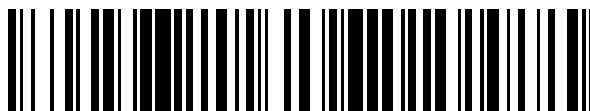


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 664 569**

51 Int. Cl.:

**C07K 14/415** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.07.2011 PCT/GB2011/051340**

87 Fecha y número de publicación internacional: **26.01.2012 WO12010872**

96 Fecha de presentación y número de la solicitud europea: **18.07.2011 E 11748702 (5)**

97 Fecha y número de publicación de la concesión europea: **06.12.2017 EP 2596012**

54 Título: **Citocromo P450 de plantas**

30 Prioridad:

**22.07.2010 GB 201012262**  
**22.12.2010 GB 201021707**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**20.04.2018**

73 Titular/es:

**SUN PHARMACEUTICAL INDUSTRIES  
(AUSTRALIA) PTY LIMITED (100.0%)**  
**Level 4, 1 Acacia Place**  
**Notting Hill VIC 3168, AU**

72 Inventor/es:

**WINZER, THILO;**  
**WALKER, TRACY CAROL y**  
**GRAHAM, IAN ALEXANDER**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

ES 2 664 569 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

Citocromo P450 de plantas

## 5 Introducción

Esta divulgación se refiere al aislamiento y secuenciación de moléculas de ácidos nucleicos que codifican para un nuevo citocromo P450s de un cultivar de *Papaver somniferum* [adormidera]; células transgénicas transformadas con dichas moléculas de ácidos nucleicos, secuenciar las variantes del gen; el uso de dichos genes/proteínas en la producción de noscapina y el uso de los genes como marcadores de plantas de adormidera que sintetizan la noscapina.

## Antecedentes

- 15 El citocromo P450s de plantas son una gran familia de enzimas responsables de la oxidación, peroxidación y reducción de un amplio número de metabolitos intermedios de la planta como alcaloides, terpenoides, lípidos, glicósidos y glucosinolatos. Se sabe que los P450s están implicado en el metabolismo y la desintoxicación de pesticidas así como en la biosíntesis de metabolitos primarios y secundarios.
- 20 Los citocromos P450s de las plantas son conocidos en la técnica y han sido clonados, expresados y caracterizados con éxito. Por ejemplo, los documentos WO2009/064771 y WO2008/070274 describen cada uno, los genes del citocromo P450 y su uso en la alteración del contenido alcaloide en la *Nicotiana tabacum*. Estas solicitudes de patente describen cómo la inhibición del P450s específico reduce la cantidad de N' nitrosomonicotina, un conocido carcinógeno, in planta. El documento WO2008/150473 describe la sobreexpresión del citocromo P450s para conferir
- 25 resistencia o tolerancia a herbicidas, en particular, benzotiadiazonas y sulfonilureas. En el documento WO2008/088161 describe unas plantas transgénicas que sobreexpresan un citocromo P450 lo cual hace aumentar el tamaño de la semilla o el contenido proteico almacenado de las semillas. La sobreexpresión confiere también una mayor resistencia al estrés hídrico. Lo es evidente es que el citocromo P450s de las plantas presenta varias funciones en la regulación de las actividades bioquímicas en las células vegetales y son conocidas en la
- 30 técnica.

La amapola del opio *P. somniferum* es la planta de la que se extrae el opio. La amapola de opio es la única amapola comercialmente explotada de la familia *Papaveraceae* y es la principal fuente de opiáceos naturales. El opio se extrae del látex cosechado de las vainas de semillas verdes. Otra fuente de alcaloides opiáceos es la paja de adormidera que es la planta madura seca. *P. somniferum* es una fuente de alcaloides opiáceos clínicamente útiles como morfina, codeína, tebaina, noscapina [también conocida como narcotina] y papaverina. La aplicación clínica de estos alcaloides opiáceos y sus derivados es amplia y tiene uso como analgésicos, antitusígenos y antiespasmódicos. Aunque no se usa como un agente farmacológico en sí mismo, la tebaina es un opiáceo particularmente útil que se puede convertir en una gama de compuestos tales como hidrocodona, oxicodona, oximorfona, naltrefina, naltrexona, buprenorfina y etorfina. Estos productos intermedios también tienen amplias aplicaciones farmacéuticas. Por ejemplo, oxicodona, oximorfona y etorfina se utilizan ampliamente como analgésicos para el dolor moderado a intenso y, a menudo, se combinan con otros analgésicos como el ibuprofeno. La buprenorfina se usa en el tratamiento de la adicción a la heroína y el dolor crónico. Naltrexona se usa en el tratamiento del alcohol y la adicción a los opiáceos.

- 45 Esta divulgación se refiere a la identificación y caracterización del citocromo P450s aislado a partir de un cultivar de *Papaver somniferum* que llamamos PSCYP1, PSCYP2 y PSCYP3. La proteína prevista codificada por PSCYP1 presenta una identidad de secuencia superior al citocromo P450 de *Coptis japonica* (Nº de acceso de GenBank: BAF98472.1, 46% de identidad). El homólogo más cercano con una asignación a la subfamilia del citocromo P450 es CYP82C4 de *Arabidopsis lyrata* (nº de referencia de la secuencia del NCBI XP\_002869304.1, 44% de identidad).
- 50 Se ha mostrado que la proteína CYP82C4 de *Arabidopsis thaliana* añade un grupo hidroxilo en la posición 5 del 8-metoxipsoraleno, una furocumarina, creando el 5-hidroxi-8-metoxipsoraleno (Kruse et al. (2008) Chemistry & Biology 15: 149-156). Los homólogos más cercanos de la proteína prevista codificada por PSCYP2 se anotan como estilopina sintasas de *Argemone mexicana* (nº de acceso de GenBank: ABR14721, 77% de identidad), *Papaver somniferum* (nº de acceso de GenBank: ADB89214, 76% de identidad) y *Eschscholzia californica* (nº de acceso de GenBank: BAD98250, 72% de identidad). Pertenecen a la subfamilia CYP719A del citocromo P450s que se ha encontrado únicamente en especies vegetales que producen alcaloide de isoquinolina donde catalizan la formación de puentes metilenodioxo (Ikezawa et al. (2009) Plant Cell Rep. 28:123-133). El homólogo más cercano de la proteína prevista codificada por PSCYP3 está anotada como protopina 6-hidroxilasa de *Eschscholzia californica* (nº de acceso de GenBank: BAK20464, 44% de identidad). El homólogo más cercano con una asignación a la
- 60

subfamilia del citocromo P450 es CYP82C4 de *Arabidopsis lyrata* mencionada anteriormente (42% de identidad). Sorprendentemente, *PSCYP1*, *PSCYP2* y *PSCYP3* son exclusivos de los cultivares de *Papaver somniferum* que producen noscapina. Los cultivares que no producen noscapina no incluyen este gen.

