, se trataron hojas medianas de las plantas cultivadas en invernadero de dos meses de edad con etileno como se describió anteriormente. Se usaron muestras de 0 y 24-48 horas para la extracción del RNA. En algunos casos, las muestras de hojas para el proceso de senescencia se tomaron de las plantas 10 días después de la remoción de la cabeza de la flor. Estas muestras se utilizaron también para la extracción. Se aisló el ARN total utilizando RNeasy Plant Mini Kit® (Qiagen, Inc., Valencia, California) de acuerdo con el protocolo del fabricante.
- 35

- 40 La muestra de tejido se molió en nitrógeno líquido hasta un polvo fino usando un mortero y mano de mortero tratados con DEPC. Se transfirieron aproximadamente 100 miligramos de tejido molido a un tubo Eppendorf estéril de 1,5 ml. Se colocó este tubo de muestra en nitrógeno líquido hasta que se recogieron todas las muestras. Luego, se añadieron 450 µl de regulador RLT como el suministrado en el kit (con la adición de mercaptoetanol) a cada tubo individual. Se agitó la muestra vigorosamente formando un vórtice y se incubó a 56 °C durante 3 minutos. Se aplicó luego el lisado a la columna de centrifugación QIAshredder® colocado en un tubo de recolección de 2 ml, y se centrifugó durante 2 minutos a velocidad máxima. Se recolectó el flujo que pasó a través de la misma y se añadieron 0,5 volúmenes de etanol al lisado aclarado. Se mezcló bien la muestra y se la transfirió a una columna de centrifugación mini RNeasy® colocada en un tubo de recolección de 2 ml. Se centrifugó la muestra durante 1 minuto a 10.000 rpm. A continuación, se pipetearon 700 µl de regulador RW1 sobre la columna de RNeasy® y se centrifugó durante 1 minuto a 10.000 rpm. Se pipeteó regulador RPE sobre la columna de RNeasy® en un nuevo tubo de recolección y se centrifugó durante 1 minuto a 10.000 rpm. Se añadió nuevamente regulador RPE a la columna de centrifugación de RNeasy® y se centrifugó durante 2 minutos a velocidad máxima para secar la membrana. Para eliminar cualquier etanol remanente, se colocó la membrana en un tubo de recolección separado y se centrifugó durante 1 minuto adicional a máxima velocidad. Se transfirió la columna de RNeasy® a un nuevo tubo de recolección de 1,5 ml, y se pipetearon 40 µl de agua libre de RNasa directamente sobre la membrana de RNeasy®. Se centrifugó este tubo de elución final durante 1 minuto a 10.000 rpm. Se analizó la calidad y cantidad de ARN total a través de gel de formaldehído desnaturalizado y espectrofotómetro.
- 50

Se aisló el ARN poli (A) utilizando el kit de purificación poli A + ARN Oligotex® (Qiagen Inc.) de acuerdo con el protocolo del fabricante. Se utilizaron alrededor de 200 µg de ARN total en un volumen máximo de 250 µl. Se añadió un volumen de 250 µl del regulador OBB y 15 µl de suspensión de Oligotex® a los 250 µl de ARN total. Se mezclaron completamente los contenidos por pipeteo y se incubaron durante 3 minutos a 70 °C en un bloque de calentamiento. A continuación, se colocó la muestra a temperatura ambiente durante aproximadamente 20 minutos. El Oligotex®: se sedimentó complejo de ARNm por centrifugación durante 2 minutos a velocidad máxima. Se removió todo menos 50 µl del sobrenadante del tubo de microcentrífuga. Se trató la muestra adicionalmente por el regulador OBB. El Oligotex®: se resuspendió el sedimento de ARNm en 400 µl de regulador OW2 mediante agitación tipo vórtice. Se transfirió esta mezcla a una columna de centrifugación pequeña colocada en un nuevo tubo y se centrifugó durante 1 minuto a velocidad máxima. Se transfirió la columna de centrifugación a un nuevo tubo y se añadió a la columna 400 µl adicionales de regulador OW2. Se centrifugó luego el tubo durante 1 minuto a velocidad máxima. Se transfirió la columna de centrifugación a un tubo final de microcentrífuga de 1,5 ml. Se eluyó la muestra con 60 µl de regulador OEB caliente (70 °C). Se analizó el producto Poli A por geles de formaldehído desnaturalizado y análisis espectrofotométrico.

15 Ejemplo 3

PCR de transcripción inversa

La primera cadena de ADNc fue producida usando la transcriptasa inversa SuperScript de acuerdo con el protocolo del fabricante (Invitrogen, Carlsbad, California). La mezcla de ARN enriquecido en poli A+ / cebador oligo dT consistió en menos de 5 µg de ARN total, 1 µl de mezcla dNTP 10 mM, 1 µl de Oligo d(T)₁₂₋₁₈ (0,5 µg / µl), y hasta 10 µl de agua tratada con DEPC. Cada muestra se incubó a 65 °C durante 5 minutos, luego se colocó sobre hielo durante al menos 1 minuto. Se preparó una mezcla de reacción añadiendo cada uno de los siguientes componentes en orden: 2 µl de regulador 10X RT, 4 µl de MgCl₂ 25 mM, 2 µl de DTT 0,1 M, y 1 µl de inhibidor de RNasa recombinante RNasa OUT. Se pipeteó una adición de 9 µl de mezcla de reacción para cada mezcla de ARN / cebador y se mezcló suavemente. Se incubó a 42 °C durante 2 minutos y se añadió 1 µl de SuperScript II RT a cada tubo. El tubo se incubó durante 50 minutos a 42 °C. Se terminó la reacción a 70 °C durante 15 minutos y se enfrió sobre hielo. Se recogió la muestra por centrifugación y se añadió 1 µl de RNasa H a cada tubo y se incubó durante 20 minutos a 37 °C. Se llevó a cabo la segunda PCR con 200 pmoles de cebador directo y 100 pmoles de cebador inverso (mezcla de oligo d(T) de 18 nt, seguido de 1 base al azar).

Las condiciones de reacción fueron de 94 °C durante 2 minutos y luego 40 ciclos de PCR a 94 °C durante 1 minuto, 45 °C a 60 °C durante 2 minutos, 72 °C durante 3 minutos, con una extensión a 72 °C 10 min adicionales. Se analizaron 10 microlitros de la muestra amplificada por electroforesis usando un gel de agarosa al 1%. Se purificaron los fragmentos de tamaño correcto del gel de agarosa.

Ejemplo 4

Generación de poblaciones de fragmentos de PCR

Se ligaron los fragmentos de PCR del Ejemplo 3 en un vector pGEM-T Easy (Promega, Madison, Wisconsin) de acuerdo con las instrucciones del fabricante. El producto ligado se transformó en células competentes JM109 y se sembró en placas sobre medio LB para selección azul / blanco. Se seleccionaron las colonias y se las cultivó en una placa de 96 pozos con 1,2 ml de medio LB durante la noche a 37 °C. Se generó una existencia para todas las colonias seleccionadas. Se purificó el ADN del plásmido de las placas utilizando robótica miniprep Biomeck 2000 de Beckman con un kit Wizard SV Miniprep (Promega). Se eluyó el ADN del plásmido con 100 µl de agua y se almacenó en una placa de 96 pozos. Se digirieron los plásmidos mediante EcoR1 y se analizaron utilizando gel de agarosa al 1% para confirmar la cantidad de ADN y el tamaño de los insertos. Se secuenciaron los plásmidos que contenían un inserto de 400-600 pb utilizando un secuenciador CEQ 2000 (Beckman, Fullerton, California). Se alinearon las secuencias con la base de datos del GenBank mediante una búsqueda con BLAST. Se identificaron y analizaron los fragmentos relacionados con p450. Alternativamente, se aislaron fragmentos de p450 a partir de bibliotecas de sustracción. También se analizaron estos fragmentos como se describió anteriormente.

Ejemplo 5

Construcción de una biblioteca de ADNc

Se construyó una biblioteca de ADNc mediante la preparación de ARN total de hojas tratadas de etileno de la siguiente manera. En primer lugar, se extrajo el ARN total de las hojas tratadas con etileno de línea de tabaco 58 - 33 utilizando un fenol modificado con ácido y el protocolo de extracción con cloroformo. El protocolo fue modificado para utilizar un gramo de tejido que se trituró y se agitó posteriormente con agitación tipo vórtice en 5 ml de regulador de extracción (Tris-HCl 100 mM, pH 8,5; NaCl 200 mM; EDTA 10 mM; 0,5% de SDS) a la que se le añadió 5 ml de fenol (pH 5,5) y 5 ml de cloroformo. La muestra extraída se centrifugó y se guardó el sobrenadante. Esta

etapa de extracción se repitió 2 - 3 veces hasta que el sobrenadante aparecía claro. Se añadieron aproximadamente 5 ml de cloroformo para eliminar las trazas de fenol. Se precipitó el ARN a partir de las fracciones de sobrenadante combinadas mediante la adición de un volumen de etanol 3 veces y 1/10 de volumen de NaOAc 3 M (pH 5,2) y almacenando a -20 °C durante 1 hora. Después de la transferencia a un recipiente de vidrio Corex, se centrifugó la fracción de ARN a 9000 RPM durante 45 minutos a 4 °C. Se lavó el sedimento con etanol al 70% y se centrifugó durante 5 minutos a 9000 rpm a 4 °C. Después de secar el sedimento, se disolvió el ARN sedimentado en 0,5 ml de agua libre de RNasa. La calidad y la cantidad del ARN total se analizó por gel de formaldehído desnaturalizado y espectrofotómetro, respectivamente.

El ARN total resultante se utilizó para aislar ARN poli A+ usando un protocolo de celulosa de oligo (dT) (Invitrogen) y columnas de centrifugación de microcentrífuga (Invitrogen) mediante el siguiente protocolo. Se sometieron aproximadamente 20 mg de ARN total a purificación dos veces para obtener ARN poli A+ de alta calidad. Se analizó el producto de ARN poli A+ mediante gel de formaldehído desnaturalizado y posterior RT-PCR de genes conocidos de longitud completa para asegurar una alta calidad del ARNm.

A continuación, se utilizó ARN poli A+ como molde para producir una biblioteca de ADNc empleando el kit de síntesis de ADNc, el kit de síntesis de ZAP-ADNc, y el kit de clonación dorado ZAP-ADNc Gigapack III (Stratagene, La Jolla, California). El método implicaba seguir el protocolo del fabricante como se especifica. Se usaron aproximadamente 8 µg de ARN poli A+ para construir la biblioteca de ADNc. El análisis de la biblioteca primaria reveló aproximadamente $2,5 \times 10^6$ - 1×10^7 pfu. Se completó una prueba de calidad de fondo de la biblioteca mediante ensayos de complementación utilizando IPTG y X-gal, donde las plaquetas recombinantes se expresaron más de 100 veces por encima de la reacción de fondo.

Un análisis más cuantitativo de la biblioteca mediante PCR al azar mostró que el tamaño promedio del ADNc del inserto fue de aproximadamente 1,2 kb. El método utiliza un método de PCR de dos etapas. Para la primera etapa, se diseñaron cebadores inversos con base en la información preliminar de la secuencia obtenida a partir de fragmentos de p450. Se usaron los cebadores inversos y los cebadores T3 (directos) diseñados para amplificar los genes correspondientes de la biblioteca de ADNc. Se sometieron las reacciones de PCR a electroforesis en agarosa y se cortaron las bandas correspondientes de alto peso molecular, se purificaron, clonaron y secuenciaron. En la segunda etapa, se usaron los nuevos cebadores diseñados a partir de 5'UTR o de la región de codificación de partida de p450 como los cebadores directos junto con los cebadores inversos (diseñados a partir de 3'UTR de p450) en la PCR subsiguiente para obtener clones de p450 de longitud completa.

Los fragmentos de p450 se generaron mediante amplificación por PCR de la biblioteca de ADNc construida como se describe en el Ejemplo 3 con la excepción del cebador inverso. El cebador T7 localizado en el plásmido secuencia abajo de los insertos de ADNc se utilizó como cebador inverso. Se aislaron fragmentos de PCR, se clonaron y se secuenciaron como se describe en el Ejemplo 4.

Se aislaron genes para p450 de longitud completa mediante este método de PCR a partir de la biblioteca de ADNc construida. Se usaron los cebadores inversos específicos de los genes (diseñados a partir de la secuencia, secuencia abajo de los fragmentos de p450) y un cebador directo (T3 en el plásmido de la biblioteca) para clonar los genes de longitud completa. Se aislaron, clonaron y secuenciaron los fragmentos de PCR. Si fuera necesario, se aplica una segunda etapa de PCR. En la segunda etapa, se utilizaron los nuevos cebadores directos diseñados a partir de 5'UTR de los p450 clonados junto con los cebadores inversos diseñados a partir de 3'UTR de clones de p450 en las reacciones PCR posteriores para obtener clones de p450 de longitud completa. Los clones se secuenciaron posteriormente.

Ejemplo 6

Identidad del ácido nucleico, relación estructural del fragmento de ácido nucleico aislado, e hibridación con GeneChip®

Se secuenciaron más de 100 fragmentos de p450 clonados junto con el análisis de transferencia tipo Northern para determinar su relación estructural. El enfoque utilizó cebadores directos con base en cualquiera de los dos motivos p450 comunes situados cerca del terminal carboxilo de los genes para p450. Los cebadores directos correspondieron a motivos del citocromo p450 FXPERF (SEQ ID NO: 2268) o GRRXCP (A/G) (SEQ ID NO: 2269). Los cebadores inversos usaron cebadores estándar ya sea del plásmido, SP6 o T7 situado en ambos brazos del plásmido pGEM, o una cola de poli A. El protocolo utilizado se describe a continuación.

Se usó espectrofotometría para estimar la concentración del ADN bicatenario de partida de acuerdo con el protocolo del fabricante (Beckman Coulter). La plantilla se diluyó con agua hasta la concentración apropiada, desnaturalizado por calentamiento a 95 °C durante 2 minutos, y posteriormente se colocó sobre hielo. La reacción de secuenciación se preparó sobre hielo utilizando 0,5 a 10 µl de la plantilla de ADN desnaturalizado, 2 µl de 1,6 pmoles del cebador directo, 8 µl de la mezcla DTC Quick Start Master y se llevó el volumen total hasta 20 ml con agua. El programa de

termociclado consistió en 30 ciclos del siguiente ciclo: 96 °C durante 20 segundos, 50 °C durante 20 segundos, y 60 °C durante 4 minutos, seguido por mantenimiento a 4 °C.

5 La reacción de secuenciación se detuvo añadiendo 5 µl de regulador de parada (igual volumen de NaOAc 3 M y EDTA 100 mM y 1 µl de glicógeno de 20 mg / ml). La muestra se precipitó con 60 µl de etanol frío al 95% y se centrifugó a 6.000 x g durante 6 minutos. El etanol se descartó. El sedimento se lavó dos veces con 200 µl de etanol frío al 70%. Después de secar el sedimento, se añadieron 40 µl de una solución de SLS y se resuspendió el sedimento. Se colocó encima una capa de aceite mineral y se colocó la muestra en el secuenciador automatizado CEQ 8000 para su posterior análisis.

10 Para verificar las secuencias de ácido nucleico, se secuenció nuevamente la secuencia de ácido nucleico en ambas direcciones utilizando cebadores directos para la región FXPERF (SEQ ID NO: 2268) o GRRXCP (A/G) (SEQ ID NO: 2269) del gen para p450 o cebadores inversos ya sea para el plásmido o la cola de poli A. Toda la secuenciación se realizó al menos dos veces en ambas direcciones.

15 Las secuencias de ácido nucleico de fragmentos de citocromo p450 se compararon entre sí a partir de la región de codificación correspondiente al primer ácido nucleico después de la región que codifica al motivo GRRXCP (A/G) (SEQ ID NO: 2269) hasta el codón de parada. Esta región se seleccionó como un indicador de la diversidad genética entre las proteínas p450. Se observaron un gran número de genes para p450 genéticamente distintos, más de 70 genes, similar a las otras especies de plantas. Tras la comparación de las secuencias de ácido nucleico, se encontró que los genes podrían ser colocados en grupos de secuencias distintas con base en su identidad de secuencia. Se encontró que la mejor agrupación única de los miembros de p450 se determinó que era aquellas secuencias con 20 75% de identidad de ácidos nucleicos o mayor. (Véase, por ejemplo, la Tabla 1 de la publicación de la solicitud de patente de los Estados Unidos No. 2004/0162420.

25 Se usó la hibridación del microarreglo GeneChip® (Affymetrix Inc.; Santa Clara, CA) para identificar los genes con patrones de expresión diferenciales entre las líneas isogénicas cercanas convertidoras y no convertidoras después de la activación con etileno. El tamaño del chip era de 18 micras y el formato del arreglo era de 100 - 2187, con capacidad para 528 conjuntos de sondas (11628 sondas). Se usaron siete pares de hibridación para obtener la verificación independiente de los resultados de los microarreglos. Estos consistían en un par (convertidor / no convertido) de muestras de tabaco Burley no tratadas 4407-33 / 4407-25, cuatro pares de muestras 4407-33 / 4407-25 tratadas con etileno, un par de tabaco oscuro tratado con etileno NL Madole/181, otro par de líneas isogénicas cercanas para la conversión de nicotina, y un par de hojas envejecidas en forma natural de 4407 = 33/25 (Tabla 1).

30 Tabla 1. Relaciones de señal normalizada convertidora:no convertidora de la hibridación de GeneChip®

	Burley no tratado (4407 - 33 / 25)	Burley tratado con etileno (4407 - 33 / 25)				Oscuro tratado con etileno (178 / NL Madole)	Burley envejecido (4407 - 33 / 25)
		Ej. 1	Ej. 2	Ej. 3	Ej. 4		
Inducida							
D121-AA8	1,03	2,143	12,90	5,17	12,19	16,60	2,57
D120-AH4	1,44	1,90	12,74	2,87	7,55	8,17	1,69
D35-BG1	1,73	2,32	13,06	22,22	19,10	28,76	3,40
Control							
Actina tipo I (5')	1,18	0,99	0,74	0,73	0,57	1,02	0,97
Actina tipo I (3')	1,09	1,12	0,81	1,08	0,79	0,93	0,85

Todos los 14 conjuntos de hibridaciones fueron exitosos como lo demuestra el Reporte de Expresión generado usando instrumentos de detección de Genome Explorations, Inc. (Memphis, TN). Los reportes principales incluían

análisis de ruido, factor de escala, fondo, conjuntos totales de sondas, el número y el porcentaje de los conjuntos de sondas presentes y ausentes, intensidad de la señal de los controles de mantenimiento. Los datos fueron analizados posteriormente y presentados utilizando el programa GCOS en combinación con otro programa. Se hicieron comparaciones de señal entre pares de tratamiento, y se compilaron los datos generales para todas las sondas respectivas para todas las hibridaciones y también se analizaron los datos de expresión. Los resultados basados en las intensidades de señal mostraron que sólo dos genes, D121-AA8 y D120-AH4 y un fragmento, D35-BG11, que es un fragmento parcial de D121-AA8, tenían inducción reproducible en líneas convertidoras tratadas con etileno en comparación con líneas no convertidoras. La señal de un gen en una línea convertidora, por ejemplo, la variedad de tabaco Burley 4407-33, se determinó como la relación con la señal de un gen en una línea isogénica no convertidora relacionada, 4407-25. Sin tratamiento con etileno, la relación de las señales del convertidor con respecto al no convertidor para todos los genes se aproximó a 1,00. Para eliminar la influencia de las diferencias de fondo, se calcularon también las relaciones de señal normalizadas. Las relaciones de señal normalizadas se obtuvieron dividiendo la relación emparejada tratada con la relación no emparejada correspondiente. Tras el tratamiento y análisis con etileno, se determinó que dos genes, D121-AA8 y D120-AH4, fueron inducidos en líneas convertidoras con relación a las líneas no convertidoras como se determinó mediante cuatro análisis independientes. Estos dos genes comparten 99,8% de homología relativa y sus señales de hibridación relativa en variedades convertidoras estaban en el intervalo de aproximadamente 2 a 22 veces mayor que las señales en sus contrapartes no convertidoras. Con base en los coeficientes normalizados, no se indujeron clones de control interno de tipo actina en líneas convertidoras. Además, un fragmento (D35-BG11), cuya región codificante está totalmente contenida tanto en los genes-D121 como D120-AA8 AH4, fue altamente inducido en las mismas muestras de líneas isogénicas convertidora y no convertidoras emparejadas. Además, los genes D121-AA8 y D120-AH4 fueron fuertemente inducidos en líneas convertidoras de pares de tabaco oscuro isogénicas, NL Madole y 181 (8 a 28 veces), lo que demuestra que la inducción con etileno de estos genes en líneas convertidoras fue una respuesta *in planta*. Los mismos genes fueron identificados en las comparaciones realizadas a partir de las hibridaciones de muestras envejecidas en forma natural de 4407-33 / 25 también. Los ensayos por RT-PCR de estos materiales utilizando cebadores específicos para D121-AA8 verificaron los resultados del microarreglo para este gen.

