

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 534 203**

51 Int. Cl.:

C12N 9/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.04.2005 E 05778860 (6)**

97 Fecha y número de publicación de la concesión europea: **18.03.2015 EP 1751278**

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:
20.04.2015

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

ES 2 534 203 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 relaciona con un cultivo de tejidos de células de tabaco regenerables que tienen una mutación en un gen de nicotina desmetilasa endógeno, en donde dicha planta de tabaco regenerada exhibe expresión reducida del gen de nicotina desmetilasa mutado o producto genético, o la nicotina desmetilasa codificada por el gen mutado exhibe actividad enzimática reducida, como se define en las reivindicaciones. Adicionalmente, la presente invención es pertinente a un método para reducir la expresión o actividad enzimática de una nicotina desmetilasa en una planta de tabaco, y a un método para producir un producto de tabaco, como se define en las reivindicaciones.

10 Antecedentes

Durante la maduración o curado del tabaco la expresión de diversos genes es alterada. Tales genes pueden afectar las rutas metabólicas involucradas en la formación de numerosos metabolitos secundarios incluyendo terpenoides, polifenoles y alcaloides que afectan las características de calidad del producto final. Por ejemplo, la bioconversión de la nicotina para formar nornicotina durante la senescencia de la planta y en el postcultivo o fase de curado de las hojas se presentan muchas especies de *Nicotiana*. La nicotina es la fuente predominante 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 y subsecuentemente durante el almacenamiento y procesamiento de la hoja.

Los genes expresados durante la maduración o curado del tabaco pueden ser expresados constitutivamente, inducidos por etileno o por genes relacionados con la senescencia, por ejemplo, genes que codifican un citocromo p450. Los citocromos p450, por ejemplo, catalizan las reacciones enzimáticas para un rango diverso de sustratos químicamente disímiles que incluyen el metabolismo oxidativo, peroxidativo y reductor 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 p450 hemotiolato, usualmente actúan como oxidasas terminales en cadenas de transferencia de electrones de componentes múltiples, denominadas sistemas de monooxigenasa que contienen p450. Reacciones específicas catalizadas por estos sistemas de enzimas incluyen desmetilación, hidroxilación, epoxidación, N-oxidación, sulfoxidación, y N-, S- y O- desalquilaciones, desulfonación, desaminación y reducción de grupos azo, nitro y N-óxido.

El papel diverso de las enzimas p450 de las plantas de *Nicotiana* ha estado implicado en la ejecución de una variedad de metabolitos vegetales tales como fenilpropanoides, alcaloides, terpenoides, lípidos, glicósidos cianogénicos, glucosinolatos, y una hueste de otras entidades químicas. Algunas enzimas p450 pueden impactar la composición de los metabolitos vegetales. Por ejemplo, se ha deseado durante mucho tiempo mejorar el sabor y aroma de ciertas plantas alterando un perfil de la planta de ácidos grasos seleccionados a través de cruzamientos; sin embargo, se sabe muy poco acerca de los mecanismos involucrados en el control de los niveles de esos constituyentes de las hojas. La subregulación o sobrerregulación de las enzimas p450 asociadas con la modificación de ácidos grasos puede facilitar la acumulación de ácidos grasos deseados que proveen cualidades fenotípicas de las hojas más preferidas.

La función de las enzimas p450 y sus papeles crecientes en los constituyentes vegetales están siendo aún descubiertos. Por ejemplo, se encontró que una clase especial de enzimas p450 catalizan la ruptura de ácidos grasos en aldehídos volátiles C6 y C9 y β -alcoholes que contribuyen de manera principal al olor "verde fresco" de frutas y vegetales. El nivel de otros p450 objetivo novedosos puede ser alterado para potenciar las cualidades de los constituyentes de las hojas modificando la composición en lípidos y en metabolitos de ruptura relacionados en las hojas de *Nicotiana*. Varios de estos constituyentes en la hoja son afectados por la senescencia que estimula la maduración de propiedades de calidad de las hojas. Todavía otros reportes han mostrado que las enzimas p450 juegan un papel funcional en la alteración de los ácidos grasos que están involucrados en las interacciones planta-patógeno y resistencia a enfermedades.

La gran multiplicidad de las formas de la enzima p450, sus estructuras y funcionamiento diferentes han hecho que su investigación acerca de las enzimas p450 de la *Nicotiana* sea muy difícil antes de la presente invención. Además, la clonación de las enzimas p450 ha sido obstaculizada al menos en parte porque estas proteínas localizadas en la membrana están presentes típicamente en baja abundancia y frecuentemente son inestables durante la purificación. Por lo tanto, existe una necesidad para identificar las enzimas p450 en plantas y las secuencias de ácidos nucleicos asociadas con estas enzimas p450. En particular, solo se han reportado unas pocas proteínas de citocromo p450 y *Nicotiana*. La divulgación descrita aquí abarca el descubrimiento de los citocromos 450 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 divulgación abarca el descubrimiento de otras secuencias constitutivas e inducidas por etileno o senescencia que apuntan a la necesidad de regular las rutas metabólicas involucradas 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 vegetales que tienen características deseables para utilización en programas de cruzamiento para desarrollar germoplasmas más deseables, y especialmente germoplasmas tipo no GMO (organismo genéticamente modificado).

Resumen de la invención

Los presentes inventores han identificado y caracterizado secuencias constitutivas, e inducidas por etileno y senescencia, incluyendo un clon genómico de la nicotina desmetilasa, del tabaco. También se describe aquí el uso de estas secuencias en métodos de cruzamiento y en 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 con respecto a una planta de control.

Una planta de tabaco puede ser identificada como una variante para la expresión del gen de nicotina desmetilasa, por ejemplo, a los niveles transcripcional, postranscripcional y de traducción o a nivel de la actividad enzimática) utilizando las secuencias descritas aquí y métodos estándar conocidos en el arte.

La planta de tabaco induce un gen endógeno de nicotina desmetilasa que tiene una mutación (por ejemplo, una eliminación, sustitución, mutación puntual, translocación, inversión, duplicación o una inserción). La planta de tabaco puede incluir un gen de nicotina desmetilasa que tiene una mutación nula, incluye un gen recombinante que silencia un gen de nicotina desmetilasa endógeno, o incluye una nicotina desmetilasa que tiene actividad enzimática reducida o alterada. El gen de nicotina desmetilasa de la planta de tabaco puede estar ausente. La planta de tabaco puede ser una planta transgénica.

La planta de tabaco incluye una molécula de ácido nucleico endógena seleccionada del grupo consistente de las secuencias de ácidos nucleicos mostradas en la Figura 1, Figuras 3 a 7 y Figuras 10-17, en donde el ácido nucleico incluye una mutación. Mutaciones de ejemplo incluyen eliminación, sustitución, mutación puntual, translocación, inversión, duplicación o una inserción.

La primera planta de tabaco del método antes descrito puede incluir una molécula de ácido nucleico endógena seleccionada del grupo consistente de las secuencias de ácidos nucleicos mostradas en la Figura 1, Figuras 3 a 7 y Figura 10-17 en donde el ácido nucleico incluye una mutación nula. La planta de tabaco puede incluir un gen recombinante que silencia la expresión de la molécula de ácido nucleico endógena una molécula de ácido nucleico endógena seleccionada del grupo consistente de las secuencias de ácidos nucleicos mostradas en la Figura 1, Figuras 3 a 7, y Figuras 10-17.

La planta de tabaco incluye, si se desea, una molécula de ácido nucleico endógena seleccionada del grupo consistente de las secuencias de ácidos nucleicos mostradas en la Figura 1, Figuras 3 a 7 y Figuras 10 a 17, en donde la molécula de ácido nucleico codifica un polipéptido que tiene actividad enzimática reducida o alterada. En todavía otras realizaciones, la planta de tabaco es una planta transgénica.

Las plantas de tabaco de ejemplo divulgadas aquí incluyen *Nicotiana africana*, *Nicotiana amplexicaulis*, *Nicotiana arentsii*, *Nicotiana benthamiana*, *Nicotiana bigelovii*, *Nicotiana corymbosa*, *Nicotiana debneyi*, *Nicotiana excelsior*, *Nicotiana exigua*, *Nicotiana glutinosa*, *Nicotiana goodspeedii*, *Nicotiana gossei*, *Nicotiana hesperis*, *Nicotiana ingulba*, *Nicotiana knightiana*, *Nicotiana maritima*, *Nicotiana megalosiphon*, *Nicotiana miersii*, *Nicotiana nesophila*, *Nicotiana noctiflora*, *Nicotiana nudicaulis*, *Nicotiana otophora*, *Nicotiana palmeri*, *Nicotiana paniculata*, *Nicotiana petunioides*, *Nicotiana pharmbaginifolia*, *Nicotiana repanda*, *Nicotiana rosulata*, *Nicotiana rotundifolia*, *Nicotiana rustica*, *Nicotiana setchelli*, *Nicotiana stocktonii*, *Nicotiana eastii*, *Nicotiana suaveolens* o *Nicotiana trigonophylla*. Otras plantas de tabaco incluyen variedades de *Nicotiana tabacum* o *Nicotiana rustica*. Todavía otra planta de tabaco es un tabaco negro oriental, tabaco amarillo o curado al aire, Virginia, o una planta de tabaco Burley.

En todavía un aspecto relacionado adicionalmente, la invención presenta un cultivo de tejidos de células de tabaco regenerables obtenidas de acuerdo con los métodos descritos aquí, como se define en las reivindicaciones. Tales cultivos de tejidos regeneran plantas de tabaco capaces de expresar todas las características fisiológicas y morfológicas de la planta de tabaco que tiene la expresión variante del gen de nicotina desmetilasa o un atributo modificado. Células regenerables de ejemplo son embriones, células meristemáticas, semillas, polen, hojas, raíces, puntas de raíces, o flores o son protoplastos de callus derivadas de las mismas.

En todavía aspectos relacionados, la invención presenta un método para producir un producto de tabaco como se define en las reivindicaciones.

Productos de tabaco de ejemplo incluyen hojas o tallos o ambos; un producto de tabaco sin humo; un rapé húmedo o seco; un tabaco para masticar; productos de cigarrillos; productos de cigarrillos; cigarrillos; tabacos para pipa; o bidis.

- Un aspecto adicional de la divulgación presenta un método para reducir la expresión o actividad enzimática de un polipéptido de tabaco constitutivo, o inducido por etileno o inducido por senescencia en una planta vegetal. Este método involucra la reducción del nivel o actividad enzimática de un polipéptido de tabaco constitutivo endógeno, o inducido por etileno o senescencia en la célula vegetal. En una realización deseable, el polipéptido de tabaco es un p450. En otras realizaciones deseables, la célula vegetal es de una especie de *Nicotiana*.
- La reducción del nivel del polipéptido de tabaco constitutivo, o inducido por etileno o senescencia puede involucrar la expresión de un transgen que codifica una molécula de ácido nucleico antisentido de una secuencia de ácidos nucleicos mostrada en la Figura 1, Figuras 3 a 7, y Figuras 10-17 o una secuencia de ácidos nucleicos que codifica un polipéptido que contiene una secuencia de aminoácidos mostrada en la Figura 1, Figuras 3 y 4, y Figuras 10-17 en la célula vegetal. El transgen puede codificar una molécula de ARN de cadena doble de un ácido nucleico de tabaco constitutivo, o inducido por etileno o senescencia o una secuencia de aminoácidos en la célula vegetal. El transgen puede ser expresado, por ejemplo, en una forma específica para el tejido, específica para la célula o específica para el órgano. Además, la reducción del nivel del polipéptido de tabaco constitutivo, o inducido por etileno o senescencia involucra de manera deseable la cosupresión del polipéptido de tabaco constitutivo, o inducido por etileno o senescencia en la célula vegetal. Al reducir el nivel del polipéptido de tabaco constitutivo, o inducido por etileno o senescencia se puede involucrar la expresión de un gen negativo dominante producido en la célula vegetal. De manera deseable, el polipéptido de tabaco constitutivo, o inducido por etileno o senescencia incluye una mutación en un gen que codifica una secuencia de aminoácidos mostrada en la Figura 1, Figuras 3 y 4, y Figuras 10 a 17. La expresión reducida puede ocurrir a nivel transcripcional, a nivel translacional, o a nivel postranslacional.
- Un aspecto adicional presenta una molécula 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 una nicotina desmetilasa de tabaco que contiene una secuencia de nucleótidos que es al menos 70% idéntica a la secuencia de nucleótidos de SEQ ID NO: 4 o SEQ ID NO: 5, o que contiene los nucleótidos 2010-2949 y/o 3947-4562 de SEQ ID NO: 4, o que contiene la secuencia de SEQ ID NO: 4 o SEQ IDE NO:5. La molécula de ácido nucleico aislada, por ejemplo, está enlazada operativamente 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. De manera deseable, la célula vegetal, tal como una célula de planta de tabaco, está incluida en una planta. La divulgación también representa una semilla, por ejemplo una semilla de tabaco, de una planta que contiene el vector de expresión, en donde la semilla incluye una molécula de ácido nucleico aislada que hibrida bajo condiciones restrictivas a la secuencia de SEQ ID NO: 4 enlazada operativamente a una secuencia promotora heteróloga. Adicionalmente, la divulgación presenta una planta derivada de una célula germinada que contiene el vector de expresión, una hoja, bien sea verde o curada, de la planta, y un artículo de manufactura hecho a partir de la hoja.
- La secuencia de nucleótidos puede contener una secuencia que hibrida bajo condiciones restrictivas al complemento de la secuencia de nucleótidos de SEQ ID NO: 4 y/o SEQ ID NO: 5, o a un fragmento de SEQ ID NO: 4 o SEQ ID NO: 5. De manera deseable, la secuencia de nucleótidos codifica una nicotina desmetilasa que es sustancialmente idéntica a la secuencia de aminoácidos de SEQ ID NO: 3. La nicotina desmetilasa puede tener al menos 70% de secuencia de identidad de aminoácidos con respecto a la secuencia de aminoácidos de la nicotina desmetilasa de SEQ ID NO: 3 o a un fragmento de una nicotina desmetilasa que tiene actividad enzimática alterada (por ejemplo, reducida) en comparación con el polipéptido de longitud completa. De manera deseable, la nicotina desmetilasa incluye la secuencia de aminoácidos de SEQ ID NO: 3.
- En otro aspecto, la divulgación presenta una molécula de ácido nucleico aislada que contiene un promotor que hibrida bajo condiciones de restricción a la secuencia de SEQ ID NO: 8, o un fragmento de la misma que conduce la transcripción. De manera deseable, el promotor (i) es inducido siguiendo tratamiento con etileno o durante la senescencia; y (ii) incluye (a) pares de bases 1-2009 de SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivas idénticos a 200 pares de bases consecutivas de la secuencia definida por los pares de bases 1-2009 de SEQ ID NO: 4, o (c) una porción de nucleótido de 20 pares de bases idéntica en secuencia a una porción de 20 pares de bases consecutivas de la secuencia fijada en los pares de bases 1-2009 de SEQ ID NO: 4.
- Un aspecto adicional presenta un promotor de ácido nucleico aislado que contiene una secuencia de nucleótidos que tiene 50% o más de identidad de secuencia con la secuencia de SEQ ID NO: 8. De manera deseable, este promotor de ácido nucleico aislado es inducido siguiendo el tratamiento con etileno o durante la senescencia y, por ejemplo, incluye la secuencia de SEQ ID NO: 8. Alternativamente, el promotor puede incluir un fragmento obtenible de SEQ ID NO: 8, en donde el fragmento guía la transcripción de un gen heterólogo o reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión del gen). En una realización deseable la secuencia del promotor está enlazada operativamente a una secuencia de ácido nucleico heteróloga, y puede, por ejemplo, estar contenida en un vector de expresión. En otras realizaciones deseables el vector de expresión está contenido en una célula, por ejemplo, una célula vegetal. De manera deseable, la célula vegetal, tal como una célula de 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 en donde la semilla incluye una molécula de ácido nucleico aislada que hibrida bajo condiciones de restricción a la secuencia de SEQ ID NO: 8

enlazado operativamente a la secuencia de ácidos nucleicos heteróloga. Adicionalmente, la divulgación presenta una planta derivada de una semilla germinada que contiene el promotor de este aspecto, una hoja, bien sea verde o curada, de la planta, y un artículo de manufactura hecho a partir de la hoja.

5 En aún otro aspecto, la descripción 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 SEQ ID NO: 8 o un fragmento obtenible de SEQ ID NO: 8 enlazado operativamente a una secuencia de ácidos nucleicos heteróloga 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).

10 En un aspecto deseable adicional, la divulgación presenta una molécula de ácido nucleico aislada que contiene un intrón que hibrida bajo condiciones de restricción a la secuencia de 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 marcador molecular para identificar secuencias de ácido nucleico de la nicotina desmetilasa. En una realización deseable, el intrón incluye (a) pares de bases 2950-3946 de SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivos idénticos a 200 pares de bases consecutivos de la secuencia definida por los pares de bases 2950-3946 de 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 consecutivos de la secuencia fijada en los pares de bases 2950-3946 de SEQ ID NO: 4.

20 Otro aspecto deseable presenta un intrón de ácido nucleico aislado que incluye una secuencia de nucleótidos que tiene 50% o más de identidad de secuencia con la secuencia de SEQ ID NO: 7, o un fragmento del mismo 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 ácidos nucleicos de la nicotina desmetilasa. El silenciamiento de la expresión del gen puede involucrar, por ejemplo, recombinación homóloga (por ejemplo utilizando la secuencia de SEQ ID NO: 188 o un fragmento de la misma) o una mutación que da como resultado un producto de gen que no tiene actividad de nicotina desmetilasa. En particular, el intrón puede incluir la secuencia de SEQ ID NO: 7 o un fragmento obtenible de SEQ ID NO: 7. De manera deseable, una molécula de ácido nucleico aislada que incluye un intrón esta enlazada operativamente a una secuencia de ácidos nucleicos heteróloga 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 vegetal puede ser una célula de tabaco. Es deseable una planta, por ejemplo, una planta de tabaco, que incluye una célula vegetal que contiene la secuencia de SEQ ID NO: 7 o un fragmento obtenible de SEQ ID NO: 7 enlazado operativamente a una secuencia de ácidos nucleicos heteróloga en un vector de expresión. Adicionalmente, una semilla, por ejemplo una semilla de tabaco, de una planta, en donde la semilla contiene un intrón que hibrida bajo condiciones restrictivas a SEQ ID NO: 7 enlazado operativamente a una secuencia de ácidos nucleicos heteróloga también es deseable.

35 Adicionalmente, la divulgación presenta una planta derivada de la semilla germinada que contiene el intrón de este aspecto, una hoja, bien sea verde o curada, de la planta, y un artículo de manufactura hecho a partir de la hoja verde o curada.

40 En aún otro aspecto, la divulgación 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 SEQ ID NO: 7 o un fragmento obtenible de SEQ ID NO: 7 enlazado operativamente a una secuencia de ácidos nucleicos heteróloga 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).

45 En un aspecto adicional, la divulgación presenta una molécula de ácido nucleico aislada que contiene una región no traducida que hibrida bajo condiciones restrictivas a la secuencia de 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 del gen) o puede ser utilizada como un marcador para identificar las secuencias de ácidos nucleicos de la nicotina desmetilasa. En una realización deseable de este aspecto, la región no traducida incluye (a) pares de bases 4563-6347 de SEQ ID NO: 4, o (b) al menos 200 pares de bases consecutivos idénticos a 200 pares de bases consecutivos de la secuencia definida por los pares de bases 4563-6347 de 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 fijada en los pares de bases 4563-6347 de SEQ ID NO: 4.

55 Un aspecto deseable adicional presenta una región no traducida de ácido nucleico aislado que contiene una secuencia de nucleótidos que tiene 50% o más de identidad de secuencia con la secuencia de SEQ ID NO: 9. De manera deseable, la región no traducida incluye la secuencia de SEQ ID NO: 9 o la región no traducida incluye un fragmento obtenible de 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 del gen) o puede ser utilizada como marcador para identificar las secuencias de ácidos nucleicos de la nicotina desmetilasa. La región no traducida de manera deseable está enlazada operativamente a una secuencia de ácidos nucleicos heteróloga y puede estar contenida en un vector de expresión. Adicionalmente, este vector de expresión está contenido de manera deseable

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 de ácidos nucleicos aislada que tiene 50% o más de identidad de secuencia con la secuencia de SEQ ID NO: 9 y esta enlazada operativamente a una secuencia de ácidos nucleicos heteróloga.

5 La divulgación también presenta una semilla, por ejemplo, una semilla de tabaco, de una planta, en donde la semilla incluye una región no traducida que hibrida bajo condiciones restrictivas a SEQ ID NO: 9 enlazado operativamente a una secuencia de aminoácidos heteróloga. Adicionalmente, esta divulgación presenta una planta derivada de una semilla germinada que contiene la región no traducida de este aspecto de la invención, una hoja, bien sea verde o curada, de la planta, y un artículo de manufactura hecho a partir de la hoja verde o curada.

10 Adicionalmente, la invención presenta un método para reducir la expresión o alterar la actividad enzimática de la nicotina desmetilasa en una planta de tabaco, tal como se define en las reivindicaciones.

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, en donde el vector es capaz de dirigir la expresión de la nicotina desmetilasa codificada por la molécula de ácido nucleico aislada. De manera deseable, el vector
15 incluye la secuencia de SEQ ID NO: 4 o SEQ ID NO: 5.

En un aspecto adicional, la divulgación presenta una planta de tabaco que tiene expresión reducida de una secuencia de ácidos nucleicos que codifica un polipéptido, por ejemplo, una que incluye la secuencia de SEQ ID NO: 3, y que desmetila la nicotina, en donde la expresión reducida (o una reducción en la actividad enzimática) reduce el nivel de nicotina en la planta. En una realización deseable, la planta de tabaco es una planta transgénica, tal como
20 una que incluye un transgen que, cuando es expresado en la planta transgénica, silencia la expresión genética de una nicotina desmetilasa de tabaco endógena.

En particular, la planta transgénica incluye de manera deseable uno o más de los siguientes: un transgen que expresa una molécula antisentido de una nicotina desmetilasa de tabaco o una molécula de ARN capaz de inducir ARN de interferencia (ARNi); un transgen, que cuando se expresa en la planta transgénica, cosuprime la expresión
25 de una nicotina desmetilasa de tabaco; un transgen que codifica un producto de gen negativo dominante, por ejemplo, una forma mutada de la secuencia de aminoácidos de SEQ ID NO: 3; una mutación puntual en un gen que codifica la secuencia de aminoácidos de SEQ ID NO: 3; una eliminación en un gen que codifica una nicotina desmetilasa de tabaco; y una inserción en un gen que codifica una nicotina desmetilasa de tabaco.

En otras realizaciones deseables, la expresión reducida de una secuencia de ácidos nucleicos que codifica un polipéptido ocurre a nivel transcripcional, a nivel de traducción, o al nivel de postraducción.
30

Un aspecto adicional presenta un método para reducir la expresión de nicotina desmetilasa de tabaco en una planta o en un componente de la planta. Este método involucra las etapas de: (a) introducir en células vegetales un transgen que codifica una nicotina desmetilasa de tabaco enlazado operativamente a un promotor funcional en las células vegetales para producir células vegetales transformadas; (b) regenerar una planta o un componente de
35 planta de las células vegetales transformadas, en donde la nicotina desmetilasa de tabaco es expresada en las células de la planta o componente de la planta, reduciendo por lo tanto la expresión de la nicotina desmetilasa de tabaco en una planta o componente de la planta. En realizaciones particulares de este aspecto, el transgen que codifica la nicotina desmetilasa de tabaco es expresado constitutivamente o expresado de manera inducible, por ejemplo, en una forma específica para el tejido, específica para la célula o específica para el órgano. En otra realización de este aspecto, la expresión del transgen cosuprime la expresión de una nicotina desmetilasa de tabaco endógena o cualquier otro polipéptido descrito aquí.
40

Un aspecto adicional de la divulgación presenta otro método para reducir la expresión de nicotina desmetilasa de tabaco o cualquiera de los otros polipéptidos descritos aquí en una planta o componente de una planta. Este método incluye las etapas de: (a) introducir en plantas vegetales un transgen que codifica una secuencia de codificación antisentido de una nicotina desmetilasa de tabaco a una molécula de ARN capaz de inducir ARN de interferencia (ARNi) enlazado operativamente a un promotor funcional en las células vegetales para producir células vegetales transformadas; (b) regenerar una planta o un componente de la planta a partir de las células vegetales transformadas, en donde la molécula antisentido o una de ARN capaz de inducir ARN de interferencia (ARNi) de la secuencia de codificación de la nicotina desmetilasa de tabaco es expresada en las células de la planta o un
45 componente de la planta, reduciendo por lo tanto la expresión de la nicotina desmetilasa de tabaco en una planta o componente de la planta. De manera deseable, el transgen que codifica una secuencia antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi) y una nicotina desmetilasa de tabaco es expresada constitutivamente o es expresada de manera inducible, por ejemplo, en una forma específica para el tejido, específica para la célula o específica para el órgano. En otra realizaciones deseables la molécula antisentido o de ARN capaz de inducir ARNi de la secuencia de codificación de la nicotina desmetilasa de tabaco contiene el complemento de SEQ ID NO: 1, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 61, SEQ ID NO: 188, o un fragmento de las mismas.
50
55

Un aspecto adicional presenta aún otro método para reducir la expresión de nicotina desmetilasa de tabaco en una planta o componente de la planta. Este método involucra las etapas de: (a) introducir en las células vegetales un transgen que codifica un producto de gen negativo dominante de una nicotina desmetilasa de tabaco enlazado operativamente a un promotor funcional en las células vegetales para producir células vegetales transformadas; y (b) regenerar una planta o componente de la planta a partir de las células vegetales transformadas, en donde el producto de gen negativo dominante de la nicotina desmetilasa de tabaco es expresado en las células de la planta o componente de la planta, reduciendo por lo tanto la expresión de la nicotina desmetilasa de tabaco en una planta o componente de la planta. En realizaciones particulares de este aspecto, el transgen que codifica el producto de gen negativo dominante es expresado constitutivamente o es expresada de manera inducible, por ejemplo, en una forma específica para el tejido, específica para la célula o específica para el órgano.

Un aspecto adicional presenta un método adicional para reducir la expresión o la actividad enzimática de la nicotina desmetilasa de tabaco en una célula vegetal. Este método involucra reducir el nivel de una nicotina desmetilasa de tabaco, o su actividad enzimática, en la célula vegetal. De manera deseable, la célula vegetal es de una dicotiledónea, una planta solanácea, o una especie de *Nicotiana*. En realizaciones deseables de este aspecto, la reducción de nivel de nicotina desmetilasa en el tabaco endógena involucra la expresión de un transgen que codifica una molécula de ácido nucleico antisentido o una molécula de ARN capaz de inducir ARN de interferencia (ARNi) de una nicotina desmetilasa de tabaco en una célula vegetal, o involucra la expresión de un transgen que codifica una molécula de ARN de cadena doble de una nicotina desmetilasa de tabaco en la célula vegetal. De manera deseable, el ARN de cadena doble es una secuencia de ARN correspondiente a la secuencia de SEQ ID NO: 1; SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 61, SEQ ID NO: 188 o un fragmento de las mismas. En una realización adicional, la reducción de nivel de nicotina desmetilasa de tabaco endógena involucra la cosupresión de la nicotina desmetilasa de tabaco endógena en la célula vegetal o involucra la expresión de un producto de gen negativo dominante en la célula vegetal. En particular, el producto de gen negativo dominante puede incluir un gen que codifica una forma mutada de la secuencia de aminoácidos de SEQ ID NO: 3 o cualquier otra secuencia de aminoácidos descrita aquí.

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

La nicotina desmetilasa tiene al menos 70% de identidad con la secuencia de aminoácidos de SEQ ID NO: 3. De manera deseable, el compuesto hace disminuir la actividad de la nicotina desmetilasa de tabaco.

Un aspecto adicional presenta una planta o componente de planta de tabaco curado 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. De manera deseable, el componente vegetal es hoja de tabaco o tallo de tabaco. En una realización deseable, la nitrosamina es nicotina, y el contenido de nicotina de manera deseable es menos de 5 mg/g, 4.5 mg/g, 4.0 mg/g, 3.5 mg/g, 3.0 mg/g, más deseablemente menor de 2.5 mg/g, 2.0 mg/g, 1.5 mg/g, 1.0 mg/g, más deseablemente menor de 750 µg/g, 500 µg/g, 250 µg/g, 100 µg/g, aun más deseablemente menor de 75 µg/g, 50 µg/g, 25 µg/g, 10 µg/g, 7.0 µg/g, 5.0 µg/g, 4.0 µg/g, y aun más deseablemente menor de 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 en la misma es menor de 90%, 70%, 50%, 30%, 10%, de manera deseable menos de 5%, 4%, 3%, 2%, 1.5%, 1%, y más deseablemente menos de 0.75%, 0.5%, 0.25%, o 0.1%. En otra realización deseable, la nitrosamina es N'-nitrosonomicotina (NNN), y el contenido de N'-NNN es de manera deseable menor de 5 mg/g, 4.5 mg/g, 4.0 mg/g, 3.5 mg/g, 3.0 mg/g, más deseablemente menor de 2.5 mg/g, 2.0 mg/g, 1.5 mg/g, 1.0 mg/g, más deseablemente menor de 750 µg/g, 500 µg/g, 250 µg/g, 100 µg/g, aun más deseablemente menor de 75 µg/g, 50 µg/g, 25 µg/g, 10 µg/g, 7.0 µg/g, 5.0 µg/g, 4.0 µg/g, y aun más deseablemente menor de 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 de la misma es menor de 90%, 70%, 50%, 30%, 10%, de manera deseable menos de 5%, 4%, 3%, 2%, 1.5%, 1%, y de manera más deseable menos de 0.75%, 0.5%, 0.25%, o 0.1%. En realizaciones deseables adicionales de este aspecto, la planta o componente de planta de tabaco curada es un tabaco oscuro, tabaco Burley, tabaco curado amarillo, tabaco Virginia curado al aire o tabaco oriental.

Adicionalmente, la planta o componente de planta de tabaco curada de manera deseable incluye un gen de nicotina desmetilasa recombinante, por ejemplo, uno que contiene la secuencia de SEQ ID NO: 4 o SEQ ID NO: 5 o un fragmento del mismo. De manera deseable, la expresión de un gen de nicotina desmetilasa endógeno, o de cualquier otra secuencia de aminoácidos descrita aquí, en la planta o componente de la planta de tabaco curada es silenciada.

Otro aspecto presenta un producto de tabaco que contiene una planta o componente de la planta de tabaco curado que incluye (i) expresión reducida de una nicotina desmetilasa o cualquier otro polipéptido descrito aquí o (ii) una

nicotina desmetilasa u otro polipéptido descrito aquí que tiene actividad alterada, y una cantidad reducida de nitrosamina. De manera deseable, el producto de tabaco es tabaco sin humo, rapé húmedo o seco, un tabaco para masticar, productos de cigarros, productos de cigarrillos, tabacos para pipa o bidis. En particular, el producto de tabaco de este aspecto puede contener tabaco oscuro, tabaco molido o incluir un componente saborizante.