## 5 **Declaraciones de la invención**

De acuerdo con un aspecto de la invención se proporciona una molécula aislada de ácido nucleico que codifica un polipéptido del citocromo P450 donde dicha molécula de ácido nucleico comprende o consiste en una secuencia de nucleótido seleccionada de entre el grupo que consiste en:

- 10 i) una secuencia de nucleótidos representada por la secuencia de la Figura 1c, 1d ó 3c.
- ii) una secuencia de nucleótidos donde dicha secuencia está degenerada como resultado del código genético de la secuencia de nucleótidos definida en (i);
- 15 iii) una molécula de ácido nucleico cuya cadena complementaria hibrida con secuencias de nucleótidos con al menos un 90% de identidad bajo condiciones de hibridación muy astringentes a la secuencia en la Figura 1c, 1d ó 3c donde dicha molécula de ácido nucleico codifica un polipéptido del citocromo P450;
- iv) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos como se representa en la Figura 4c;
- 20 v) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos donde dicha secuencia de aminoácidos se modifica por adición, delección o sustitución de al menos un residuo de aminoácido como se representa en iv) anterior y que presenta al menos un 85% de identidad de la secuencia en relación a la secuencia de longitud completa representada en la Figura 4c que ha retenido la actividad del citocromo P450.
- 25 La hibridación de una molécula de ácido nucleico se produce cuando dos moléculas de ácido nucleico complementarias experimentan un número de enlaces de hidrógeno entre sí. La rigurosidad de la hibridación puede variar de acuerdo con las condiciones ambientales que rodean a los ácidos nucleicos, la naturaleza del método de hibridación y la composición y longitud de las moléculas de ácido nucleico utilizadas. Los cálculos con respecto a las condiciones de hibridación requeridas para alcanzar grados particulares de rigor se discuten en Sambrook et al.,
- 30 Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 2001); y Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes Part I, Chapter 2 (Elsevier, New York, 1993). La  $T_m$  es la temperatura a la que el 50% de una cadena dada de una molécula de ácido nucleico se hibrida con su cadena complementaria. Lo siguiente es un conjunto ejemplar de condiciones de hibridación y no es restrictivo:

### 35 **Rigor Muy Alto (permite secuencias para compartir al menos 90% de identidad para hibridar)**

Hibridación:	5x SSC a 65 ° C durante 16 horas
Lavar dos veces:	2x SSC a temperatura ambiente (RT) durante 15 minutos cada uno
Lavar dos veces:	0.5x SSC a 65 ° C durante 20 minutos cada uno

### 40 **Rigor Alto (permite secuencias para compartir al menos 80% de identidad para hibridar)**

Hibridación:	5x-6x SSC a 65 ° C-70 ° C durante 16-20 horas
Lavar dos veces:	2x SSC a temperatura ambiente durante 5-20 minutos cada uno
Lavar dos veces:	1x SSC a 55 ° C-70 ° C durante 30 minutos cada uno

### **Rigor Bajo (permite secuencias para compartir al menos 50% de identidad para hibridar)**

Hibridación:	6x SSC a temperatura ambiente a 55 ° C durante 16-20 horas
Lavar al menos dos veces:	2x-3x SSC a temperatura ambiente a 55 ° C durante 20-30 minutos cada uno

- 45 En una realización preferida de la invención, dicha molécula de ácido nucleico comprende o consiste en una secuencia de nucleótidos como se representa en la Figura 1c ó 1d.

De acuerdo con un aspecto adicional de la invención, se proporciona un polipéptido aislado seleccionado de entre el grupo que consiste en:

50

i) un polipéptido que comprende o que consiste en una secuencia de aminoácidos como se representa en la Figura 4c; o

ii) un polipéptido modificado donde los polipéptidos variantes presentan al menos un 85% de identidad en relación a la secuencia de longitud completa de aminoácidos representada en la Figura 4c y que retiene la actividad del citocromo P450.

En una realización preferida dicho polipéptido aislado comprende o consiste en una secuencia de aminoácidos que es un 85% idéntica a la secuencia de aminoácidos de longitud completa en la Figura 4c y que ha retenido la actividad del citocromo P450.

10

Un polipéptido modificado como se describe en el presente documento puede diferir en la secuencia de aminoácidos por una o más sustituciones, adiciones, deleciones, truncamientos que pueden estar presentes en cualquier combinación. Entre las variantes preferidas están aquellas que varían de un polipéptido de referencia por sustituciones conservadoras de aminoácidos. Tales sustituciones son aquellas que sustituyen un aminoácido dado por otro aminoácido de características similares. La siguiente lista no excluyente de aminoácidos se considera reemplazos conservadores (similares): a) alanina, serina y treonina; b) ácido glutámico y ácido aspártico; c) asparagina y glutamina d) arginina y lisina; e) isoleucina, leucina, metionina y valina y f) fenilalanina, tirosina y triptófano. Las más preferidas son las variantes que retienen o potencian la misma función y actividad biológica que el polipéptido de referencia del que varía.

20

En una realización, los polipéptidos variantes presentan al menos un 90%, 95% de identidad y al menos un 99% de identidad con la secuencia de aminoácidos de longitud completa ilustrada en el presente documento. De acuerdo con otro aspecto de la invención, se proporciona un vector que comprende una molécula de ácido nucleico que codifica un polipéptido citocromo P450 de acuerdo con la invención donde dicha molécula de ácido nucleico está unida operativamente a una molécula de ácido nucleico que comprende una secuencia promotora.

25

En una realización preferida de la invención, dicha secuencia de ácido nucleico comprende un promotor que confiere una expresión constitutiva sobre dicho polipéptido citocromo P450.