Con base en estos resultados, el gen D121-AA8 (cuya secuencia de ADNc es la secuencia de la SEQ ID NO: 5; Figura 4) fue identificado como el gen para la nicotina desmetilasa de tabaco de interés. En vista de la regla de nomenclatura de p450, se determinó que D121-AA8 es más similar a los p450 en la familia CYP82E (The Arabidopsis Genome Initiative (AGI) y The Arabidopsis Information Resource (TAIR); Frank, Plant Physiol. 110: 1035 - 1046, 1996; Whitbred et al., Plant Physiol. 124: 47 - 58, 2000); Schopfer y Ebel, Mol. Gen. Genet. 258: 315 - 322, 1998; y Takemoto et al., Plant Cell Physiol. 40: 1232 - 1242, 1999).

Ejemplo 7

Análisis bioquímico de la actividad enzimática

El análisis bioquímico, por ejemplo, como se describe en las solicitudes presentadas anteriormente, determinó que la secuencia de la SEQ ID NO: 5 codifica una nicotina desmetilasa de tabaco (SEQ ID NO: 3; Figuras 3 y 4).

En particular, la función del clon candidato D121-AA8 fue confirmado como el gen que codifica para la nicotina desmetilasa, mediante el ensayo de actividad enzimática de p450 expresada de forma heteróloga en células de levadura como sigue.

1. Construcción del vector de expresión de levadura

Las secuencias putativas codificadoras de la proteína del ADNc que codifica la nicotina desmetilasa del tabaco (D121-AA8), D120-AH4, D121-AA8, 208-AC-8, y D208-AD9, se clonaron en el vector de expresión de levadura pYeDP60. Se introdujeron los sitios apropiados de BamHI y MfeI (subrayados más abajo) a través de cebadores de PCR que contienen estas secuencias ya sea secuencia arriba del codón de inicio de la traducción (ATG) o secuencia abajo del codón de parada (TAA). El MfeI en el producto de PCR amplificado es compatible con el sitio de EcoRI en el vector. Los cebadores utilizados para amplificar el ADNc de D121-AA8 fueron 5'-TAGCTA-CGCGGATCCATGCTTTCTCCCATAGAAGCC-3' (SEQ ID NO: 2194) y 5'-CTGGATCACAAATTGTTAGTGATG-GTGATGGTGATGCGATCCTCTATAAAGCTCAGGTGCCAGGC-3' (SEQ ID NO: 2297). Un segmento de la secuencia que codifica nueve aminoácidos adicionales en el terminal C de la proteína, incluyendo seis histidinas, se incorporó en el cebador inverso para facilitar expresión de p450 etiquetado con 6-His después de la inducción. Los productos de PCR se ligaron en el vector pYeDP60 después de digestiones con enzimas en orientación sentido con referencia al promotor GAL10-CYC1. La construcción apropiada de los vectores de expresión de levadura se verificó mediante análisis con enzimas de restricción y secuenciación de ADN. Además, la expresión de las proteínas p450 se visualizó en la electroforesis en gel de SDS-PAGE para la fase detergente de los microsomas de levadura. El tamaño predicho de las proteínas p450 es de 59 kD, con base en la secuencia del gen; un resultado que fue confirmado por el análisis en gel.

2. Transformación de la levadura

La línea de levadura WAT11, modificada para expresar NADPH-citocromo p450 reductasa ATR1 de Arabidopsis, se transformó con los plásmidos de ADNc pYeDP60-p450. Se mezclaron 50 microlitros de suspensión de células de levadura WAT11 con ~ 1 µg de ADN de plásmido en una cubeta con 0,2 cm de separación de los electrodos. Se aplicó un pulso de 2,0 kV por un electroporador Eppendorf (Modelo 2510). Las células se esparcieron en placas de SGI (5 g / L de ácidos bactocasamino, 6,7 g / L de base nitrogenada de levadura sin aminoácidos, 20 g / L de glucosa, 40 mg / L de DL-triptófano, 20 g / L de agar). Los transformantes fueron confirmados mediante análisis PCR realizado directamente sobre las colonias seleccionadas al azar.

3. Expresión de p450 en células de levadura transformadas

Se utilizaron colonias individuales de levadura para inocular 30 mL de medio SGI (5 g / L de ácidos bactocasamino, 6,7 g / L de base nitrogenada de levadura sin aminoácidos, 20 g / L de glucosa, 40 mg / L de DL-triptófano) y se cultivaron a 30 °C durante aproximadamente 24 horas. Una alícuota de este cultivo se diluyó en una proporción 1:50 en 1000 ml de medio YPGE (10 g / L de extracto de levadura, 20 g / L de Bacto peptona, 5 g / L de glucosa, 30 ml / L de etanol) y se cultivó hasta que se consumió completamente la glucosa, como se indica por el cambio colorimétrico de una tira reactiva para análisis de orina Diastix (Bayer, Elkhart, IN). La inducción de p450 clonado se inició mediante la adición de DL-galactosa hasta una concentración final del 2%. Los cultivos se desarrollaron durante 20 horas adicionales antes de utilizarlos para el ensayo de actividad *in vivo* o para la preparación de microsomas.

Las células de levadura WAT11 que expresan pYeDP60-CYP71D20 (un p450 que cataliza la hidroxilación de 5-epi-aristolóqueno y 1-deoxicapsidiol en *Nicotiana tabacum*) se utilizaron como control para los ensayos de expresión de p450 y de la actividad enzimática.

Para evaluar la efectividad de la expresión en la levadura del p450 con gran detalle, se realizó la espectroscopía de diferencia de CO reducido. El espectro de CO reducido exhibió un pico a 450 nm de proteínas de todas las cuatro líneas de levadura transformadas con p450. No se observaron picos similares en microsomas de control derivados de células de levadura no transformadas de control o el blanco, células de levadura de control del vector. Los resultados indicaron que las proteínas p450 se expresaron de manera efectiva en líneas de levadura que albergan las pYeDP60-CYP 450. Las concentraciones de proteína p450 expresadas en microsomas de levadura oscilaron desde 45 a 68 nmol / mg de proteína total.

4. Ensayo enzimático *in vivo*

La actividad de la nicotina desmetilasa en las células de levadura transformadas se ensayaron mediante la alimentación del cultivo de levadura con DL-Nicotina (pirrolidina-2-¹⁴C). Se añadió nicotina marcada con ¹⁴C (54 mCi / mmol) a 75 µl del cultivo inducido con galactosa para una concentración final de 55 µM. Se incubó el cultivo de ensayo con agitación en tubos de polipropileno de 14 ml durante 6 horas y se extrajo con 900 µl de metanol. Después de la centrifugación, se separaron 20 µl del extracto de metanol con rp-HPLC y se cuantificó la fracción de nornicotina por LSC.

El cultivo de control de WAT11 (pYeDP60-CYP71D20) no convirtió nicotina para nornicotina, mostrando que la cepa de levadura WAT11 no contiene actividad de enzima endógena que pueda catalizar la etapa de bioconversión de nicotina en nornicotina. En contraste, la levadura que expresa el gen para la nicotina desmetilasa de tabaco produjo una cantidad detectable de nornicotina, lo que indica la actividad de la nicotina desmetilasa del producto de traducción de la SEQ ID NO: 4 o SEQ ID NO: 5.

5. Preparación del microsoma de levadura

Después de la inducción por galactosa durante 20 horas, se recolectaron las células de levadura por centrifugación y se lavaron dos veces con regulador TES-M (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, sorbitol 0,6 M, 2-mercaptoetanol 10 mM). Se resuspendió el sedimento en regulador de extracción (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, sorbitol 0,6 M, 2-mercaptoetanol 2 mM, albúmina de suero bovino al 1%, cóctel inhibidor de proteasa (Roche) a razón de 1 tableta / 50 ml). Se rompieron luego las células con perlas de vidrio (0,5 mm de diámetro, Sigma) y se centrifugó el extracto celular durante 20 min a 20.000 x g para eliminar los restos celulares. Se sometió el sobrenadante a ultracentrifugación a 100.000 x g durante 60 min y el sedimento resultante contenía la fracción microsomal. Se suspendió la fracción microsomal en regulador TEG-M (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, 20% de glicerol y 2-mercaptoetanol 1,5 mM) a una concentración de proteína de 1 mg / mL. Se almacenaron las preparaciones microsomales en un congelador de nitrógeno líquido hasta su uso.

6. Ensayo de actividad enzimática en preparaciones microsomales de levadura

Se realizaron ensayos de actividad de nicotina desmetilasa con preparaciones microsomales de levadura. En particular, se obtuvo DL-Nicotina (Pirrolidina-2-¹⁴C) a través de Moravek Bioquímicos y tenía una actividad específica de 54 mCi / mmol. La clorpromazina (CPZ) y el citocromo c oxidado (cyt. C), ambos inhibidores de p450, se adquirieron de Sigma. La forma reducida de la nicotinamida adenina dinucleótido fosfato (NADPH) es el donador de electrones típico para el citocromo p450 a través de la NADPH: citocromo p450 reductasa. La NADPH se omitió para la incubación de control. El ensayo enzimático de rutina incluyó proteínas microsomales (alrededor de 1 mg / ml), NADPH 6 mM, y 55 mM de nicotina marcada con ¹⁴C. La concentración de CPZ y Cyt. C, cuando se utilizó, fue de 1 mM y 100 μM, respectivamente. La reacción se llevó a cabo a 25 °C durante 1 hora y se detuvo con la adición de 300 μl de metanol por cada 25 μl de mezcla de reacción. Después de centrifugación, se separaron 20 μl del extracto de metanol con un sistema de cromatografía líquida de alta resolución (HPLC) de fase reversa (Agilent) utilizando una columna de cromatografía Inertsil ODS-3 3 μm (150 x 4,6 mm) de Varian. La fase móvil isocrática era la mezcla de metanol y regulador de fosfato de potasio 50 mM, pH 6,25, con una relación de 60:40 (v / v) y la velocidad de flujo era de 1 ml / min. El pico de nornicotina, tal como se determinó por comparación con la auténtica nornicotina no marcada, se recogió y se lo sometió a un Contador de Centelleo Líquido 2.900 tri-carb (LSC) (Perkin Elmer) para la cuantificación. Se calcula la actividad de la nicotina desmetilasa con base en la producción de nornicotina marcada con ¹⁴C después de más de 1 hora de incubación.

Se observó actividad como la de p450 en preparaciones microsomales de células de levadura de control que expresan CYP71D20 y los tres cultivos de levadura con p450 de pruebas transformados con los genes D120-AH4, D208-AC8, y D208-AD9. Sin embargo, el control y los tres p450 de prueba no mostraron ninguna formación de conversión de nornicotina lo que sugiere que no contenían una enzima endógena o inducida que pueda catalizar la desmetilación de nicotina. En contraste, los resultados de los análisis por HPLC y LSC mostraron cantidades detectables de nornicotina producida a partir de la desmetilación de nicotina usando muestras microsomales obtenidas a partir de células de levadura que expresan el gen para la nicotina desmetilasa de tabaco (D121-AA8). Estos resultados indican que la actividad de la nicotina desmetilasa resulta del producto del gen D121-AA8. La actividad de nicotina desmetilasa requiere NADPH y se ha demostrado que es inhibida por inhibidores específicos de p450, de conformidad con la nicotina desmetilasa de tabaco siendo un p450. La actividad enzimática de la nicotina desmetilasa de tabaco (D121-AA8) fue de aproximadamente 10,8 pKat / mg de proteína como se calculó por la intensidad de la radiación y las concentraciones de proteína. Un conjunto típico de resultados de ensayo de la enzima obtenidos para las células de levadura se muestra en la siguiente tabla (Tabla 2).

Tabla 2. Actividad de la desmetilasa en microsomas de células de levadura que expresan genes para p450 D121-AA8 y de control

Muestra	Microsomas	Microsomas + clorpromazina 1 mM	Microsomas + citocromo C 100 μM	Microsomas - NADPH
D121-AA8	10,8 ± 1,2* pkat/mg de proteína	1,4 ± 1,3 pkat/mg de proteína	2,4 ± 0,7 pkat/mg de proteína	0,4 ± 0,1 pkat/mg de proteína
Control (CYP71D20)	No detectado	No detectado	No detectado	No detectado
*n = 12, otros n = 3				

La omisión de NADPH del ensayo utilizando microsomas derivados de células de levadura D121-AA8 resultó en la abolición de la actividad de nicotina desmetilasa; por lo tanto, no se formó nornicotina (Tabla 2). Cuando se añadieron dos inhibidores de p450 conocidos, clorpromazina (CPZ, 1 mM) y citocromo c oxidado (cyt C, 100 μM), en las mezclas de ensayo enzimático en forma separada y se incubaron durante 1 hora antes de añadir la solución de detención de metanol, disminuyeron significativamente las actividades de la nicotina desmetilasa (Tabla 2). Juntos, estos experimentos demostraron que D121-AA8 codifica una proteína citocromo p450 que cataliza la conversión de nicotina en nornicotina cuando se expresa en levadura.

40 Ejemplo 8

Identidad de la secuencia de aminoácidos relacionada de clones de longitud completa

La secuencia de ácido nucleico de genes de Nicotiana de longitud completa clonados en el Ejemplo 5 se dedujo por su secuencia completa de aminoácidos. Los genes para el citocromo p450 fueron identificados por la presencia de tres motivos conservados del dominio de p450, que correspondían a UXXRXXZ (SEQ ID NO: 2274), PXRFXF (SEQ ID NO: 2275) o GXRXC (SEQ ID NO: 2276) en el terminal carboxilo donde U es E o K, X es cualquier aminoácido y

Z es P, T, S o M. Todos los genes para p450 se caracterizaron para la identidad de los aminoácidos usando un programa BLAST que compara sus secuencias de longitud completa entre sí y con genes de tabaco conocidos. El programa utiliza la herramienta BLAST especial del NCBI (Alineación de dos secuencias (b12seq), <http://www.ncbi.nlm.nih.gov/blast/b12seq/b12.html>). Se alinearon dos secuencias bajo BLASTN sin filtro para secuencias de ácido nucleico y BLASTP para las secuencias de aminoácidos. Con base en su identidad porcentual de aminoácidos, se agrupó cada secuencia en grupos de identidad, donde la agrupación contenía a los miembros que comparten al menos un 85% de identidad con otro miembro. Utilizando estos criterios, se identificaron grupos únicos y se describen en la Tabla 3. Se dedujo que la secuencia de aminoácidos del gen para la nicotina desmetilasa de longitud completa tiene la secuencia proporcionada en la SEQ ID NO: 5 (Figura 4).

10 Tabla 3: Grupos de identidad de secuencia de aminoácidos de genes para p450 de nicotiana de longitud completa

1 D208-AD9 (SEQ ID NO: 233); D120-AH4 (SEQ ID NO: 189); D121-AA8 (SEQ ID NO: 191), D122-AF10 (SEQ ID NO: 193); D103-AH3 (SEQ ID NO: 231); D208-AC8 (SEQ ID NO: 227); D235-AB1 (SEQ ID NO: 255)

15 Los genes de longitud completa se agruparon además con base en la homología de aminoácidos altamente conservada entre el dominio de p450 UXXRXXZ (SEQ ID NO: 2274) y el dominio de p450 GXRXC (SEQ ID NO: 2276) cerca del extremo del terminal carboxilo. Los clones individuales se alinearon con base en la homología de secuencia entre los dominios conservados y se colocaron en grupos de identidad distintos. En varios casos, aunque la secuencia de ácido nucleico del clon era única, la secuencia de aminoácidos para la región era idéntica. La agrupación final fue similar a aquella basada en el porcentaje de identidad para la secuencia de aminoácidos completa de los clones.

20 Tabla 4: Grupos de identidad de secuencia de aminoácidos de regiones entre dominios conservados de genes para p450 de Nicotiana

1 D208-AD9 (SEQ ID NO: 233); D120-AH4 (SEQ ID NO: 189); D121-AA8 (SEQ ID NO: 191), D122-AF10 (SEQ ID NO: 193); D103-AH3 (SEQ ID NO: 231); D208-AC8 (SEQ ID NO: 227); D235-AB1 (SEQ ID NO: 255)

Ejemplo 9

25 Uso de fragmentos y clones del citocromo p450 de nicotiana en la regulación alterada de las cualidades del tabaco

El uso de fragmentos de ácido nucleico de p450 de tabaco o de genes completos es útil en la identificación y selección de aquellas plantas que tienen fenotipos alterados de tabaco o constituyentes del tabaco y, más importante aún, metabolitos alterados. Las plantas de tabaco transgénico son generadas por una variedad de sistemas de transformación que incorporan fragmentos de ácido nucleico o genes de longitud completa, seleccionados de los reportados en este documento, en las orientaciones, ya sea para la subregulación, por ejemplo orientación antisentido, o sobreexpresión por ejemplo, orientación sentido y similares. Para la sobreexpresión con genes de longitud completa, es deseable cualquier secuencia de ácido nucleico que codifica la totalidad o una parte funcional o secuencia de aminoácidos de los genes de longitud completa descritos en esta invención. Tales secuencias de ácido nucleico deseablemente son eficaces para aumentar la expresión de una cierta enzima y resultando por lo tanto en un efecto fenotípico dentro de Nicotiana. Las líneas de Nicotiana que son homocigotas se obtienen a través de una serie de retrocruzamientos y se evalúan los cambios fenotípicos incluyendo, pero sin limitarse a, el análisis de ARN endógeno para p450, transcriptos, péptidos p50 expresados y concentraciones de metabolitos vegetales utilizando técnicas comúnmente disponibles para una persona normalmente capacitada en la técnica. Los cambios exhibidos en las plantas de tabaco proporcionan información sobre el papel funcional del gen de interés seleccionado o son de uso como especies de plantas preferidas de Nicotiana.