- 5 La invención también presenta un método para hacer 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 involucra proveer una planta o componente de planta de tabaco curada 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 la preparación del producto de tabaco a partir de la planta o componente de la planta de tabaco curada, como se define en las reivindicaciones.

Definiciones

- 15 “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 grupos azo, nitro, N-óxido, y tales otros grupos químicos enzimáticamente reactivos. La actividad enzimática alterada se refiere a un descenso en la actividad enzimática (por ejemplo, de una nicotina desmetilasa de tabaco) en al menos 10-20%, preferiblemente en al menos 25-50%, y más preferiblemente en al menos 55-95% o más con respecto a la actividad de una enzima de control (por ejemplo, una nicotina desmetilasa de tabaco de una planta de tabaco tipo silvestre). La actividad de una enzima, tal como una nicotina desmetilasa puede ser probada utilizando métodos estándar en la técnica, por ejemplo, utilizando ensayos de microsoma de levadura descritos aquí.

- 25 El término “ácido nucleico” se refiere a un polímero de desoxirribonucleótido o ribonucleótido bien sea en forma de cadena sencilla o doble, o sentido o antisentido, y a menos que se indique otra cosa, abarca análogos conocidos de nucleótidos naturales que hibridan a ácidos nucleicos de una manera similar a los nucleótidos de origen natural. A menos que se indique otra cosa, una secuencia de ácidos nucleicos particular incluye la secuencia complementaria de la misma. Los términos “enlazado operativamente”, “en combinación operable” y “en orden operable” se refieren a un enlace funcional entre una secuencia de control de expresión de ácidos nucleicos (tal como un promotor, secuencia de señalización, o arreglo de sitios de enlazamiento del factor de transcripción) y una segunda secuencia de ácidos nucleicos, en donde la secuencia de control de expresión afecta la transcripción y/o traducción del correspondiente ácido nucleico a la segunda secuencia. De manera deseable, una secuencia de ácidos nucleicos enlazado operativamente se refiere a un fragmento de un gen que esta enlazado a otras secuencias del mismo gen para formar un gen de longitud completa.

- 35 El término “recombinante” cuando se utiliza 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 codificados por un ácido nucleico heterólogo. Las células recombinantes pueden expresar genes o fragmentos de genes bien sea en formas en 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 son modificados y reintroducidos en la célula por medios artificiales.

- 40 Un “gen estructural” es una porción de un gen que comprende un segmento de ADN que codifica una proteína, polipéptido o una porción del mismo, y que excluye, por ejemplo, la secuencia 5´ que guía la iniciación de la transcripción o el 3´ UTR. El gen estructural puede codificar alternativamente un producto no traducible. El gen estructural puede ser uno que normalmente se encuentra en la célula o que normalmente no se encuentra en la célula o localización celular en donde es introducido, en cuyo caso se denomina un “gen heterólogo”. Un gen heterólogo puede ser derivado en todo o en parte de una fuente conocida en la técnica, incluyendo un genoma bacteriano o episoma, un ADN eucariótico, 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 manera de control de expresión. Tales modificaciones incluyen, pero no se limitan a, mutaciones, inserciones, eliminaciones y sustituciones de uno o más nucleótidos.

- 50 El gen estructural puede constituir una secuencia de codificación no interrumpida o puede incluir uno o más intrones, enlazados por las uniones de empalme apropiadas. El gen estructural puede ser traducible o no traducible, incluyendo una molécula en antisentido o de ARN capaz de inducir ARN de interferencia (ARNi). El gen estructural puede ser una composición de segmentos derivada de una pluralidad de fuentes y de una pluralidad de secuencias de genes (de origen natural o sintético, en donde sintético se refiere a un ADN que es sintetizado químicamente).

- 55 Un “exón” tal como se utiliza aquí en referencia a una secuencia de ácidos nucleicos significa una porción de la secuencia de ácidos nucleicos de un gen, en donde la secuencia de ácidos nucleicos del exón codifica al menos un aminoácido del producto genético. Un exón es típicamente adyacente a un segmento de ADN no codificador tal como un intrón.

Un "intrón" tal como se utiliza aquí en referencia a una secuencia de ácidos nucleicos significa una región no codificadora de un gen que es flanqueada por regiones codificadoras. Un intrón típicamente es una región no codificadora de un gen que es transcrita en una molécula de ARN pero es escindida por el ARN que se empalma durante la producción del ARN mensajero u otro ARN estructural funcional.

5 Un "3' UTR" tal como se utiliza aquí en referencia a una secuencia de ácidos nucleicos indica una secuencia de ácidos nucleicos no codificadora próxima a un codón de detención de un exón.

10 "Derivado de" se utiliza para indicar 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 limitándose a, sustitución, adición, inserción, eliminación, extracción, aislamiento, mutación y replicación) de la fuente original.

15 "Sintetizada químicamente", tal como se relaciona con una secuencia de ADN, significa que las porciones de los nucleótidos componentes fueron ensambladas *in vitro*. La síntesis química manual de ADN puede ser logrado utilizando procedimientos bien establecidos (Caruthers, *Methodology of DNA and RNA Sequencing*, (1983), Weissman (ed.), Praeger Publishers, New York, Chapter 1); la síntesis química automática puede llevarse a cabo utilizando cualquiera de un cierto número de máquinas disponibles comercialmente.

20 El alineamiento óptimo de las secuencias para la comparación puede llevarse a cabo, por ejemplo, por el algoritmo de homología local de Smith y Waterman, *Adv. Appl. Math.* 2:482 (1981), por el algoritmo de alineamiento de homología de Needleman y Wunsch, *J. Mol. Biol.* 48:443 (1970), por la búsqueda según el método de similitud de Pearson y Lipman *Proc. Natl. Acad. Sci. (U.S.A.)* 85: 2444 (1988), por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o por inspección.

25 La NCBI Basic Local Alignment Search Tool (BLAST) (Altschul et al., 1990) está disponible a partir de varias fuentes, incluyendo el National Center for Biological Information (NCBI, Bethesda, Md.) y en internet, para uso en conexión con los programas de análisis de secuencias blastp, blastn, blastx; tblastn, y tblastx. Puede tenerse acceso en <http://www.ncbi.nlm.nih.gov/BLAST/>. Una descripción de cómo determinar la identidad de secuencia utilizando este programa está disponible en http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.

30 Los términos "identidad de aminoácidos sustancial" o "identidad de secuencia de aminoácidos sustancial" tal como se aplica a las secuencias de aminoácidos si se utiliza aquí denota una característica de un polipéptido en donde el péptido comprende una secuencia que tiene al menos 70 por ciento de identidad en secuencia, por ejemplo preferiblemente 80 por ciento de identidad de secuencia de aminoácidos, más preferiblemente 90 por ciento de identidad de secuencia de aminoácidos, y lo más preferiblemente al menos 99 a 100 por ciento de identidad de secuencia en comparación con la secuencia de proteínas mostrada en las Figuras 10 a 17. De manera deseable, para una nicotina desmetilasa, la comparación de secuencias se efectúa de manera deseable en una región que sigue al motivo GXRXCX(G/A) (SEQ ID NO: 2265) del citocromo p450 hasta el codón de detención del péptido traducido.

35 Los términos "identidad sustancial de ácidos nucleicos" o "identidad sustancial de secuencia de ácidos nucleicos" tal como se aplica a las secuencias de ácidos nucleicos tal 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 preferiblemente 81 o 91 por ciento de identidad de secuencia de ácidos nucleicos, y lo más preferiblemente al menos 95, 99, o incluso 100 por ciento de identidad de secuencia en comparación con una región de un grupo de referencia correspondiente al primer ácido nucleico que sigue a la región de codificación del motivo GXRXCX(G/A) (SEQ ID NO: 2265) de citocromo p450 hasta el codón de detención del péptido traducido.

40 Otra indicación de que las secuencias de nucleótidos son sustancialmente idénticas es si las dos moléculas se hibridan una a otra bajo condiciones de restricción. Las condiciones de restricción son dependientes de la secuencia y serán diferentes en diferentes circunstancias. De manera general, se seleccionan las condiciones de restricción para que sean de alrededor de 5°C a aproximadamente 20°C, de manera usual aproximadamente 10°C hasta aproximadamente 15°C, inferior al punto de fusión térmico (Tm) para la secuencia específica a una fuerza iónica y pH definidos. La Tm es la temperatura (bajo fuerza iónica y pH definidos) a la cual el 50% de la secuencia objetivo hibrida a una sonda comparada. Típicamente, las condiciones de restricción serán aquellas en las cuales la concentración de sal es aproximadamente 0.02 molar a pH 7 y la temperatura es al menos aproximadamente 60°C. Por ejemplo, en el procedimiento de hibridación Southern estándar, las condiciones de restricción incluirán un lavado inicial en 6xSSC a 42°C seguido por uno o más lavados adicionales en 0.2xSSC a una temperatura de al menos aproximadamente 55°C, típicamente alrededor de 60°C, y frecuentemente alrededor de 65°C.

55 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, cuando una secuencia de ácidos nucleicos codifica esencialmente el mismo polipéptido que una segunda secuencia

de ácidos nucleicos, las dos secuencias de ácidos nucleicos son sustancialmente idénticas incluso si no hibridan bajo condiciones restrictivas debido a la degeneración permitida por el código genético (véase, Darnell et al. (1990) Molecular Cell Biology, Second Edition Scientific American Books W. H. Freeman and Company New York para una explicación de la degeneración de los codones y el código genético). La pureza u homogeneidad de la proteína pueden ser indicadas mediante un cierto número de medios bien conocidos en el arte, tales como electroforesis en gel de poli(acrilamida) de una muestra de proteína, seguida por visualización por tinción. Para ciertos propósitos puede requerirse de alta resolución y pueden utilizarse HPLC o un medio similar para la purificación.

Por un anticuerpo que “se enlaza específicamente” o “reconoce específicamente” un polipéptido particular, tal como una nicotina desmetilasa de tabaco, se entiende una afinidad incrementada del anticuerpo por el polipéptido con respecto a 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, Figuras 3 y 4, o 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 SEQ ID NO: 3 tiene de manera deseable 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 un cierto número de métodos estándar en la técnica, por ejemplo, análisis Western, ELISA o inmunoprecipitación. Los anticuerpos que se enlazan específicamente a un polipéptido, por ejemplo una nicotina desmetilasa, también son útiles para purificar el polipéptido.

Tal como se utiliza aquí, el término “vector” es utilizado en referencia a moléculas de ácidos nucleicos que transfieren segmentos de ADN hacia una célula. Un vector puede actuar para replicar ADN y puede reproducirse independientemente en una célula anfitriona. El término “vehículos” se utiliza algunas veces de manera intercambiable con “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 de ácidos nucleicos apropiadas necesarias para la expresión de la secuencia de codificación enlazado operativamente en un organismo anfitrión particular. Las secuencias de ácidos nucleicos necesarias para la expresión en procariontes incluyen usualmente un promotor, un operador (opcional), y un sitio de enlazamiento al ribosoma, frecuentemente junto con otras secuencias. De manera deseable, el promotor incluye la secuencia de SEQ ID NO: 8 o un fragmento de la misma que guía 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 guían la transcripción. Las células eucarióticas son conocidas por utilizar promotores, potenciadores y señales de terminación y poliadenilación, tales como la secuencia 3' UTR de SEQ ID NO: 9. En algunos casos, se ha observado que los vectores de expresión vegetal requieren la presencia de intrones derivados de las plantas, tales como un intrón que tiene la secuencia de SEQ ID NO: 7, para tener expresión estable. Como tal, la secuencia de SEQ ID NO: 7 o cualquier intrón que tenga una unión de empalme de ARN apropiada puede ser utilizada como se describe aquí más adelante. Vectores deseables incluyen una secuencia de ácido nucleico mostrada en las Figuras 1, 3 a 7, 10 a 17.

Para el propósito de regenerar plantas manipuladas genéticamente completas con raíces, un ácido nucleico puede ser insertado en las células vegetales, por ejemplo, por cualquier técnica tal como inoculación *in vivo* o por cualquiera de las técnicas de cultivo de tejidos *in vitro* conocidas para producir células vegetales transformadas que pueden ser regeneradas en plantas completas. Así, por ejemplo, la inserción en células vegetales puede ser por inoculación *in vitro* por *A. tumefaciens* patogénico o no patogénico. También pueden emplearse otras tales técnicas de cultivo de tejidos.

“Tejido vegetal”, “componente de plantas” o “célula vegetal” incluyen tejidos de plantas diferenciados y no diferenciados, incluyendo, pero no limitándose a, raíces, brotes, hojas, polen, semillas, tejidos tumorales y diversas formas de células en cultivo, tales como células individuales, protoplastos, embriones y tejido de callus. El tejido vegetal puede estar *in planta* o en un cultivo de órganos, tejidos o células.

“Célula vegetal” tal como se utiliza aquí incluye células de plantas *in planta* y células de plantas y protoplastos en cultivo. “ADNc” o “ADN complementario” se refiere en general a una molécula de ADN de cadena sencilla con una secuencia de nucleótidos que es complementaria a una molécula de ARN no procesada que contiene un intrón, o a una ARNm procesada que carece de intrones. El ADNc es formado por la acción de la enzima reversa transcriptasa sobre una plantilla de ARN.

“Tabaco” tal como se utiliza aquí, incluye curado amarillo, Virginia, Burley, oscuro, oriental y otros tipos de plantas dentro del género *Nicotiana*. La semilla del género *Nicotiana* es fácilmente disponible comercialmente en la forma de *Nicotiana tabacum*.

“Artículos de manufactura” o “productos de tabaco” incluyen productos tales como rapé húmedo y seco, tabacos para mascar, productos de cigarrillos, productos de cigarros, cigarrillos, tabacos para pipa, bidis y productos similares derivados del tabaco.

Por “silenciamiento del gen” se entiende un descenso en el nivel de la expresión del gen (por ejemplo, expresión de un gen que codifica nicotina desmetilasa de tabaco) en al menos 30-50%, preferiblemente en al menos 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 tipo silvestre). La reducción de tales niveles de expresión puede lograrse empleando métodos estándar que son conocidos en la técnica incluyendo, sin limitación, ARN de interferencia, interferencia de cadena triple, ribozimas, recombinación homóloga, silenciamiento de genes inducido por virus, tecnologías antisentido y de cosupresión, expresión de un producto de gen negativo dominante, o a través de la generación de genes mutados utilizando técnicas de mutagénesis estándar, tales como las descritas aquí. Los niveles de un polipéptido o transcripto de nicotina desmetilasa de tabaco, o ambos, son monitorizados de acuerdo con cualquier técnica estándar incluyendo, pero no limitándose a, inmunoprecipitación Northern, protección de RNasa, o inmunoprecipitación.
- 5
- 10 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 a 17. Fragmentos deseables de ejemplo son los aminoácidos 1-313 de la secuencia de SEQ ID NO: 3 y los aminoácidos 314-517 de la secuencia de SEQ ID NO: 3, así como la secuencia de SEQ ID NO: 2 y 63. Además, con respecto a un fragmento o porción de una secuencia de ácidos nucleicos, los fragmentos deseables incluyen al menos 100, 250, 500, 750, 1000 o 1500 de ácidos nucleicos contiguos de cualquiera de las secuencias de ácidos nucleicos mostradas en las figuras 1, 3 a 7, y 10 a 17. Fragmentos deseables de ejemplos son los ácidos nucleicos 1-2009, 2010-2949, 2950-3946, 3947-4562, 4563-6347 y 4731-6347 de la secuencia de SEQ ID NO: 4.
- 15
- 20 Por un "polipéptido sustancialmente puro" se entiende un polipéptido que ha sido separado de la mayoría de los componentes que lo acompañan de manera natural; sin embargo, otras proteínas encontradas en la fracción microsómica asociadas con una preparación que tienen una actividad enzimática de al menos 8.3 pKat/mg de proteína también se consideran como un polipéptido sustancialmente puro. Típicamente, el polipéptido sustancialmente puro cuando está al menos 60%, en peso, libre de las proteínas y moléculas de origen orgánico con las cuales está asociado de manera natural. Preferiblemente, la preparación es al menos 75%, más preferiblemente al menos 90%, y lo más preferiblemente al menos 99%, en peso, del polipéptido deseado. Un polipéptido sustancialmente puro puede ser obtenido, por ejemplo, por extracción a partir de una fuente natural (por ejemplo, una célula vegetal de tabaco); por expresión de un ácido nucleico recombinante que codifica el polipéptido; o sintetizando químicamente la proteína. La pureza puede ser medida por cualquier método apropiado, por ejemplo, cromatografía de columna, electroforesis en gel de poliacrilamida o análisis por HPLC.
- 25
- 30 Por "molécula de ácido nucleico aislada" se entiende una secuencia de ácidos nucleicos libre de las secuencias de ácidos nucleicos que flanquean de manera natural la secuencia de la molécula de ácido nucleico en el genoma de un organismo.
- 35 Por "célula transformada" se entiende una célula en la cual (o en un ancestro de la cual) se ha introducido, por medio de técnicas de ADN recombinante, una molécula de ADN, por ejemplo una molécula de ADN que codifica nicotina desmetilasa de tabaco o cualquiera de las secuencias de ácidos nucleicos divulgadas aquí (por ejemplo, las secuencias de ácidos nucleicos mostradas en las figuras 1, 3 a 7 y 10 a 17).
- 40 Por una "nicotina desmetilasa de tabaco" o "nicotina desmetilasa" tal como se utiliza aquí, se entiende un polipéptido que es sustancialmente idéntico a la secuencia de SEQ ID NO: 3. De manera deseable, una nicotina desmetilasa de tabaco es capaz de convertir la nicotina ($C_{10}H_{14}N_2$), también denominada como 3-(1-metil-2pirrolidinil)piridina en nornicotina ($C_9H_{12}N_2$). La actividad de una nicotina desmetilasa de tabaco puede ser probada utilizando métodos estándar en el arte, tales como la medición de la desmetilación de nicotina radioactiva por microsomas expresados en levadura, tal como se describe aquí.
- 45 Tal como se proveen aquí, los términos "citocromo p450" y "p450" se utilizan indistintamente.
- 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 una secuencia de ácidos nucleicos 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 de nicotina desmetilasa de tabaco.
- 50 La figura 3 es una secuencia de ácidos nucleicos de nicotina desmetilasa de tabaco genómica (SEQ ID NO: 4) y su producto de traducción (SEQ ID NO: 3)
- La figura 4 es la secuencia de ácidos nucleicos de la región de codificación del gen de nicotina desmetilasa de tabaco (SEQ ID NO: 5), y su producto de traducción (SEQ ID NO: 6).

La figura 5 es la secuencia de ácidos nucleicos de un intrón (SEQ ID NO: 7) presente en la secuencia genómica de nicotina desmetilasa de tabaco.

La figura 6 es la secuencia de ácidos nucleicos del gen promotor de desmetilasa de tabaco (SEQ ID NO: 8).

5 La figura 7 es la secuencia de ácidos nucleicos de la 3'UTR del gen de nicotina desmetilasa de tabaco (SEQ ID NO: 9).

La figura 8 es una imagen de electroforesis que muestra los productos por PCR de líneas de tabaco con el conjunto de cebador de longitud completa ("FL") Geno.

10 La figura 9 es una imagen de electroforesis que muestra los productos por PCR de líneas de tabaco con conjuntos de cebadores (1), (2), (3), y (4) como se establece en el Ejemplo 17. Los tamaños aproximados de las bandas son 3,500 nucleótidos (nt) para FL, 2,600 nt para (1), 1,400 nt para (2), 600 nt para (3), y 1,400 nt para (4).

La figura 10 es la secuencia de ácidos nucleicos de D35-BG11 (SEQ ID NO: 62), y la secuencia de aminoácidos de D35-BG11 (SEQ ID NO: 63), la cual es la traducción de D35-BG11 (SEQ ID NO: 62).

La Figura 11 es la secuencia de ácidos nucleicos D120-AH4 (SEQ ID NO: 188), y la secuencia de aminoácidos de D120-AH4 (SEQ ID NO: 189), la cual es la traducción de D120-AH4 (SEQ ID NO: 188).

15 La Figura 12 es la secuencia de ácidos nucleicos D121-AA8 (SEQ ID NO: 190), y la secuencia de aminoácidos de D121-AA8 (SEQ ID NO: 191), la cual es la traducción de D121-AA8 (SEQ ID NO: 190).

La Figura 13 es la secuencia de ácidos nucleicos D122-AF10 (SEQ ID NO: 192), y la secuencia de aminoácidos de D122-AF10 (SEQ ID NO: 193), la cual es la traducción de D122-AF10 (SEQ ID NO: 192).

20 La Figura 14 es la secuencia de ácidos nucleicos D208-AC8 (SEQ ID NO: 226), y la secuencia de aminoácidos de D208-AC8 (SEQ ID NO: 227), la cual es la traducción de D208-AC8 (SEQ ID NO: 226).

La Figura 15 es la secuencia de ácidos nucleicos D103-AH3 (SEQ ID NO: 230), y la secuencia de aminoácidos de D103-AH3 (SEQ ID NO: 231), la cual es la traducción de D103-AH3 (SEQ ID NO: 230).

La Figura 16 es la secuencia de ácidos nucleicos D208-AD9 (SEQ ID NO: 232), y la secuencia de aminoácidos de D208-AD9 (SEQ ID NO: 233), la cual es la traducción de D208-AD9 (SEQ ID NO: 232).

25 La Figura 17 es la secuencia de ácidos nucleicos D235-AB1 (SEQ ID NO: 254), y la secuencia de aminoácidos de D235-AB1 (SEQ ID NO: 255), la cual es la traducción de D235-AB1 (SEQ ID NO: 254).

30 La figura 18 es un conjunto de alineamientos de secuencias de aminoácidos consistentes del alineamiento de secuencias 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); el alineamiento de secuencia de D244-AD4 (SEQ ID NO: 2203), D244-AB6 (SEQ ID NO: 2204), D285-AA8 (SEQ ID NO: 2205), D285-4B9 (SEQ ID NO: 2206), y D268-AE2 (SEQ ID NO: 2207); y el alineamiento de secuencia de D100A-AC3 (SEQ ID NO: 2208) y D100A-BE2 (SEQ ID NO: 2209).

35 La figura 19 es un diagrama que muestra la clonación de fragmentos de ADNc de citocromo p450 por PCR. Los cebadores utilizados para la clonación aparecen 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 muestra las secuencias del conjunto de sondas (SEQ ID NO: 385 - SEQ ID NO: 445) de los clones sobre el microarreglo GeneChip®.

Descripción detallada

40 Tradicionalmente, se involucraron etapas numerosas en el desarrollo de cualquier germoplasma de planta novedoso deseable. El cruzamiento de plantas comienza con el análisis y definición de los problemas y la debilidad del germoplasma actual, el establecimiento de metas del programa, y la definición de objetivos de cruzamiento específicos. La siguiente etapa es la selección de germoplasma que posee las características que satisfacen las metas del programa. La meta es combinar en una variedad individual una combinación mejorada de características deseables del germoplasma progenitor. Las características deseables incluyen, por ejemplo, un rendimiento en semillas más alto, resistencia a enfermedades e insectos, tolerancia a la sequía y al calor, y mejores cualidades agronómicas. Sin embargo, estos procesos, que llevan a la etapa final de comercialización y distribución, pueden tomar de seis a doce años a partir del momento en que se hace el inicio. De acuerdo con lo anterior, el desarrollo de

nuevas variedades es un proceso que consume tiempo que requiere una planificación hacia adelante precisa, uso eficiente de recursos, y un mínimo de cambios en la dirección.

5 La mejora de variedades vegetales a través de la transformación genética se ha hecho crecientemente importante para el cruzamiento vegetal moderno. Los genes de interés comercial potencial, tales como los genes que confieren características específicas, deseadas a las plantas de resistencia a enfermedades, resistencia a insectos o calidad mejorada, pueden ser incorporados en especies en cultivos a través de diversas tecnologías de transferencia de genes. La capacidad de manipular la expresión genética provee un medio para producir nuevas características en plantas transformadas. En algunas situaciones pueden desearse niveles altos o incrementados de expresión genética. Por ejemplo, es deseable incrementar la producción de una proteína que por sí misma maximiza la resistencia a una enfermedad, rendimiento, sabor o cualquier otro atributo comercialmente deseable de una planta. De la misma manera, la regulación de la expresión genética endógena por ejemplo, por silenciamiento de genes puede dar como resultado plantas o productos de plantas más valiosos.

15 Durante la maduración o curado del tabaco, la activación, sobreexpresión o subregulación de cualquiera de los genes identificados como inducidos por etileno o relacionados con la senescencia (por ejemplo, los que tienen la secuencia de SEQ ID NOS: 4, 188 y 226) pueden afectar aquellas rutas metabólicas involucradas en la formación de numerosos metabolitos secundarios incluyendo terpenoides, polifenoles, alcaloides, etc., que afectan las características de calidad del producto final (por ejemplo, resistencia a enfermedades, resistencia a insectos, calidad mejorada, aroma modificado, sabor modificado y similares). De la misma forma afectadas por los genes identificados aquí pueden ser las rutas metabólicas asociadas con la rata y tipo de materia seca acumulada durante la senescencia o la partición de la materia seca dentro de la planta durante la senescencia. Los cambios en la rata y tipo de acumulación de almidón, formación de lignina, deposición de celulosa y translocación de azúcares podrían ser demostrados. El control de los genes identificados aquí también puede afectar aquellas rutas metabólicas involucradas en la determinación de las tasas de senescencia, la uniformidad de la senescencia dentro de una hoja y entre hojas de una planta individual, y la inducción de la senescencia por medios artificiales o naturales. Los agentes o actividades inductores de la senescencia que estimulan o activan los genes identificados aquí incluyen, por ejemplo, agente químicos tales como peróxidos diluidos, pesticidas, herbicidas, reguladores del crecimiento, tratamientos por calor, lesiones, o gases tales como ozono y concentraciones elevadas de dióxido de carbono.

Identificación de secuencias de tabaco expresadas constitutivamente o inducidas por etileno o senescencia

30 El ARN fue extraído de tejido de *Nicotiana* de líneas de *Nicotiana* convertidoras y no convertidoras. El ARN extraído fue utilizado entonces para crear ADNc. Las secuencias de ácido nucleico fueron generadas entonces utilizando dos estrategias.

35 En la primera estrategia, el ARN enriquecido con poli A fue extraído de tejido vegetal y se hizo el ADNc por PCR en transcripción reversa. La cadena sencilla de ADNc fue utilizada entonces para crear poblaciones por PCR específicas de p450 utilizando cebadores degenerados más un cebador reverso oligo d(T). El primer diseño estuvo basado en los motivos altamente conservados de otras secuencias de genes de citocromo p450 de otra planta. Ejemplos de cebadores degenerados específicos se presentan en la figura 1 de las Publicaciones de Solicitud de Patente US 2004/0103449 A1, US 2004/0111759 A1, y US 2004/0117869 A1. La secuencia de fragmentos de plásmidos que contienen insertos de tamaño apropiado fue analizada posteriormente. Los tamaños de estos insertos variaban típicamente desde aproximadamente 300 hasta aproximadamente 800 nucleótidos dependiendo de los cebadores usados.

40 En una segunda estrategia, se construyó inicialmente una biblioteca de ADNc. El ADNc en los plásmidos fue utilizado para crear poblaciones de PCR específicas de p450 utilizando cebadores degenerados más el cebador T7 sobre plásmido como cebador reverso. Como sucedió en la primera estrategia, la secuencia de fragmentos de plásmidos que contenía insertos de tamaño apropiado fue analizada posteriormente.

45 Las líneas de plantas de *Nicotiana* conocidas por producir altos niveles de nornicotina (convertidora) y líneas de plantas que tenían bajos niveles de nornicotina pueden ser utilizadas como materiales de partida. Las hojas pueden ser retiradas de las plantas y tratadas con etileno para activar las actividades enzimáticas de p450 definidas aquí. El ARN total es extraído utilizando técnicas conocidas en el arte. Los fragmentos de ADNc pueden ser generados entonces utilizando PCR (RT-PCR) con el oligo cebador d(T) (SEQ ID NO: 2260) como se describe en la figura 19. La biblioteca de ADNc puede ser construida entonces como se describe de manera más completa en los ejemplos aquí.

55 La región conservada de las enzimas tipo p450 se utiliza como plantilla para cebadores degenerados, ejemplos de los cuales se muestran en la figura 19. Utilizando cebadores degenerados, las bandas específicas de p450 fueron amplificadas por PCR. Las bandas indicadoras de enzimas similares a p450 fueron identificadas por secuenciamiento de ADN. Los fragmentos del PCR fueron caracterizados utilizando búsqueda BLAST, alineamiento u otras herramientas para identificar candidatos apropiados.

La información de secuencias a partir de los fragmentos identificados fue utilizada para desarrollar los cebadores para PCR. Estos cebadores en combinación con cebadores de plásmidos en la biblioteca de ADNc fueron utilizados para clonar genes de p450 de longitud completa. Se llevó a cabo un análisis reverso por Southern a gran escala para examinar la expresión diferencial para todos los clones de fragmentos obtenidos y en algunos casos los clones de longitud completa. Estos ensayos Southern reversos a gran escala pueden ser llevados a cabo utilizando ADNc total de diferentes tejidos como una sonda para hibridar con fragmentos de ADN clonados con el fin de seleccionar todos los insertos clonados. Los ensayos de inmunoprecipitación Northern no radiactiva y radiactiva (P^{32}) fueron utilizados también para caracterizar los fragmentos de p450 clonados y clones de longitud completa.

Una vez que se obtienen las células vegetales que expresan el nivel deseado de la enzima p450, los tejidos vegetales y las plantas completas pueden ser regenerados a partir de las mismas utilizando métodos y técnicas bien conocidos en el arte. Las plantas regeneradas son reproducidas entonces por medios convencionales y los genes introducidos pueden ser transferidos a otras cepas y cultivares por técnicas de cruzamiento de plantas convencionales.

Los genes inducidos por etileno o relacionados con la senescencia, por ejemplo, los identificados en SEQ ID NOS: 4, 188 y 226 pueden codificar enzimas que son determinantes importantes de los parámetros de calidad de la hoja del tabaco importantes para una variedad de productos de tabaco. Los productos de tabaco incluyen rapé húmedo y seco, tabacos para mascar, cigarrillos, tabacos para pipa, bidis y productos similares para fumar. Los parámetros de calidad de las hojas pueden incluir: atributos visuales tales como color, uniformidad de la superficie, textura o variegación; características estructurales o físicas tales como las ejemplificadas por la relación lámina a tallo, aceitosidad, potencial para llenado de cigarrillos, densidad aparente, retención de humedad, y flexibilidad; características químicas o bioquímicas relacionadas con sabor, aroma, capacidad de fermentación, ratas de quema, temperaturas de quema, absorción y liberación de sabores artificiales; y generación de constituyentes del humo incluyendo alquitrán o materia en partículas, alcaloides y otros atributos similares. Las reacciones enzimáticas que resultan de estos genes inducidos por etileno o relacionados con la senescencia pueden producir también metabolitos secundarios que influyen en las interacciones con patógenos o insectos que afectan el rendimiento y calidad de las hojas de tabaco. Por ejemplo, Wagner et al. (Nature Biotechnology, 19:371-374, 2001) demostró que la supresión del gen p450 hidroxilasa incrementa grandemente la acumulación de cembratrien-ol, un metabolito secundario que influye en la resistencia de los áfidos.