30

En una realización preferida alternativa de la invención, dicha secuencia de ácido nucleico comprende un promotor que confiere una expresión regulada sobre dicho polipéptido citocromo P450.

En una realización preferida de la invención, dicha expresión regulada es una expresión regulada por tejido o mediante el desarrollo.

35

En otra realización alternativa de la invención dicha expresión regulada es una expresión inducible.

En una realización alternativa de la invención un vector que incluye una molécula de ácido nucleico de acuerdo con la invención no necesita incluir un promotor u otra secuencia reguladora, particularmente si el vector se debe usar para introducir la molécula de ácido nucleico en las células para la recombinación en el gen.

40

Preferiblemente, la molécula de ácido nucleico en el vector está bajo el control de, y operativamente unida a, un promotor apropiado u otros elementos reguladores para la transcripción en una célula huésped tal como una célula microbiana (por ejemplo, bacteriana, de levadura) o vegetal. El vector puede ser un vector de expresión bifuncional que funciona en múltiples huéspedes. En el caso del ADN genómico del citocromo P450, este puede contener su propio promotor u otros elementos reguladores y, en el caso del ADNc, puede estar bajo el control de un promotor adecuado u otros elementos reguladores para la expresión en la célula huésped.

45

Por "promotor" se entiende una secuencia de nucleótidos ascendente del sitio de iniciación de la transcripción y que contiene todas las regiones reguladoras requeridas para la transcripción. Los promotores adecuados incluyen promotores constitutivos, específicos de tejido, inducibles, de desarrollo u otros promotores para la expresión en células vegetales comprendidas en plantas que dependen del diseño. Tales promotores incluyen promotores virales, fúngicos, bacterianos, animales y derivados de plantas capaces de funcionar en células vegetales.

50

Los promotores constitutivos incluyen, por ejemplo el promotor CaMV 35S (Odell et al. (1985) Nature 313, 9810-812); actina del arroz (McElroy et al. (1990) Plant Cell 2: 163-171); ubiquitina (Christian et al. (1989) Plant Mol. Biol. 18: (675-689); pEMU (Last et al. (1991) Theor Appl. Genet. 81: 581-588); MAS (Velten et al. (1984) EMBO J. 3. 2723-2730); promotor ALS (solicitud de Estados Unidos n.º de Serie 08/409,297), y similares. Otros promotores constitutivos incluyen los descritos en las patentes de Estados Unidos Nos. 5.608.149; 5.608.144; 5.604.121; 5.569.597; 5.466.785; 5.399.680; 5.268.463; y 5.608.142, cada uno de los cuales se

60

incorpora por referencia.

Los promotores regulados químicamente se pueden usar para modular la expresión de un gen en una planta mediante la aplicación de un regulador químico exógeno. Dependiendo del objetivo, el promotor puede ser un promotor inducible por productos químicos, donde la aplicación de la expresión génica inducida por productos químicos, o un promotor reprimible químico, donde la aplicación del químico reprime la expresión génica. Los promotores inducibles por productos químicos son conocidos en la técnica e incluyen, pero no se limitan a, el promotor In2-2 de maíz, que se activa mediante herbicidas benzenesulfonamida, el promotor GST de maíz, que se activa mediante compuestos electrofílicos hidrófobos que se usan como pre-herbicidas de emergencia, y el promotor de tabaco PR-1a, que se activa con ácido salicílico. Otros promotores regulados químicamente de interés incluyen promotores que responden a esteroides (véase, por ejemplo, el promotor inducible por glucocorticoides en Schena et al. (1991) Proc. Natl. Acad. Sci. USA 88: 10421-10425 y McNellis et al. (1998) Plant J. 14(2): 247-257) y promotores inducibles por tetraciclina y reprimibles por tetraciclina (ver, por ejemplo, Gatz et al. (1991) Mol. Gen. Genet. 227: 229-237, y Números de Patentes de EE.UU. 5.814.618 y 5.789.156, incorporadas aquí como referencia.

15 Cuando se desea una expresión mejorada en tejidos particulares, pueden utilizarse promotores específicos de tejido. Los promotores específicos de tejido incluyen los descritos por Yamamoto et al. (1997) Plant J. 12(2): 255-265; Kawamata et al. (1997) Plant Cell Physiol. 38(7): 792-803; Hansen et al. (1997) Mol. Gen. Genet. 254(3): 337-343; Russell et al. (1997) Transgenic Res. 6(2): 157-168; Rinehart et al. (1996) Plant Physiol. 112(3): 1331-1341; Van Camp et al. (1996) Plant Physiol. 112(2): 525-535; Canevaschi et al. (1996) Plant Physiol. 112(2): 513-524; Yamamoto et al. (1994) Plant Cell Physiol. 35(5): 773-778; Lam (1994) Results Probl. Cell Differ. 20: 181-196; Orozco et al. (1993) Plant Mol. Biol. 23(6): 1129-1138; Mutsuoka et al. (1993) Proc. Natl. Acad. Sci. USA 90: 9586-9590; y Guevara-Garcia et al. (1993) Plant J. 4(3): 495-50.

25 "Unido operativamente" significa unido como parte de la misma molécula de ácido nucleico, adecuadamente posicionada y orientada para que se inicie la transcripción a partir del promotor. El ADN operativamente unido a un promotor está "bajo la regulación de iniciación transcripcional" del promotor. En un aspecto preferido, el promotor es un promotor específico de tejido, un promotor inducible o un promotor regulado en el desarrollo.

30 Particularmente interesante en el presente contexto son los constructos de ácidos nucleicos que funcionan como vectores de plantas. Los procedimientos y vectores específicos usados previamente con gran éxito en las plantas los describe Guerineau and Mullineaux (1993) (Plant transformation and expression vectors. En: Plant Molecular Biology Labfax (Croy RRD ed) Oxford, BIOS Scientific Publishers, páginas 121-148. Los vectores adecuados pueden incluir vectores derivados de plantas víricas (véase por ejemplo EP194809).

35 Si se desea, pueden incluirse marcadores genéticos seleccionables en la construcción, tales como aquellos que confieren fenotipos seleccionables tales como resistencia a herbicidas (por ejemplo, kanamicina, higromicina, fosfotricina, clorsulfurón, metotrexato, gentamicina, espectinomycin, imidazolinonas y glifosato).