Ejemplo 10

Clonación de la nicotina desmetilasa genómica de tabaco Burley convertidor

45 Se extrajo ADN genómico de la planta de tabaco Burley convertidora línea 4407-33 (una variedad de *Nicotiana tabacum* línea 4407) usando el kit de Qiagen Plant Easy como se describe en los Ejemplos anteriores (véase también el procedimiento del fabricante).

50 Los cebadores se diseñaron con base en el promotor 5' y la región 3' UTR clonados en los ejemplos anteriores. Los cebadores directos fueron 5'-GGC TCT AGA TAA ATC TCT TAA GTT ACT AGG TTC TAA-3' (SEQ ID NO: 2280) y 5'-TCT CTA AAG TCC CCT TCC-3' (SEQ ID NO: 2288) y los cebadores inversos fueron 5'-GGC TCT AGA AGT CAA TTA TCT TCT ACA AAC CTT TAT ATA TTA GC-3' (SEQ ID NO: 2281), y 5'-CCA GCA TTC CTC AAT TTC-3' (SEQ ID NO: 2289). Se aplicó la PCR al ADN genómico 4407-33 con 100 µl de mezcla de reacción. Se utilizó enzima de alta fidelidad Pfx para la amplificación por PCR. El producto de la PCR fue visualizado en gel de agarosa

al 1% después de la electroforesis. Se observó una única banda con un peso molecular de aproximadamente 3,5 kb y se la cortó del gel. La banda resultante se purificó usando un kit de purificación de gel (Qiagen; con base en el procedimiento del fabricante). El ADN purificado fue digerido por la enzima Xba I (NEB; que se utiliza de acuerdo con las instrucciones del fabricante). El plásmido pBluescript se digirió con Xba I usando el mismo procedimiento. El fragmento se purificó en gel y se ligó al plásmido pBluescript. La mezcla de ligación fue transformada en la célula competente GM109 y se sembraron sobre la placa LB que contenía 100 mg / l de ampicilina con selección azul / blanco. Se recolectaron las colonias blancas y se cultivaron en 10 ml de medio líquido que contenía ampicilina. Se extrajo el ADN mediante miniprep. El ADN del plásmido que contenía el inserto fue secuenciado usando un secuenciador CEQ 2000 (Beckman, Fullerton, California) con base en el procedimiento del fabricante. Se usaron los cebadores T3 y T7 y otros 8 cebadores internos para la secuenciación. La secuencia se ensambló y analizó, proporcionando así la secuencia genómica (SEQ ID NO: 4; Figura 3). Con base en la secuencia genómica, se determinó que el gen para la nicotina desmetilasa tanto en las líneas de tabaco convertidoras como no convertidoras no contiene un elemento transponible.

La comparación de la secuencia de la SEQ ID NO: 5 con la secuencia de la SEQ ID NO: 4 permitió la determinación de un solo intrón dentro de la porción de codificación del gen (identificado como la secuencia de la SEQ ID NO: 7; Figura 5). Como se muestra en la Figura 2, la estructura genómica de la nicotina desmetilasa de tabaco incluye dos exones que flanquean un solo intrón. El primer exón abarca los nucleótidos 2010 a 2949 de la SEQ ID NO: 4, que codifica los aminoácidos 1 a 313 de la SEQ ID NO: 3, y el segundo exón abarca los nucleótidos 3947 hasta 4562 de la SEQ ID NO: 4, que codifica los aminoácidos 314 a 517 de la SEQ ID NO: 3. Por consiguiente, el intrón abarca los nucleótidos 2950-3946 de la SEQ ID NO: 4. La secuencia de intrón se muestra en la Figura 5 y es aquella de la SEQ ID NO: 7. El producto de traducción de la secuencia de ADN genómico se muestra en la Figura 3 como la secuencia de la SEQ ID NO: 3. La secuencia de aminoácidos de la nicotina desmetilasa de tabaco contiene un motivo de anclaje a la membrana del retículo endoplasmático.

Ejemplo 11

25 Clonación de las secuencias que flanquean a 5' (SEQ ID NO: 8) y 3'UTR (SEQ ID NO: 9) del tabaco convertidor

A. Aislamiento de ADN total del tejido de las hojas de tabaco convertidor

Se aisló el ADN genómico de las hojas del tabaco convertidor 4407-33. Se realizó el aislamiento del ADN utilizando un kit DNeasy Plant Mini de la compañía Qiagen, Inc. (Valencia, CA) de acuerdo con el protocolo del fabricante (Dneasy' Plant Mini and DNeasy Plant Maxi Handbook, Qiagen, enero de 2004). El procedimiento para la preparación del ADN incluye las siguientes etapas: se trituró tejido de la hoja de tabaco (aproximadamente 20 mg en peso seco) hasta un polvo fino bajo nitrógeno líquido durante 1 minuto. El tejido en polvo fue transferido a un tubo de 1,5 ml. Se añadieron regulador AP1 (400 µl) y 4 µl de solución patrón de RNasa (100 mg / ml) hasta un máximo de 100 mg de tejido de hoja molido y se agitó vigorosamente con agitación tipo vórtice. La mezcla se incubó durante 10 min a 65 °C y se mezcló 2-3 veces durante la incubación por inversión del tubo. Se añadió luego regulador AP2 (130 µl) al lisado. Se revolvió la mezcla y se incubó durante 5 min sobre hielo. Se aplicó el lisado a una columna QIAshredder Mini Spin y se centrifugó durante 2 min (14.000 rpm). Se transfirió la fracción que fluyó a través de la columna a un nuevo tubo sin alterar el sedimento de residuos celulares. Se añadió luego regulador AP3/E (1,5 volúmenes) al lisado aclarado y se mezcló mediante pipeteo. Se aplicó la mezcla (650 µl) de la etapa anterior incluyendo cualquier precipitado a una columna DNeasy Mini Spin. Se centrifugó la mezcla durante 1 min a > 6000 x g (> 8000 rpm) y se descartó el flujo a través de la columna. Este se repitió con el resto de la muestra y se descartaron el flujo a través de la columna y el tubo de recolección. Se colocó la columna DNeasy Mini Spin en un nuevo tubo de recolección de 2 ml. Luego se añadió regulador AW (500 µl) a la columna de DNeasy y se centrifugó durante 1 min (> 8000 rpm). Se descartó el flujo a través de la columna. Se reutilizó el tubo de recolección en la siguiente etapa. Se añadió luego regulador AW (500 µl) a la columna DNeasy y se centrifugó durante 2 min (> 14.000 rpm) con el fin de secar la membrana. La columna DNeasy se transfirió a un tubo de 1,5 ml. Entonces se pipeteó el regulador AE (100 µl) sobre la membrana DNeasy. Se incubó la mezcla durante 5 min a temperatura ambiente (15-25 °C) y luego se centrifugó durante 1 min (> 8000 rpm) para eluir.

La calidad y cantidad del ADN se estimó mediante corrimiento de las muestras en un gel de agarosa.

B. Clonación de secuencias de flanco 5' del gen estructural

50 Se utilizó un método PCR inverso modificado para clonar 750 nucleótidos de las secuencias de flanco 5' del gen estructural de la SEQ ID NO: 5. En primer lugar, se seleccionaron las enzimas de restricción apropiadas con base en el sitio de restricción en el fragmento de secuencia conocido y la distancia de los sitios de restricción secuencia abajo de las secuencias de flanco 5'. Se diseñaron dos cebadores con base en este fragmento conocido. El cebador directo se ubicó secuencia abajo del cebador inverso. El cebador inverso se ubicó en la porción 3' del fragmento conocido.

El procedimiento de clonación incluía las siguientes etapas:

- 5 Se digirió el ADN genómico purificado (5 µg) con 20-40 unidades de la enzima de restricción apropiada (EcoRI y SpeI) en una mezcla de reacción de 50 µl. Se realizó una electroforesis en gel de agarosa con 1/10 del volumen de la mezcla de reacción para determinar si se digirió completamente el ADN. Se realizó una ligadura directa después de la digestión completa por ligación durante la noche a 4 °C. Una mezcla de reacción de 200 µl que contenía 10 µl de ADN digerido y 0,2 µl de ADN ligasa T4 (NEB) se ligó durante la noche a 4 °C. Se realizó una PCR en la reacción de ligación después de obtener un pequeño genoma circular artificial. La PCR se realizó con 10 µl de la reacción de ligación y 2 cebadores a partir de fragmentos conocidos en dos diferentes direcciones en 50 µl de mezcla de reacción. Se aplicó un programa de PCR en gradiente con temperaturas de hibridación de 45 - 56 °C.
- 10 Se realizó una electroforesis en gel de agarosa para controlar la reacción PCR. Se cortó la banda deseada del gel y se usó un kit de purificación del gel QIAquick de QIAGEN para purificar la banda. Los fragmentos de la PCR purificados se ligaron en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se extrajeron los plásmidos de ADN transformados por miniprep usando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. El ADN del plásmido que contenía el inserto fue secuenciado usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Aproximadamente 758 nt (nucleótidos 1241 - 2009 de la SEQ ID NO: 4) de la secuencia de flanqueo 5' se clonaron por el método descrito anteriormente.
- 15

C. Clonación de las secuencias de flanqueo 5' más largas (SEQ ID NO: 8; Figura 6) del gen estructural

- 20 Se usó el kit BD GenomeWalker Universal (Clontech Laboratories, Inc., Palo Alto, CA) para la clonación de una secuencia de flanqueo 5' adicional del gen estructural, D121-AA8 de acuerdo con el manual de usuario del fabricante (BD GenomeWalker de agosto de 2004). El tamaño y la pureza del ADN genómico de tabaco se analizaron corriendo las muestras en un gel de agarosa al 0,5%. Se prepararon un total de 4 reacciones de extremo romo (DRA I, STU I, ECOR V, PVU II) para una construcción de desplazamiento sobre el genoma de 33 bibliotecas del tabaco. Después de la purificación de los ADN digeridos, se ligaron los ADN genómicos digeridos al adaptador de desplazamiento sobre el genoma. Se aplicaron las reacciones PCR primarias a los cuatro ADN digeridos mediante el uso de cebador adaptador AP1 y el cebador específico del gen de D121-AA8 (CTCTATTGATACTAGCTGGTTTTGGAC; SEQ ID NO: 2282). Se utilizaron los productos primarios de la PCR directamente como plantillas para la PCR anidada. El cebador adaptador anidado proporcionado por el kit y el cebador anidado del clon conocido D121-AA8 (SEQ ID NO: 5) (GGAGGGAGAGTATAACTTACGGATTC; SEQ ID NO: 2283) se utilizaron en la reacción PCR. Los productos de la PCR se controlaron corriendo la electroforesis en gel. Se cortaron las bandas deseadas del gel, y se purificaron los fragmentos de PCR usando el kit de purificación en gel QIAquick de QIAGEN. Los fragmentos de la PCR se ligaron en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Los plásmidos de ADN transformados se extrajeron mediante miniprep utilizando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. El ADN del plásmido que contenía el inserto fue secuenciado usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA).
- 25
- 30 Otros aproximadamente 853 nt de la secuencia de flanqueo 5', incluyendo los nucleótidos 399 - 1240 de la SEQ ID NO: 4, se clonaron mediante el método descrito anteriormente.

- 40 Una segunda ronda de desplazamiento sobre el genoma se realizó de acuerdo con el mismo método con la diferencia de que se usaron los siguientes cebadores GWR1A (5'-AGTAACCGATTGCTCACGTTATCCTC-3') (SEQ ID NO: 2284) y GWR2A (5'-CTCTATTCAACCCACACGTAAGT-3') (SEQ ID NO: 2285). Otros aproximadamente 398 nt de la secuencia de flanqueo, incluyendo los nucleótidos 1-398 de la SEQ ID NO: 4, se clonaron mediante este método.

- 45 Una búsqueda de elementos reguladores reveló que, además de la caja "TATA", cajas "CAAT" y cajas "GAGA", varios sitios de reconocimiento similares a MYB y elementos de especificidad de órganos están presentes en la región promotora de nicotina desmetilasa de tabaco. Elementos sensibles inductores putativos y elementos regulados por nitrógeno, identificados utilizando métodos convencionales, también están presentes en la región promotora.

D. Clonación de secuencias de flanqueo 3' del gen estructural

- 50 Se usó el kit BD GenomeWalker Universal (Clontech Laboratories, Inc., Palo Alto, CA) para la clonación de una secuencia de flanqueo 3' del gen estructural, D121-AA8 de acuerdo al manual de usuario del fabricante. El procedimiento de clonación es el mismo que se describe en la Sección C anterior de este ejemplo, excepto por los cebadores específicos del gen. El primer cebador se diseñó desde cerca del extremo del gen estructural-D121 AA8 (5'-CTA TCT GGT AAC CTG ATC CTG ATA CTT-3') (SEQ ID NO: 2286). El cebador anidado fue diseñado secuencia más abajo del cebador 1 del gen estructural-D121 AA8 (CTA TAC GTA AGG TAA ATC CTG TGG AAC) (SEQ ID NO: 2287). Los productos finales de la PCR fueron controlados por electroforesis en gel. Se cortaron las bandas deseadas del gel. Se purificaron los fragmentos de PCR usando el kit de purificación en gel QIAquick de QIAGEN. Los fragmentos de la PCR se ligaron en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las
- 55

instrucciones del fabricante. Los plásmidos de ADN transformados se extrajeron mediante miniprep utilizando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. El ADN del plásmido que contenía el inserto fue secuenciado usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Aproximadamente 1617 nucleótidos de la secuencia de flanqueo 3' adicional (nucleótidos 4731 - 6347 de la SEQ ID NO: 4) se clonaron mediante el método descrito anteriormente. La secuencia de ácido nucleico de la región 3'UTR se muestra en la Figura 7.

Ejemplo 12

Cribado del genero *Nicotiana* por la presencia o ausencia de un gen para la nicotina desmetilasa

Cuarenta y tres especies de *Nicotiana*, cuarenta y nueve líneas de *Nicotiana rustica*, y aproximadamente seiscientas líneas de *Nicotiana tabacum* se sembraron en macetas y las plantas resultantes se cultivaron en el invernadero.

Se tomaron muestras de hojas de plantas de seis semanas de edad. Se llevaron a cabo extracciones de ADN de las hojas utilizando el mini kit DNeasy Plant (Qiagen, Inc., Valencia, CA) de acuerdo con el protocolo del fabricante.

Los cebadores fueron diseñados con base en el promotor 5' y las regiones 3' UTR descritos en la presente invención. El cebador directo fue 5'-GGC TCT AGA TAA ATC TCT TAA GTT ACT AGG TTC TAA-3' (SEQ ID NO: 2290) y el cebador inverso fue 5'-GGC TCT AGA AGT CAA TTA TCT TCT ACA AAC CTT TAT ATA TTA GC-3' (SEQ ID NO: 2291) (desde -750 de la región de flanqueo 5' hasta 180 nt de 3' UTR). Se extrajo ADN genómico de todas las líneas de *Nicotiana* anteriormente mencionadas para el análisis de PCR. Se usaron una mezcla de reacción de 100 µl y la enzima de alta fidelidad Pfx para la amplificación por PCR. La temperatura de hibridación utilizada fue de 54 °C debido a la menor homología entre las especies (esta temperatura es 2 °C más baja que la temperatura utilizada para la secuencia genómica de clonación de tabaco convertidor 4407 como se describió anteriormente). El producto de la PCR fue visualizado en gel de agarosa al 0,8% después de la electroforesis. Una única banda con un peso molecular de aproximadamente 3,5 kb estaba o bien presente o ausente sobre el gel. Las líneas con una banda positiva se calificaron como aquellas que tienen el gen objetivo. Para las líneas que carecían de bandas positivas, se realizaron cuatro reacciones PCR adicionales usando cuatro conjuntos más de cebadores. Estos conjuntos de cebadores fueron seleccionados de diferentes regiones del gen. Los cuatro conjuntos de cebadores fueron:

(1) desde el codón de inicio (5'-GCC CAT CCT ACA GTT ACC TAT AAA AAG GAA G -3') (SEQ ID NO: 2292) hasta el codón de parada (5'-ACC ATG AAG AAA GAT CTT AGG TTT TAA -3') (SEQ ID NO: 2293),

(2) desde 570 nt secuencia abajo del codón de inicio (5'-CTG AAG ATC GTG ATG A -3') (SEQ ID NO: 2294) hasta el final del intrón (5'-TGC TGC ATC CAA GAC CA -3') (SEQ ID NO: 2295),

(3) desde 300 nt secuencia abajo del comienzo del intrón (5'-GGG CTA TAT GGA TTC GC-3') (SEQ ID NO: 2296) hasta el final del intrón (5'-TGC TGC ATC GAC CAA CA -3 ') (SEQ ID NO: 2295), y

(4) desde 300 nt secuencia abajo del comienzo del intrón (5'-GGG CTA TAT GGA TTC GC-3') (SEQ ID NO: 2296) hasta la 3' UTR (5'-AGT CAA TTA TCT TCT ACA AAC CTT TAT ATA TTA GC -3') (SEQ ID NO: 2195).

Si las cinco reacciones PCR antes mencionadas, mostraron todas bandas no correctas, la línea se calificó como carente del gen objetivo. Los ejemplos de la cantidad de ADN genómico y los productos de PCR para el gen objetivo para la nicotina desmetilasa son representados en las Figuras 8 y 9.

El germoplasma identificado como carente del gen para la nicotina desmetilasa se utiliza como el material fuente para fitomejoramiento con tabacos cultivados. Sin embargo, cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10-17, o un fragmento de la misma, se puede utilizar de una manera similar. Métodos de hibridación interespecífica o intraespecífica combinados con métodos convencionales de fitomejoramiento, tales como el retrocruzamiento o el método de pedigrí, se pueden utilizar para transferir el gen para la nicotina desmetilasa aberrante o ausente o cualquier secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, desde la fuente donadora hasta los tabacos cultivados. Los resultados de los experimentos de cribado para la nicotina desmetilasa se exponen en la Tabla 5 a continuación. Una línea negativa para nicotina desmetilasa pueden ser reproducida con la misma línea o con otra línea negativa (por ejemplo, *Nicotiana africana* x *Nicotiana africana* o *Nicotiana africana* x *Nicotiana amplexicaulis* o cualquier combinación de fitomejoramiento adecuada). También se reproducen líneas negativas con cualquier variedad comercial de tabaco de acuerdo con técnicas convencionales de fitomejoramiento de tabaco conocidos en la técnica. Las líneas de tabaco pueden ser reproducidas con cualquier otra planta compatible de acuerdo con procedimientos convencionales en la técnica.