Generación de anticuerpos

Se hicieron anticuerpos específicos de péptidos derivando su secuencia de aminoácidos y seleccionando regiones de péptidos que fueran antigénicas y únicas con respecto a otros clones. Los anticuerpos de conejo fueron hechos para que péptidos sintéticos se conjugaran con una proteína portadora. Los análisis por inmunoprecipitación Western u otros métodos inmunológicos fueron llevados a cabo sobre tejidos de plantas utilizando estos anticuerpos. Además, los anticuerpos específicos para péptidos fueron hechos para varios clones de longitud completa derivando su secuencia de aminoácidos y seleccionando regiones de péptidos que eran potencialmente antigénicas y eran únicas con respecto a otros clones. Los anticuerpos de conejo fueron hechos en péptidos sintéticos conjugados a una proteína portadora. Los análisis por inmunoprecipitación Western fueron llevados a cabo utilizando estos anticuerpos.

Subregulación de la expresión genética y alteración de la actividad enzimática

Las plantas que tienen expresión disminuida de un polipéptido se generan de acuerdo con métodos de silenciamiento de genes estándares. (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, las secuencias de ácidos nucleicos de la nicotina desmetilasa de tabaco (por ejemplo, SEQ ID NO: 4, 5, 7, 8 y 9, o fragmentos de las mismas tales como la secuencias de SEQ ID NOS: 1 y 62) así como secuencias de ácidos nucleicos sustancialmente idénticas (por ejemplo, la secuencia de SEQ ID NO: 188) pueden ser utilizadas para alterar los fenotipos del tabaco o metabolitos del tabaco, por ejemplo, nornicotina en cualquier especie de *Nicotiana*. La expresión reducida de un gen de nicotina desmetilasa de tabaco puede ser lograda usando, por ejemplo, ARN de interferencia (ARNi) (Smith et al., 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); virusinduced gene silencing ("VIGS") (Baulcombe, Current Opinions in Plant Biology, 2:109-113, 1999; Cogoni and Macino, Genes Dev 10: 638-643, 2000; Ngelbrecht et al., PNAS 91:10502-10506, 1994); silenciamiento del gen objetivo transfiriendo un gen endógeno de la planta en la orientación en 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, The Netherlands, 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 January, 1999); atrapamiento de gen y etiquetado de T-ADN (Bums et al., Genes Dev. 8: 1087-1105, 1994; Spradling, et al., PNAS 92:10824-10830, 1995; Skames et al., Bio/Technology 8, 827-831, 1990; Sundaresan, et al., Genes Dev. 9: 1797-1810, 1995); y cualquiera de los otros posibles sistemas de silenciamiento de genes que están disponibles en las

áreas científicas que dan como resultado la subregulación de la expresión del polipéptido de tabaco o una reducción en su actividad enzimática. Como se provee adicionalmente aquí, cualquiera de las secuencias de aminoácidos provistas aquí pueden ser subreguladas o sobrerreguladas utilizando técnicas descritas aquí y otras tecnologías encontradas en el arte. Más abajo se describen en detalle métodos de ejemplo.

5 Interferencia de ARN

La interferencia de ARN ("ARNi") es un proceso aplicable generalmente para inducir silenciamiento de genes postraduccion potente y específico en muchos organismos incluyendo plantas (véase, por ejemplo Boshier et al., Nat. Cell Biol. 2: E31-36, 2000; y Tavernarakis et al., Nat. Genetics 24: 180-183, 2000). El ARNi involucra la introducción de ARN con carácter de cadena doble parcial o completamente en la célula o en el ambiente extracelular. La inhibición es específica porque una secuencia de nucleótidos de una porción del gen objetivo (por ejemplo una nicotina desmetilasa de tabaco) se escoge para producir el ARN inhibidor. La porción escogida generalmente abarca exones del gen objetivo, pero la porción escogida también puede incluir regiones no traducidas (UTR), así como intrones (por ejemplo, la secuencia de SEQ ID NO: 7, o una secuencia de ácidos nucleicos a partir de un gen vegetal deseado, tal como una secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17).

Por ejemplo, para construir vectores de transformación que producen ARN capaz de formación dúplex, dos secuencias de ácidos nucleicos, una en la orientación en sentido y la otra en antisentido, pueden ser enlazadas operativamente, y colocadas bajo el control de un promotor viral fuerte, tal como CaMV 35S o el promotor aislado a partir del virus de color polvoso pardo de la cassava (CBSV). Sin embargo, el uso del promotor endógeno, tal como el promotor de la nicotina desmetilasa que tiene la secuencia de SEQ ID NO: 8, o un fragmento del mismo que guía la transcripción, también puede ser deseable. La longitud de las secuencias de ácidos nucleicos de nicotina desmetilasa de tabaco incluidas en tal constructo es de manera deseable al menos de 25 nucleótidos, pero puede abarcar una secuencia que incluye hasta el gen de nicotina desmetilasa de tabaco de longitud completa.

Los constructos que producen ARN capaz de formar dúplex pueden ser introducidos en el genoma de una planta, tal como una planta de tabaco, por transformación mediada por *Agrobacterium* (Chuang et al., Proc. Natl. Acad. Sci. USA 97: 4985-4990, 2000), produciendo interferencia genética específica y heredable en la nicotina desmetilasa de tabaco. El ARN de cadena doble también puede ser introducido directamente en la célula (esto es, intracelularmente) o inducido extracelularmente, por ejemplo, bañando una semilla, brote o planta en una solución que contiene el ARN de cadena doble.

Dependiendo de la dosis de material de ARN de cadena doble administrado, el ARNi puede proveer una pérdida de función parcial o completa para el gen objetivo. Puede obtenerse una reducción o pérdida de la expresión genética en al menos 99% de las células objetivo. En general, dosis menores de material infectado y tiempos más largos de administración de ARNs da como resultado la inhibición en una fracción más pequeña de células.

El ARN utilizado en ARNi puede comprender una o más cadenas de ribonucleótido polimerizado; puede incluir modificaciones bien sea al esqueleto fosfato-azúcar o al nucleósido. La estructura de cadena doble puede ser formada mediante una cadena de ARN sencilla autocomplementaria o por dos cadenas de ARN complementarias y la formación de ARN dúplex puede ser iniciada bien sea dentro o fuera de la célula. El ARN puede ser introducido en una cantidad que permite la administración de al menos una copia por célula. Sin embargo, dosis más altas (por ejemplo, al menos, 5, 10, 100, 500 o 1000 copias por células) de material de cadena doble puede producir más inhibición efectiva, la inhibición es específica de la secuencia porque las secuencias de nucleótidos correspondientes a la región doble del ARN son direccionadas para la inhibición genética. El ARN que contiene secuencias de nucleótido idénticas a una porción del gen objetivo es preferido para la inhibición. Las secuencias de ARN con inserciones, eliminaciones y mutaciones puntuales sencillas con respecto a la secuencia objetivo también pueden ser efectivas para la inhibición. Así, la identidad de secuencia puede ser optimizada por algoritmos de alineamiento conocidos en el arte y calculando el porcentaje de diferencia entre las secuencias de nucleótidos. Alternativamente, la región dúplex del ARN puede ser definida funcionalmente como una secuencia de nucleótidos que es capaz de hibridar con una porción del transcrito del gen objetivo.

Además, el ARN utilizado para ARNi puede ser sintetizado bien sea *in vivo* o *in vitro*. Por ejemplo, la ARN polimerasa endógena en la célula puede mediar en la transcripción *in vivo*, o la polimerasa de ARN clonada puede ser utilizada para la transcripción *in vivo* o *in vitro*. Para la transcripción a partir de un transgen *in vivo* o un constructo de expresión, puede utilizarse una región reguladora para transcribir la cadena (o cadenas) de ARN.

Interferencia por cadena triple

La expresión del gen de nicotina desmetilasa de tabaco endógeno o la expresión de un fragmento de ácido nucleico a partir de un gen de planta deseado, tal como cualquier secuencia de ácidos nucleicos en las figuras 1, 3 a 7 y 10-17, también puede ser subregulada direccionando secuencias de desoxirribonucleótido complementarias a la región reguladora de un gen de tabaco (por ejemplo, regiones promotoras o potenciadoras) para formar estructuras helicoidales triples que evitan la transcripción del gen de tabaco en 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).

5 Las moléculas de ácidos nucleicos utilizadas en la formación de hélice triple para la inhibición de la transcripción son preferiblemente de cadena sencilla y están compuestas de desoxirribonucleótidos. La composición en bases de estos oligonucleótidos debería promover la formación de hélice triple a través de las reglas de apareamiento de bases de Hoogsteen, las cuales generalmente requieren estiramientos dimensionables bien sea de purinas o pirimidinas que estén presentes en una cadena de un dúplex. Las secuencias de nucleótidos pueden ser basadas en pirimidina, las cuales darán como resultado tripletes TAT y CGC a través de las tres cadenas asociadas de la hélice triple resultante. Las moléculas ricas en pirimidina proveen complementariedad en bases a una región rica en purina de una cadena sencilla del dúplex en una orientación paralela a la cadena. Además, las moléculas de ácido nucleico pueden ser escogidas de tal manera que sean ricas en purina, por ejemplo, que contengan un estiramiento de residuos G. Estas moléculas formarán una hélice triple con un dúplex de ADN que es rico en pares GC, en los cuales la mayoría de los residuos de purina están localizados sobre una cadena sencilla del dúplex direccionado, dando como resultado tripletes CGC a través de estas tres cadenas en el triplete.

15 Alternativamente, las secuencias potenciales que pueden ser direccionadas para la formación de hélices triples pueden ser incrementadas creando una molécula de ácido nucleico de "retroconmutación". Las moléculas de retroconmutación son sintetizadas de una forma alternante 5'-3', 3'-5', de tal manera que aparean bases con la primera cadena de un dúplex y luego la otra, eliminando la necesidad de un estiramiento dimensionable de purinas o pirimidinas que estén presentes en una cadena de un dúplex.

20 Ribozimas

Las ribozimas son moléculas de ARN que actúan como enzimas y que pueden ser manipuladas para escindir otras moléculas de ARN. Una ribozima puede ser diseñada para aparearse específicamente con cualquier ARN objetivo y escindir el esqueleto fosfodiéster en una localización específica, inactivando por lo tanto funcionalmente el ARN objetivo. La ribozima misma no es consumida en este proceso y puede actuar catalíticamente para escindir múltiples copias de moléculas de ARNm objetivo. De acuerdo con lo anterior, las ribozimas también pueden ser utilizadas como un medio para subregular la expresión de una nicotina desmetilasa de tabaco. El diseño y uso de ribozimas específicas para un ARN objetivo está descrito 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 ácidos nucleicos de un gen vegetal deseado, tal como cualquier secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17) a cada lado del sitio activo de la ribozima.

Además, las secuencias de ribozimas también pueden ser incluidas dentro de un ARN antisentido para conferir actividad de escisión de ARN sobre el ARN antisentido y, por lo tanto, incrementar la efectividad del constructo antisentido.

35 Recombinación homóloga

La tecnología de reemplazo de genes es otro método deseable para subregular la expresión de un gen dado. La tecnología de reemplazo de genes está basada en la recombinación homóloga (véase, Schnable et al., Curr. Opinions Plant Biol. 1: 123-129, 1998). La secuencia de ácidos nucleicos de la enzima de interés tal como nicotina desmetilasa de tabaco o un polipéptido codificado por cualquier secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7 y 10-17 puede ser manipulada por mutagénesis (por ejemplo, inserciones, eliminaciones, duplicaciones o reemplazos) para hacer disminuir la función enzimática. La secuencia alterada puede ser introducida entonces en el genoma para reemplazar el gen existente, por ejemplo tipo silvestre a través de recombinación homóloga (Putchá et al., Proc. Natl. Acad. Sci, USA 93: 5055-5060, 1996; y Kempin et al., Nature 389: 802-803, 1997). Alternativamente, un gen de nicotina desmetilasa de tabaco endógeno puede ser reemplazado con un gen que no tenga actividad de desmetilasa, por ejemplo, la secuencia de SEQ ID NO: 188.

Cosupresión

Un método adicional deseable para silenciar la expresión genética es la cosupresión (también denominada como supresión en sentido). Está técnica, que involucra la introducción de un ácido nucleico, por ejemplo, un fragmento de ácido nucleico a partir de un gen vegetal deseado, tal como una secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7 y 10-17, configurada en la orientación en sentido, ha demostrado bloquear efectivamente la transcripción de los genes objetivo (véase, por ejemplo Napoli et al., Plant Cell, 2: 279-289, 990 y Jorgensen et al., Patente de los Estados Unidos No. 5,034,323).

En general, la supresión en sentido involucra la transcripción de la secuencia introducida. Sin embargo, la supresión también puede ocurrir cuando la secuencia introducida no contiene secuencias de codificación *per se*, pero solamente intrones o secuencias 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 general será sustancialmente idéntica al gen endógeno direccionado para la represión. Tal identidad es típicamente mayor que aproximadamente 50%, pero se prefieren identidades mayores (por ejemplo, 80% o incluso 95%) por que dan como resultado una represión más efectiva. El efecto de la cosupresión también puede ser aplicado a otras proteínas dentro de una familia similar de genes que exhiben homología u homología sustancial.

5 Los segmentos de un gen de una planta pueden ser utilizados directamente, por ejemplo, para inhibir la expresión de genes homólogos en especies de plantas vegetales.

En la supresión en sentido, la secuencia introducida, que requiere menos que la identidad absoluta, no necesita ser de longitud completa, con respecto bien sea al producto de transcripción primario o al ARNm procesado completamente. Un grado más alto de identidad de secuencia en una secuencia más corta que de longitud completa compensa una secuencia más larga de menor identidad. Adicionalmente, la secuencia introducida no necesita tener el mismo patrón de intrón o exón, y la identidad de los segmentos no codificadores pueden ser igualmente efectiva. Se prefieren secuencias de al menos 50 pares de bases, con secuencias introducidas de mayor longitud como las más preferidas (véase, por ejemplo, los métodos descritos por Jorgensen et al., Patente de los Estados Unidos No. 5,034,323)

15 Supresión antisentido

En la tecnología antisentido, un segmento de ácidos nucleicos del gen vegetal deseado, tal como una secuencia de ácidos nucleicos mostrados en las figuras 1, 3 a 7 y 10-17, es clonada y enlazado operativamente a una región de control de expresión tal que la cadena de antisentido de ARN es sintetizada. El constructo es transformado entonces en plantas y se produce la cadena antisentido de ARN. En células vegetales, se ha demostrado que el ARN antisentido inhibe la expresión genética.

El segmento de ácidos nucleicos que va a ser introducido en la supresión antisentido es generalmente sustancialmente idéntico a al menos una porción del gen o genes endógenos que van a ser reprimidos, pero no necesitan ser idénticos. Las secuencias de ácidos nucleicos de la nicotina desmetilasa de tabaco divulgada aquí pueden ser incluidas 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 al gen objetivo. Pueden usarse segmentos de un gen de otra planta, por ejemplo, directamente para inhibir la expresión de genes homólogos en diferentes variedades de tabaco.

La secuencia introducida también no necesita ser de longitud completa con respecto bien sea al producto de transcripción primario o al ARNm procesado completamente. En general, puede utilizarse una homología más alta para compensar el uso de una secuencia más corta. Además, la secuencia introducida no necesita tener el mismo patrón de intrón o exón, y la homología de segmentos no codificadores será igualmente efectiva. En general, tal secuencia antisentido tendrá al menos usualmente 15 pares de bases, preferiblemente alrededor de 15-200 pares de bases, y más preferiblemente 200-2,000 pares de bases de longitud o más. La secuencia antisentido puede ser complementaria a todo o a una porción del gen que va a ser suprimido, y, según es apreciado por los experimentados en la técnica, el sitio o sitios en particular a los cuales se enlaza la secuencia antisentido así como la longitud de la secuencia antisentido variarán, dependiendo del grado de inhibición deseado y de la exclusividad de la secuencia antisentido. Un constructo transcripcional que expresa una secuencia de nucleótidos antisentido reguladora negativa en plantas incluye, en la dirección de la transcripción, un promotor, la secuencia codificadora para el ARN antisentido sobre la cadena en sentido, y una región de terminación transcripcional. 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 and Nilsson (Nature 377: 495-500, 1995); Cheung et al., (Cell 82: 383-393, 1995); y Shewmaker et al. (U.S. Pat. No. 5,107,065).

Negativos dominantes

Las plantas transgénicas que expresan un transgen que codifica un producto genético negativo dominante de un producto genético de tabaco pueden ser probadas en ambientes artificiales o en el campo para demostrar que el transgen confiere subregulaciones a un producto genético de tabaco en la planta transgénica. Los transgenes negativos dominantes son construidos de acuerdo con métodos conocidos en el arte. Típicamente, un gen negativo dominante incluye un polipéptido regulador negativo mutante de un producto genético de tabaco el cual, cuando se sobreexpresa, perturba la actividad de la enzima tipo silvestre.

Mutantes

Las plantas que tienen expresión o actividad enzimática disminuida de un producto genético de tabaco también pueden ser generadas utilizando metodologías de mutagénesis estándar. Tales métodos de mutagénesis incluyen, sin limitación, el tratamiento de semillas con metilsulfato de etilo (Hildering and Verkerk, In, The use of induced mutations in plant breeding. Pergamon press, pp 317-320, 1965) o de radiación con UV, rayos X e irradiación con neutrones rápidos (véase, por ejemplo, Verkerk, Neth. J. Agric. Sci. 19:197-203, 1971; y Poehlman, Breeding Field Crops, Van Nostrand Reinhold, New York (3.sup.rd ed), 1987), uso de transposones (Fedoroff et al., 1984; U.S. Pat.

No. 4,732,856 y U.S. Pat. No. 5,013,658), así como metodologías de inserción de T-ADN (Hoekema et al., 1983; U.S. Pat. No. 5,149,645). Los tipos de mutaciones que pueden estar presentes en un gen de tabaco incluyen, por ejemplo, mutaciones puntuales, eliminaciones, inserciones, duplicaciones e inversiones. Tales mutaciones de manera deseable están presentes en la región de codificación de un gen de tabaco; sin embargo las mutaciones en la región promotora, y el intrón, o una región no traducida de un gen de tabaco también pueden ser deseables.

Por ejemplo, la mutagénesis por inserción de T-ADN puede ser utilizada para generar mutaciones por inserción en un gen de tabaco para subregular la expresión del gen. Teóricamente, se requieren aproximadamente 100,000 inserciones de T-ADN independientes para una probabilidad del 95% de obtener 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 de T-ADN en plantas pueden ser seleccionadas utilizando el análisis por reacción en cadena de polimerasa (PCR). Por ejemplo, un cebador puede ser diseñado para un extremo del T-ADN y otro cebador puede ser diseñado para el gen de interés y ambos cebadores pueden ser utilizados en el análisis por PCR. Si no se obtiene un producto por PCR, entonces no hay inserción en el gen de interés. En contraste, si se obtiene un producto por PCR, entonces hay una inserción en el gen de interés.

La expresión de un producto genético de tabaco mutado puede ser evaluada de acuerdo con procedimientos estándar (por ejemplo, los descritos aquí) 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 expresión disminuida de un gen que codifica un producto genético de tabaco son realizaciones deseables de la presente invención. Una planta que tiene una mutación en cualquiera de las secuencias de ácidos nucleicos descritas aquí puede ser utilizada en un programa de cruzamiento tal como se describe aquí.

Promotores de plantas

Un promotor deseable es un promotor de caulimovirus, por ejemplo, un promotor del virus mosaico de la coliflor (CaMV) o un promotor del virus mosaico de la vena de cassava (CsVMV). Estos promotores confieren altos niveles de expresión en la mayor parte de los tejidos vegetales, y la actividad de estos promotores no es dependiente de las proteínas codificadas viralmente. La CaMV es una fuente para los promotores 35S y 19S. Ejemplos de constructos de expresión en plantas que utilizan estos promotores son conocidos en la técnica. En la mayor parte de los tejidos de plantas transgénicas, el promotor CaMV 35S es un promotor fuerte. El promotor CaMV también es altamente activo en monocotiledóneas. Además, la actividad de este promotor puede ser incrementada adicionalmente (esto es, entre 2-10 veces) por duplicación del promotor de CaMV 35S.

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

Promotores de monocotiledóneas de ejemplo incluyen, sin limitación, el promotor del virus amarillo moteado de comelina, el promotor del virus badna de la caña de azúcar, el promotor del virus baciliforme tungro del arroz, el elemento de virus polvoso de maíz, y el promotor del virus enano de trigo.

Para ciertas aplicaciones, puede ser deseable producir un producto genético de tabaco, tal como un producto genético mutante negativo dominante, en un tejido apropiado, a un nivel apropiado, o en un tiempo de desarrollo apropiado. Para este propósito, hay un surtido de promotores genéticos, cada uno con sus propias características distintivas incorporadas en sus secuencias reguladoras, de las que se ha demostrado son reguladas en respuesta a señales inducibles tales como el ambiente, hormonas y/o señales de desarrollo. Estos incluyen, sin limitación, promotores de genes que son responsables por la expresión del gen regulada por calor, expresión de gen regulada por la luz (por ejemplo el guisante rbcS-3a; el promotor de maíz rbcS; el gen de la proteína de enlazamiento de clorofila a/b encontrado en el guisante; o el promotor Arabssu), expresión de gen regulada por hormonas (por ejemplo, las secuencias que responden al ácido abscísico (ABA) del gen Em del trigo; los promotores inducibles por ABA HVA1 y HVA22 y rd29A de cebada y Arabidopsis; y la expresión del gen inducida por lesiones (por ejemplo de wun1) la expresión de gen específica de órganos (por ejemplo, el gen de la proteína de almacenamiento específica de tubérculos; el gen de zeína de 23-kDa de maíz descrito; o el gen de la β -faseolina de alubia francesa, o los promotores inducibles por patógenos (por ejemplo, los promotores PR-1, prp-1 o de β -1,3 glucanasa, el promotor wirla inducible por hongos del trigo, y los promotores inducibles por nematodos, TobRB7-5A Y Hmg-1, de tabaco y perejil, respectivamente).

Vectores de expresión en plantas

Típicamente, los vectores de expresión en plantas incluyen (1) un gen vegetal clonado bajo el control transcripcional de la secuencia reguladora de 5' y 3' y (2) un marcador seleccionable dominante. Tales vectores de expresión en plantas también pueden contener, si se desea, una región reguladora promotora (por ejemplo, una que confiere expresión inducible o constitutiva, inducida por patógenos o lesiones, regulada por el ambiente o el desarrollo, o específica para células o tejidos), un sitio de arranque de la iniciación de la transcripción, un sitio de enlazamiento a

ribosoma, una señal de procesamiento de ARN, un sitio de terminación de la transcripción y/o una señal de poliadenilación.

5 Los vectores de expresión en plantas también pueden incluir opcionalmente señales de procesamiento de ARN, por ejemplo, intrones, los cuales han demostrado ser importantes para la síntesis y acumulación eficiente de ARN. La localización de las secuencias de empalme del ARN pueden influir dramáticamente en el nivel de la expresión transgénica en plantas. A la vista de este hecho, un intrón puede ser posicionado corriente arriba o corriente abajo de una secuencia de codificación de la nicotina desmetilasa de tabaco en el transgen para alterar los niveles de la expresión genética.

10 Además de las secuencias de control reguladoras de 5' mencionadas más arriba, los vectores de expresión también pueden incluir regiones de control reguladoras que están presentes generalmente en las regiones 3' de los genes vegetales. Por ejemplo, la región terminadora 3' puede ser incluida en el vector de expresión para incrementar la estabilidad del ARNm. Tal región terminadora puede ser derivada de la región terminadora PI-II de la patata. Además, otros terminadores utilizados comúnmente son derivados de las señales de la octopina o nopalina sintasa.

15 El vector de expresión en plantas también contiene típicamente un gen marcador seleccionable dominante utilizado para identificar aquellas células que han sido transformadas. Los genes seleccionables útiles para sistemas vegetales incluyen el gen de la aminoglicósido fosfotransferasa del transposón Tn5 (Aph II), los genes que codifican genes de resistencia a antibióticos, por ejemplo, los que codifican resistencia a higromicina, kanamicina, bleomicina; neomicina, G418, estreptomycin y espectinomycin. Los genes requeridos para la fotosíntesis también pueden ser utilizados como marcadores seleccionables en cepas deficientes en fotosíntesis. Finalmente, los genes que codifican la resistencia a los herbicidas pueden ser utilizados como marcadores seleccionables; los genes de resistencia a herbicidas útiles incluyen el gen bar que codifica la enzima fosfinotricina acetil transferasa y confiere resistencia al herbicida de amplio espectro Basta® (Bayer Cropscience Deutschland GmbH, Langenfeld, Alemania). Otros marcadores seleccionables incluyen genes que proveen resistencia a tales otros herbicidas tales como glifosfato y similares, y las imidazolinonas, sulfonilureas, herbicidas de triazolpirimidina, tales como clorosulfon, bromoxinil, dalapon y similares. Adicionalmente, los genes que codifican la dihidrofolato reductasa pueden ser utilizados en combinación con moléculas tales como metatrexato.

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

30 Además de un marcador seleccionable, puede ser deseable utilizar un gen informador. En algunos casos un gen informador puede ser utilizado sin un marcador seleccionable. Los genes informadores son genes que típicamente no están presentes o no son expresados en el organismo o tejido receptor. El gen informador codifica típicamente para una proteína que provee algún cambio fenotípico o una propiedad enzimática. Ejemplos de tales genes se proveen en Weising et al. (Ann. Rev. Genetics 22:421, 1988). Genes informadores preferidos incluyen sin limitación el gen de la glucuronidasa (GUS) y los genes de GFP.

40 Por construcción del vector de expresión en plantas, hay disponibles varios métodos estándar para la introducción del vector en un anfitrión vegetal, generando por lo tanto una planta transgénica. Estos métodos incluyen (1) transformación mediada por *Agrobacterium* (*A. tumefaciens* o *A. rhizogenes*) (véase, por ejemplo, Lichtenstein and Fuller In: Genetic Engineering, vol 6, PWJ Rigby, ed, London, Academic Press, 1987; y Lichtenstein, C.P., y Draper, J., In: DNA Cloning, Vol II, D.M. Glover, ed, Oxford, IRI Press, 1985; Patentes de los Estados Unidos Números, 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, 5,463,174, 5,469,976, y 5,464,763; y Patentes Europeas Números 0131624, 0159418, 0120516, 0176112, 0116718, 0290799, 0292435, 0320500, y 0627752, y Solicitudes de Patente Europea Números 0267159 y 0604622.), (2) el sistema de administración en partículas (véase, por ejemplo, los números de Patente de los Estados Unidos 4,945,050 y 5,141,131), (3) protocolos de microinyección, (4) procedimientos con polietileno glicol (PEG), (5) consumo de ADN mediado por liposomas, (6) protocolos de electroporación (véase, por ejemplo, WO 87/06614 y Patentes de los Estados Unidos Números 5,384,253, 5,472,869, 5,641,664, 5,679,558, 5,712,135, 6,002,070, y 6,074,877, (7) el método de vórtex, o (8) la así llamada metodología de los filamentos (véase, por ejemplo, Coffee et al., Patente de los Estados Unidos Números 5,302,523 y 5,464,765). El tipo de tejido vegetal que puede ser transformado con un vector de expresión incluye tejido de embrión, tejido de callus tipo I y II, hipocotilos, meristema y similares.

55 Una vez introducida en el tejido vegetal, la expresión del gen estructural puede ser probada por cualquier medio conocido en la técnica, y la expresión puede ser medida como ARNm transcrito, proteína sintetizada o la cantidad de silenciamiento de gen que ocurre determinada por la monitorización de metabolitos a través de análisis químico de alcaloides secundarios en tabaco (tal como se describe aquí; véase también la Patente de los Estados Unidos No. 5,583,021). Se conocen técnicas para el cultivo *in vitro* del tejido vegetal, y en un cierto número de casos, para la

regeneración en plantas completas (véase, por ejemplo, las Patentes de los Estados Unidos Números 5,595,733 y 5,766,900). Los procedimientos para transferir el complejo de expresión introducido a cultivares útiles comercialmente es conocido por los experimentados en la técnica.

5 Una vez que se obtienen las células vegetales que expresan el nivel deseado de un producto genético deseable, los tejidos vegetales y las plantas completas pueden ser regenerados a partir de las mismas utilizando métodos y técnicas bien conocidas en el arte. Las plantas regeneradas son reproducidas entonces por medios convencionales y los genes introducidos pueden ser transferidos a otras cepas y cultivares mediante técnicas de cruzamiento vegetal convencionales.

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

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

15 La expresión genética puede ser medida, por ejemplo, por análisis de inmunoprecipitación Northern estándar (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, (2001), y Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., (1989)) utilizando un gen o un fragmento de gen de tabaco como sonda de hibridación. La determinación de los niveles de expresión de ARN también puede ser auxiliada por PCR de transcripción reversa (rtPCR), incluyendo rtPCR cuantitativo (véase, por ejemplo, Kawasaki et al., in PCR Technology: Principles and Applications of RNA Amplification (H. A. Erlich, Ed.) Stockton Press (1989); Wang et al., in 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 bien conocidas adicionales para la determinación de la expresión de un gen de tabaco incluyen hibridación *in situ*, e hibridación fluorescente *in situ* (véase, por ejemplo, Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, (2001)). Las técnicas estándar anteriores también son útiles para comparar el nivel de expresión entre plantas, por ejemplo, entre una planta que tiene una mutación en un gen de tabaco y una planta de control.

20

25

Si se desea, la expresión de un gen de tabaco, (por ejemplo, una secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas) puede ser medida al nivel de producción de proteína utilizando la misma metodología general y técnicas de análisis de proteínas estándar incluyendo los ensayos de Bradford, ensayos espectrofotométricos, y técnicas de detección inmunológicas, tales como precipitación Western o inmunoprecipitación con un antibiótico específico para el polipéptido deseable (Ausubel et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, (2001), y Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., (1989)).

30

La actividad de cualquier polipéptido descrito aquí puede ser probada utilizando métodos estándar en el arte. Por ejemplo, la actividad de un p450 se prueba típicamente utilizando ensayos basados en fluorescencia (véase, por ejemplo, Donato et al. Drug Metab Dispos. 32: 699-706, 2004). En particular, la actividad de la nicotina desmetilasa puede ser probada como se describe aquí utilizando ensayos de microsomas de levadura.