40 De acuerdo con un aspecto adicional de la invención, se proporciona una célula transgénica transformada o transfectada con una molécula o vector de ácido nucleico de acuerdo con la invención.

En una realización preferida de la invención, dicha célula es una célula vegetal.

45 En una realización preferida de la invención, dicha célula vegetal es de la familia *Papaveraceae*.

En una realización preferida de la invención, dicha célula vegetal es una célula de *Papaver somniferum*.

De acuerdo con un aspecto adicional de la presente invención, se proporciona una planta que comprende una célula vegetal de acuerdo con la invención.

50 En una realización preferida de la invención, dicha planta es de la familia *Papaveraceae*; preferiblemente *Papaver somniferum*.

55 En una realización alternativa preferida de la invención, dicha célula es una célula microbiana; preferiblemente una célula bacteriana o fúngica [por ejemplo levadura, *Saccharomyces cerevisiae*].

En una realización preferida de la invención, dicha célula está adaptada de manera que la molécula de ácido nucleico que codifica el citocromo P450 se sobreexpresa cuando se compara con una célula no transgénica de la misma especie.

De acuerdo con un aspecto adicional de la invención, se proporciona una molécula de ácido nucleico de acuerdo con la invención que comprende un casete de transcripción seleccionado de entre el grupo consistente en:

- i) un casete que está adaptado de modo que las moléculas de ácido ribonucleico sentido y antisentido se transcriben a partir de dicho casete, donde dichas moléculas de ácido nucleico sentido y antisentido están adaptadas para reconocer al menos una parte o la totalidad de su longitud para formar un ARN de interferencia pequeño [siARN] o ARN de horquilla corta [ARNhc].
- ii) un casete tal como se define en (i) anterior que está provisto de al menos dos promotores adaptados para transcribir cadenas sentido y antisentido de dicha molécula de ácido ribonucleico.
- iii) un casete tal como se define en (i) anterior que comprende una molécula de ácido nucleico, donde dicha molécula comprende una primera parte unida a una segunda parte, donde dichas primera y segunda parte son complementarias de al menos parte de su secuencia y además donde la transcripción de dicha molécula de ácido nucleico produce una molécula de ácido ribonucleico que forma una región bicatenaria por apareamiento de bases complementarias de dichas primera y segunda parte, formando así un ARNhc,
- iv) un casete de acuerdo con cualquiera de i) a iii) que comprende una molécula de ácido nucleico que es parte de un vector adaptado para la expresión en una célula vegetal.

Una técnica para eliminar específicamente la función del gen es mediante la introducción de ARN bicatenario, también denominado ARN inhibidor /de interferencia (siARN) pequeño o ARN de horquilla corta [ARNhc], en una célula que da como resultado la destrucción del ARNm complementario a la secuencia incluida en la molécula de siARN/ ARNhc. La molécula de siARN comprende dos cadenas complementarias de ARN (una cadena sentido y una cadena antisentido) unidas entre sí para formar una molécula de ARN de doble cadena. La molécula de siARN normalmente se deriva de los exones del gen que se debe extirpar. El mecanismo de la interferencia de ARN está siendo dilucidado. Muchos organismos responden a la presencia de ARN bicatenario activando una cascada que conduce a la formación de siARN. La presencia de ARN bicatenario activa un complejo de proteína que comprende RNasa III que procesa el ARN bicatenario en fragmentos más pequeños (siARNs, aproximadamente de 21-29 nucleótidos de longitud) que se convierten en parte de un complejo de ribonucleoproteína. El siARN actúa como una guía para el complejo RNasa para escindir ARNm complementario a la cadena antisentido del siARN dando como resultado la destrucción del ARNm.

En una realización preferida de la invención, dicha molécula de ácido nucleico es parte de un vector adaptado para la expresión en una célula vegetal.

De acuerdo con un aspecto adicional de la invención, se proporciona una célula vegetal transfectada con una molécula o vector de ácido nucleico de acuerdo con la invención, donde dicha célula presenta una expresión reducida de dicho polipéptido citocromo P450.

De acuerdo con un aspecto de la invención, se proporciona un proceso para la modificación de un alcaloide opiáceo que comprende:

- i) proporcionar una célula vegetal transgénica de acuerdo con la invención;
  - ii) cultivar dicha célula vegetal para producir una planta transgénica; y opcionalmente
  - iii) cosechar dicha planta transgénica, o parte de la misma.
- En un procedimiento preferido de la invención, dicho material vegetal cosechado es paja seca y se extrae dicho alcaloide opiáceo.

De acuerdo con un aspecto alternativo de la invención, se proporciona un proceso para la modificación de un alcaloide opiáceo que comprende:

- i) proporcionar una célula microbiana transgénica de acuerdo con la invención que expresa un citocromo P450 de acuerdo con la invención en cultivo con al menos un alcaloide opiáceo;
- ii) cultivar la célula microbiana en unas condiciones que modifican uno o más alcaloides opiáceos; y opcionalmente
- iii) aislar dicho alcaloide opiáceo de la célula microbiana o cultivo celular.

En un procedimiento preferido de la invención, dicha célula microbiana es una célula bacteriana o célula fúngica / de levadura.

Si se usan células microbianas como organismos en el proceso de acuerdo con la invención, crecen o se cultivan de la manera en la que el experto en la materia está familiarizado, dependiendo del organismo anfitrión. Como regla

general, los microorganismos se cultivan en un medio líquido que comprende una fuente de carbono, generalmente en forma de azúcares, una fuente de nitrógeno, generalmente en forma de fuentes de nitrógeno orgánico como extracto de levadura o sales como sulfato de amonio, oligoelementos como sales de hierro, manganeso y magnesio y, en su caso, vitaminas, a temperaturas de entre 0 ° C y 100 ° C, preferiblemente entre 10 ° C y 60 ° C, mientras gasean en oxígeno.

