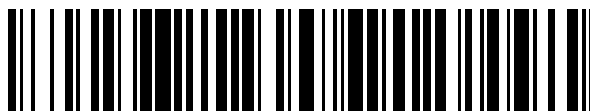


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 373 881**

51 Int. Cl.:
C12N 9/10

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **05779758 .1**
96 Fecha de presentación: **20.04.2005**
97 Número de publicación de la solicitud: **1751280**
97 Fecha de publicación de la solicitud: **14.02.2007**

54 Título: **CLON GENÓMICO DE NICOTINA DESMETILASA DE TABACO Y USOS DEL MISMO.**

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
19.04.2005 US 110062

45 Fecha de publicación de la mención BOPI:
09.02.2012

45 Fecha de la publicación del folleto de la patente:
09.02.2012

73 Titular/es:
U.S. Smokeless Tobacco Company LLC
6603 West Broad Street
Richmond, VA 23230, US

72 Inventor/es:
XU, Dongmei

74 Agente: **Carvajal y Urquijo, Isabel**

ES 2 373 881 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

Clon genómico de nicotina desmetilasa de tabaco y usos del mismo

La presente invención se refiere a secuencias de ácido nucleico que codifican para nicotina desmetilasa y a métodos para usar esas secuencias de ácido nucleico para alterar fenotipos vegetales.

5 Antecedentes

El citocromo p450 cataliza reacciones enzimáticas para una gama diversa de sustratos químicamente distintos que incluyen el metabolismo oxidativo, peroxidativo y reductor de sustratos xenobióticos y endógenos. En plantas, p450 participa en rutas bioquímicas que incluyen la síntesis de productos vegetales tales como fenilpropanoides, alcaloides, terpenoides, lípidos, glicósidos cianogénicos y glucosilatos (Chappell, Annu. Rev. Plant Physiol. Plant Mol. Biol. 46: 521-547, 1995). El citocromo p450s, también conocido como proteínas hemo-tiolato p450, actúa habitualmente como oxidasas terminales en cadenas de transferencia de electrones de múltiples componentes, denominadas sistemas de monooxigenasas que contienen p450. Las reacciones específicas catalizadas por estos sistemas enzimáticos incluyen desmetilación, hidroxilación, epoxidación, N-oxidación, sulfooxidación, N-, S- y O-desalquilaciones, desulfatación, desaminación y reducción de grupos N-óxido, nitro y azo.

El papel diverso de las enzimas p450 de la planta *Nicotiana* se ha implicado en la generación de una variedad de metabolitos vegetales tales como fenilpropanoides, alcaloides, terpenoides, lípidos, glicósidos cianogénicos, glucosilatos y una gran cantidad de otras entidades químicas. Algunas enzimas p450 pueden tener un impacto sobre la composición de metabolitos vegetales. Por ejemplo, se ha deseado durante mucho tiempo mejorar el sabor y aroma de ciertas plantas alterando el perfil de una planta de ácidos grasos seleccionados a través de cultivo; sin embargo, se conoce muy poco sobre los mecanismos implicados en el control de los niveles de estos constituyentes de la hoja. La regulación por disminución o regulación por incremento de enzimas p450 asociadas con la modificación de los ácidos grasos puede facilitar la acumulación de ácidos grasos deseados que proporcionan cualidades fenotípicas más preferidas de la hoja.

La función de las enzimas p450 y sus papeles crecientes en constituyentes vegetales están descubriéndose aún. Por ejemplo, se encontró que una clase especial de enzimas p450 catalizaba la descomposición de ácido graso en β -alcoholes y aldehídos C6 y C9 volátiles que son contribuyentes principales del olor "verde fresco" de frutas y verduras. El nivel de otras p450 seleccionadas como diana novedosas puede alterarse para potenciar las cualidades de los constituyentes de la hoja modificando la composición lipídica y metabolitos de descomposición relacionados en la hoja de *Nicotiana*. Varios de estos constituyentes en la hoja se ven afectados por la senescencia que estimula la maduración de propiedades de calidad de la hoja. Todavía otros informes han mostrado que las enzimas p450 desempeñan un papel funcional en la alteración de ácidos grasos que están implicados en la resistencia a enfermedad e interacciones planta-patógeno.

En otros ejemplos, se ha sugerido que las enzimas p450 están implicadas en la biosíntesis de alcaloides. Tal como se proporciona en las solicitudes de patente por el solicitante, de las que la presente solicitud reivindica prioridad, la nornicotina, un alcaloide minoritario encontrado en *Nicotiana tabacum*, se produce mediante la desmetilación mediada por p450 de nicotina seguida por acilación y nitrosación en la posición N produciendo de ese modo una serie de N-acilnornicotinas y N-nitrosnornicotinas. Se cree que la N-desmetilación, catalizada por una p450 desmetilasa, es una fuente primaria de biosíntesis de nornicotina en *Nicotiana*.

Existe una necesidad en la técnica de reactivos y métodos para modificar fenotipos vegetales. En particular, existe una necesidad de reactivos y métodos para modificar la nicotina desmetilasa. La presente invención proporciona varias estrategias para modificar la expresión de una nicotina desmetilasa.

Sumario de la invención

Los presentes inventores han identificado y caracterizado un clon genómico de la nicotina desmetilasa a partir de tabaco. Se incluyen en el presente documento secuencias para la región codificante de la proteína nicotina desmetilasa, la región no traducida en 3' ("UTR en 3'"), un intrón individual y el promotor del gen de nicotina desmetilasa junto con sus secuencias reguladoras de la transcripción (figura 1). Se describe además el uso de estas secuencias para crear plantas transgénicas que tienen niveles alterados de nornicotina o N'-nitrosnornicotina ("NNN") o ambos con respecto a una planta control.

Por consiguiente, la invención describe una molécula de ácido nucleico aislada, una planta de tabaco transgénica, un componente vegetal de la planta de tabaco, un producto de tabaco, una semilla de la planta de tabaco, una nicotina desmetilasa de tabaco recombinante y diversos métodos que implican lo mismo, tal como se define en las reivindicaciones.

Por consiguiente, en el primer aspecto, la descripción describe una molécula de ácido nucleico aislada, por ejemplo, una secuencia de ADN, que contiene una secuencia de nucleótidos que codifica para una nicotina desmetilasa. En realizaciones deseables, la secuencia de nucleótidos del primer aspecto es sustancialmente idéntica a una secuencia de nucleótidos que codifica para una nicotina desmetilasa de tabaco, tal como una nicotina desmetilasa de tabaco que contiene una secuencia de nucleótidos que es al menos el 70% idéntica a la secuencia de nucleótidos de SEQ ID NO:1 o SEQ ID NO:3, o que contiene los nucleótidos 2010-2949 y/o 3947-4562 de SEQ ID NO: 1, o que contiene la secuencia de SEQ ID NO:1 o SEQ ID NO:3. La molécula de ácido nucleico aislada del primer aspecto de la descripción, por ejemplo, está operativamente unida a un promotor funcional en una célula vegetal y deseablemente está contenida en un vector de expresión. En otras realizaciones deseables, el vector de expresión está contenido en una célula, por ejemplo, una célula vegetal. Deseablemente, la célula vegetal, tal como una célula vegetal de tabaco, está incluida en una planta. En otra realización deseable, la descripción describe una semilla, por ejemplo, una semilla de tabaco, a partir de una planta que contiene el vector de expresión, en la que la semilla incluye una molécula de ácido nucleico aislada que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:3 operativamente unida a una secuencia promotora heteróloga. Además, la descripción describe una planta derivada de una semilla germinada que contiene el vector de expresión, una hoja, o bien verde o bien curada, de la planta, y un artículo de fabricación preparado a partir de la hoja.

En otra realización deseable, la secuencia de nucleótidos contiene una secuencia que hibrida en condiciones rigurosas con el complemento de la secuencia de nucleótidos de SEQ ID NO:1 y/o SEQ ID NO:3, o a un fragmento de SEQ ID NO:1 o SEQ ID NO:3. Deseablemente, la secuencia de nucleótidos codifica para una nicotina desmetilasa que es sustancialmente idéntica a la secuencia de aminoácidos de SEQ ID NO:2. En una realización deseable adicional del primer aspecto de la invención, la nicotina desmetilasa tiene una identidad de secuencia de aminoácidos de al menos el 70% con la secuencia de aminoácidos de la nicotina desmetilasa de SEQ ID NO:2 o con un fragmento de una nicotina desmetilasa que tiene actividad enzimática alterada (por ejemplo, reducida) en comparación con el polipéptido de longitud completa. Deseablemente, la nicotina desmetilasa incluye la secuencia de aminoácidos de SEQ ID NO:2.

En otro aspecto, la descripción describe una molécula de ácido nucleico aislada que contiene un promotor que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:6, o un fragmento de la misma que dirige la transcripción. Deseablemente, el promotor (i) se induce tras tratamiento con etileno o durante la senescencia; y (ii) incluye (a) los pares de bases 1-2009 de SEQ ID NO:1, 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 1-2009 de SEQ ID NO:1, o (c) una parte de nucleótidos de 20 pares de bases idéntica en secuencia a una parte de 20 pares de bases consecutivos de la secuencia expuesta en los pares de bases 1-2009 de SEQ ID NO:1.

Un aspecto adicional de la descripción describe un promotor de ácido nucleico aislado que contiene una secuencia de nucleótidos que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:6. Deseablemente, este promotor de ácido nucleico aislado se induce tras tratamiento con etileno o durante la senescencia y, por ejemplo, incluye la secuencia de SEQ ID NO:6. Alternativamente, el promotor puede incluir un fragmento que puede obtenerse a partir de SEQ ID NO:6, en el que el fragmento dirige la transcripción de un gen heterólogo o reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica). En una realización deseable la secuencia promotora está operativamente unida 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. Deseablemente, la célula vegetal, tal como una célula vegetal de tabaco, está incluida en una planta. En otra realización deseable, la descripción describe una semilla, por ejemplo, una semilla de tabaco, de una planta que contiene el vector de expresión, en la que la semilla incluye una molécula de ácido nucleico aislada que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:6 operativamente unida a una secuencia de ácido nucleico heteróloga. Además, la descripción describe una planta derivada de una semilla germinada que contiene el promotor de este aspecto de la descripción, una hoja, o bien verde o bien curada, de la planta, y un artículo de fabricación preparado a partir de la hoja.

Otro aspecto de la descripción describe un método de expresión de un gen heterólogo en una planta. Este método implica (i) introducir en una célula vegetal un vector que contiene una secuencia promotora tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:6 operativamente unida a una secuencia de ácido nucleico heteróloga; y (ii) regenerar una planta a partir de la célula. Además, este método puede implicar la transmisión sexual del vector a la progenie y, además, puede incluir la etapa de recoger la semilla producida por la progenie.

Aún en otro aspecto, la descripción describe un método de reducción de la expresión de 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:6 o un fragmento que puede obtenerse a partir de SEQ ID NO:6 operativamente unido a una secuencia de ácido nucleico heteróloga y (ii) expresar el vector en la planta de tabaco. En una realización deseable de este método, se silencia la expresión de la nicotina desmetilasa. En otra realización deseable, el vector expresa ARN, tal como ARN antisentido o una molécula de ARN que puede inducir interferencia por ARN (ARNi).

En un aspecto deseable adicional, la descripción describe una molécula de ácido nucleico aislada que contiene un intrón que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:5, o un fragmento de la misma que reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica) o puede servir como marcador molecular para identificar secuencias de ácido nucleico de la nicotina desmetilasa. En una realización deseable, el intrón incluye (a) los pares de bases 2950-3946 de SEQ ID NO:1, o (b) al menos 200 pares de bases consecutivos idénticos a los 200 pares de bases consecutivos de la secuencia definida por los pares de bases 2950-3946 de SEQ ID NO:1, o (c) una parte de nucleótidos de 20 pares de bases idéntica en secuencia a una parte de 20 pares de bases consecutivos de la secuencia expuesta en los pares de bases 2950-3946 de SEQ ID NO:1.

Otro aspecto deseable de la descripción describe un intrón de ácido nucleico aislado que incluye una secuencia de nucleótidos que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:5, o un fragmento de la misma que reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica) o puede servir como marcador molecular para identificar secuencias de ácido nucleico de la nicotina desmetilasa. El silenciamiento de la expresión génica puede implicar, por ejemplo, recombinación homóloga o una mutación que da como resultado un producto génico que no tiene actividad nicotina desmetilasa. En particular, el intrón puede incluir la secuencia de SEQ ID NO:5 o un fragmento que puede obtenerse a partir de SEQ ID NO:5. Deseablemente, una molécula de ácido nucleico aislada que incluye un intrón está operativamente unida a una secuencia de ácido nucleico 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 puede ser una célula de tabaco. Una planta, por ejemplo, una planta de tabaco, que incluye una célula vegetal que contiene la secuencia de SEQ ID NO:5 o un fragmento que puede obtenerse a partir de SEQ ID NO:5 operativamente unido a una secuencia de ácido nucleico heteróloga en un vector de expresión es otra realización deseable de la presente invención. Además, una semilla, por ejemplo, una semilla de tabaco, de una planta, en la que la semilla contiene un intrón que hibrida en condiciones rigurosas con SEQ ID NO:5 operativamente unido a una secuencia de ácido nucleico heteróloga también es deseable. Además, la descripción describe una planta derivada de la semilla germinada que contiene el intrón de este aspecto de la invención, una hoja, o bien verde o bien curada, de la planta, y un artículo de fabricación preparado a partir de la hoja verde o curada.

Un aspecto adicional de la descripción describe un método de expresión de un intrón en una planta. Este método implica (i) introducir en una célula vegetal un vector de expresión que contiene la secuencia de SEQ ID NO:5 o un fragmento que puede obtenerse a partir de SEQ ID NO:5 operativamente unido a una secuencia de ácido nucleico heteróloga; y (ii) regenerar una planta a partir de la célula. En una realización deseable, este método también implica (iii) transmitir sexualmente el vector a la progenie, y puede incluir la etapa adicional de recolectar la semilla producida por la progenie. El método incluye deseablemente, por ejemplo, regenerar una planta a partir de la semilla germinada, una hoja, o bien verde o bien curada, de la planta, y un método de preparación de un artículo de fabricación a partir de la hoja.

Aún en otro aspecto, la descripción describe un método de reducción de la expresión de 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:5 o un fragmento que puede obtenerse a partir de SEQ ID NO:5 operativamente unido a una secuencia de ácido nucleico heteróloga y (ii) expresar el vector en la planta de tabaco. En una realización deseable de este método, se silencia la expresión de la nicotina desmetilasa. En otra realización deseable, el vector expresa ARN, tal como ARN antisentido o una molécula de ARN que puede inducir interferencia por ARN (ARNi).

En un aspecto adicional, la descripción describe una molécula de ácido nucleico aislada que contiene una región no traducida que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:7 o un fragmento de la misma que puede alterar el patrón de expresión de un gen, reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica), o puede usarse como marcador para identificar secuencias de ácido nucleico de la nicotina desmetilasa. En una realización deseable de este aspecto de la descripción, la región no traducida incluye (a) los pares de bases 4563-6347 de SEQ ID NO:1, o (b) al menos 200 pares de bases consecutivos idénticos a los 200 pares de bases consecutivos de la secuencia definida por los pares de bases 4563-6347 de SEQ ID NO:1, o (c) una parte de nucleótidos de 20 pares de bases idéntica en secuencia a una parte de 20 pares de bases consecutivos de la secuencia expuesta en los pares de bases 4563-6347 de SEQ ID NO:1.

Un aspecto deseable adicional de la descripción describe una región no traducida de ácido nucleico aislada que contiene una secuencia de nucleótidos que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:7. Deseablemente, la región no traducida incluye la secuencia de SEQ ID NO:7 o la región no traducida incluye un fragmento que puede obtenerse a partir de SEQ ID NO:7 que puede alterar el patrón de expresión de un gen, reduce o altera la actividad enzimática de la nicotina desmetilasa (por ejemplo, silencia la expresión génica), o puede usarse como marcador para identificar secuencias de ácido nucleico de la nicotina desmetilasa. La región no traducida deseablemente está operativamente unida a una secuencia de ácido nucleico heteróloga y puede estar contenida en un vector de expresión. Además, este vector de expresión está contenido deseablemente en una célula, tal como una célula vegetal, por ejemplo, una célula de tabaco. Otra realización deseable de la descripción describe una planta, tal como una planta de tabaco, que incluye una célula vegetal que contiene un vector que

incluye una secuencia de ácido nucleico aislada que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:7 y está operativamente unida a una secuencia de ácido nucleico heteróloga.

5 La descripción también describe una semilla, por ejemplo, una semilla de tabaco, de una planta, en la que la semilla incluye una región no traducida que hibrida en condiciones rigurosas con SEQ ID NO:7 operativamente unida a una secuencia de ácido nucleico heteróloga. Además, la descripción describe una planta derivada de una semilla germinada que contiene la región no traducida de este aspecto de la invención, una hoja, o bien verde o bien curada, de la planta, y un artículo de fabricación preparado a partir de la hoja verde o curada.

10 En un aspecto adicional, la descripción describe un método de expresión de una región no traducida en una planta. Este método implica (i) introducir en una célula vegetal un vector que contiene una secuencia de ácido nucleico aislada que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO:7 y está operativamente unida a una secuencia de ácido nucleico heteróloga; y (ii) regenerar una planta a partir de la célula. Además, este método también puede implicar (iii) transmitir sexualmente el vector a la progenie, y deseablemente, incluye la etapa adicional de recoger la semilla producida por la progenie. El método incluye deseablemente regenerar una planta a partir de la semilla germinada, una hoja, o bien verde o bien curada, de la planta, y un método de preparación de un artículo de fabricación preparado a partir de la hoja verde o curada.

20 Además, la descripción describe un método de reducción de la expresión o alteración de la actividad enzimática de nicotina desmetilasa en una planta de tabaco. Este método incluye las etapas de (i) introducir en la planta de tabaco un vector que contiene una secuencia de ácido nucleico aislada que tiene una identidad de secuencia del 50% o más con la secuencia de SEQ ID NO: 7 y está operativamente unida a una secuencia de ácido nucleico heteróloga y (ii) expresar el vector en la planta de tabaco. Deseablemente, se silencia la expresión de la nicotina desmetilasa. En otras realizaciones deseables el vector expresa ARN, por ejemplo, ARN antisentido o una molécula de ARN que puede inducir interferencia por ARN (ARNi).

25 Otro aspecto de la descripción describe un vector de expresión que incluye una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica para una nicotina desmetilasa, en la que el vector puede dirigir la expresión de la nicotina desmetilasa codificada por la molécula de ácido nucleico aislada. Deseablemente, el vector incluye la secuencia de SEQ ID NO: 1 o SEQ ID NO:3. En otras realizaciones deseables, la descripción describe una planta o componente vegetal, por ejemplo, una planta de tabaco o componente vegetal (por ejemplo, un tallo u hoja de tabaco), que incluye una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica para un polipéptido que desmetila la nicotina.

30 Un aspecto adicional de la descripción describe una célula que contiene una molécula de ácido nucleico aislada que incluye una secuencia de nucleótidos que codifica para una nicotina desmetilasa. Deseablemente esta célula es una célula vegetal o una célula bacteriana, tal como un *Agrobacterium*.

35 Otro aspecto de la descripción describe una planta o componente vegetal (por ejemplo, un tallo u hoja de tabaco) que contiene una molécula de ácido nucleico aislada que codifica para una nicotina desmetilasa, en la que la molécula de ácido nucleico se expresa en la planta o el componente vegetal. Deseablemente, la planta o componente vegetal es un angiosperma, una dicotiledónea, una planta solanácea o una especie de *Nicotiana*. Otras realizaciones deseables de este aspecto son una semilla o una célula de la planta o componente vegetal, así como una hoja, o bien verde o bien curada, derivada de la planta y un artículo de fabricación preparado a partir de la misma.

40 En un aspecto adicional, la descripción describe una planta de tabaco que tiene expresión reducida de una secuencia de ácido nucleico que codifica para un polipéptido, por ejemplo, uno que incluye la secuencia de SEQ ID NO:2, y que desmetila la nicotina, en la que la expresión reducida (o una reducción en actividad enzimática) reduce el nivel de nor nicotina en la planta. En una realización deseable, la planta de tabaco es una planta transgénica, tal como una que incluye un transgén que, cuando se expresa en la planta transgénica, silencia la expresión génica de una nicotina desmetilasa de tabaco endógena.

45 En particular, la planta transgénica incluye deseablemente uno o más de los siguientes: un transgén que expresa una molécula antisentido de una nicotina desmetilasa de tabaco o una molécula de ARN que puede inducir interferencia por ARN (ARNi); un transgén que, cuando se expresa en la planta transgénica, suprime conjuntamente la expresión de una nicotina desmetilasa de tabaco; un transgén que codifica para un producto génico dominante negativo, por ejemplo, una forma mutada de la secuencia de aminoácidos de SEQ ID NO:2; una mutación puntual en un gen que codifica para la secuencia de aminoácidos de SEQ ID NO:2; una delección en un gen que codifica para una nicotina desmetilasa de tabaco; y una inserción en un gen que codifica para una nicotina desmetilasa de tabaco.

En otras realizaciones deseables, la expresión reducida de una secuencia de ácido nucleico que codifica para un polipéptido se produce al nivel transcripcional, al nivel traduccional o al nivel postraduccional.

Otro aspecto de la descripción describe una planta de tabaco que contiene un casete de expresión recombinante integrado de manera estable en el genoma de la misma, en la que el casete puede efectuar una reducción en la actividad nicotina desmetilasa. Las semillas de esta planta de tabaco se describen en una realización deseable. Otras realizaciones deseables incluyen una hoja, o bien verde o bien curada, derivada de esta planta y un artículo de fabricación preparado a partir de la misma.

Un aspecto adicional de la descripción describe un método de expresión de una nicotina desmetilasa de tabaco en una planta. Este método implica (i) introducir en una célula vegetal un vector de expresión que incluye una molécula de ácido nucleico que contiene una secuencia de nucleótidos que codifica para una nicotina desmetilasa; y (ii) regenerar una planta a partir de la célula. En una realización deseable, este método describe la transmisión sexual del vector a la progenie, y deseablemente también incluye la etapa adicional de recoger la semilla producida por la progenie. Las realizaciones deseables adicionales incluyen una planta derivada de la semilla germinada, una hoja, o bien verde o bien curada, de la planta, y un artículo de fabricación preparado a partir de la hoja verde o curada.

Un aspecto adicional de la descripción describe una nicotina desmetilasa de tabaco sustancialmente pura. Deseablemente, esta nicotina desmetilasa de tabaco incluye una secuencia de aminoácidos que tiene una identidad de al menos el 70% con la secuencia de aminoácidos de SEQ ID NO:2 o incluye la secuencia de aminoácidos de SEQ ID NO:2. En una realización deseable, la nicotina desmetilasa de tabaco, tras su expresión en una célula vegetal, convierte la nicotina en nornicotina. En otras realizaciones deseables, la nicotina desmetilasa de tabaco, tras su expresión en una célula vegetal, se localiza predominantemente en las hojas, o la nicotina desmetilasa de tabaco se induce mediante etileno o se expresa durante la senescencia de la planta.

En un aspecto adicional, la descripción describe un anticuerpo sustancialmente puro que reconoce específicamente y se une a una nicotina desmetilasa de tabaco. Deseablemente, el anticuerpo reconoce y se une a una nicotina desmetilasa de tabaco recombinante, por ejemplo, una que contiene la secuencia de SEQ ID NO:2 o un fragmento de la misma.

Otro aspecto de la descripción describe un método de producción de una nicotina desmetilasa de tabaco. Este método implica las etapas de: (a) proporcionar una célula transformada con una molécula de ácido nucleico aislada que contiene una secuencia de nucleótidos que codifica para un polipéptido que desmetila la nicotina; (b) cultivar la célula transformada en condiciones para expresar la molécula de ácido nucleico aislada; y (c) recuperar la nicotina desmetilasa de tabaco. La descripción también describe una nicotina desmetilasa de tabaco recombinante producida según este método.

En un aspecto adicional, la descripción describe un método de aislamiento de una nicotina desmetilasa de tabaco o fragmento de la misma. Este método implica las etapas de: (a) poner en contacto la molécula de ácido nucleico de SEQ ID NOS:1, 3, 5, 6 ó 7 o una parte de la misma con una preparación de ácido nucleico a partir de una célula vegetal en condiciones de hibridación que proporcionan la detección de secuencias de ácido nucleico que tienen una identidad de secuencia de al menos el 70% o superior con la secuencia de ácido nucleico de SEQ ID NOS: 1, 3, 5, 6 ó 7; y (b) aislar las secuencias hibridantes de ácido nucleico.

En un aspecto adicional, la descripción describe otro método de aislamiento de una nicotina desmetilasa de tabaco o fragmento de la misma. Este método incluye las etapas de: (a) proporcionar una muestra de ADN de célula vegetal; (b) proporcionar un par de oligonucleótidos que tienen identidad de secuencia con una región de una molécula de ácido nucleico que tiene la secuencia de SEQ ID NOS: 1, 3, 5, 6 ó 7; (c) poner en contacto el par de oligonucleótidos con el ADN de célula vegetal en condiciones adecuadas para la amplificación de ADN mediada por la reacción en cadena de la polimerasa; y (d) aislar la nicotina desmetilasa de tabaco amplificada o fragmento de la misma. En una realización deseable de este aspecto, la etapa de amplificación se lleva a cabo usando una muestra de ADNc preparada a partir de una célula vegetal. En otra realización deseable, la nicotina desmetilasa de tabaco codifica para un polipéptido que es al menos el 70% idéntico a la secuencia de aminoácidos de SEQ ID NO:2.

Un aspecto adicional de la descripción describe un método para reducir la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. Este método implica las etapas de: (a) introducir en células vegetales un transgén que codifica para una nicotina desmetilasa de tabaco operativamente unido a un promotor funcional en las células vegetales para proporcionar células vegetales transformadas; y (b) regenerar una planta o componente vegetal a partir de las células vegetales transformadas, en las que la nicotina desmetilasa de tabaco se expresa en las células de la planta o componente vegetal, reduciendo de ese modo la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. En realizaciones particulares de este aspecto de la descripción, el transgén que codifica para la nicotina desmetilasa de tabaco se expresa de manera constitutiva o se expresa de manera inducible, por ejemplo, de una manera específica de tejido, específica de célula o específica de órgano. En otra realización de este aspecto de la descripción, la expresión del transgén suprime conjuntamente la expresión de una nicotina desmetilasa de tabaco endógena.

Un aspecto adicional de la descripción describe otro método para reducir la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. Este método incluye las etapas de: (a) introducir en células vegetales

un transgén que codifica para una secuencia codificante antisentido de una nicotina desmetilasa de tabaco o una molécula de ARN que puede inducir la interferencia por ARN (ARNi) operativamente unido a un promotor funcional en las células vegetales para proporcionar células vegetales transformadas; y (b) regenerar una planta o componente vegetal a partir de las células vegetales transformadas, en las que el antisentido o una molécula de ARN que puede inducir la interferencia por ARN (ARNi) de la secuencia codificante de la nicotina desmetilasa de tabaco se expresa en las células de la planta o componente vegetal, reduciendo de ese modo la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. Deseablemente, el transgén que codifica para una secuencia antisentido o una molécula de ARN que puede inducir la interferencia por ARN (ARNi) de una nicotina desmetilasa de tabaco se expresa de manera constitutiva o se expresa de manera inducible, por ejemplo de una manera específica de tejido, específica de célula o específica de órgano.

Un aspecto adicional de la descripción describe aún otro método para reducir la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. Este método implica las etapas de: (a) introducir en células vegetales un transgén que codifica para un producto génico dominante negativo de una nicotina desmetilasa de tabaco operativamente unido a un promotor funcional en las células vegetales para proporcionar células vegetales transformadas; y (b) regenerar una planta o componente vegetal a partir de las células vegetales transformadas, en las que el producto génico dominante negativo de la nicotina desmetilasa de tabaco se expresa en las células de la planta o componente vegetal, reduciendo de ese modo la expresión de la nicotina desmetilasa de tabaco en una planta o componente vegetal. En realizaciones particulares de este aspecto de la descripción, el transgén que codifica para el producto génico dominante negativo se expresa de manera constitutiva o se expresa de manera inducible, por ejemplo, de una manera específica de tejido, específica de célula o específica de órgano.

Un aspecto adicional de la descripción describe un método adicional para reducir la expresión o la actividad enzimática de la nicotina desmetilasa de tabaco en una célula vegetal. Este método implica reducir el nivel de una nicotina desmetilasa de tabaco endógena, o su actividad enzimática, en la célula vegetal. Deseablemente, la célula vegetal es de una dicotiledónea, una planta solanácea o una especie de *Nicotiana*. En realizaciones deseables de este aspecto, reducir el nivel de la nicotina desmetilasa de tabaco endógena implica expresar un transgén que codifica para una molécula de ácido nucleico antisentido o una molécula de ARN que puede inducir la interferencia por ARN (ARNi) de una nicotina desmetilasa de tabaco en la célula vegetal, o implica expresar un transgén que codifica para una molécula de ARN bicatenario de una nicotina desmetilasa de tabaco en la célula vegetal. Deseablemente, el ARN bicatenario es una secuencia de ARN que corresponde a la secuencia de SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:6, o SEQ ID NO:7, o un fragmento de la misma. En una realización adicional, reducir el nivel de la nicotina desmetilasa de tabaco endógena implica la cosupresión de la nicotina desmetilasa de tabaco endógena en la célula vegetal o implica expresar un producto génico dominante negativo en la célula vegetal. En particular, el producto génico dominante negativo puede incluir un gen que codifica para una forma mutada de la secuencia de aminoácidos de SEQ ID NO:2.