35

Usos

La regulación del gen endógeno correspondiente a cualquiera de las secuencias descritas aquí, por ejemplo, por silenciamiento de genes puede dar como resultado plantas o productos de plantas más valiosos. En particular, las secuencias identificadas aquí como secuencias inducidas por etileno o relacionadas con la senescencia (por ejemplo, aquellas que tienen la secuencia de SEQ ID NO: 4, 188 y 226, o una secuencia de ácido nucleico mostrado en las figuras 1, 3 a 7 y 10-17 o un fragmento de las mismas pueden ser utilizadas para afectar las rutas metabólicas involucradas en la formación de numerosos metabolitos secundarios, incluyendo terpenoides, polifenoles, alcaloides, etc., que afectan las características de calidad del producto final. De la misma manera los genes identificados aquí pueden ser utilizados para regular rutas metabólicas asociadas con la rata y tipo de materia seca acumulada durante la senescencia o la partición de la materia seca dentro de la planta durante la senescencia. La regulación de los genes identificados aquí también pueden ser utilizada para afectar las rutas metabólicas involucradas en la determinación de las ratas de senescencia, la uniformidad de la senescencia dentro de una hoja y entre hojas de una planta individual, y la inducción de la senescencia por agentes o actividades que estimulen o activen los genes identificados aquí, y por lo tanto, controlar la calidad de un producto o artículo de manufactura que incluya una hoja u otro componente de la planta.

40

45

50

La región promotora de un gen descrita aquí puede ser utilizada para guiar la expresión de cualquier producto genético deseable para mejorar la calidad de un cultivo o potenciar características específicas. La regulación de una secuencia promotora también puede ser utilizada para subregular genes de tabaco endógenos, incluyendo genes involucrados en la biosíntesis de alcaloides y/o otras rutas. De manera deseable, una región promotora de gen de tabaco u otra región reguladora transcripcional se utiliza para alterar propiedades químicas tales como el contenido

55

de nornicotina y los niveles de nitrosamina en una planta. Además, los motivos promotores, los cuales pueden ser identificados fácilmente en una secuencia promotora utilizando métodos estándar en la técnica, pueden ser utilizados para identificar factores que se asocian con o regulan la expresión de un producto genético de tabaco, por ejemplo, un p450.

5 Además, cualquiera de las secuencias de la presente invención (por ejemplo, las secuencias de ácidos nucleicos mostradas en la figura 1, figuras 3 a 7, y 10-17, o fragmentos de las mismas) pueden ser utilizadas en métodos que reducen la expresión genética o alteran la actividad enzimática de un producto genético, tal como un p450, utilizando técnicas estándar descritas aquí. Tales técnicas incluyen, sin limitación, interferencia por ARN, interferencia con cadena triple, ribozimas, recombinación homóloga, silenciamiento de genes inducido por virus, tecnologías antisentido de cosupresión, expresión de un producto genético negativo dominante, 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 pueden utilizarse para alterar los ácidos grasos que están involucrados en las interacciones planta-patógeno y en la resistencia a enfermedades o pueden utilizarse para alterar el perfil de la planta de ácidos grasos seleccionados y por lo tanto alterar el sabor o aroma de la planta o componente de la planta.

Adicionalmente, utilizando métodos estándar, cualquier porción de un gen de tabaco, incluyendo el promotor, la secuencia codificadora, un intrón o una 3' UTP, o un fragmento del mismo, pueden ser utilizados como marcador genético para aislar genes, promotores o regiones reguladoras relacionados, para seleccionar el gen relacionado en otra especie de tabaco o de *Nicotiana*, o para determinar si una planta tiene una mutación en un gen endógeno correspondiente. Una porción de un gen de tabaco también puede ser utilizada para monitorizar el flujo de genes a través de un esfuerzo de cruzamiento para seguir la intergresión o pérdida de un gen en particular.

Por ejemplo, la *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 cual reside el gen original. Un marcador para el gen relacionado también podría ser utilizado para seleccionar germoplasma de tabaco existente, segregar o sintetizar poblaciones creadas por hibridaciones, poblaciones creadas a partir de tratamientos mutagénicos o a partir de diversos métodos de cultivo de tejidos. Como tales, las secuencias de ácidos nucleicos descritas aquí (por ejemplo, las secuencias de ácidos nucleicos mostradas en la figura 1, figuras 3 a 7 y 10-17 o fragmentos de las mismas) pueden ser utilizadas para identificar o afectar genes involucrados en resistencia a enfermedades o a insectos, propiedades de sabor y aroma, tolerancia a herbicidas, factores de calidad relacionados con constituyentes indeseables, o que incrementan el rendimiento de la hoja, o afectan los componentes de la hoja o la planta, tales como ligninas, celulosa, etc., relacionados con características estructurales o contenido de fibra.

Productos

Los productos de tabaco que tienen una cantidad reducida de contenido de nitrosamina son manufacturados utilizando cualquiera de los materiales de la planta de tabaco descritos aquí de acuerdo con métodos estándar conocidos en la técnica. En una realización, los productos de tabaco son manufacturados utilizando material vegetal de tabaco obtenido a partir de una planta de tabaco curada. La planta de tabaco curada puede contener o haber sido cruzada para contener actividad reducida de nicotina desmetilasa. Por ejemplo, la planta de tabaco curada puede ser una planta de tabaco resultante de un cruce que incluye una planta de tabaco identificada por tener expresión variante de nicotina desmetilasa. De manera deseable, 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 en ella es menor de 90%, 70%, 50%, 30%, 10%, 5%, 4%, 3%, 2%, 1.5%, 1%, 0.75%, 0.5%, 0.25%, o 0.1%. La expresión "una cantidad reducida" se refiere a una cantidad de nornicotina o NNN o de ambas en una planta o componente de una planta de tabaco o un producto de tabaco que es menor de lo que se encontraría en una planta o componente de planta o producto de tabaco tipo silvestre a partir de la misma variedad de tabaco, procesada de la misma manera, que no ha sido hecha transgénica para buscar nornicotina o NNN reducidas. En un ejemplo, una planta de tabaco tipo silvestre de la misma variedad ha sido procesada de la misma manera y utilizada como control para medir si una reducción de nornicotina o NNN o ambas ha sido obtenida por los métodos descritos aquí. En otro ejemplo, las plantas que tienen una cantidad reducida de contenido de nitrosamina son evaluadas utilizando métodos estándar, por ejemplo, monitorizando la presencia o ausencia de un gen o un producto genético, por ejemplo, una nicotina desmetilasa, o una mutación particular en un gen. En todavía otro ejemplo, el contenido de nitrosamina de las plantas resultantes de un programa de cruzamiento se compara con el contenido de nitrosamina de la línea receptora o línea donante, o ambas, utilizadas para cruzar 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 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 deberían entenderse como ilustrativos de, pero no limitantes del alcance de la invención el cual está definido en las reivindicaciones anexas.

Ejemplo 1Desarrollo de tejido vegetal y tratamiento con etilenoCrecimiento de la planta

5 Las plantas fueron sembradas en macetas y cultivadas en un invernadero durante 4 semanas. Los brotes de 4 semana de edad fueron trasplantados a macetas individuales y cultivados en el invernadero durante 2 meses. Las plantas fueron regadas 2 veces al día con agua que contenía 150 ppm de fertilizante NPK durante el crecimiento. Las hojas verdes expandidas fueron desprendidas de las plantas para realizar el tratamiento con etileno descrito más adelante.

Línea celular 78379

10 La línea 78379 de tabaco que es una línea de tabaco Burley producida por la Universidad de Kentucky fue utilizada como una fuente de material vegetal. Se cultivaron 100 plantas como ese estándar en el arte de cultivo del tabaco, trasplantadas y etiquetadas con un número distintivo (1-100). La fertilización y el manejo de campo fueron llevados a cabo según se recomienda.

15 Tres cuartos de las 100 plantas convirtieron entre 20 y 100% de la nicotina a nornicotina. Un cuarto de las cien plantas convirtió menos del 5% de la nicotina a nornicotina. La planta número 87 tuvo la menor conversión (2%) mientras que la planta número 21 tuvo 100% de conversión. Las plantas que convirtieron menos del 3% fueron clasificadas como no convertidoras. Se hicieron semillas autopolinizadas de la planta número 87 y de la planta número 21 así como semillas cruzadas (21 x 87 y 87 x 21) para estudiar las diferencias genéticas y fenotípicas. Las plantas de 21 autopolinizada fueron convertidoras, y el 99% de las autopolinizadas de 87 fueron no convertidoras. El otro 1% de las plantas de 87 mostraron baja conversión (5-15%). Las plantas de cruzamiento recíprocos fueron todas convertidoras.

Línea celular 4407

25 La línea 4407 de *Nicotiana*, que es una línea Burley, fue utilizada como fuente de material vegetal. Se seleccionaron plantas uniformes y representativas (100) y se etiquetaron. De las 100 plantas 97 fueron no convertidoras y tres fueron convertidoras. La planta número 56 tuvo la menor cantidad de conversión (1.2%) y la planta número 58 tuvo el nivel más alto de conversión (96%). Con estas dos plantas se hicieron semillas autopolinizadas y semillas cruzadas.

30 Las plantas autopolinizadas 58 se segregaron con convertidor 3:1 a una relación no convertidora. Las plantas 58-33 y 58-25 fueron identificadas como líneas de plantas convertidoras y no convertidoras homocigóticas, respectivamente. La conversión estable de 58-33 fue confirmada por análisis de su progenie.

Línea celular PBLB01

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

Procedimientos de tratamiento con etileno

35 Las hojas verdes fueron desprendidas de plantas cultivadas en invernadero de 2-3 meses y asperjadas con 0.3% de solución de etileno (Prep brand Ethephon (Rhone-Poulenc)). Cada hoja asperjada fue colgada en una gradilla de curado equipada con humidificador y cubierta con plástico. Durante el tratamiento, las hojas de muestra fueron asperjadas periódicamente con la solución de etileno. Aproximadamente 24-48 horas después del tratamiento con etileno, las hojas fueron recolectadas para la extracción de ARN. Otra submuestra fue tomada para análisis de constituyentes metabólicos para determinar la concentración de metabolitos en la hoja y constituyentes más específicos de interés tales como una variedad de alcaloides.

45 Como ejemplo, el análisis de alcaloides pudo ser llevado a cabo como sigue. Las muestras (0.1 g) fueron agitadas a 150 rpm con 0,5 ml de NaOH 2N, y 5 ml de una solución de extracción que contenía quinolina como estándar interno y metil t-butil éter. Las muestras fueron analizadas en un GC HP 6890 equipado con un detector FID. Se utilizó una temperatura de 250°C para el detector y el inyector. Se utilizó una columna HP (30 m – 0.32 mm-ID) consistente de sílica fusionada entrecruzada con fenol al 5% y metil silicona al 95% a un gradiente de temperatura de 110-185°C a 10°C por minuto. La columna fue operada a 100°C con una tasa de flujo de 1.7 cm³ minuto⁻¹ con una relación de división de 40:1 y un volumen de inyección de 2:1 utilizando helio como gas transportador.

Ejemplo 2

50 Aislamiento de ARN

Para las extracciones de ARN, se trataron hojas medias de plantas cultivadas en invernadero de dos meses de edad con etileno como se describió más arriba. Las muestras de 0 y 24-48 horas fueron utilizadas para la extracción de ARN. En algunos casos, se tomaron muestras de hojas bajo el proceso de senescencia de las plantas 10 días después de la remoción de la cabeza de la flor. Estas muestras fueron utilizadas también para extracción. El ARN total fue aislado utilizando el Rneasy Plant Mini Kit[®] (Qiagen, Inc., Valencia, California) de acuerdo con el protocolo del fabricante.

La muestra de tejido fue triturada bajo nitrógeno líquido hasta un polvo fino utilizando un mortero y pistilo tratados con DEPC. Se transfirieron aproximadamente 100 miligramos de tejido triturado a un tubo Eppendorf de 1.5 ml estéril. Este tubo de muestra fue colocado en nitrógeno líquido hasta que se recolectaron todas las muestras. Luego, se agregaron a cada tubo individual 450 µl de regulador RLT según se suministra en el kit (con la adición de mercaptoetanol). La muestra fue sometida a vórtex vigorosamente e incubada a 56°C durante 3 minutos. El lisado fue aplicado entonces al asiento de la columna de rotación QIAshredder[®] en un tubo de recolección de 2 ml, y fue centrifugado durante 2 minutos a velocidad máxima. El flujo fue recolectado y se agregaron 0.5 volúmenes de etanol al lisado clarificado. La mezcla fue bien mezclada y transferida a una columna Rneasy[®] mini giratoria en un tubo de recolección de 2 ml. La muestra fue centrifugada durante 1 minuto a 10,000 rpm. A continuación se agregaron mediante pipeta 700 µl de regulador RW1 sobre la columna Rneasy[®] y se centrifugó a 10,000 rpm. Se agregó con pipeta regulador RPE sobre la columna Rneasy[®] en un nuevo tubo de recolección y se centrifugó durante 1 minuto a 10,000 rpm. Se agregó de nuevo regulador RPE a la columna Rneasy[®] giratoria y se centrifugó durante 2 minutos a velocidad máxima para secar la membrana. Para eliminar cualquier etanol arrastrado, la membrana fue colocada en un tubo de recolección separado y centrifugado durante 1 minuto adicional a velocidad máxima. La columna Rneasy[®] fue transferida a un nuevo tubo de recolección de 1.5 ml y se agregaron con pipeta 40 µl de agua libre de RNasa sobre la membrana Rneasy[®]. Este tubo eluido final fue centrifugado durante 1 minuto a 10,000 rpm. La calidad y la cantidad del ARN total fueron analizadas mediante gel desnaturalizado con formaldehído y espectrofotometría.

El poli(A) ARN fue aislado utilizando el kit de purificación Oligotex[®] poli A + ARN (Qiagen Inc.) siguiendo el protocolo del fabricante. Se utilizaron aproximadamente 200 µg de ARN total en un volumen máximo de 250 µl. Se agregó un volumen de 250 µl de regulador OBB y 15 µl de suspensión de Oligotex[®] a los 250 µl de ARN total. Los contenidos fueron mezclados exhaustivamente mediante pipeta y se incubaron durante 3 minutos a 70°C sobre un bloque de calentamiento. La muestra fue colocada entonces a temperatura ambiente durante aproximadamente 20 minutos. El complejo Oligotex[®]:ARNm fue convertido en pellas por centrifugación durante 2 minutos a velocidad máxima. Todo excepto 50 µl del sobrenadante fue retirado del tubo de microcentrífuga. La muestra fue tratada adicionalmente con regulador OBB. La pella de Oligotex[®]:ARNm fue resuspendida en 400 µl de regulador OW2 mediante vórtex. Esta mezcla fue transferida sobre una columna de rotación pequeña colocada en un nuevo tubo y centrifugada durante 1 minuto a velocidad máxima. La columna de rotación fue transferida a un nuevo tubo y se agregaron 400 µl adicionales del regulador OW2 a la columna. El tubo fue centrifugado entonces durante 1 minuto a velocidad máxima. La columna de rotación fue transferida a un tubo de microcentrífuga final de 1.5 ml. La muestra fue eluida con 60 µl de regulador OEB caliente (70°C). El producto poli A fue analizado por geles de formaldehído desnaturalizados y análisis espectrofotométrico.

Ejemplo 3

PCR de transcripción reversa

La primera cadena de ADNc fue producida utilizando la transcriptasa reversa SuperScript siguiendo el protocolo del fabricante (Invitrogen, Carlsbad, California). La mezcla de ARN enriquecido en poli A+/cebador oligo dT consistía de menos de 5 µg de ARN total, 1 µl de mezcla de 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 fue incubada a 65°C durante 5 minutos, luego colocadas sobre hielo durante al menos 1 minuto. Se preparó una mezcla de reacción agregando 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 RNasa OUT Recombinant RNase Inhibitor. Una adición de 9 µl de mezcla de reacción fue transferida con pipeta a cada mezcla ARN/cebador y se mezcló suavemente. Se incubó a 42°C durante 2 minutos y se agregó 1 µl de Super Script II RT a cada tubo. El tubo fue incubado durante 50 minutos a 42°C. La reacción fue terminada a 70°C durante 15 minutos y enfriada sobre hielo. La muestra fue recolectada por centrifugación y se agregó 1 µl de RNasa H a cada tubo y se incubó durante 20 minutos a 37°C. El segundo PCR fue llevado a cabo con 200 pmoles del cebador de avance y 100 pmoles del cebador reverso (mezcla de 18 nt oligo d(T) seguida por una base aleatoria).

Las condiciones de reacción fueron 94°C durante 2 minutos y luego 40 ciclos de PCR a 94°C durante 1 minuto, 45°C hasta 60°C durante 2 minutos, 72°C durante 3 minutos, con una extensión a 72°C durante 10 minutos extra. Se analizaron diez microlitros de la muestra amplificada por electroforesis utilizando un gel de agarosa al 1%. Los fragmentos de tamaño correcto fueron purificados a partir del gel de agarosa.

Ejemplo 4

Generación de las poblaciones de fragmentos de PCR

Los fragmentos de PCR del Ejemplo 3 fueron ligados sobre un pGEM-T Easy Vector (Promega, Madison, Wisconsin) siguiendo las instrucciones del fabricante. El producto ligado fue transformado en células competentes JM 109 y sembrados sobre placas de medio LB para selección azul/blanco. Las colonias fueron seleccionadas y cultivadas en una placa de 96 pozos con 1.2 ml de medio LB durante la noche a 37°C. Se generó una reserva congelada para todas las colonias seleccionadas. El ADN de plásmido fue purificado de las placas utilizando el Beckman Biomeck 200 miniprep robotics con Wizard SV Miniprep kit (Promega). El ADN en el plásmido fue eludido con 100 µl de agua y almacenado en una placa de 96 pozos. Los plásmidos fueron digeridos con EcoRI y luego analizados utilizando gel de agarosa al 1% para confirmar la cantidad de ADN y el tamaño de los insertos. Los plásmidos que contenían un inserto de 400-600 bp fueron secuenciados utilizando un secuenciador CEQ 2000 (Beckman, Fullerton, California). Las secuencias fueron alineadas con la base de datos GenBank con búsqueda BLAST. Los fragmentos relacionados con p450 fueron identificados y analizados posteriormente. Alternativamente, se aislaron fragmentos de p450 a partir de bibliotecas de sustracción. Estos fragmentos también fueron analizados como se describió anteriormente.

Ejemplo 5

Construcción de una biblioteca de ADNc

Se construyó una biblioteca ADNc preparando ARN total a partir de hojas tratadas con etileno como sigue. Primero, se extrajo el ARN total a partir de las hojas tratadas con etileno de la línea de tabaco 58-33 utilizando un protocolo de extracción modificado con fenol ácido y cloroformo. El protocolo fue modificado para utilizar un gramo de tejido que fue triturado y subsecuentemente sometido a vórtex en 5 ml de regulador de extracción (Tris-HCl 100 mM, pH 8.5; NaCl 200 mM; EDTA 10 mM; SDS al 0,5%) al cual se agregaron 5 ml de fenol (pH 5.5) y 5 ml de cloroformo. La muestra extraída fue centrifugada y se conservó el sobrenadante. Esta etapa de extracción fue repetida 2-3 veces hasta que el sobrenadante apareció claro. Se agregaron aproximadamente 5 ml de cloroformo para remover cantidades traza de fenol. El ARN fue precipitado a partir de las fracciones sobrenadantes combinadas agregando 3 veces el volumen de etanol y 1/10 veces el volumen de NaOAc 3M (pH 5.2) y almacenando a -20°C durante 1 hora. Después de transferir a un contenedor de vidrio Corex la fracción de ARN fue centrifugada a 9,000 RPM durante 45 minutos a 4°C. La pella fue lavada con etanol al 70% y se hizo rotar durante 5 minutos a 9,000 RPM a 4°C. Después de secar la pella, el ARN en forma de pella fue disuelta en 0.5 ml de agua libre de RNasa. La calidad y cantidad del ARN total fue analizada por gel desnaturalizado con formaldehído y espectrofotómetro, respectivamente.

El ARN total resultante fue utilizado para aislar poli A+ ARN utilizando un protocolo de celulosa oligo (dT) (Invitrogen) y columnas de rotación de microcentrífuga (Invitrogen) mediante el siguiente protocolo. Aproximadamente veinte mg de ARN total fueron sometidos dos veces a purificación para obtener poli A+ ARN de alta calidad. El producto de poli A+ ARN fue analizado ejecutando un gel desnaturalizado con formaldehído y subsecuente RT-PCR de genes de longitud completa conocidos para asegurar la alta calidad del ARNm.

A continuación se utilizó poli A+ ARN como plantilla para producir una biblioteca de ADNc empleando el kit de síntesis de ADNc, kit de síntesis ZAP-cDNA, y el kit de clonación ZAP-cDNA Gigapack III gold (Stratagene, La Jolla, California). El método involucró seguir el protocolo del fabricante según se especifica. Se utilizaron aproximadamente 8 µg de poli A+ ARN para construir la biblioteca de ADNc. El análisis de la biblioteca primaria reveló aproximadamente $2.5 \times 10^6 - 1 \times 10^7$ pfu. Una prueba de fondo de calidad de la biblioteca fue completada por ensayos de complementación utilizando IPTG y X-gal, donde placas recombinantes fueron expresadas a más de 100 veces por encima de la reacción de fondo.

Un análisis más cuantitativo de la biblioteca por PCR aleatorio mostró que el tamaño promedio del ADNc inserto fue aproximadamente 1.2 kb. El método utilizó un método de PCR de dos etapas. Para la primera etapa, los cebadores reversos fueron diseñados con base en la información de secuencia preliminar obtenida a partir de los fragmentos p450. Los cebadores reversos diseñados y los cebadores T3 (de avance) fueron utilizados para amplificar los genes correspondientes de la biblioteca de ADNc. Las reacciones de PCR fueron sometidas a electroforesis en agarosa y las bandas correspondientes de alto peso molecular fueron seccionadas, purificadas, clonadas y secuenciadas. En la segunda etapa, los nuevos cebadores diseñados para 5' UTR o la región de inicio de codificación de p450 como los cebadores de avance juntos con los cebadores reversos (diseñados a partir de 3' UTR de p450) fueron utilizados en la PCR subsecuente para obtener clones de p450 de longitud completa.

Los fragmentos de p450 fueron generados por amplificación por PCR a partir de la biblioteca de ADNc construida tal como se describe en el Ejemplo 3 con excepción del cebador reverso. EL cebador T7 localizado en el plásmido corriente abajo de los insertos de ADNc fue utilizado como cebador reverso. Los fragmentos de PCR fueron aislados, clonados y secuenciados como se describe en el Ejemplo 4.

Los genes p450 de longitud completa fueron aislados por este método de PCR a partir de la biblioteca de ADNc construida. Los cebadores reversos específicos del gen (diseñados a partir de la secuencia corriente debajo de los fragmentos p450) y un cebador de avance (T3 sobre el plásmido de la biblioteca) fueron utilizados para clonar los genes de longitud completa. Los fragmentos de PCR fueron aislados, clonados y secuenciados. En caso necesario, se aplicó una segunda etapa de PCR. En la segunda etapa, nuevos cebadores de avance diseñados a partir de la 5' UTR de p450 clonados junto con los cebadores reversos diseñados a partir de 3' UTR de clones de p450 fueron

utilizados en las reacciones de PCR subsecuentes para obtener clones de p450 de longitud completa. Los clones fueron secuenciados subsecuentemente.

Ejemplo 6

5 Identidad de ácidos nucleicos, relación estructural de fragmentos de ácidos nucleicos aislados e hibridación en GeneChip®

10 Fueron secuenciados más de 100 fragmentos de p450 clonados en conjunción con análisis por inmunoprecipitación Northern para determinar su relación estructural. La metodología utilizó cebadores de avance basados en cualquiera de los dos motivos de p450 comunes localizados cerca del terminal carboxilo de los genes de p450. Los cebadores de avance corresponden a los motivos del citocromo p450 FXPERF (SEQ ID NO: 2268) o GRRXCP (A/G) (SEQ ID NO: 2269). Los cebadores reversos utilizaron cebadores estándar de cualquiera del plásmido, SP6 o T7 localizado en ambos brazos del plásmido pGEM, o una cola poli A. A continuación se describe el protocolo usado.

15 Se utilizó espectrofotometría para estimar la concentración del ADN de cadena doble de partida siguiendo el protocolo del fabricante (Beckman Coulter). La plantilla fue eluida con agua a la concentración apropiada, desnaturalizada por calentamiento a 95°C durante 2 minutos, y subsecuentemente colocada sobre hielo. La reacción de secuenciamiento fue preparada sobre hielo utilizando 0.5 a 10 µl de la plantilla de ADN desnaturalizada, 2 µl de 1.6 pmol del cebador de avance, 8 µl del DTCS Quick Start Master Mix y el volumen total fue llevado a 20 µl con agua. El programa de termociclado consistió de 30 ciclos del ciclo siguiente: 96°C durante 20 segundos, 50°C durante 20 segundos y 60°C durante 4 minutos seguido por sostenimiento a 4°C.

20 La reacción de secuenciamiento fue detenida agregando 5 µl de regulador de detención (volumen igual de NaOAc 3M y EDTA 100 mM y 1 µl de 20 mg/ml de glicógeno. La muestra fue precipitada con 60 µl de etanol frío al 95% y centrifugada a 6000 x g durante 6 minutos. Se descartó el etanol. La pella fue lavada dos veces con 200 µl de etanol frío al 70%. Después de que la pella estaba seca, se agregaron 40 µl de solución SLS y la pella fue resuspendida. Se sobrepuso una capa de aceite mineral y la muestra fue colocada en el CEQ 8000 Automated Sequencer para análisis posterior.

25 Para verificar las secuencias de ácidos nucleicos, la secuencia de ácidos nucleicos fue resecuenciada en ambas direcciones utilizando cebadores de avance para la región FXPERF (SEQ ID NO: 2268) o GRRXCP(A/G)(SEQ ID NO: 2269) del gen p450 o cebadores reversos para cualquiera del plásmido o de la cola poli A. Todo el secuenciamiento fue llevado a cabo al menos dos veces en ambas direcciones.

30 Las secuencias de ácidos nucleicos de los fragmentos del citocromo p450 fueron comparadas una con otra a partir de la región de codificación correspondiente al primer ácido nucleico después de la región que codifica el motivo GRRXC(A/G) (SEQ ID NO: 2269) a través del codón de detención. Esta región fue seleccionada como indicadora de la diversidad genética entre proteínas de p450. Un gran número de genes de p450 distinguibles genéticamente, por encima de 70 genes, fueron observados, similares a los de otras especies de plantas. Por comparación de las secuencia de ácidos nucleicos, se encontró que los genes podrían ser colocados en distintos grupos de secuencias con base en su identidad de secuencias. Se encontró que el mejor agrupamiento único de miembros de p450 fue determinado como aquellas secuencias con 75% o más de identidad en ácidos nucleicos. (Véase, por ejemplo, Tabla 1 de la Publicación de Solicitud de Patente US 2004/0162420).

40 Se utilizó hibridación en microarreglo GeneChip® (Affymetrix Inc.; Santa Clara, CA) para identificar genes con diferentes patrones de expresión entre las líneas isogénicas cercanas convertidoras y no convertidoras después de la activación con etileno. El tamaño del chip fue de 18 micrones y el formato de arreglo fue de 100-2187, acomodando 528 conjuntos de sondas (11,628 sondas). Se utilizaron 7 pares de hibridación para obtener una verificación independiente de los resultados del microarreglo. Estos consistieron de un par (convertidor/no convertidor) de muestras de tabaco Burley no tratado 4407-33/4407-25, cuatro pares de muestras 4407-33/4407-25 tratadas con etileno, un par de muestras 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 senescentes de manera natural de 45 4407=33/25 (Tabla 1).

Tabla 1 Relaciones de señal normalizada convertidor:no convertidor de hibridación por GeneChip®

	Burley no tratado (4407-33/25)	Burley tratado con etileno (4407-33/25)				Oscuro tratado con etileno (178/NL Madole)	Burley senescente (4407-33/25)
		Exp 1	Exp 2	Exp 3	Exp 4		
Inducido							
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-BG11	1.73	2.32	13.06	22.22	19.10	28.76	3.40
Control							
I similar a actina (5´)	1.18	0.99	0.74	0.73	0.57	1.02	0.97
I similar a actina (3´)	1.09	1.12	0.81	1.08	0.79	0.93	0.85

5 Todos los 14 conjuntos de hibridaciones fueron exitosos según fue evidenciado por el Reporte de Expresión generado utilizando instrumentos de detección de Genome Explorations, Inc. (Memphis, TN). Los reportes principales incluyen análisis de ruido, factor de escala, fondo, conjuntos de sonda totales, número y porcentaje de conjuntos de sondas presentes y ausentes, intensidad de señal de controles de mantenimiento. Los datos fueron analizados subsecuentemente presentados utilizando el software GCOS en combinación con otro software. Se hicieron comparaciones de señal entre pares de tratamiento, y los datos globales para todas las sondas respectivas para todas las hibridaciones fueron compilados y los datos de expresión también fueron analizados. Los resultados basados en las intensidades de señal mostraron que solamente dos genes, D121-AA8 y D120-AH4 y un fragmento D35-BG11, el cual es un fragmento parcial de D121-AA8, tenían inducción reproducible en líneas convertidoras tratadas con etileno en comparación con líneas no convertidoras. Las señales de un gen en una línea convertidora, por ejemplo, la variedad de tabaco Burley 4407-33, fue determinada como la relación de la señal de un gen en una línea isogénica no convertidora relacionada, 4407-25. Sin tratamiento con etileno, la relación de las señales de convertidor a no convertidor para todos los genes se aproximó a 1.00. Para eliminar la influencia de las diferencias de fondo, también se calcularon relaciones de señal normalizadas. Las relaciones de señal normalizadas son obtenidas dividiendo la relación del par tratado con la correspondiente relación del no par correspondiente. Después del tratamiento con etileno y análisis, se determinó que dos genes, D121-AA8 y D120-AH4, fueron inducidos en líneas convertidoras con respecto a líneas no convertidoras según se determinó por cuatro análisis independientes.

10 Estos dos genes comparten un 99.8 de homología relativa y sus señales de hibridación relativas en variedades convertidoras variaron desde aproximadamente 2 hasta 22 veces más que las señales de sus contrapartes no convertidoras. Con base en las relaciones normalizadas, dos clones de control internos similares a actina no fueron inducidos en líneas convertidoras. Además, un fragmento (D35-BG11) cuya región de codificación está completamente contenida en los genes D121-AA8 y D120-AH4, fue altamente inducido en las mismas muestras de líneas convertidoras y no convertidoras isogénicas pareadas. Adicionalmente, los genes D121-AA8 y D120-AH4 fueron inducidos fuertemente en líneas convertidoras de pares de tabaco oscuro isogénicas, NL Madole y 181 (8 a 28 veces), demostrando así que la inducción por etileno de estos genes en líneas convertidoras fue en una respuesta *in planta*. Los mismos genes fueron identificados en las comparaciones hechas a partir de hibridaciones de muestras senescentes de manera natural de 4407-33/25 también. Los ensayos de RT-PCR de estos materiales utilizando cebadores específicos para D121-AA8 verificaron los resultados en microarreglo para este gen

25 Con base en estos resultados, el gen D121-AA8 (la secuencia ADNc de la cual es la secuencia de SEQ ID NO: 5; figura 4) fue identificado como el gen de nicotina desmetilasa de tabaco de interés. A la vista de las reglas de nomenclatura de p450, se determinó que D121-AA8 es el más similar al p450 en la familia CYP82E (The Arabidopsis Genome Initiative (AGT) and The Arabidopsis Information Resource (TAIR); Frank, Plant Physiol. 110:1035-1046, 1996; Whitbred et al., Plant Physiol. 124:47-58,2000); Schopfer and 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, tal como se describe en solicitudes presentadas previamente, determinó que la secuencia de SEQ ID NO: 5 codifica una nicotina desmetilasa de tabaco (SEQ ID NO: 3; figuras 3 y 4).