El pH del medio líquido puede mantenerse constante, es decir, regularse durante el período de cultivo, o no. Los cultivos se pueden cultivar de forma discontinua, semidiscontinua o continua. Los nutrientes pueden proporcionarse al comienzo de la fermentación o alimentarse en forma semicontinua o continua. Los alcaloides opiáceos metilados producidos se pueden aislar de los organismos como se describió anteriormente mediante procesos conocidos por los expertos en la técnica, por ejemplo mediante extracción, destilación, cristalización, si es apropiado, precipitación con sal, y / o cromatografía. Con este fin, los organismos pueden ser interrumpidos ventajosamente de antemano. En este proceso, el valor del pH se mantiene ventajosamente entre pH 4 y 12, preferiblemente entre pH 6 y 9, especialmente preferiblemente entre pH 7 y 8.

El medio de cultivo que se utilizará debe cumplir adecuadamente los requisitos de las cepas en cuestión. Las descripciones de los medios de cultivo para diversos microorganismos se pueden encontrar en el libro de texto "Manual of Methods for General Bacteriology" of the American Society for Bacteriology (Washington D.C., USA, 1981).

Como se ha descrito anteriormente, estos medios que se pueden utilizar de acuerdo con la invención, normalmente comprenden una o más fuentes de carbono, fuentes de nitrógeno, sales inorgánicas, vitaminas y / u oligoelementos.

Las fuentes de carbono preferidas son azúcares, tales como mono-, di- o polisacáridos. Ejemplos de fuentes de carbono son glucosa, fructosa, manosa, galactosa, ribosa, sorbosa, ribulosa, lactosa, maltosa, sacarosa, rafinosa, almidón o celulosa. Los azúcares también se pueden agregar a los medios a través de compuestos complejos como la melaza u otros subproductos de la refinación de azúcar. La adición de mezclas de una variedad de fuentes de carbono también puede ser ventajosa. Otras posibles fuentes de carbono son aceites y grasas como, por ejemplo, aceite de soja, aceite de girasol, aceite de cacahuete y / o grasa de coco, ácidos grasos como, por ejemplo, ácido palmítico, ácido esteárico y / o ácido linoleico, alcoholes y / o polialcoholes tales como, por ejemplo, glicerol, metanol y / o etanol, y / o ácidos orgánicos tales como, por ejemplo, ácido acético y / o ácido láctico.

Las fuentes de nitrógeno son habitualmente compuestos de nitrógeno orgánicos o inorgánicos o materiales que comprenden estos compuestos. Ejemplos de fuentes de nitrógeno comprenden amoníaco en forma líquida o gaseosa o sales de amonio como sulfato de amonio, cloruro de amonio, fosfato de amonio, carbonato de amonio o nitrato de amonio, nitratos, urea, aminoácidos o fuentes complejas de nitrógeno como licor de maíz, harina de soja, proteína desoja, extracto de levadura, extracto de carne y otros. Las fuentes de nitrógeno se pueden usar individualmente o como una mezcla.

Los compuestos de sales inorgánicas que pueden estar presentes en los medios comprenden las sales de cloruro, fósforo y sulfato de calcio, magnesio, sodio, cobalto, molibdeno, potasio, manganeso, zinc, cobre y hierro.

Los compuestos que contienen azufre inorgánico tales como, por ejemplo, sulfatos, sulfitos, ditionitos, tetratiónatos, tiosulfatos, sulfuros, o compuestos de azufre orgánicos tales como mercaptanos y tioles pueden usarse como fuentes de azufre para la producción de productos químicos finos que contienen azufre, en particular de metionina.

Se puede usar ácido fosfórico, dihidrogenofosfato de potasio o hidrogenofosfato dipotásico o las correspondientes sales que contienen sodio como fuentes de fósforo.

Se pueden añadir agentes quelantes al medio para mantener los iones metálicos en solución. Los agentes quelantes particularmente adecuados comprenden dihidroxifenoles tales como catecol o protocatecuato y ácidos orgánicos tales como ácido cítrico.

Los medios de fermentación usados de acuerdo con la invención para cultivar microorganismos generalmente también comprenden otros factores de crecimiento tales como vitaminas o promotores del crecimiento, que incluyen, por ejemplo, biotina, riboflavina, tiamina, ácido fólico, ácido nicotínico, pantotenato y piridoxina. Los factores de crecimiento y las sales se derivan a menudo de componentes de medios complejos tales como extracto de levadura, melazas, licor de salsa de maíz y similares. Además, es posible agregar precursores adecuados al medio de cultivo. La composición exacta de los compuestos de los medios depende en gran medida del experimento particular y se decide individualmente para cada caso específico. La información sobre la optimización de los medios se puede

encontrar en el libro de texto "Applied Microbiol. Physiology, A Practical Approach" (Editors P.M. Rhodes, P.F. Stanbury, IRL Press (1997) páginas 53-73, ISBN 0 19 963577 3). Los medios de crecimiento también pueden obtenerse de proveedores comerciales, por ejemplo, Standard 1 (Merck) o BHI (infusión de cerebro y corazón, DIFCO) y similares.

5

Todos los componentes del medio se esterilizan, ya sea por calor (20 min a 1,5 bar y 121°C) o por esterilización por filtración. Los componentes se pueden esterilizar juntos o, si es necesario, por separado. Todos los componentes del medio pueden estar presentes al comienzo del cultivo o pueden agregarse de forma continua o por lotes, según se desee.

10

La temperatura de cultivo está normalmente entre 15°C y 45°C, preferiblemente entre 25°C y 40°C, y puede mantenerse constante o puede alterarse durante el experimento. El pH del medio debe estar en el intervalo de 5 a 8.5, preferiblemente alrededor de 7.0. El pH para el cultivo puede controlarse durante el cultivo mediante la adición de compuestos básicos como hidróxido de sodio, hidróxido de potasio, amoníaco y amoníaco acuoso o compuestos ácidos como ácido fosfórico o ácido sulfúrico. La formación de espuma puede controlarse empleando antiespumantes tales como, por ejemplo, ésteres de poliglicol de ácido graso. Para mantener la estabilidad de los plásmidos, es posible añadir al medio sustancias adecuadas que tengan un efecto selectivo, por ejemplo, antibióticos. Las condiciones aeróbicas se mantienen introduciendo oxígeno o mezclas de gases que contienen oxígeno tales como, por ejemplo, aire ambiente en el cultivo. La temperatura del cultivo es normalmente de 20°C a 45°C y preferiblemente de 25°C a 40°C. El cultivo continúa hasta que la formación del producto deseado es máxima. Este objetivo normalmente se logra dentro de 10 a 160 horas.