En otras realizaciones deseables de este aspecto de la descripción, la nicotina desmetilasa de tabaco endógena incluye una mutación puntual en un gen que codifica para la secuencia de aminoácidos de SEQ ID NO:2. En otras realizaciones deseables, reducir el nivel de expresión de una nicotina desmetilasa de tabaco endógena implica una delección en un gen que codifica para una nicotina desmetilasa de tabaco o implica una inserción en un gen que codifica para una nicotina desmetilasa de tabaco. La reducción de la expresión puede producirse al nivel transcripcional, al nivel de traduccional o al nivel postraduccional.

Un aspecto adicional de la descripción describe un método para identificar un compuesto que altera la expresión de una nicotina desmetilasa de tabaco en una célula. Este método implica las etapas de: (a) proporcionar una célula que contiene un gen que codifica para una nicotina desmetilasa de tabaco; (b) aplicar un compuesto candidato a la célula; y (c) medir la expresión del gen que codifica para la nicotina desmetilasa de tabaco, en el que un aumento o una disminución en la expresión con respecto a una muestra control sin tratar es una indicación de que el compuesto altera la expresión de la nicotina desmetilasa de tabaco.

En una realización deseable de este método, el gen de la parte (a) codifica para una nicotina desmetilasa de tabaco que tiene una identidad de al menos el 70% con la secuencia de aminoácidos de SEQ ID NO:2. Deseablemente, el compuesto disminuye o aumenta la expresión del gen que codifica para la nicotina desmetilasa de tabaco.

En otro aspecto, la descripción describe otro método para identificar un compuesto que altera la actividad de una nicotina desmetilasa de tabaco en una célula. Este método implica las etapas de: (a) proporcionar una célula que expresa un gen que codifica para una nicotina desmetilasa de tabaco; (b) aplicar un compuesto candidato a la célula; y (c) medir la actividad de la nicotina desmetilasa de tabaco, en el que un aumento o una disminución en la actividad con respecto a una muestra control sin tratar es una indicación de que el compuesto altera la actividad de la nicotina desmetilasa de tabaco. En una realización deseable de este aspecto de la descripción, el gen de la etapa (a) codifica para una nicotina desmetilasa de tabaco que tiene una identidad de al menos el 70% con la secuencia de aminoácidos de SEQ ID NO:2. Deseablemente, el compuesto disminuye o aumenta la actividad de la nicotina desmetilasa de tabaco.

Un aspecto adicional de la descripción describe una planta de tabaco o componente vegetal curado 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. Deseablemente, el componente vegetal es una hoja de tabaco o tallo de tabaco. En una realización deseable, la nitrosamina es nornicotina, y el contenido de nornicotina deseablemente es inferior a 5 mg/g, 4,5 mg/g, 4,0 mg/g, 3,5 mg/g, 3,0 mg/g, más deseablemente inferior a 2,5 mg/g, 2,0 mg/g, 1,5 mg/g, 1,0 mg/g, más deseablemente inferior a 750 µg/g, 500 µg/g, 250 µg/g, 100 µg/g, incluso más deseablemente inferior a 75 µg/g, 50 µg/g, 25 µg/g, 10 µg/g, 7,0 µg/g, 5,0 µg/g, 4,0 µg/g, e incluso más deseablemente inferior a 2,0 µg/g, 1,0 µg/g, 0,5 µg/g, 0,4 µg/g, 0,2 µg/g, 0,1 µg/g, 0,05 µg/g, o 0,01 µg/g o en el que el porcentaje de alcaloides secundarios con respecto al contenido en alcaloides total en el mismo es inferior al 90%, el 70%, el 50%, el 30%, el 10%, deseablemente inferior al 5%, el 4%, el 3%, el 2%, el 1,5%, el 1% y más deseablemente inferior al 0,75%, el 0,5%, el 0,25% o el 0,1%. En otra realización deseable, la nitrosamina es N'-nitrosonornicotina (NNN), y el contenido de N'-NNN deseablemente es inferior a 5 mg/g, 4,5 mg/g, 4,0 mg/g, 3,5 mg/g, 3,0 mg/g, más deseablemente inferior a 2,5 mg/g, 2,0 mg/g, 1,5 mg/g, 1,0 mg/g, más deseablemente inferior a 750 µg/g, 500 µg/g, 250 µg/g, 100 µg/g, incluso más deseablemente inferior a 75 µg/g, 50 µg/g, 25 µg/g, 10 µg/g, 7,0 µg/g, 5,0 µg/g, 4,0 µg/g, e incluso más deseablemente inferior a 2,0 µg/g, 1,0 µg/g, 0,5 µg/g, 0,4 µg/g, 0,2 µg/g, 0,1 µg/g, 0,05 µg/g o 0,01 µg/g o en el que el porcentaje de alcaloides secundarios con respecto al contenido en alcaloides total contenido en el mismo es inferior al 90%, el 70%, el 50%, el 30%, el 10%, deseablemente inferior al 5%, el 4%, el 3%, el 2%, el 1,5%, el 1% y más deseablemente inferior al 0,75%, el 0,5%, el 0,25% o el 0,1%. En realizaciones deseables adicionales de este aspecto de la descripción, la planta de tabaco o componente vegetal curado es un tabaco oscuro, tabaco Burley, tabaco curado en chimenea, tabaco curado al aire o tabaco oriental.

Además, la planta de tabaco o componente vegetal curado de la descripción incluye deseablemente un gen de nicotina desmetilasa recombinante, por ejemplo, uno que contiene la secuencia de SEQ ID NO:1 o SEQ ID NO:3, o un fragmento de la misma. Deseablemente, se silencia la expresión de un gen de nicotina desmetilasa endógena en la planta de tabaco o componente vegetal curado.

Otro aspecto de la descripción describe un producto de tabaco que contiene una planta de tabaco o componente vegetal curado que incluye (i) expresión reducida de una nicotina desmetilasa o (ii) una nicotina desmetilasa que tiene actividad alterada, y una cantidad reducida de una nitrosamina. Deseablemente, el producto de tabaco es tabaco sin humo, rapé seco o húmedo, un tabaco de mascar, cigarrillo, cigarro, puro delgado, tabaco de pipa o bidis. En particular, el producto de tabaco de este aspecto de la descripción puede contener tabaco oscuro, tabaco molido o incluye un componente aromatizante.

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

Definiciones

"Actividad enzimática" pretende incluir pero sin limitarse a desmetilación, hidroxilación, epoxidación, N-oxidación, sulfooxidación, N-, S- y O-desalquilaciones, desulfuración, desaminación y reducción de azo, nitro, N-óxido y otros grupos químicos enzimáticamente reactivos de este tipo. Actividad enzimática alterada se refiere a una disminución en la actividad enzimática (por ejemplo, de una nicotina desmetilasa de tabaco) en al menos el 10-20%, preferiblemente en al menos el 25-50% y más preferiblemente en al menos el 55-95% o superior con respecto a la actividad de una enzima control (por ejemplo, una nicotina desmetilasa de una planta de tabaco de tipo natural). La actividad de una nicotina desmetilasa de tabaco puede someterse a ensayo usando métodos convencionales en la técnica, tal como midiendo la desmetilación de nicotina radioactiva mediante microsomas expresados en levadura, tal como se describe en el presente documento.

La expresión "ácido nucleico" se refiere a un polímero de desoxirribonucleótidos o ribonucleótidos en forma o bien mono o bien bicatenaria, o sentido o antisentido, y a menos que se limite de otra forma, abarca análogos conocidos de nucleótidos naturales que se hibridan con ácido nucleicos de una manera similar a nucleótidos que se producen de manera natural. A menos que se indique de otra forma, una secuencia de ácido nucleico particular incluye la secuencia complementaria de la misma. Las expresiones "operativamente unido", "en combinación operativa" y "en orden operativo" se refieren al enlace funcional entre una secuencia de control de la expresión de ácidos nucleicos (tal como un promotor, secuencia señal o alineamiento de sitios de unión de factores de transcripción) y una segunda secuencia de ácido nucleico, en la que la secuencia de control de la expresión afecta a la transcripción y/o traducción del ácido nucleico que corresponde a la segunda secuencia.

El término "recombinante" cuando se usa con referencia a una célula indica que la célula replica un ácido nucleico

heterólogo, expresa el ácido nucleico o expresa un péptido, péptido heterólogo o proteína codificada por un ácido nucleico heterólogo. Las células recombinantes pueden expresar genes o fragmentos génicos en forma o bien sentido o bien antisentido o una molécula de ARN que puede inducir interferencia por ARN (ARNi) que no se encuentra dentro 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 la que los genes se modifican y se reintroducen en la célula por medios artificiales.

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

El gen estructural puede constituir una secuencia codificante ininterrumpida o puede incluir uno o más intrones, unidos mediante las uniones de corte y empalme apropiadas. El gen estructural puede ser traducible o no traducible, incluyendo un antisentido o una molécula de ARN que puede inducir interferencia por ARN (ARNi). El gen estructural puede ser una amalgama de segmentos derivados de una pluralidad de fuentes y de una pluralidad de secuencias génicas (que se producen de manera natural o sintética, en las que sintética se refiere a ADN que se sintetiza químicamente).

Un “exón” tal como se usa en el presente documento en referencia a una secuencia de ácido nucleico se refiere a una parte de la secuencia de ácido nucleico de un gen, en la que la secuencia de ácido nucleico del exón codifica para al menos un aminoácido del producto génico. Un exón es normalmente adyacente a un segmento de ADN no codificante tal como un intrón. Deseablemente, un exón codifica para una parte de la secuencia de aminoácidos de la nicotina desmetilasa de tabaco de SEQ ID NO:2, tal como los aminoácidos 1-313 y/o 314-517 de la secuencia de SEQ ID NO:2.

Un “intrón” tal como se usa en el presente documento en referencia a una secuencia de ácido nucleico se refiere a una región no codificante de un gen que está flanqueada por regiones codificantes. Un intrón es normalmente una región no codificante de un gen que se transcribe en una molécula de ARN pero luego se corta mediante corte y empalme de ARN durante la producción del ARN mensajero u otro ARN estructural funcional. Deseablemente, un intrón incluye la secuencia de SEQ ID NO:5, o un fragmento de la misma.

Una “UTR en 3'” tal como se usa en el presente documento en referencia a una secuencia de ácido nucleico se refiere a una secuencia de ácido nucleico no codificante proximal a un codón de terminación de un exón. Deseablemente, una UTR en 3' incluye la secuencia de SEQ ID NO:7, o un fragmento de la misma.

“Derivado de” se usa para referirse a tomado, obtenido, recibido, indicado, replicado o descendiente de una fuente (química y/o biológica). Un derivado puede producirse mediante manipulación química o biológica (incluyendo, pero sin limitarse a, sustitución, adición, inserción, deleción, extracción, aislamiento, mutación y replicación) de la fuente original.

“Químicamente sintetizado”, en relación con una secuencia de ADN, significa que partes de los nucleótidos componentes se ensamblaron *in vitro*. Puede lograrse la síntesis química manual del ADN usando procedimientos bien establecidos (Caruthers, *Methodology of DNA and RNA Sequencing*, (1983), Weissman (ed.), Praeger Publishers, Nueva York, capítulo 1); la síntesis química automatizada puede realizarse usando uno de varias máquinas disponibles comercialmente.

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

La herramienta de búsqueda de alineación local básica del NCBI (“*NCBI Basic Local Alignment Search Tool*”) (BLAST) (Altschul *et al.*, 1990) está disponible de varias fuentes, incluyendo el National Center for Biological Information (NCBI, Bethesda, Md.) y en Internet, para su uso en conexión con los programas de análisis de

secuencias `blastp`, `blastn`, `blastx`, `tblastn` y `tblastx`. Puede accederse al mismo en <http://www.ncbi.nlm.nih.gov/BLAST/>. Una descripción de cómo determinar la identidad de secuencia usando este programa está disponible en http://www.ncbi.nlm.nih.gov/BLAST/blast_help.html.

Las expresiones “identidad de aminoácidos sustancial” o “identidad de secuencia de aminoácidos sustancial” tal como se aplica a secuencias de aminoácidos y tal como se usa en el presente documento indican una característica de un polipéptido, comprendiendo el péptido una secuencia que tiene una identidad de secuencia de al menos el 70 por ciento, preferiblemente una identidad de secuencia de aminoácidos del 80 por ciento, más preferiblemente una identidad de secuencia de aminoácidos del 90 por ciento y lo más preferiblemente una identidad de secuencia de al menos el 99 al 100 por ciento tal como se compara con la secuencia de proteína de SEQ ID NOS:2 y/o 4. Deseablemente, para una nicotina desmetilasa, la comparación de secuencias compara deseablemente una región que sigue al motivo del citocromo p450 (GXRXCX(G/A); SEQ ID NO:29) con el codón de terminación del péptido traducido.

Las expresiones “identidad de ácido nucleico sustancial” o “identidad de secuencia de ácido nucleico sustancial” tal como se aplica a secuencias de ácido nucleico y tal como se usa en el presente documento indican una característica de una secuencia de polinucleótidos, comprendiendo el polinucleótido una secuencia que tiene una identidad de secuencia de al menos el 50 por ciento, preferiblemente el 60, el 65, el 70 o el 75 por ciento, más preferiblemente una identidad de secuencia de ácido nucleico del 81 o el 91 por ciento y lo más preferiblemente una identidad de secuencia de al menos el 95, el 99 o incluso 100 por ciento tal como se compara con SEQ ID NOS:1, 3, 5, 6, y/o 7. Deseablemente, para una secuencia de ácido nucleico de nicotina desmetilasa, la comparación compara deseablemente una secuencia que codifica para una región que sigue al motivo del citocromo p450 (GXRXCX (G/A); SEQ ID NO:29) con el codón de terminación del péptido traducido.

Otra indicación de que las secuencias de nucleótidos son sustancialmente idénticas es si dos moléculas hibridan entre sí en condiciones rigurosas. Las condiciones rigurosas son dependientes de la secuencia y serán diferentes en diferentes circunstancias. Generalmente, las condiciones rigurosas se seleccionan para que sean de aproximadamente 5°C a aproximadamente 20°C, habitualmente de aproximadamente 10°C a aproximadamente 15°C, inferiores al punto de fusión térmica (T_m) para la secuencia específica a un pH y una fuerza iónica definidos. La T_m es la temperatura (en pH y fuerza iónica definidos) a la que el 50% de la secuencia diana hibrida con una sonda apareada. Normalmente, condiciones rigurosas serán aquellas en las que la concentración salina es aproximadamente 0,02 molar a pH 7 y la temperatura es al menos de aproximadamente 60°C. Por ejemplo, en un procedimiento de hibridación de tipo Southern convencional, las condiciones rigurosas incluirán un lavado inicial en 6x SSC a 42°C seguido por uno o más lavados adicionales en 0,2x SSC a una temperatura de al menos aproximadamente 55°C, normalmente de aproximadamente 60°C, y a menudo de aproximadamente 65°C.

Las secuencias de nucleótidos son también sustancialmente idénticas para los fines de esta invención cuando dichas secuencias de nucleótidos codifican para polipéptidos y/o proteínas que son sustancialmente idénticas. Por tanto, cuando una secuencia de ácido nucleico codifica esencialmente para el mismo polipéptido que una segunda secuencia de ácido nucleico, las dos secuencias de ácido nucleico son sustancialmente idénticas incluso si no hibridaran en condiciones rigurosas debido a la degeneración permitida por el código genético (véase, Darnell *et al.* (1990) *Molecular Cell Biology*, segunda edición Scientific American Books W. H. Freeman and Company Nueva York para una explicación de la degeneración de codones y el código genético). La pureza u homogeneidad de proteínas puede indicarse mediante varios medios bien conocidos en la técnica, tales como electroforesis en gel de poliacrilamida de una muestra de proteína, seguido por visualización tras tinción. Para ciertos fines, puede necesitarse alta resolución y pueden usarse HPLC o medios similares para la purificación.

Por un anticuerpo que “se une específicamente” o “reconoce específicamente” una nicotina desmetilasa de tabaco quiere decirse un aumento de la afinidad del anticuerpo por una nicotina desmetilasa de tabaco con respecto a una cantidad igual de cualquier otra proteína. Por ejemplo, un anticuerpo que se une específicamente a una nicotina desmetilasa de tabaco que contiene la secuencia de aminoácidos de SEQ ID NO:2 tiene deseablemente una afinidad por su antígeno que es al menos 2 veces, 5 veces, 10 veces, 30 veces o 100 veces superior que para una cantidad igual de cualquier otro antígeno, incluyendo antígenos relacionados. La unión de un anticuerpo a un antígeno, por ejemplo, nicotina desmetilasa de tabaco, puede determinarse mediante cualquiera de varios métodos convencionales en la técnica, por ejemplo, análisis de inmunotransferencia de tipo Western, ELISA o inmunoprecipitación conjunta.

Tal como se usa en el presente documento, el término “vector” se usa en referencia a moléculas de ácido nucleico que transfieren segmento(s) de ADN a una célula. Un vector puede actuar replicando el ADN y puede reproducirse independientemente en una célula huésped. El término “vehículo” se usa a veces de manera intercambiable con “vector.” El término “vector de expresión” tal como se usa en el presente documento se refiere a una molécula de ADN recombinante que contiene una secuencia codificante deseada y secuencias de ácido nucleico apropiadas necesarias para la expresión de la secuencia codificante operativamente unida en un organismo huésped particular. Las secuencias de ácido nucleico necesarias para la expresión en procariotas incluyen habitualmente un promotor, un operador (opcional) y un sitio de unión al ribosoma, a menudo junto con otras secuencias. Deseablemente, el

- promotor incluye la secuencia de SEQ ID NO:6, o un fragmento de la misma que dirige la transcripción. También son deseables secuencias promotoras que tienen una identidad de secuencia de al menos el 50%, el 60%, el 75%, el 80%, el 90%, el 95% o incluso el 99% con la secuencia de SEQ ID NO:6 y que dirigen la transcripción. Se sabe que las células eucariotas utilizan promotores, potenciadores y señales de poliadenilación y terminación, tales como la
- 5 la secuencia de UTR en 3' de SEQ ID NO:7. En algunos casos, se ha observado que los vectores de expresión de plantas requieren la presencia de intrones derivados de plantas, tales como el intrón que tiene la secuencia de SEQ ID NO:5, para tener expresión estable. Como tal, puede usarse la secuencia de SEQ ID NO: 5, o cualquier otro intrón que tiene una unión de corte y empalme de ARN apropiada tal como se describe adicionalmente en el presente documento.
- 10 Para el fin de regenerar plantas modificadas mediante ingeniería genética completas con raíces, puede insertarse un ácido nucleico en las células vegetales, por ejemplo, mediante cualquier técnica tal como inoculación *in vivo* o mediante cualquiera de las técnicas de cultivo tisular *in vitro* conocidas para producir células vegetales transformadas que pueden regenerarse para dar plantas completas. Por tanto, por ejemplo, la inserción en células vegetales puede ser mediante inoculación *in vitro* mediante *A. tumefaciens* patógeno o no patógeno. También
- 15 pueden emplearse otras técnicas de cultivo tisular de este tipo.
- "Tejido vegetal", "componente vegetal" o "célula vegetal" incluye tejidos diferenciados y no diferenciados de plantas, incluyendo, pero sin limitarse a, raíces, brotes, hojas, polen, semillas, tejido tumoral y diversas formas de células en cultivo, tales como células individuales, protoplastos, embriones y tejido calloso. El tejido vegetal puede estar en la planta o en un órgano, tejido o cultivo celular.
- 20 "Célula vegetal" tal como se usa en el presente documento incluye células vegetales en la planta y células vegetales y protoplastos en cultivo. "ADNc" o "ADN complementario" se refiere generalmente a una molécula de ADN monocatenario con una secuencia de nucleótidos que es complementaria a una molécula de ARN no procesado que contiene un intrón, o un ARNm procesado que carece de intrones. El ADNc se forma mediante la acción de la enzima transcriptasa inversa sobre un molde de ARN.
- 25 "Tabaco" tal como se usa en el presente documento incluye curado en chimenea, Virginia, Burley, oscuro, oriental y otros tipos de planta dentro del género *Nicotiana*. La semilla del género *Nicotiana* está fácilmente disponible comercialmente en forma de *Nicotiana tabacum*.
- "Artículos de fabricación" o "productos de tabaco" incluyen productos tales como rapé seco y húmedo, tabacos de mascar, cigarrillos, puros, puros delgados, tabacos de pipa, bidis y productos derivados de tabaco similares.
- 30 Por "silenciamiento génico" quiere decirse una disminución en el nivel de expresión génica (por ejemplo, la expresión de un gen que codifica para una nicotina desmetilasa de tabaco) en al menos el 30-50%, preferiblemente en al menos el 50-80% y más preferiblemente en al menos el 80-95% o superior con respecto al nivel en una planta control (por ejemplo, una planta de tabaco de tipo natural). La reducción de tales niveles de expresión puede lograrse empleando métodos convencionales que se conocen en la técnica incluyendo, sin limitación, interferencia
- 35 por ARN, interferencia por triple hebra, ribozimas, recombinación homóloga, silenciamiento génico inducido por virus, tecnologías antisentido y de cosupresión, expresión de un producto génico dominante negativo o a través de la generación de genes mutados usando técnicas de mutagénesis convencionales, tales como las descritas en el presente documento. Los niveles de un transcrito o polipéptido de nicotina de desmetilasa de tabaco, o ambos, se monitorizan según cualquier técnica convencional incluyendo, pero sin limitarse a, transferencia de tipo Northern, protección de ARNasa o inmunotransferencia.
- 40 Por una "nicotina desmetilasa de tabaco" o "nicotina desmetilasa" tal como se usa en el presente documento, quiere decirse un polipéptido que es sustancialmente idéntico a la secuencia de SEQ ID NO:2. Deseablemente, una nicotina desmetilasa de tabaco puede convertir la nicotina ($C_{10}H_{14}N_2$, también denominada 3-(1-metil-2-pirrolidinil)piridina) en nor nicotina ($C_9H_{12}N_2$). La actividad de una nicotina desmetilasa de tabaco puede someterse a ensayo usando métodos convencionales en la técnica, tal como midiendo la desmetilación de nicotina radioactiva mediante microsomas expresados en levadura, tal como se describe en el presente documento.
- 45 Por un "fragmento" o "parte" de una secuencia de aminoácidos de nicotina desmetilasa de tabaco quiere decirse al menos por ejemplo, 20, 15, 30, 50, 75, 100, 250, 300, 400 ó 500 aminoácidos contiguos de la secuencia de aminoácidos de SEQ ID NO:2. Fragmentos deseables a modo de ejemplo son los aminoácidos 1-313 de la secuencia de SEQ ID NO:2 y los aminoácidos 314-517 de la secuencia de SEQ ID NO:2. Además, con respecto a un fragmento o parte de una secuencia de ácido nucleico de nicotina desmetilasa de tabaco, los fragmentos deseables incluyen al menos 100, 250, 500, 750, 1000 ó 1500 ácidos nucleicos contiguos de la secuencia de ácido nucleico de SEQ ID NO:1. Fragmentos deseables a modo de ejemplo son los ácidos nucleicos 1-2009, 2010-2949, 2950-3946, 3947-4562, 4563-6347 y 4731-6347 de la secuencia de SEQ ID NO:1. Otros fragmentos deseables
- 50 incluyen los ácidos nucleicos 1-100, 101-250, 251-500, 501-750 y 751-998 de SEQ ID NO:5, los ácidos nucleicos 1-398, 1-1400, 1401-2009, 1840-2009, 1940-2009, 399-1240 y 1241-2009 de SEQ ID NO:6, y los ácidos nucleicos 1-100, 101-250, 251-500, 501-750, 751-1000, 1001-1250, 1251-1500 y 1501-1786 de SEQ ID NO:7.

Por un "polipéptido sustancialmente puro" quiere decirse una nicotina desmetilasa de tabaco que se ha separado de la mayoría de los componentes que la acompañan de manera natural; sin embargo, otras proteínas encontradas en la fracción microsomal asociadas con una preparación que tiene una actividad nicotina desmetilasa de al menos 8,3 pKat/mg de proteína, 9 pKat/mg de proteína, 9,5 pKat/mg de proteína, 10 pKat/mg de proteína, 10,5 pKat/mg o 10,8 pKat/mg de proteína también se considera que son un polipéptido sustancialmente puro. Normalmente, el polipéptido es sustancialmente puro cuando está libre al menos al 60% en peso de las proteínas y moléculas orgánicas que se producen de manera natural con las que está asociado de manera natural. Preferiblemente, la preparación es al menos el 75%, más preferiblemente al menos el 90% y lo más preferiblemente al menos el 99%, en peso, una nicotina desmetilasa de tabaco. Puede obtenerse una nicotina desmetilasa de tabaco sustancialmente pura, por ejemplo, mediante extracción a partir de una fuente natural (por ejemplo, una célula vegetal de tabaco); mediante expresión de un ácido nucleico recombinante que codifica para una nicotina desmetilasa de tabaco; o mediante síntesis química de la proteína. La pureza puede medirse mediante cualquier método apropiado, por ejemplo, cromatografía en columna, electroforesis en gel de poliacrilamida o mediante análisis de HPLC.

Por "molécula de ácido nucleico aislada" quiere decirse una secuencia de ácido nucleico libre de las secuencias de ácido nucleico que flanquean de manera natural la secuencia de la molécula de ácido nucleico en el genoma de un organismo.

Por "célula transformada" quiere decirse una célula dentro de la que (o dentro de un ancestro de la que) se ha introducido, por medio de técnicas de ADN recombinante, una molécula de ADN, por ejemplo, una molécula de ADN que codifica para una nicotina desmetilasa de tabaco.

Tal como se proporciona en el presente documento, los términos "citocromo p450" y "p450" se usan de manera intercambiable.

Breve descripción de los dibujos

La figura 1 es un diagrama esquemático de la estructura genómica del gen de nicotina desmetilasa de tabaco.

Las figuras 2-1 a 2-3 son la secuencia de ácido nucleico genómica de la nicotina desmetilasa de tabaco (SEQ ID NO:1) y su producto de traducción (SEQ ID NO:2).

La figura 3 es la secuencia de ácido nucleico de la región codificante de la nicotina desmetilasa de tabaco (SEQ ID NO:3) (también denominada D121-AA8) y su producto de traducción (SEQ ID NO:4).

La figura 4 es la secuencia de ácido nucleico de un intrón (SEQ ID NO:5) presente en la secuencia genómica de la nicotina desmetilasa de tabaco.

La figura 5 es la secuencia de ácido nucleico de la región promotora de la nicotina desmetilasa de tabaco (SEQ ID NO:6).

La figura 6 es la secuencia de ácido nucleico de la UTR en 3' del gen de nicotina desmetilasa de tabaco (SEQ ID NO:7).

Descripción detallada

Identificación de un clon genómico que codifica para una nicotina desmetilasa de tabaco

Según la presente invención, el ARN se extrajo de tejido de *Nicotiana* de líneas de *Nicotiana* convertidoras y no convertidoras. El ARN extraído se usó entonces para crear ADNc. Las secuencias de ácido nucleico de la presente invención se generaron entonces usando dos estrategias.

En la primera estrategia, el ARN enriquecido en poli A se extrajo de tejido vegetal y se preparó ADNc mediante PCR con transcripción inversa. Entonces, el ADNc monocatenario se usó para crear poblaciones de PCR específicas de p450 usando cebadores degenerados más un cebador inverso de oligo d(T). El diseño de cebadores fue basándose en los motivos altamente conservados de secuencias génicas de citocromo p450 de otras plantas. Se exponen ejemplos de cebadores degenerados específicos en la figura 1 de las publicaciones de solicitud de patente US 2004/0103449 A1 y US 2004/0111759 A1. Se analizó adicionalmente la secuencia de fragmentos de plásmidos que contenían insertos de tamaño apropiado. Estos insertos de tamaño normalmente oscilan entre aproximadamente 300 y aproximadamente 800 nucleótidos dependiendo de qué cebadores se usaron.

En una segunda estrategia, se construyó inicialmente una biblioteca de ADNc. El ADNc en los plásmidos se usó para crear poblaciones de PCR específicas de p450 usando cebadores degenerados más un cebador de T7 en el

plásmido como cebador inverso. Como en la primera estrategia, se analizó adicionalmente la secuencia de fragmentos de los plásmidos que contenían insertos de tamaño apropiado.