5 En particular, la función del clon candidato D121-AA8 fue confirmada como gen de codificación para la nicotina desmetilasa, probando la actividad de la enzima p450 expresado de manera heteróloga en células levadura como sigue:

1. Construcción del vector de expresión en levadura

10 La secuencia de codificación de proteína putativa del ADNc que codifica la nicotina desmetilasa de tabaco (D121-AA8), D120-AH4, D121-AA8, 208-AC-8, y D208-AD9, fueron clonados en el vector de expresión en levaduras pYeDP60. Se introdujeron sitios apropiados BamHI y MfeI (subrayados más abajo) a través de cebadores de PCR que contenían estas secuencias bien sea corriente arriba del codón de inicio de la traducción (ATG) o corriente abajo del codón de detención (TAA). El MfeI en el producto de PCR amplificado es compatible con el sitio EcoRI del vector. Los cebadores utilizados para amplificar el ADNc de D121-AA8 fueron 5'-TAGCTACGCGGATCCATGCTTTCTCCCATAGAAGCC-3' (SEQ ID NO: 2194) y 5'-CTGGATCACAATTGTTAGTGATGGTGATGGTGATGCGATCCTCTATAAAGCTCAGGTGCCAGGC-3' (SEQ ID NO: 2297). Un segmento de secuencia que codifica nueve aminoácidos extra en el terminal C de la proteína incluyendo seis histidinas fue incorporado en el cebador reverso para facilitar la expresión del p450 etiquetado con 6-His por inducción. Los productos de PCR fueron ligados en el vector pYeDP60 después de las digestiones enzimáticas en la orientación en sentido con referencia al promotor GAL10-CYC1. La construcción apropiada de los vectores de expresión en levadura fue verificada por el análisis con enzima de restricción y secuenciamiento de ADN. Además, la expresión de las proteínas de p450 fue visualizada en electroforesis en gel SDS-PAGE para la fase detergente de los microsomas de levadura. El tamaño predicho de las proteínas p450 es 59 kD, con base en la secuencia genética; se confirmó un resultado mediante el análisis en gel.

2. Transformación de levadura

25 La línea de levadura WAT11, modificada para expresar la NADPH-citocromo p450 reductasa de Arabidopsis ATR1, fue transformada con los plásmidos de ADNc pYeDP60-p450. Se mezclaron cincuenta microlitros de suspensión de células de levadura WAT11 con ~ 1 µg de ADN de plásmido en una cubeta con una brecha de electrodo de 0.2 cm. Se aplicó un pulso a 2.0 kV mediante un electroporador Eppendorf (Modelo 2510). Las células fueron esparcidas sobre placas SGI (5 g/L de bactocasamino ácidos, 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 por análisis por PCR llevado a cabo directamente sobre colonias seleccionadas aleatoriamente.

3. Expresión de p450 en células de levadura transformadas

35 Se utilizaron colonias de levadura individuales para inocular 30 mL de medio SGI (5 g/L de bactocasamino ácidos, 6.7 g/L de bases nitrogenadas 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. Un alícuota de este cultivo fue diluida 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 cultivo hasta que la glucosa fue consumida completamente como lo indicó el cambio colorimétrico de una tira de reactivo de urianálisis Diastix (Bayer, Elkhart, IN). La inducción del p450 clonado fue iniciada agregando DL-galactosa a una concentración final de 2%. Los cultivos se dejaron crecer durante 20 horas adicionales antes de ser utilizados para el ensayo de actividad *in vivo* o para la preparación de microsomas.

40 Las células de levadura WAT11 que expresaban pYeDP60-CYP71D20 (un p450 que cataliza la hidroxilación de 5-epi-aristolóqueno y 1-desoxicapsidiol en *Nicotiana tabacum*) fueron utilizadas como control para la expresión de p450 y los ensayos de actividad enzimática.

45 Para evaluar la efectividad de la expresión en levadura de p450 en gran detalle, se llevó a cabo espectroscopia diferencial con CO reducido. El espectro con CO reducido exhibió un pico a 450 nm de proteínas para todas las cuatro líneas de levadura transformadas con p450. No se lograron picos similares en los microsomas de control derivados de las células de levadura no transformadas de control o las células de levadura con control de vector blanco. Los resultados indicaron que las proteínas de p450 fueron expresadas efectivamente en líneas celulares que alojan el pYeDP60-CYP 450. Las concentraciones de la proteína de p450 expresada en microsomas de levadura varió de 45 a 68 nmol/mg de proteína total.

4. Ensayo enzimático *in vivo*

50 La actividad de la nicotina desmetilasa en las células de levadura transformadas fue probada alimentando el cultivo de levadura con DL-Nicotina (Pirrolidina-2-¹⁴C). La nicotina marcada con ¹⁴C (54 mCi/mmol) fue agregada a 75 µl del cultivo inducido con galactosa hasta una concentración final de 55 µM. El cultivo de ensayo fue incubado con

agitación en tubos de polipropileno de 14 ml durante 6 horas y luego extraído con 900 µl de metanol. Después de rotar, se separaron 20 µl del extracto en metanol con una rp-HPLC y la fracción de nornicotina fue cuantificada por LSC.

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

5. Preparación de microsomas de levadura

10 Después de la inducción con galactosa durante 20 horas, las células de levadura fueron recolectadas por centrifugación y lavadas dos veces con regulador TES-M (Tris-HCl 50 mM, pH 7.5, EDTA 1 mM, sorbitol 0.6 M, 2-mercaptoethanol 10 mM). La pella fue resuspendida en regulador de extracción (Tris-HCl 50 mM, pH 7.5, EDTA 1 mM, sorbitol 0.6 M, 2-mercaptoethanol 2 mM, albumina de suero bovino al 1%, Cóctel Inhibidor de Proteasa (Roche) a 1 tableta/50 ml). Las células fueron rotas entonces con perlas de vidrio (0.5 mm de diámetro, Sigma) y el extracto celular fue centrifugado durante 20 minutos a 20,000 x g para remover los residuos celulares. El sobrenadante fue sometido a ultracentrifugación a 100,000 x g durante 60 minutos y la pella resultante contenía la fracción microsómica. La fracción microsómica fue suspendida en el regulador TEG-M (Tris-HCl 50 mM, pH 7.5, EDTA 1 mM, glicerol al 20% y 2-mercaptoethanol 1.5 mM) a una concentración de proteína de 1 mg/mL. Las preparaciones microsómicas fueron almacenadas en un congelador con nitrógeno líquido hasta el uso.

20 6. Ensayo de actividad enzimática en preparaciones microsómicas de levadura

Se llevaron a cabo ensayos de actividad de la nicotina desmetilasa con preparaciones microsómicas de levadura. En particular la DL-nicotina (pirrolidina-2-¹⁴C) fue obtenida a partir de Moravek Biochemicals y tenía una actividad específica de 54 mCi/mmol. La clorpromazina (CPZ) y el citocromo oxidado c (cyt. C) ambos inhibidores de P450, fueron comprados de Sigma. La forma reducida de nicotinamida adenina dinucleótido fosfato (NADPH) es el donante de electrones típico para citocromo P450 a través de la NADPH:citocromo P450 reductasa. El NADPH fue obtenido para la incubación de control. El ensayo enzimático de rutina incluyó proteínas microsómicas (alrededor de 1 mg/ml), NADPH 6 mM, y nicotina marcada con ¹⁴C 55 µM. La concentración de CPZ y Cyt. C, cuando se utilizaron, fue de 1 mM y 100 µM, respectivamente. La reacción fue llevada a cabo a 25°C durante 1 hora y fue detenida con la adición de 300 µl de metanol a cada 25 µl de mezcla de reacción. Después de la centrifugación, se separaron 20 µl del extracto metanólico con un sistema de Cromatografía Líquida de Alto Rendimiento en fase reversa (HPLC) (Agilent) utilizando una columna de cromatografía Inertsil ODS-33 µ (150 x 4.6 mm) de Varian. La fase móvil isocrática fue 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 rata de flujo fue de 1 ml/minuto. El pico de la nornicotina, tal como se determino por comparación con nornicotina autentica no marcada, fue recolectada y sometida a un 2900 tri-carb Liquid Scintillation Counter (LSC) (Perkin Elmer) para su cuantificación. La actividad de la nicotina desmetilasa se calcula con base en la producción de nornicotina marcada con ¹⁴C después de 1 hora de incubación.

Se observó actividad similar a la de p450 en preparaciones microsómicas a partir de células de levadura de control que expresa CYP71D20 y los tres cultivos de levadura con p450 de prueba transformados con los genes D120-AH4, D208-AC8, y D208-AD9. Sin embargo, el control y las tres p450 de prueba no mostraron ninguna formación por conversión en nornicotina sugiriendo que no contienen una enzima endógena o inducida que pueda catalizar la desmetilación de la 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 utilizando muestras microsómicas obtenidas a partir de células de levadura que expresan el gen de la nicotina desmetilasa de tabaco (D121-AA8). Estos resultados indican que la actividad de la nicotina desmetilasa es el resultado del producto genético de D121-AA8. La actividad de la nicotina desmetilasa requería NADPH y demostró ser inhibida por inhibidores específicos de p450, consistentes con la nicotina desmetilasa de tabaco que es una p450. La actividad enzimática para la nicotina desmetilasa de tabaco (D121-AA8) fue aproximadamente 10.8 pKat/mg de proteína calculada por intensidad radioactiva y concentraciones de proteína. Un conjunto típico de resultados de ensayo enzimático obtenidos para las células de levaduras se muestran en la siguiente tabla. (Tabla 2).

50 Tabla 2. Actividad de desmetilasa en microsomas de células de levadura que expresan D121-AA8 y genes P450 de control

Muestras	Microsomas	Microsomas + clorpromazina 1 mM	Microsomas + citocromo C 100 µM	Microsomas - NADPH
D121-AA8	10.8 ± 1.2* pkat/mg proteína	1.4 ± 1.3 pkat/mg proteína	2.4 ± 0.7 pkat/mg proteína	0.4 ± 0.1 pkat/mg proteína

Muestras	Microsomas	Microsomas clorpromazina 1 mM +	Microsomas citocromo C 100 µM +	Microsomas NADPH -
Control (CYP71D20)	No detectado	No detectado	No detectado	No detectado
*n=12, otros n=3				

5 La omisión del NADPH del ensayo que utiliza microsomas derivados de las células de levadura D121-AA8 dio como resultado la abolición de la actividad de nicotina desmetilasa; por lo tanto no se formó normicotina (Tabla 2). Cuando dos inhibidores conocidos de P450, la clorpromazina (CPZ, 1 mM) y citocromo C oxidado (cyt C, 100 µM) fueron agregados a las mezclas del ensayo enzimático separadamente y se incubaron durante 1 hora antes de agregar la solución de detención de metanol, las actividades de nicotina desmetilasa decrecieron significativamente (Tabla 2). Juntos estos experimentos demostraron que la D121-AA8 codifica una proteína de citocromo p450 que cataliza la conversión de nicotina a normicotina cuando se expresan levaduras.

Ejemplo 8

10 Identidad de secuencias de aminoácidos relacionadas de clones de longitud completa

Las secuencias de ácidos nucleicos de genes de nicotina de longitud completa clonados en el Ejemplo 5 fueron deducidas en cuanto a su secuencia completa de aminoácidos. Los genes de citocromo p450 fueron identificados por la presencia de tres motivos de dominio p450 conservados, los cuales correspondían a UXXRXXZ (SEQ ID NO: 2274), PXRFXF (SEQ ID NO: 2275) o GXRXC (SEQ ID NO: 2276) en el terminal carboxil donde U es E o K, X es cualquier aminoácido y Z es P, T, S o M. Todos los genes de p450 fueron caracterizados en cuanto a la identidad del aminoácido utilizando un programa BLAST que compara subsecuencias de longitud completa una con otra y con genes de tabaco conocidos. El programa utilizó la herramienta BLAST especial NCBI (Alineamiento de dos secuencias (b12seq), <http://www.ncbi.nlm.nih.gov/blast/b12seq/b12.html>). Dos secuencias fueron alineadas bajo BLASTN sin filtro para secuencias de ácidos nucleicos y BLASTP para secuencias de aminoácidos. Con base en su porcentaje de identidad en aminoácidos, cada secuencia fue agrupada en grupos de identidad en donde el agrupamiento contenía miembros que compartían al menos 85% de identidad con otro miembro.

Utilizando estos criterios, se identificaron grupos únicos y se representan en la Tabla 3. Se dedujo que la secuencia de aminoácidos del gen de nicotina desmetilasa de longitud completa tenía la secuencia provista en SEQ ID NO: 5 (Figura 4) *

25 Tabla 3 Grupos de identidad de secuencia de aminoácido de genes 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)

30 Los genes de longitud completa fueron agrupados posteriormente con base en la homología de aminoácidos altamente conservada entre el dominio UXXRXXZ p450 (SEQ ID NO: 2274) y el dominio de p450 GXRXC (SEQ ID NO: 2276) cerca al extremo del terminal carboxilo. Los clones individuales fueron alineados con base en la homología de secuencia entre los dominios conservados y colocados en diferentes grupos de identidad. En varios casos, aunque la secuencia de ácidos nucleicos del clon era única, la secuencia de aminoácidos para la región fue idéntica. El agrupamiento final fue similar al basado en el porcentaje de identidad para la secuencia de aminoácidos completa de los clones.

Tabla 4 Grupos de identidad de secuencia en aminoácidos de regiones entre dominios conservados de genes p450 de *Nicotiana*

40 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

Uso de fragmentos de citocromo p450 de *Nicotiana* y clones en regulación alterada de las cualidades del tabaco

El uso de los fragmentos de ácidos nucleicos de p450 de tabaco o genes completos es útil en la identificación y selección de aquellas plantas que tienen fenotipos de tabaco o constituyentes de tabaco alterados y, de manera más importante, metabolitos alterados. Las plantas de tabaco transgénicas son generadas por una variedad de sistemas de transformación que incorporan fragmentos de ácidos nucleicos o genes de longitud completa, seleccionados de los reportados aquí, en orientaciones para subregulación, por ejemplo orientación antisentido, o sobreexpresión, por ejemplo, orientación en sentido y similares. Para la sobreexpresión en genes de longitud completa, es deseable cualquier secuencia de ácidos nucleicos que codifique la totalidad o una parte funcional o una secuencia de aminoácidos de los genes de longitud completa descrita en esta invención. Tales secuencias de ácidos nucleicos de manera deseable son efectivas para incrementar la expresión de una cierta enzima y da como resultado así el efecto fenotípico dentro de la *Nicotiana*. La líneas de *Nicotiana* que son homocigóticas se obtienen a través de una serie de retrocruzamientos y se establecen los cambios fenotípicos incluyendo, pero no limitándose a, análisis de ARN de p450 endógeno, transcritos, péptidos expresados en p450 y concentraciones de metabolitos vegetales utilizando técnicas comúnmente disponibles para una persona de habilidad normal en la técnica. Los cambios exhibidos en las plantas de tabaco proveen información sobre el papel funcional del gen seleccionado de interés o son de utilización como especies de plantas de *Nicotiana* preferidas.

Ejemplo 10Clonación de la nicotina desmetilasa genómica de tabaco de Tabaco Burley convertidor

El ADN genómico fue extraído de una línea de plantas de tabaco Burley convertidoras 4407-33 (una variedad de *Nicotiana tabacum* línea 4407) usando el kit tal como se describió más arriba en los Ejemplos (véase también el procedimiento del fabricante).

Los cebadores fueron diseñados con base en el promotor 5' y en la región 3' UTR clonada en los ejemplos previos. Los cebadores de avance 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 reversos 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 PCR al ADN genómico 4407-33 con 100 ml de mezcla de reacción. Se utilizó la enzima de alta fidelidad Pfx para a amplificación por PCR. El producto de PCR fue visualizado sobre gel de agarosa al 1% después de la electroforesis. Se observó y seccionó del gel una banda con peso molecular de aproximadamente 3.5 kb. La banda resultante fue purificada usando un kit para purificación del geles (Qiagen; con base en el procedimiento del fabricante). Tel ADN purificado fue digerido mediante la enzima Xba I (NEB; usada de acuerdo con las instrucciones del fabricante). El plásmido pBluescript fue digerido mediante Xba I usando el mismo procedimiento. El fragmento fue purificado en gel y ligado al plásmido pBluescript. La mezcla de ligazón fue transformada en célula GM109 competente y sembrada sobre una placa LB que contenía 100 mg/l de ampicilina con selección azul/blanco. Se extrajeron las colonias blancas y se cultivaron en 10 mL de medio líquido que contenía ampicilina. El ADN fue extraído por minipreparación. El ADN plásmido que contenía el inserto fue secuenciado usando un secuenciador CEQ 2000 (Beckman, Fullerton, California) con base en el procedimiento del fabricante. Los cebadores T3 y T7 y otros 8 cebadores internos fueron usados para el secuenciamiento. La secuencia fue ensamblada y analizada, proveyendo así la secuencia genómica (SEQ ID NO: 4; Figura 3). Con base en la secuencia genómica, se determinó que el gen de la nicotina desmetilasa en ambas líneas de tabaco convertidora y no convertidora, no contiene un elemento transponible.

La comparación de la secuencia de SEQ ID NO: 5 con la secuencia de SEQ ID NO: 4 permitió la determinación de un intrón sencillo con la porción de codificación del gen (identificada como la secuencia de 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 intrón sencillo. El primer exón cubre los nucleótidos 2010 a 2949 de SEQ ID NO: 4, los cuales codifican los aminoácidos 1-313 de SEQ ID NO: 3, y el segundo exón cubre los nucleótidos 3947 a 4562 de SEQ ID NO: 4, los cuales codifican los aminoácidos 314-517 de SEQ ID NO: 3. De acuerdo con lo anterior, el intrón cubre los nucleótidos 2950-3946 de SEQ ID NO: 4. En la Figura 5 se provee la secuencia del intrón y es la de SEQ ID NO: 7. El producto de traducción del ADN genómico se provee en la Figura 3 como la secuencia de SEQ ID NO: 3. La secuencia de aminoácidos de la nicotina desmetilasa del tabaco contiene un motivo de anclaje de la membrana del retículo endoplasmático.

Ejemplo 11Clonación de la secuencias de flanqueo de 5' (SEQ ID NO: 8) y 3' UTR (SEQ ID NO: 9) de tabaco convertidorA. Aislamiento de ADN total de tejido de hojas de tabaco convertidor

El ADN genómico fue aislado a partir de las hojas de tabaco convertidor 4407-33. El aislamiento del ADN fue llevado a cabo utilizando un DNeasy Plant Mini Kit de la compañía Qiagen, Inc. (Valencia, Ca) de acuerdo con el protocolo del fabricante (Dneasy' Plant Mini and DNeasy Plant Maxi Handbook, Qiagen January 2004). El procedimiento para

la preparación del ADN incluyó las siguientes etapas: el tejido de hojas de tabaco (aproximadamente 20 mg en peso seco) fue triturado hasta un polvo fino bajo nitrógeno líquido durante 1 minuto. El polvo del tejido fue transferido a un tubo de 1.5 ml. Se agregaron regulador AP1 (400 µl) y 4 µl de solución de reserva de RNasa (100 mg/ml) a un máximo de 100 mg de tejido de hojas triturado y se sometió a vórtex vigorosamente. La mezcla fue incubada durante 5 10 minutos a 65°C y mezclada 2-3 veces durante la incubación por inversión del tubo. El regulador AP2 (130 µl) fue agregado entonces al lisado. La mezcla fue mezclada e incubada durante 5 minutos sobre hielo. El lisado fue aplicado a un QIAshredder Mini Spin Column y centrifugado durante 2 minutos (14,000 rpm). La fracción que fluyó fue transferida a un nuevo tubo sin perturbar la pella de residuos celulares. El regulador AP3/E (1.5 volúmenes) fue agregado entonces al lisado clarificado y mezclado con una pipeta. La mezcla (650 µl) de la etapa precedente incluyendo cualquier precipitado fue aplicada a una DNeasy Mini Spin Column. La mezcla fue centrifugada durante 1 minuto a >8000 x g (>8000 rpm) y se descartó la parte que fluyó. Esto fue repetido con la muestra restante y se descartaron la parte que fluyó y el tubo de recolección. La DNeasy Mini Spin Column fue colocada en un nuevo tubo de recolección de 2 ml. Se agregó entonces regulador AW (500 µl) a la columna de DNeasy y se centrifugó durante 1 minuto (>8000 rpm). Se descartó la parte que fluyó. El tubo de recolección fue reutilizado en la siguiente etapa. Se agregó entonces regulador AW (500 µl) a la columna DNeasy y se centrifugó durante 2 minutos (>14,000 rpm) con el fin de secar la membrana. La columna DNeasy fue transferida a un tubo de 1.5 ml. El regulador AE (100 µl) fue transferido con pipeta sobre la membrana DNeasy. La mezcla fue incubada durante 5 minutos a temperatura ambiente (15-25°C) y luego centrifugada durante 1 minuto (>8000 rpm) para eluir.

La calidad y cantidad del ADN fue estimada corriendo las muestras sobre un gel de agarosa.

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

Se utilizó un método de PCR inversa para clonar 750 nucleótidos de las secuencias de flanco de 5' del gen estructural de SEQ ID NO: 5. Primero, 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 corriente abajo de las secuencias de flanco 5'. Los dos cebadores fueron diseñados con base en este fragmento conocido. El cebador de avance fue localizado corriente abajo del cebador reverso. El cebador reverso fue localizado en la porción 3' del fragmento conocido.

El procedimiento de clonación incluyó las siguientes etapas:

El ADN genómico purificado (5 µg) fue digerido con 20-40 unidades de la enzima de restricción apropiada (EcoRI y SpeI) en una mezcla de reacción de 50 µl. Se ejecutó una electroforesis en gel de agarosa con 1/10 volúmenes de la mezcla de reacción para determinar si el ADN fue digerido completamente. Se llevó a cabo una ligazón directa después de la digestión completa ligando 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 T4 CADN ligasa (NEB) fue ligado durante la noche a 4°C. El PCR sobre la reacción de ligazón fue ejecutado después de que se obtuvo un genoma circular pequeño artificial. La PCR fue llevada a cabo con 10 µl de reacción de ligazón y dos cebadores a partir de fragmentos conocidos en 2 direcciones diferentes en 50 µl de mezcla de reacción. Se aplicó un programa de PCR gradiente con temperaturas de fusión de 45-56°C.

Se llevó a cabo la electroforesis en gel de agarosa para verificar la reacción de PCR. La banda deseada fue cortada del gel y se utilizó un kit de purificación de gel QIAquick de QIAGEN para purificar la banda. Los fragmentos de PCR purificados fueron ligados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Los plásmidos de ADN transformados fueron extraídos por minipreparación utilizando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. El ADN del plásmido que contenía el inserto fue secuenciado utilizando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Aproximadamente 758 nt (nucleótidos 1241-2009 de SEQ ID NO: 4) de la secuencia de flanco de 5' fueron clonados por el método descrito más arriba.

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

Se utilizó el BD GenomeWalker Universal Kit (Clontech laboratories, Inc., Palo Alto, CA) para clonar una secuencia de flanco 5' adicional del gen estructural, D121-AA8 de acuerdo con el manual de usuario del fabricante (BD GenomeWalker Agosto de 2004). El tamaño y la pureza del ADN genómico de tabaco fueron probados corriendo muestras sobre un gel de agarosa al 0.5%. Se fijaron un total de 4 reacciones de extremo romo (DRA I, STU I, ECOR V, PVU II) para la construcción en marcha del genoma de biblioteca 33 de tabaco. Después de la purificación de los ADN digeridos, los ADN genómicos digeridos fueron ligados al adaptador paseante de genoma. Las reacciones PCR primarias fueron aplicadas a los cuatro ADN digeridos utilizando el cebador adaptador AP1 y el cebador específico del gen de D121-AA8 (CTCTATTGACTAGCTGGTTTTGGAC; SEQ ID NO: 2282). Los productos de PCR primarios fueron utilizados directamente como plantillas para el PCR anidado. El cebador anidado adaptador provisto por el kit y el cebador anidado del clon conocido D121-AA8 (SEQ ID NO: 5) (GGAGGGAGAGTATAACTTACGGATTC; SEQ ID NO: 2283) fueron utilizados en la reacción de PCR. Los productos de PCR fueron verificados corriendo electroforesis en gel. Las bandas deseadas fueron seccionadas del gel y los fragmentos PCR fueron purificados utilizando el kit de purificación de gel QIAquick de QIAGEN. Los fragmentos de PCR purificados fueron ligados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las

instrucciones del fabricante. Los plásmidos de ADN transformados fueron extraídos por minipreparación utilizando el kit SV Miniprep (Promega, Madison, WI) y siguiendo las instrucciones del fabricante. El ADN de plásmido que contenía el inserto fue secuenciado utilizando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Aproximadamente otros 853 nt de la secuencia de flanqueo de 5', incluyendo nucleótidos 399-1240 de SEQ ID NO: 4, fueron clonados por el método descrito anteriormente.

Una segunda ronda del genoma paseante fue llevada a cabo de acuerdo con el mismo método con la diferencia de que fueron utilizados los siguientes cebadores GWR1A (5'-AGTAACCGATTGCTCACGTTATCCTC -3') (SEQ ID NO: 2284) y GWR2A (5'-CTCTATTCAACCCACACGTAAGT -3') (SEQ ID NO: 2285). Otra secuencia de flanqueo de aproximadamente 398 nt, incluyendo los nucleótidos 1-398 de SEQ ID NO: 4, fue clonado por este método.

Una búsqueda de elementos reguladores reveló que, además de la caja "TATA", las cajas "CAAT" y las cajas "GAGA", están presentes varios sitios de reconocimiento similares a MYB y elementos de especificidad del órgano en la región promotora de la nicotina desmetilasa del tabaco. Elementos de respuesta de obtención putativos y elementos regulados por nitrógeno, identificados utilizando métodos estándar, también están presentes en la región promotora.

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

Se utilizó el BD GenomeWalker Universal Kit (Clontech laboratories, Inc., PaloAlto, CA) para clonar la secuencia flanqueadora de 3' del gen estructural, D121-AA8 de acuerdo con el manual de usuario del fabricante. El procedimiento de clonación es el mismo que se describió en la sección C precedente de este ejemplo, excepto para los cebadores específicos del gen. El primer cebador fue diseñado de cerca al extremo del gen estructural D121-AA8 (5'-CTA AAC TCT GGT CTG AAC CTG ATA CTT -3') (SEQ ID NO: 2286). El cebador anidado fue diseñado posteriormente corriente abajo del cebador 1 del gen estructural de D121-AA8 (CA TAC GTA AGG TAA ATC CTG TGG AAC) (SEQ ID NO: 2287). Los productos de PCR finales fueron verificados por electroforesis en gel. Las bandas deseadas fueron seccionadas del gel. Los fragmentos de PCR fueron purificados utilizando el kit de purificación de gel QIAquick de QIAGEN. Los fragmentos de PCR purificados fueron ligados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Los plásmidos de ADN transformados fueron extraídos por minipreparación utilizando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. El ADN del plásmido que contenía el inserto fue secuenciado utilizando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Aproximadamente 1617 nucleótidos de la secuencia de flanqueo de 3' adicional (nucleótidos 4731-6347 de SEQ ID NO: 4) fueron clonados por el método descrito más arriba. La secuencia de ácido nucleicos de la región 3'UTR se muestra en la figura 7.

Ejemplo 12

Filtración del género *Nicotiana* para la presencia o ausencia de un gen de nicotina desmetilasa

Cuarenta y tres especies de *Nicotiana*, cuarenta y nueve líneas de *Nicotiana* rústica y aproximadamente seiscientas líneas de *Nicotiana tabacum* fueron sembradas en macetas y las plantas resultantes fueron cultivadas en el invernadero.

Las muestras de hojas fueron tomadas de plantas de seis semanas de edad. Las extracciones de ADN de las hojas fueron llevadas a cabo utilizando el DNeasy Plant Mini Kit (Qiagen, Inc., Valencia, CA) de acuerdo con el protocolo del fabricante.

Los cebadores fueron diseñados con base en las regiones de promotor 5' y 3' UTR descritas aquí. El cebador de avance fue 5'-GGC TCT AGA AAA ATC TCT TAA GTT ACT AGG TTC TAA-3' (SEQ ID NO: 2290) y el cebador reverso 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 flanqueante de 5' a 180 nt de 3' UTR). El ADN genómico extraído de todas las líneas de *Nicotiana* antes mencionadas fue utilizado para el análisis por PCR. Una mezcla de reacción de 100 µl y la enzima de alta fidelidad Pfx fueron utilizadas para la amplificación por PCR. La temperatura de fusión usada fue de 54°C debido a la menor homología entre las especies (esta temperatura es 2°C inferior a la temperatura utilizada para la clonación de la secuencia genómica del tabaco convertidor 4407 como se describió más arriba). El producto de PCR fue visualizado sobre gel de agarosa al 0.8% después de la electroforesis. Una banda sencilla con peso molecular de aproximadamente 3.5 kb estaba presente o ausente en el gel. Las líneas con una banda positiva fueron calificadas como poseedoras del gen objetivo. Para las líneas que carecieron de bandas positivas, se llevaron a cabo cuatro reacciones adicionales de PCR utilizando cuatro conjuntos más de cebadores. Estos conjuntos de cebadores fueron seleccionados de las 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 detención (5'-ACC AAG ATG AAA GAT CTT AGG TTT TAA -3') (SEQ ID NO: 2293),

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

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

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

- 5 Si las cinco reacciones de PCR antes mencionadas no mostraron bandas correctas, la línea fue calificada como carente del gen objetivo. Ejemplos de la cantidad de ADN genómico y productos de PCR del gen de nicotina desmetilasa objetivo se representan en las figuras 8 y 9.

10 El germoplasma identificado como carente del gen de nicotina desmetilasa se utiliza como material fuente para cruzar con tabacos cultivados. Sin embargo, cualquier secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, puede ser utilizada de manera similar. Los métodos de hibridación interespecíficos o intraespecíficos combinados con métodos estándar de cruzamiento, tales como retrocruzamiento o el método del pedigrí, pueden ser utilizados para transferir el gen de nicotina desmetilasa aberrante o ausente o cualquier secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, desde la fuente donante a los tabacos cultivados. Los resultados de los experimentos de filtración para la nicotina desmetilasa se presentan en la Tabla 5 a continuación. Una línea negativa para nicotina desmetilasa puede ser cruzada consigo misma o con otra línea negativa (por ejemplo, *Nicotiana africana* x *Nicotiana africana* o *Nicotiana africana* x *Nicotiana amplexicaulis* o cualquier combinación de cruzamiento adecuada). Las líneas negativas también se cruzan con cualquier otra variedad comercial de tabaco de acuerdo con técnicas de cruzamiento de tabaco estándar conocidas en el arte. Las líneas de tabaco pueden ser cruzadas con cualquier otra planta compatible de acuerdo con procedimientos estándar en el arte.