15

20

El caldo de fermentación puede procesarse posteriormente. La biomasa puede, según los requisitos, eliminarse total o parcialmente del caldo de fermentación por métodos de separación tales como, por ejemplo, centrifugación, filtración, decantación o una combinación de estos métodos o dejarse completamente en dicho caldo. Es ventajoso procesar la biomasa después de su separación.

25

Sin embargo, el caldo de fermentación también puede espesarse o concentrarse sin separar las células, usando métodos conocidos tales como, por ejemplo, con la ayuda de un evaporador rotatorio, un evaporador de película delgada, un evaporador de película descendente, mediante ósmosis inversa o por nanofiltración. Finalmente, este caldo de fermentación concentrado se puede procesar para obtener los alcaloides opiáceos presentes en el mismo.

30

De acuerdo con un aspecto adicional de la invención, se proporciona el uso de un gen codificado por una molécula de ácido nucleico representada por la secuencia de ácido nucleico en la Figura 1d o una molécula de ácido nucleico que se hibrida en condiciones de hibridación rigurosas con la secuencia de nucleótidos en la Figura 1d y codifica un polipéptido con actividad del citocromo P450 como medio para identificar la presencia o ausencia de un gen que codifica dicho citocromo P450 en una planta *Papaveraceae*.

35

De acuerdo con un aspecto adicional de la invención, se proporciona un método para determinar la presencia o ausencia de un gen de acuerdo con la invención en una variedad *Papaveraceae* que comprende:

40

- i) obtener una muestra de una planta *Papaveraceae*;
- ii) extraer el ADN genómico de la planta; y
- iii) analizar el ADN genómico para la presencia de un gen que comprende o consiste en una secuencia de nucleótidos tal como se representa en la Figura 3c.

45

Los métodos para analizar el ADN genómico son conocidos en la técnica. Por ejemplo, los métodos de reacción en cadena de la polimerasa usando cebadores de oligonucleótidos de secuencia específica para amplificar regiones específicas del gen de acuerdo con la invención. Los análisis de extracción, aislamiento y restricción usando endonucleasas con restricción de una secuencia específica seguido de la separación y transferencia tipo Southern blot para analizar la estructura genómica han sido establecidos durante más de treinta años. Los análisis se pueden dirigir a la estructura del intrón o exón o las regiones ascendente o descendente del gen, por ejemplo las regiones promotoras.

50

De acuerdo con un aspecto adicional de la invención, se proporciona el uso de un gen codificado por una molécula de ácido nucleico representada por la secuencia de ácido nucleico en la Figura 3a, 3b ó 3c, o una molécula de ácido nucleico que se hibrida en condiciones de hibridación rigurosas con la secuencia de nucleótidos en la Figura 3a, 3b ó 3c y codifica un polipéptido con actividad del citocromo P450 como medio para identificar un locus donde dicho locus está asociado con la expresión o actividad de dicho citocromo P450. La mutagénesis como un medio para inducir cambios fenotípicos en organismos es bien conocida en la técnica e incluye, pero no se limita a, el uso de

60



agentes mutagénicos tales como mutágenos químicos [por ejemplo. análogos de bases, agentes desaminantes, agentes de intercalación de ADN, agentes alquilantes, transposones, bromo, azida sódica] y mutágenos físicos [por ejemplo radiación ionizante, exposición al psoraleno combinada con irradiación UV].

5 De acuerdo con un aspecto adicional de la invención, se proporciona una planta que presenta una expresión alterada de un polipéptido de citocromo P450 codificado por una molécula de ácido nucleico de acuerdo con la invención que comprende las etapas de:

- i) mutagénesis de semillas de tipo silvestre de una planta que expresa dicho polipéptido de citocromo P450;
- 10 ii) cultivo de la semilla en i) para producir las primeras y posteriores generaciones de plantas;
- iii) obtener la semilla de la planta de la primera generación y las siguientes generaciones de plantas;
- iv) determinar si la semilla de dicha primera y siguientes generaciones de plantas presenta una secuencia de nucleótidos alterada y/o una expresión alterada de dicho polipéptido de citocromo P450.

15 De acuerdo con un aspecto de la invención, se proporciona un método para el análisis de una planta que comprende

- i) obtener una muestra de la planta de acuerdo con la invención y analizar la secuencia de ácido nucleico de una molécula de ácido nucleico seleccionada del grupo que consiste en:
  - a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se representa en la Figura 3c;
  - 20 b) una molécula de ácido nucleico que se hibrida con la molécula de ácido nucleico en a) bajo condiciones de hibridación astringentes y que codifica un polipéptido con actividad polipéptido citocromo P450; y opcionalmente
- ii) comparar la secuencia de nucleótidos de la molécula de ácido nucleico en dicha muestra con una secuencia de un
  - 25 ácido nucleico de la planta original de tipo silvestre.
- iii) extraer el ácido nucleico de dichas plantas mutadas;
- iv) amplificación de una parte de dicha molécula de ácido nucleico mediante una reacción en cadena de la
  - 30 polimerasa;
- v) formar una preparación que comprende el ácido nucleico amplificado y el ácido nucleico extraído de la semilla de tipo silvestre para formar ácido nucleico heterodúplex;
- 35 vi) incubar dicha preparación con una nucleasa monocatenaria que corta en una región de ácido nucleico heterodúplex para identificar la falta de coincidencia en dicho heterodúplex; y
- vii) determinar el sitio del desapareamiento en dicho heterodúplex de ácido nucleico.

40 En un método preferido de la invención, dicha variedad de planta presenta una expresión y/o actividad del polipéptido de citocromo P450 aumentadas.

De acuerdo con un aspecto adicional de la invención, se proporciona una planta obtenida mediante el método de acuerdo con la invención.

45 De acuerdo con un aspecto de la invención, se proporciona una planta, donde dicha planta comprende un vector viral que incluye todo o parte de un gen que comprende una molécula de ácido nucleico de acuerdo con la invención.