Pueden usarse como materiales de partida líneas de plantas de *Nicotiana* que se sabe que producen altos niveles de nornicotina (convertidoras) y líneas vegetales que tienen bajos niveles de nornicotina. Entonces, pueden quitarse las hojas de las plantas y tratarse con etileno para activar las actividades enzimáticas de p450 definidas en el presente documento. El ARN total se extrae usando técnicas conocidas en la técnica. Entonces, pueden generarse fragmentos de ADNc usando PCR (RT-PCR) con el cebador de oligo d(T). Entonces, puede construirse la biblioteca de ADNc tal como se describe más completamente en los ejemplos en el presente documento.

La región conservada de las enzimas de tipo p450 se usó como molde para cebadores degenerados. Usando cebadores degenerados, se amplificaron bandas específicas de p450 mediante PCR. Se identificaron bandas indicativas para enzimas similares a p450 mediante secuenciación de ADN. Los fragmentos de PCR se caracterizaron usando búsqueda BLAST, alineación u otras herramientas para identificar candidatos apropiados.

Se usó información de secuencia a partir de fragmentos identificados para desarrollar los cebadores de PCR. Estos cebadores, en combinación con cebadores de plásmido en la biblioteca de ADNc, se usaron para clonar genes de p450 de longitud completa. Se realizó análisis inverso de tipo Southern a gran escala para examinar la expresión diferencial de todos los clones de fragmentos obtenidos y en algunos casos clones de longitud completa. En este aspecto de la invención, estos ensayos de tipo Southern inversos a gran escala pueden realizarse usando ADNc totales marcados de diferentes tejidos como sonda para hibridar con fragmentos de ADN clonados con el fin de seleccionar todos los insertos clonados. También se usaron ensayos de transferencia de tipo Northern no radioactivos y radioactivos (P^{32}) para caracterizar los fragmentos de p450 clonados y clones de longitud completa.

Se prepararon anticuerpos específicos de péptidos derivando su secuencia de aminoácidos y seleccionando regiones peptídicas que eran antigénicas y únicas con respecto a otros clones. Se prepararon anticuerpos de conejo frente a péptidos sintéticos conjugados con una proteína portadora. Se realizaron análisis de inmunotransferencia de tipo Western u otros métodos inmunológicos sobre tejido vegetal usando estos anticuerpos. Además, se prepararon anticuerpos específicos de péptidos para varios clones de longitud completa derivando su secuencia de aminoácidos y seleccionando regiones peptídicas que eran potencialmente antigénicas y eran únicas con respecto a otros clones. Se prepararon anticuerpos de conejo frente a péptidos sintéticos conjugados con una proteína portadora. Se realizaron análisis de inmunotransferencia de tipo Western usando estos anticuerpos.

Regulación por disminución de la nicotina desmetilasa de tabaco

Se generan plantas que tienen una disminución de la expresión de una nicotina desmetilasa de tabaco según métodos de silenciamiento génico convencionales. (Para revisiones, 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, el promotor del gen de nicotina desmetilasa de tabaco (por ejemplo, SEQ ID NO:6), el gen estructural (SEQ ID NO:3), el intrón (SEQ ID NO:5) o la UTR en 3' (SEQ ID NO:7) o el clon genómico entero (SEQ ID NO:1) pueden usarse para alterar fenotipos del tabaco o metabolitos del tabaco, por ejemplo, nornicotina en cualquier especie de *Nicotiana*. La disminución de la expresión de un gen de nicotina desmetilasa de tabaco puede conseguirse usando, por ejemplo, interferencia por ARN (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); silenciamiento génico inducido por virus ("VIGS") (Baulcombe, *Current Opinions in Plant Biology*, 2:109-113, 1999; Cogoni y Macino, *Genes Dev* 10: 638-643, 2000; Ngelbrecht *et al.*, *PNAS* 91: 10502-10506, 1994); silenciamiento del gen diana transfiriendo un gen endógeno vegetal en la orientación sentido (Jorgensen *et al.*, *Plant Mol Biol* 31: 957-973, 1996); expresión de gen antisentido; recombinación homóloga (Ohl *et al.*, *Homologous Recombination and Gene Silencing in Plants* (Kluwer, Dordrecht, Países Bajos, 1994); sistemas Cre/lox (Qin *et al.*, *PNAS* 91: 1706-1710, 1994; Koshinsky *et al.*, *The Plant Journal* 23: 715-722, 2000; Chou, *et al.*, *Plant and Animal Genome VII Conference Abstracts*. San Diego, CA, 17-21 de enero de 1999); atrapamiento génico y etiquetado de ADN-T (Burns *et al.*, *Genes Dev.* 8: 1087-1105, 1994; Spradling, *et al.*, *PNAS* 92:10824-10830, 1995; Skarnes *et al.*, *Bio/ Technology* 8, 827-831, 1990; Sundaresan, *et al.*, *Genes Dev.* 9: 1797-1810, 1995); y cualquiera de los otros sistemas de silenciamiento génico posibles que están disponibles en las áreas de la ciencia que dan como resultado la regulación por disminución de la expresión de una nicotina desmetilasa de tabaco o una reducción en su actividad enzimática. Se describen en más detalle a continuación métodos a modo de ejemplo.

Interferencia por ARN

La interferencia por ARN ("ARNi") es un proceso generalmente aplicable para inducir silenciamiento génico postraducciona 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). La ARNi implica la introducción de ARN con carácter parcial o completamente bicatenario en la célula o en el entorno extracelular. La

inhibición es específica porque se elige una secuencia de nucleótidos de una parte del gen diana (por ejemplo, una nicotina desmetilasa de tabaco) para producir ARN inhibidor. La parte elegida abarca generalmente exones del gen diana, pero la parte elegida también puede incluir regiones no traducidas (UTR), así como intrones (por ejemplo, las secuencias de SEQ ID NO:5 ó 7).

5 Por ejemplo, para construir vectores de transformación que producen ARN que pueden formar dúplex, dos secuencias de ácido nucleico de la nicotina desmetilasa de tabaco, una en la orientación sentido y la otra en la antisentido, pueden estar operativamente unidas, y colocadas bajo el control de un promotor viral fuerte, tal como CaMV 35S o el promotor aislado del virus de la raya parda de Cassava (CBSV). Sin embargo, el uso del promotor endógeno, tal como el promotor de la nicotina desmetilasa de tabaco que tiene la secuencia de SEQ ID NO:6, o un
10 fragmento de la misma que dirige la transcripción, también puede ser deseable. La longitud de las secuencias de ácido nucleico de la nicotina desmetilasa de tabaco incluidas en un constructo de este tipo es deseablemente de al menos 25 nucleótidos, pero puede abarcar una secuencia que incluye hasta el gen de nicotina desmetilasa de tabaco de longitud completa.

15 Pueden introducirse constructos que producen ARN que pueden formar dúplex en el genoma de una planta, tal como una planta de tabaco, mediante transformación mediada por *Agrobacterium* (Chuang *et al.*, Proc. Natl. Acad. Sci. USA 97:4985-4990, 2000), produciendo interferencia genética heredable y específica en una nicotina desmetilasa de tabaco. El ARN bicatenario también puede introducirse directamente en la célula (es decir, intracelularmente) o introducirse extracelularmente, por ejemplo, bañando una semilla, plántula o planta en una disolución que contiene el ARN bicatenario.

20 Dependiendo de la dosis de material de ARN bicatenario suministrada, la ARNi puede proporcionar pérdida parcial o completa de la función del gen diana. Puede obtenerse una reducción o pérdida de la expresión génica en al menos el 99% de las células seleccionadas como diana. En general, dosis inferiores de material inyectado y tiempos más largos tras la administración de ARNbc dan como resultado la inhibición en una fracción más pequeña de células.

25 El ARN usado en la ARNi puede comprender una o más hebras de ribonucleótido polimerizado; puede incluir modificaciones en o bien la estructura principal de fosfato-azúcar o bien el nucleósido. La estructura bicatenaria puede estar formada por una única hebra de ARN autocomplementaria o por dos hebras de ARN complementarias y puede iniciarse la formación de dúplex de ARN o bien dentro o bien fuera de la célula. El ARN puede introducirse en una cantidad que permite el suministro de al menos una copia por célula. Sin embargo, dosis superiores (por ejemplo, al menos 5, 10, 100, 500 ó 1000 copias por célula) de material bicatenario pueden proporcionar una
30 inhibición más eficaz. La inhibición es específica de secuencia porque las secuencias de nucleótidos correspondientes a la región de dúplex del ARN se seleccionan como diana para la inhibición genética. Se prefiere para la inhibición el ARN que contiene una secuencia de nucleótido idéntica a una parte del gen diana. También pueden ser eficaces para la inhibición secuencias de ARN con inserciones, deleciones y mutaciones puntuales individuales con respecto a la secuencia diana. Por tanto, puede optimizarse la identidad de secuencia mediante
35 algoritmos de alineación conocidos en la técnica y calculando el porcentaje de diferencias entre las secuencias de nucleótidos. Alternativamente, puede definirse funcionalmente la región de dúplex del ARN como una secuencia de nucleótidos que puede hibridar con una parte del transcrito del gen diana.

Además, el ARN usado para la ARNi puede sintetizarse o bien *in vivo* o bien *in vitro*. Por ejemplo, la ARN polimerasa endógena en la célula puede mediar la transcripción *in vivo*, o la ARN polimerasa clonada puede usarse para la
40 transcripción *in vivo* o *in vitro*. Para la transcripción a partir de un transgén *in vivo* o un constructo de expresión, puede usarse una región reguladora para transcribir la hebra de ARN (o hebras).

Interferencia por triple hebra

45 La expresión génica de la nicotina desmetilasa de tabaco endógena también puede regularse por disminución seleccionando como diana secuencias de desoxirribonucleótidos complementarias a la región reguladora de un gen de nicotina desmetilasa de tabaco (por ejemplo, regiones potenciadoras o promotoras) para formar estructuras de triple hélice que impiden la transcripción del gen de nicotina desmetilasa de tabaco en células diana. (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.)

50 Las moléculas de ácido nucleico usadas en la formación de triple hélice para la inhibición de la transcripción son preferiblemente monocatenarias y se componen de desoxirribonucleótidos. La composición de bases de estos oligonucleótidos debe promover la formación de triple hélice mediante las reglas de apareamiento de bases de Hoogsteen, que generalmente requieren que estén presentes tramos considerables de o bien purinas o bien pirimidinas en una hebra de un dúplex. Las secuencias de nucleótidos pueden ser a base de pirimidinas, lo que dará como resultado tripletes TAT y CGC a través de las tres hebras asociadas de la triple hélice resultante. Las
55 moléculas ricas en pirimidina proporcionan complementariedad de bases con una región rica en purinas de una hebra individual del dúplex en una orientación en paralelo a esa hebra. Además, pueden elegirse moléculas de ácido nucleico que son ricas en purinas, por ejemplo, que contienen un tramo de residuos de G. Estas moléculas formarán

una triple hélice con un dúplex de ADN que es rico en pares GC, en el que la mayoría de los residuos de purina están ubicados en una hebra individual del dúplex seleccionado como diana, dando como resultado tripletes CGC a través de las tres hebras en el tríplex.

- 5 Alternativamente, las posibles secuencias que pueden seleccionarse como diana para la formación de la triple hélice pueden aumentarse creando una molécula de ácido nucleico "de cambio de sentido". Las moléculas de cambio de sentido se sintetizan de una manera 5'-3', 3'-5' alternante, de manera que aparean sus bases con en primer lugar una hebra de un dúplex y luego con la otra, eliminando la necesidad de que esté presente un tramo considerable de o bien purinas o bien pirimidinas en una hebra de un dúplex.

Ribozimas

- 10 Las ribozimas son moléculas de ARN que actúan como enzimas y pueden diseñarse mediante ingeniería genética para escindir otras moléculas de ARN. Una ribozima puede diseñarse para aparearse específicamente con prácticamente cualquier ARN diana y escindir la estructura principal de fosfodiéster en una ubicación específica, inactivando funcionalmente de ese modo el ARN diana. La propia ribozima no se consume en este proceso y puede actuar catalíticamente para escindir copias múltiples de moléculas diana de ARNm. Por consiguiente, las ribozimas
- 15 también pueden usarse como medios para regular por disminución la expresión de una nicotina desmetilasa de tabaco. El diseño y uso de ribozimas específicas de ARN diana se describe en Haseloff *et al.* (Nature 334:585-591, 1988). Preferiblemente, la ribozima incluye al menos aproximadamente 20 nucleótidos continuos complementarios a la secuencia diana (por ejemplo, una nicotina desmetilasa de tabaco) a cada lado del sitio activo de la ribozima.

- 20 Además, también pueden incluirse secuencias de ribozima dentro de un ARN antisentido para conferir actividad de escisión de ARN al ARN antisentido y, de ese modo, aumentar la eficacia del constructo antisentido.

Recombinación homóloga

- 25 La tecnología de sustitución génica es otro método deseable para la regulación por disminución de la expresión de un gen dado, por ejemplo, una nicotina desmetilasa de tabaco. La tecnología de sustitución génica está basada en la recombinación homóloga (véase, Schnable *et al.*, Curr. Opinions Plant Biol. 1:123-129, 1998). La secuencia de ácido nucleico de la enzima de interés tal como una nicotina desmetilasa de tabaco puede manipularse mediante mutagénesis (por ejemplo, inserciones, delecciones, duplicaciones o sustituciones) para disminuir la función enzimática. La secuencia alterada puede introducirse entonces en el genoma para sustituir el gen existente, por ejemplo, de tipo natural, mediante recombinación homóloga (Puchta *et al.*, Proc. Natl. Acad. Sci. USA 93:5055-5060, 1996; y Kempin *et al.*, Nature 389: 802-803, 1997).

30 Cosupresión

Un método deseable adicional de silenciamiento de la expresión génica es la cosupresión (también denominada supresión sentido). Se ha mostrado que esta técnica, que implica la introducción de un ácido nucleico configurado en la orientación sentido, bloquea eficazmente la transcripción de genes diana (véase, por ejemplo, Napoli *et al.*, Plant Cell, 2:279-289, 1990 y Jorgensen *et al.*, patente estadounidense n.º 5.034.323).

- 35 Generalmente, la supresión sentido implica la transcripción de la secuencia introducida. Sin embargo, la cosupresión también puede producirse cuando la secuencia introducida no contiene ninguna secuencia codificante *per se*, sino sólo un intrón (por ejemplo, la secuencia de SEQ ID NO:5) o secuencias no traducidas tales como la secuencia de SEQ ID NO:7 u otras secuencias de este tipo sustancialmente idénticas a secuencias presentes en el transcrito primario del gen endógeno que va a reprimirse. La secuencia introducida generalmente será sustancialmente
- 40 idéntica al gen endógeno seleccionado como diana para la represión. Tal identidad es normalmente superior a aproximadamente el 50%, pero se prefieren identidades superiores (por ejemplo, del 80% o incluso el 95%) debido a que dan como resultado una represión más eficaz. El efecto de la cosupresión también puede aplicarse a otras proteínas dentro de una familia de genes similar que muestran homología u homología sustancial. Los segmentos de un gen de una planta pueden usarse directamente, por ejemplo, para inhibir la expresión de genes homólogos en
- 45 diferentes especies vegetales.

- En la supresión sentido, la secuencia introducida, requiere menos que identidad absoluta, no necesita ser de longitud completa, con respecto a o bien el producto de transcripción primario o bien ARNm completamente procesado. Un grado superior de identidad de secuencia en una secuencia de longitud completa más corta compensa una secuencia más larga de identidad inferior. Además, la secuencia introducida no necesita tener el mismo patrón de intrones o exones, y la identidad de segmentos no codificantes puede ser igualmente eficaz. Se prefieren secuencias de al menos 50 pares de bases, prefiriéndose más secuencias introducidas de mayor longitud (véanse, por ejemplo, los métodos descritos por Jorgensen *et al.*, patente estadounidense n.º 5.034.323).
- 50

Supresión antisentido

En la tecnología antisentido, un segmento de ácido nucleico del gen de plantas deseado, tal como el que se encuentra en SEQ ID NOS:1, 3, 5, 6 ó 7, se clona y se une operativamente a una región de control de la expresión de manera que se sintetiza la hebra antisentido del ARN. Entonces, el constructo se transforma en plantas y se produce la hebra antisentido del ARN. En células vegetales, se ha mostrado que el ARN antisentido inhibe la expresión génica.

Generalmente, el segmento de ácido nucleico que va a introducirse en la supresión antisentido es sustancialmente idéntico a al menos una parte del gen o genes endógenos que van a reprimirse, pero no necesita ser idéntico. Las secuencias de ácido nucleico de la nicotina desmetilasa de tabaco dadas a conocer en el presente documento pueden estar incluidas en vectores diseñados de manera que el efecto inhibitor se aplica a otras proteínas dentro de una familia de genes que muestran homología u homología sustancial con el gen diana. Pueden usarse segmentos de un gen de una planta, por ejemplo, directamente para inhibir la expresión de genes homólogos en diferentes variedades de tabaco.

La secuencia introducida tampoco necesita ser de longitud completa con respecto a o bien el producto de transcripción primario o bien el ARNm completamente procesado. Generalmente, puede usarse una homología superior para compensar el uso de una secuencia más corta. Además, la secuencia introducida no necesita tener el mismo patrón de intrones o exones, y la homología de segmentos no codificantes será igualmente eficaz. En general, una secuencia antisentido de este tipo tendrá habitualmente al menos 15 pares de bases, preferiblemente de manera aproximada 15-200 pares de bases y más preferiblemente 200-2.000 pares de bases de longitud o superior. La secuencia antisentido puede ser complementaria a todo o una parte del gen que va a suprimirse (por ejemplo, un promotor de nicotina desmetilasa de tabaco (SEQ ID NO:6), exón, intrón (SEQ ID NO:5) o UTR (SEQ ID NO:7) y, tal como aprecian los expertos en la técnica, el sitio o sitios particulares a los que la secuencia antisentido se une así como la longitud de la secuencia antisentido oscilarán, dependiendo del grado de inhibición deseado y la singularidad de la secuencia antisentido. Un constructo transcripcional que expresa una secuencia de nucleótidos antisentido reguladora negativa de plantas incluye, en la dirección de transcripción, un promotor, la secuencia que codifica para el ARN antisentido en la hebra sentido y una región de terminación de la transcripción. Las secuencias antisentido pueden construirse y expresarse tal como se describe, por ejemplo, en van der Krol *et al.* (Gene 72: 45-50, 1988); Rodermeil *et al.* (Cell 55: 673-681, 1988); Mol *et al.* (FEBS Lett. 268: 427-430, 1990); Weigel y Nilsson (Nature 377: 495-500, 1995); Cheung *et al.*, (Cell 82: 383-393, 1995); y Shewmaker *et al.* (patente estadounidense n.º 5.107.065).

Dominantes negativos

Pueden someterse a ensayo plantas transgénicas que expresan un transgén que codifica para un producto génico dominante negativo de una nicotina desmetilasa de tabaco en entornos artificiales o en el campo para demostrar que el transgén confiere regulación por disminución de una nicotina desmetilasa de tabaco en la planta transgénica. Se construyen transgenes dominante negativos según métodos conocidos en la técnica. Normalmente, un gen dominante negativo codifica para un polipéptido regulador negativo mutante de una nicotina desmetilasa de tabaco que, cuando se sobreexpresa, altera la actividad de la enzima de tipo natural.

Mutantes

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

Por ejemplo, puede usarse mutagénesis insercional de ADN-T para generar mutaciones insercionales en un gen de nicotina desmetilasa de tabaco para regular por disminución la expresión del gen. Teóricamente, se requieren aproximadamente 100.000 inserciones de ADN-T 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). Pueden seleccionarse líneas de plantas etiquetadas con ADN-T usando análisis de reacción en cadena de la polimerasa (PCR). Por ejemplo, puede diseñarse un cebador para un extremo del ADN-T y puede diseñarse otro cebador para el gen de interés y ambos cebadores pueden usarse en el análisis de PCR. Si no se

obtiene un producto de PCR, entonces no hay inserción en el gen de interés. En cambio, si se obtiene un producto de PCR, entonces hay una inserción en el gen de interés.

Puede evaluarse la expresión de una nicotina desmetilasa de tabaco mutada según procedimientos convencionales (por ejemplo, los descritos en el presente documento) y, opcionalmente, puede compararse con la expresión de la enzima no mutada. Cuando se compara con plantas no mutadas, plantas mutadas que tienen expresión disminuida de un gen que codifica para una nicotina desmetilasa de tabaco son realizaciones deseables de la presente invención.

Promotores vegetales

Un ejemplo de un promotor vegetal útil según la invención es el promotor de nicotina desmetilasa que tiene la secuencia de SEQ ID NO:6, o un fragmento de la misma que dirige la transcripción. Otros promotor deseable es un promotor de caulimovirus, por ejemplo, un promotor del virus del mosaico de la coliflor (CaMV) o el promotor del virus del mosaico de la vena de Cassava (CsVMV). Estos promotores confieren altos niveles de expresión en la mayoría de tejidos vegetales, y la actividad de estos promotores no es dependiente de proteínas codificadas viralmente. CaMV es una fuente para los promotores tanto 35S como 19S. Se conocen en la técnica ejemplos de constructos de expresión vegetales que usan estos promotores. En la mayoría de tejidos de plantas transgénicas, el promotor 35S de CaMV es un promotor fuerte. El promotor de CaMV también es altamente activo en monocotiledóneas. Además, la actividad de este promotor puede aumentarse adicionalmente (es decir, entre 2-10 veces) mediante duplicación del promotor 35S de CaMV.

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

Los promotores de monocotiledóneas a modo de ejemplo incluyen, sin limitación, el promotor del virus del moteado amarillo de *Commelina*, el promotor del badnavirus de la caña de azúcar, el promotor del virus baciliforme del tungro del arroz, el elemento del virus de la raya del maíz y el promotor del virus del enanismo del trigo.

Para ciertas aplicaciones, puede ser deseable producir una nicotina desmetilasa de tabaco, tal como una nicotina desmetilasa de tabaco mutante dominante negativa, en un tejido apropiado, a un nivel apropiado, o en un momento del desarrollo apropiado. Para este fin, se ha mostrado que una colección de promotores génicos, cada uno con sus propias características distintas incorporadas en sus secuencias reguladoras, se regulan en respuesta a señales inducibles tales como el entorno, hormonas y/o señales de desarrollo. Estos incluyen, sin limitación, promotores génicos que son responsables de la expresión génica regulada por calor, la expresión génica regulada por luz (por ejemplo, el gen *rbcS-3A* del guisante; el promotor *rbcS* del maíz; el gen de la proteína de unión a clorofila a/b encontrado en el guisante; o el promotor Arabssu), la expresión génica regulada por hormonas (por ejemplo, las secuencias sensibles a ácido abscísico (ABA) del gen *Em* del trigo; los promotores HVA1 y HVA22 inducibles por ABA, y rd29A de cebada y *Arabidopsis*; y la expresión génica inducida por heridas (por ejemplo, de *wun1*), la expresión génica específica de órganos (por ejemplo, del gen de la proteína de almacenamiento específica de tubérculos; el gen de zeína de 23 KDa del maíz descrito por; o el gen β -faseolina del guisante francés), o promotores inducibles por patógenos (por ejemplo, los promotores *PR-1*, *prp-1*, o 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 de plantas

Normalmente, los vectores de expresión plantas incluyen (1) un gen de plantas clonado (por ejemplo, un gen de nicotina desmetilasa de tabaco) bajo el control transcripcional de secuencias reguladoras en 5' y 3' (por ejemplo, el promotor de la nicotina desmetilasa de tabaco (SEQ ID NO:6) y las regiones de UTR en 3' (SEQ ID NO:7)) y (2) un marcador seleccionable dominante. Tales vectores de expresión de plantas también pueden contener; si se desea, una región reguladora del promotor (por ejemplo, una que confiere expresión inducible o constitutiva, específica de tejido o célula, regulada por el desarrollo o por el entorno, inducida por heridas o por patógenos), un sitio de inicio de la iniciación de la transcripción, un sitio de unión al ribosoma, una señal de procesamiento de ARN, un sitio de terminación de la transcripción y/o una señal de poliadenilación.

Los vectores de expresión de plantas también pueden incluir opcionalmente señales de procesamiento del ARN, por ejemplo, intrones, que se ha mostrado que son importantes para la síntesis y acumulación de ARN eficaz. La ubicación de las secuencias de corte y empalme de ARN puede influir drásticamente en el nivel de expresión transgénica en plantas. En vista de este hecho, un intrón puede colocarse en el sentido de 5' o en el sentido de 3' de una secuencia codificante de nicotina desmetilasa de tabaco en el transgén para alterar los niveles de expresión génica.

Además de las secuencias de control reguladoras en 5' mencionadas anteriormente, los vectores de expresión también pueden incluir regiones de control reguladoras que generalmente están presentes en las regiones 3' de los genes vegetales. Por ejemplo, la región terminadora en 3' (por ejemplo, la secuencia de SEQ ID NO:7) puede incluirse en el vector de expresión para aumentar la estabilidad del ARNm. Una región terminadora de este tipo puede derivarse de la región terminadora PI-II de la patata. Además, otros terminadores comúnmente usados se derivan de las señales de octopina o nopalina sintasa.

El vector de expresión de plantas también contiene normalmente un gen marcador seleccionable dominante usado para identificar las células que se han transformado. Los genes seleccionables útiles para sistemas vegetales incluyen el gen de aminoglicósido fosfotransferasa del transposón Tn5 (Aph II), genes que codifican para genes de resistencia a antibióticos, por ejemplo, los que codifican para resistencia a higromicina, kanamicina, bleomicina, neomicina, G418, estreptomycin o espectinomycin. También pueden usarse genes requeridos para la fotosíntesis como marcadores seleccionables en cepas de fotosíntesis deficiente. Finalmente, pueden usarse genes que codifican para resistencia a herbicidas como marcadores seleccionables; los genes de resistencia a herbicida útiles incluyen el gen *bar* que codifica para la enzima fosfinotricin acetiltransferasa y que confiere resistencia al herbicida de amplio espectro Basta® (Bayer Cropscience Deutschland GmbH, Langenfeld, Alemania). Otros marcadores seleccionables incluyen genes que proporcionan resistencia a otros herbicidas de este tipo tales como glifosato y similares, y herbicidas de imidazolinonas, sulfonilureas, triazopirimidina, tales como clorosulfurón, bromoxinil, dalapón y similares. Además, pueden usarse genes que codifican para dihidrofolato reductasa en combinación con moléculas tales como metatrexato.

El uso eficaz de marcadores seleccionables se facilita mediante una determinación de la susceptibilidad de una célula vegetal a un agente seleccionable particular y una determinación de la concentración de este agente que destruye eficazmente la mayoría, si no todas, de las células transformadas. Algunas concentraciones útiles de antibióticos para la transformación de 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 a herbicidas se describe, por ejemplo, por Vasil (Cell Culture and Somatic Cell Genetics of Plants, vol. I, II, III Laboratory Procedures and Their Applications Academic Press, Nueva York, 1984).

Además de un marcador seleccionable, puede ser deseable usar un gen indicador. En algunos casos, puede usarse un gen indicador sin un marcador seleccionable. Los genes indicadores son genes que normalmente no están presentes ni se expresan en el tejido u organismo del receptor. El gen indicador normalmente codifica para una proteína que proporciona algún cambio fenotípico o propiedad enzimática. Ejemplos de tales genes se proporcionan en Weising *et al.* (Ann. Rev. Genetics 22:421, 1988). Los genes indicadores preferidos incluyen sin limitación el gen de glucuronidasa (GUS) y los genes GFP.

Tras la construcción del vector de expresión de plantas, están disponibles varios métodos convencionales para la introducción del vector en un huésped vegetal, generando de ese modo una planta transgénica. Estos métodos incluyen (1) transformación mediada por *Agrobacterium* (*A. tumefaciens* o *A. rhizogenes*) (véanse, por ejemplo, Lichtenstein y Fuller en: Genetic Engineering, vol. 6, PWJ Rigby, ed, Londres, Academic Press, 1987; y Lichtenstein, C.P., y Draper, J., En: ADN Cloning, vol. II, D.M. Glover, ed, Oxford, IRI Press, 1985; las patentes estadounidenses 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 las solicitudes de patente europea números 0131624B1, 0159418B1, 0120516, 0176112, 0116718, 0267159, 0290799, 0292435, 0320500, 0604622 y 0627752), (2) el sistema de suministro de partículas (véanse, por ejemplo, las patentes estadounidenses números 4.945.050 y 5.141.131), (3) protocolos de microinyección, (4) procedimientos con polietilenglicol (PEG), (5) captación de ADN mediada por liposomas, (6) protocolos de electroporación (véanse, por ejemplo, los documentos WO 87/06614, WO 92/09696 y WO 93/21335; y las patentes estadounidenses números 5.384.253 y 5.472.869), (7) el método de la agitación con vórtex u (8) la denominada metodología de "fibras cortas monocristalinas" (véanse, por ejemplo, Coffee *et al.*, patentes estadounidenses números 5.302.523 y 5.464.765). El tipo de tejido vegetal que puede transformarse con un vector de expresión incluye tejido embrionario, tejido calloso tipo I y II, hipocotilos, meristemo y similares.