Tabla 5. Resultados de ejemplos de cribado del género *Nicotiana* para el gen para la nicotina desmetilasa

Nombre científico o nombre común u (origen)	Número de inventario	Resultados del cribado
<i>Nicotiana africana</i>	TW6	Negativo
<i>Nicotiana amplexicaulis</i>	TW10	Negativo
<i>Nicotiana arentsii</i>	TW12	Negativo
<i>Nicotiana benthamiana</i>	TW16	Negativo
<i>Nicotiana bigelovii</i>	TW18	Negativo
<i>Nicotiana corymbosa</i>	TW35	Negativo
<i>Nicotiana debneyi</i>	TW36	Negativo
<i>Nicotiana excelsior</i>	TW46	Negativo
<i>Nicotiana exigua</i>	TW48	Negativo
<i>Nicotiana glutinosa</i>	TW58	Negativo
<i>Nicotiana goodspeedii</i>	TW67	Negativo
<i>Nicotiana gossei</i>	TW68	Negativo
<i>Nicotiana hesperis</i>	TW69	Negativo
<i>Nicotiana ingulba</i>	TW71	Negativo
<i>Nicotiana knightiana</i>	TW73	Negativo
<i>Nicotiana maritima</i>	TW82	Negativo
<i>Nicotiana megalosiphon</i>	TW83	Negativo
<i>Nicotiana miersii</i>	TW85	Negativo
<i>Nicotiana nesophila</i>	TW87	Negativo
<i>Nicotiana noctiflora</i>	TW88	Negativo
<i>Nicotiana nudicaulis</i>	TW90	Negativo
<i>Nicotiana otophora</i>	TW94	Positivo
<i>Nicotiana palmeri</i>	TW98	Negativo
<i>Nicotiana paniculata</i>	TW99	Negativo
<i>Nicotiana petunioides</i>	TW105	Negativo
<i>Nicotiana plumbaginifolia</i>	TW106	Negativo
<i>Nicotiana repanda</i>	TW110	Negativo

ES 2 535 643 T3

Nombre científico o nombre común u (origen)	Número de inventario	Resultados del cribado
<i>Nicotiana rosulata</i>	TW112	Negativo
<i>Nicotiana rotundifolia</i>	TW114	Negativo
<i>Nicotiana rustica</i>	TW116	Negativo
<i>Nicotiana setchellii</i>	TW121	Negativo
<i>Nicotiana stocktonii</i>	TW126	Negativo
<i>Nicotiana eastii</i>	TW127	Negativo
<i>Nicotiana suaveolens</i>	TW128	Negativo
<i>Nicotiana thrysiflora</i>	TW139	Positivo
<i>Nicotiana tomentosa</i>	TW140	Positivo
<i>Nicotiana tomentosiformis</i>	TW142	Positivo
<i>Nicotiana trigonophylla</i>	TW143	Negativo
NL Madole	Semilla base	Positivo
KY 14	Semilla base	Positivo
TN 86	Semilla base	Positivo
Coker 176	Semilla base	Positivo
KY21	TC62	Positivo
KY22	TC63	Positivo
KY24	TC64	Positivo
KY26	TC65	Positivo
KY33	TC66	Positivo
KY34	TC67	Positivo
KY35	TC68	Positivo
KY41A	TC69	Positivo
KY54	TC71	Positivo
KY52	TC70	Positivo
Virginia 528	TC85	Positivo
Virginia B-29	TC86	Positivo
401 Rojo Cereza	TC227	Positivo

ES 2 535 643 T3

Nombre científico o nombre común u (origen)	Número de inventario	Resultados del cribado
401 Rojo Cereza Libre	TC228	Positivo
KY170	TC474	Positivo
KY171	TC475	Positivo
Maryland 609	TC505	Positivo
Maryland Mamut	TC507	Positivo
VA403	TC580	Positivo
KY908	TC630	Positivo
Earl Jennett Madole	TC642	Positivo
Kavala	TC533	Positivo
Kavala No 15A	TC534	Positivo
GR 10	TC 19	Positivo
GR 10A	TC20	Positivo
GR 24	TC27	Positivo
NOD 9	TI 1745	Positivo
NOD 12	TI 1747	Positivo
NOD 17	TI 1749	Positivo
80111 Pudawski 66CMS	TI 1661	Positivo
84160 Pudawski 66	TI 1683	Positivo
MII 109	TI 1715	Positivo
Mississippi Reliquia de Familia	TI 1716	Positivo
Ovens 62	TI 1741	Positivo
BT 101	TI 1594	Positivo
Kentucky MI 429	TI 1595	Positivo
Shiroenshu 201	TI 1604	Positivo
Shiroenshu 202	TI 1605	Positivo
Ostrolist 2747 II	TI 1568	Positivo
Ergo	TI 1349	Positivo
Burley 323	TI 1535	Positivo

ES 2 535 643 T3

Nombre científico o nombre común u (origen)	Número de inventario	Resultados del cribado
Burley Rusa	TI 1534	Positivo
Puremozhetz 83	TI 1569	Positivo
Bulsunov 80	TI 1537	Positivo
Amarillo Riogrande	TI74	Positivo
Espado	TI151	Positivo
Crillo Saltono	TI1082	Positivo
Kutsaga E-1	TI1552	Positivo
Beinhart 1000-1	TI1561	Positivo
Kelly Brownleaf	TC50	Positivo
KY9	TC54	Positivo
Negro Mamut	TC460	Positivo
Cola de Lagarto Orinoco	TC477	Positivo
Bel MS-2	TC493	Positivo
Maryland 201	TC503	Positivo
Perique	TC556	Positivo
NC-BMR 90	TC571	Positivo
LN KY 171	TC605	Positivo
Samsun	TC536	Positivo
Xanthi-Parental	TC554	Positivo
(Turquía)	TI 1222	Positivo
Hongrois (España)	TI 1246	Positivo
(Etiopia)	TI 1269	Positivo
Ravajk(Yugoslavia)	TI 1284	Positivo
(Bolivia)	TI 1301	Positivo
Adjectifolia (Nueva Zelanda)	TI 1317	Positivo
NO. 6055 (Cuba)	TI 1375	Positivo
(Bulgaria)	TI 1386	Positivo
Grande Reditto (Italia)	TI 1414	Positivo

Nombre científico o nombre común u (origen)	Número de inventario	Resultados del cribado
(Alemania)	TI 1459	Positivo
(Suiza)	TI 1506	Positivo
Sirone (Australia)	TI 1508	Positivo
Dubek 566 (Polonia)	TI 1567	Positivo
Kagoshima Maruba (Japón)	TI 158	Positivo
Erzegovina Lecce MI 411 (Italia)	TI 1602	Positivo
(Colombia)	TI 291	Positivo
Okso (antiguamente Unión Soviética)	TI 86	Positivo

Ejemplo 13

Creación o generación de mutaciones y cribado de la variación genética en el gen para la nicotina desmetilasa

5 Una variación genética o mutaciones preexistente en la secuencia que codifica para la nicotina desmetilasa o cualquier otro de los genes representados por una secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10-17, o un fragmento de la misma, se criban mediante el uso de tecnologías moleculares, incluyendo lesiones locales inducidas dirigidas en genomas (TILLING), métodos de dactiloscopia de ADN tales como polimorfismos amplificados de longitud de fragmentos (AFLP), y polimorfismos de nucleótido individuales (SNP). En la práctica, se usan poblaciones de plantas que representan una variación genética preexistente tales como una planta transgénica 10 (por ejemplo, cualquiera de las descritas en la presente invención) o aquellas creadas por la exposición de los tejidos reproductivos, semillas, u otros tejidos vegetales a mutágenos químicos tales como agentes de alquilación, etano metil sulfonato (EMS), por ejemplo, o a radiación, tal como rayos X o rayos gamma. Para poblaciones mutadas, se determina experimentalmente la dosis de la sustancia química mutagénica o la radiación para cada tipo de tejido vegetal de tal manera que se obtiene una frecuencia de mutación que está por debajo de un nivel de 15 umbral que se caracteriza por letalidad o esterilidad reproductiva. El número de las semillas de la generación M1 o el tamaño de las poblaciones de plantas M1 resultante de los tratamientos mutagénicos se estiman con base en la frecuencia esperada de mutaciones. La progenie, la generación M2, de las plantas M1 representa la población que deseablemente se evalúa para una mutación en un gen, por ejemplo, el gen para la nicotina desmetilasa.

20 TILLING, dactiloscopia de ADN, SNP o tecnologías similares pueden ser utilizadas para detectar una variación genética inducida o de origen natural en un gen deseable tal como el gen para nicotina desmetilasa. La variación puede deberse a supresiones, sustituciones, mutaciones puntuales, translocaciones, inversiones, duplicaciones, inserciones o mutaciones nulas completas. Estas tecnologías podrían utilizarse en una selección asistida por 25 marcador (programa de fitomejoramiento MA) para transferir o reproducir los alelos nulos o diferentes del gen para la nicotina desmetilasa o cualquier secuencia de ácido nucleico mostrados en las Figuras 1, 3 a 7, y 10-17, o un fragmento de la misma, en otros tabacos. Un fitomejorador podría crear poblaciones segregantes de hibridaciones de un genotipo que contiene el alelo nulo o diferente con un genotipo agrónomicamente deseable. Las plantas en las generaciones F2 o de retrocruzamiento podría ser cribadas utilizando un marcador desarrollado a partir de la 30 secuencia de la nicotina desmetilasa o una secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, y 10-17, o un fragmento de la misma, usando una de las técnicas enumeradas anteriormente. Las plantas identificadas por poseer los alelos nulos o disímiles podría ser retrocruzada o autopolinizada para crear la siguiente población que pudiera ser cribada. Dependiendo del patrón de herencia esperado o de la tecnología MAS utilizada, puede ser necesaria la autopolinización de las plantas seleccionadas antes de cada ciclo de retrocruzamiento para ayudar a la identificación de las plantas individuales deseadas. El retrocruzamiento u otro procedimiento de fitomejoramiento se pueden repetir hasta que se recupera el fenotipo deseado del progenitor recurrente.

35

REIVINDICACIONES

- 5 1. Una planta de tabaco que tiene una mutación en un gen endógeno para nicotina desmetilasa que tiene la secuencia de nucleótidos expuesta en SEQ ID NO: 4, en donde la mutación se selecciona del grupo que consiste en una mutación puntual, una supresión, una inserción, una duplicación y una inversión, y en donde dicha planta de tabaco exhibe expresión reducida del gen mutado para nicotina desmetilasa o el producto génico en relación con una planta de control no mutada, o la nicotina desmetilasa codificada por el gen mutado exhibe actividad enzimática reducida con relación a una nicotina desmetilasa no mutada.
2. La planta de la reivindicación 1, en donde la expresión de la SEQ ID NO: 4 es silenciada.
- 10 3. La planta de cualquiera de las reivindicaciones precedentes, en donde dicha mutación es una mutación puntual o una supresión.
4. La planta de cualquiera de las reivindicaciones precedentes, en donde la planta de tabaco es un tabaco negro, tabaco Burley, tabaco curado al humo, tabaco Virginia, curado al aire, o tabaco oriental.
5. Tabaco curado elaborado a partir de la planta de tabaco de cualquiera de las reivindicaciones precedentes.
- 15 6. Un método para reducir la expresión de la nicotina desmetilasa en una planta de tabaco, dicho método comprendiendo expresar un transgén que codifica una molécula de ARN bicatenario que inhibe la expresión de una nicotina desmetilasa codificada por la SEQ ID NO: 4 en dicha planta, dicha molécula de ARN bicatenario comprendiendo al menos 25 nucleótidos consecutivos que tienen 91% o más de identidad de secuencia con la SEQ ID NO: 4, en donde la expresión de dicha nicotina desmetilasa es inhibida en dicha planta en relación con una planta de tabaco que carece de dicho transgén.
- 20 7. El método de la reivindicación 6, en donde dicha planta de tabaco es de una especie seleccionada del grupo que consiste de *Nicotiana tabacum*, *Nicotiana sylvestris*, *Nicotiana tomentosiformis* y *Nicotiana glauca*.
8. El método de la reivindicación 6, en donde dicha planta de tabaco es *Nicotiana tabacum*.
9. El método de la reivindicación 6, 7, u 8, en donde dicho transgén comprende un promotor constitutivo.
10. El método de la reivindicación 6, 7, u 8, en donde dicho transgén comprende un promotor inducible.
- 25 11. El método de cualquiera de las reivindicaciones 6-10, en donde dicha molécula de ARN bicatenario comprende al menos 100 nucleótidos que tienen 91% o más de identidad de secuencia con un intrón o un exón de la SEQ ID NO: 4.
12. El método de cualquiera de las reivindicaciones 6-10, en donde dicha molécula de ARN bicatenario comprende al menos 100 nucleótidos que tienen 100% de identidad de secuencia con un intrón o un exón de la SEQ ID NO: 4.
- 30 13. El método de cualquiera de las reivindicaciones 6-10, en donde dicha molécula de ARN bicatenario comprende al menos 250 nucleótidos que tienen 91% o más de identidad de secuencia con un intrón o un exón de la SEQ ID NO: 4.
14. El método de cualquiera de las reivindicaciones 6-10, en donde dicha molécula de ARN bicatenario comprende al menos 250 nucleótidos que tienen 100% de identidad de secuencia con un intrón o un exón de la SEQ ID NO: 4.
- 35 15. El método de cualquiera de las reivindicaciones 6-10, en donde dicha molécula de ARN bicatenario comprende al menos 500 nucleótidos que tienen 91% o más de identidad de secuencia con un intrón o un exón de la SEQ ID NO: 4.
- 40 16. El método de cualquiera de las reivindicaciones 12-15, en donde dicha molécula de ARN bicatenario comprende al menos 100 nucleótidos que tienen 100% de identidad de secuencia con los nucleótidos 2010 a 2949 o 3947 a 4562 de la SEQ ID NO: 4.

SEQ ID NO.: 1 (D35-BG11)

atgacttatg cattgcaagt ggaacactta acaatggcac attgatcca aggttcaat
tacagaactic caaatgacga gcccttggat atgaaggaag gtgcaggcat aactatacgt
aaggtaaatac ctgtggaact gataatagcg cctcgctgg cacctgagct ttattaa

SEQ ID NO.: 2 (D35-BG11)

Met Thr Tyr Ala Leu Gln Val Glu His Leu Thr Met Ala His Leu Ile
Gln Gly Phe Asn Tyr Arg Thr Pro Asn Asp Glu Pro Leu Asp Met Lys
Glu Gly Ala Gly Ile Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile
Ile Ala Pro Arg Leu Ala Pro Glu Leu Tyr

FIGURA 1

Estructura Genómica de 33-L
(Gen Genómico de D121-AA8)

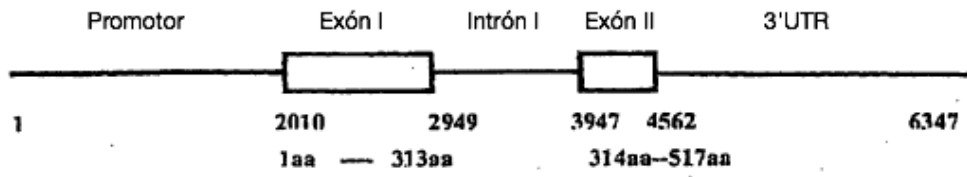


FIGURA 2

SBQ ID NO:3

CARACTERÍSTICAS

Ubicación / Calificadores

Total 1..6347
 Secuencias de flanqueo 5' 1..2009
 CDS unión (2010..2948,3947..4561)
 Intrón 2949..3946
 Producto Citocromo p450
 3' UTR 4562..6347
 Traducción

MLSPIEAIIVGLVTFITFLFFFLWTKKSQKPSKPLPPKIPGGWPFVIGHLFHFNDGDDRPLARKLGDLDADKYGFVP
 TFRLLGLPLVLVSSYEAVKDCFSINDAIFSNRPAFLYGDYLGYNAMFLANYGPYWRKRNKLVIQEVLASARL
 EKFKHVRFARIQASIKNLYTRIDGNSSTINLTDNLEELNFGGLIVKMIAGKNYESGKGDQEVERFKKAFKDFMIL
 SMEFVLWDAPPIPLFKWVDFQGHVIGMKRTFKDIDSVFQNWLEERINKREKMEVNAEGNEQDFIDVVLKMSNE
 YLGEYSRDTVIKATVPSLVLDAAADTVALHINWGMALLINNQKALTKAQEIDTKVGRDRWVESDIKDLVYLQ
 AIVKEVLRLLYPPGPLLVPHEVVEDCVVSGYHYPKGTRLFANVMKLRDPKLVSDPDTDFPERFIATDIDFRGQY
 YKIYIFGSGRRSCPGMTYALQVEHLTMAHLIQGFNYRTPNDEPLDMKEGAGITIRKVNVPVELIAPRLAPELY

SEQ ID NO:4

RECuento BASE 2046 a 984 c 1163 g 2154 t

1 TCTCTAAAGT CCCCTTCCAC TTTATCTTAG CTGTGTGATT TCTTTCAGAC AACCTTATTT
 61 TTATTCAGAC TCTTATTTGT ATTATTTCTAG AAGCTCGTGT ACTTGTGACA CCAGTCTCTGG
 121 GATGGTATTT AGATATCGCT ATTATTTTGG CTTATTCACT TCAGTTCAGA TTTTATTTCCA
 181 GTTATTTGAT TTCCTTATTA TTAATCAAAT TGAATGTGTA AAAATGGTTA AAATTACTCC
 241 AATGTTGGCT TTCCTAGTAA GCGAAATATC AGGCGCCATC ACGGTACCCG AAGGTGAGAA
 301 TTTCAGATCG TGACAGCCGC ATCTCAAGGG GTGTGATGTA AACAGTTTAC GATGGTGCAA
 361 GCATTAGTGG CTGCTTCGAC GACTTAAATC CGTAACTTAT AGATCACACG AATACAACCT
 421 TACTATTTTA ACACCCAGCA AATTCCTGAT AAAAACAAAT CTAACATAGC ACATCAAAT
 481 GTAATGATT GAAGAAAAG ATGACTTTTA TAGACAGAGA AAAAACAGT TACGTGTGGG
 541 GTTGAATAGA GATTGTGGCT ATGCTATTTT TAATATTGAA ATTCACCCGAC TTTTATAGTT
 601 CAATACGAAA AGAGTAAGTG AAAAGGTTCTG AAAAGGAAAA GGACAATGCC TAAAGGACA
 661 CATTACAGAC ACATACACTG AATGATTTTA ATTTCTAGTC CGAAGATTTC TAGTTCGAGG
 721 ATAACGTGAG CAATCGGTTA CTTCCCATTA GCAATTGCCA ACTGGATGTT TGAATATTTA
 781 TGTTTCTGGC CAATAGAGGA GAGGAATACT ACGTTACGTA TGGAGTTTGA ACCCTTCACA
 841 TCAACTTATT AAGTGAGTTA TCCCTCAATC ACGATTCAAC TATGATTCCG CTAACCTCAA
 901 AGAATATTGA GTTAATTATT CAAATGATTA GTCCAAAAT ATTAAATAAA GTTATACTAT
 961 TTCCTTTATT TGTAAACATC ATCTTTTGT TACATATTTA GTTAAATCTT AGCCCAACCC
 1021 TATCGTTGGT TCGACTTTTT TCTTTTAAAT TTTGATTTAT TCTGTTCCGGT AATTTCCGCTT
 1081 TGTTTGGCTT GAATAATAAC TAGTGCATAA AGTCATATAC TCTAATATTT TTAATTGAAG
 1141 TACTCAGAAA TACAAAATA AAAACATTC TAAGCTCACA TGATAGTTGA CAAATCTTT
 1201 ATCCAAAACA AGAGGCGGAG CCAGGATTTG AAACCTATGG GTTCAGAAAT CTAATCTCT
 1261 TAAGTTACTA GGTCTAAAT TAATAATTA TACATGTTCA ATGAATTTCT TAAGACAAAT
 1321 ACATAGTTTG AACGAAAGCT ACTGGGTTTC GCCGAATCCG TAAGTTATAC TCTCCCTCCG
 1381 CCCCAGTCCA AAACAGCTA GTATCAATAG AGAGAGAGAG AGAGAGAGAG AGATAATAAA
 1441 TTTGACCAAT GACAATGGCT TATTACTTGC TTAGAGTTAA TTGGTGAAC TTAGAATAT
 1501 AATAAGGAAT ATTTAAACAG ATACGTATC AATCCACGAG TAACGAAGTA AGAAATACCC
 1561 TAAAATCGTA GAAACATTAC GTTAAATTGC TTGACAGCCT ATCTAGTAAG AGTCAAATC
 1621 TACTATCTAT CTTGTTCCGC CATTTCCTTA AAGAAGTACA TGAGCTTTAT CATCCACCTC
 1681 AACATGAATG CAAAAGAAAA TTATTGTGCA ACTTAATATG TTATAATCAA TGATATGTGT
 1741 CTTGTGTAAC AAAGTATATA TTTGATACG ATATTAAATAT GTAGGTGTTA TATTTTAAA
 1801 TATCAAATAT CATACTTAA ACCGATTTTT TAAAACTTA GGCCAATTAC CCTACCAACT
 1861 AAAATACTGT ATATCAAACA CTAATGTTTT CTATTTCGGT ACGACAGTTC TCTATTTACC
 1921 ATATTATGGA ATTATGCCCA TCCTACAGTT ACCTATAAAA AGGAAGTTGC CGATAGTTAT

FIGURA 3

SEQ ID NO:4 CONT'D.