15

20

Tabla 5: Resultados de ejemplo de la filtración del género *Nicotiana* para el gen de nicotina desmetilasa

Nombre científico o nombre común o (origen)	Número de inventario	Resultados de la filtración
<i>Nicotiana africana</i>	TW6	Negativo
<i>Nicotiana anplexicaulis</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 glauca</i>	TW69	Negativo
<i>Nicotiana ingulba</i>	TW71	Negativo
<i>Nicotiana knighliana</i>	TW73	Negativo
<i>Nicotiana maritima</i>	TW82	Negativo
<i>Nicotiana megalosiphon</i>	TW83	Negativo
<i>Nicotiana miersii</i>	TW85	Negativo

ES 2 534 203 T3

Nombre científico o nombre común o (origen)	Número de inventario	Resultados de la filtración
<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
<i>Nicotiana rosulata</i>	TW112	Negativo
<i>Nicotiana rotundifolia</i>	TW114	Negativo
<i>Nicotiana rustica</i>	TW116	Negativo
<i>Nicotiana setchelli</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 básica	Positivo
KY 14	Semilla básica	Positivo
TN 86	Semilla básica	Positivo
Coker 176	Semilla básica	Positivo
KY21	TC62	Positivo
KY22	TC63	Positivo
KY24	TC64	Positivo
KY26	TC65	Positivo
KY33	TC66	Positivo
KY34	TC67	Positivo
KY35	TC68	Positivo

ES 2 534 203 T3

Nombre científico o nombre común o (origen)	Número de inventario	Resultados de la filtración
KY41A	TC69	Positivo
KY54	TC71	Positivo
KY52	TC70	Positivo
Virginia 528	TC85	Positivo
Virginia B-29	TC86	Positivo
401 Rojo cereza	TC227	Positivo
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 Heirloom	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

ES 2 534 203 T3

Nombre científico o nombre común o (origen)	Número de inventario	Resultados de la filtración
Ostrolist 2747 II	TI 1568	Positivo
Ergo	TI 1349	Positivo
Burley 323	TI 1535	Positivo
Burley ruso	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 hoja parda	TC50	Positivo
KY9	TC54	Positivo
Mamut negro	TC460	Positivo
Orinoco cola de lagarto	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
Adjuctifolia (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 o (origen)	Número de inventario	Resultados de la filtración
(Alemania)	TI 1459	Positivo
(Suiza)	TI 1506	Positivo
Sironer (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 (Antigua Unión Soviética)	TI 86	Positivo

Ejemplo 13

Creación o generación de mutaciones y filtración por variación genética en el gen de nicotina desmetilasa

5 Las variaciones o mutaciones genéticas preexistentes en la secuencia que codifica la nicotina desmetilasa o cualquier otro gen representado por una secuencia de ácidos nucleicos mostrada en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, son filtrados utilizando tecnologías moleculares que incluyen lesiones locales inducidas direccionadas en genomas (LABRADO), métodos de huellas dactilares de ADN tales como polimorfismo de longitud de fragmento amplificada (AFLP) y polimorfismo de nucleótidos individuales (SNP). En la práctica, se usan las poblaciones de plantas que representan variaciones genéticas preexistentes tales como una planta transgénica, (por ejemplo, cualquiera de las descritas aquí) o las creadas por la exposición de tejidos reproductivos, semillas u otros tejidos vegetales a mutágenos químicos tales como agentes alquilantes, sulfonato de etano metilo (EMS) por ejemplo, o a radiación tal como rayos X o rayos gamma. Para poblaciones mutagenizadas la dosificación de los agentes químicos mutagénicos o radiación se determina experimentalmente para cada tipo de tejido vegetal de tal manera que se obtiene una frecuencia de mutación que está por debajo de un nivel de umbral caracterizado por la letalidad o esterilidad reproductiva. El número de semillas de generación MI o el tamaño de las poblaciones de plantas MI resultantes de los tratamientos mutagénicos se estiman con base en la frecuencia esperada de mutaciones. La progenie, la generación M2 de las plantas MI representa la población que de manera deseable es evaluada para una mutación en un gen, por ejemplo, el gen de la nicotina desmetilasa.

20 El labrado, las huellas dactilares de ADN, SNP o técnicas similares pueden ser utilizados para detectar variaciones genéticas inducidas o de origen natural en un gen deseable tal como el gen de la nicotina desmetilasa. La variación puede ser el resultado de eliminaciones, sustituciones, mutaciones puntuales, translocaciones, inversiones, duplicaciones, inserciones o mutaciones nulas completas. Estas tecnologías podrían ser utilizadas en una selección asistida por marcador (programa de cruzamiento MA) para transferir o cruzar los alelos nulos o disímiles del gen de la nicotina desmetilasa o cualquier secuencia de ácidos nucleicos mostradas en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, en otros tabacos. Un experto en cruzamientos podría crear poblaciones segregables a partir de hibridaciones de un genotipo que contiene el alelo nulo o disímil con un genotipo agrónomicamente deseable. Las plantas en las generaciones F2 o de retrocruzamiento podrían ser filtradas utilizando un marcador seleccionado de la secuencia de la nicotina desmetilasa o una secuencia de ácidos nucleicos mostradas en las figuras 1, 3 a 7, y 10-17, o un fragmento de las mismas, utilizando una de las técnicas listadas previamente. Las plantas identificadas por poseer los alelos nulos o disímiles podrían ser retrocruzadas o autopolinizadas para crear la siguiente población que podría ser filtrada. Dependiendo del patrón de herencia esperado o de la tecnología MAS utilizada, puede ser necesario autopolinizar las plantas seleccionadas antes de cada ciclo de retrocruzamiento para ayudar en la identificación de las plantas individuales deseadas. El retrocruzamiento u otro procedimiento de cruzamiento pueden ser repetidos hasta que el fenotipo deseado del progenitor recurrente sea recuperado.

REIVINDICACIONES

1. Un cultivo de tejidos de células de tabaco regenerables, teniendo dichas células de tabaco una mutación en un gen de nicotina desmetilasa endógeno que tiene la secuencia definida en SEQ INO: 4, en donde la mutación es seleccionada del grupo consistente de una mutación puntual, una eliminación, una inserción, una duplicación y una inversión, en donde dicho cultivo de tejidos regenera plantas de tabaco capaces de expresar todas las características fisiológicas y morfológicas de una planta de tabaco que tiene dicha mutación, y en donde dichas plantas de tabaco regeneradas exhiben expresión del gen o del producto genético de nicotina desmetilasa con respecto a una planta de control no mutada, o la nicotina desmetilasa codificada por el gen mutado exhibe actividad enzimática reducida con respecto a una nicotina desmetilasa no mutada.
2. El cultivo de tejidos de células de tabaco regenerables de la reivindicación 1, en donde las células regenerables son embriones, células meristemáticas, semillas, polen, hojas, raíz, puntas de raíz, o flores o son protoplastos o callus derivados de los mismos.
3. Un método para producir un producto de tabaco que comprende:
- (a) proveer tabaco curado a partir de una planta de tabaco que tiene una mutación en un gen de nicotina desmetilasa endógeno que tiene la secuencia fijada en SEQ ID NO: 4, en donde la mutación es seleccionada del grupo consistente de una mutación puntual, una eliminación, una inserción, una duplicación y una inversión, y en donde dicha planta de tabaco exhibe expresión reducida del gen o del producto genético de nicotina desmetilasa mutado con respecto a una planta de control no mutada, o la nicotina desmetilasa codificada por el gen mutado exhibe actividad enzimática reducida con respecto a una nicotina desmetilasa no mutada; y
- (b) preparar un producto de tabaco utilizando dicho tabaco curado; particularmente en donde dicho producto de tabaco es una hoja o tallo curados, un rapé húmedo o seco, un tabaco para mascar, un producto de cigarrillo, un producto de cigarro, un cigarrillo, un tabaco para pipa o bidis.
4. Un método para reducir la expresión o actividad enzimática de una nicotina desmetilasa en una planta de tabaco con respecto a una planta de control, comprendiendo dicho método:
- (a) expresar un transgen que codifica un ácido nucleico antisentido que exhibe expresión de un ácido nucleico endógeno seleccionado del grupo consistente de las secuencias de ácidos nucleicos SEQ ID NOS: 4 o 7 en dicha planta, en donde la planta de control carece de transgen; o
- (b) expresar un transgen que cosuprime un ácido nucleico endógeno seleccionado del grupo consistente de SEQ ID NOS: 4 y 7 en dicha planta, en donde la planta de control carece del transgen; o
- (c) expresar un producto de gen negativo dominante en dicha planta que inhibe la expresión de un ácido nucleico endógeno seleccionado del grupo consistente de las secuencias de ácidos nucleicos de SEQ ID NOS: 4 o 7, en donde dicha expresión reducida ocurre a nivel transcripcional, a nivel de traducción, o a nivel postraducción, en donde la planta de control carece del producto de gen negativo dominante.
5. El cultivo de tejidos, o método de la reivindicación 1, 2 o 3, en donde dicha mutación es una mutación puntual.
6. El cultivo de tejidos, o un método de la reivindicación 1, 2 o 3, en donde dicha mutación es una eliminación.
7. El cultivo de tejido, o método de la reivindicación 1, 2 o 3, en donde dicho tabaco es *Nicotiana tabacum*, *Nicotiana sylvestris*, *Nicotiana tomentosiformis* o *Nicotiana glauca*.
8. El cultivo de tejidos, o método de la reivindicación 1, 2, 3 o 4, en donde la expresión del gen de nicotina desmetilasa endógeno mutante que tiene la secuencia fijada en SEQ ID NO: 4 es reducida en al menos 30% en la planta de tabaco mutante con relación al nivel de expresión en la planta de control no mutante.
9. El cultivo de tejidos, o método de la reivindicación 1, 2, 3 o 4, en donde la actividad enzimática de la nicotina desmetilasa mutante codificada por el gen de nicotina desmetilasa endógeno que tiene la secuencia fijada en SEQ ID NO: 4 es disminuido en al menos 10% en la planta de tabaco mutante con respecto a la nicotina desmetilasa no mutada.

SEQ ID NO.: 1 (D35-BG11)

atgacttatg cattgcaagt ggaacactta acaatggcac attgatcca aggttcaat
tacagaactic caaatgacga gcccttggat atgaaggaag gtgcaggcat aactatcgt
aaggtaaac cigtggaact gataatagcg cctgcctgg 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 del Genoma de 33-L
(Gen Genómico de D121-AA8)

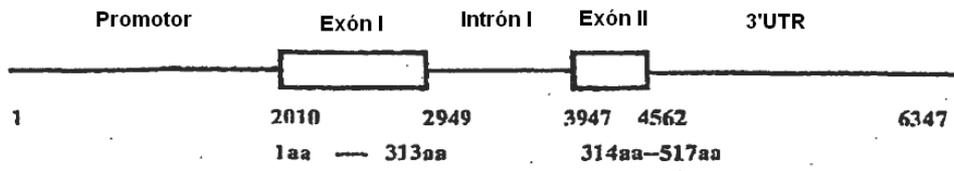


FIGURA 2

SEQ ID NO:3

CARACTERÍSTICAS	Localización/ Calificador
Total	1..6347
SECUENCIAS DE FLANQUEO DE 5'	1..2009
CDS	unión (2010..2948,3947..4561)
Intrón	2949..3946
Producto	Citocromo P450
3' UTR	4562..6347
Traducción	

```

MLSPIEAIVGLVTFITLFFFLWTKKSQKPSKPLPPKIPGGWPVIGELFHFNDGDDRPLARKLGLDLADKYGF
TFRLGLPLVLVSSYEAVKDCFSINDAIFSNRPAFLYGDYLGYNNAFLANYGPYWRKNRKLVIQEVLSAS:
EKFKHVRFARIQASIKNLYTRIDGNSSTINLTDWLEELNFGGLIVKMIAGKNYESGGKGDQVERFKKAFKDFM:
SMFEVLWDAFPIPLFKWVDFQGHVKAMKRTFKDIDSVPQNWLEEHINKREKMEVNAEGNEQDFIDVVLSKMS:
YLGEQYSRDTVIKATVFSVLDAADTVLHINWGMALLINNQKALTKAQEEIDTKVQKDRWVEESDIKDLVY:
AIVKEVLRLLYFPGPLLVPHENVEDCVVSGYHIIKGRTRLFANVMKLRDPKLVSDPDTTFPERFIATDIDPRG:
YKYIPFGSGRRSCPGMTYALQVEHLTMAHLIQGFNYRTPNDEPLDMKEGAGITTRKVNPFVELIAPRIAPL:
    
```

SEQ ID NO:4

RECuento DE BASES 2046 a 984 c 1163 g 2154 t

```

1 TCTCTAAAGT CCCCTCCAC TTTATCTTAG CTGTGTGATT TCTTTCAGAC AACCTTATT
61 TTATTCAGAC TCTTATTTGT ATTATCTAG AAGCTCGTGT ACTTGTGACA CCAGTCTGG
121 GATGGTATT AGATATCGCT ATTATTTGG CTTATTCACT TCAGTTCAGA TTTTATCCA
181 GTTATTGAT TCTTTATTA TTAATCAAAT TGAATGTGA AAAATGGTTA AAATTACTCC
241 AATGTTGGCT TTCCTAGTAA GCGAAATATC AGGCGCCATC ACGGTACCCG AAGGTGAGAA
301 TTTCAGATCG TGACAGCCGC ATCTCAGGGG GTGTGATGTA AACAGTTTAC GATGGTGCA
361 GCATTAGTGG CTGCTTCGAC GACTTAAATC CGTAACCTAT AGATCACACG AATACAACCT
421 TACTATTTTA ACACCCAGCA AATTCCTGAT AAAACAAT CTACATAGC ACATCAAAAT
481 GTAAATGAT GAAGAAAAG ATGACTTTTA TAGACAGAGA AAAAAACAGT TACGTGTGGG
541 GTTGAATAGA GATTGTGGCT ATGCTATTC TAATATTGAA ATTCACCGAC TTTTTTAGTT
601 CAATACGAAA AGAGTAAAGT AAAAGGCTG AAAAGGAAA GGACAATGCC TAAAGGACA
661 CATTCAGAAC ACATACACTG AATGATCTTA ATTTCTAGT CGAAGATTC TAGTTCGAGG
721 ATAACGTGAG CAATCGGTTA CTCCCAATTA GCAATGCCA ACTGGATGTT TGACTATTTA
781 TGTTCTGGC CAATAGAGGA GAGGAATACT ACGTTACGTA TGGAGTTTGA ACCCTTACA
841 TCAACTTATT AAGTGAGTTA TCCCTCAATC ACGATTCAAC TATGATCCG CTAACCTCAA
901 AGAATATTGA GTTAATATT CAAATGATTA GTCCAAAATT ATTTAATAAA GTTATACTAT
961 TCTTTTATT TGTAAACATC ATCTTTTGT TACATATTIA GTTAAATCTT AGCCCAACCC
1021 TATCGTTGGT TCGACTTTT TTCTTTAAT TTTGATTTAT TCTGTTCGGT AATTCGGCTT
1081 TGTTTGGCTT GAATAATAAC TAGTGATAA AGTCATATAC TCTAATATT TTAATTGAAG
1141 TACTCACAAA TACAAAATA AAAACATTC TAAGCTCACA TGATAGTTGA CAAATCTTT
1201 ATCCAAAACA AGAGGCGGAG CCAGGATTTG AACTTATGG GTTCAGAAAT CTAATCTCT
1261 TAAGTTACTA GGTCTAAAT TAATAATTTA TACATGTTCA ATGAATTTCT TAAGCAAAT
1321 ACATAGTTG AACGAAAGCT ACTGGGTTCC GCCGAATCCG TAAGTATAC TCTCCCTCCG
1381 CCCCGGTCCA AAACAGCTA GTATCAATAG AGAGAGAGAG AGAGAGAGAG AGATAATAA
1441 TTTGACCATT GACAATGGCT TATTACTTGC TTAGAGTTAA TTGGTGAAC TAGAGAATAT
1501 AATAAGGAAT ATTTAAACAG ATACGTCATC AATCCAGGAG TAACGAAATA AGAAATACCC
1561 TAAAATCGTA GAAACATTAC GTTAAATTGC TTGACAGCCT ATCTAGTAAG AGTCAAAATC
1621 TACTATCTAT CTGTTCGGC CATTTCCTTA AAGAAGTACA TGAGCTTTAT CATCCACCTC
1681 AACATGAATG CAAAAGAAA TTATTGTGCA ACTTAATATG TTATAATCAA TGATATGTT
1741 CTGTGTAAAC AAAGTATATA TTTGATACG ATATTAATAT GTAGGTGTTA TATTTTAAA
1801 TATCAAAAT CATACTAAC ACCGATTTTT TAAAACCTTA GGCCAATTAC CCTACCAACT
1861 AAAATACTGT ATATCAACA CTAATGTTT CTATTTCGGT ACGACAGTTC TCTATTTACC
1921 ATATTATGGA ATTATGCCCA TCCTACAGTT ACCTATAAAA AGGAAGTTGC CGATAGTTAT
    
```

FIGURA 3
HOJA 1 DE 3

SEQ ID NO: 4 CONTINUACIÓN

```

1981 AITCTCAACT TCTTATCTAA AAATCCATAA TGCTTTCTCC CATAGAAGCC ATTGTAGGAC
2041 TAGTAACCTT CACATTTCTC TTCTTCTTCC TATGGACAAA AAAATCTCAA AAACCTTCAA
2101 AAQCCTTACC ACCGAAAATC CCCGGAGGAT GGCCCGTAAT CGGCCATCTT TTCCACTTCA
2161 ATGACGACGG CGACGACCGT CCATTAGCTC GAAAACCTCG AGACTTAGCT GACAAATACG
2221 GCCCCGTTTT CACTTTTCGG CTAGGCCTTC CCCTTGTCTT AGTTGTAGC AGTTACGAG
2281 CTGTAAGA CTGTTTCTCT ACAAATGACG CCATTTTTC CAATCGTCCA GCTTTTCTTT
2341 ACGCGGATTA CCTTGGCTAC AATAATGCCA TGCTATTTTT GGCCAAITAC GGACCTTACT
2401 GCGGAAAAA TCGAAAATTA GTTATTGAGG AAGTCTCTC CGTAGTCTGT CTCGAAAAAT
2461 TCAAACACGT GAGATTTGCA AGAATTCAGG CGAGCATTAA GAATTTATAT ACTCGAATTG
2521 ATGGAAATTC GAGTACGATA AATTTAACTG ATTGGTTAGA AGAATTGAAT TTTGGTCTGA
2581 TCGTGAAGAT GATCGCTGGA AAAAATTATG AATCCGGTAA AGGAGATGAA CAAGTGGAGA
2641 GATTTAAGAA AGCGTTTAA GATTTTATGA TTTTATCAAT GGAGTTTGTG TTATGGGATG
2701 CATTTCCAAT TCCATTATTT AAATGGGTGG ATTTTCAAGG GCATGTTAAG GCTATGAAA
2761 GGACTTTTAA AGATATAGAT TCTGTTTTTC AGAATTGGTT AGAGGAACAT ATTAATAAAA
2821 GAGAAAAAAT GGAGGTTAAT GCAGAAGGGA ATGAAACAGA TTTCATTGAT GTGGTGCCTT
2881 CAAAAATGAG TAATGAATAT CTGGTGAAG GTTACTCTCG TGATACTGTC ATTAAGCAA
2941 CGGTGTTTGT AAGTTCACCT GTCATTTTC ATTTATTAC TTTTATTTT AGGAGCAGAC
3001 ATGTTAATAA TAATTTGGAG CAACTGTAAA GTTATCTATG TGTACAGGTT CGAGCCTCAG
3061 GTGCAACCAC TAATGCTTGT ATTAGATTAT GTTGTCTGCA TCATACCCTT AATTGGAGTG
3121 TGGCTCTTCC CGAACCCTGC AATGCTGGAT GCTGGATGCT TTATGTATCA GACTGACCTT
3181 TTTGTTAAAC TATCTAAATA CTAAGGATGA TTTAATAAAA ATATAGAAAT GTAACAGAA
3241 AAAGATGAGA TTATTTTGG GGCTATATGG ATTCGCCCCG GCTTTGGGAG GTAAAACGGT
3301 ATCTACCAGT TGAGACTTTA CTCCAGAACT TTATCTCGAG AGCTCTGAAAT AAAAATGAAA
3361 TAGTATTAC CACTCCAAA TCTTTGATGG TAAAAAGATG AGATATAACC TCTTATAATT
3421 GATTGAACCA CGTGTATAGA ATAAAACTTC TTTACTCCCA TTCAGCATAA GAAAATAGAA
3481 ACCAAACGGA ATTCCTCTCT TTTTATGGGG GAAATTCCTT AATGCTGTG TGAATATAGA
3541 TFCATGTCGT TATCTATTT TTAATATGTA TGAAAATCAA TATAGTCAA GTTAATACTT
3601 ATGTCATTTG GTTTGCGGAC AAGTTATATT GGAACATAT ATACGCTCA TTATAGATA
3661 GTGATTATTT AGAGGATATA CATTTTTTTT GGATAAATAT TTGATTTATT GGATTAATA
3721 TAGAATATAC AGGTAAGGTC TAAAACTGTG GTTTGCTTTT AACTAATAA AACTTGACCT
3781 CGTACAATTC TAAAGAAATA TTTGAARTAA ATGAATTAT TTATGTATA TCAATTAATA
3841 AAATCATAGT ATAGATGAGA TGTGTGCATA CTGACAATA ACTATACTAA CTAACAACAG
3901 GTATGTGAAT AATTGATATT CTTTTTTTAA TTATCTTTT TTCCAGAGTT TGGTCTTGG
3961 TGCAGCAGAC ACAGTTGCTC TTCACATAAA TTGGGGAATG GCATTTATGA TAAACAATCA
4021 AAAGGCTCTG ACCGAAAGCAC AAGAAGAGAT AGACACAAA GTTGGTAAGG ACAGATGGGT
4081 AGAAGAGAGT GATATTAAGG ATTTGGTATA CCTCCAAGCT ATTGTTAAAG AAGTGTACG
4141 ATTATATCCA CCAGGACCTT TGTTAGTACC ACACGAAAT GTAGAGAIT GTGTTGTTAG
4201 TGGATATCAC ATTCTAAAG GGACAAGATT ATTCGCAAC GTCATGAAAC TGCAACGTTA
4261 TCCTAAACTC TGGTCTGATC CTGATACTTT CGATCCAGAG AGATTCATTG CTACTGATAT
4321 TGACTTTCGT GGTGACTACT ATAAGTATAT CCCGTTTGGT TCTGGAGAC GATCTGTCC
4381 AGGGATGACT TATGCATTGC AAGTGGAAAC CTAACAATG GCACATTTGA TCCAAGGTTT
4441 CAATTACAGA ACTCCAATG ACGAGCCCTT GGATATGAAG GAAGGTGCG GCATAACTAT
4501 ACGTAAGGTA AATCCTGTGG AACTGATAAT AGCGCCTCGC CTGGCACCTG AGCTTTATTA
4561 AAACCTAAGA TCTTTTATCT TGGTTGATCA TTGTATAATA CTCTAAATG GATATTCATT
4621 TACCTTTTAT CAATTAATG TCAGTACGAG TTTTCTAAT TTGGTACATT TGAATAATA
4681 AGTAAAGAA ATTTGTGCTA ATATATAAAG GTTTGTAGAA GATAATTGAC TGGTTGTACC
4741 ACAATCTCCA GTGAAAGTGT TAATTAITTA CTTGATCCAC AGCTTATTCT ATGTTTGAAA
4801 TTTGCTTAGT GTCATGATAT TACTCCATCA AATTCAGAA ATAATCATTT CCAACTTTTG
4861 CTGGACTGGA CGATCTTTCA ATAATAAAGG ATCTTTAATT TGCCAAAGTT GAGATCAAAA
4921 TACTGGTTCG TTTACCAATA AGAATGAAAT GTGATGGAAA TTATGTACGT TGGGATAAGG
4981 GAACCAACT ATCAGGAGA CTAAGAAGTA CGTAAAGGAA ARGAAAAAT TTGCCATTGA
5041 TTGCTACTAA GTAACCTAAC AAAATCTTTC AGAAAAGAA CACTTGTATA AGTCCGGGTT
5101 GAAAGTTTTG GTGTCTCTTT TCTTATGTAT TGTGTCTTT AGACAGTAT GTACTTAGTT
5161 ATTCAGAA TTTATTTTCG TATTAGAGCT CAAGACTCTG TATTTATTAG TTCTGGGAGA
5221 ATTATCATGT ATTTTCAGTC TTTTGTATT TCTGTAAAT CCGCTATTTT GGTTCFTTAT
5281 TGCTATGTC GGCTTCTTA GAAAAATGTT TAGGCGCTAT CAGACTGAT TGAGATTTTG
5341 TATCGTGACT GATAATTACA CGGTTTAGTA AATTTTGATA TTTTCAAAA GAGTTTTTTT

```

FIGURA 3
HOJA 2 DE 3

SEQ ID NO: 4 CONTINUACIÓN

```

5401 AATAAATAT GCAACTTCAG TCAAAACATA CAACGTTTTG TTGTATAAAT CCGATCAAAA
5461 CATATAACTT ACATAAAACT TGCATATGAA TTTTTTGTG TATAAACATT TGGTTAAAAC
5521 ATATAACTTA CATATAATTT GCATACAACT TACATAAAAC TTGCATATAA ACAATTTATT
5581 TATGTCTTTG TTTTGGAGTA TCAATTTGAA ATTCCAACAA AAACAAACTC TAATTTTAC
5641 CAAACTCTCT CAAAATTGAG TTATAGATTT CAAAAGATAT CCITAATCGT TTGCAATTGT
5701 AACAAATCCGA TCGGCCGTTT TGAGATCTAG CGTGTGTGTT GCGCGTTTGA GACCTTGAGT
5761 AACTTCACTT TATGTTGTAT GACTTGTATA TGTGGTCGGA ATTAAATTC GGGAGTTCA
5821 GAGTTGATTC GGATGAAAAA TTCTAATTC GGAAGTTTTA AGATGGAATG ATTGACTAAG
5881 GATTGACGTT TGAGTAAACG ATCTCGGAAT CGAGATTGA AGGTCCCAAT AGGTTCTAT
5941 GATGATTTCA GACTTGAGCG TATGTTTGGG TTGAGTATCG GGTGGTCCCG GAGCATTTC
6001 ACGCTGATTA TAGAAAATG GCATCTTAAA GGTTTTAGAA TTTTATAAGT TTGGTTTGAA
6061 GTGGATTTTG ATATTATCGG TGTCCATTTG GAGTTTCGAG CCTTGAATA GGTTCGTATC
6121 GTAAATTTTG ACTTTAGTGT AAAGTTCCGG GTCATCCGG AGTGTTTTGA TAAGATTCTG
6181 ATGCGTTCGT CGAAGTTTGG AAGTTTGAAA GTTGAAAAGA AGATTTTAA TAGGCGATTC
6241 ATGATTTTGA TGTATTGTGT GTCGAGCCTT TGGATAAGTT TGTGTGAGGT ATGGGACTTG
6301 TTGGTATGAA TGGACGAGCT CTACGGGGC CTCGAGTAAG TTTCCGA
    
```

FIGURA 3
HOJA 3 DE 3

SEQ ID NO:5

Secuencias de Codificación del Gen de Nicotina Desmetilasa de Tabaco

RECUENTO DE BASES 489 a 275 c 333 g 457 t

```

1 ATGCTTTCTC CCATAGAAGC CATTGTAGGA CTAGTAACCT TCACATTTCT CTTCTTCTTC
61 CTATGGACAA AAAAATCTCA AAAACCTTCA AAACCCTTAC CACCGAAAAT CCCCAGGAGGA
121 TGGCCGGTAA TCGGCCATCT TTTCCACTTC AATGACGACG GCGACGACOG TCCATTAGCT
181 CGAAAACCTCG GAGACTTAGC TGACAAATAC GGCCCCGTTT TCACTTTTTCG GCTAGGCCTT
241 CCCCTTGTCT TAGTTGTAAG CAGTTACGAA GCTGTAAGAAG ACTGTTTCTC TACAAATGAC
301 GCCATTTTTT CCAATCGTCC AGCITTTCTT TAGCGCGATT ACCTTGGCTA CAATAATGCC
361 ATGCTATTTT TGGCCAATTA CGGACCTTAC TGGCGAAAAA ATCGAAAATT AGTTATTTCAG
421 GAAGTTCTCT CCGCTAGTCC TCTCGAAAAA TTCAAACACG TGAGATTTCG AAGAATTCAA
481 GCGAGCATT AAGAAATTTATA TACTCGAATT GATGGAAATT CGAGTACGAT AAATTTAACT
541 GATTGGTTAG AAGAATGAA 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 TGTGGTCTT TCAAAAATGA GTAATGAATA TCTTGGTGAA
901 GGTACTCTC GTGATACTGT CATTAAAGCA ACGGTGTTTA GTTTGGTCTT CGATCGACGA
961 GACACAGTTG CTCTTACAT AAATGGGGA ATGCCATTAT TGATAAACA TCAAAAGGCC
1021 TTGACGAAG CAAGAAGA GATAGACACA AAGTTGGTA AGGACAGATG GGTAGAAGAG
1081 AGTGATATTA AGGATTGGT ATACCTCCAA GCTATTGTTA AAGAAGTGTT ACGATTATAT
1141 CCACCAGGAC CTTTGTAGT ACCACACGAA AATGTAGAAG ATTGTGTGT TAGTGGATAT
1201 CACATTCCTA AAGGGACAAG ATTATTCCCA AACGTCATGA AACTGCACG TGATCCATAA
1261 CTCTGGTCTG ATCCTGATAC TTTGATCCA GAGAGATTCA TTGCTACGA TATTGACTTT
1321 CGTGGTCAGT ACTATAAGTA TATCCGTTT GGTCTCGAA GACGATCTTG TCCAGGGATG
1381 ACTTATGCAT TGCAAGTGG AACTTAACA ATGGCACATT TGATCCAAGG TTTCAATTAC
1441 AGPACTCCAA ATGACGAGCC CTTGGATATG AAGGAAGGTG CAGGCATAAC TATACGTAAG
1501 GTAAATCCTG TGGAACTGAT AATAGCCCT CGCTGGCAC CTGAGCTTTA TTAA
    
```