Una vez introducido en el tejido vegetal, la expresión del gen estructural puede someterse a ensayo mediante cualquier medio conocido en la técnica, y la expresión puede medirse como ARNm transcrito, proteína sintetizada o la cantidad de silenciamiento génico que se produce tal como se determina mediante la monitorización de metabolitos por medio de análisis químico de alcaloides secundarios en el tabaco (tal como se describe en el presente documento; véase también la patente estadounidense n.º 5.583.021). Se conocen técnicas para el cultivo *in vitro* del tejido vegetal, y en varios casos, para la regeneración para dar plantas completas (véanse, por ejemplo, las patentes estadounidenses números 5.595.733 y 5.766.900). Los expertos en la técnica conocen procedimientos para transferir el complejo de expresión introducido a cultivares comercialmente útiles.

Una vez que se obtienen células vegetales que expresan el nivel deseado de nicotina desmetilasa (o nor nicotina o NNN o ambas), pueden regenerarse tejidos vegetales y plantas completas a partir de las mismas usando métodos y técnicas bien conocidas en la técnica. Entonces, las plantas regeneradas se reproducen por medios convencionales y los genes introducidos pueden transferirse a otras cepas y cultivares mediante técnicas de cultivo de plantas

convencionales.

- 5 Las plantas de tabaco transgénicas pueden incorporar un ácido nucleico de cualquier parte del gen genómico en diferentes orientaciones para o bien regulación por disminución, por ejemplo, orientación antisentido o bien en una forma para inducir ARNi, o sobreexpresión, por ejemplo, orientación sentido. La sobreexpresión de la secuencia de ácido nucleico que codifica para la secuencia de aminoácidos completa o una parte funcional de la misma de un gen de nicotina desmetilasa de tabaco de longitud completa es deseable para aumentar la expresión de la nicotina desmetilasa dentro de las líneas de *Nicotiana*.

Determinación de los niveles transcripcionales o traduccionales de una nicotina desmetilasa de tabaco

- 10 La expresión de la nicotina desmetilasa de tabaco puede medirse, por ejemplo, mediante análisis de transferencia de tipo Northern convencional (Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001), y Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., (1989)) usando una nicotina desmetilasa de tabaco (o fragmento de ADNc) como sonda de hibridación. La determinación de los niveles de expresión de ARN también puede ayudarse mediante PCR con transcripción inversa (rtPCR), incluyendo rtPCR cuantitativa (véanse, por ejemplo, Kawasaki *et al.*, en PCR Technology: Principles and Applications of DNA amplification (H. A. Erlich, Ed.) Stockton Press (1989); Wang *et al.* en PCR Protocols: A Guide to Methods and Applications (M. A. Innis, *et al.*, Eds.) Academic Press (1990); y Freeman *et al.*, Biotechniques 26:112-122 y 124-125, 1999). Técnicas bien conocidas adicionales para determinar la expresión de un gen de nicotina desmetilasa de tabaco incluyen hibridación *in situ*, e hibridación *in situ* fluorescente (véase, por ejemplo, Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001)). Las técnicas convencionales 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 nicotina desmetilasa de tabaco y una planta control.

- 25 Si se desea, la expresión de un gen de nicotina desmetilasa de tabaco puede medirse al nivel de la producción de proteína de nicotina desmetilasa de tabaco usando el mismo enfoque general y técnicas de análisis de proteínas convencionales que incluyen ensayos de Bradford, ensayos espectrofotométricos y técnicas de detección inmunológica, tales como inmunotransferencia de tipo Western o inmunoprecipitación con un anticuerpo específico de nicotina desmetilasa de tabaco (Ausubel *et al.*, Current Protocols in Molecular Biology, John Wiley & Sons, Nueva York, NY, (2001), y Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., (1989)).

Identificación de moduladores de una nicotina desmetilasa de tabaco

- 30 El aislamiento de un ADNc de nicotina desmetilasa de tabaco también facilita la identificación de moléculas que aumentan o disminuyen la expresión de una nicotina desmetilasa de tabaco. Según un enfoque, se añaden moléculas candidatas a concentraciones variables a un medio de cultivo de células (por ejemplo, células procariotas tales como *E. coli* o células eucariotas tales como células de levadura, de mamífero, de insecto o vegetales) que expresan un ARNm de nicotina desmetilasa de tabaco. Entonces, se mide la expresión de la nicotina desmetilasa de tabaco en presencia y ausencia de una molécula candidata usando métodos convencionales tales como los expuestos en el presente documento.

- 40 Los moduladores candidatos pueden ser moléculas purificadas (o sustancialmente purificadas) o pueden ser un componente de una mezcla de compuestos. En un ensayo de compuestos mezclados, la expresión de la nicotina desmetilasa de tabaco se somete a prueba frente a subgrupos progresivamente más pequeños del conjunto de compuestos candidatos (por ejemplo, producidos mediante técnicas de purificación convencionales, por ejemplo, HPLC) hasta que se demuestra que un único compuesto o una mezcla de compuestos mínima altera la expresión génica de la nicotina desmetilasa de tabaco. Una molécula que promueve una disminución en la expresión de la nicotina desmetilasa de tabaco se considera particularmente útil. Los moduladores que se encuentran que son eficaces al nivel de la expresión o actividad de la nicotina desmetilasa de tabaco pueden confirmarse como útiles en la planta.

Para usos agrícolas, las moléculas, los compuestos o los agentes identificados usando los métodos dados a conocer en el presente documento pueden usarse como productos químicos aplicados como aerosoles o polvos sobre el follaje de las plantas. Las moléculas, los compuestos o los agentes también pueden aplicarse a las plantas en combinación con otra molécula que produce algún beneficio a la planta.

- 50 Usos del promotor del gen de nicotina desmetilasa y regiones no traducidas

La región promotora del gen de nicotina desmetilasa descrito en el presente documento es inducible con etileno o está relacionado con la senescencia de la planta. Por consiguiente, este promotor podría usarse para dirigir la expresión de cualquier producto génico deseable para mejorar la calidad de los cultivos o potenciar rasgos específicos. Ya que el promotor de la nicotina desmetilasa de tabaco (por ejemplo, SEQ ID NO:6) es inducible y se

expresa durante un periodo particular del ciclo de vida de la planta, pueden introducirse constructos que contienen este promotor en la planta para expresar genes únicos implicados en la biosíntesis de productos de sabor y aroma que resultan de metabolitos secundarios. Ejemplos de tales compuestos son compuestos en la ruta terpenoide, otros alcaloides, hormonas vegetales, flavonoides o restos que contienen azúcar. También puede usarse un promotor de nicotina desmetilasa de tabaco para aumentar o modificar la expresión de proteínas o hidratos de carbono estructurales que afectan a las propiedades de uso final. Además, también podría combinarse un promotor de nicotina desmetilasa de tabaco con genes heterólogos incluyendo genes implicados en la biosíntesis de productos nutricionales, agentes farmacéuticos o materiales industriales. El promotor puede usarse para dirigir la regulación por disminución de genes endógenos en tabaco incluyendo nicotina desmetilasa u otros genes implicados en la biosíntesis de alcaloides y o en otras rutas.

Además, también puede usarse la región promotora (por ejemplo, SEQ ID NO:6) de un gen de nicotina desmetilasa de tabaco o la UTR en 3' (por ejemplo, SEQ ID NO:7) de un gen de nicotina desmetilasa de tabaco en cualquier método de silenciamiento génico dirigido al sitio tal como etiquetado de ADN-T, atrapamiento génico y recombinación homóloga para alterar el patrón de expresión de un gen diana, tal como se describe en el presente documento. También pueden usarse motivos promotores, que pueden identificarse fácilmente en una secuencia promotora, tal como la secuencia de SEQ ID NO:6 usando métodos convencionales en la técnica, para identificar factores que se asocian con o regulan la expresión de una nicotina desmetilasa de tabaco. Deseablemente, se usa una región promotora de la nicotina desmetilasa de tabaco u otra región reguladora de la transcripción para alterar propiedades químicas tales como el contenido en nornicotina y niveles de nitrosamina en una planta.

Además, puede usarse cualquier parte de un gen de nicotina desmetilasa de tabaco como marcador genético para aislar genes relacionados, regiones reguladoras o promotoras, o para seleccionar el gen de desmetilasa en otra especie de tabaco o *Nicotiana*.

Productos

Se preparan productos de tabaco que tienen una cantidad reducida de contenido en nitrosamina usando cualquier material de la planta de tabaco descrita en el presente documento según métodos convencionales conocidos en la técnica. En una realización, los productos de tabaco se preparan usando material vegetal obtenido de un tabaco curado genéticamente modificado que tiene una cantidad reducida de nornicotina o NNN inferior a 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 el que los porcentajes de alcaloides secundarios con respecto al contenido en alcaloides total contenidos en el mismo es inferior al 90%, el 70%, el 50%, el 30%, el 10%, el 5%, el 4%, el 3%, el 2%, el 1,5%, el 1%, el 0,75%, el 0,5%, el 0,25% o el 0,1%. Es decir, el tabaco curado se prepara a partir de una planta de tabaco genéticamente modificada. La frase "una cantidad reducida" pretende referirse a una cantidad de nornicotina o NNN o ambas en una planta de tabaco, tabaco o un producto de tabaco transgénico que es inferior a la que se encontraría en una planta de tabaco, tabaco o un producto de tabaco que se produce de manera natural de la misma variedad de tabaco, procesado de la misma manera, que no se convirtió en material transgénico para nornicotina o NNN reducida. Por tanto, en algunos contextos, se usa un tabaco que se produce de manera natural de la misma variedad que se ha procesado de la misma manera como control mediante el cual medir si se ha obtenido una reducción de nornicotina o NNN mediante los métodos descritos en el presente documento. Los niveles de nornicotina y NNN se miden según métodos bien conocidos en la técnica del tabaco.

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

Ejemplo 1

Desarrollo de tejido vegetal y tratamiento con etileno

Crecimiento vegetal

Se sembraron plantas en macetas y se hicieron crecer en un invernadero durante 4 semanas. Se transplantaron las plántulas de 4 semanas de edad a macetas individuales y se hicieron crecer en el invernadero durante 2 meses. Se regaron las plantas 2 veces al día con agua que contenía 150 ppm de fertilizante NPK durante el crecimiento. Se desprendieron las hojas verdes expandidas de las plantas para realizar el tratamiento con etileno descrito a continuación.

Línea celular 78379

Se usó la línea de tabaco 78379, que es una línea de tabaco Burley cedida por la Universidad de Kentucky como fuente de material vegetal. Se cultivaron cien plantas como patrón en la técnica de tabaco en crecimiento, se

transplantaron y se etiquetaron con un número distintivo (1-100). Se realizaron la fertilización y el tratamiento de campo tal como se recomienda.

Tres cuartos de las 100 plantas convirtieron entre el 20 y el 100% de la nicotina en nornicotina. Un cuarto de las 100 plantas convirtieron menos del 5% de la nicotina en nornicotina. La planta número 87 tenía la menor conversión (2%) mientras que la planta número 21 tenía un 100% de conversión. Se clasificaron las plantas que convertían menos del 3% como no convertidoras. Se prepararon 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 diferencias genéticas y fenotípicas. Las plantas de la 21 autofecundada eran convertidoras, y el 99% de las autofecundadas de la 87 eran no convertidoras. El otro 1% de las plantas de la 87 mostraron baja conversión (5-15%). Las plantas de cruces recíprocos eran todas convertidoras.

Línea celular 4407

La línea de *Nicotiana* 4407, que es una línea Burley, se usó como fuente de material vegetal. Se seleccionaron y etiquetaron plantas uniformes y representativas (100). De las 100 plantas, 97 eran no convertidoras y tres eran convertidoras. La planta número 56 tenía la menor cantidad de conversión (1,2%) y la planta número 58 tenía el nivel más alto de conversión (96%). Se prepararon semillas autopolinizadas y semillas cruzadas con estas dos plantas.

Se segregaron plantas de la 58 autofecundada con una razón de convertidora con respecto a no convertidora 3:1. Se identificaron las plantas 58-33 y 58-25 como líneas de plantas convertidoras y no convertidoras homocigotas, respectivamente. Se confirmó la conversión estable de 58-33 mediante análisis de su progenie.

Línea celular PBLB01

PBLB01 es una línea Burley desarrollada por ProfiGen, Inc. y se usó como fuente de material vegetal. Se seleccionó la planta convertidora a partir de semillas de fundación de PBLB01.

Procedimiento de tratamiento con etileno

Se desprendieron hojas verdes de plantas hechas crecer en invernadero de 2-3 meses y se pulverizaron con disolución de etileno al 0,3% (Prep brand Ethephon (Rhône-Poulenc)). Se colgó cada hoja pulverizada en un estante de curado equipado con humidificador y se cubrió con plástico. Durante el tratamiento, se pulverizaron periódicamente las hojas de muestra con la disolución de etileno. Aproximadamente 24-48 horas tras el tratamiento con etileno, se recogieron las hojas para la extracción de ARN. Se tomó otra submuestra para el análisis de constituyentes metabólicos para determinar la concentración de metabolitos de las hojas y constituyentes de interés más específicos tales como una variedad de alcaloides.

Como ejemplo, el análisis de alcaloides puede realizarse tal como sigue. Se agitaron muestras (0,1 g) a 150 rpm con 0,5 ml de NaOH 2 N y una disolución de extracción de 5 ml que contenía quinolina como patrón interno y metil t-butil éter. Se analizaron las muestras en un aparato HP 6890 GC equipado con un detector FID. Se usó una temperatura de 250°C para el detector y el inyector. Se usó una columna HP (30 m-0,32 mm-1 mm) que consistía en sílice fundida reticulada con un 5% de fenol y un 95% de metilsilicona a un gradiente de temperatura de 110-185°C a 10°C por minuto. Se hizo funcionar la columna a 100°C con una velocidad de flujo de $1,7 \text{ cm}^3 \text{ min}^{-1}$ con una razón de reparto de 40:1 con un volumen de inyección 2:1 usando helio como gas portador.

Ejemplo 2

Aislamiento de ARN

Para extracciones de ARN, se trataron hojas intermedias de plantas hechas crecer en invernadero de dos meses de edad con etileno tal como se describió anteriormente. Se usaron las muestras de 0 y 24-48 horas para la extracción de ARN. En algunos casos, se tomaron muestras de hojas en el proceso de senescencia de las plantas 10 días tras la eliminación del capítulo floral. Se usaron también estas muestras para la extracción. Se aisló el ARN total usando el kit RNeasy Plant Mini® (Qiagen, Inc., Valencia, California) según el protocolo del fabricante.

Se molió la muestra de tejido en nitrógeno líquido hasta obtener un polvo fino usando un mortero y mano de mortero tratados con DEPC. Se transfirieron aproximadamente 100 miligramos de tejido molido a un tubo Eppendorf de 1,5 ml estéril. Se colocó este tubo de muestra en nitrógeno líquido hasta que se recogieron todas las muestras. Entonces, se añadieron 450 µl de tampón RLT tal como se proporciona en el kit (con la adición de mercaptoetanol) a cada tubo individual. Se agitó con vórtex la muestra vigorosamente y se incubó a 56°C durante 3 minutos. Entonces, se aplicó el lisado a la columna de centrifugación QIAshredder® que descansa en un tubo de recogida de 2 ml, y se centrifugó durante 2 minutos a velocidad máxima. Se recogió la fracción no retenida y se añadió 0,5 veces el

volumen de etanol al lisado clarificado. Se mezcló bien la mezcla y se transfirió a una minicolumna de centrifugación Rneasy® que descansa en un tubo de recogida de 2 ml. Se centrifugó la muestra durante 1 minuto a 10.000 rpm. A continuación, se pipetearon 700 µl de tampón RW1 sobre la columna Rneasy® y se centrifugó durante 1 minuto a 10.000 rpm. Se pipeteó tampón RPE sobre la columna Rneasy® en un nuevo tubo de recogida y se centrifugó durante 1 minuto a 10.000 rpm. Se añadió de nuevo tampón RPE a la columna de centrifugación Rneasy® y se centrifugó durante 2 minutos a velocidad máxima para secar la membrana. Para eliminar cualquier remanente de etanol, se colocó la membrana en un tubo de recogida separado y se centrifugó durante un 1 minuto adicional a velocidad máxima. Se transfirió la columna Rneasy® a un nuevo tubo de recogida de 1,5 ml y se pipetearon 40 µl de agua libre de ARNasa directamente sobre la membrana Rneasy®. Se centrifugó este tubo de elución final durante 1 minuto a 10.000 rpm. Se analizó la calidad y cantidad del ARN total mediante gel con formaldehído desnaturalizado y espectrofotómetro.

Se aisló ARN de poli(A) usando el kit de purificación de ARN de poli A+ Oligotex® (Qiagen Inc.) siguiendo el protocolo del fabricante. Se usaron aproximadamente 200 µg de ARN total en 250 µl de volumen máximo. Se añadieron un volumen de 250 µl de tampón OBB y se añadió 15 µl de suspensión Oligotex® a los 250 µl de ARN total. Se mezcló meticulosamente el contenido pipeteando e incubando durante 3 minutos a 70°C sobre un bloque de calentamiento. Entonces se colocó la muestra a temperatura ambiente durante aproximadamente 20 minutos. Se sedimentó el complejo de Oligotex®:ARNm mediante centrifugación durante 2 minutos a velocidad máxima. Se eliminó todo a excepción de 50 µl del sobrenadante del tubo de microcentrífuga. Se trató la muestra adicionalmente mediante tampón OBB. Se resuspendió el sedimento de Oligotex®:ARNm en 400 µl de tampón OW2 mediante agitación con vórtex. Se transfirió esta mezcla sobre una columna de centrifugación pequeña colocada en un tubo nuevo y se centrifugó durante 1 minuto a velocidad máxima. Se transfirió la columna de centrifugación a un tubo nuevo y se añadieron 400 µl adicionales de tampón OW2 a la columna. Entonces, se centrifugó el tubo durante 1 minuto a máxima velocidad. Se transfirió la columna de centrifugación a un tubo de microcentrífuga de 1,5 ml final. Se eluyó la muestra con 60 µl de tampón OEB caliente (70°C). Se analizó el producto de poli A mediante geles con formaldehído desnaturalizados y análisis espectrofotométrico.

Ejemplo 3

PCR con transcripción inversa

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

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

Ejemplo 4

Generación de poblaciones de fragmentos de PCR

Se ligaron fragmentos de PCR del ejemplo 3 en un vector pGEM-T Easy (Promega, Madison, Wisconsin) siguiendo las instrucciones del fabricante. Se transformó el producto ligado en células competentes JM109 y se sembraron en placa en placas con medio LB para la selección azul/blanco. Se seleccionaron colonias y se hicieron crecer en una placa de 96 pocillos con 1,2 ml de medio LB durante la noche a 37°C. Se generó una reserva congelada para todas las colonias seleccionadas. Se purificó ADN de plásmido de las placas usando sistemas robóticos de minipreparación Biomeck 2000 de Beckman con el kit Wizard SV Miniprep (Promega). Se eluyó el ADN de plásmido con 100 µl de agua y se almacenó en una placa de 96 pocillos. Se digirieron los plásmidos mediante EcoR1 y se analizaron usando gel de agarosa al 1% para confirmar la cantidad de ADN y el tamaño de los insertos. Se secuenciaron plásmidos que contenían un inserto de 400-600 pb usando un secuenciador CEQ 2000 (Beckman, Fullerton, California). Se alinearon las secuencias con la base de datos GenBank mediante búsqueda BLAST. Se identificaron fragmentos relacionados con p450 y se analizaron adicionalmente. Alternativamente, se aislaron

fragmentos de p450 a partir de bibliotecas de sustracción. También se analizaron estos fragmentos tal como se describió anteriormente.

Ejemplo 5

Construcción de una biblioteca de ADNc

- 5 Se construyó una biblioteca de ADNc preparando ARN total a partir de hojas tratadas con etileno tal como sigue. En primer lugar, se extrajo ARN total a partir de hojas tratadas con etileno de la línea de tabaco 58-33 usando un protocolo de extracción con cloroformo y fenol ácido modificado. Se modificó el protocolo para usar un gramo de tejido que se molió y posteriormente se agitó con vórtex en 5 ml de tampón de extracción (Tris-HCl 100 mM, pH 8,5; NaCl 200 mM; EDTA 10 mM; SDS al 0,5%) al que se añadieron 5 ml de fenol (pH 5,5) y 5 ml de cloroformo. Se
- 10 centrifugó la muestra extraída y se guardó el sobrenadante. Se repitió esta etapa de extracción 2-3 veces hasta que el sobrenadante apareció transparente. Se añadieron aproximadamente 5 ml de cloroformo para eliminar cantidades traza de fenol. Se precipitó el ARN de las fracciones de sobrenadante combinadas añadiendo un volumen de 3 veces de etanol y 1/10 del volumen de NaOAc 3 M (pH 5,2) y almacenándolo a -20°C durante 1 hora. Tras transferir a un recipiente de vidrio Corex, se centrifugó la fracción de ARN a 9.000 rpm durante 45 minutos a 4°C. Se lavó el
- 15 sedimento con etanol al 70% y se centrifugó durante 5 minutos a 9.000 rpm a 4°C. Tras secar el sedimento, se disolvió el ARN sedimentado en 0,5 ml de agua libre de ARNasa. Se analizó la calidad y cantidad del ARN total mediante gel con formaldehído desnaturalizado y espectrometría, respectivamente.

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

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

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

- Se generaron los fragmentos de p450 mediante amplificación por PCR a partir de la biblioteca de ADNc construida tal como se describe en el ejemplo 3 con la excepción del cebador inverso. Se usó el cebador de T7 ubicado en el
- 40 plásmido en el sentido de 3' de los insertos de ADNc como cebador inverso. Se aislaron, se clonaron y se secuenciaron fragmentos de PCR tal como se describe en el ejemplo 4.

- Se aislaron genes de p450 de longitud completa mediante este método de PCR a partir de la biblioteca de ADNc construida. Se usaron cebadores inversos específicos de gen (diseñados a partir de la secuencia en el sentido de 3' de los fragmentos de p450) y un cebador directo (T3 en el plásmido de la biblioteca) para clonar los genes de longitud completa. Se aislaron, se clonaron y se secuenciaron fragmentos de PCR. Si es necesario, se aplicó una
- 45 segunda etapa de PCR. En la segunda etapa, se usaron nuevos cebadores directos diseñados a partir de la UTR en 5' de los P450 clonados junto con los cebadores inversos diseñados a partir de la UTR en 3' de clones de p450 en las reacciones de PCR posteriores para obtener clones de p450 de longitud completa. Posteriormente se
- 50 secuenciaron los clones.

Ejemplo 6

Caracterización de fragmentos clonados – análisis de transferencia de tipo Southern inversa

Se llevaron a cabo ensayos de transferencia de tipo Southern inversa a gran escala no radioactivos en todos los

clones de p450 identificados en ejemplos anteriores para detectar la expresión diferencial. Se observó que el nivel de expresión entre diferentes agrupamientos de p450 era muy diferente. Se llevó a cabo detección en tiempo real adicional en aquéllos con alta expresión.

Se realizaron procedimientos de transferencia de tipo Southern no radioactivos tal como sigue.

- 5 1) Se extrajo el ARN total a partir de hojas convertidoras (58-33) y no convertidoras (58-25) tratadas y no tratadas con etileno usando el kit Rnaeasy de Qiagen tal como se describe en el ejemplo 2.
- 10 2) Se produjo una sonda marcando con cola de biotina un ADNc monocatenario derivado de ARN enriquecido en poli A+ generado en la etapa anterior. Se generó este ADNc monocatenario marcado mediante RT-PCR del ARN total convertidor y no convertidor (Invitrogen) tal como se describe en el ejemplo 3 con la excepción del uso de oligo dT biotinilado como cebador (Promega). Se usaron éstos como sonda para hibridar con el ADN clonado.
- 3) Se digirió el ADN de plásmido con la enzima de restricción EcoR1 y se procesó en geles de agarosa. Se secaron simultáneamente los geles y se transfirieron a dos membranas de nailon (Biodyne B). Se hibridó una membrana con sonda convertidora y la otra con sonda no convertidora. Se reticularon las membranas con luz UV (parámetro de autorreticulación, 254 nm, Stratagene, Stratalinker) antes de la hibridación.
- 15 Alternativamente, se amplificaron los insertos mediante PCR a partir de cada plásmido usando las secuencias ubicadas en ambos brazos del plásmido p-GEM, T3 y SP6, como cebadores. Se analizaron los productos de PCR mediante procesamiento en geles de agarosa Ready-to-run de 96 pocillos. Se sometieron los insertos confirmados a transferencia puntual sobre dos membranas de nailon. Se hibridó una membrana con sonda convertidora y la otra con sonda no convertidora.
- 20 4) Se hibridaron las membranas y se lavaron siguiendo las instrucciones del fabricante con la modificación de rigurosidad de lavado (Enzo MaxSence kit, Enzo Diagnostics, Inc, Farmingdale, NY). Se hibridaron previamente las membranas con tampón de hibridación (formamida tamponada con SSC 2x, que contiene detergente y potenciadores de la hibridación) a 42°C durante 30 min. y se hibridaron con 10 µl de sonda desnaturalizada durante la noche a 42°C. Entonces, se lavaron las membranas con tapón de lavado e hibridación 1X 1 vez a temperatura ambiente durante 10 min. y 4 veces a 68°C durante 15 min. Las membranas estaban listas para el procedimiento de detección.
- 25 5) Se detectaron las membranas lavadas mediante marcaje con fosfatasa alcalina mediante detección colorimétrica con NBT/BCIP tal como se describe en el procedimiento de detección del fabricante (Enzo Diagnostics, Inc.). Se bloquearon las membranas durante una hora a temperatura ambiente con disolución de bloqueo 1x, se lavaron 3 veces con reactivos de detección 1X durante 10 min., se lavaron 2 veces con tampón de reacción de prerrevelado 1x durante 5 min. y entonces se revelaron los puntos en disolución de revelado durante 30-45 min. hasta que aparecieron los puntos. El fabricante (Enzo Diagnostics, Inc) proporcionó todos los reactivos. Además, también se realizó un ensayo de tipo Southern inverso a gran escala usando el kit de detección e hibridación KLP Southern siguiendo las instrucciones del fabricante (KPL, Gaithersburg, Maryland).

35 Ejemplo 7

Caracterización de clones - Análisis de transferencia de tipo Northern

Como alternativa al análisis de transferencia de tipo Southern, se hibridaron algunas membranas y se detectaron tal como se describe en el ejemplo de ensayos de transferencia tipo Northern. Se usó hibridación de tipo Northern para detectar ARNm expresado de manera diferencial en *Nicotiana* tal como sigue.

- 40 Se usó un método de cebado al azar para preparar sondas a partir de p450 clonado (sistemas de marcaje de ADN Megaprime, Amersham Biosciences). Se mezclaron los siguientes componentes: 25 ng de molde de ADN desnaturalizado; 4 µl de cada dTTP, dGTP y dCTP no marcado; 5 µl de tampón de reacción; dATP marcado con P³² y 2 µl de Klenow I; y H₂O, para llevar la reacción a 50 µl. Se incubó la mezcla en 37°C durante 1-4 horas, y se detuvo con 2 µl de EDTA 0,5 M. Se desnaturalizó la sonda mediante incubación a 95°C durante 5 minutos antes de usar.
- 45

Se prepararon muestras de ARN a partir de hojas recientes tratadas y no tratadas con etileno de varios pares de líneas de tabaco. En algunos casos, se usó ARN enriquecido en poli A+. Se llevaron aproximadamente 15 µg de ARN total o 1,8 µg de ARNm (métodos de extracción de ARN y ARNm tal como se describe en el ejemplo 5) a igual volumen con DEPC H₂O (5-10 µl). Se añadieron el mismo volumen de tampón de carga (MOPS 1x; formaldehído al 18,5%; formamida al 50%; Ficoll 400 al 4%; azul de bromofenol) y 0,5 µl de EtBr (0,5 µg/µl). Posteriormente se desnaturalizaron las muestras en la preparación para la separación del ARN mediante electroforesis.

- 50

Se sometieron a electroforesis las muestras sobre un gel con formaldehído (agarosa al 1%, MOPS 1x, formaldehído 0,6 M) con tampón MOP 1X (ácido morfolinopropanosulfónico 0,4 M; Na-acetato-3 x H₂O 0,1 M; EDTA 10 mM; ajustar a pH 7,2 con NaOH). Se transfirió el ARN a una membrana Hybond-N+ (nylon, Amersham Pharmacia Biotech) mediante el método de capilaridad en tampón SSC 10X (NaCl 1,5 M; Na-citrato 0,15 M) durante 24 horas.

5 Se reticularon con UV muestras de membranas con ARN (parámetro de autorreticulación, 254 nm, Stratagene, Stratallinker) antes de la hibridación.

Se hibridó previamente la membrana durante 1-4 horas a 42°C con 5-10 ml de tampón de prehibridación (SSC 5x; formamida al 50%; disolución de Denhardt 5x; SDS al 1%; ADN no homólogo fragmentado desnaturalizado por calor 1 00 µg/ml). Se desechó el tampón de prehibridación antiguo, y se añadieron nuevo tampón de prehibridación y sonda. Se llevó a cabo la hibridación durante la noche a 42°C. Se lavó la membrana durante 15 minutos con SSC 2x a temperatura ambiente, seguido por un lavado con SSC 2x.