1981 ATTCTCAACT TCTTATCTAA AAATCCATAA TGCTTCTCC CATAGAAGCC ATTGTAGGAC
 2041 TAGTAAACCTT CACATTTCTC TTCTTCTTCC TATGGACAAA AAAATCTCAA AAACCTTCAA
 2101 AACCCCTTACC ACCGAAAATC CCCGGAGGAT GGCCGGTAAT CGGCCATCTT TTCCACTTCA
 2161 ATGACGACGG CGACGACCGT CCATTAGCTC GAAAACTCGG AGACTTAGCT GACAAATACG
 2221 GCCCCGTTTT CACTTTTCGG CTAGGCCCTC CCCTTGTCTT AGTTGTAAAG AGTTACGAAG
 2281 CTGTAAAAGA CTGTTTCTCT ACAAATGACG CCATTTTTTC CAATCGTCCA GCTTTTCTTT
 2341 ACGGCGATTA CCTTGGCTAC AATAATGCCA TGCTATTTTT GGCCAATTAC GGACCTTACT
 2401 GCGGAAAAAA TCGAAAAATTA GTTATTCAGG AAGTTCCTC CGCTAGTCGT CTCGAAAAAT
 2461 TCAAACACGT GAGATTTGCA AGAATTCAGG CGAGCATTA GAATTTATAT ACTCGAATTG
 2521 ATGGAAATTC GAGTACGATA AATTTAACTG ATTGGTTAGA AGAATTGAAT TTTGGTCTGA
 2581 TCGTGAAGAT GATCGCTGGA AAAAATTATG AATCCGGTAA AGGAGATGAA CAAGTGGAGA
 2641 GRTTAAAGAA AGCGTTTAAAG GATTTTATGA TTTTATCAAT GGAGTTTGTG TTAGGGGATG
 2701 CATTTCCAAAT FCCATTATTT AAAATGGGTGG ATTTTCAAGG GCATGTTAAG GCTATGAAAA
 2761 GGACTTTTAA AGATATAGAT TCTGTTTTTC AGAATTTGGT AGAGGAACAT ATTAATAAAA
 2821 GAGAAAAAAT GGAGGTTAAT GCAGAAAGGA ATGAACAAGA TTTCATTGAT TGTGTGCTTT
 2881 CAAAAATGAG TAATGAATAT CTGGTGAAG GTTACTCTCG TGACTACTGC ATTAAGGCAA
 2941 CCGTGTTTGT AAGTTCATCT GTCATTTTTC ATTTATTCAC TTTTATTTTG AGGAGCAGAC
 3001 ATGTTAATAA TAATTTGGAG CAACTGTAAA GTTACTATG TGTACAGGTT CGAGCCTCAG
 3061 GTGCAACCAC TAATGCTTGT ATTAGATTAT GTTGTCTGCA TCATACCCCT AATTTGGAGTG
 3121 TGGCTCTTCC CGAACCCCTG AATGCTGGAT GCTGGATGCT TTATGTATCA GACTGACCTT
 3181 TTTGTTAAAC TATCTAATA CTAAAGGATGA TTTAATAAAA ATATAGAATG GTAAACAGAA
 3241 AAAGATGAGA TTTATTTTGG GGCTATATGG ATTCGCCCGG GCTTTGGGAG GTAAAACGGT
 3301 ATCTACCAGT TGAGACTTTA CTCCAGAACT TTATCTCGAG AGCTCTGAAT AAAAATGAAA
 3361 TAGTATTTAC CACTCCAAA TCTTTGATGG TAAAAAGATG AGATATAACC TCTTATAATT
 3421 GATTGAACCA CGTTGATAGA ATAAAACTTC TTTACTCCCA TTCAGCATAA GAAAAATGAA
 3481 ACCAAACCGA ATTCCTCTCT TTTTGGGGG GAAATTCCTT AATTTGCTGT TGAATATAGA
 3541 TTCAATGCTT TATTCATTT TTAATAATGA TGAAAATCAA TATAGTCAA GTTAATCCTT
 3601 ATGTCATTTG GTTTGCGGAC AAGTTATATT GGAACTATAT AATACGCTA TTATAGAATA
 3661 GTGATTTATT AGAGGATATA CATTTTTTTT GGATAAATAT TTGATTTATT GGATTA AAAA
 3721 TAGAATATAC AGGTAAGGTC TAAAACGTTG GTTTGCTTTT AACTAAATA AACTTGACCT
 3781 CGTACAATTC TAAAGAAATA TTTGAAATAA ATGAAATATT TTATTGTTAA TCAATTA AAAA
 3841 AAATCATAGT ATAGATGAGA TGTGTGCATA CTTGACAAAT ACTATACTAA CTAAAACAAG
 3901 GTATGTGAAT AATGATATT CCTTTTTTAA TTATTCTTTT TTCCAGAGTT TGGCTTTGGA
 3961 TGCAOCAGAC ACAGTTGCTT TTCACATAAA TTGGGGAATG GCATTATGTA TAAACAATCA
 4021 AAAGGCCCTG ACGAAAGCAC AAGAAGAGAT AGACACAAA GTTGGTAAAG ACAGATGGGT
 4081 AGAAGAGAGT GATATTAAAG ATTTGGTATA CCTCCAGCT ATTGTTAAAG AAGTGTACG
 4141 ATTATATCCA CCAGGACCTT TGTTAGTACC ACACGAAAAT GTAGAAGATT GTTTGTTAG
 4201 TGGATATCAC ATTCCATAAG GGACAGATT ATTCGCAAC GTCATGAAAC TGCAACGTGA
 4261 TCCTAAACTC TGGTCTGATC CTGATACTTT CGATCCAGAG AGATTCATG CTACTGATAT
 4321 TGACTTTTGT GGTCACTACT ATAAGTATAT CCCGTTTGGT TCTGGGAAGC GATCTTGTCC
 4381 AGGGATGACT TATGCATTGC AAGTGGACA CTTAACRAAT GCACATTTGA TCCAAGGTTT
 4441 CAATTACAGA ACTCCAAAAT ACGAGCCCTT GGATATGAAG GAAGGTGCAG GCATAACTAT
 4501 ACGTAAGGTA AATCCTGTGG AACTGATAAT AGCGCCTCGC CTGGCACCTG AGCTTTATTA
 4561 AAACCTAAGA TCTTTCATCT TGGTTGATCA TTGTATRAATA CTCTAAATG GATATTCATT
 4621 TACCTTTTAT CAATTAATTG TCAGTACGAG TTTTCTAAT TTGGTACATT TGAATAATA
 4681 AGTAAAGAAAT AATTGTGCTA ATATATAAAG GTTTGTGAAA GATAATTGAC TGGTGTACC
 4741 ACAATCTCCA GTGAAAGTGT TAATTATTTA CTTGATCCAC AGCTTATTCT ATGTTTGA AA
 4801 TTTGCCTAGT GTCATGATAT TACTCCATCA AATTCAGAA ATAATCATTT CCAACTTTTG
 4861 CTGGACTGGA CGATCTTTCA ATAATAAAGG ATCTTTAATT TGCCAAAGTT GAGATCAAAA
 4921 TACTGGTCCG TTTACCAATA AGAATGAAAT GTGATGGAAA TTATGTACGT TGGGATAAGG
 4981 GAACACAACCT ATCAAGGAGA CTAAAAGGTA CGTAAAGGAA AAGAAAAAAT TTGCCATTGA
 5041 TTGCTACTAA GTAAACCTAAC AAAATCTTTC AGAAAAAGAA CACTTGTATA AGTCGGGGTT
 5101 GAAAGTTTTG GTGTCTCTTT TCTTATGTAT TGTGTCTTT AGACAGTATT GACTTAGTT
 5161 ATTCAGAAAT TTTATTTTGG TATTAGAGCT CAAGACTCTG TATTTATTAG TTTCTGGGAGA
 5221 ATTATCATGT ATTTTCAGTC TTTTGTATT TCTGTAAAT CCGCTATTTT GGTCTTTTAT
 5281 TGCTATGTTT GGCTTTCCTA GAAAATGTTG TAGGCGCTAT CACGACTGAT TGAGATTTTG
 5341 TATCGTGACT GATAATTACA CGGTTTAGTA AATTTTGATA TTTTCAAAA GAGTTTTTTT

FIGURA 3

SEQ ID NO:4 CONT'D.

```

5401 AATAAAATAT GCAACTTCAG TCAAAACATA CAACGTTTTG TTGTATAAAT CCGATCAAAA
5461 CATATAACTT ACATAAAACT TCGGTATGAA TTTTTGTTG TATAAACATT TGGTTAAAAC
5521 ATATAACTTA CATATAATTT GCATACAACT TACATAAAAC TTGCATATAA ACAAATTTATT
5581 TATGTCTTTG TTTTGGAGTA TCAATTTGAA ATTCCAACAA AAACAAACTC TAATTTTAC
5641 CAAACTCTCT CAAAATTGAG TTRTAGATTT CAAAAGATAT CCTAATCGT TTGCAATTGT
5701 AACAAATCGA TCGGCCGTTT TGAGATCTAG CGTGTGTTT GCGGTTTGA GACCTTGAGT
5761 AACITCACTT TATGTTGTAT GACTTGTATA TGTGGTCGGA ATTAATTC GGGGAGTCA
5821 GAGTTGATTC GGATGAAAAA TTCTAATTC GGAAGTTTTA AGATGGAATG ATTGACTAAG
5881 GATTGACGTT TGAGTAAACG ATCTCGGAAT CGAGATTGA AGGTCCAAT AGGTTCGTAT
5941 GATGATTTCA GACTTGAGCG TATGTTTGGG TTGAGTATCG GGTGGTCCGG GAGCAITTC
6001 ACGCTGATTA TAGAAAATTG GCATCTTAAA GGTITTAGAA TTTCATAAGT TTGGTTTGAA
6061 GTGGATTTTG ATATTATCGG TGTCATTTG GAGTTTCGAG CCTTGGAATA GGTTCTGATC
6121 GTAAATTTTG ACTTTAGTGT AAAGTTCGGC GTCATTCGG AGTGTTTTGA TAAGATTCTG
6181 ATGCGTTCGT CGAAGTTTGG AAGTTTGAAA GTTGAAAAGA AGATTTTAA TAGGCGATTG
6241 ATGATTTTGA TGTTATTTGT GTCGAGCCTT TGGATAAGTT TGTGTGAGGT ATGGGACTTG
6301 TTGGTATGAA TGGACGAGCT CTACGGGGC CTCGAGTAAG TTTCCGA
    
```

FIGURA 3

SEQ ID NO:5

Secuencias que Codifican al Gen para la Nicotina Desmetilasa de Tabaco

RECUENTO BASE	489 a	275 c	333 g	457 t		
1	ATGCTTTCTC	CCATAGAAGC	CATGTAGGA	CTAGTAACCT	TCACATTCT	CTTCTTCTTC
61	CTATGGACAA	AAAAATCTCA	AAAACCTTCA	AAACCTTAC	CACCGAAAAT	CCCCGGAGGA
121	TGGCCGGTAA	TGGCCCATCT	TTTCCACTTC	AATGACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAACCTCG	GAGACTTAGC	TGACAAATAC	GGCCCCGTTT	TCACCTTTTCG	GCTAGGCCTT
241	CCCCTTGCTC	TAGTTGTAAG	CAGTTACGAA	GCTGTAAAAG	ACTGTTTCTC	TACAAATGAC
301	GCCATTTTTT	CCAATCGTCC	AGCCTTCTT	TACGGCGAAT	ACCTGGCTA	CAATAATGCC
361	ATGCTATTTT	TGGCCRAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGAAAATT	AGTTATTTCAG
421	GAAGTTCTCT	CCGCTAGTCG	TCTCGAAAAA	TTCAAACACG	TGAGATTTCG	AAGAATTCAA
481	GCGAGCATT	AGAATTTATA	TACTCGAATT	GATGGAAATT	CGAGTACGAT	AAATTTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAAGA	TGATCGCTGG	AAAAAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAG	AGATTTAAGA	AAGCGTTTAA	GGATTTTATG
661	ATTTTATCAA	TGGAGTTTGT	GTTATGGGAT	GCATTTCCAA	TTCCATTATT	TAAATGGGTG
721	GATTTTCAAG	GGCATGTTAA	GGCTATGAAA	AGGACTTTTA	AAGATATAGA	TTCTGTTTTT
781	CAGAATTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTTAA	TGCAGAAGGG
841	AATGAACAAG	ATTTCAATGA	TGTGGTGCCT	TCAAAAATGA	GTAATGAATA	TCTTGGTGAA
901	GGTACTCTC	GTGATACTGT	CATTAAAGCA	ACGGTGTTTA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTACAT	AAATGGGGA	ATGGCATTAT	TGATAAACAA	TCRAAAGGCC
1021	TTGACGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGGACAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTGGT	ATACCTCCAA	GCTATTGTTA	AAGAAGTGT	ACGATTATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACACGAA	AATGTAGAAG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGGACAAG	ATTATTCCCA	AACGTCATGA	AACTGCAACG	TGATCCTAAA
1261	CTCTGGTCTG	ATCCTGATAC	TTTCGATCCA	GAGAGATFCA	TTGCTACTGA	TATTGACTTT
1321	CGTGGTCAGT	ACTATAAGTA	TATCCCGTTT	GGTCTCGGAA	GAAGATCTTG	TCCAGGGATG
1381	ACTTATGCAT	TGCAAGTGG	ACACTTAACA	ATGGCACATT	TGATCCAAGG	TTTCAATTAC
1441	AGAACTCCAA	ATGACGAGCC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	TATACGTAAG
1501	GTAATTCCTG	TGGAACTGAT	AATAGCGCCT	CGCCTGGCAC	CTGAGCTTTA	TTAA

SEQ ID NO:6

Aminoácidos Deducidos del Gen para la Nicotina Desmetilasa de Tabaco

1	MLSPIEIAIVG	LVTFTFLFFF	LWTKKSQKPS	KPLPPKIPGG	WPVIGHLPHF	NDDGDDRPLA
61	RLGLDLADKY	GPVFTFRLGL	PLVLVSSYE	AVKDCPSTND	AIFSNRPAPL	YGDYLYGYNNA
121	MLFLANYGPY	WRKNRKLVIQ	EVLSASRLEK	FKHVRFARIQ	ASIKNLYTRI	DGNSSTINLT
181	DWLEELNFGI	IVKMIAGKNY	ESGKGDEQVE	RFKAKKDFM	ILSMFVLWD	APPPIPLFKWV
241	DPQGHVKAMK	RTFKDIDSVF	QNWLEEHINK	REKMEVNAEG	NEQDFIDVVL	SKMSNEYLGE
301	GYSRDTVIKA	TVFSLVLDA	DTVALHINWG	MALLINNQKA	LITKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRLY	PPGPLLVPHE	NVEDCVVSGY	HIPKGRFLFA	NVMKLRDPK
421	LWSDPDFTFD	ERPIATDIDF	RQYYKYIPF	GSGRRSCPGM	TYALQVEHLT	MAHLIQGFVY
481	RTFNDEPLDM	KEGAGITIRK	VNPVELIIAP	RLAPELY		

FIGURA 4

SEQ ID NO:7

Intrón del gen para Nicotina Desmetilasa de Tabaco (998 nt)

```

1  GTAAGTTCAT CTGTCATTTT TCATTTATTC ACTTTTATTT TGAGGAGCAG ACATGTTAAT
61  AATAATTGG AGCAACTGTA AAGTTATCTA TGTGTACAGG TTCGAGCCTC AGGTGCAACC
121 ACTAATGCTT GTATTAGATT ATGTTGTCTG CATCATACCC CTAATTGGAG TGTGGCTCTT
181 CCCGAACCCT GCAATGCTGG ATGCTGGATG CTTTATGTAT CAGACTGACC TTTTGTGTTAA
241 ACTATCTAAA TACTAAGGAT GATTTAATAA AAATATAGAA TGGTAAACAG AAAAAGATGA
301 GATTATTTTT GGGGCTATAT GGATTCGCCC GGGCTTTGGG AGGTAAAACG GTATCTACCA
361 GTTGAGACTT TACTCCAGAA CTTTATCTCG AGAGCTCTGA ATAAAAATGA AATAGTATTT
421 ACCACTCCAA AATCTTTGAT GGTAAAAGA TGAGATATAA CCTCTTATAA TTGATTGAAC
481 CACGTTGATA GAATAAAACT TCTTACTCC CATTGAGCAT AAGAAAAATG AAACCAAACG
541 GAATTCCTCT CTTTTTAGG GGGAAATTC TTAATTGCTT GTTGAATATA GATTCATGTC
601 GTTATTCTAT TTTAATAAT GATGAAAATC AATRTAGTCA AAGTTAATAC TTATGTCATT
661 TGGTTTGGCG ACAAGTTATA TTGGAACTAT ATAATACGTC TATTATAGAA TAGTGATTAT
721 TTAGAGGATA TACAATTTTT TTGGATAAAT ATTTGATTTA TTGGATTAAA AATAGAATAT
781 ACAGGTAAGG TCTAAAACGT GTGTTTGCTT TTACACTAAA TAACTTGAC CTCGTACAAT
841 TCTAAGAAAA TATTTGAAAT AAATGAATTA TTTTATGTT AATCAATTAA AAAAATCATA
901 GTATAGATGA GATGTGTGCA TACTTGACAA TAACATACT AACTAAAACA AGGTATGTGA
961 ATAATTGATA TTCCTTTTTT AATTATTCTT TTTCCAG
    
```