SEQ ID NO:6

Aminoácidos Deducidos del Gen de Nicotina Desmetilasa de Tabaco

```

1 MLSPIEAIVG LVTFTLFFF LWTKKSQKPS KPLPPKIPGG WFIGHLFHP NDDGDDRPLA
61 RKLGLADKY GPVFTFRGL PLVLVSSYE AVKDCFSIND AIFSNRPAFL YGDYLGYNNA
121 HLFLANYGPF WRKNRKLVIQ EVLSASRLEK FKHVRFARIQ ASIKNLYTRI DGNSSSTINLT
181 DWLEELNPGI IVKMIAGRWY ESGKGDEQVE RFRKAFKDFM ILSMEFVLWD APPIPLFKWV
241 DPOGHVKAMK RTFKDIDSVF QNWLEEHINK REKMEVNAEG NEQDFIDVVL SKMSNEYLGE
301 GYSRDTVIKA TVPSLVLDAA DTVALHINWG MALLINNQKA LTKAQEIDT KVGKDRWVEE

361 SDIKDLVYLQ AIVKEVLRLY PPGPLLVPEH NVEDCVVSGY HIPKGTRLF A NVMKLRDPK
421 LWSDPDTFDP ERPIATDIDP RGQYKYIPF GSGRRSCPQM TYALQVEHLT MAHLIQGFY
481 RTPNDEPLDM KEGAGITIRK VNPVELIAP RLAPELY
    
```

FIGURA 4

SEQ ID. NO: 7

Intrón del gen de Nicotina Desmetilasa de Tabaco

(998 nt)

```

1  GTAAGTTCAT  CTGTCATTTT  TCATTTATTC  ACTTTTATTT  TGAGGAGCAG  ACATGTTAAT
61  AATAAATTTGG  AGCRAACTGTA  AAGTTATCTA  TGTGTACAGG  TTCGAGCCTC  AGGTGCAACC
121  ACTAATGCTT  GTATTAGATT  ATGTTGTCTG  CATCATACCC  CTAATTGGAG  TGTGGCTCCT
181  CCCGAACCCT  GCAATGCTGG  ATGCTGGATG  CTTTATGTAT  CAGACTGACC  TTTTGTAA
241  ACTATCTAAA  TACTAAGGAT  GATTTAATAA  AAATATAGAA  TGGTAAACAG  AAAAAGATGA
301  GATTATTTTT  GGGGCTATAT  GGATTCGCCC  GGGCTTTGGG  AGGTAAAACG  GTATCTACCA
361  GTGAGACTT  TACTCCAGAA  CTTTATCTCG  AGAGCTCTGA  ATAAAAATGA  AATAGTATTT
421  ACCACTCCAA  AATCTTTGAT  GGTA AAAAGA  TGAGATATAA  CCTCTTATAA  TTGATTGAAC
481  CACGTTGATA  GAATAAACT  TCTTTACTCC  CATTGAGCAT  AAGAAAAATG  AAACCAACG
541  GAATTCCTCT  CTTTTTTAGG  GGGAAATTC  TTAATTGCTT  GTTGAATATA  GATTCATGTC
601  GTTATTCTAT  TTTTAATAAT  GATGAAAATC  AATATAGTCA  AAGTTAATAC  TTATGTCATT
661  TGGTTTGCGG  ACAAGTTATA  TTGGAACAT  ATAATACGTC  TATTATAGAA  TAGTGATTAT
721  TTAGAGGATA  TACATTTTT  TTGGATAAAT  ATTGATTTA  TTGGATTA  AATAGAATAT
781  ACAGGTAAGG  TCTAAAACGT  GTGTTTGCCT  TTACACTRAA  TAAACTTGAC  CTCGTACAAT
841  TCTAAGAAA  TATTGAAAT  AAATGAATTA  TTTTATTGTT  AATCAATTAA  AAAAATCATA
901  GTATAGATGA  GATGTGTGCA  TACTTGACAA  TAACTATACT  AACTAAAACA  AGGTATGTGA
961  ATAATTGATA  TTCCTTTTT  AATTATTCTT  TTTCCAG
    
```

FIGURA 5

SEQ ID NO:8

Promotor del gen de Nicotina Desmetilasa de Tabaco (2009 nt)

```

1 TCTCTAAAGT CCCCTCCAC TTTATCTTAG CTGTGTGATT TCTTTCAGAC AACCTTATTT
61 TTATTCAGAC TCATTATTGT ATTATICTAG AAGCTCGTGT ACTTGTGACA CCAGTTCGG
121 GATGGTATTT AGATATCGCT ATTATTTTGG CTTATTCACT TCAGTTCAGA TTTTATCCA
181 GTTATTTGAT TTCTTTATTA TTAATCAAAAT TGAATTGTTA AAAATGGTTA AAATTACTCC
241 AATGTTGGCT TTCTAGTAA GCGAAATATC AGGCGCCATC ACGGTACCCG AAGGTGAGAA
301 TTTCAGATCG TGACAGCCGC ATCTCAAGGG GTGTGATGTA AACAGTTTAC GATGGTGCAA
361 GCATTAGTGG CTGCTTCGAC GACTTAAATC CGTAACTTAT AGATCACACG AATACAACCTT
421 TACTATTTTA ACACCCAGCA AATTCCTGAT AAAAACAATT CTAACATAGC ACATCAAAAT
481 GTAATGATT GAAGAAAAG ATGACTTTTA TAGACAGAGA AAAAAACAGT TACGTGTGGG
541 GTTGAATAGA GATTGTGGCT ATGCTATTTT TAATATTGAA ATTCACCGAC TTTTTAGTT
601 CAATACGAAA AGAGTAAGTG AAAAGGTCTG AAAAGGAAA GGACAATGCC TAAAAGGACA
661 CATTCAGAAC ACATACACTG AATGATTTCTA ATTTCTAGTC CGAAGATTTT TAGTTCGAGG
721 ATAACGTGAG CAATCGGTTA CTTCCCATTA GCAATTGCCA ACTGGATGTT TGACTATTTA
781 TGTTCCTGGC CAATAGAGGA GAGGAATACT ACGTTACGTA TGGAGTTTGA ACCCTTCACA
841 TCAACTTATT AAGTGAOTTA TCCCTCAATC ACGATTCAAC TATGATTCGG CTAACCTCAA
901 AGAATATTGA GTTAATTATT CAAATGATTA GTCCAAAATT ATTTAATAAA GTTATACTAT
961 TTCTTTTATT TGTAAACATC ATCTTTTGT TACATATTTA GTTAAATCTT AGCCCAACCC
1021 TATCGTTGGT TCGACTTTTT TTCTTTTAAAT TTGATTTAT TCTGTTCCGT AATTCGCTT
1081 TGTTCGGCTT GAATAATAAC TAGTGCATAA AGTCATATAC TCTAATATTT TTAATTGAAG
1141 TACTCACAAA TACAAAATA AAAAACATTC TAAGCTCACA TGATAGTTGA CAAAATCTTT
1201 ATCCAAAACA AGAGGCGGAG CCAGGATTTG AAACCTATGG GTTCAGAAAT CTAATCTCT
1261 TAAGTTACTA GGTCTAAAT TAATAATTTA TACATGTTCA ATGAATTTCT TAAGACAAAT
1321 ACATAGTTTG AACGAAAGCT ACTGGGTTCC GCCGAATCCG TAAGTTATAC TCTCCCTCCG
1381 CCCCAGTCCA AAACCAGCTA GTATCAATAG AGAGAGAGAG AGAGAGAGAG AGATAATAAA
1441 TTTGACCATT GACAAATGGCT TATTACTTGC TTAGAGTTAA TTGGTGAAC TTAGAATAT
1501 AATAAGGAAT ATTTAAACAG ATACGTCATC AATCCACGAG TAACGAAGTA AGAAATACCC
1561 TAAAATCGTA GAAACATTAC GTTAAATTGC TTGACAGCCT ATCTAGTAAG AGTCAAAATC
1621 TACTATCTAT CTGTTCGCGC CATTITCTTA AAGAAGTACA TGAGCTTTAT CATCCACCTC
1681 AACATGAATG CAAAAGAAA TTATTGTGCA ACTTAATATG TTATAATCAA TGATATGTGT
1741 CTTGTGTAAC AAAGTATATA TTTGATACG ATATTAATAT GTAGGTGTTA TATTTTAAA
1801 TATCAAATAT CATACTTAAC ACCGATTTTT TAAAACTTA GGCCAATTAC CCTACCAACT
1861 AAAATACTGT ATATCAACA CTAATGTTT CTATTTCCGT ACGACAGTTC TCTATTTACC
1921 ATATTATGGA ATTATGCCA TCCTACAGTT ACCTATAAAA AGGAAGTTGC CGATAGTTAT
1981 ATTCCAACT TCTTATCTAA AAATCCATA
    
```

FIGURA 6

SEQ ID NO: 9

3' UTR del Gen de Nicotina Desmetilasa de Tabaco

```

1 AACCTAAGAT CTTTCATCTT GGTGATCAT TGTATAATAC TCCTAAATGG ATATTCATT
61 ACCTTTTATC AATTAAITGT CAGTACGAGT TTTTCTAATT TGGTACATTT GTAATAATAA
121 GTAAAGAATA ATTGTGCTAA TATATAAAGG TTTGTAGAAG ATAATTGACT GGTGTACCA
181 CAATCTCCAG TGAAGTGTT AATTATTTAC TTGATCCACA GCTTATTCTA TGTTTGAAAT
241 TTGCCTAGTG TCATGATATT ACTCCATCAA ATCCAAGAAA TAATCATTTT CAACITTTTG
301 TGGACTGGAC GATCTTTCAA TAATAAAGGA TCTTTAATTT GCCAAAGTTG AGATCAAAAT
361 ACTGCTGCCT TTACCAATAA GAATGAAATG TGATGGAAT TATGTACGTT GGGATAAGGG
421 AACCAACTA TCAAGGAGAC TAAAGGTAC GTAAGGAAA AGAAAAAATT TGCCATTGAT
481 TGCTACTAAG TAACCTAACA AAATCTTTCA GAAAAGAATC ACTTGTATAA GTCCGGGTTG
541 AAAGTTTTGG TGCTCTTTT CTTATGTATT GTTGTCTTTA GACAGTATTG TACTTAGTTA
601 TTTCAGAATT TTATTTTCGT ATTAGAGCTC AAGACTCTGT ATTTATTAGT TCTGGGAGAA
661 TTATCATGTA TTTTCAGTCT TTTGTTAATT CTGTAAATTC CGCTATTTTG GTTCTTTATT
721 GCTATGTTCC GCTTTCCTAG AAAATGTGTT AGGCGCTATC ACGACTGATT GAGATTTTGT
781 ATCGTCACTG ATAATTACAC GGTTTAGTAA ATTTTGATAT TTTCAAAAAG AGTTTTTTTA
841 ATAAATATG CAACTTCAGT CAAACATAC AACGTTTTGT TGTATAAATC CGATCAAAAC
901 ATATAACTTA CATAAACTT GCGTATGAAT TTTTGTGTTG ATAAACATTT GGTAAAAACA
961 TATAACTTAC ATATAATTTG CATACAACCT ACATAAACTT TGCATATAAA CAATTTATTT
1021 ATGTCITTTGT TTTGAGTAT CAATTTGAAA TTCCAACAAA AACAACTCT AATTTTACC
1081 AAACTCTCTC AAAATTGAGT TATAGATTTC AAAAGATATC CTTAATCGTT TGCAATTGTA
1141 ACAATCCGAT CGGCCGTTTT GAGATCTAGC GTGTGTGTTG GCGGTTTGAG ACCTTGAGTA
1201 ACTTCACTTT ATGTTGATG ACTTGTATAT GTGCTCGGAA TTAATTTTCG GGAAGTTCAG
1261 AGTTGATTCC GATGAAAAAT TCTAATTTCC GAAGTTTAA GATGGAATGA TTGACTAAGG
1321 ATTGACGTTT GAGTAAACGA TCTCGGAATC GAGATTGAA GGTCCAATA GGTTCGTATG
1381 ATGATTTCCG ACTTGAGCGT ATGTTTGGGT TGAGTATCCG GTGGTCCGGG AGCATTTCAA
1441 CGCTGATTAT AGAAAAATGG CATCTAAAG GTTTTGAAT TTCATAAGT TGSTTTGAAG
1501 TGGATTTTGA TATTATCGGT GTCCATTTGG AGTTTCGAGC CTGGAATAG GTTCGTATCG
1561 TAAATTTTGA CTTTATGTA AAGTTCGGCG TCATTCCGGA GTGTTTGTAT AAGATTCTGA
1621 TGCGTTCGTC GAAGTTTGA AGTTTGAAG TTGAAAAGAA GATTTTAAAT AGGCGATTCA
1681 TGATTTTGTAT GTTATTTGTG TCGAGCCTTT GGATAAGTTT GTGTGAGGTA TGGGACTTGT
1741 TGSTATGAAT GGACGAGCTC TACGGGGGCC TCGAGTAAGT TTCGGA
    
```

FIGURA 7

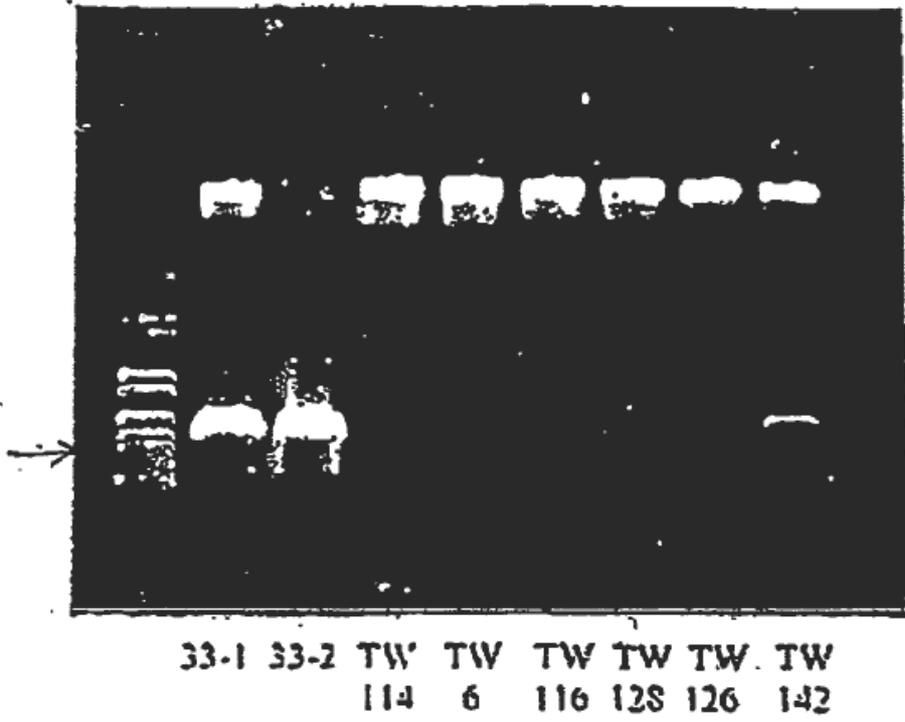


FIGURA 8

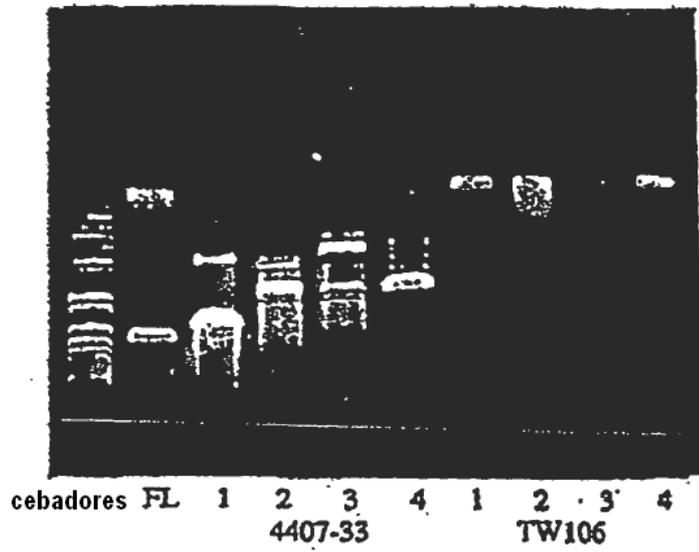


FIGURA 9

FIG. 10

SEQ ID 62 D35-B611
 1 ATGACTTAT GCATTGCAAG TGGAACACTT AACAAATGGCA-
 61 CATTGTATCC AAGGTTTCAA TTACAGAACT CCAAATGACG AGCCCTTGGG TATGAAGGAA
 121 GGTCCAGGCA TAACTATACG TAAGGTAAAT CCTGTGGAAC TGATAATAGC GCCTCGCCTG
 181 GCACCTGAGC TTTATTAA
 SEQ ID 63
 MTYALQVEHLTMAHLIQGFNYRTPNDEPLDMKEGAGITIRKVNVELIIPRLAPELY

FIG. 11

NOMBRE D120-AH4
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 188
 1 ATAATGCTTT CTCCATAGA AGCCATTGTA GGAAGTAGTAA CCTTCACATT TCTCTCTCTC
 61 TTCCTATGGA CAAAAAATC TCAAAAACCT TCAAAACCTT TACCACCGAA AATCCCCGGA
 121 GGATGCCCCG TAATCGGCCA TCTTTTCCAC TTCAATGACG ACGCGCAGCA CCGTCCATTA
 181 GCTCGAAAAA TCGGAGACTT AGCTGACAAA TACGGCCCCG TTTTCACTTT TCGGCTAGGC
 241 CTTCCTCCCTG TCTTAGTGTG AAGCAGTTAC GAAGCTGTAA AAGACTGTTT CTCTACAAAT
 301 GACGCCATTT TTTCCAATCG TCCAGCTTTT CTTTACGGCG ATTACCTTGG CTACATAAT
 361 GCCATGCTAT TTTTGGCCAA TTACGGACCT TACTGGCGAA AAAATCGAAA ATTAGTATT
 421 CAGGAAGTTC TCTCCGCTAG TCGTCTCGAA AAATTCAAAC ACGTGAGATT TGCAAGAATT
 481 CAAGCGAGCA TTAAGAATT ATATACTCGA ATTGATGGAA ATTCGAGTAC GATAAATTTA
 541 ACTGATTGGT TAGAAGAATT GAATTTTGGT CTGATCTGTA AGATGATCGC TGGAAAAAT
 601 TATGAATCCG CTAAGGAGA TGAACAAGTG GAGAGATTTA AGAAAGCGTT TAAGGATTTT
 661 ATGATTTTAT CAATGGAGTT TGTGTTATGG GATGCATTTT CAATCCATT ATTTAATGG
 721 GTGGATTTTC AAGGCCATGT TAAGGCTATG AAAAGGACTT TTAAGATAT AGATTCTGTT
 781 TTTGAGATT GGTAGGGGA ACATATTAAT AAAAGAGAAA AAATGAGGT TAATGCAGAA
 841 GGGAAATGAA AAGATTTTAT TGATGTGGTG CTTTCAAAAA TGAGTAAATGA ATATCTTGGT
 901 GAAGGTACT CTCGTGATAC TGTCAATAAA GCAACGGTGT TTAGTTTGGT CTGGATGCA
 961 GCAGACACAG TTGCTCTTCA CATAAATGG GGAATGCCAT TATGATAAAA CAATCAAAAG
 1021 GCCTTGACGA AAGCACAAGA AGAGATAGAC ACAAAGTTG GTAAGGACAG ATGGGTAGAA
 1081 GAGAGTGATA TTAAGGATTT GGTATACCTC CAAGCTATTG TTAAGAAAT GTTACGATTA
 1141 TATCCACCAG GACCTTTGTT AGTACCACAC GAAATGTAG AAGATTGTGT TGTTAGTGA
 1201 TATCACATTC CTAAGGGAC AAGATTATTC GCAAACGTCA TGAAACTGCT ACGTGATCCT
 1261 AAACCTTGGC CTGATCTTGA TACTTTTGGT CCAGAGAGAT TCATTGCTAC TGATATTGAC
 1321 TTTCTGGTTC AGTACTATAA GTATATCCCG TTTGGTTCTG GAAGACGATC TTTCCAGGG
 1381 ATGACTTATG CATTGCAAGT GGAACACTTA ACAATGGCAC ATTGATCCA AGTTTCAAT
 1441 TACAGAATTC CAAATGACGA GCCCTTGGAT ATGAAGGAAG GTGCAGGCAT AACTATACGT
 1501 AAGGTAATTC CTGTGGAATC GATAATAGCG CCGCCCTGG CACCTGAGCT TTATTAATC
 1561 CTAAGATCTT TCACTTGGT TGATCATTG? ATAATACTCC TAAATGGATA TTCATTTACC
 1621 TTTTATCAAT TAA
 SEQ. ID. NO. 189
 1 MLSPIEAVG LVTFTPLFFP LWTRKSQKPS KPLPKIPGG WPVIGHLFHP NDDGDDRPLA
 61 RULGLADKY GPVFTPLGL PLVLVVSSE AVKDCFSTRD AIPSNRPAFL YGDYLGYNNA
 121 MFLANLYGFI WRKNRKLVIQ EVLSASRLEK FKHVRFARIQ ASIKRLYTRI DGNSSINLNT
 181 DWLEELNPLG IVKNIAGKNY ESGRGEQVE RFKAFKDFM ILSMEFLVMD AFPIPLFRWV
 241 DFQGHVKAMK RIFKIDISVF QNWLGEHINK REKMEVNAEG NEQDFIDVVL SKMSNEYLGE
 301 GYSRDTVIKA TVPSLVLDAA DTVALHINWG MALLINNRKA LTKAQEEIDT KVKQRWVEE
 361 SDIKDLVYLQ AIVKEVLRLY PPGFLVPEH NVEDCVVSGY HIFKGTRLFV NVMKLLRDFK
 421 LWPDPDTFDP ERFIATDIDF RGQYKYIPF GSRRRSCPGM TYALQVEHLT MAHLIQGFNY
 481 RTPNDEPLDM KEGAGITIRK VNPVELIIP RLAPELY

FIG. 12

NOMBRE D121-AA8
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 190

1	AATCCATAAT	GCTTCTCC	ATAGAAGCCA	TTGTAGGACT	AGTAACCTTC	ACATTTCTCT
61	TCTTCTCCT	ATGGACAAA	AAATCTCAA	AACCTTCAA	ACCCTTACCA	CCGAAAATCC
121	CCGGAGGATG	GCCGGTAATC	GGCCATCTTT	TCCACTTCAA	TGACGACGGC	GACGACCGTC
181	CATTAGCTCG	AAAACCTCGA	GACTTAGCTG	ACAAATACGG	CCCGTTTTTC	ACTTTTCGGC
241	TAGGCCTTCC	CCTTGTCTTA	GTTGTAAGCA	GTTACGAAGC	TGTAAGGAC	TGTTTCTCTA
301	CAAAATGACGC	CATTTTTTCC	AATCGTCCAG	CTTTTCTTTA	CGGCGATTAC	CTTGGGTACA
361	ATAATGCCAT	GCTATTTTTG	GCCAATTACG	GACCTTACTG	GCGAAAAAAT	CGAAAAATAG
421	TTATTAGGA	AGTCTCTCC	GCTAGTCGTC	TCGAAAAAAT	CAAACACGTG	AGATTTGCAA
481	GAATTCAAGC	GAGCATTAA	AATTTATATA	CTCGAATTGA	TGGAAATTCG	AGTACGATAA
541	ATTTAACTGA	TTGGTTAGAA	GAATTGAATT	TTGGTCTGAT	CGTGAAGATG	ATCGCTGGAA
601	AAAATTATGA	ATCCGGTAAA	GGAGATGAAC	AAAGTGGAGAG	ATTTAAGAAA	GCCTTTAAGG
661	ATTTTATGAT	TTTATCAATG	GAGTTTGTGT	TATGGGATGC	ATTTCCAATT	CCATTATTTA
721	AATGGGTGGA	TTTTCAAGGG	CATGTTAAGG	CTATGAAAAG	GACTTTTAAA	GATATAGATT
781	CTGTTTTTCA	GAATTGGTTA	GAGGAACATA	TTAATAAAG	AGAAAAAATG	GAGGTTAATG
841	CAGAAGGGAA	TGAACAAGAT	TTCATTGATG	TGGTCTTTC	AAAAATGAGT	AATGAATATC
901	TTGGTGAAGG	TTACTCTCGT	GATACGTCA	TTAAAGCAAC	CGTGTTTAGT	TTGGTCTTGG
961	ATGCAGCAGA	CACAGTTGCT	CTTCACATAA	ATTGGGGAAT	GGCATTATTG	ATAAACAAATC
1021	AAAAGGCCCT	GACGAAAAGCA	CAAGAAGAGA	TAGACACAAA	AGTTGGTAA	GACAGATGGG
1081	TAGAAGAGAG	TGATATTAA	GATTTGGTAT	ACCTCCAAGC	TATTGTTAAA	GAAGTGTAC
1141	GATTATATCC	ACCAGGACCT	TTGTTAGTAC	CACACGAAA	TGTAGAAGAT	TGTGTTGTTA
1201	GTGGATATCA	CATTCCTAAA	GGGACAAGAT	TATTCGCAA	CGTCATGAAA	CTGCAACGTG
1261	ATCCTAAACT	CTGGTCTGAT	CCTGATACTT	TCGATCCAGA	GAGATTCATT	GCTACTGATA
1321	TTGACTTTCG	TGGTCACTAC	TATAAGTATA	TCCCGTTTGG	TTCTGGAAGA	CGATCTTGTG
1381	CAGGGATGAC	TTATGCATTG	CAAGTGGAA	ACTTAACAAT	GGCACATTG	ATCCAAGGTT
1441	TCATTACAG	AACTCCAAT	GACGAGCCCT	TGGATATGAA	GGAAGGTGCA	GGCATAACTA
1501	TACGTAAGGT	AAATCCTGTG	GAAGTGTATA	TAGCGCCTCG	CCTGGCACCT	GAGCTTTATT
1561	AAAACCTAAG	ATCATCTTGC	TTGAT			

SEQ. ID. NO. 191

1	MLSPIEAIVG	LVTFTFLFFF	LWTKKSQKPS	KPLPPKIPGG	WPVIGHLPHF	NDGDDDRPLA
61	RKLGDLADKY	GPVFTFRLGL	PLVLVVSSYE	AVKDCFSTND	AIFSNRPAPL	YGDYLYNNA
121	MLFLANYGFPY	WRKQRKLVIQ	EVLSASRLEK	FKHVRFARIQ	ASIKNLYTRI	DGNSSSTINLT
181	DWLEELNPL	IVKMIAGINY	ESGKGEQVE	RFKKAFKDFM	ILSMFVLWD	APPIPLFKWV
241	DFQGHVKAMK	RTFKDIDSVF	QNWLEEHINK	REKMEVNAEG	NEQDFLDVVL	SKMSNEYLGE
301	GYSRDTVIKA	TVFSLVLDAA	DTVALHINWG	MALLINNQKA	LTKAQEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRLY	PPGPLLVPHE	NVEDCVVSGY	HIPKGTRLPA	NVMKLRDPK
421	LWSDFDTFDP	ERFIATDIDF	RGQYKYIPF	GSGRRSCPDM	TYALQVEHIT	MAHLIQGFNY
481	RTPNDEPLDM	KEGAGITIRK	VNFVELIAP	RLAPELY		

FIG. 13

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

1	CTAAA	ACTCC	ATAATGGTTT	CTCCC	GATAGA	AGCCATTGTA	GGACTAGTAA	CCCTTACACT
61	TCTCT	TCTAC	TTCCATGGC	CCAAAAA	AAT	TCAAATACCT	TCAAAACCAT	TACCACCGAA
121	AATTCCC	GGA	GGGTGGCCGG	TAATCGG	CCA	TCTTTTCTAC	TTCGATGATG	ACGGCGACGA
181	CCGTCC	ATTA	GCTCGAAAAC	TCGGAGACTT	AGCTGACAAA	TACGGCCCCG	TTTTCACTTT	
241	CCGGCT	AGGC	CTTCCGCTTG	TGTTAATTGT	AAGCAGTTAC	GAAGCTGTAA	AAGACTGCTT	
301	CTCTACA	AAT	GACGCCATTT	TCTCCAATCG	TCCAGCTTTT	CTTTACGGTG	AATACCTTGG	
361	CTACAATA	AAT	GCCATGCTAT	TTTTGACAAA	ATACGGACCT	TATTGCGGAA	AAAATAGAAA	
421	ATTAGT	CATT	CAGGAAGTTC	TCTCTGCTAG	TCGTC	TCGAA	AAATTGAAGC	ACGTGAGATT
481	TGGTAAA	AAT	CAAACGAGCA	TTAAGAGTTT	ATACACTCGA	ATTGATGGAA	ATTTCGAGTAC	
541	GATAAAT	CTA	ACTGATTGGT	TAGAAGAATT	GAATTTTGGT	CTGATCGTGA	AAATGATCGC	
601	TGGGAAAA	AAT	TATGAATCCG	GTAAGGAGA	TGAAC	AAGTG	GAGAGATTTA	GGAAAGCGTA
661	TAAGGAT	TTTT	ATAATTTTAT	CAATGGAGTT	TGTGTTATGG	GATGCTTTTC	CAATCCATT	
721	GTTCAA	ATGG	GTGGATTTTC	AAGGCTATGT	TAAGGCCATG	AAAAGGACAT	TTAAGGATAT	
781	AGATTCT	GTT	TTTCAGAAAT	GGTTAGAGGA	ACATGTC	AAG	AAAAGAGAAA	AAATGGAGGT
841	TAATGC	ACAA	GGGAATGAC	AAGATTT	CAAT	TGATGTGGTG	CTTTC	AAAAA
901	ATATCT	TGAT	GAAGGTTACT	CTCGTGATAC	TGTCATA	AAAA	GCAACAGTGT	TTAGTTTGGT
961	CTTGG	ATGCT	GCGGACACAG	TTGCTCTTCA	CATGAATGG	GGAA	TGGCAT	TACTGATAAA
1021	CAATCA	AACAT	GCCTTGAAGA	AAGCACAAGA	AGAGATCGAT	AAGAA	AGTTG	GTAAGGAAAG
1081	ATGGG	TAGAA	GAGAGTGATA	TTAAGGATTT	GGTCTACCTC	CAAGCTATTG	TTAAAGAA	AGT
1141	GTTAC	GATTA	TATCCACCAG	GACCTTTATT	AGTACCTCAT	GA	AAATGTAG	AGGATTGTGT
1201	TGTTAG	TGGA	TATCACATTC	CTAAAGGGAC	TAGACTATTC	GCGAACGTTA	TGAAATGCA	
1261	GCGGAT	CCCT	AAACTCTGGT	CAAATCCTGA	TAAGTTGAT	CCAGAGAGAT	TCTTCGCTGA	
1321	TGATA	TTGAC	TACCGTGGTC	AGCACTATGA	GTTTATCCCA	TTTGGTCTG	GAAGACGATC	
1381	TTGTC	CGGGG	ATGACTTATG	CATTACAAGT	GGAACACCTA	ACAATAGCAC	ATTTGATCCA	
1441	GGGTT	TCAAT	TACAAAACTC	CAAATGACGA	GCCCTTGGAT	ATGAAGGAAG	GTGCAGGATT	
1501	AAC	TACGT	AAAGTAAATC	CTGTAGAAGT	GACAATTACG	GCTCGCTGG	CACCTGAGCT	
1561	TTATTA	AAAC	CTTAGATGTT	TTATCTTGAT	TGTACTAATA	TATATATGCA	GAAAAAATG	