10 Se llevó a cabo análisis de tipo Northern usando clones de longitud completa en tejido de tabaco obtenido de líneas Burly convertidoras y no convertidoras que se indujeron por el tratamiento con etileno. El fin era identificar los clones de longitud completa que mostraban expresión elevada en líneas convertidoras inducidas con etileno con respecto a

15 líneas convertidoras inducidas con etileno con respecto a líneas Burley no convertidoras inducidas con etileno. Haciendo esto, puede determinarse la relación de funcionalidad de clones de longitud completa comparando diferencias bioquímicas en constituyentes de la hoja entre líneas convertidoras y no convertidoras.

Ejemplo 8

Immunodetección de p450 codificados por los genes clonados

20 Se seleccionaron regiones peptídicas correspondientes a 20-22 aminoácidos de longitud de tres clones de p450 para que 1) tuviesen inferior homología o no tuviesen homología con otros clones y 2) tuviesen buena hidrofiliidad y antigenicidad. La secuencia de aminoácidos de las regiones peptídicas seleccionadas de los clones respectivos de p450 se enumeran a continuación. Se conjugaron los péptidos sintetizados con KHL y entonces se inyectaron en conejos. Se recogieron antisueros 2 y 4 semanas tras la 4^a inyección (Alpha Diagnostic Intl. Inc. San Antonio, TX).

25 D234-AD1 DIDGSKSLVKAHRKIDEILG (SEQ ID NO: 8)

D90a-BB3 RDAFREKETFDENDVEELNY (SEQ ID NO: 9)

D89-AB1 FKNNGDEDRHFSQKLGLADKY (SEQ ID NO: 10)

30 Se examinaron los antisueros para determinar la reactividad cruzada con proteínas diana de tejido de planta de tabaco mediante análisis de inmunotransferencia de tipo Western. Se obtuvieron extractos de proteínas brutos a partir de hojas intermedias tratadas con etileno (de 0 a 40 horas) de líneas convertidoras y no convertidoras. Se determinaron las concentraciones de proteína de los extractos usando kit de ensayo de proteína RC DC (BIO-RAD) siguiendo el protocolo del fabricante.

35 Se cargaron dos microgramos de proteína en cada carril y se separaron las proteínas en geles con gradiente del 10% - 20% usando el sistema de SDS-PAGE Laemmli. Se transfirieron las proteínas desde los geles hasta membranas de transferencia de nitrocelulosa PROTRAN (Schleicher & Schuell) con la celda Trans-Blot Semi-Dry (BIO-RAD). Se detectaron proteínas de p450 diana y se visualizaron con el kit de detección de inmunotransferencia de tipo Western ECL Advance (Amersham Biosciences). Se prepararon en conejos anticuerpos primarios frente a los conjugados de péptidos sintéticos-KLH. Se adquirieron de Sigma anticuerpos secundarios frente a IgG de conejo, acoplados con peroxidasa. Se usaron los anticuerpos tanto primario como secundario a diluciones 1:1000. Los anticuerpos mostraron fuerte reactividad frente a una única banda en las inmunotransferencias de tipo Western indicando que los antisueros eran monoespecíficos con respecto al péptido diana de interés. Los antisueros también reaccionaban de manera cruzada con péptidos sintéticos conjugados con KLH.

40

Ejemplo 9

Identidad de ácido nucleico y relación de estructura de fragmentos de ácido nucleico aislados

45 Se secuenciaron más de 100 fragmentos de p450 clonados conjuntamente con el análisis de transferencia de tipo Northern para determinar su relación estructural. El enfoque usó cebadores directos basados en cualquiera de dos motivos de p450 comunes ubicados cerca del extremo carboxilo terminal de los genes de p450. Los cebadores directos correspondían a motivos de citocromo p450 FXPERF (SEQ ID NO:11) o GRRXCP(A/G) (SEQ ID NO:12). Los cebadores inversos usaban cebadores convencionales de o bien el plásmido, SP6 o bien T7 ubicados en ambos

50 brazos del plásmido pGEM, o una cola de poli A. El protocolo usado se describe a continuación.

Se usó espectrometría para estimar la concentración del ADN bicatenario de partida siguiendo el protocolo del fabricante (Beckman Coulter). Se diluyó el molde con agua a la concentración apropiada, se desnaturalizó mediante calentamiento a 95°C durante 2 minutos y posteriormente se colocó en hielo. Se preparó en hielo la reacción de secuenciación usando de 0,5 a 10 µl de molde de ADN desnaturalizado, 2 µl de 1,6 pmoles del cebador directo, 8 µl de DTCS Quick Start Master Mix y el volumen total se llevó hasta 20 µl con agua. El programa de termociclado consistió en 30 ciclos del siguiente ciclo: 96°C durante 20 segundos, 50°C durante 20 segundos y 60°C durante 4 minutos seguido por mantenimiento a 4°C.

Se detuvo la reacción de secuenciación añadiendo 5 µl de tampón de parada (volumen igual de NaOAc 3 M y EDTA 100 mM y 1 µl de glicógeno 20 mg/ml). Se precipitó la muestra con 60 µl de etanol al 95% frío y se centrifugó a 6000 x g durante 6 minutos. Se desechó el etanol. Se lavó el sedimento 2 veces con 200 µl de etanol al 70% frío. Tras secar el sedimento, se añadieron 40 µl de disolución de SLS y se resuspendió el sedimento. Se cubrió con una capa de aceite mineral. Entonces, la muestra se colocó en el secuenciador automatizado CEQ 8000 para su análisis adicional.

Con el fin de verificar las secuencias de ácido nucleico, se volvió a secuenciar la secuencia de ácido nucleico en ambas direcciones usando cebadores directos para la región FXPERF (SEQ ID NO:11) o GRRXCP(A/G) (SEQ ID NO:12) del gen de p450 o cebadores inversos para o bien el plásmido o bien la cola de poli A. Se llevó a cabo toda la secuenciación al menos dos veces en ambas direcciones.

Se compararon las secuencias de ácido nucleico de fragmentos de citocromo p450 entre sí desde la región codificante correspondiente al primer ácido nucleico tras la región que codifica para el motivo GRRXCP(A/G) (SEQ ID NO:12) hasta el codón de terminación. Se seleccionó esta región como indicador de diversidad genética entre las proteínas p450. Se observó un gran número de genes de p450 genéticamente distintos, en exceso de 70 genes, similar al de otras especies de plantas. Con la comparación de secuencias de ácido nucleico, se encontró que los genes podían colocarse en distintos grupos de secuencias basados en su identidad de secuencia. Se encontró que la mejor agrupación única de miembros de p450 se determinaba que eran las secuencias con una identidad de ácido nucleico del 75% o superior. (Véase por ejemplo, la tabla 1 de la publicación de solicitud de patente US 2004/0162420).

La reducción del porcentaje de identidad dio como resultado grupos significativamente más grandes. Se observó una agrupación preferida para las secuencias con una identidad de ácido nucleico del 81% o superior, una agrupación más preferida con una identidad de ácido nucleico del 91% o superior y una agrupación lo más preferida para las secuencias con una identidad de ácido nucleico del 99% o superior. La mayoría de los grupos contenían al menos dos miembros y frecuentemente tres o más miembros. Otros no se descubrieron repetidamente, lo que sugiere que el enfoque tomado podía aislar ARNm tanto de alta como de baja expresión en el tejido usado.

Usando la tecnología de chips génicos para identificar genes que se expresan de manera diferencial en líneas de tabaco convertidoras frente a no convertidoras, se determinó que D121-AA8 tenía inducción reproducible en líneas convertidoras tratadas con etileno. Basándose en estos resultados, se identificó el gen D121-AA8 (cuya secuencia de ADNc es la secuencia de SEQ ID NO:3; figura 3) como el gen de nicotina desmetilasa de tabaco de interés.

En vista de la regla de nomenclatura de p450, el gen de nicotina desmetilasa de tabaco es novedoso y pertenece a la clase CYP82E (The Arabidopsis Genome Initiative (AGI) y The Arabidopsis Information Resource (TAIR); Frank, Plant Physiol. 110:1035-1046, 1996; Whitbred *et al.*, Plant Physiol. 124:47-58, 2000); Schopfer y Ebel, Mol. Gen. Genet. 258: 315-322, 1998; y Takemoto *et al.*, Plant Cell Physiol. 40:1232-1242, 1999).

Ejemplo 10

Análisis bioquímico de la nicotina desmetilasa de tabaco

Los análisis bioquímicos, por ejemplo, tal como se describió en solicitudes presentadas previamente, determinaron que la secuencia de SEQ ID NO:3 codifica para una nicotina desmetilasa de tabaco (SEQ ID NO:4; figura 3).

En particular, se confirmó la función del clon candidato (D121-AA8), como el gen codificante para nicotina desmetilasa, sometiendo a ensayo la actividad enzimática de p450 expresado de manera heteróloga en células de levadura tal como sigue.

1. Construcción del vector de expresión de levadura

Se clonó la supuesta secuencia codificante de proteína del ADNc que codifica para la nicotina desmetilasa de tabaco (D121AA8) en el vector de expresión de levadura pYeDP60. Se introdujeron sitios BamHI y MfeI apropiados (subrayados a continuación) mediante cebadores de PCR que contienen estas secuencias o bien en el sentido de 5'

del codón de inicio de la traducción (ATG) o bien en el sentido de 3' del codón de terminación (TAA). El MfeI en el producto de PCR amplificado es compatible con el sitio EcoRI en el vector. Los cebadores usados para amplificar el ADNc fueron 5'-TAGCTACGCGGATCCATGCTTTCTCCCATAGAAGCC-3' (SEQ ID NO:27) y 5'-CTGGATCACAAATTGTTAGTGATGGTGATGGTGATGCGATCCTCTATAAAG CTCAGGTGCCAGGC-3' (SEQ ID NO:28). Se incorporó un segmento de secuencia que codifica para nueve aminoácidos extra en el extremo C-terminal de la proteína, que incluye seis histidinas, en el cebador inverso para facilitar la expresión de p450 etiquetado con 6-His tras la inducción. Se ligaron los productos de PCR en el vector YeDP60 tras digestiones enzimáticas en la orientación sentido con referencia al promotor GAL10-CYC1. Se verificó la construcción apropiada del vector de expresión de levadura mediante análisis con enzimas de restricción y secuenciación de ADN.

10 2. Transformación de levadura

Se transformó la línea de levadura WAT11, modificada para expresar NADPH-citocromo p450 reductasa ATR1 de *Arabidopsis*, con los plásmidos de ADNc pYeDP60-p450. Se mezclaron cincuenta microlitros de suspensión celular de levadura WAT11 con ~1 µg de ADN de plásmido en una cubeta con separación de electrodo de 0,2 cm. Se aplicó un pulso a 2,0 kV mediante un electroporador Eppendorf (modelo 2510). Se extendieron las células sobre placas de SGI (bactocasaminoácidos 5 g/l, base nitrogenada de levadura sin aminoácidos 6,7 g/l, glucosa 20 g/l, DL-triptófano 40 mg/l, agar 20 g/l). Se confirmaron los transformantes mediante análisis de PCR llevado a cabo directamente en colonias seleccionadas al azar.

3. Expresión de p450 en células de levadura transformadas

Se usaron colonias de levadura individuales para inocular 30 ml de medio SGI (bactocasaminoácidos 5 g/l, base nitrogenada de levadura sin aminoácidos 6,7 g/l, glucosa 20 g/l, DL-triptófano 40 mg/l) y se hicieron crecer a 30°C durante aproximadamente 24 horas. Se diluyó 1:50 una alícuota de este cultivo en 1000 ml de medio YPGE (extracto de levadura 10 g/l, bactopectona 20 g/l, glucosa 5 g/l, etanol 30 ml/l) y se hicieron crecer hasta que se consumió completamente la glucosa tal como se indica mediante el cambio colorimétrico de una tira reactiva de urianálisis Diastix (Bayer, Elkhart, IN). Se inició la inducción de p450 clonado añadiendo DL-galactosa hasta una concentración final del 2%. Se hicieron crecer los cultivos durante 20 horas adicionales antes de usarse para el ensayo de la actividad *in vivo* o para la preparación de microsomas.

De usaron células de levadura WAT11 que expresan pYeDP60-CYP71D20 (un p450 que cataliza la hidroxilación de 5-epi-aristoloqueno y 1-desoxicapsidiol en *Nicotiana tabacum*) como control para la expresión de p450 y ensayos de actividad enzimática.

Para evaluar la eficacia de la expresión en levadura del p450 en gran detalle, se llevó a cabo espectroscopia de diferencia de CO reducido. El espectro de CO reducido presentaba un pico a 450 nm de proteínas de las cuatro líneas de levadura transformadas con p450. No se observaron picos similares en los microsomas de la levadura control o la levadura control con vector. Los resultados indicaron que las proteínas p450 se expresaban eficazmente en las líneas de levaduras que albergaban el pYeDP60-CYP 450. Las concentraciones de proteína p450 expresada en microsoma de levadura oscilaba entre 45 y 68 nmoles/mg de proteína total.

4. Ensayo enzimático *in vivo*

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

El cultivo control de WAT11 (pYeDP60-CYP71D20) no convirtió la nicotina en nornicotina, lo que muestra que la cepa de levadura WAT11 no contiene actividades enzimáticas endógenas que puedan catalizar la etapa de bioconversión de nicotina en nornicotina. En cambio, la levadura que expresa el gen de nicotina desmetilasa de tabaco produjo una cantidad detectable de nornicotina, indicando la actividad nicotina desmetilasa del producto de traducción de SEQ ID NO:3.

5. Preparación de microsomas de levadura

Tras la inducción con galactosa durante 20 horas, se recogieron células de levadura mediante centrifugación y se lavaron dos veces con tampón TES-M (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, sorbitol 0,6 M, 2-mercaptoetanol 10 mM). Se resuspendió el sedimento en tampón de extracción (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, sorbitol 0,6 M, 2-mercaptoetanol 2 mM, albúmina sérica bovina al 1%, cóctel de inhibidores de proteasas (Roche) a 1 comprimido/50 ml). Entonces se rompieron las células con perlas de vidrio (0,5 mm de diámetro, Sigma) y se centrifugó el extracto

celular durante 20 min. a 20.000 x g para eliminar residuos celulares. El sobrenadante se sometió a ultracentrifugación a 100.000 x g durante 60 min. y el sedimento resultante contenía la fracción microsomal. Se suspendió la fracción microsomal en tampón TEG-M (Tris-HCl 50 mM, pH 7,5, EDTA 1 mM, glicerol al 20% y 2-mercaptoetanol 1,5 mM) a una concentración de proteína de 1 mg/ml. Se almacenaron las preparaciones

6. Ensayo de actividad enzimática en preparaciones microsomales de levadura

Se llevaron a cabo ensayos de actividad nicotina desmetilasa con preparaciones microsomales de levadura. En particular, se obtuvo DL-Nicotina (pirrolidina-2-¹⁴C) de Moravek Biochemicals y tenía una actividad específica de 54 mCi/mmol. Se adquirieron clorpromazina (CPZ) y citocromo c oxidado (cit. C), ambos inhibidores de P450, de Sigma. La forma reducida de nicotinamida adenina dinucleótido fosfato (NADPH) es el donador de electrones típico para citocromo P450 mediante la NADPH:citocromo P450 reductasa. Se omitió NADPH para la incubación control. El ensayo enzimático de rutina incluía proteínas microsomales (aproximadamente 1 mg/ml), NADPH 6 mM y nicotina marcada con ¹⁴C 55 µM. La concentración de CPZ y cit. C, cuando se usaron, fue de 1 mM y 100 µM, respectivamente. La reacción se llevó a cabo a 25°C durante 1 hora y se detuvo con la adición de 300 µl de metanol a cada mezcla de reacción de 25 µl. Tras la centrifugación, se separaron 20 µl del extracto de metanol con un sistema de cromatografía líquida de alta resolución (HPLC) (Agilent) de fase inversa usando una columna cromatográfica Inertsil ODS-3 3 µ (150 x 4,6 mm) de Varian. La fase móvil isocrática era la mezcla de metanol y tampón fosfato de potasio 50 mM, pH 6,25, con una razón de 60:40 (v/v) y la velocidad de flujo era de 1 ml/min. Se recogió el pico de nornicotina, tal como se determinó mediante comparación con nornicotina no marcada auténtica y se sometió a contador de centelleo líquido 2900 tri-carb (LSC) (Perkin Elmer) para la cuantificación. Se calculó la actividad de la nicotina desmetilasa basándose en la producción de nornicotina marcada con ¹⁴C durante 1 hora de incubación.

Las preparaciones microsomales a partir de células de levadura control que expresaban CYP71D20 no tenían ninguna actividad nicotina desmetilasa microsomal. Por el contrario, las muestras microsomales obtenidas de células de levadura que expresaban el gen de nicotina desmetilasa de tabaco mostraron niveles significativos de actividad nicotina desmetilasa. La actividad nicotina desmetilasa requería NADPH y se mostró que se inhibía por inhibidores específicos de p450, lo que concuerda con que la nicotina desmetilasa de tabaco es un p450. La actividad enzimática para nicotina desmetilasa de tabaco (D121-AA8) era de aproximadamente 10,8 pKat/mg de proteína tal como se calculó mediante la intensidad radioactiva y las concentraciones de proteína. Un conjunto típico de resultados de ensayos enzimáticos obtenidos para las células de levadura se muestra en la tabla 1.

TABLA 1: ACTIVIDAD DESMETILASA EN MICROSOMAS DE CÉLULAS DE LEVADURA QUE EXPRESAN D121-AA8 y P450 CONTROL

Muestra	Microsomos	Microsomos + clorpromazina 1 mM	Microsomos + citocromo C 100 µM	Microsomos - NADPH
D121-AA8	10,8 ± 1,2* pkat/mg de proteína	1,4 ± 1,3 pkat/mg de proteína	2,4 ± 0,7 pkat/mg de proteína	0,4 ± 0,1 pkat/mg de proteína
Control (CYP71D20)	No detectado	No detectado	No detectado	No detectado
*Resultados promedio de 3 réplicas.				

Conjuntamente, estos experimentos demostraron que la nicotina desmetilasa de tabaco de longitud completa clonada (SEQ ID NO:3; D121-AA8) codifica para una proteína de citocromo p450 que cataliza la conversión de nicotina en nornicotina cuando se expresa en levadura.

Ejemplo 11

Identidad de secuencia de aminoácidos relacionada de fragmentos de ácido nucleico aislados

Se dedujeron las secuencias de aminoácidos de secuencias de ácido nucleico obtenidas para fragmentos de citocromo p450 del ejemplo 8. La región deducida correspondía al aminoácido inmediatamente tras el motivo de secuencia GXRXC(P/A/G) (SEQ ID NO:13) hasta el extremo carboxilo terminal, o codón de terminación. Al comparar la identidad de secuencia de los fragmentos, se observó una agrupación única para las secuencias con una identidad de aminoácidos del 70% o superior. Se observó una agrupación preferida para las secuencias con una

identidad de aminoácidos del 80% o superior, más preferida con una identidad de aminoácidos del 90% o superior, y una agrupación lo más preferida para las secuencias con una identidad de aminoácidos del 99% o superior. Se encontró que varias de las secuencias de ácido nucleico únicas tienen identidad de aminoácidos completa con respecto a otros fragmentos y por tanto sólo se notificó un miembro con el aminoácido idéntico.

- 5 Se seleccionó al menos un miembro de cada grupo de identidad de aminoácidos para estudios funcionales y de clonación génica usando plantas. Además, se seleccionaron miembros de los grupos que se ven afectados de manera diferencial por el tratamiento con etileno u otras diferencias biológicas tal como se evaluaron mediante análisis de tipo Northern y Southern para estudios funcionales y de clonación génica. Para ayudar en la clonación génica, estudios de expresión y evaluaciones de plantas completas, pueden prepararse anticuerpos específicos de péptido basándose en la identidad de secuencia y la secuencia diferencial.

Ejemplo 12

Identidad de secuencia de aminoácidos relacionada de clones de longitud completa

- Se dedujo la secuencia de ácido nucleico de genes de *Nicotiana* de longitud completa clonados en el ejemplo 5 para determinar su secuencia de aminoácidos entera. Se identificaron genes de citocromo p450 mediante la presencia de tres motivos de dominio p450 conservados, que correspondían a UXXRXXZ (SEQ ID NO:14), PXRFXF (SEQ ID NO:15) o GXRXC (SEQ ID NO:16) en el extremo carboxilo terminal en los que U es E o K, X es cualquier aminoácido y Z es P, T, S o M. Se caracterizaron todos los genes de p450 para determinar su identidad de aminoácidos usando un programa BLAST que comparaba sus secuencias de longitud completa entre sí y frente a genes de tabaco conocidos. El programa usó la herramienta BLAST especial del NCBI (alineación de dos secuencias (b12seq), <http://www.ncbi.nlm.nih.gov/blast/b12seq/b12.html>). Se alinearon dos secuencias en BLASTN sin filtro para las secuencias de ácido nucleico y BLASTP para las secuencias de aminoácidos. Basándose en su porcentaje de identidad de aminoácidos, se agrupó cada secuencia en grupos de identidad en los que la agrupación contenía miembros que compartían una identidad de al menos el 85% con otro miembro. Se observó una agrupación preferida para las secuencias con una identidad de aminoácidos del 90% o superior, una agrupación más preferida tenía una identidad de aminoácidos del 95% o superior y una agrupación lo más preferida tenía las secuencias con una identidad de aminoácidos del 99% o superior. Se dedujo que la secuencia de aminoácidos del gen nicotina desmetilasa de longitud completa tenía la secuencia proporcionada en SEQ ID NO:4 (figura 3).

Ejemplo 13

Clones de citocromo p450 de *Nicotiana* que carecen de uno o más de los dominios específicos de P450 de tabaco

- 30 Cuatro clones tenían alta homología de ácido nucleico, que oscilaba entre el 90% y el 99% de homología de ácido nucleico. Sin embargo, debido a un desplazamiento del marco de lectura de nucleótidos estos genes no contenían uno o más de los tres dominios de citocromo p450 del extremo C terminal y se excluyeron de los grupos de identidad.

Ejemplo 14

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

- El uso de fragmentos de ácido nucleico de p450 o genes enteros de tabaco es útil en la identificación y selección de las plantas que tienen fenotipos de tabaco o constituyentes de tabaco alterados y, de manera más importante, metabolitos alterados. Se generan plantas de tabaco transgénicas mediante una variedad de sistemas de transformación que incorporan fragmentos de ácido nucleico o genes de longitud completa, seleccionados de los notificados en el presente documento, en orientaciones para o bien regulación por disminución, por ejemplo orientación antisentido, o bien sobreexpresión, por ejemplo, orientación sentido y similares. Para la sobreexpresión en genes de longitud completa, es deseable cualquier secuencia de ácido nucleico que codifica para la totalidad o una parte funcional o secuencia de aminoácidos de los genes de longitud completa descritos en esta invención. Tales secuencias de ácido nucleico son eficaces deseablemente para aumentar la expresión de una determinada enzima y por tanto dan como resultado un efecto fenotípico dentro de *Nicotiana*. Se obtienen líneas de *Nicotiana* que son homocigotas a través de una serie de retrocruzamientos y se evalúan para determinar cambios fenotípicos incluyendo, pero sin limitarse a, análisis de ARN de p450 endógeno, transcritos, péptidos expresados de p450 y concentraciones de metabolitos vegetales usando técnicas comúnmente disponibles para un experto habitual en la técnica. Los cambios presentados en las plantas de tabaco proporcionan información sobre el papel funcional del gen de interés seleccionado o son de uso como especies de plantas de *Nicotiana* preferidas.

Ejemplo 15Clonación de la nicotina desmetilasa de tabaco genómica a partir de tabaco Burley convertidor

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

- 5 Se diseñaron los cebadores basándose en la región promotora 5' y UTR en 3' clonadas en ejemplos anteriores. Los cebadores directos fueron 5'-GGC TCT AGA TAA ATC TCT TAA GTT ACT AGG TTC TAA-3' (SEQ ID NO:17) y 5'-TCT CTA AAG TCC CCT TCC-3' (SEQ ID NO:25) y los cebadores inversos fueron 5'-GGC TCT AGA AGT CAA TTA TCT TCT ACA AAC CTT TAT ATA TTA GC-3' (SEQ ID NO:18), y 5'-CCA GCA TTC CTC AAT TTC-3' (SEQ ID NO:26). Se aplicó PCR al ADN genómico 4407-33 con 100 µl de mezcla de reacción. Se usó la enzima de alta fidelidad Pfx para la amplificación por PCR. Se visualizó el producto de PCR en gel de agarosa al 1% tras la electroforesis. Se observó una única banda con un peso molecular de aproximadamente 3,5 kb y se cortó del gel. Se purificó la banda resultante usando un kit de purificación de gel (Qiagen; basándose en el procedimiento del fabricante). Se digirió el ADN purificado mediante la enzima Xba I (NEB; usada según las instrucciones del fabricante). Se digirió el plásmido pBluescript con Xba I usando el mismo procedimiento. Se purificó en gel el fragmento y se ligó en el plásmido pBluescript. Se transformó la mezcla de ligación en la célula competente GM109 y se sembró en placa sobre placas con LB que contenían 100 mg/l de ampicilina con selección azul/blanco. Se recogieron las colonias blancas y se hicieron crecer en 10 ml de medio líquido LB que contenía ampicilina. Se extrajo el ADN mediante miniprep. Se secuenció el ADN de plásmido que contenía el inserto usando un secuenciador CEQ 2000 (Beckman, Fullerton, California) basándose en el procedimiento del fabricante. Se los cebadores de T3 y T7 y otros 8 cebadores internos para secuenciar. Se ensambló la secuencia y se analizó, proporcionando así la secuencia genómica (SEQ ID NO:1; figuras 2-1 a 2-3).

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

Ejemplo 16Clonación de secuencias flanqueantes en 5' (SEQ ID NO:6) y UTR en 3' (SEQ ID NO:7) a partir de tabaco convertidor

- 35 A. Aislamiento de ADN total a partir de tejido de hojas de tabaco convertidor

- Se aisló ADN genómico a partir de hojas de tabaco convertidor 4407-33. Se llevó a cabo el aislamiento del ADN usando un kit DNeasy Plant Mini de la compañía Qiagen, Inc. (Valencia, Ca) según el protocolo del fabricante (véase manual del fabricante Dneasy' Plant Mini and DNeasy Plant Maxi Handbook, Qiagen enero de 2004). El procedimiento para la preparación del ADN incluía las siguientes etapas: se molió tejido de hojas de tabaco (aproximadamente 20 mg de peso seco) para dar un polvo fino bajo nitrógeno líquido durante 1 minuto. Se transfirió el polvo de tejido a un tubo de 1,5 ml. Se añadieron tampón AP1 (400 µl) y 4 µl de disolución madre de ARNasa (100 mg/ml) a un máximo de 100 mg de tejido de hoja molido y se agitó con vortex vigorosamente. Se incubó la mezcla durante 10 min. a 65°C y se mezcló 2-3 veces durante la incubación mediante inversión del tubo. Entonces, se añadió al lisado tampón AP2 (130 µl). Se mezcló la mezcla y se incubó durante 5 min. sobre hielo. Se aplicó el lisado a una columna QIAshredder Mini Spin y se centrifugó durante 2 min. (14,000 rpm). Se transfirió la fracción no retenida a un nuevo tubo sin alterar el sedimento de residuo celular. Entonces, se añadió tampón AP3/E (1,5 volúmenes) al lisado clarificado y se mezcló con pipeteo. Se aplicó la mezcla (650 µl) de la etapa precedente incluyendo cualquier precipitado a una columna DNeasy Mini Spin. Se centrifugó la mezcla durante 1 min. a >6000 x g (>8000 rpm) y se desechó la fracción no retenida. Esto se repitió con la muestra restante y se desecharon la fracción no retenida y el tubo de recogida. Se colocó la columna DNeasy Mini Spin en un nuevo tubo de recogida de 2 ml. Entonces, se añadió tampón AW (500 µl) a la columna DNeasy y se centrifugó durante 1 min. (>8000 rpm). Se desechó la fracción no retenida. Se volvió a usar el tubo de recogida en la siguiente etapa. Entonces, se añadió tampón AW (500 µl) a la columna DNeasy y se centrifugó durante 2 min. (>14.000 rpm) con el fin de secar la membrana. Se transfirió la columna DNeasy a un tubo de 1,5 ml. Entonces, se pipeteó tampón AE (100 µl) sobre la membrana DNeasy. Se incubó la mezcla durante 5 min. a temperatura ambiente (15-25°C) y entonces se centrifugó

durante 1 min. (>8000 rpm) para eluir.

Se estimó la calidad y la cantidad del ADN procesando las muestras en un gel de agarosa.