FIGURA 5

SEQ ID NO: 8

Promotor del gen para Nicotina Desmetilasa de tabaco (2009 nt)

```

1 TCTCTAAAGT CCCCTTCCAC TTTATCTTAG CTGTGTGATT TCTTTCAGAC AACCTTATTT
61 TTATTCAGAC TCTTATTTGT ATTATCTAG AAGCTCGTGT ACTTGTGACA CCAGTTCGTG
121 GATGGTATTT AGATATCGCT ATTATTTTGG CTTATTCACT TCAGTTCAGA TTTTATTCCA
181 GTTATTTGAT TTCTTTATTA TTAATCAAAT TGAATTGTTA AAAATGGTTA AAATTACTCC
241 AATGTTGGCT TTCCTAGTAA GCGAAATATC AGGCGCCATC ACGGTACCCG AAGGTGAGAA
301 TTTCAGATCG TGACAGCCGC ATCTCAAGGG GTGTGATGTA AACAGTTTAC GATGGTGCAA
361 GCATTAGTGG CTGCTTCGAC GACTTAAATC CGTAACTTAT AGATCACACG AATACAACCT
421 TACTATTTTA ACACCCAGCA AATTCCTGAT AAAACAATT CTAACATAGC ACATCAAAT
481 GTAAATGATT GAAGAAAAAG ATGACTTTTA TAGACAGAGA AAAAAACAGT TACGTGTGGG
541 GTTGAATAGA GATTGTGGCT ATGCTATTTT TAATATTGAA ATTCACCGAC TTTTTTAGTT
601 CAATACGAAA AGAGTAAGTG AAAAGGTCTG AAAAGGAAAA GGACAATGCC TAAAAGGACA
661 CATTCAAGAA ACATACACTG AATGATTCTA ATTTCTAGTC CGAAGATTTT TAGTTCGAGG
721 ATAACTGAG CAATCGGTTA CTTCCATTA GCAATGGCCA ACTGGATGTT TGACTATTTA
781 TGTTTCTGGC CAATAGAGGA GAGGAATACT ACGTTACGTA TGGAGTTTGA ACCCTTACA
841 TCAACTTATT AAGTGAGTTA TCCCTCAATC ACGATTCAAC TATGATTCCG CTAACCTCAA
901 AGAATATTGA GTTAATTATT CAAATGATTA GTCCAAAATT ATTTAATAAA GTTATACTAT
961 TTCTTTTATT TGTAACATAC ATCTTTTGT TACATATTTA GTTAAATCTT AGCCCAACCC
1021 TATCGTTGGT TCGACTTTTT TTCTTTTAAAT TTGATTTAT TCTGTTGCGT AATTTGCTT
1081 TGTTTGGCTT GAATAATAAC TAGTGCATAA AGTCATATAC TCTAATATTT TTAATTGAAG
1141 TACTCAGAAA TACAAAAATA AAAAACATTC TAAGCTCACA TGATAGTTGA CAAAATCTTT
1201 ATCCAAAACA AGAGGCGGAG CCAGGATTTG AACTTATGG GTTCAGAATT CTAATCTCT
1261 TAAGTTACTA GGTTCTAAAT TAATAATTTA TACATGTTCA ATGAATTTCT TAAGACAAAT
1321 ACATAGTTTG AACGAAAGCT ACTGGGTTG GCCGAATCCG TAAGTTATAC TCTCCCTCCG
1381 CCCCGTCCA AAACAGCTA GTATCAATAG AGAGAGAGAG AGAGAGAGAG AGATAATAAA
1441 TTTGACCATT GACAAATGGT TATTACTTGC TTAGAGTTAA TTGGTGAAC TAGAGAATAT
1501 AATAAGGAAT ATTTAAACAG ATACGTCATC AATCCACGAG TAACGAAGTA AGAAATACC
1561 TAAAATCGTA GAAACATTAC GTTAAATTGC TTGACAGCCT ATCTAGTAAG AGTCAAAATC
1621 TACTATCTAT CTTGTCCGC CATTTTCTTA AAGAAGTACA TGAGCTTTAT CATCCACCTC
1681 AACATGAATG CAAAAGAAAA TTATTGTGCA ACTTAATATG TTATAATCRA TGATATGTGT
1741 CTGTGTAAAC AAAGTATATA TTTGATACG ATATTAATAT GTAGGTGTTA TATTTTAA
1801 TATCAAATAT CATACTAAC ACCGATTTT TAAAACTTA GGCCAATTAC CCTACCAACT
1861 AAAATACTGT ATATCAAACA CTAATGTTTT CTATTTGCGT ACGACAGTTC TCTATTTACC
1921 ATATTATGGA ATTATGCCA TCCTACAGTT ACCTATAAAA AGGAAGTTGC CGATAGTTAT
1981 ATTCTCAACT TCTTATCTAA AAATCCATA

```

FIGURA 6

SEQ ID NO: 9

3' UTR del gen para Nicotina Desmetilasa de tabaco

```

1 AACCTAAGAT CTTTCATCTT GGTGATCAT TGTATAATAC TCCTAAATGG ATATTCATTT
61 ACCTTTTATC AATTAATTGT CAGTACGAGT TTTTCTAATT TGGTACATTT GTAATAATAA
121 GTAAAGAATA ATTGTGCTAA TATATAAAGG TTTGTAGAAG ATAATTGACT GGTGTGACCA
181 CAATCTCCAG TGAAAGTGTT AATTATTTAC TTGATCCACA GCTTATTCTA TGTTTGAAT
241 TTGCCTAGTG TCATGATATT ACTCCATCAA ATTCAGAAA TAATCATTTT CAACTTTTC
301 TGGACTGGAC GATCTTTCAA TAATAAAGGA TCTTTAATTT GCCAAGTTG AGATCAAAAT
361 ACTGGTCGCT TTACCAATAA GAATGAAATG TGATGGAAAT TATGTACGTT GGGATAAGGG
421 AACACAATA TCAAGGAGAC TAAAAGGTAC GTAAAGGAAA AGAAAAAATT TGCCATTGAT
481 TGCTACTAAG TAACCTAACA AAATCTTTCA GAAAAGAATC ACTTGATATA GTCGGGGTTG
541 AAAGTTTTGG TGTCTCTTTT CTTATGTATT GTTGTCTTTA GACAGTATTG TACTTAGTTA
601 TTTCAGAATT TTATTTTCGT ATTAGAGCTC AAGACTCTGT ATTTATTAGT TCTGGGAGAA
661 TTATCATGTA TTTTCAGTCT TTTGTTATTT CTGTAAATTC CGCTATTTTG GTTCTTTTAT
721 GCTATGTTCC GCTTTCCTAG AAAATGTGTT AGGCGCTATC ACGACTGATT GAGATTTTGT
781 ATCGTGACTG ATAATTACAC GGTTTAGTAA ATTTTGATAT TTTCAAAAAG AGTTTTTTTA
841 ATAAAATATG CAACTTCAGT CAAAACATAC AACGTTTTGT TGTATAAATC CGATCAAAAC
901 ATATAACTTA CATAAACTT CCGTATGAAT TTTTGTGTT ATAACAATT GGTAAAAACA
961 TATAACTTAC ATATAATTTG CATACAACTT ACATAAAACT TGCATATAAA CAATTTATTT
1021 ATGTCTTTGT TTTTGAGTAT CAATTTGAAA TTCCAACAAA AACAACTCT AATTTTTTACC
1081 AAACCTCTCT AAAAATTGAGT TATAGATTTT AAAAGATATC CTTAATCGTT TGCAATTGTA
1141 ACAATCCGAT CGGCCGTTTT GAGATCTAGC GTGTTGTTTG GCGGTTTGAG ACCTTGAGTA
1201 ACTTCACTTT ATGTTGTATG ACTTGATAT ATGTTGTTTG GTGTTGTTTG GGAAGTTCAG
1261 AGTTGATTCG GATGAAAAAT TCTAATTTTC GAAGTTTTTA GATGGAATGA TTGACTAAGG
1321 ATTGACGTTT GAGTAAACGA TCTCCGAATC GAGATTTGAA GGTCCRAATA GGTTCGTATG
1381 ATGATTTTCA ACTTGAGCGT ATGTTTGGGT TGAGTATCGG GTGTTCCGGG AGCATTTCAA
1441 CGCTGATTAT AGAAAAATGG CATCTTAAAG GTTTTAGAAT TTCATAAGTT TGGTTTGAAG
1501 TGGATTTTGA TATTATCGGT GTCCATTTGG AGTTTTCGAG CTGGAATAG GTTCGTATCG
1561 TAAATTTTGA CTTTAGTGT AAGTTTCGGC TCATTCGGA GTGTTTTCAT AAGATTCTGA
1621 TCCGTTTCGT GAAGTTTGA AGTTTGAAG TTGAAAAGAA GATTTTAAAT AGGCGATTCA
1681 TGATTTTGT GTTATTGTG TCGAGCCTTT GGATAAGTTT GTGTGAGGTA TGGGACTTGT
1741 TGGTATGAAT GGACGAGCTC TACGGGGGCC TCGAGTAAAT TTCGGA

```

FIGURA 7

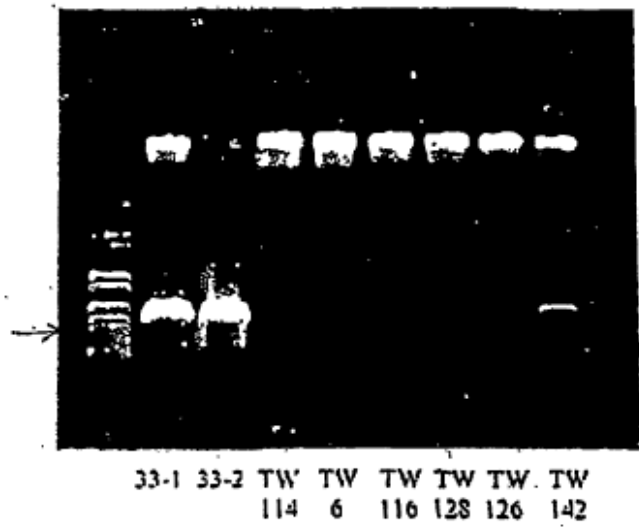


FIGURA 8

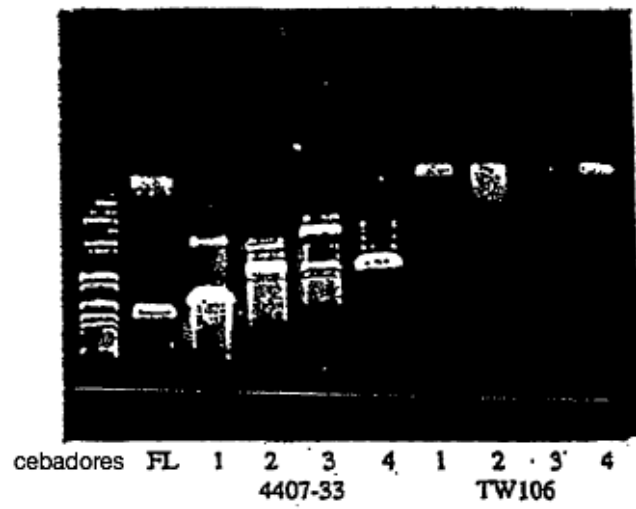


FIGURA 9

FIG. 10

SEQ ID 62 D35-BG11
 1 ATGACTTAT GCATTGCAAG TGAACACTT AACAAATGGCA-
 61 CATTTGATCC AAGGTTTCAA TTACAGAACT CCAAATGACG AGCCCTTGGG TATGAAGGAA
 121 GGTGCAGGCA TAACATACG TAAGGTAAT CCTGTGGAAC TGATAATAGC GCCTCGCCTG
 181 GCACCTGAGC TTTATTAA
 SEQ ID 63
 MTYALQVEHLTMAHLIQGFNYRTPNDEPLDMKEGAGITIRKVNPFVELIIPRLAPELY

FIG. 11

NOMBRE D120-AH4
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 188

1	ATAATGCTTT	CTCCCATAGA	AGCCATGTGA	GGACTAGTAA	CCTTCACATT	TCTCTCTTC
61	TTCCATATGGA	CAAAAAATC	TCAAAAACCT	TCAAAAACCT	TACCACCAGAA	AATCCCCGGA
121	GGATGGCCGG	TAATCGGCCA	TCTTTTCCAC	TTCAATGACG	ACGGCGACGA	CCGTCCATTA
181	GCTCGAAAAC	TCGGAGACTT	AGCTGACAAA	TACGGCCCCG	TTTTCACTTT	TCGGCTAGGC
241	CTTCCCCTTG	TCTTAGTTGT	AAGCAGTTAC	GAAGCTGTAA	AAGACTGTTT	CTCTACAAAT
301	GACGCCATTT	TTTCCAATCG	TCCAGCTTTT	CTTTACGGCG	ATTACCTTGG	CTACAATAAT
361	GCCATGCTAT	TTTTGGCCAA	TTACGGACCT	TACTGGCGAA	AAAATCGAAA	ATTAGTTATT
421	CAGGAAGTTC	TCTCCGCTAG	TCGTCTCGAA	AAATTCAAAC	ACGTGAGATT	TGCAAGAATT
481	CAAGCGAGCA	TTAAGAAATT	ATATACTCGA	ATTGATGGAA	ATTCGAGTAC	GATAAATTTA
541	ACTGATGGT	TAGAAGAATT	GAATTTTGGT	CTGATCGTGA	AGATGATCGC	TGGAAAAAAT
601	TATGAATCCG	GTAAGGAGA	TGAACAAGTG	GAGAGATTTA	AGAAAGCGTT	TAAGGATTTT
661	ATGATTTTAT	CAATGGAGTT	TGTGTTATGG	GATGCATTTT	CAATTCATT	ATTAAATGG
721	GTGGATTTTC	AAGGGCATGT	TAAGGCTATG	AAAAGGACTT	TTAAAGATAT	AGATTCGTGT
781	TTTCAGAATT	GGTTAGGGGA	ACATATTAAT	AAAAGAGAAA	AAATGGAGGT	TAATGCAGAA
841	GGGAATGAAC	AAGATTTTCA	TGATGTGGTG	CTTCAAAAA	TGAGTAATGA	ATATCTTGGT
901	GAAGGTACT	CTCGTGATAC	TGTCATTAAT	GCAACGGTGT	TTAGTTTGGT	CTTGGATGCA
961	GCAGACACAG	TTGCTCTTCA	CATAAATTGG	GGAAATGGCAT	TATGATATAA	CAATCAAAAG
1021	GCCTTGACGA	AAGCACAAGA	AGAGATAGAC	ACAAAAGTTG	GTAAGGACAG	ATGGGTAGAA
1081	GAGAGTGATA	TTAAGGATTT	GGTATACCTC	CAAGCTATTG	TTAAAGAAGT	GTTACGATTA
1141	TATCCACCAG	GACCTTTGTT	AGTACCACAC	GAAAATGTAG	AAGATTGTGT	TGTTAGTGA
1201	TATCACATTC	CTAAAGGGAC	AAGATTATTC	GCAAACGTCA	TGAAACTGCT	ACGTGATCCT
1261	AAACTCTGGC	CTGATCCTGA	TACTTTCGAT	CCAGAGAGAT	TCATTGCTAC	TGATATTGAC
1321	TTTCGTGGTC	AGTACTATAA	GTATAACCCG	TTTGGTTCTG	GAAGACGATC	TTGTCCAGGG
1381	ATGACTTATG	CATTGCAAGT	GGAACACTTA	ACAATGGCAC	ATTTGATCCA	AGGTTTCAAT
1441	TACAGAACTC	CAAATGACGA	GCCCTTGGAT	ATGAAGGAAG	GTGCAGGCAT	AACATACGT
1501	AAGGTAATC	CTGTGGAACT	GATAATAGCG	CCTCGCCTGG	CACCTGAGCT	TTATTAATAAC
1561	CTAAGATCTT	TCATCTTGGT	TGATCATTGT	ATAATACTCC	TAAATGGATA	TTCATTTACC
1621	TTTTATCAAT	TAA				

SEQ. ID. NO. 189

1	MLSPIERAVG	LVTFTFLFFF	LWTKKSQKPS	KPLPPKIPGG	WVVIHGLFHF	NDDGDDRPLA
61	RKLGDLADKY	GPVFTFRLGL	PLVLVSSYE	AVKDCFSTND	AIFSNRPAPL	YGDYLYGYNNA
121	HLFLANYGPF	WRQNRKLVIQ	EVLASARLEK	FKHVRPARIQ	ASIKHLYTRI	DGNSSTINLT
181	DWLEELNFGI	IVKMIAGKNY	ESGRGDEQVE	RFKKAFKDFM	ILSMEFVLWD	AFPIPLFKWV
241	DFQGHVKANK	RTFKDIDSVF	QNWLGHEHMK	REKMEVNAEG	NEQDFIDVVL	SKMSNEYLGE
301	GYSRDTVIKA	TVFSLVLDAA	DIVALHINWG	MALLINNQKA	LTKAQBEIDT	KVQKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRLY	PPGFLLVPHE	NVEDCVVSGY	HIPKGTRLF	NVMKLLRDPK
421	LWPDFDTDF	ERFIATDIDF	RGQYKYIPF	GSGRRSCPDM	TYALQVEHLT	MAHLIQGFNY
481	RTPNDEPLDM	KEGAGITIRK	VNPVELLIAP	RLAPELY		

FIG. 12

NOMBRE	D121-AA8					
ORGANISMO	NICOTIANA TABACUM					
SEQ. ID. NO.	190					
1	AATCCATAAT	GCTTTCCTCC	ATAGAAGCCA	TTGTAGGACT	AGTAACCTTC	ACATTTCTCT
61	TCCTTCCTCC	ATGGACAAAA	AAATCTCAAA	AACCTTCAAA	ACCCTTACCA	CCGAAAAATCC
121	CCGGAGGATG	GCCGGTAATC	GGCCATCTTT	TCCACTTCAA	TGACGACGGC	GACGACCGTC
181	CATTAGCTCG	AAAACCTCGG	GACTTAGCTG	ACAAATACGG	CCCCGTTTC	ACTTTTCGGC
241	TAGGCCTTCC	CCTTGTCTTA	GTGTAAGCA	GTTACGAAGC	TGTA AAAAGAC	TGTTTCTCTA
301	CAAAATGACGC	CATTTTTTCC	AATCGTCCAG	CTTTTCTTTA	CGGCGATTAC	CTTGGCTACA
361	ATAATGCCAT	GCTATTTTGG	GCCAATTACG	GACCTTACTG	GCGAAAAAT	CGAAAATTAG
421	TTATTCAGGA	AGTTCCTCTC	GCTAGTCTGT	TCGAAAAAT	CAAACACGTG	AGATTTGCAA
481	GAATTCAAGC	GAGCATTAA	AATTTATATA	CTCGAATGA	TGGAAAATCG	AGTACGATAA
541	ATTTAACTGA	TTGGTTAGAA	GAATTGAATT	TTGGTCTGAT	CGTGAAGATG	ATCGCTGGAA
601	AAAAATTATGA	ATCCGGTAAA	GGAGATGAAC	AAGTGGAGAG	ATTTAAGAAA	GCCTTAAAGG
661	ATTTTATGAT	TTTATCAATG	GAGTTTGTGT	TATGGGATGC	ATTTCCAATT	CCATTATTTA
721	AATGGGTGGA	TTTTCAAGGG	CATGTTAAGG	CTATGAAAAG	GACTTTTAAA	GATATAGATT
781	CTGTTTTTCA	GAATTGGTTA	GAGGAACATA	TTAATAAAA	AGAAAAAATG	GAGGTTAATG
841	CAGAAGGGAA	TGAACAAGAT	TTCATTGATG	TGGTGCTTTC	AAAAATGAGT	AATGAAATATC
901	TTGGTGAAGG	TTACTCTCGT	GATACTGTCA	TTAAAGCAAC	GGTGTTTAGT	TTGGTCTTGG
961	ATGCAGCAGA	CACAGTTGCT	CTTCACATAA	ATTGGGGAAT	GGCATTATTG	ATAACAATC
1021	AAAAGGCCTT	GACGAAAAGCA	CAAGAAGAGA	TAGACACAAA	AGTTGGTAAG	GACAGATGGG
1081	TAGAAGAGAG	TGATATTAAG	GATTTGGTAT	ACCTCCAAGC	TATGTTAAA	GAAGTGTAC
1141	GATTATATCC	ACCAGGACCT	TTGTTAGTAC	CACACGAAA	TGTAGAAGAT	TGTGTTGTTA
1201	GTGGATATCA	CATTCCTAAA	GGGACAAGAT	TATTCGCAA	CGTCATGAAA	CTGCAACGTG
1261	ATCCTAAACT	CTGGTCTGAT	CCTGATACTT	TCGATCCAGA	GAGATTCATT	GCTACTGATA
1321	TTGACTTTCG	TGGTCAGTAC	TATAAGTATA	TCCCGTTTGG	TTCTGGAAGA	CGATCTTGTC
1381	CAGGGATGAC	TTATGCATG	CAAGTGGAAC	ACTTAACRAAT	GGCACATTTG	ATCCAAGGTT
1441	TCAATTACAG	AACTCCAAT	GACGAGCCCT	TGGATATGAA	GGAAAGGTGCA	GGCATAACTA
1501	TACGTAAGGT	AAATCCTGTG	GAACCTGATAA	TAGCGCCTCG	CCTGGCACCT	GAGCTTTATT
1561	AAAACCTAAG	ATCATCTTGC	TTGAT			
SEQ. ID. NO.	191					
1	MLSPIEAIVG	LVTFTFLFFP	LWTKKSQKPS	KPLPPKIPGG	WVVIHGLFHP	NDDGDDRPLA
61	RKLGDLADKY	GPVPTFRGL	PLVLVSSYE	AVKDCFSTND	AIFSNRPAPL	YGDYLGYNNA
121	MLFLANYGYP	WRKNRKLVIQ	EVLASARLEK	FKHVRFARIQ	ASIKNLYTRI	DGNSSTINLT
181	DWLEELNPGL	IVKMIAGKNY	ESGKGDEQVE	RFKKAFKDFM	ILSMEFVLWD	APPIPLFKWV
241	DFQGHVKAMK	RTPKDIDSVF	QNWLEEHINK	REKMEVNAEG	NEQDFIDVVL	SKMSNEYLGE
301	GYSRDTVIKA	TVPSLVLDAA	DTVALHINWG	MALLINNQKA	LTKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	ALVKEVLRLY	PPGPILLPHE	NVEDCVVSGY	HIFKGTRLFPA	NVMKLRDPK
421	LWSDFDTFDP	ERFIATDIDF	RGQYKYIIPF	GSGRRSCPDM	TYALQVEHLT	MAHLIQGFNY
481	RTPNDEPLDM	KEGAGITIRK	VNPVELIAP	RLAPELY		