50 En una realización preferida de la invención, dicho gen está codificado por una molécula de ácido nucleico que comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:

- i) una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se representa en la Figura 1c ó 1d;
- 55 ii) una molécula de ácido nucleico que comprende una secuencia de nucleótidos que se hibrida en condiciones de hibridación rigurosas con una molécula de ácido nucleico en (i) y que codifica un polipéptido de citocromo P450;
- iii) una molécula de ácido nucleico que codifica una variante del polipéptido que varía en un polipéptido que comprende la secuencia de aminoácido como se representa en la Figura 4c.

60 En una realización preferida de la invención, dicha molécula de ácido nucleico comprende o consiste en una

secuencia de nucleótidos como se representa en la Figura 1c

En una realización preferida de la invención, dicha molécula de ácido nucleico comprende o consiste en una secuencia de nucleótidos como se representa en la Figura 1d.

5

En una realización preferida de la invención, dicha molécula de ácido nucleico comprende o consiste en una secuencia de nucleótidos como se representa en la Figura 13.

De acuerdo con un aspecto adicional de la invención, se proporciona un vector viral que comprende la totalidad o parte de una molécula de ácido nucleico de acuerdo con la invención.

10

De acuerdo con un aspecto de la invención, se proporciona el uso de un vector viral de acuerdo con la invención en el silenciamiento génico inducido viral en una planta.

15 En una realización preferida de la invención, dicha planta es de la familia *Papaveraceae*.

El silenciamiento génico inducido por virus [VIGS] es conocido en la técnica y explota un mecanismo de defensa antiviral mediado por ARN. Las plantas que están infectadas con un virus no modificado inducen un mecanismo que se dirige específicamente al genoma viral. Sin embargo, los vectores virales que se modifican genéticamente para incluir moléculas de ácido nucleico derivadas de genes de plantas hospedadoras también inducen la inhibición específica de la expresión del vector viral y adicionalmente el ARNm diana del huésped. Esto permite silenciar genes genéticamente específicos sin modificación genética del genoma de la planta y es esencialmente una modificación no transgénica.

20

A lo largo de la descripción y las reivindicaciones de esta memoria descriptiva, las palabras "comprender" y "contener" y variaciones de las palabras, por ejemplo "que comprende" y "comprende", significan "que incluyen pero no se limitan a", y no están destinadas (y no) excluyen otros restos, aditivos, componentes, enteros o etapas.

25

A lo largo de la descripción y las reivindicaciones de esta memoria descriptiva, el singular abarca el plural a menos que el contexto requiera lo contrario. En particular, cuando se usa el artículo indefinido, debe entenderse que la especificación contempla tanto la pluralidad como la singularidad, a menos que el contexto requiera lo contrario.

30

Las características, enteros, particularidades, compuestos, restos químicos o grupos descritos junto con un aspecto particular, realización o ejemplo de la invención deben entenderse aplicables a cualquier otro aspecto, realización o ejemplo descrito en la presente, a menos que sea incompatible con los mismos.

35

Ahora a continuación, se describirá una realización de la invención solo a modo de ejemplo y con referencia a las siguientes figuras:

40 La Figura 1a es una secuencia de nucleótidos de un ADNc que codifica para *PSCYP1*, la Figure 1b es una secuencia de nucleótidos de un ADNc que codifica para *PSCYP2*, la Figure 1c es una secuencia de nucleótidos de un ADNc que codifica para *PSCYP3*; la Figure 1d es una secuencia de nucleótidos de otra realización de un ADNc que codifica para *PSCYP3*;

La Figura 2 ilustra la frecuencia de EST del gen *PSCYP1* en bibliotecas EST derivadas de la secuenciación 454 de tejidos de tallos y cápsulas de cultivares de GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1. Las 16 librerías EST se generaron por pirosecuenciación usando bibliotecas de ADNc preparadas a partir de tallos (S) y cápsulas (C) de dos etapas de desarrollo 'cosecha-temprana' EH, 1-3 días después de la caída de los pétalos) y 'cosecha-tardía' (LH, 4-6 días después de la caída de los pétalos) de cada uno de los cuatro cultivares de *P. somniferum*; Figura 2b ilustra la frecuencia de EST del gen *PSCYP2*; Figura 2c ilustra la frecuencia de EST del gen *PSCYP3*;

50

La Figure 3a es la secuencia de nucleótidos del gen que codifica para *PSCYP1*, la Figure 3b es la secuencia de nucleótidos del gen que codifica para *PSCYP2*; la Figure 3c es la secuencia de nucleótidos del gen que codifica para *PSCYP3*;

La Figura 4a es la secuencia de aminoácidos deducida de *PSCYP1*; la Figura 4b es la secuencia de aminoácidos deducida de *PSCYP2*; la Figura 4c es la secuencia de aminoácidos deducida de *PSCYP3*; la Figura 4d es la secuencia de aminoácidos deducida de *PSCYP3*;

55

La Figura 5 ilustra que la secuencia del gen *PSCYP1* sólo está presente en el cultivar de GSK NOSCAPINA CVS1 y está ausente en los cultivares de GSK MORFINA CVS1, GSK MORFINA CVS2 y GSK TEBAINA CVS1;

La Figura 6 ilustra que la secuencia del gen *PSCYP2* sólo está presente en el cultivar de GSK NOSCAPINA CVS1 y está ausente en los cultivares de GSK MORFINA CVS1, GSK MORFINA CVS2 y GSK TEBAINA CVS1;

60

La Figura 7 ilustra que la secuencia del gen *PSCYP3* sólo está presente en el cultivar de GSK NOSCAPINA CVS1 y está ausente en los cultivares de GSK MORFINA CVS1, GSK MORFINA CVS2 y GSK TEBAINA CVS1;

La Figura 8a es una representación tabular de la segregación del gen *PSCYP1* en un mapeo de población F2 derivado de un cruce parental de cultivares de GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1 junto con cosegregación de *PSCYP1* y la acumulación de noscapina en plantas F2 individuales, la Figura 8b es la representación equivalente de la segregación del gen *PSCYP2*, la Figura 8c es la representación equivalente de la segregación del gen *PSCYP3*, el ensayo de genotipado de *PSCYP3* falló en 16 muestras (tal como indica el fracaso a la hora amplificar el control positivo interno), estas muestras se excluyeron del análisis de cosegregación de *PSCYP3*;