B. Clonación de secuencias flanqueantes en 5' del gen estructural

Se usó un método de PCR inversa modificado para clonar 750 nucleótidos de las secuencias flanqueantes en 5' del gen estructural de SEQ ID NO:1. En primer lugar, se seleccionaron enzimas de restricción apropiadas basándose en el sitio de restricción en el fragmento de secuencia conocida y la distancia de los sitios de restricción en el sentido de 3' de las secuencias flanqueantes en 5'. Se diseñaron los cebadores basándose en este fragmento conocido. Se ubicó el cebador directo en el sentido de 3' del cebador inverso. Se ubicó el cebador inverso en la parte 3' del fragmento conocido.

10 El procedimiento de clonación incluyó las siguientes etapas:

Se digirió el ADN genómico purificado (5 µg) con 20-40 unidades de la enzima de restricción apropiada (EcoRI y SpeI) en una mezcla de reacción de 50 µl. Se llevó a cabo una electroforesis en gel de agarosa con un volumen 1/10 de la mezcla de reacción para determinar si se digirió el ADN hasta el final. Se llevó a cabo una ligación directa tras la digestión completa mediante ligación durante la noche a 4°C. Se ligó una mezcla de reacción de 200 µl que contenía 10 µl de ADN digerido y 0,2 µl de ADN ligasa de T4 (NEB) durante la noche a 4°C. Se llevó a cabo PCR sobre la reacción de ligación tras obtenerse un genoma circular artificial pequeño. Se llevó a cabo la PCR con 10 µl de reacción de ligación y 2 cebadores de fragmentos conocidos en dos direcciones diferentes en 50 µl de mezcla de reacción. Se aplicó un programa de PCR en gradiente con temperaturas de apareamiento de 45-56°C.

20 Se llevó a cabo electroforesis en gel de agarosa para comprobar la reacción de PCR. Se cortó la banda deseada del gel y se usó un kit de purificación de gel QIAquick de QIAGEN para purificar la banda. Se ligaron fragmentos de PCR purificados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se extrajeron los plásmidos de ADN transformados mediante miniprep. usando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se secuenció el ADN de plásmido que contenía el inserto usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Se clonaron aproximadamente 758 nt (nucleótidos 1241-2009 de SEQ ID NO:1) de la secuencia flanqueante en 5' mediante el método descrito anteriormente.

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

Se usó el kit BD GenomeWalker Universal (Clontech laboratories, Inc., PaloAlto, CA) para clonar una secuencia flanqueante en 5' adicional del gen estructural, D121-AA8 según el manual de usuario del fabricante (véase el manual del fabricante BD GenomeWalker agosto de 2004). Se sometieron a prueba el tamaño y la pureza del ADN genómico de tabaco procesando muestras en un gel de agarosa al 0,5%. Se establecieron un total de 4 reacciones de extremos romos (DRA I, STU I, ECOR V, PVU II) para la construcción de paseo genómico de 33 bibliotecas de tabaco. Tras la purificación de los ADN digeridos, se ligaron los ADN genómicos digeridos al adaptador para paseo genómico. Se aplicaron reacciones de PCR primarias a los cuatro ADN digeridos usando cebador adaptador AP1 y el cebador específico de gen de D121-AA8 (CTCTATTGATACTAGCTGGTTTTGGAC; SEQ ID NO: 19). Se usaron los productos de PCR primarios directamente como moldes para la PCR anidada. Se usaron el cebador anidado adaptador proporcionado por el kit y el cebador anidado del clon conocido D121-AA8 (SEQ ID NO:3) (GGAGGGAGAGTATAACTTACGGATTC; SEQ ID NO:20) en la reacción de PCR. Se comprobaron los productos de PCR mediante procesamiento por electroforesis en gel. Las bandas deseadas se cortaron del gel, y se purificaron los fragmentos de PCR usando el kit de purificación de gel QIAquick de QIAGEN. Se ligaron los fragmentos de PCR purificados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se extrajeron los plásmidos de ADN transformados mediante miniprep. usando el kit SV Miniprep (Promega, Madison, WI) y siguiendo las instrucciones del fabricante. Se secuenció el ADN de plásmido que contenía el inserto usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Se clonaron otros aproximadamente 853 nt de la secuencia flanqueante en 5', incluyendo los nucleótidos 399-1240 de SEQ ID NO:1, mediante el método descrito anteriormente.

45 Se realizó una segunda ronda del paseo genómico según el mismo método con la diferencia de que se usaron los siguientes cebadores GWR1A (5'-AGTAACCGATTGCTCACGTATCCTC-3') (SEQ ID NO:21) y GWR2A (5'-CTCTATTCAACCCACACGTAAGT-3') (SEQ ID NO:22). Se clonaron otros aproximadamente 398 nt de la secuencia flanqueante, incluyendo los nucleótidos 1-398 de SEQ ID NO:1, mediante este método.

50 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 de órgano en la región promotora de la nicotina desmetilasa de tabaco. También están presentes supuestos elementos de respuesta a estimulantes y elementos regulados por nitrógeno, identificados usando métodos convencionales, en la región promotora.

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

Se usó el kit BD GenomeWalker Universal (Clontech laboratories, Inc., PaloAlto, CA) para clonar la secuencia flanqueante en 3' del gen estructural, D121-AA8 según el manual de usuario del fabricante. El procedimiento de clonación es el mismo que se describe en la sección C precedente de este ejemplo, excepto por los cebadores específicos de gen. Se diseñó el primer cebador a partir de las proximidades al extremo del gen estructural D121-AA8 (5'-CTA AAC TCT GGT CTG ATC CTG ATA CTT-3') (SEQ ID NO:23). Se diseñó el cebador anidado adicionalmente en el sentido de 3' del cebador 1 del gen estructural D 121-AA8 (CTA TAC GTA AGG TAA ATC CTG TGG AAC) (SEQ ID NO:24). Se comprobaron los productos de PCR finales mediante electroforesis en gel. Se cortaron las bandas deseadas del gel. Se purificaron los fragmentos de PCR usando el kit de purificación de gel QIAquick de QIAGEN. Se ligaron los fragmentos de PCR purificados en un vector pGEM-T Easy (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se extrajeron los plásmidos de ADN transformados mediante miniprep. usando el kit SV Miniprep (Promega, Madison, WI) siguiendo las instrucciones del fabricante. Se secuenció el ADN de plásmido que contenía el inserto usando un secuenciador CEQ 2000 (Beckman, Fullerton, CA). Se clonaron aproximadamente 1617 nucleótidos de la secuencia flanqueante en 3' adicional (nucleótidos 4731-6347 de SEQ ID NO:1) mediante el método descrito anteriormente. La secuencia de ácido nucleico de la región de UTR en 3' se muestra en la figura 6.

Los documentos WO 03/078577, WO 2004/035745, PCT/US/2004/034218 y PCT/US/2004/034065 se citan en el presente documento.

Lista de secuencias

<110> U. S. Smokeless Tobacco Company *et al.*

<120> Clon genómico de nicotina desmetilasa de tabaco y usos del mismo

<130> 07678/141WO13

<150> 11/xxx.xxx

<151> 19-04-2005

<150> 60/665.451

<151> 24-03-2005

<150> 60/665.097

<151> 24-03-2005

<150> 60/646.764

<151> 25-01-2005

<150> 60/607.357

<151> 03-09-2004

<150> 60/566.235

<151> 29-04-2004

<150> 10/934.944

<151> 03-09-2004

<150> 10/943.507

<151> 17-09-2004

<150> Documento PCT/US04/034218

<151> 15-10-2004

<150> Documento PCT/US04/034065

<151> 15-10-2004

<160> 29

5 <170> PatentIn versión 3.3

<210> 1

<211> 6347

<212> ADN

<213> *Nicotiana tabacum*

10 <400> 1

tctctaaagt	ccccttcac	tttatcttag	ctgtgtgatt	tctttcagac	aaccttattt	60
ttattcagac	tcttatttgt	attattctag	aagctcgtgt	acttgtgaca	ccagttctgg	120
gatggatatt	agatatcgct	attattttgg	cttattcact	tcagttcaga	ttttattcca	180
gttatttgat	ttctttatta	ttaatcaa	tgaattgtta	aaaatggtta	aaattactcc	240

aatggttggt	ttcctagtaa	gcgaaatata	aggcgccata	acggtacccg	aaggtgagaa	300
tttcagatcg	tgacagccgc	atctcaaggg	gtgtgatgta	aacagtttac	gatggtgcaa	360
gcattagtgg	ctgcttcgac	gacttaaata	cgtaacttat	agatcacacg	aatacaactt	420
tactatttta	acaccagca	aattcctgat	aaaaacaatt	ctaacatagc	acatcaaaat	480
gtaaatgatt	gaagaaaaag	atgactttta	tagacagaga	aaaaaacagt	tacgtgtggg	540
gttgaataga	gattgtggct	atgctatttc	taatattgaa	attcacccgac	tttttttagtt	600
caatacgaaa	agagtaagtg	aaaagggtctg	aaaaggaaaa	ggacaatgcc	taaaaggaca	660
catcagaac	acatacactg	aatgattcta	atttctagtc	cgaagatttc	tagttcgagg	720
ataacgtgag	caatcgggta	cttcccatta	gcaattgcca	actggatggt	tgactattta	780
tgtttctggc	caatagagga	gaggaatact	acgttacgta	tggagtttga	acccttcaca	840
tcaacttatt	aagtgagtta	tccctcaata	acgattcaac	tatgattccg	ctaacttcaa	900
agaatattga	gttaattatt	caaataatta	gtccaaaatt	atttaataaa	gttatactat	960
ttcttttatt	tgtaacatac	atcttttggt	tacatattta	gttaaatact	agcccaaccc	1020
tatcggttgg	tcgacttttt	ttcttttaata	tttgatttat	tctgttcggg	aatttcgctt	1080
tgtttggctt	gaataataac	tagtgcataa	agtcataata	tctaataatt	ttaattgaag	1140
tactcacaaa	tacaaaaata	aaaaacattc	taagctcaca	tgatagttga	caaaatactt	1200
atccaaaaca	agaggcggag	ccaggatttg	aaacttatgg	gttcagaatt	ctaaatactc	1260
taagttacta	ggttctaaat	taataattta	tacatgttca	atgaatttct	taagacaaat	1320
acatagtttg	aacgaaagct	actgggttcg	gcgaatccg	taagttatac	tctccctccg	1380
ccccggtcca	aaaccagcta	gtatcaatag	agagagagag	agagagagag	agataataaa	1440
tttgaccatt	gacaatggct	tattacttgc	ttagagttaa	ttggtgaact	tagagaatat	1500
aataaggaat	atttaaacag	atacgtcata	aatccacgag	taacgaagta	agaaataccc	1560
taaaatcgta	gaaacattac	gttaaattgc	ttgacagcct	atctagtaag	agtcaaaata	1620
tactatctat	cttggttcgc	cattttctta	aagaagtaca	tgagctttat	catccacctc	1680
aacatgaatg	caaaagaaaa	ttattgtgca	acttaatatg	ttataatcaa	tgatattgtg	1740
cttgtgtaac	aaagtatata	tttcgatacg	atattaatat	gtaggtgtta	tattttttaa	1800
tatcaaatat	catacttaac	accgattttt	taaaaactta	ggccaattac	cctaccaact	1860
aaaatactgt	atatcaaaca	ctaattgttt	ctatttcggg	acgacagttc	tctattttacc	1920

atattatgga attatgcccc	tcctacagtt acctataaaa	aggaagttgc cgatagttat	1980
atttctcaact tcttatctaa	aaatccataa tgctttctcc	catagaagcc attgtaggac	2040
tagtaacctt cacatttctc	ttctttctcc tatggacaaa	aaaatctcaa aaaccttcaa	2100
aacctttacc accgaaaatc	cccggaggat ggccggtaat	cggccatctt ttcacttca	2160
atgacgacgg cgacgaccgt	ccattagctc gaaaactcgg	agacttagct gacaaatacg	2220
gccccgtttt cacttttcgg	ctaggccttc cccttgtctt	agttgtaagc agttacgaag	2280
ctgtaaaaga ctgtttctct	acaaatgacg ccattttttc	caatcgacca gcttttcttt	2340
acggcgatta ccttggtac	aataatgcca tgctattttt	ggccaattac ggaccttact	2400
ggcgaaaaaa tcgaaaatta	gttattcagg aagttctctc	cgctagtcgt ctcgaaaaat	2460
tcaaacacgt gagatttgca	agaattcaag cgagcattaa	gaatttatat actcgaattg	2520
atggaaattc gagtacgata	aatttaactg attggttaga	agaattgaat tttggtctga	2580
tcgtgaagat gatcgctgga	aaaaattatg aatccggtaa	aggagatgaa caagtggaga	2640
gatttaagaa agcgtttaag	gattttatga ttttatcaat	ggagtttgtg ttatgggatg	2700
catttccaat tccattattt	aaatgggtgg attttcaagg	gcatgttaag gctatgaaaa	2760
ggacttttaa agatatagat	tctgtttttc agaattgggt	agaggaacat attaataaaa	2820
gagaaaaaat ggaggttaat	gcagaaggga atgaacaaga	tttcattgat gtggtgcttt	2880
caaaaatgag taatgaatat	cttggtgaag gttactctcg	tgatactgtc attaaagcaa	2940
cgggtgttgt aagttcatct	gtcatttttc atttattcac	ttttattttg aggagcagac	3000
atgttaataa taatttggag	caactgtaaa gttatctatg	tgtacagggt cgagcctcag	3060
gtgcaaccac taatgcttgt	attagattat gttgtctgca	tcatacccct aattggagtg	3120
tggctcttcc cgaaccctgc	aatgctggat gctggatgct	ttatgtatca gactgacctt	3180
tttggtaaac tatctaaata	ctaaggatga ttttaataaaa	atatagaatg gtaaacagaa	3240
aaagatgaga ttatttttgg	ggctatatgg attcgccggg	gctttgggag gtaaaacggt	3300
atctaccagt tgagacttta	ctccagaact ttatctcgag	agctctgaat aaaaatgaaa	3360
tagtatttac cactccaaaa	tctttgatgg taaaaagatg	agatataacc tcttataatt	3420
gattgaacca cgttgataga	ataaaacttc ttactccca	ttcagcataa gaaaaatgaa	3480
acaaaacgga attcttctct	tttttagggg gaaattcctt	aattgcttgt tgaatataga	3540
ttcatgtcgt tattctattt	ttaataatga tgaaaatcaa	tatagtcaaa gttaatactt	3600

atgtcatttg	gtttgcggac	aagttatat	ggaactatat	aatacgtcta	ttatagaata	3660
gtgattat	agaggatata	catttttttt	ggataaatat	ttgattttat	ggattaaaaa	3720
tagaatatac	aggtaaggtc	taaaacgtgt	gtttgctttt	acactaaata	aacttgacct	3780
cgtacaattc	taagaaaata	tttgaaataa	atgaattatt	ttattgttaa	tcaattaaaa	3840
aatcatagt	atagatgaga	tgtgtgcata	cttgacaata	actatactaa	ctaaaacaag	3900
gtatgtgaat	aattgatatt	ccttttttaa	ttattctttt	ttccagagtt	tggtcttgga	3960
tgcagcagac	acagttgctc	ttcacataaa	ttggggaatg	gcattattga	taaacaatca	4020
aaaggccttg	acgaaagcac	aagaagagat	agacacaaaa	gttggttaagg	acagatgggt	4080
agaagagagt	gatattaagg	atttggtata	cctccaagct	attgttaaag	aagtgttacg	4140
attatatcca	ccaggacctt	tgttagtacc	acacgaaaat	gtagaagatt	gtgttgtag	4200
tggatatcac	attcctaaag	ggacaagatt	attcgcaaac	gtcatgaaac	tgcaacgtga	4260
tcctaaactc	tggtctgate	ctgatacttt	cgatccagag	agattcattg	ctactgatat	4320
tgactttcgt	ggtcagtact	ataagtatat	cccgtttggg	tctggaagac	gatcttgtcc	4380
agggatgact	tatgcattgc	aagtggaaca	cttaacaatg	gcacatttga	tccaagggtt	4440
caattacaga	actccaaatg	acgagccctt	ggatatgaag	gaagggtgcag	gcataactat	4500
acgtaaggta	aatcctgtgg	aactgataat	agcgccctgc	ctggcacctg	agctttatta	4560
aaacctaaga	tctttcatct	tggttgatca	ttgtataata	ctcctaaatg	gatattcatt	4620
taccttttat	caattaattg	tcagtacgag	tttttcta	ttggtacatt	tgtaataata	4680
agtaaagaat	aattgtgcta	atatataaag	gtttgtagaa	gataattgac	tggttgtagc	4740
acaatctcca	gtgaaagtgt	taattattta	cttgatccac	agcttattct	atgtttgaaa	4800
tttgccatgt	gtcatgatat	tactccatca	aattcaagaa	ataatcattt	ccaacttttg	4860
ctggactgga	cgatctttca	ataataaagg	atctttaatt	tgccaaagtt	gagatcaaaa	4920
tactgggtgc	tttaccaata	agaatgaaat	gtgatggaaa	ttatgtacgt	tggtgataagg	4980
gaacacaact	atcaaggaga	ctaaaaggta	cgtaaaggaa	aagaaaaaat	ttgccattga	5040
ttgctactaa	gtaacctaac	aaaatctttc	agaaaagaat	cacttgata	agtcgggggt	5100
gaaagttttg	gtgtctcttt	tcttatgtat	tggtgtcttt	agacagtatt	gtacttagtt	5160
atttcagaat	tttattttcg	tattagagct	caagactctg	tatttattag	ttctggggaga	5220
attatcatgt	atcttcagtc	ttttgttatt	tctgtaaatt	ccgctatttt	ggttctttat	5280

tgctatgttc ggctttccta gaaaatgtgt taggcgctat cacgactgat tgagattttg 5340
 tatcgtgact gataattaca cggtttagta aattttgata ttttcaaaaa gagttttttt 5400
 aataaaatat gcaacttcag tcaaaacata caacgttttg ttgtataaat ccgatcaaaa 5460
 catataactt acataaaaact tgcgtatgaa ttttttgttg tataaacatt tgggttaaac 5520
 atataactta catataattt gcatacaact tacataaaac ttgcatataa acaattttatt 5580
 tatgtctttg tttttgagta tcaatttgaa attccaacaa aaacaaactc taattttttac 5640
 caaactctct caaaattgag ttatagattt caaaagatat ccttaatcgt ttgcaattgt 5700
 aacaatccga tcggccggtt tgagatctag cgtgttggtt ggcggtttga gaccttgagt 5760
 aacttcactt tatgttgat gacttgata tgtggtcgga attaaatttc gggaagttca 5820
 gagttgattc ggatgaaaaa ttctaatttc ggaagtttta agatggaatg attgactaag 5880
 gattgacgtt tgagtaaacg atctcggaat cgagatttga aggttccaat aggttcgtat 5940
 gatgatttca gacttgagcg tatgtttggg ttgagtatcg ggtggtcgag gagcatttca 6000
 acgctgatta tagaaaattg gcatcttaaa ggtttttagaa tttcataagt ttggtttgaa 6060
 gtggattttg atattatcgg tgtccatttg gagtttcgag ccttgaata gggttcgtatc 6120
 gtaaattttg acttttagtgt aaagttcggc gtcattccgg agtgttttga taagattctg 6180
 atgcgttcgt cgaagtttg aagtttgaaa gttgaaaaga agatttttaa taggcgattc 6240
 atgattttga tgttatttgt gtcgagcctt tggataagtt tgtgtgaggt atgggacttg 6300
 ttggtatgaa tggacgagct ctacgggggc ctcgagtaag tttcggga 6347

<210> 2

<211> 517

<212> PRT

5 <213> *Nicotiana tabacum*

<400> 2

Met Leu Ser Pro Ile Glu Ala Ile Val Gly Leu Val Thr Phe Thr Phe
 1 5 10 15

Leu Phe Phe Phe Leu Trp Thr Lys Lys Ser Gln Lys Pro Ser Lys Pro
 20 25 30

Leu Pro Pro Lys Ile Pro Gly Gly Trp Pro Val Ile Gly His Leu Phe
 35 40 45

His Phe Asn Asp Asp Gly Asp Asp Arg Pro Leu Ala Arg Lys Leu Gly
 50 55 60

Asp Leu Ala Asp Lys Tyr Gly Pro Val Phe Thr Phe Arg Leu Gly Leu
 65 70 75 80

Pro Leu Val Leu Val Val Ser Ser Tyr Glu Ala Val Lys Asp Cys Phe
 85 90 95

Ser Thr Asn Asp Ala Ile Phe Ser Asn Arg Pro Ala Phe Leu Tyr Gly
 100 105 110

Asp Tyr Leu Gly Tyr Asn Asn Ala Met Leu Phe Leu Ala Asn Tyr Gly
 115 120 125

Pro Tyr Trp Arg Lys Asn Arg Lys Leu Val Ile Gln Glu Val Leu Ser
 130 135 140

Ala Ser Arg Leu Glu Lys Phe Lys His Val Arg Phe Ala Arg Ile Gln
 145 150 155 160

Ala Ser Ile Lys Asn Leu Tyr Thr Arg Ile Asp Gly Asn Ser Ser Thr
 165 170 175

Ile Asn Leu Thr Asp Trp Leu Glu Glu Leu Asn Phe Gly Leu Ile Val
 180 185 190

Lys Met Ile Ala Gly Lys Asn Tyr Glu Ser Gly Lys Gly Asp Glu Gln
 195 200 205

Val Glu Arg Phe Lys Lys Ala Phe Lys Asp Phe Met Ile Leu Ser Met
 210 215 220

Glu Phe Val Leu Trp Asp Ala Phe Pro Ile Pro Leu Phe Lys Trp Val
 225 230 235 240

Asp Phe Gln Gly His Val Lys Ala Met Lys Arg Thr Phe Lys Asp Ile
 245 250 255

Asp Ser Val Phe Gln Asn Trp Leu Glu Glu His Ile Asn Lys Arg Glu
 260 265 270

Lys Met Glu Val Asn Ala Glu Gly Asn Glu Gln Asp Phe Ile Asp Val
275 280 285

Val Leu Ser Lys Met Ser Asn Glu Tyr Leu Gly Glu Gly Tyr Ser Arg
290 295 300

Asp Thr Val Ile Lys Ala Thr Val Phe Ser Leu Val Leu Asp Ala Ala
305 310 315 320

Asp Thr Val Ala Leu His Ile Asn Trp Gly Met Ala Leu Leu Ile Asn
325 330 335

Asn Gln Lys Ala Leu Thr Lys Ala Gln Glu Glu Ile Asp Thr Lys Val
340 345 350

Gly Lys Asp Arg Trp Val Glu Glu Ser Asp Ile Lys Asp Leu Val Tyr
355 360 365

Leu Gln Ala Ile Val Lys Glu Val Leu Arg Leu Tyr Pro Pro Gly Pro
370 375 380

Leu Leu Val Pro His Glu Asn Val Glu Asp Cys Val Val Ser Gly Tyr
385 390 395 400

His Ile Pro Lys Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gln
405 410 415

Arg Asp Pro Lys Leu Trp Ser Asp Pro Asp Thr Phe Asp Pro Glu Arg
420 425 430

Phe Ile Ala Thr Asp Ile Asp Phe Arg Gly Gln Tyr Tyr Lys Tyr Ile
435 440 445

Pro Phe Gly Ser Gly Arg Arg Ser Cys Pro Gly Met Thr Tyr Ala Leu
450 455 460

Gln Val Glu His Leu Thr Met Ala His Leu Ile Gln Gly Phe Asn Tyr
465 470 475 480

Arg Thr Pro Asn Asp Glu Pro Leu Asp Met Lys Glu Gly Ala Gly Ile
485 490 495

Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile Ile Ala Pro Arg Leu
500 505 510

Ala Pro Glu Leu Tyr
515

<210> 3

<211> 1554

5 <212> ADN

<213> Secuencia codificante del gen de nicotina desmetilasa de tabaco

<400> 3

atgctttctc ccatagaagc cattgtagga ctagtaacct tcacatttct cttcttcttc	60
ctatggacaa aaaaatctca aaaaccttca aaacccttac caccgaaaat ccccgaggga	120
tggccggtaa tcggccatct tttccacttc aatgacgacg gcgacgaccg tccattagct	180
cgaaaactcg gagacttagc tgacaaatac ggccccgttt tcacttttcg gctaggcctt	240
ccccttgtct tagttgtaag cagttaacgaa gctgtaaaag actgtttctc taaaaatgac	300
gccatttttt ccaatcgcc agcttttctt tacggcgatt accttggcta caataatgcc	360
atgctatttt tggccaatta cggaccttac tggcgaaaaa atcgaaaatt agttattcag	420
gaagtctctc ccgctagtcg tctcgaaaaa ttcaaacacg tgagatttgc aagaattcaa	480
gcgagcatta agaatttata tactcgaatt gatggaaatt cgagtacgat aaatttaact	540
gattgggttag aagaattgaa ttttggctctg atcgtgaaga tgatcgctgg aaaaaattat	600
gaatccggta aaggagatga acaagtggag agatttaaga aagcgtttaa ggattttatg	660
attttatcaa tggagtttgt gttatgggat gcatttccaa ttccattatt taaatgggtg	720
gattttcaag ggcattgttaa ggctatgaaa aggactttta aagatataga ttctgttttt	780
cagaattggg tagaggaaca tattaataaa agagaaaaaa tggagggttaa tgcagaaggg	840
aatgaacaag atttcattga tgtggtgctt tcaaaaatga gtaatgaata tcttgggtgaa	900
ggttactctc gtgatactgt cattaaagca acggtgttta gtttggctct ggatgcagca	960
gacacagttg ctcttcacat aaattgggga atggcattat tgataaacia tcaaaaggcc	1020
ttgacgaaag cacaagaaga gatagacaca aaagttggta aggacagatg ggtagaagag	1080
agtgatatta aggatttggg atacctccaa gctattgtta aagaagtgtt acgattatat	1140
ccaccaggac ctttgttagt accacacgaa aatgtagaag attgtgttgt tagtggatat	1200
cacattccta aagggacaag attattcgca aacgtcatga aactgcaacg tgatcctaaa	1260
ctctgggtctg atcctgatac tttegatcca gagagattca ttgctactga tattgacttt	1320
cgtgggtcagt actataagta tatcccgttt gggtcttgaa gacgatcttg tccagggatg	1380
acttatgcat tgcaagtgga acacttaaca atggcacatt tgatccaagg tttcaattac	1440
agaactccaa atgacgagcc cttggatatg aaggaagggtg caggcataac tatacgtaag	1500
gtaaatacctg tggaactgat aatagcgctt cgctggcac ctgagcttta ttaa	1554

<210> 4

<211> 517

5 <212> PRT

ES 2 373 881 T3

<213> Gen de nicotina desmetilasa de tabaco

<400> 4

Met	Leu	Ser	Pro	Ile	Glu	Ala	Ile	Val	Gly	Leu	Val	Thr	Phe	Thr	Phe	1	5	10	15
Leu	Phe	Phe	Phe	Leu	Trp	Thr	Lys	Lys	Ser	Gln	Lys	Pro	Ser	Lys	Pro	20	25	30	
Leu	Pro	Pro	Lys	Ile	Pro	Gly	Gly	Trp	Pro	Val	Ile	Gly	His	Leu	Phe	35	40	45	
His	Phe	Asn	Asp	Asp	Gly	Asp	Asp	Arg	Pro	Leu	Ala	Arg	Lys	Leu	Gly	50	55	60	
Asp	Leu	Ala	Asp	Lys	Tyr	Gly	Pro	Val	Phe	Thr	Phe	Arg	Leu	Gly	Leu	65	70	75	80
Pro	Leu	Val	Leu	Val	Val	Ser	Ser	Tyr	Glu	Ala	Val	Lys	Asp	Cys	Phe	85	90	95	
Ser	Thr	Asn	Asp	Ala	Ile	Phe	Ser	Asn	Arg	Pro	Ala	Phe	Leu	Tyr	Gly	100	105	110	
Asp	Tyr	Leu	Gly	Tyr	Asn	Asn	Ala	Met	Leu	Phe	Leu	Ala	Asn	Tyr	Gly	115	120	125	
Pro	Tyr	Trp	Arg	Lys	Asn	Arg	Lys	Leu	Val	Ile	Gln	Glu	Val	Leu	Ser	130	135	140	
Ala	Ser	Arg	Leu	Glu	Lys	Phe	Lys	His	Val	Arg	Phe	Ala	Arg	Ile	Gln	145	150	155	160