FIG. 13

NOMBRE D122-AF10
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 192

```

1 CTAAAACCTCC ATAATGGTTT CTCCCGTAGA AGCCATTGTA GGACTAGTAA CCCTTACACT
61 TCTCTTCTAC TTCCCTATGGC CCAAAAAATT TCAAATACCT TCAAAACCAT TACCACCGAA
121 AATTCCCAGGA GGGTGGCCGG TAATCGGCCA TCITTTCTAC TTCGATGATG ACGGCGACGA
181 CCGTCCATTA GCTCGAAAAC TCGGAGACTT AGCTGACAAA TACGGCCCGG TTTTCACITT
241 CCGGCTAGGC CTCCCGCTTG TGTTAATTGT AAGCAGTTAC GAAGCTGTAA AAGACTGCTT
301 CTCTACAAAT GACGCCATTT TCTCCAATCG TCCAGCTTTT CTTTACGGTG AATACCTTGG
361 CTACAATAAT GCCATGCTAT TTTTGACAAA ATACGGACCT TATTGGCGAA AAAATAGAAA
421 ATTAGTCATT CAGGAAGTTC TCTCTGCTAG TCGTCTCGAA AAATTGAAGC ACGTGAGATT
481 TGGTAAAAAT CAAACGAGCA TTAAGAGTIT ATACACTCGA ATTGATGGAA ATTCGAGTAC
541 GATAAATCTA ACTGATTGGT TAGAAGAATT GAATTTTGGT CTGATCGTGA AAATGATCGC
601 TGGGAAAAAT TATGAATCCG GTAAGGAGA TGAACAAGTG GAGAGATTTA GGAAAGCGTA
661 TAAGGATTTT ATAATTTTAT CAAATGGAGTT TGTTGTTATGG GATGCTTTTC CAATTCATT
721 GTTCAAATGG GTGGATTTTC AAGGCTATGT TAAGGCCATG AAAAGGACAT TTAAGGATAT
781 AGATTCTGTT TTTCAGAAAT GGTAGAGGA ACATGTCAAG AAAAGAGAAA AAATGGAGGT
841 TAATGCACAA GGGAAATGAAC AAGATTTTCAT TGATGTGGTG CTTTCAAAA TGAGTAAATGA
901 ATATCTTGAT GAAGGTACT CTCTGTGATC TGTCATAAAA GCAACAGTGT TTAGTTTGGT
961 CTTGGATGCT GCGGACACAG TTGCTCTTCA CATGAATTGG GGAATGGCAT TACTGATAAA
1021 CAATCAACAT GCCTTGAAGA AAGCACAGA AGAGATCGAT AAGAAAGTTG GTAAGGAAAAG
1081 ATGGGTAGAA GAGAGTGATA TTAAGGATT GGTCTACCTC CAAGCTATTG TTAAGAAAGT
1141 GTTACGATTA TATCCACCAG GACCTTTATT AGTACCTCAT GAAAATGTAG AGGATTGTGT
1201 TGTTAGTGGG TATCACATTC CTAAGGGGAC TAGACTATTC GCGAACGTTA TGAATTGCA
1261 GCGCGATCCT AACTCTGGT CAAATCCTGA TAAGTTTGAT CCAGAGAGAT TCTTCGCTGA
1321 TGATATTGAC TACCGTGGTC AGCACTATGA GTTTATCCCA TTTGGTTCTG GAAGACGATC
1381 TTGTCCGGGG ATGACTTATG CATTACAAGT GGAACACCTA ACAATAGCAC ATTTGATCCA
1441 GGGTTTCAAT TACAAAACCT CAAATGACGA GCCCTTGGAT ATGAAGGAAG GTGCAGGATT
1501 AACTATACGT AAAGTAAATC CTGTAGAAGT GACAATTACG GCTCGCCTGG CACCTGAGCT
1561 TTATTAACAC CTTAGATGTT TTATCTTGAT TGTAATAATA TATATATGCA GAAAAATTG
    
```

SEQ. ID. NO. 193

```

1 MVSPVEAIVG LVTLLLPYF LWPKKFQIPS KPLPPKIPGG WPVIGHLPYF DDDGDDRPLA
61 RKLGDLDLADKY GPVFTFRLGL PLVLIVSSYE AVKDCPSTND AIFSNRPAFL YGEYLGYNNA
121 MDFLTQYGPY WRKRNKLVIQ EVLSASRLEK LKHVRFQKIQ TSIKSLYTRI DGNSSSTINLT
181 DWLEELNPLG IVMIAIGKTY ESGKGDEQVE RFRKAYKDFI ILSMEFVLWD AFPPIPLFKWV
241 DFQGYVKAMK RTFKDIDSVF QNMLEEHVKK REKMEVNAQG NEQDFIDVVL SKMSNEYLDE
301 GYSRDTVIKA TVFSLVLDAA DTVALHMHWG MALLINNQHA LKKAQEIDK KVGKERWVEE
361 SDIKDLVILQ AIVKEVLRLY PPGPLLVPHE NVEDCVVSGY HIPKGRFLFA NVMKLRDPK
421 LWSNPKRFDK ERFVADDIDY RQHYEFIFP GSGRRSCPFGM TYALQVEHLT LAHLIQGPNY
481 KTFNDEFLDM KEGAGLTIRK VNPVEVTITA RLAPELY
    
```

FIG. 14

NOMBRE D208-AC8
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 226

1	ATGCTTTCTC	CCATAGAAGC	CTTTGTAGGA	CTAGTAACCT	TCACATTTCT	CTTATACTTC
61	CTATGGACAA	AAAAATCTCA	AAAACCTCCA	AAACCCTTAC	CACCGAAAAT	CCCCGGAGGA
121	TGGCCGGTAA	TCGGCCATCT	TTTTCACTTC	AATAACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAGCTCG	GAGACTTAGC	TGATAAATAC	GGCCCCGTTT	TCACTTTTCG	GCTAGGCTCT
241	CCCCTTTGTC	TAGTTGTAAG	CAGTTACGAA	GCTATAAAAG	ATTGCTTCTC	TACAAATGAT
301	GCCATTTTCT	CCAATCGTCC	AGCTCTTCTT	TACGGCGAAT	ACCTTGGCTA	CAATAATACA
361	ATGCTTTTTC	TAGCAAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGTAAATT	AGTCATTACAG
421	GAAGTTCCTC	CTGCTAGTCG	TCTCGAAAAA	TTCAAACAAG	TGAGATTAC	CAGAATTCAA
481	ACGAGCATT	AGAATTTATA	CACTCGAATT	AATGGAAATT	CGAGTACGAT	AAATCTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAATA	TGATCGCTGG	GAAAAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAA	AGATTTAAGA	ATGCGTTTAA	GGATTTTATG
661	GTTTATCAA	TGGAATTTGT	ATTATGGGAT	GCATTTCCAA	TTCCATTATT	TAATGGGGTG
721	GATTTTCAAG	GTCAATTTAA	GGCAATGAAA	AGGACATTTA	AGGATATAGA	TTCTGTTTTT
781	CAGAACTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TAGAGGTTGG	TGCAGAAGGG
841	AATGAACAAG	ATTTCAATTGA	TGTGGTGCCT	TCAAAATTGA	GTAAGAATA	TCCTGATGAA
901	GGTACTCTC	GTGATACTGT	CATTAAGCA	ACAGTTTTTA	GTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTCACAT	AAATTGGGGA	ATGACATTAT	TGATAAACAA	TCAAAATGCC
1021	TTGATGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGGATAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTAGT	ATACCTCCAA	GCTATTGTTA	AAAAGGTGTT	ACGATTATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACATGAA	AATGTAAGG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGGACTAG	ATTATTCGCA	AACGTCAATGA	AACTGCAGCG	CGATCCTAAA
1261	CTCTTGTC	ATCCTGATAA	GTTGATCCA	GAGAGATTCA	TCGCTGGTGA	TATTGACTTC
1321	CGTGGTCACC	ACTATGAGTT	TATCCCATT	GGTTCGGAA	GACGATCTTG	TCCGGGGATG
1381	ACTTATGCAT	TGCAAGTGA	ACACCTAACA	ATGGCACATT	TAATCCAGGG	TTTCAATTAC
1441	AAAACCTCAA	ATGACGAGGC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	AAATACGTAAG
1501	GTAATCCAG	TGGAATTGAT	AATAACGCCT	CGCTTGGCAC	CTGAGCTTTA	CTAAAACCTA
1561	AGATGTTTCA	TCTTGGTTGA	TCATTGT			

SEQ. ID. NO. 227

1	MLSPIEAFVG	LVTFTFLLYF	LWTKKSQKLP	KPLPPKIPGG	WVIGHLHFH	NNDGDDRPLA
61	RKLGDLADKY	GPVFTFRLGL	PLVLVSSYE	AIKDCPSTND	AIFSNRPALL	YGEYLGYNNT
121	MLFLANYGPY	WRKNRKLVIQ	EVLASRLEK	FKQVRFTRIQ	TSIRNLYTRI	NGNSSTINLT
181	DWLEELNFGL	IVKMLAGKNY	ESGRGDEQVE	RFKNAPKDFM	VLSMEFVLWD	APPIPLFKWV
241	DFQGHKAMK	RTPKDIDSVP	QNWLEEHINR	REKIEVGAEG	NEQDFIDVVL	SKLSKEYLDE
301	GYSRDTVIKA	TVFSLVLDAA	DTVALHINWG	MTLLINNQNA	LHKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKVLRLY	PPGPLLVPH	NVKDCVVSFY	HIPKGTRLF	NVMKLRDPK
421	LLSNPKFDP	ERFIAGDIDF	RGHHYEFIPF	GSGRRSCPGM	TYALQVEHLT	MAHLIQFPNY
481	RTPNDEALDM	KEGAGITIRK	VNPVELIITP	RLAPELY		

FIG. 15

NOMBRE D103-AH3
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 230

```

1 ATGGTTTTTC CCATAGAAGC CTTTGTAGGA CTAGTAACCT TCACATTTCT CTTATACTTC
61 CTATGGACAA AAAAATCTCA AAAACTTCCA AAACCCTTAC CACCGAAAAAT CCCC GGAGGA
121 TGGCCCGTAA TCGGCCACCT TTTTCACTTC AATAACGACG GCGACGACCG TCCATTAGCT
181 CGAAAACCTCG GAGACTTAGC TGATAAATAC GGCCCCGTTT TCACCTTTTCG GCTAGGTCTT
241 CCCCTTGTGC TAGTTGTAAG CAGTTACGAA GCTACAAAAG ATTGCTTCTC TACAAATGAC
301 GCCATTTTCT CCAATCGTCC AGCTTTTCTT TACGGCGAAT ACCTTGGCTA CAATAATACA
361 ATGCTTTTTC TAGCAAATTA CCGACCTTAC TGGCGAAAAA ATCGTAAAT AGTCATTACAG
421 GAAGTTCCTCT CTGCTAGTCG TCTCGAAAAA TTCAAACAAG TGAGATTCAC CAGAATTCOA
481 ACGAGCATTAA AGAATTTATA CACTCGAATT AATGGAAAT CGAGTACGAT AAATCTAACT
541 GATTGGTTAG AAGAATTGAA TTTTGGTCTG ATCGTGAAA TGATCGCTGG GAAAAATTAT
601 GAATCCGGTA AAGGAGATGA ACAAGTGGAA AGATTTAAGA ATGCGTTTAA GGATTTTATG
661 GTTTTATCAA TGGAAATTGT ATTATGGGAT GCATTTCOA TTCCATTATT TAAATGGGTG
721 GATTTTCAAG GTCATATTAA GACAATGAAA AGGACATTTA AGGATATAGA TTCTGTTTTT
781 CAGAACTGGT TAGAGGAACA TATTAATAAA AGAGAAAAAA TGGAGGTGG TGCAGAAGGG
841 AATGAACAAG ATTTCAATGA TGTGGTGTCT TCAAATTTGA GTAAAGAATA TCTTGATGAA
901 GGTACTCTC GTGATACTGT CAITAAAGCA ACAGTTTTTA GTTTGGTCTT GGATGCAGCA
961 GACACAGTTG CTCTTCACAT AAATTGGGGA ATGACATTAT TGATAACAA TCAAAATGCC
1021 TTGATGAAAG CACAAGAAGA GATAGACACA AAAGTTGGTA AGGATAGATG GGTAGAAGAG
1081 AGTGATATTA AGGATTTAGT ATACCTCAA GCTATTGTTA AAAAGGTGTT ACGATTATAT
1141 CCACCAGGAC CTTTGTAGT ACCACATGAA AATGTAAGG ATTGTGTTGT TAGTGGATAT
1201 CACATTCCTA AAGGGACTAG ATTATTCGCA AACGTATGA AACTGCAGCG CGATCCTAAA
1261 CTCTTGTCOA ATCTGATAA GTTCGATCCA GAGAGATTCA TCGCTGGTGA TATTGACTTC
1321 CGTGGTCAAC ACTATGAGTT TATCCATCT GGTTCGGAA GACGATCTTG TCCGGGGATG
1381 ACTTATGCAT TGCAAGTGA ACACCTAACA ATGGCACATT TAATCCAGGG TTTCAATTAC
1441 AAAACTCAA ATGACGAGGT CTTGGATATG AAGGAAGGTG CAGGCATAAC AATACGTAAG
1501 GTAAATCCAG TGGAAATTGAT AATAACGCT CGCTTGGCAC CTGAGCTTTA CTA AACCTA
1561 AGATCTTTCA TCTTGGTTGA TCATTGTTA ATA
    
```

SEQ. ID. NO. 231

```

1 MVFPPIEAFVG LVTFTFLLYF LWTKKSQKLP RPLPPKIPGG WPVIGHLPFH NNDGDDRPLA
61 RKLGLDLADKY GPVFTFRLGL PLVLVVSSYE ATKDCPSTND AIFSNRPAPL YGEYLGYNNT
121 MLFLANYGPY WRKNRKLVIQ EVLSASRLEK FKQVRFTRIQ TSIKNLYTRI NGNSSTINLT
181 DWLEELNPGF IVKMIAGKNY ESGKGDEQVE RPKNAPKDFM VLSMEFVLWD AFPPIPLFKWV
241 DFQGHIKIMK RTFKDIDSVF QNWLEEHINK REKMEVGAEG NEQDFLDVVL SKLSKEYLDE
301 GYSRDTVIKA TVFSLVLDAA DTVALHINWG MLLINNQNA LMKRQEEIDT KVGKDRWVEE
361 SDIKDLVYLQ AIVKVLRLY PPGPLLPHE NVKDCVVSFY HIPKGTRLFV NVMKLQRDPK
421 LLSNPKDFDP ERFIAGDIDF RGHYEFIPS GSGRRSCPVM TYALQVEHLT MAHLIQGFNY
481 KTFNDEVLDL KEGAGITIRK VNFVELIITP RLAPELY
    
```


FIG. 16

NOMBRE D208-AD9
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 232

1	ATGCTTTCTC	CCATAGAAGC	CATTGTAGGA	CTAGTAACCT	TCACATTTCT	CTTCTTCTTC
61	CTATGGACAA	AAAAATCTCA	AAAACCTTCA	AAACCCTTAC	CACCGAAAAT	CCCCGGAGGA
121	TGGCCGGTAA	TCGGCCATCT	TTTCCACTTC	AATGACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAACCTCG	GAGACTTAGC	TGACAAATAC	GGCCCCGTTT	TCACTTTTCG	GCTAGGCCTT
241	CCCCTGTGCT	TAGTTGTAAG	CAGTTACGAA	GCTGTAAGA	ACTGTTTCTC	CACAAATGAC
301	GCCATTTTTT	CCAATCGTCC	AGCTTTTCTT	TACGGCGATT	ACCTTGGCTA	CAATAATGCC
361	ATGCTATTTT	TGGCCAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGAAAAAT	AGTTATTTCAG
421	GAAGTTCTCT	CCGCTAGTCC	TCTCGAAAAA	TTCAAACACG	TGAGATTTGC	AAGAATTCAA
481	GCGAGCATGA	AGAAATTATA	TACTCGAATT	GATGGAAAT	CGAGTACGAT	AAATTTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAAGA	TGATCGCTGG	AAAAAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAG	AGATTTAAGA	AAGCGTTTAA	GGATTTTATG
661	ATTTTATCAA	TGGAGTTTGT	GTTATGGGAT	GCATTCCAA	TTCCATTATT	TAAATGGGTG
721	GATTTTCAAG	GGCATGTTAA	GGCTATGAAA	AGGACTTTTA	AAGATATAGA	TTCTGTTTTT
781	CAGAAITGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTTAA	TGCAGAAGGG
841	AATGAACAAG	ATTTCAATTGA	TGTGGTGTCT	TCAAAAATGA	GTAATGAATA	TCTTGGTGAA
901	GGTTACTCTC	GTGATACTGT	CATTGAAGCA	ACGGTGTTTA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTCACAT	AAATGGGGGA	ATGGCATTAT	TGATAAACAA	TCAAAGGCC
1021	TTGACGAAAAG	CACAAGAAGA	GATAGACACA	AAAGTTTGTA	AGGCAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTGGT	ATACCTCCAA	GCTATGTGTA	AAGAAGTGT	ACGATATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACACGAA	AATGTAGAAG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGGACAAG	ATTATTCGCA	AACGTCATGA	AACTGCAACG	TGATCCTAAA
1261	CTCTGGTCTG	ATCCTGATAC	TTTCGATCCA	GAGAGATTCA	TTGCTACTGA	TATTGACTTT
1321	CGTGGTCAGT	ACTATAAGTA	TATCCCCTTT	GGTCTGGAA	GACGATCTTG	TCCAGGGATG
1381	ACTTATGCAT	TGCAAGTGGG	ACACTTAACA	ATGGCACATT	TGATCCAAGG	TTTCAATTAC
1441	AGAACTCCAA	ATGACGAGCC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	TATACGTAAG
1501	GTAARTCTCG	TGGAECTGAT	AATAGCGCCT	CGCCTGGCAC	CTGAGCTTTA	TTAAAACCTA
1561	AGATGTTTCA	TCTTGGTTGA				