SEQ. ID. NO. 193

1	MVSPVE	AI	VG	LVTLL	LLFYF	LWPKK	FQIPS	KPLPK	IPGG	WPVIGH	LFYF	DDGD	DRPLA
61	RKLGDL	ADKY	GPVFT	FRLGL	FLVLIV	SSYE	AVKDC	FSTND	AIPSNR	PAFL	YGEYL	GYNNA	
121	MLFLTK	YGPY	WRKNR	KLVIQ	EVLSAS	RLEK	LKHVR	FGKIQ	TSIKSL	YTRI	DGNS	SSTINLT	
181	DWLEEL	NFGL	IVKMI	AGKNY	ESGKG	DEQVE	RFRKAY	KDFI	ILSME	FVLD	AFPI	PLFKWV	
241	DFQGY	VKAMK	RTFKD	IDSVF	QNWLE	EEHVK	REKME	VNAQG	NEQDF	IDVVL	SKMS	NEYLDE	
301	GYSRDT	VIKA	TVFSL	VLDAA	DTVAL	HMNWG	MALLINN	QHA	LKKAQ	EEIDK	KVGKER	WVRE	
361	SDIKDL	VYLQ	AIVKE	VLRLY	BPGPL	LVPH	NVEDC	VVSGY	HIPKG	TRLFA	NVMKL	QRDPK	
421	LWSNFD	KFDP	ERFFA	DDIDY	RGQHY	EPIF	GSGRR	SCPGM	TYALQ	VEHLT	LAHLI	IQGFNY	
481	KTFNDE	PLDM	KEGAG	LTI	RK	VNPVE	VTTITA	RLAPE	LY				

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	TTTTCAC TTC	AATAACGACG	GGCAGGACCG	TCCATTAGCT
181	CGAAAGCTCG	GAGACTTAGC	TEATAAATAC	GGCCCCGTTT	TCAC TTTTCG	GCTAGGTCTT
241	CCCCTTGTGC	TAGTTGTAAG	CAGTTACGAA	GCTATAAAAAG	ATTGCTTCTC	TACAAATGAT
301	GCCATTTTCT	ECAATCGTCC	AGCTCTTCTT	TACGGCGAAT	ACCTTGGCTA	CAATAATACA
361	ATGCTTTTTT	TAGCAAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGTAAATT	AGTCATT CAG
421	GAAGTTCTCT	CTGCTAGTCG	TC TCGAAAAA	TTCAAACAAG	TGAGATT CAC	CAGAATTCAA
481	ACGAGCATT A	AGAATTTATA	CACTCGAATT	AATGGAAATT	CGAGTACGAT	AAATCTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAAAA	TGATCGCTGG	GAAAAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAA	AGATTTAAGA	ATGCGTTTAA	GGATTTTATG
661	GTTTTATCAA	TGGAATTTGT	ATTATGGGAT	GCATTTCCAA	TTCCATTATT	TAATGGGTG
721	GATTTTCAAG	GTCATATTAA	GGCAATGAAA	AGGACATTTA	AGGATATAGA	TTC TGT TTTT
781	CAGAACTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TAGAGGTTGG	TGCAGAAGGG
841	AATGAACAAG	ATTTCAATTGA	TGTGGTCTT	TCAA AATTGA	GTAAGAATA	TCTTGATGAA
901	GGTTACTCTC	GTGATACTGT	CATTAAGCA	ACA G TTTT TA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTCACAT	AAATTGGGGA	ATGACATTAT	TGATAAACAA	TCAAAATGCC
1021	TTGATGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGGATAGATT	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTAGT	ATACCTCCAA	GCTATTGTTA	AAAAGGTGTT	ACGATTATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACATGAA	AATGTAAAGG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGGACTAG	ATTATTGCGA	AACGTCATGA	AACTGCAGCG	CGATCCTAAA
1261	CTCTTGFCAA	ATCTGATATA	GTTGATCCA	GAGAGATTCA	TCCGTGGTGA	TATTGACTTC
1321	CGTGGTCACC	ACTATGAGTT	TATCCCATTT	GGTCTGGAA	GACGATCTTG	TCCGGGGATG
1381	ACTTATGCAT	TGCAAGTGGA	ACACCTAACA	ATGGCACATT	TAATCCAGGG	TTTCAATTAC
1441	AAAAC TCCAA	ATGACGAGGC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	AATACGTAAG
1501	GTA AATCCAG	TGGAATTGAT	AA TAACGCCCT	CGCTTGGCAC	CTGAGCTTTA	CTAAAACCTA
1561	AGATGTTTCA	TCTTGGTTGA	TCATTGT			

SEQ. ID. NO. 227

1	MLSPIEAFVG	LVTFTFLLYF	LWTKKSQKLP	KPLPPKIPGG	WPIVGHLEFHF	NNDGDDRFLA
61	RKLGDLADKY	GPVFTFRLGL	PLVLVVSSEY	AIKDCPSTND	AIFSNRPALL	YGEYLGYNNT
121	MLFLANYGPY	WRKNRKLVIQ	EVLSASRLEK	FKQVRFTRI Q	TSIRKLYTRI	NGNSSTINLT
181	DWLEELNPGF	IVKMIAGKNY	ESGKGDEQVE	RFKNAFKDFM	VLSMEFVLWD	APP I PLFKWV
241	DFQGHKAMK	RTPKDIDSVF	QNWLEEHINK	REKIEVGAEG	NEQDFIDVVL	SKLSKEYLDE
301	GYSRDTV IKA	TVFSLVLDA A	DTVALHINWG	MTLLINNQNA	LMKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKVLRLRY	PPGPLLV PHE	NVKDCVVS GY	HIPRGTRLFA	NVMKLRDPRK
421	LLSNPKFDFP	ERFIAGDIDF	RGHHYEF I PF	GSGRRSCP GM	TYALQVEHLT	MAHLIQGFNY
481	KTFNDEALDM	KEGAGITIRK	VNPFVELIITP	RLAPELY		

FIG. 15
 NOMBRE D103-AH3
 ORGANISMO NICOTIANA TABACUM
 SEQ. ID. NO. 230

1	ATGGTTTTTC	CCATAGAAGC	CTTTGTAGGA	CTAGTAACCT	TCACATTTCT	CTTATACTTC
61	CTATGGACAA	AAAAATCTCA	AAAACCTTCCA	AAACCCTTAC	CACCGAAAAT	CCCCGGAGGA
121	TGGCCGGTAA	TCGGCCACCT	TTTTCACTTC	AATAACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAAC TCG	GAGACTTAGC	TGATAAATAC	GGCCCCGTTT	TCACTTTTCG	GCTAGGTCCT
241	CCCTTTGTGC	TAGTTGTAAAG	CAGTTACGAA	GCTACAAAAG	ATTGCTTCTC	TACAAATGAC
301	GCCATTTTCT	CCAATCGTCC	AGCTTTTCTT	TACGGCGAAT	ACCTTGGGCTA	CAATAATACA
361	ATGCTTTTTC	TAGCAAATTA	CGSACCTTAC	TGGCGAAAAA	ATCGTAAAT	AGTCATTACG
421	GAAGTTCTCT	CTGCTAGTCG	TCTCGAAAAA	TTCAAACAAG	TGAGATTAC	CAGAATTCAA
481	ACGAGCATT	AGAATTATA	CACTCGAATT	AATGGAAAT	CGAGTACGAT	AAATCTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAATA	TGATCGCTGG	GAATAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAA	AGATTTAAGA	ATGCGTTTAA	GGATTTTATG
661	GTTTTATCAA	TGGAATTTGT	ATTATGGGAT	GCATTTCCAA	TTCCATTATT	TAAATGGGTG
721	GATTTTCAAG	GTCATATTAA	GACAAATGAA	AGGACATTTA	AGGATATAGA	TTCTGTTTTT
781	CAGAACTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTGG	TGCAGAAGGG
841	AATGAACAAG	ATTTCAATGA	TGTGGTGTCT	TCAAATTTGA	GTAAGAATA	TCTTGATGAA
901	GGTACTCTC	GTGATACTGT	CATTAAGCA	ACAGTTTTTA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTCACAT	AAATGGGGGA	ATGACATTAT	TGATAAACAA	TCAAAATGCC
1021	TTGATGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGGATAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTAGT	ATACCTCCAA	GCTATTGTGA	AAAAGGTGTT	ACGATTATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACATGAA	AATGTAAAGG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGGACTAG	ATTATCGCA	AACGTACATGA	AACCTGCAGCG	CGATCCTAAA
1261	CTCTTGTC	ATCCTGATAA	GTTTCGATCCA	GAGAGATTCA	TCGCTGGTGA	TATTGACTTC
1321	CGTGGTCACC	ACTATGAGTT	TATCCCATCT	GTTCTGGAA	GACGATCTTG	TCCGGGGATG
1381	ACTTATGCAT	TGCAAGTGG	ACACTAACA	ATGGCACATT	TAATCCAGGG	TTTCAATTAC
1441	AAAACCTCAA	ATGACGAGGT	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	AATACGTAAG
1501	GTAAATCCAG	TGGAATTTGAT	AATAACGCCT	CGCTTGGCAC	CTGAGCTTTA	CTAAAACCTA
1561	AGATCTTTCA	TCTTGGTTGA	TCATGTGTTA	ATA		

SEQ. ID. NO. 231

1	MVFPIEAFVG	LVTFTFLLYF	LWTKKSQKLP	KPLPPKIPGG	WVIGHLHFH	NNDGDDRPLA
61	RKLGDLADKY	GPVFTFRGL	PLVLVSSYE	ATKDCPSTND	AIPSNRPAPL	YGEYLYNNT
121	MLFLANYGPY	WRKNRKLVIQ	EVLSASRLEK	FKQVRFTRIQ	TSIKNLYTRI	NGNSSITNLT
181	DWLEELNFGL	IVKMIAGKNY	ESGKDEQVE	RFKNAFKDFM	VLSMEFVLWD	AFPIPLFKWV
241	DFQGHIRTKM	RTFKDIDSVP	QNWLEEHINK	REKMEVGAEG	NEQDFIDVVL	SKLSKEYLDE
301	GYSRDTVIKA	TVFSLVLDAA	DTVALHINWG	MLLINNQNA	LMKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKVLRLY	PFGPLVFPHE	NVKDCVVSFY	HIPKGTRLF	NVMKLRDPK
421	LLSNFDFKDP	ERFLAGDIDF	RGHHYEFIPS	GSGRRSCP	TYALQVEHLT	MAHLIQGFNY
481	KTPNDEVLD	KEGAGITIRK	VNFVELIITP	RLAPELY		

FIG. 16

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

1	ATGCTTCTC	CCATAGAAGC	CATTGTAGGA	CTAGTAACCT	TCACATTCT	CCTCTTCTC
61	CTATGGACAA	AAAAATCTCA	AAAACCTTCA	AAACCCTTAC	CACCGAAAT	CCCCGGAGGA
121	TGGCCGGTAA	TCGGCCATCT	TTTCCACTTC	AATGACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAACCTCG	GAGACTTAGC	TGACAAATAC	GCSCCCGTTT	TCACTTTTCG	GCTAGGCCCTT
241	CCCCTTGTCT	TAGTTGTAAG	CAGTTACGAA	GCTGTAAAAG	ACTGTTTCTC	CACAAAATGAC
301	GCCATTTTTT	CCAATCGTCC	AGCTTTTCTT	TACGGCGATT	ACCTTGGCTA	CAATAATGCC
361	ATGCTATTTT	TGGCCAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGAAAATT	AGTTATTTCAG
421	GAAGTTCTCT	CCGCTAGTCT	TCTCGAAAAA	TTCAAACACG	TGAGATTTGC	AAGAATTCAA
481	GCGAGCATGA	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	CAGAAATGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGTTTAA	TGCGAAGGGG
841	AATGAACAAG	ATTTCAATGA	TGTGGTCTT	TCAAAAATGA	GTAATGAATA	TCCTGGTGAA
901	GGTTACTCTC	GTGATACTGT	CATTGAAGCA	ACGGTGTTTA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CCTTTCACAT	AAATTGGGGA	ATGGCATTAT	TGATAAACAA	TCAAAAAGCC
1021	TTGACGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTTGTA	AGGACAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTGGT	ATACCTCCAA	GCTATTGTTA	AAGAAGTGTI	ACGATTATAT
1141	CCACCAGGAC	CTTGTGTTAGT	ACCACACGAA	AATGTAGAAG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCCTA	AAGGCACAAG	ATTATTGCGA	AACGTATGA	AACGTCAACG	TGATCCTAAA
1261	CTCTGGTCTG	ATCCTGATAC	TPTCGATCCA	GAGAGATTCA	TTGCTACTGA	TATTGACTTT
1321	CGTGGTCACT	ACTATAAGTA	TATCCCCTTT	GGTCTGGAA	GACGATCTTG	TCCAGGGATG
1381	ACTTATGCAT	TGCAAGTGGG	ACACTTAACA	ATGGCACATT	TGATCCAAGG	TTTCAATTAC
1441	AGAATCCAA	ATGACGAGCC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	TATACGTAAG
1501	GTAATCCTG	TGGAACCTGAT	AATAGCGCCT	CGCCTGGCAC	CTGAGCTTTA	TTAAAACCTA
1561	AGATGTTTCA	TCTTGGTTGA				

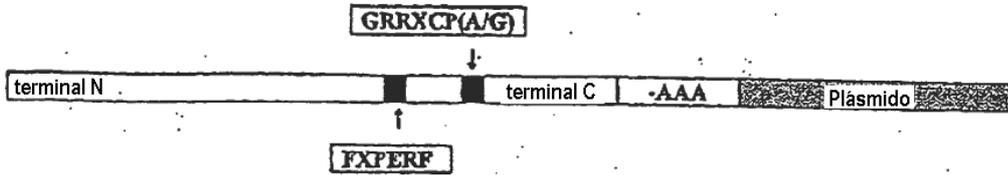
SEQ. ID. NO. 233

1	MLSPIEAIVG	LVTFTPLFFF	LWTKKSQKPS	RFLPPKIPGG	WVVIHGLPHF	NDGDDRPLA
61	RKIGDLADRY	GPVFTFRLGL	PLVLVVSSYE	AVKDCFSTND	AIFSNRPAPL	YGDYLGYNNA
121	MLFLANYGPY	WRKNRKLVIQ	EVLASARLEK	PKHVRFARIQ	ASMKNLYTRI	DGNSSTINLT
181	DWLEELNFGI	IVKNIAGKNY	ESGKGDEQVE	RFKKAFFKDFN	ILSMFVFLWD	APPIPLFKWV
241	DFQGHVKAMK	RTPKIDISVF	QNWLEEHINK	REKMEVNAEG	NEQDFIDVVL	SKMSNEYLGE
301	GYSRDTVIER	TVPSLVLDAA	DTVALHINWG	MALLINNQKA	LTKAQEEIDT	KVCKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRBY	PPGPLLVPHE	NVEDCVVSGY	HIPKGRLEFA	NVNRLQRDPK
421	LWSDPDTFDP	ERFIATDIDF	RGQYYKVIIPF	GPGRRSRCPGM	TYALQVEHLT	MAHLIQGPNY
481	RTPNDEPLDM	KEGAGITIRK	VNPVELIIP	RLAPELY		

FIG. 17

NOMBRE	D235-AB1					
ORGANISMO	NICOTIANA TABACUM					
SEQ. ID. NO.	254					
1	AAAATTCATA	ATGGTTTTTC	CCATAGAAGC	CTTTGTAGGA	CTAGTAACCT	TCACATTTCT
61	CTTATACTTC	CTATGGACAA	AAAAATCTCA	AAAAC TTCCA	AAACCCTTAC	TACCGAAAAT
121	CCCCGGAGGA	TGGCCGGTAA	TCGGCCATCT	TTTTCACTTC	AATAACGACG	GCGACGACCG
181	TCCATTAGCT	CGAAAAC TCG	GAGACTTAGC	TGATAAATAC	GGCCCCGTTT	TCACTTTTCG
241	GCTAGGTCTT	CCCCTTGTGC	TAGTTGTAAG	CAGTTACGAA	GCTATAAAAAG	ATTGCTTCTC
301	TACAAATGAC	GCCATTTTCT	CCAATCGTCC	AGCTTTTCTT	TACGGCGAAT	ACCTTGGCTA
361	CAATAATACA	ATGCTTTTTC	TAGCAAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGTA AATT
421	AGTCATT CAG	GAAGTCTCT	CTGCTAGTCG	TCTCGAAAAA	TTCAAACAAG	TGAGATT CAC
481	CAGAATTCAA	ACGAGCATT A	AGAATTTATA	CACTCGAATT	AATGGAAATT	CGAGTACGAT
541	AAATCTA ACT	GATTGGTTAG	AAGAATGGA	TTTTGGTCTG	ATCGTAAAA	TGATCGCTGG
601	GAAAAATTAT	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAA	AGATTTAAGA	ATGCGTTTAA
661	GGATTTTATG	GTTTTATCAA	TGGAATTTGT	ATTATGGGAT	GCATTTCCAA	TTCCATTATT
721	TAAATGGGTG	GATTTTCAAG	GTCATATTA	GGCAATGAAA	AGGACATTTA	AGGATATAGA
781	TTCTGTTTTT	CAGAACTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTTGG
841	TGCAGAAGGG	AATGAACAAG	ATTTCAATGA	TGTGGTGCTT	TCAAAATTGA	GTAAGAATA
901	TCTTGATGAA	GGTACTCTC	GTGATACTGT	CATTAAGCA	ACAGTTTTTA	GTTTGGTCTT
961	GGATGCAGCA	GACACAGTTG	CTCTCACAT	AAATTGGGGA	ATGACATTAT	TGATAAACAA
1021	TCAAAATGCC	TTGATGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGTATAGATG
1081	GGTAGAAGAG	AGTGATATTA	AGGATTTAGT	ATACCTCAA	GCTATTGTTA	AAAAGGTGTT
1141	ACGATTATAT	CCACCAGGAC	CTTTGTTAGT	ACCACATGAA	TATGTAAGG	ATTGTGTTGT
1201	TAGTGGATAT	CACATTCCTA	AAGGGACTAG	ATTATTCGCA	AACGTCATGA	AACTGCAGCG
1261	CGATCCTAAA	CTCTTGTCAA	ATCCTGATAA	GTTCCGATCCA	GAGAGATTCA	TCGCTGGTGA
1321	TATCGACTTC	CGTGGTCACC	ACTATGAGTT	TATCCCAATT	GGTCTCGAA	GACGATCTTG
1381	TCCGGGGATG	ACTTATGCAT	TGCAAGTGG A	ACACCTAACA	ATGGCACATT	TAATCCAGGG
1441	TTTCAATTAC	AAAAC TCCAA	ATGACGAGGC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC
1501	AATACGTAAG	GTA AATCCGG	TGGAATTGAT	AATAACGCCT	CGCTTGGCAC	CTGAGCTTTA
1561	CTAAACCTA	AGATCTTTCA	TCTTGGTTGA	TCATTGTTTA	ATACTCCTAG	ATAGATGGGT
1621	ATTCATC					
SEQ. ID. NO.	255					
1	MVPPIEAFVG	LVTFTFLLYF	LWTKKSQKLP	KELLPKIPGG	WPVIGHLPHF	NNDGDDRPLA
61	RKLGDLADKY	GFVFTFRLGL	PLVLVVSSE	AIKDCFSTND	AIFSNRPAFL	YGEYLGYNNT
121	MLFLANYGPI	WRKNRKLVIQ	EVLSASRLEK	FKQVRFTRIQ	TSIKNLYTRI	NGNSSTINLT
181	DWLEELDFGL	IVKMIAGKNY	ESGKGDEQVE	RPKNAFKDFM	VLSMEFVLWD	AFPIPLFKWV
241	DFQGHKAMK	RTFKDIDSVP	QNWLEEHINK	REKMEVGAEG	NEQDFIDVVL	SKLSKEYLDE
301	GYSRDTVIKA	TVPSLVLDAA	DTVALHINWG	MTLLINNQNA	LMKAQEEIDT	KVGKYRWVZE
361	SDIKDLVYLQ	AIVKKVLRLY	PPGFLVPHIE	YVKDCVVSQY	HIPKGRLEFA	NVMKLRDPK
421	LLSNPKDFDP	ERFIAGDIDF	RGHHEPIPF	GSGRRSCPQM	TYALQVEHLT	MAHLIQGFNY
481	KTFNDEALDM	KEGAGITIRK	VNPFVELIITP	RLAPELY		

FIG. 13 : Clonación de Fragmentos de ADNc de Citocromo p450 por PCR



DM	FXPERF -for	5'TTYIICCIGARMGTTY-3'	SEQ ID NO:2255
DM4	GRRXCP(A/G)-for	5'-GGIMGIMGIIITGYCCIGS-3'	SEQ ID NO:2256
DM12	FKPERF-for	5'-TTYAARCCTGAGAGATT-3'	SEQ ID NO:2257
DM13	PERFL-for	5'-CCAGARAGATTCTTG-3'	SEQ ID NO:2258
DM17	GRRMCP-for	5'-GGRMGRMGRATGTGYCC-3'	SEQ ID NO:2259
OLIGO d(T)		5'-TTTTTTTTTTTTTTTTTIN-3'	SEQ ID NO:2260
T7		5'-ATTATGCTGAGTGATATCCC-3'	SEQ ID NO:2261
SP6		5'-ATTJAGGTGACACTATAG-3'	SEQ ID NO:2262

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

Nombre del Conjunto de Sondas	Sonda		Posición de Interrogación de la Sonda	Secuencia de la Sonda	SEQ ID
	X	Y			
GEN1018_x_at	D121-AA8	107 105	58	TATGAATCCGGTAAGGAGATGAAC	414
GEN1018_x_at	D121-AA8	69 1	84	TCCGGTAAAGGAGATGAACAAGTGG	415
GEN1018_x_at	D121-AA8	1 15	80	AACAAGTGGAGAGATTTAAGAAAGC	416
GEN1018_x_at	D121-AA8	111 91	120	GATTTTATCAATGGAGTTTGTGTTA	417
GEN1018_x_at	D121-AA8	63 115	142	TTATGGGATGCATTTCCAATTCAT	418
GEN1018_x_at	D121-AA8	54 111	145	TGGGATGCATTTCCAATTCATTT	419
GEN1018_x_at	D121-AA8	14 97	147	GGATGCATTTCCAATTCATTTATTT	420
GEN1018_x_at	D121-AA8	35 85	148	GATGCATTTCCAATTCATTTATTTA	421
GEN1018_x_at	D121-AA8	98 117	150	TGCATTTCCAATTCATTTATTTAAA	422
GEN1018_x_at	D121-AA8	98 59	184	CATTTTAAATGGTGGATTTTCA	423
GEN1018_x_at	D121-AA8	12 23	182	ATTTTCAAGGCAATGTTAAGGCTAT	424
GEN1018_x_at	D121-AA8	27 53	187	CAAGGCAATGTTAAGGCTATGAAA	425
GEN1018_x_at	D121-AA8	45 79	185	TGTTAAGGCTATGAAAAGGACTTTT	426
GEN1018_x_at	D121-AA8	94 47	359	AAGGTTACTCTCGTGATACTGTCT	427
GEN1018_x_at	D121-AA8	2 73	418	GCAGACACAGTTGCTCTTCACATAA	428
GEN1018_x_at	D121-AA8	74 83	588	GTTACGTTTATATCCACCAGGACCT	429
GEN1018_x_at	D121-AA8	88 3	604	CCAGGACCTTTGTTAGTACCACACG	430
GEN1018_x_at	D121-AA8	31 3	610	CCFTTGTAGTACCACACGAAATG	431
GEN1018_x_at	D121-AA8	111 11	719	AACCTGGTCTGATCCTGATACTTT	432
GEN1018_x_at	D121-AA8	32 55	775	GACTTTCGGTCCAGTACTATAGT	433
GEN1018_x_at	D121-AA8	102 103	783	TGGTCAGTACTATAAGTATATCCCG	434
GEN2012_x_at	D35-BG11	32 7	58	GATCCAGGTTTCAATTCAGAACT	435
GEN2012_x_at	D35-BG11	118 101	114	GTCCAGGCATACTATACGTAAGGT	436
GEN2012_x_at	D35-BG11	21 27	140	AATCCTGTGGAACTGATAATAGCGC	437
GEN2012_x_at	D35-BG11	66 21	141	ATCCTGTGGAACTGATAATAGCGCC	438
GEN2012_x_at	D35-BG11	108 17	143	CCCTGTGGAACTGATAATAGCGCCTC	439
GEN2012_x_at	D35-BG11	71 105	145	TGTGGAACCTGATAATAGCGCCTCGC	440
GEN2012_x_at	D35-BG11	64 98	148	GGAACTGATAATAGCGCCTCGCCTG	441
GEN2012_x_at	D35-BG11	14 77	149	GAACTGATAATAGCGCCTCGCCTGG	442

FIGURA 20
PÁGINA 2 DE 3

Nombre del Conjunto de Sondas	Sonda		Posición de Interrogación de la Sonda	Secuencia de la Sonda	SEQ ID
	X	Y			
GEN1018_x_at	107	105	58	TATGATCCGGTAAAGGAGATGAAC	414
GEN1019_x_at	69	1	84	TCCGGTAAGCAGAGATGAACAAGTGG	415
GEN1019_x_at	1	15	80	AACAAGTGGAGAGATTTTAGNAACC	416
GEN1019_x_at	111	81	120	GATTTTATCAATGGAGTTTGTTA	417
GEN1019_x_at	63	115	142	TTATGGGATGCAATCCAAATCCAT	418
GEN1019_x_at	54	111	145	TGGGATGCATTTCCAAATCCATAT	419
GEN1019_x_at	14	87	147	GGATGCATTTCCAAATCCATATATTT	420
GEN1019_x_at	35	85	148	GATGCATTTCCAAATCCATATATTA	421
GEN1019_x_at	98	117	150	TGCATTTCCAAATCCATATATTA	422
GEN1019_x_at	96	59	184	CAATATTTAAATGGGTGGATTTCA	423
GEN1019_x_at	12	23	182	ATTTCAAGGGCATTGTAAGGCTAT	424
GEN1019_x_at	27	53	187	CAAGGGCATTGTAAGGCTATGAAA	425
GEN1019_x_at	45	79	185	TGTTAAGGCTATGAAAAGGACTTTT	426
GEN1019_x_at	94	47	359	AAGGTACTCTCGTGAFACTGTCT	427
GEN1019_x_at	2	73	418	GCAGACACAGTTGCTCTTCACATA	428
GEN1019_x_at	74	83	588	GTTACGATATATCCACCAGGACCT	429
GEN1019_x_at	66	3	604	CCAGGACCTTTGTTAGTACCACACG	430
GEN1019_x_at	31	3	610	CCTTTGTAGTACCACACGAAATG	431
GEN1019_x_at	111	11	719	AACCTGTCTGATCTGTGATACTTTT	432
GEN1019_x_at	32	55	775	GACTTTCGTGGTCAGTACTATTAAGT	433
GEN1019_x_at	102	103	783	TGGTCAGTACTATATGATATATCCCG	434
GEN2012_x_at	32	7	58	GATCCAAGGTTTCAATTACAGAACT	435
GEN2012_x_at	119	101	114	GTCCAGGCATAACTATATAGTAAAGT	436
GEN2012_x_at	21	27	140	AATCCTGTGGAACTGATAAATAGCGC	437
GEN2012_x_at	68	21	141	ATCCTGTGGAACTGATAAATAGCGCC	438
GEN2012_x_at	108	17	143	CCTGTGGAACTGATAAATAGCGCTC	439
GEN2012_x_at	71	105	146	TGTGGAACTGATAAATAGCGCTCGC	440
GEN2012_x_at	64	99	148	GGAAGTATATATAGCGCTCGCTG	441
GEN2012_x_at	14	77	149	GAACTGATAAATAGCGCTCGCTGG	442

FIGURA 20
PÁGINA 2 DE 3

Nombre del Conjunto de Sondas	Sonda		Posición de Interrogación de la Sonda	Secuencia de la Sonda	SEQ ID
	X	Y			
GEN1018_x_at	7	85	14	GAAAGCCCTTAAAGGATTTATGAT	385
GEN1018_x_at	110	91	35	GATTTATCAATGGAGTTTGTTA	386
GEN1018_x_at	85	115	57	TTATGGGATGCATTTCCAATGCCAT	387
GEN1018_x_at	52	111	60	TGGGATGCATTTCCAATGCCAT	388
GEN1018_x_at	18	97	82	GGATGCATTTCCAATGCCATTTT	389
GEN1018_x_at	38	85	63	GATGCATTTCCAATGCCATTTTA	390
GEN1018_x_at	99	117	65	TGCATTTCCAATTTCCATTTATTA	391
GEN1018_x_at	18	69	66	GCATTTCCAATTTCCATTTATTAAT	392
GEN1018_x_at	59	3	82	TATTTAATGGTGGATTTTCAAGG	393
GEN1018_x_at	13	23	97	AATTTCAAGGGCATGTTAAGGCTAT	394
GEN1018_x_at	28	53	102	CAAGGGCATGTTAAGGCTATGNA	395
GEN1018_x_at	10	93	108	GGCATGTTAAGGCTATGAAAAGGAC	396
GEN1018_x_at	44	79	110	TGTTAAGGCTATGAAAAGGACTTTT	397
GEN1018_x_at	95	47	274	AAGTTACTCTCGTATGACTGTCAT	398
GEN1018_x_at	1	73	333	GCAGACACAGTTGCTCTTCACATA	399
GEN1018_x_at	63	41	339	ACAGTTGCTCTTCACATAAATGGG	400
GEN1018_x_at	84	81	473	GGTATACCTCCAGCTATGTTAAA	401
GEN1018_x_at	25	95	501	GTGTTACGATTAATCCACCAGGAC	402
GEN1018_x_at	69	21	505	TACGATTAATCCACCAGGACCTTT	403
GEN1018_x_at	67	3	519	CCAGGACCTTTGTTAGTACCACAG	404
GEN1018_x_at	30	3	525	CCTTTGTAGTACCACACGAAATG	405
GEN1018_x_at	1	107	622	TACGTATCCTAACTCTGGCCCTGA	406
GEN1018_x_at	72	18	643	CTGATCCTGATACTTTCGATCCAGA	407
GEN1018_x_at	33	55	690	GACTTTCGTGGTCCAGTACTATAACT	408
GEN1018_x_at	103	103	698	TGGTCAGTACTATAAGTATATCCCG	409
GEN1018_x_at	107	73	13	GAATTTGAATTTTGGTCTGATCGTGA	410
GEN1018_x_at	65	107	24	TGGTCTGATCGTGAAGATGATCGCT	411
GEN1018_x_at	118	25	31	ATCGTGAAGATGATCGCTGGAAAA	412
GEN1018_x_at	72	71	39	GATGATCGCTGGAAAAAATTTATGAA	413

FIGURA 20
PÁGINA 1 DE 3