Ala	Ser	Ile	Lys	Asn	Leu	Tyr	Thr	Arg	Ile	Asp	Gly	Asn	Ser	Ser	Thr	165	170	175
Ile	Asn	Leu	Thr	Asp	Trp	Leu	Glu	Glu	Leu	Asn	Phe	Gly	Leu	Ile	Val	180	185	190
Lys	Met	Ile	Ala	Gly	Lys	Asn	Tyr	Glu	Ser	Gly	Lys	Gly	Asp	Glu	Gln	195	200	205
Val	Glu	Arg	Phe	Lys	Lys	Ala	Phe	Lys	Asp	Phe	Met	Ile	Leu	Ser	Met	210	215	220
Glu	Phe	Val	Leu	Trp	Asp	Ala	Phe	Pro	Ile	Pro	Leu	Phe	Lys	Trp	Val	225	230	235
Asp	Phe	Gln	Gly	His	Val	Lys	Ala	Met	Lys	Arg	Thr	Phe	Lys	Asp	Ile	245	250	255
Asp	Ser	Val	Phe	Gln	Asn	Trp	Leu	Glu	Glu	His	Ile	Asn	Lys	Arg	Glu	260	265	270
Lys	Met	Glu	Val	Asn	Ala	Glu	Gly	Asn	Glu	Gln	Asp	Phe	Ile	Asp	Val	275	280	285
Val	Leu	Ser	Lys	Met	Ser	Asn	Glu	Tyr	Leu	Gly	Glu	Gly	Tyr	Ser	Arg	290	295	300
Asp	Thr	Val	Ile	Lys	Ala	Thr	Val	Phe	Ser	Leu	Val	Leu	Asp	Ala	Ala	305	310	315
Asp	Thr	Val	Ala	Leu	His	Ile	Asn	Trp	Gly	Met	Ala	Leu	Leu	Ile	Asn	325	330	335
Asn	Gln	Lys	Ala	Leu	Thr	Lys	Ala	Gln	Glu	Glu	Ile	Asp	Thr	Lys	Val	340	345	350
Gly	Lys	Asp	Arg	Trp	Val	Glu	Glu	Ser	Asp	Ile	Lys	Asp	Leu	Val	Tyr	355	360	365
Leu	Gln	Ala	Ile	Val	Lys	Glu	Val	Leu	Arg	Leu	Tyr	Pro	Pro	Gly	Pro	370	375	380

Leu Leu Val Pro His Glu Asn Val Glu Asp Cys Val Val Ser Gly Tyr
385 390 395 400

His Ile Pro Lys Gly Thr Arg Leu Phe Ala Asn Val Met Lys Leu Gln
405 410 415

Arg Asp Pro Lys Leu Trp Ser Asp Pro Asp Thr Phe Asp Pro Glu Arg
420 425 430

Phe Ile Ala Thr Asp Ile Asp Phe Arg Gly Gln Tyr Tyr Lys Tyr Ile
435 440 445

Pro Phe Gly Ser Gly Arg Arg Ser Cys Pro Gly Met Thr Tyr Ala Leu
450 455 460

Gln Val Glu His Leu Thr Met Ala His Leu Ile Gln Gly Phe Asn Tyr
465 470 475 480

Arg Thr Pro Asn Asp Glu Pro Leu Asp Met Lys Glu Gly Ala Gly Ile
485 490 495

Thr Ile Arg Lys Val Asn Pro Val Glu Leu Ile Ile Ala Pro Arg Leu
500 505 510

Ala Pro Glu Leu Tyr
515

<210> 5

<211> 998

<212> ADN

5 <213> Intrón del gen de nicotina desmetilasa de tabaco

<400> 5

gtaagttcat ctgtcatttt tcattttatc actttttat	60
aataatttgg agcaactgta aagttatcta tgtgtacagg ttcgagcctc aggtgcaacc	120
actaatgctt gtattagatt atgttgtctg catcataccc ctaattggag tgtggctctt	180
cccgaaccct gcaatgctgg atgctggatg ctttatgtat cagactgacc tttttgttaa	240
actatctaaa tactaaggat gatttaataa aaatatagaa tggtaaacag aaaaagatga	300
gattatTTTT ggggtatat ggattcgccc gggctttggg aggtaaaacg gtatctacca	360

gttgagactt tactccagaa ctttatctcg agagctctga ataaaaatga aatagtattt	420
accactccaa aatctttgat ggtaaaaaga tgagatataa cctcttataa ttgattgaac	480
cacgttgata gaataaaact tctttactcc cattcagcat aagaaaaatg aaaccaaacg	540
gaattcttct cttttttagg gggaaattcc ttaattgctt gttgaatata gattcatgtc	600
gttattctat ttttaataat gatgaaaatc aatatagtca aagttaatac ttatgtcatt	660
tggtttgcg acaagttata ttggaactat ataatacgtc tattatagaa tagtgattat	720
ttagaggata tacatttttt ttggataaat atttgattta ttggattaaa aatagaatat	780
acaggtaagg tctaaaacgt gtgtttgctt ttacactaaa taaacttgac ctctgacaat	840
tctaagaaaa tatttgaaat aaatgaatta ttttattggt aatcaattaa aaaaatcata	900
gtatagatga gatgtgtgca tacttgacaa taactatact aactaaaaca aggtatgtga	960
ataattgata ttcccttttt aattattctt ttttccag	998

<210> 6

<211> 2009

<212> ADN

5 <213> Promotor del gen de nicotina desmetilasa de tabaco

<400> 6

tctctaaagt ccccttccac tttatcttag ctgtgtgatt tctttcagac aaccttattt	60
ttattcagac tcttatttgt attattctag aagctcgtgt acttgtgaca ccagttctgg	120
gatggtattt agatatcgtt attattttgg cttattcact tcagttcaga ttttattcca	180
gttatttgat ttctttatta ttaatcaaatt tgaattgtta aaaatgggta aaattactcc	240
aatgttggtt ttcttagtaa gcgaaatata aggcgccatc acggtaccgc aaggtgagaa	300
tttcagatcg tgacagccgc atctcaaggg gtgtgatgta aacagtttac gatggtgcaa	360
gcattagtgg ctgcttcgac gacttaaata cgtaacttat agatcacacg aatacaactt	420
tactatttta acaccagca aattcctgat aaaaacaatt ctaacatagc acatcaaaat	480
gtaaatgatt gaagaaaaag atgactttta tagacagaga aaaaaacagt tacgtgtggg	540
gttgaataga gattgtggct atgctatttc taatattgaa attcaccgac ttttttagtt	600
caatacgaaa agagtaagtg aaaaggctctg aaaaggaaaa ggacaatgcc taaaaggaca	660
cattcagaac acatacactg aatgattcta atttctagtc cgaagatttc tagttcgagg	720
ataacgtgag caatcgggta cttcccatta gcaattgcc actggatgtt tgactattta	780
tgtttctggc caatagagga gaggaatact acgttacgta tggagtttga acccttcaca	840

tcaacttatt aagtgagtta tccctcaatc acgattcaac tatgattccg ctaacttcaa	900
agaatattga gttaattatt caaatgatta gtccaaaatt atttaataaa gttatactat	960
ttcttttatt tgtaacatac atcttttgggt tacatattta gttaaatctt agcccaaccc	1020
tatcggttgggt tgcactttttt ttctttttaat tttgattttat tctgttcgggt aatttcgctt	1080
tgtttgggctt gaataataac tagtgcataa agtcatatac tctaataattt ttaattgaag	1140
tactcacaaa tacaaaaata aaaaacattc taagctcaca tgatagttga caaatctttt	1200
atccaaaaca agaggcggag ccaggatttg aaacttatgg gttcagaatt ctaaactctct	1260
taagttacta ggttctaaat taataattta tacatgttca atgaatttct taagacaaat	1320
acatagtttg aacgaaagct actgggttcg gccgaatccg taagttatac tctccctccg	1380
ccccggtcca aaaccagcta gtatcaatag agagagagag agagagagag agataataaa	1440
tttgaccatt gacaatgggt tattacttgc ttagagttaa ttggtgaact tagagaatat	1500
aataaggaat atttaaacag atacgtcatc aatccacgag taacgaagta agaaatacc	1560
taaaatcgta gaaacattac gttaaattgc ttgacagcct atctagtaag agtcaaaatc	1620
tactatctat cttgttcgcg catttttctta aagaagtaca tgagctttat catccacctc	1680
aacatgaatg caaaagaaaa ttattgtgca acttaatatg ttataatcaa tgatatgtgt	1740
cttgtgtaac aaagtatata ttctgatacg atattaatat gtaggtgtta tattttttaa	1800
tatcaaatat catacttaac accgattttt taaaaactta ggccaattac cctaccaact	1860
aaaatactgt atatcaaaca ctaatgtttt ctatttcgggt acgacagttc tctatttacc	1920
atattatgga attatgccca tcctacagtt acctataaaa aggaagttgc cgatagttat	1980
attctcaact tcttatctaa aaatccata	2009

<210> 7

<211> 1786

<212> ADN

5 <213> UTR en 3' del gen de nicotina desmetilasa de tabaco

<400> 7

aacctaagat ctttcatctt ggttgatcat tgtataatac tcctaaatgg atattcattt	60
acctttttatc aattaattgt cagtacgagt ttttctaatt tggtagattt gtaataataa	120
gtaaagaata attgtgctaa tatataaagg tttgtagaag ataattgact ggttgtagca	180
caatctccag tgaaagtgtt aattatttac ttgatccaca gcttattcta tgtttgaaat	240

ttgcctagtg tcatgatatt actccatcaa attcaagaaa taatcatttc caacttttgc	300
tggactggac gatctttcaa taataaagga tctttaattt gccaaagttg agatcaaaat	360
actggtcgct ttaccaataa gaatgaaatg tgatggaaat tatgtacgtt gggataaggg	420
aacacaacta tcaaggagac taaaaggtag gtaaaggaaa agaaaaaatt tgccattgat	480
tgctactaag taacctaacaa aaatctttca gaaaagaatc acttgataaa gtcgggggtg	540
aaagtttttg tgtctctttt cttatgtatt gttgtcttta gacagtattg tacttagtta	600
tttcagaatt ttattttcgt attagagctc aagactctgt atttattagt tctgggagaa	660
ttatcatgta ttttcagtct tttgttattt ctgtaaaattc cgctattttg gttctttatt	720
gctatgttcg gctttcctag aaaatgtgtt aggcgctatc acgactgatt gagattttgt	780
atcgtgactg ataattacac ggtttagtaa attttgatat tttcaaaaag agttttttta	840
ataaaatatg caacttcagt caaaacatac aacgttttgt tgtataaatc cgatcaaaac	900
atataactta cataaaactt gcgtatgaat tttttgttgt ataaacattt gggtaaaaca	960
tataacttac atataatttg catacaactt acataaaaact tgcataataa caattttattt	1020
atgtctttgt ttttgagtat caatttgaaa ttccaacaaa aacaaactct aatttttacc	1080
aaactctctc aaaattgagt tatagatttc aaaagatatc cttaatcgtt tgcaattgta	1140
acaatecgat cggccgtttt gagatctagc gtgttgtttg gcggtttgag acctgagta	1200
acttcacttt atgttgtagt acttgatat gtggtcggaa ttaaatttcg ggaagttcag	1260
agttgattcg gatgaaaaat tctaatttcg gaagttttta gatggaatga ttgactaagg	1320
attgacgttt gagtaaacga tctcggaatc gagatttgaa ggtccaata gggtcgtagt	1380
atgatttcag acctgagcgt atgtttgggt tgagtatcgg gtggtcggg agcatttcaa	1440
cgctgattat agaaaattgg catcttaaag gttttagaat ttcataagtt tggtttgaag	1500
tggattttga tattatcgggt gtccatttgg agtttcgagc cttggaatag gttcgtatcg	1560
taaattttga ctttagtgta aagttcggcg tcatcccgga gtgttttgat aagattctga	1620
tgcgttcgtc gaagtttgga agtttgaaag ttgaaaagaa gatttttaat aggcgattca	1680
tgattttgat gttatttgtg tcgagccttt ggataagttt gtgtgaggta tgggacttgt	1740
tggtagaat ggacgagctc tacggggggcc tcgagtaagt ttcgga	1786

<210> 8

<211> 21

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Péptido sintetizado

<400> 8

Asp Ile Asp Gly Ser Lys Ser Lys Leu Val Lys Ala His Arg Lys Ile
1 5 10 15

Asp Glu Ile Leu Gly
20

5 <210> 9

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

10 <223> Péptido sintetizado

<400> 9

Arg Asp Ala Phe Arg Glu Lys Glu Thr Phe Asp Glu Asn Asp Val Glu
1 5 10 15

Glu Leu Asn Tyr
20

<210> 10

<211> 22

15 <212> PRT

<213> Secuencia artificial

<220>

<223> Péptido sintetizado

<400> 10

Phe Lys Asn Asn Gly Asp Glu Asp Arg His Phe Ser Gln Lys Leu Gly
1 5 10 15

Asp Leu Ala Asp Lys Tyr
20

20

<210> 11

<211> 6

<212> PRT

<213> Secuencia artificial

5 <220>

<223> Secuencia consenso del motivo p450 conservado

<220>

<221> MISC_FEATURE

<222> (2)..(2)

10 <223> Xaa = cualquier aminoácido

<400> 11

Phe Xaa Pro Glu Arg Phe
1 5

<210> 12

<211> 7

15 <212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia consenso del motivo p450 conservado

<220>

20 <221> MISC_FEATURE

<222> (4)..(4)

<223> Xaa = cualquier aminoácido

<220>

<221> MISC_FEATURE

25 <222> (7)..(7)

<223> Xaa = Ala o Gly

<400> 12

Gly Arg Arg Xaa Cys Pro Xaa
1 5

<210> 13
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 5 <220>
 <223> Secuencia consenso del motivo p450 conservado
 <220>
 <221> MISC_FEATURE
 <222> (2)..(2)
 10 <223> Xaa = cualquier aminoácido
 <220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 <223> Xaa = cualquier aminoácido
 15 <220>
 <221> MISC_FEATURE
 <222> (7)..(7)
 <223> Xaa = Ala o Gly
 <400> 13

 Gly Xaa Arg Xaa Cys Pro Xaa
 1 5
 20
 <210> 14
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 25 <220>
 <223> Secuencia consenso del motivo p450 conservado
 <220>
 <221> MISC_FEATURE
 <222> (1).. (1)
 30 <223> Xaa = Glu o Lys

<220>

<221> MISC_FEATURE

<222> (2)..(3)

<223> Xaa = cualquier aminoácido

5 <220>

<221> MISC_FEATURE

<222> (5)..(6)

<223> Xaa = cualquier aminoácido

<220>

10 <221> MISC_FEATURE

<222> (7)..(7)

<223> Xaa = Pro, Thr, Ser, o Met

<400> 14

Xaa Xaa Xaa Arg Xaa Xaa Xaa
1 5

15 <210> 15

<211> 6

<212> PRT

<213> Secuencia artificial

<220>

20 <223> Secuencia consenso del motivo p450 conservado

<220>

<221> MISC_FEATURE

<222> (2).. (2)

<223> Xaa = cualquier aminoácido

25 <220>

<221> MISC_FEATURE

<222> (5)..(5)

<223> Xaa = cualquier aminoácido

<400> 15

Pro Xaa Arg Phe Xaa Phe
1 5

<210> 16
 <211> 5
 <212> PRT

5 <213> Secuencia artificial
 <220>
 <223> Secuencia consenso del motivo p450 conservado
 <220>
 <221> MISC_FEATURE

10 <222> (2).. (2)
 <223> Xaa = cualquier aminoácido
 <220>
 <221> MISC_FEATURE
 <222> (4)..(4)

15 <223> Xaa = cualquier aminoácido
 <400> 16

Gly Xaa Arg Xaa Cys
1 5

<210> 17
 <211> 36

20 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador
 <400> 17

25 **ggctctagat aaatctctta agttactagg ttctaa** 36
 <210> 18
 <211> 44
 <212> ADN
 <213> Secuencia artificial

<220>

<223> Cebador

<400> 18

ggctctagaa gtcaattatc ttctacaaac ctttatatat tagc 44

5 <210> 19

<211> 27

<212> ADN

<213> Secuencia artificial

<220>

10 <223> Cebador

<400> 19

ctctattgat actagctggt ttggac 27

<210> 20

<211> 26

15 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador

<400> 20

20 ggagggagag tataacttac ggattc 26

<210> 21

<211> 26

<212> ADN

<213> Secuencia artificial

25 <220>

<223> Cebador

<400> 21

agtaacogat tgctacggtt atcttc 26

<210> 22

30 <211> 25

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador

5 <400> 22

ctctattcaa cccacacgt aactg 25

<210> 23

<211> 27

<212> ADN

10 <213> Secuencia artificial

<220>

<223> Cebador

<400> 23

ctaaactctg gctgatcct gatactt 27

15 <210> 24

<211> 27

<212> ADN

<213> Secuencia artificial

<220>

20 <223> Cebador

<400> 24

ctatacgtaa ggtaaatcct gtggaac 27

<210> 25

<211> 18

25 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador

<400> 25

30 tctctaaagt ccccttc 18

<210> 26

<211> 18

<212> ADN

<213> Secuencia artificial

5 <220>

<223> Cebador

<400> 26

ccagcattod tcaattc 18

<210> 27

10 <211> 36

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador

15 <400> 27

tagctaogcg gatccatgct ttctoccata gaagcc 36

<210> 28

<211> 64

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Cebador

<400> 28

ctggatcaca attggttagtg atggtgatgg tgatgcgato ctctataaag ctcaggtgcc 60

aggc 64

25 <210> 29

<211> 7

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia consenso para el motivo p450 conservado

<220>

<221> MISC_FEATURE

<222> (2)..(2)

5 <223> Xaa = cualquier aminoácido

<220>

<221> MISC_FEATURE

<222> (4).. (4)

<223> Xaa = cualquier aminoácido

10 <220>

<221> MISC_FEATURE

<222> (6)..(6)

<223> Xaa = cualquier aminoácido

<220>

15 <221> MISC_FEATURE

<222> (7).. (7)

<223> Xaa = Gly o Ala

<400> 29

Gly Xaa Arg Xaa Cys Xaa Xaa
1 5

20

REIVINDICACIONES

1. Molécula de ácido nucleico aislada que codifica para una nicotina desmetilasa, comprendiendo dicha molécula de ácido nucleico (a) la secuencia de SEQ ID NO:1, o (b) al menos 200 pares de bases consecutivas idénticas a 200 pares de bases consecutivas de SEQ ID NO:5, SEQ ID NO:6 o SEQ ID NO:7.
- 5 2. Molécula de ácido nucleico aislada según la reivindicación 1,
 - (a) siendo dicha molécula de ácido nucleico ADN;
 - (b) comprendiendo dicha molécula de ácido nucleico la secuencia de SEQ ID NO:1;
 - (c) codificando dicha molécula de ácido nucleico para una nicotina desmetilasa que comprende la secuencia de aminoácidos de SEQ ID NO:2;
 - 10 (d) estando dicha molécula de ácido nucleico aislada operativamente unida a un promotor funcional en una célula vegetal; o
 - (e) comprendiendo un promotor que hibrida en condiciones rigurosas con la secuencia de SEQ ID NO:6, o un fragmento que comprende al menos 200 pares de bases consecutivas idénticas a 200 pares de bases consecutivas de SEQ ID NO:6, en el que dicho promotor dirige la transcripción, particularmente en el que el promotor se induce tras tratamiento con etileno o durante la senescencia.
 - 15 3. Ácido nucleico aislado según la reivindicación 1 ó 2,
 - (a) en el que la expresión de la nicotina desmetilasa se induce tras tratamiento con etileno o durante la senescencia; o
 - (b) en el que la molécula de ácido nucleico comprende la secuencia de SEQ ID NO:6.
 - 20 4. Molécula de ácido nucleico aislada según la reivindicación 1, comprendiendo la molécula de ácido nucleico la secuencia de SEQ ID NO:5 o un fragmento que tiene al menos 200 pares de bases consecutivas que puede obtenerse a partir de SEQ ID NO:5.
 5. Molécula de ácido nucleico aislada según la reivindicación 1, comprendiendo la molécula de ácido nucleico la secuencia de SEQ ID NO:7 o un fragmento que tiene al menos 200 pares de bases consecutivas que puede obtenerse a partir de SEQ ID NO:7.
 - 25 6. Planta de tabaco transgénica que tiene con respecto al nivel en una planta control una expresión reducida o actividad enzimática alterada de una nicotina desmetilasa codificada por una molécula de ácido nucleico según la reivindicación 1,
 - (a) comprendiendo dicha planta transgénica un transgén que tiene una identidad de secuencia del 91% o superior con SEQ ID NO:1, o con un fragmento de la misma, que, cuando se expresa en dicha planta transgénica, silencia la expresión génica de una nicotina desmetilasa de tabaco endógena;
 - 30 (b) comprendiendo dicha planta transgénica un transgén que expresa una molécula antisentido de una nicotina desmetilasa de tabaco;
 - (c) comprendiendo dicha planta transgénica un transgén que codifica para una molécula de ARN bicatenario de una nicotina desmetilasa de tabaco;
 - 35 (d) comprendiendo dicha planta transgénica un transgén que, cuando se expresa en dicha planta transgénica, suprime conjuntamente la expresión de una nicotina desmetilasa de tabaco; o
 - (e) comprendiendo dicha planta transgénica una delección en una nicotina desmetilasa endógena codificada por la secuencia de SEQ ID NO:1; y
 - 40 en la que dicha expresión reducida o actividad enzimática alterada reduce el nivel de nornicotina en dicha planta.
 7. Planta de tabaco transgénica según la reivindicación 6, en la que dicha nicotina desmetilasa comprende la secuencia de aminoácidos de SEQ ID NO:2.

8. Componente vegetal de la planta de tabaco según la reivindicación 6, particularmente siendo dicho componente vegetal una semilla, tallo u hoja de tabaco, especialmente en el que dicha hoja está curada.
9. Producto de tabaco que comprende el componente vegetal según la reivindicación 8.
10. Semilla de la planta de tabaco según la reivindicación 6, comprendiendo dicha semilla dicho transgén o dicha delección.
11. Método de producción de una nicotina desmetilasa de tabaco, comprendiendo dicho método las etapas de:
 - (a) proporcionar una célula transformada con una molécula de ácido nucleico aislada según la reivindicación 1;
 - (b) cultivar dicha célula transformada en condiciones para expresar dicha molécula de ácido nucleico aislada; y
 - (c) recuperar dicha nicotina desmetilasa de tabaco.
12. Nicotina desmetilasa de tabaco recombinante producida según el método de la reivindicación 11.
13. Método de aislamiento de una nicotina desmetilasa de tabaco o fragmento que tiene al menos 200 pares de bases consecutivas de la misma, comprendiendo dicho método las etapas de:
 - (a) poner en contacto la molécula de ácido nucleico según la reivindicación 1 con una preparación de ácido nucleico de una célula vegetal en condiciones de hibridación que proporcionan la detección de secuencias de ácido nucleico que tienen una identidad de secuencia de al menos el 70% o superior con un ácido nucleico según la reivindicación 1; y
 - (b) aislar dichas secuencias de ácido nucleico hibridantes.
14. Método de aislamiento de un ácido nucleico que comprende SEQ ID NO:5, SEQ ID NO:6 o SEQ ID NO:7, o fragmento que tiene al menos 200 pares de bases consecutivas del mismo, comprendiendo dicho método las etapas de:
 - (a) proporcionar una preparación de ácido nucleico a partir de una célula vegetal;
 - (b) poner en contacto un par de oligonucleótidos que tienen identidad de secuencia con una región de SEQ ID NO:5, SEQ ID NO:6 o SEQ ID NO:7 con dicha preparación de ácido nucleico en condiciones adecuadas para la amplificación de ADN mediada por reacción en cadena de la polimerasa para producir un producto de reacción en cadena de la polimerasa; y
 - (c) aislar el ácido nucleico o fragmento del mismo producido en la etapa (c).
15. Método según la reivindicación 14, en el que dicha preparación de ácido nucleico es ADN genómico.
16. Método según la reivindicación 14, en el que dicha preparación de ácido nucleico es ADNc.
17. Molécula de ácido nucleico aislada que comprende SEQ ID NO:5, SEQ ID NO:6 o SEQ ID NO:7.
18. Molécula de ácido nucleico aislada según la reivindicación 17, estando dicha molécula nucleica aislada operativamente unida a una secuencia de ácido nucleico heteróloga.

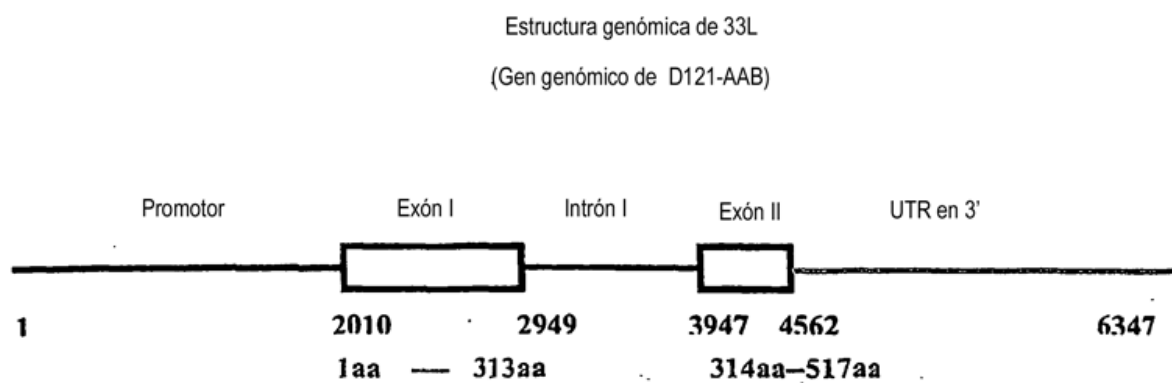


FIGURA 1

CARACTERÍSTICAS	Ubicación/Calificadores
Total	1..6347
Secuencias flanqueantes en 5'	1..2009
CDS	unión (2010..2948,3947..4561)
Intrón	2949..3946
Producto	Citocromo P450
UTR en 3'	4562..6347

Traducción

MLSPIEAIIVGLVTFITLFFFLWTKKSQKPSKPLPPKIPGGWPFVIGHLFHFNDGDDDRPLARKLGDLDADKYGPVF
TFRLGLPLVLVSSYEAVKDCFSTNDAIFS NRPAFLYGDYLGYNNAFLANYGPYWRKNRKLVIQEVLSASRL
EKPKHVRFARIQASIKNLYTRIDGNSSTINLTDWLEELNPLGLIVKMIAGKNYESGKGDEQVERPKAFKDFMIL
SMEFVLWDAFPIPLFKWVDFQGHVKAMKRTFKDIDS VFQNWLEEHINKREKMEVNAEGNEQDPIDVVL SKMSNE
YLGE GYSRDTVIKATVFSVLDAADTVALHINWGMALLINQKALTKAQEEIDTKVKGDRWVEESDIKDLVYLQ
AIVKEVLRLLYPPGPLLVPHENVEDCVVSGYHIPKGT RLFPANVMKLQRDPKLWSDPDTFDPERFIATDIDFRGQY
YKYIPFGSGRRSCPGMTYALQVEH LTMALHIQGFNYRTPNDEPLDMKEGAGITIRKVNVELI IAPRLAPELY

RECuento DE BASES 2046 a 984 c 1163 g 2154 t

1	TCTCTAAAGT	CCCCTTCCAC	TTTATCTTAG	CTGTGTGATT	TCTTTCAGAC	AACCTTATTT
61	TTATTCAGAC	TCTTATTTGT	ATTATTCTAG	AAGCTCGTGT	ACTTGTGACA	CCAGTCTCTGG
121	GATGGTATTT	AGATATCGCT	ATTATTTTGG	CTTATTCACT	TCAGTTCAGA	TTTTATTCCA
181	GTTATTTGAT	TTCTTTATTA	TTAATCAAAT	TGAATTGTTA	AAAATGGTTA	AAATTACTCC
241	AATGTTGGCT	TTCTAGTAA	GCGAAATATC	AGGCGCCATC	ACGGTACCCG	AAGGTGAGAA
301	TTTCAGATCG	TGACAGCCGC	ATCTCAAGGG	GTGTGATGTA	AACAGTTTAC	GATGGTGCAA
361	GCATTAGTGG	CTGCTTCGAC	GACTTAAATC	CGTAACTTAT	AGATCACACG	AATACAACTT
421	TACTATTTTA	ACACCCAGCA	AATTCCTGAT	AAAAACAATT	CTAACATAGC	ACATCAAAAT
481	GTAATGATT	GAAGAAAAAG	ATGACTTTTA	TAGACAGAGA	AAAAAACAGT	TACGTGTGGG
541	GTTGAATAGA	GATTGTGGCT	ATGCTATTTT	TAATATTGAA	ATTCAACGAC	TTTTTTAGTT
601	CAATACGAAA	AGAGTAAGTG	AAAAGGTCTG	AAAAGGAAAA	GGACAATGCC	TAAAAGGACA
661	CATTACGAAC	ACATACACTG	AATGATTCTA	ATTTCTAGTC	CGAAGATTTT	TAGTTCGAGG
721	ATAACGTGAG	CAATCGGTGA	CTTCCCATTG	GCAATTGCCA	ACTGGATGTT	TGACTATTTA
781	TGTTTCTGGC	CAATAGAGGA	GAGGAATACT	ACGTTACGTA	TGGAGTTTGA	ACCCCTCACA
841	TCAACTTATT	AAGTGAGTTA	TCCCTCAATC	ACGATTCAAC	TATGATTCCG	CTAACTTCAA
901	AGAATATTGA	GTTAATTATT	CAAATGATTA	GTCCAAAATT	ATTTAATAAA	GTTATACTAT
961	TTCTTTTATT	TGTAACATAC	ATCTTTTGTG	TACATATTTA	GTTAAATCTT	AGCCCAACCC
1021	TATCGTTGGT	TCGACTTTTT	TTCTTTTAAAT	TTTGATTTAT	TCTGTTCCGT	AATTTTCGCTT
1081	TGTTTGGCTT	GAATAATAAC	TAGTGCATAA	AGTCATATAC	TCTAATATTT	TTAATTGAAG
1141	TACTCACAAA	TACAAAAATA	AAAAACATTC	TAAGCTCACA	TGATAGTTGA	CAAAATCTTT
1201	ATCCAAAACA	AGAGGCGGAG	CCAGGATTTG	AAACTTATGG	GTTTCAGAAAT	CTAAATCTCT
1261	TAAGTTACTA	GGTTCTAAAT	TAATAATTTA	TACATGTTCA	ATGAATTTCT	TAAGACAAAT
1321	ACATAGTTTG	AACGAAAGCT	ACTGGGTTCG	GCCGAATCCG	TAAGTTATAC	TCTCCCTCCG
1381	CCCCGGTCCA	AAACCAGCTA	GTATCAATAG	AGAGAGAGAG	AGAGAGAGAG	AGATAATAAA
1441	TTTGACCATT	GACAATGGCT	TATTACTTGC	TTAGAGTTAA	TTGGTGAAC	TAGAGAATAT
1501	AATAAGGAAT	ATTTAAACAG	ATACGTCATC	AATCCACGAG	TAACGAAGTA	AGAAATACCC
1561	TAAAATCGTA	GAAACATTAC	GTTAAATTGC	TTGACAGCCT	ATCTAGTAAG	AGTCAAAATC
1621	TACTATCTAT	CTTGTTCCGC	CATTTTCTTA	AAGAAGTACA	TGAGCTTTAT	CATCCACCTC
1681	AACATGAATG	CAAAAGAAAA	TTATTGTGCA	ACTTAATATG	TTATAATCAA	TGATATGTGT
1741	CTTGTTGTAAC	AAAGTATATA	TTTCGATACG	ATATTAATAT	GTAGGTGTTA	TATTTTATAA
1801	TATCAAAATAT	CATACTTAAC	ACCGATTTTG	TAAAAACTTA	GGCCAATTAC	CCTACCAACT
1861	AAAATACTGT	ATATCAAAACA	CTAATGTTTT	CTATTTTCGGT	ACGACAGTTC	TCTATTTACC
1921	ATATTATGGA	ATTATGCCCA	TCCTACAGTT	ACCTATAAAA	AGGAAGTTGC	CGATAGTTAT