SEQ. ID. NO. 233

1	NLSPIEAIVG	LVTFTFLFFF	LWTKKSQKPS	KPLPPKIPGG	WPVIGHLFHF	NDDGDDRPLA
61	RKLGDLADKY	GPVFTFRLGL	PLVLVVSSYE	AVKDCFSTND	AIFSNRPAFL	YGDYLYGYNNA
121	MLFLANYGPFY	WRKNRKLVIQ	EVLSASRLEK	PKHVRFARIQ	ASMKNLRYTRI	DGNSSTINLT
181	DWLEELNFGL	IVKMIAGKNY	ESGKGDEQVE	RFKKAFFKDFM	ILSMFVLWD	APPPIPLFKWV
241	DFQGHVKAMK	RTFKDIDSVF	QNWLEEHINK	REKMEVNAEG	NEQDFIDVVV	SKMSNEYLGE
301	GYSRDTVIEA	TVFSLVLDAA	DTVALHINWG	MALLINNQKA	LTKAQEEIDT	KVCKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRBY	PPGPLLVPHE	NVEDCVVSGY	HIPKGTRLFPA	NVMKLRQDPK
421	LWSDPDTFDP	ERFIATDIDF	RGQYKYIIPF	GPGRRSCPGM	TYALQVEHLT	MAHLIQGFNY
481	RTFNDEPLDM	KEGAGITIRK	VNFVELIIP	RLAPELY		

FIG. 17

NOMBRE D235-AB1
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 254

1	AAAATTCATA	ATGOTTTTTC	CCATAGAAGC	CTTTGTAGGA	CTAGTAACCT	TCACATTTCT
61	CTTATACTTC	CTATGGACAA	AAAAATCTCA	AAAACTTCCA	AAACCCTTAC	TACCGAAAAAT
121	CCCCGGAGGA	TGGCCGGTAA	TCGGCCATCT	TTTTCACTTC	AATAACGACG	GCGACGACCG
181	TCCATTAGCT	CGAAAACCTCG	GAGACTTAGC	TGATAAATAC	GGCCCCGTTT	TCACATTTTCG
241	GCTAGGTCTT	CCCCTTGTGC	TAGTTGTAAG	CAGTTACGAA	GCTATAAAAAG	ATTGCTTCTC
301	TACAAATGAC	GCCATTTTCT	CCAATCGTCC	AGCTTTTCTT	TACGGCGAAT	ACCTTGGCTA
361	CAATAATACA	ATGCTTTTTC	TAGCAAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGTAAATT
421	AGTCAATCAG	GAAGTCTCT	CTGCTAGTCG	TCTCGAAAAA	TTCAAACAAG	TGAGATTCAC
481	CAGAATTCAA	ACGAGCATT	AGAATTTATA	CACCTCGAAT	AATGGAAATT	CGAGTACGAT
541	AAATCTAAC	GATTGGTTAG	AAGAATGGA	TTTTGGTCTG	ATCGTGAAAA	TGATCGCTGG
601	GAATAATTAT	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAA	AGATTTAAGA	ATGCGTTTAA
661	GGATTTTATG	GTTTTATCAA	TGGAATTTGT	ATTATGGGAT	GCATTTCCAA	TCCATTATT
721	TAAATGGGTG	GATTTTCAAG	GTCATATTAA	GGCAATGAAA	AGGACATTTA	AGGATATAGA
781	TTCTGTTTTT	CAGAACTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTTGG
841	TGCAGAAGGG	AATGAACAAG	ATTTCAATGA	TGTGGTGTCT	TCAAAATFGA	GTAAGAATA
901	TCTTGATGAA	GGTTACTCTC	GTGATACTGT	CATTAAGCA	ACAGTTTTTA	GTTTGGTCTT
961	GGATGCAGCA	GACACAGTTG	CTCTCACAT	AAATTGGGGA	ATGACATTAT	TGATAAACAA
1021	TCAAAATGCC	TTGATGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGTATAGATG
1081	GGTAGAAGAG	AGTGATATTA	AGGATTTAGT	ATACCCTCAA	GCTATTGTTA	AAAAGGTGTT
1141	ACGATTATAT	CCACCAGGAC	CTTTGTTAGT	ACCACATGAA	TATGTAAAGG	ATTGTGTTGT
1201	TAGTGGATAT	CACATTCTTA	AAGGGACTAG	ATTATTCGCA	AACGTCAATG	AACTGCAGCG
1261	CGATCCTAAA	CTCTTGTCAA	ATCCTGATAA	GTTCCGATCCA	GAGAGATPCA	TCGCTGGTGA
1321	TATCGACTTC	CGTGGTCACC	ACTATGAGTT	TATCCCATT	GGTCTGGAA	GACGATCTTG
1381	TCCGGGGATG	ACTTATGCAT	TGCAAGTGG	ACACCTAACA	ATGGCACATT	TAATCCAGGG
1441	TTTCAATTAC	AAAACCTCAA	ATGACGAGGC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC
1501	AATACGTAAG	GTAATCCGG	TGGAATTGAT	AATAACGCCT	CGCTTGGCAC	CTGAGCTTTA
1561	CTAAACCTA	AGATCTTCA	TCTTGGTTGA	TCATTGTTTA	ATACTCCTAG	ATAGATGGGT
1621	ATTTCATC					

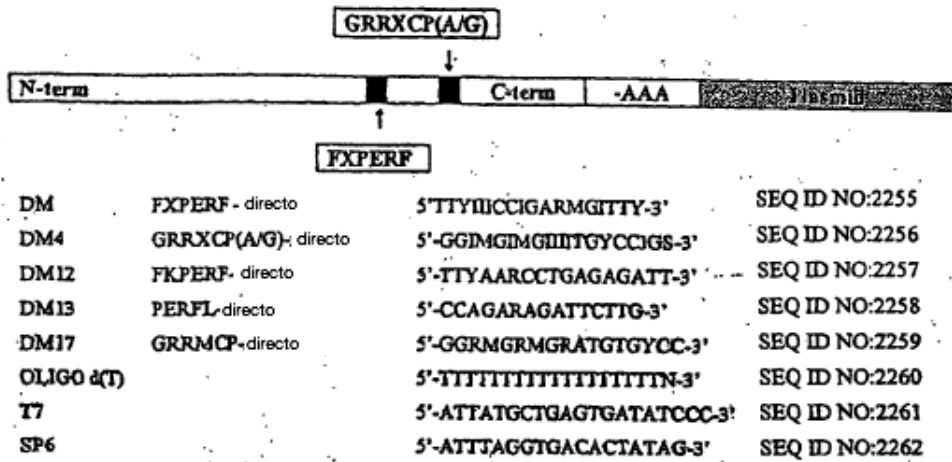
SEQ. ID. NO. 255

1	MVPPIEAFVG	LVTFTFLLYF	LWTKKSQKLP	KPLLPKIPGG	WPVIGHLPHF	NNDGDDRPLA
61	RKLGDLADKY	GPVPTFRLGL	PLVLVSSYE	AIKDCFTND	AIFSNRPAFL	YGEYLGYNNT
121	HLFLANYGPY	WRKNRKLVIQ	EVLSASRLEK	FKQVRFTRIQ	TSIKNLYTRI	NGNSSTINLT
181	DWLEELDFGL	IVKMIAGKNY	ESGKGDEQVE	RPKNAFKDFM	VLSMEFVLWD	AFPIPLFKWV
241	DFQGHKAMK	RTFKDIDSVP	QNWLEEHINK	REKMEVGABG	NEQDFIDVVL	SKLSKEYLDE
301	GYSRDTVIKA	TVFSLVLDRA	DTVALHINWG	MTLLINNOA	LMKAQEEIDT	KVGKYRWVEE
361	SDIKDLVYLQ	AIVKVLRLY	PPGFLVPHF	YVKDCVVSFY	HIPKGTFLFA	NVMKLRDPK
421	LLSNPDKFDP	ERFIAGDIDF	RGHHYEPFP	GSGRRSCPFG	TYALQVEHLT	MAHLIQGFNY
481	KTFNDEALDM	KEGAGITIRK	VNPVELIITP	RLAPELY		

GRUPO 1	Excepción		Excepción	Gx RxC
D208-AD9	98.8	EVLRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2196		
D120-AH4	97.6	EVLRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2197		
D121-AA8	91.6	EVLRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2198		
D122-AF10	91.6	EVLRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2199		
D103-AH3	98.8	KVRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2200		
D208-AC8	98.8	KVRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2201		
D235-AB1		KVRLYPPGP LVPHEVVED CWSGYHIPK GTRLFANVMK LQDPKLWSD EDTFDPERFI ATDIDFRGQY KYIIFGSGR RSC SEQ. ID. No. 2202		

FIGURA 18

FIG. 19 : Clonación de fragmentos de ADNc de citocromo p450 por PCR



I = Desoxilinosina; Y = C, T; M = A, C; R = A, G; S = C, G; N = A, T, C, G

Nombre del grupo de la sonda	Sonda Sonda		Posición de interrogación		Secuencia de la Sonda	SEQ ID
	X	Y	X	Y		
GEN1018_x_at	107	105	58		TATGAATCCGGTAAAGGAGATGAAC	414
GEN1019_x_at	69	1	64		TCCGGTAAAGGAGATGAACAAGTGG	415
GEN1019_x_at	1	15	80		AACAAGTGGAGAGATTTTAAAGAAAGC	416
GEN1019_x_at	111	91	120		GATTTTATCAATGGAGTTTGTGTTA	417
GEN1019_x_at	63	115	142		TTATGGGATGCANTTCCAAATCCAT	418
GEN1019_x_at	54	111	145		TGGGATGCATTTCCAAATCCANTTAT	419
GEN1019_x_at	14	97	147		GGATGCATTTCCAAATCCANTTATTT	420
GEN1019_x_at	35	85	148		GATGCATTTCCAAATCCANTTATTTA	421
GEN1019_x_at	98	117	150		TGCATTTCCAAATCCANTTATTTAAA	422
GEN1019_x_at	96	59	164		CATTAATTTAAATGGTGGATTTTCA	423
GEN1019_x_at	12	23	182		ATTTTCAAGGGCATGTTAAGGCTAT	424
GEN1019_x_at	27	53	187		CAAGGGCAATGTTAAGGCTATGAAA	425
GEN1019_x_at	45	79	195		TGTTAAGGCTATGAAAAGGACTTTT	426
GEN1019_x_at	94	47	359		AAGGTACTCTCGTGATACTGTAT	427
GEN1019_x_at	2	73	418		GCAGACACAGTTGCTCTTTCACATTA	428
GEN1019_x_at	74	83	588		GTTACGATTTATATCCACCAGGACCT	429
GEN1019_x_at	66	3	604		CCAGGACCTTTGTAGTACCACACAG	430
GEN1019_x_at	31	3	610		CCTTTGTAGTACCACACGAAATG	431
GEN1019_x_at	111	11	719		AACCTGGTCTGATCCTGTACTTTT	432
GEN1019_x_at	32	55	775		GACTTTCGTGGTCACTACTATAAGT	433
GEN1019_x_at	102	103	783		TGGTCAGTACTATAAGTATATCCCG	434
GEN2012_x_at	32	7	58		GNCCAAAGGTTTCAATTTACAGACT	435
GEN2012_x_at	119	101	114		GTGCAGGCATTAATACGTAAGGT	436
GEN2012_x_at	21	27	140		AATCCTGTGGAACTGATATAAGCGC	437
GEN2012_x_at	68	21	141		ATCCTGTGGAACTGATATAAGCGCC	438
GEN2012_x_at	108	17	143		CCTGTGGAACTGATATAAGCGCCTC	439
GEN2012_x_at	71	105	145		TGTGGAACTGATATAAGCGCCTGCG	440
GEN2012_x_at	64	89	148		GGAACTGATATAAGCGCCTGCGCCTG	441
GEN2012_x_at	14	77	149		GAACCTGATATAAGCGCCTGCGCCTGG	442

FIGURA 20

Nombre del grupo de la sonda	Sonda		Posición de Interrogación de la Sonda	Secuencia de la Sonda	SEQ ID
	X	Y			
GEN1018_x_at	D121-AA8	107 105	58	TATGAATCCGGTAAAGGAGATGAAC	414
GEN1019_x_at	D121-AA8	69 1	64	TCCGGTAAAGGAGATGAACAAGTGG	415
GEN1019_x_at	D121-AA8	1 15	80	AACAAGTGGAGAGNTTTTAAGAAAGC	416
GEN1019_x_at	D121-AA8	111 91	120	GATTTTATCAATGGGTTTGTGTGTTA	417
GEN1019_x_at	D121-AA8	63 115	142	TTATGGGATGCATTTCCAATTCAT	418
GEN1019_x_at	D121-AA8	54 111	145	TGGGATGCATTTCCAATTCATTTAT	419
GEN1018_x_at	D121-AA8	14 97	147	GGATGCATTTCCAATTCATTTATTT	420
GEN1019_x_at	D121-AA8	35 85	148	GATGCATTTCCAATTCATTTATTTA	421
GEN1019_x_at	D121-AA8	98 117	150	TGCATTTCCAATTCATTTATTTAAA	422
GEN1019_x_at	D121-AA8	96 59	164	CATTTATTTAAAATGGGTGGATTTTCA	423
GEN1019_x_at	D121-AA8	12 23	182	ATTTTCAAGGCATGTTAAGGCTAT	424
GEN1019_x_at	D121-AA8	27 53	187	CAAGGCCATGTTAAGGCTATGAAAA	425
GEN1019_x_at	D121-AA8	45 79	195	TGTTAAGGCATGAAAAGGACTTTT	426
GEN1019_x_at	D121-AA8	94 47	359	AAAGTTACTCTCGTGTACTGTCT	427
GEN1018_x_at	D121-AA8	2 73	418	GCAGACACAGTTGCTCTTCACATAA	428
GEN1019_x_at	D121-AA8	74 83	588	GTTACGATTATATCCACCAGGACCT	429
GEN1019_x_at	D121-AA8	66 3	604	CCAGGACCTTTGTTAGTACCACACCG	430
GEN1019_x_at	D121-AA8	31 3	610	CCTTTGTTAGTACCACACGAAANTG	431
GEN1019_x_at	D121-AA8	111 11	719	AACTCTGGTCTGATCCTGATFACCTTT	432
GEN1019_x_at	D121-AA8	32 55	775	GACTTTGCTGGTCACTACTATFAGT	433
GEN1019_x_at	D121-AA8	102 103	783	TGGTCAGTACTATFAGTATATFCCCG	434
GEN2012_x_at	D35-BG11	32 7	58	GATCCAAGGTTTCAATTTACAGAACT	435
GEN2012_x_at	D35-BG11	119 101	114	GTCCAGGCNTAACTATACGTAAGGT	436
GEN2012_x_at	D35-BG11	21 27	140	AAATCCTGTGGACTGATATAFAGCCG	437
GEN2012_x_at	D35-BG11	66 21	141	ATCCTGTGGAACCTGATATAFAGCCGC	438
GEN2012_x_at	D35-BG11	108 17	143	CCTGTGGAACCTGATATAFAGCCCTC	439
GEN2012_x_at	D35-BG11	71 105	146	TGTGGAACCTGATATAFAGCCCTCGC	440
GEN2012_x_at	D35-BG11	64 99	148	GGAACTGATATAFAGCCCTCGCCTG	441
GEN2012_x_at	D35-BG11	14 77	149	GAACTGATATAFAGCCCTCGCCTGG	442

FIGURA 20

Nombre del Grupo de la Sonda	Sonda		Posición de Interrogación de la Sonda	Secuencia de la Sonda	SEQ ID
	X	Y			
GEN1018_x_at	7	65	14	GAAGCGTTAAGGATTTTATGATT	385
GEN1018_x_at	110	91	35	GATTTTATCAATGGAGTTTGTGTTA	386
GEN1018_x_at	85	115	57	TTATGGGATGCAATTTCCAATTTCCAT	387
GEN1018_x_at	52	111	60	TGGGATGCAATTTCCAATTTCCATTTAT	388
GEN1018_x_at	18	97	62	GGATGCAATTTCCAATTTCCATTTATTT	389
GEN1018_x_at	38	85	63	GATGCAATTTCCAATTTCCATTTATTTA	390
GEN1018_x_at	99	117	65	TGCAATTTCCAATTTCCATTTATTTAAA	391
GEN1018_x_at	18	69	66	GCATTTCCAATTTCCATTTATTTAAAT	392
GEN1018_x_at	59	3	82	TATTTAAATGGGTGGATTTTCAAGG	393
GEN1018_x_at	13	23	97	ATTTTCAAGGGCATTGTTAAGGCTAT	394
GEN1018_x_at	28	53	102	CAAGGGCATTGTTAAGGCTATGAAA	395
GEN1018_x_at	10	93	106	GGCATGTTAAGGCTATGAAAAGGAC	396
GEN1018_x_at	44	78	110	TGTTAAGGCTATGAAAAGGACTTTT	397
GEN1018_x_at	95	47	274	AAGGTTACTCTCGTATGACTGTGAT	398
GEN1018_x_at	1	73	333	GCAGACACAGTTGCTTTCACATAA	399
GEN1018_x_at	63	41	338	ACAGTTGCTTTCACATAAATTTGGG	400
GEN1018_x_at	84	81	473	GGTATACCTCCANGCTATTTGTTAAA	401
GEN1018_x_at	25	95	501	GTGTTACGATTATATCCACCAGGAC	402
GEN1018_x_at	69	21	505	TACGATTATATCCACCAGGACCTTT	403
GEN1018_x_at	67	3	518	CCAGGACCTTTGTTAGTACCACACG	404
GEN1018_x_at	30	3	525	CCTTTGTTAGTACCACACGMAAATG	405
GEN1018_x_at	1	107	622	TACGTGATCCTAAACTCTGGCCTGA	406
GEN1018_x_at	72	19	643	CTGATCCTGATACTTTTCGATCCAGA	407
GEN1018_x_at	33	55	690	GACTTTCGTTGGTCACTACTATTAAGT	408
GEN1018_x_at	103	103	698	TGGTCAGTACTATTAAGTATATCCCG	409
GEN1018_x_at	107	73	13	GAATTTGAATTTTGGTCTGATCGTGA	410
GEN1018_x_at	65	107	24	TGGTCTGATCGTGTGAGATGATCGCT	411
GEN1018_x_at	119	25	31	ATCGTGAAGATGATCGCTGGAAAAA	412
GEN1018_x_at	72	71	39	GATGATCGCTGGAAAAAATTTATGAA	413

FIGURA 20