FIGURA 2-1

```

1981 ATTCTCAACT TCTTATCTAA AAATCCATAA TGCTTTCTCC CATAGAAGCC ATTGTAGGAC
2041 TAGTAACCTT CACATTTCTC TTCTTCTTCC TATGGACAAA AAAATCTCAA AAACCTTCAA
2101 AACCTTTACC ACCGAAAATC CCCGGAGGAT GGCCGGTAAT CGGCCATCTT TTCCACTTCA
2161 ATGACGACGG CGACGACCGT CCATTAGCTC GAAAACTCGG AGACTTAGCT GACAAATACG
2221 GCCCCGTTTT CACTTTTCGG CTAGGCCCTC CCCTTGTCTT AGTTGTAAGC AGTTACGAAG
2281 CTGTAAAAGA CTGTTTCTCT ACAAATGACG CCATTTTTTTC CAATCGTCCA GCTTTTCTTT
2341 ACGGCGATTA CCTTGGCTAC AATAATGCCA TGCTATTTTT GGCCAATTAC GGACCTTACT
2401 GCGAAAAAAA TCGAAAATTA GTTATTCAGG AAGTTCTCTC CGCTAGTCGT CTCGAAAAAT
2461 TCAAACACGT GAGATTTGCA AGAATTCAGG CGAGCATTAA GAATTTATAT ACTCGAATTG
2521 ATGGAAATTC GAGTACGATA AATTTAATCG ATTGGTTAGA AGAATTGAAT TTTGGTCTGA
2581 TCGTGAAGAT GATCGCTGGA AAAAATTATG AATCCGGTAA AGGAGATGAA CAAGTGGAGA
2641 GATTTAAGAA AGCGTTTAAG GATTTTATGA TTTTATCAAT GGAGTTTGTG TTATGGGATG
2701 CATTTCCAAT TCCATTATTT AAATGGGTGG ATTTTCAAGG GCATGTTAAG GCTATGAAAA
2761 GGACTTTTAA AGATATAGAT TCTGTTTTTC AGAATTGGTT AGAGGAACAT ATTAATAAAA
2821 GAGAAAAAAT GGAGGTTAAT GCAGAAGGGA ATGAACAAGA TTTTATTGAT GTGGTGCTTT
2881 CAAAAATGAG TAATGAATAT CTTGGTGAAG GTTACTCTCG TGACTACTGC ATTAAGCAA
2941 CGGTGTTTGT AAGTTCATCT GTCATTTTTC ATTTATTCAC TTTTATTTG AGGAGCAGAC
3001 ATGTTAATAA TAATTTGGAG CAACTGTAAA GTTATCTATG TGTACAGGTT CGAGCCTCAG
3061 GTGCAACCAC TAATGCTTGT ATTAGATTAT GTTGTCTGCA TCATACCCCT AATTGGAGTG
3121 TGGCTCTTCC CGAACCTTGC AATGCTGGAT GCTGGATGCT TTATGTATCA GACTGACCTT
3181 TTTGTTAAAC TATCTAAATA CTAAGGATGA TTTAATAAAA ATATAGAATG GTAAACAGAA
3241 AAAGATGAGA TTATTTTTGG GGCTATATGG ATTCGCCCCG GCTTTGGGAG GTAAACGGT
3301 ATCTACCAGT TGAGACTTTA CTCCAGAATC TTATCTCGAG AGCTCTGAAT AAAAATGAAA
3361 TAGTATTTAC CACTCCAAAA TCTTTGATGG TAAAAAGATG AGATATAACC TCTTATAATT
3421 GATTGAACCA CGTTGATAGA ATAAACTTTC TTTACTCCCA TTCAGCATAA GAAAAATGAA
3481 ACCAAACGGA ATTCTTCTCT TTTTLAGGGG GAAATTCCTT AATTGCTTGT TGAATATAGA
3541 TTCATGTCGT TATTCTATTT TTAATAATGA TGAAAATCAA TATAGTCAAA GTTAATACTT
3601 ATGTCATTTG GTTTCGCGAC AAGTTATATT GGAACATAT ATACAGTCTA TTATAGAATA
3661 GTGATTATTT AGAGGATATA CATTTTTTTT GGATAAATAT TTGATTTATT GGATTAAAAA
3721 TAGAATATAC AGGTAAGGTC TAAAACGTGT GTTTGCTTTT AACTAAATA AACTTGACCT
3781 CGTACAATTC TAAGAAAAATA TTTGAAATAA ATGAATTATT TTATTGTTAA TCAATTAAAA
3841 AAATCATAGT ATAGATGAGA TGTGTGCATA CTTGACAATA ACTATACTAA CTAAACAAAG
3901 GTATGTGAAT AATTGATATT CCTTTTTTAA TTATCTTTT TTCCAGAGTT TGGTCTTGGA
3961 TGCAGCAGAC ACAGTTGCTC TTCACATAAA TTGGGGAATG GCATTATTGA TAAACAATCA
4021 AAAGGCCTTG ACGAAAGCAC AAGAAGAGAT AGACACAAAA GTTGGTAAGG ACAGATGGGT
4081 AGAAGAGAGT GATATTAAGG ATTTGGTATA CCTCCAAGCT ATTGTTAAAG AAGTGTACG
4141 ATTATATCCA CCAGGACCTT TGTTAGTACC ACACGAAAAT GTAGAAGATT GTGTTGTTAG
4201 TCGATATCAC ATTCTAAAG GGACAAGATT ATTCGCAAAC GTCATGAAT TGCAACGTGA
4261 TCCTAAACTC TGGTCTGATC CTGATACTTT CGATCCAGAG AGATTCAATT CTACTGATAT
4321 TGACTTTCGT GGTCACTACT ATAAGTATAT CCCGTTTGGT TCTGGAAGAC GATCTTGTCC
4381 AGGGATGACT TATGCATTGC AAGTGGAACA CTTAACAATG GCACATTTGA TCCAAGGTTT
4441 CAATTACAGA ACTCCAAATG ACGAGCCCTT GGATATGAAG GAAGGTGCAG GCATAACTAT
4501 ACGTAAGGTA AATCCTGTGG AACTGATAAT AGCGCCTCGC CTGGCACCTG AGCTTTATTA
4561 AAACCTAAGA TCTTTCATCT TGGTTGATCA TTGTATAATA CTCCTAAATG GATATTCATT
4621 TACCTTTTAT CAATTAATTG TCAGTACGAG TTTTCTAAT TTGGTACATT TGAATAATA
4681 AGTAAAGAAT AATTGTGCTA ATATATAAAG GTTTGTAGAA GATAATTGAC TGGTTGTACC
4741 ACAATCTCCA GTGAAAGTGT TAATTATTTA CTTGATCCAC AGCTTATTCT ATGTTTGAAA
4801 TTTGCCTAGT GTCATGATAT TACTCCATCA AATCAAGAA ATAATCATTT CCAACTTTTG
4861 CTGGACTGGA CGATCTTTCA ATAATAAAGG ATCTTTAATT TGCCAAAGTT GAGATCAAAA
4921 TACTGGTCGC TTTACCAATA AGAATGAAAT GTGATGGAAA TTATGTACGT TGGGATAAGG
4981 GAACACAAC ATCAAGGAGA CTAAAAGGTA CGTAAAGGAA AAGAAAAAAT TTGCCATTGA
5041 TTGCTACTAA GTAACCTAAC AAAATCTTTC AGAAAAGAAT CACTTGTATA AGTCGGGGTT
5101 GAAAGTTTTG GTGTCTCTTT TCTTATGTAT TGTGTCTTT AGACAGTATT GTACTTAGTT
5161 ATTTCAGAAT TTTATTTTCG TATTAGAGCT CAAGACTCTG TATTTATTAG TTCTGGGAGA
5221 ATTATCATGT ATTTTCAGTC TTTTGTATT TCTGTAAATT CCGCTATTTT GGTTCTTTAT
5281 TGCTATGTTT GGCTTTCCTA GAAAATGTGT TAGGCGCTAT CACGACTGAT TGAGATTTTG
5341 TATCGTGACT GATAATTACA CGGTTTAGTA AATTTTGATA TTTTCAAAA GAGTTTTTTT

```

FIGURA 2-2

```

5401 AATAAAATAT GCAACTTCAG TCAAAACATA CAACGTTTTG TTGTATAAAT CCGATCAAAA
5461 CATATAACTT ACATAAAACT TCGGTATGAA TTTTTTGTTG TATAAACATT TGGTTAAAAAC
5521 ATATAACTTA CATATAATTT GCATACAACT TACATAAAAC TTGCATATAA ACAATTTATT
5581 TATGTCTTTG TTTTGTAGTA TCAATTTGAA ATTCCAACAA AAACAAACTC TAATTTTAC
5641 CAAACTCTCT CAAAATTGAG TTATAGATTT CAAAAGATAT CCTTAATCGT TTGCAATTGT
5701 AACAATCCGA TCGGCCGTTT TGAGATCTAG CGTGTGTGTTT GCGCGTTTGA GACCTTGAGT
5761 AACTTCACPT TATGTTGTAT GACTTGTATA TGTGGTCGGA ATTAAATTTT GGAAGTTCA
5821 GAGTTGATTC GGATGAAAAA TTCTAATTTT GGAAGTTTGA AGATGGAATG ATTGACTAAG
5881 GATTGACGTT TGAGTAAACG ATCTCGGAAT CGAGATTTGA AGGTTCCAAT AGGTTTCGTAT
5941 GATGATTTCA GACTTGAGCG TATGTTTGGG TTGAGTATCG GGTGGTCCGG GAGCATTTCA
6001 ACGCTGATTA TAGAAAATTG GCATCTTAAA GGTTTTAGAA TTTCATAAGT TTGGTTTGAA
6061 GTGGATTTTG ATATTATCGG TGTCCATTTG GAGTTTCGAG CCTTGGAATA GGTTCGTATC
6121 GTAAATTTTG ACTTTAGTGT AAAAGTTCGGC GTCATTCCGG AGTGTTTGA TAAGATTCTG
6181 ATGCGTTCGT CGAAGTTTGG AAGTTTGAAA GTTGAAAAGA AGATTTTAA TAGGCGATTC
6241 ATGATTTTGA TGTTATTTGT GTCGAGCCTT TGGATAAGTT TGTGTGAGGT ATGGGACTTG
6301 TTGGTATGAA TGGACGAGCT CTACGGGGGC CTCGAGTAAG TTTTCGGA

```

FIGURA 2-3

Secuencias codificantes 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	CCCCGGAGGA
121	TGGCCGGTAA	TCGGCCATCT	TTTCCACTTC	AATGACGACG	GCGACGACCG	TCCATTAGCT
181	CGAAAACTCG	GAGACTTAGC	TGACAAATAC	GGCCCCGTTT	TCACTTTTCTG	GCTAGGCCTT
241	CCCCTTGCTC	TAGTTGTAAG	CAGTTACGAA	GCTGTAAAAG	ACTGTTTCTC	TACAAATGAC
301	GCCATTTTTT	CCAATCGTCC	AGCTTTTCTT	TACGGCGATT	ACCTTGGCTA	CAATAATGCC
361	ATGCTATTTT	TGGCCAATTA	CGGACCTTAC	TGGCGAAAAA	ATCGAAAATT	AGTTATTTCAG
421	GAAGTTCTCT	CCGCTAGTCG	TCTCGAAAAA	TTCAAACACG	TGAGATTGTC	AAGAATTCAA
481	GCGAGCATTA	AGAATTTATA	TACTCGAATT	GATGGAAAAT	CGAGTACGAT	AAATTTAACT
541	GATTGGTTAG	AAGAATTGAA	TTTTGGTCTG	ATCGTGAAGA	TGATCGCTGG	AAAAAATTAT
601	GAATCCGGTA	AAGGAGATGA	ACAAGTGGAG	AGATTTAAGA	AAGCGTTTAA	GGATTTTATG
661	ATTTTATCAA	TGGAGTTTGT	GTTATGGGAT	GCATTTCCAA	TTCCATTATT	TAAATGGGTG
721	GATTTTCAAG	GGCATGTTAA	GGCTATGAAA	AGGACTTTTA	AAGATATAGA	TTCTGTTTTT
781	CAGAATTGGT	TAGAGGAACA	TATTAATAAA	AGAGAAAAAA	TGGAGGTTAA	TGCAGAAGGG
841	AATGAACAAG	ATTTTCATTGA	TGTGGTGCTT	TCAAAAATGA	GTAATGAATA	TCTTGGTGAA
901	GGTTACTCTC	GTGATACTGT	CATTAAAGCA	ACGGTGTTTA	GTTTGGTCTT	GGATGCAGCA
961	GACACAGTTG	CTCTTCACAT	AAATTGGGGA	ATGGCATTAT	TGATAAACAA	TCAAAAGGCC
1021	TTGACGAAAG	CACAAGAAGA	GATAGACACA	AAAGTTGGTA	AGGACAGATG	GGTAGAAGAG
1081	AGTGATATTA	AGGATTTGGT	ATACCTCCAA	GCTATTGTTA	AAGAAGTGTT	ACGATTATAT
1141	CCACCAGGAC	CTTTGTTAGT	ACCACACGAA	AATGTAGAAG	ATTGTGTTGT	TAGTGGATAT
1201	CACATTCTTA	AAGGGACAAG	ATTATTCGCA	AACGTCATGA	AACTGCAACG	TGATCCTAAA
1261	CTCTGGTCTG	ATCCTGATAC	TTTCGATCCA	GAGAGATTCA	TTGCTACTGA	TATTGACTTT
1321	CGTGGTCAGT	ACTATAAGTA	TATCCCGTTT	GGTTCGGA	GACGATCTTG	TCCAGGGATG
1381	ACTTATGCAT	TGCAAGTGGA	ACACTTAACA	ATGGCACATT	TGATCCAAGG	TTTCAATTAC
1441	AGAACTCCAA	ATGACGAGCC	CTTGGATATG	AAGGAAGGTG	CAGGCATAAC	TATACGTAAG
1501	GTAAATCCTG	TGGAAGTGAT	AATAGCGCCT	CGCCTGGCAC	CTGAGCTTTA	TTAA

Aminoácidos deducidos del gen de nicotina desmetilasa de tabaco

1	MLSPIEAIVG	LVTFTFLFFF	LWTKKSQKPS	KPLPPKIPGG	WPVIGHLFHF	NDDGDDRPLA
61	RKLGLADKY	GPVFTFRLGL	PLVLVSSYE	AVKDCFSTND	AIFSNRPAPL	YGDYLGYNNA
121	MLFLANYGPY	WRKNRKLVIQ	EVLSASRLEK	FKHVRFARIQ	ASIKNLYTRI	DGNSSTINLT
181	DWLEELNFGL	IVKMIAGKNY	ESGKGDEQVE	RFKKAFKDFM	ILSMEFVLWD	AFPPIPLFKWV
241	DFQGHVKAMK	RTFKDIDSVF	QNWLEEHINK	REKMEVNAEG	NEQDFIDVVL	SKMSNEYLGE
301	GYSRDTVICA	TVFSLVLDA	DTVALHINWG	MALLINNQKA	LTKAQEEIDT	KVGKDRWVEE
361	SDIKDLVYLQ	AIVKEVLRLY	PPGPILLVPH	NVEDCVVSGY	HIPKGTRLF	NVMKLQRDPK
421	LWSDPDTFDP	ERFIATDIDF	RGQYKYIPF	GSGRRSCPGM	TYALQVEHLT	MAHLIQGFNY
481	RTPNDEPLDM	KEGAGITIRK	VNPVELIIAP	RLAPELY		

FIGURA 3

Intrón del gen de nicotina desmetilasa de tabaco (998nt)

```

1  GTAAGTTCAT CTGTCATTTT TCATTTATTC ACTTTTATTT TGAGGAGCAG ACATGTTAAT
61  AATAATTTGG AGCAACTGTA AAGTTATCTA TGTGTACAGG TTCGAGCCTC AGGTGCAACC
121 ACTAATGCTT GTATTAGATT ATGTTGTCTG CATCATACCC CTAATTGGAG TGTGGCTCTT
181 CCCGAACCCT GCAATGCTGG ATGCTGGATG CTTTATGTAT CAGACTGACC TTTTGTGTTAA
241 ACTATCTAAA TACTAAGGAT GATTTAATAA AAATATAGAA TGGTAAACAG AAAAAGATGA
301 GATTATTTTT GGGGCTATAT GGATTCGCCC GGGCTTTGGG AGGTAAAACG GTATCTACCA
361 GTTGAGACTT TACTCCAGAA CTTTATCTCG AGAGCTCTGA ATAAAAATGA AATAGTATTT
421 ACCACTCCAA AATCTTTGAT GGTAAAAAGA TGAGATATAA CCTCTTATAA TTGATTGAAC
481 CACGTTGATA GAATAAAACT TCTTTACTCC CATTTCAGCAT AAGAAAAATG AAACCAAACG
541 GAATTCTTCT CTTTTTAGG GGGAAATTCC TTAATTGCTT GTTGAATATA GATTCATGTC
601 GTTATTCTAT TTTTAATAAT GATGAAAATC AATATAGTCA AAGTTAATAC TTATGTCATT
661 TGGTTTGCGG ACAAGTTATA TTGGAATAT ATAATACGTC TATTATAGAA TAGTGATTAT
721 TTAGAGGATA TACATTTTTT TTGGATAAAT ATTTGATTTA TTGGATTAAA AATAGAATAT
781 ACAGGTAAGG TCTAAAACGT GTGTTTGCTT TTACACTAAA TAAACTTGAC CTCGTACAAT
841 TCTAAGAAAA TATTTGAAAT AAATGAATTA TTTTATTGTT AATCAATTAA AAAAATCATA
901 GTATAGATGA GATGTGTGCA TACTTGACAA TAACATATACT AACTAAAACA AGGTATGTGA
961 ATAATTGATA TTCCTTTTTT AATTATTCTT TTTTCCAG

```

FIGURA 4

Promotor del gen de nicotina desmetilasa de tabaco (2009 nt)

```

1  TCTCTAAAGT CCCCTTCCAC TTTATCTTAG CTGTGTGATT TCTTTCAGAC AACCTTATTT
61  TTATTCAGAC TCTTATTTGT ATTATTTCTAG AAGCTCGTGT ACTTGTGACA CCAGTTCTGG
121 GATGGTATTT AGATATCGCT ATTATTTTGG CTTATTCACT TCAGTTCAGA TTTTATTCCA
181 GTTATTTGAT TTCTTTATTA TTAATCAAAT TGAATTGTTA AAAATGGTTA AAATTACTCC
241 AATGTTGGCT TTCCTAGTAA GCGAAATATC AGGCGCCATC ACGGTACCCG AAGGTGAGAA
301 TTTAGATCG TGACAGCCGC ATCTCAAGGG GTGTGATGTA AACAGTTTAC GATGGTGCAA
361 GCATTAGTGG CTGCTTCGAC GACTTAAATC CGTAACTTAT AGATCACACG AATACAACCTT
421 TACTATTTTA ACACCCAGCA AATTCCTGAT AAAACAATT CTAACATAGC ACATCAAAAT
481 GTAAATGATT GAAGAAAAAG ATGACTTTTA TAGACAGAGA AAAAAACAGT TACGTGTGGG
541 GTTGAATAGA GATTGTGGCT ATGCTATTTT TAATATTGAA ATTCACCGAC TTTTTTAGTT
601 CAATACGAAA AGAGTAAGTG AAAAGGTCTG AAAAGGAAAA GGACAATGCC TAAAAGGACA
661 CATTCAGAAC ACATACACTG AATGATTCTA ATTTCTAGTC CGAAGATTTT TAGTTCGAGG
721 ATAACGTGAG CAATCGGTTA CTTCCCATTA GCAATTGCCA ACTGGATGTT TGACTATTTA
781 TGTTTCTGGC CAATAGAGGA GAGGAATACT ACGTTACGTA TGAGATTTGA ACCCTTCACA
841 TCAACTTATT AAGTGAGTTA TCCCTCAATC ACGATTCAAC TATGATTCCG CTAACCTCAA
901 AGAATATTGA GTTAATTATT CAAATGATTA GTCCAAAATT ATTTAATAAA GTTATACTAT
961 TTCTTTTATT TGTAACATAC ATCTTTTGTG TACATATTTA GTTAAATCTT AGCCCAACCC
1021 TATCGTTGGT TCGACTTTTT TTCTTTTAAAT TTTGATTTAT TCTGTTCCGT AATTTCCGCTT
1081 TGTTTGGCTT GAATAATAAC TAGTGCATAA AGTCATATAC TCTAATATTT TTAATTGAAG
1141 TACTCACAAA TACAAAAATA AAAAACATTC TAAGCTCACA TGATAGTTGA CAAAATCTTT
1201 ATCCAAAACA AGAGGCGGAG CCAGGATTTG AAACCTATGG GTTCAGAATT CTAAATCTCT
1261 TAAGTTACTA GGTTCTAAAT TAATAATTTA TACATGTTCA ATGAATTTCT TAAGACAAAT
1321 ACATAGTTTG AACGAAAGCT ACTGGGTTCTG GCCGAATCCG TAAGTTATAC TCTCCCTCCG
1381 CCCCGGTCCA AAACCAGCTA GTATCAATAG AGAGAGAGAG AGAGAGAGAG AGATAATAAA
1441 TTTGACCATT GACAAATGGC TATTACTTGC TTAGAGTTAA TTGGTGAAC TTAGAGAAAT
1501 AATAAGGAAT ATTTAAACAG ATACGTCATC AATCCACGAG TAACGAAGTA AGAAATACCC
1561 TAAAATCGTA GAAACATTAC GTTAAATTGC TTGACAGCCT ATCTAGTAAG AGTCAAAATC
1621 TACTATCTAT CTTGTTCCGC CATTTCCTTA AAGAAGTACA TGAGCTTTAT CATCCACCTC
1681 AACATGAATG CAAAAGAAAA TTATTGTGCA ACTTAATATG TTATAATCAA TGATATGTGT
1741 CTTGTGTAAC AAAGTATATA TTTCGATACG ATATTAATAT GTAGGTGTTA TATTTTAAAA
1801 TATCAAATAT CATACTTAAC ACCGATTTT TAAAACTTA GGCCAATTAC CCTACCAACT
1861 AAAATACTGT ATATCAAACA CTAATGTTTT CTATTTCCGT ACGACAGTTC TCTATTTACC
1921 ATATTATGGA ATTATGCCCA TCCTACAGTT ACCTATAAAA AGGAAGTTGC CGATAGTTAT
1981 ATTCTCAACT TCTTATCTAA AAATCCATA

```

FIGURA 5

UTR en 3' del gen de nicotina desmetilasa de tabaco

```

1 AACCTAAGAT CTTTCATCTT GGTGATCAT TGTATAATAC TCCTAAATGG ATATTCATTT
61 ACCTTTTATC AATTAATTGT CAGTACGAGT TTTTCTAATT TGGTACATTT GTAATAATAA
121 GTAAAGAATA ATTGTGCTAA TATATAAAGG TTTGTAGAAG ATAATTGACT GGTGTACCA
181 CAATCTCCAG TGAAAGTGTT AATTATTTAC TTGATCCACA GCTTATTCTA TGTTTGAAAT
241 TTGCCTAGTG TCATGATATT ACTCCATCAA ATTCAAGAAA TAATCATTTT CAACCTTTTGC
301 TGGACTGGAC GATCTTTCAA TAATAAAGGA TCTTTAATTT GCCAAAGTTG AGATCAAAAT
361 ACTGGTCGCT TTACCAATAA GAATGAAATG TGATGGAAT TATGTACGTT GGGATAAGGG
421 AACACAAC TAACCTAACA AAATCTTTCA GAAAAGGAAA AGAAAAAATT TGCCATTGAT
481 TGCTACTAAG TAACCTAACA AAATCTTTCA GAAAAGGAAA AGAAAAAATT TGCCATTGAT
541 AAAGTTTGG TGTCTCTTTT CTTATGTATT GTTGTCTTTA GACAGTATTG TACTTAGTTA
601 TTTCAGAAAT TTATTTTCGT ATTAGAGCTC AAGACTCTGT ATTTATTAGT TCTGGGAGAA
661 TTATCATGTA TTTTCAGTCT TTTGTTATTT CTGTAAATTC CGCTATTTTG GTTCTTTATT
721 GCTATGTTTC GCTTTCCTAG AAAATGTGTT AGGCGCTATC ACGACTGATT GAGATTTTGT
781 ATCGTGAATG ATAATTACAC GGTTTAGTAA ATTTTGATAT TTTCAAAAAG AGTTTTTTTA
841 ATAAAAATAG CAACCTCAGT CAAAACATAC AACGTTTGT TGTATAAATC CGATCAAAAC
901 ATATAACTTA CATAAACTT GCGTATGAAT TTTTGTGTTG ATAAACATTT GGTAAAACA
961 TATAACTTAC ATATAATTG CATACAACCT ACATAAACT TGCAATATAA CAATTTATTT
1021 ATGTCTTGT TTTTGAGTAT CAATTTGAAA TTCCAACAAA AACAACTCT AATTTTACC
1081 AAACCTCTCT AAAATTGAGT TATAGATTTT AAAAGATATC CTTAATCGTT TGCAATTGTA
1141 ACAATCCGAT CGGCCGTTTT GAGATCTAGC GTGTTGTTTG GCGGTTTGAG ACCTTGAGTA
1201 ACTTCACTTT ATGTTGTATG ACTTGTATAT GTGGTCGGAA TTAAATTTTC GGAAGTTCAG
1261 AGTTGATTCG GATGAAAAAT TCTAATTTTC GAAGTTTAA GATGGAATGA TTGACTAAGG
1321 ATTGACGTTT GAGTAAACGA TCTCGGAATC GAGATTTGAA GGTTCGAATA GGTTCGTATG
1381 ATGATTTTCA ACTTGAGCGT ATGTTTGGGT TGAGTATCGG GTGGTCCGGG AGCATTTCAA
1441 CGCTGATTAT AGAAAATTGG CATCTTAAAG GTTTTAGAAT TTCATAAGTT TGGTTTGAAG
1501 TGGATTTTGA TATTATCGGT GTCCATTTGG AGTTTCGAGC CTTGGAATAG GTTCGTATCG
1561 TAAATTTTGA CTTTAGTGTA AAGTTCGGCG TCATTCCGGA GTGTTTGTAT AAGATTCTGA
1621 TCGGTTTCGTC GAAGTTTGGG AGTTTGAAA TGGAAAAGAA GATTTTAAAT AGGCGATTCA
1681 TGATTTTGAT GTTATTTGTG TCGAGCCTTT GGATAAGTTT GTGTGAGGTA TGGGACTTGT
1741 TGGTATGAAT GGACGAGCTC TACGGGGGCC TCGAGTAAGT TTCGGA

```

FIGURA 6