10 La Figura 9 ilustra un cromatograma UPLC típico para una solución estándar;

La Figura 10 ilustra un cromatograma UPLC típico para una noscapina que contiene una variedad de adormidera;

La Figura 11 es una parte con una longitud de 622 bases de la secuencia del gen del fitoeno desaturasa amplificada a partir del ADNc de GSK NOSCAPINA CVS1. La secuencia del tramo de 129 bases usado para silenciar el gen de fitoeno desaturasa está subrayada;

15 La Figura 12 es la parte de la secuencia del ADNc usada para silenciar *PSCYP2*; y

La Figura 13 es la parte de la secuencia del ADNc usada para silenciar *PSCYP3*;

La Figura 14 muestra el área del pico normalizado de la tetrahidrocolumbamina putativa en los cromatogramas UPLC obtenidos a partir de cápsulas de látex y maduras de plantas que mostraban el fenotipo de blanqueamiento tras la infección con los constructos de silenciamiento pTRV2-PDS-PSCYP2, pTRV2-PDS-PSCYP3 o pTRV2-PDS, respectivamente. También se muestra el área del pico de la tetrahidrocolumbamina putativa obtenida a partir de plantas no infectadas;

La Figura 15 muestra el área del pico normalizado de una secoberbina alcaloide putativa (en los cromatogramas UPLC obtenidos a partir de cápsulas de látex y maduras de plantas que mostraban el fenotipo de blanqueamiento tras la infección con los constructos de silenciamiento pTRV2-PDS-PSCYP2, pTRV2-PDS-PSCYP3 o pTRV2-PDS, respectivamente. También se muestra el área del pico de la secoberbina putativa obtenida a partir de plantas no infectadas. La masa, la fórmula molecular y el patrón de fragmentación del compuesto es consistente con el demetoxidroximacrantaldehído o el demetoximacrantoridina; y

La Figura 16 muestra el área del pico normalizado de otra secoberbina alcaloide putativa en los cromatogramas UPLC obtenidos a partir de cápsulas de látex y maduras de plantas que mostraban el fenotipo de blanqueamiento tras la infección con los constructos de silenciamiento pTRV2-PDS-PSCYP2, pTRV2-PDS-PSCYP3 o pTRV2-PDS, respectivamente. También se muestra el área del pico de la secoberbina putativa obtenida a partir de plantas no infectadas. La masa, la fórmula molecular y el patrón de fragmentación del compuesto es consistente con el demetoxinarcotinadiol y el narcotolinol.

## 35 Materiales y procedimientos

### Generación de librerías EST

#### a) Aislamiento de ARN y síntesis de ADNc

40 El material se recogió de tallos y cápsulas en dos etapas de desarrollo de cuatro cultivares de adormidera. El ARN se preparó individualmente a partir de cinco plantas por cultivar, etapa de desarrollo y órgano. El material cosechado se molió en nitrógeno líquido usando un mortero. El ARN se aisló de preparaciones de tallo o cápsula molida usando un método basado en CTAB (bromuro de hexadeciltrimetilamonio) como está descrito en Chang et al. (1993) Plant  
45 Molecular Rep. 11: 113-116. Con ligeras modificaciones (tres extracciones con cloroformo:alcohol isoamílico, precipitación del ARB con cloruro de litio a -20°C durante la noche). El ARN se cuantificó espectrofotométricamente antes de mezclar cantidades iguales de ARN de cinco plantas por cultivar, etapa y órgano. Las muestras se sometieron a una etapa de purificación final usando un Neasy Plus MicroKit (Qiagen, Crawley, UK) para eliminar cualquier ADN genómico restante de las preparaciones. El ARN se eluyó normalmente en 30-100 µl de agua. El  
50 ADNc se preparó usando el kit de construcción de la Biblioteca SMART cDNA (Clontech, Saint-Germain-en-Laye, Francia) de acuerdo con las instrucciones del fabricante, pero utilizando la transcriptasa inversa SuperScript II (Invitrogen, Paisley, Reino Unido) para la síntesis de la primera cadena. El cebador CDSIII PCR se modificó para: 5' ATT CTA GAT CCR ACA TGT TTT TTT TTT TTT TTT TVN 3' donde R = A or G, V = A, C o G; N = A/T or C/G. El ADNc se digirió con Mmel (New England Biolabs Inc., Hitchin, UK) seguido de una purificación final usando un kit  
55 de purificación QIAquick PCR (Qiagen, Crawley, Reino Unido).

#### b) Pirosecuenciación del ADNc

La plataforma de secuenciación de Roche 454 GS-FLX (Branford, CT, EE.UU.) se usó para realizar la  
60 pirosecuenciación de muestras de ADNc preparadas a partir de los siguientes materiales para cada uno de los

cuatro cultivares de *P. somniferum* - GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAÍNA CVS1.

1. Tallo, 1-3 días después de la caída del pétalo (cosecha temprana)
- 5 2. Tallo, 4-6 días después de la caída del pétalo (cosecha tardía)
3. Cápsula, 1-3 días después de la caída del pétalo (cosecha temprana)
4. Tallo, 4-6 días después de la caída del pétalo (cosecha tardía)

### c) Análisis de secuencia bruta, ensamblaje de secuencia contigua y anotación

- 10 Los conjuntos de datos de secuencia en bruto se derivaron de secuenciación etiquetada paralela en la plataforma de secuenciación 454 (Meyer et al. (2008) Nature Protocols 3: 267-278). Las secuencias de cebador y etiqueta se eliminaron primero de todas las lecturas de secuencia individuales. El ensamblaje de secuencia contigua solo se realizó en secuencias de más de 40 nucleótidos y que contienen menos del 3% de residuos desconocidos (N). Esas
- 15 secuencias EST de alta calidad se ensamblaron en secuencias contiguas únicas con el programa de ensamblaje de secuencias CAP3 (Huang and Madan (1999) Genome Research 9: 868-877), y los contigs resultantes se anotaron localmente usando el programa BLAST2 (Altschul et al. (1997) Nucleic Acids Res. 25: 3389-3402) contra la base de datos de péptidos no redundante descargada del NCBI.

### 20 d) Determinación del perfil de expresión de los genes del citocromo P450

El número de EST asociadas con las respectivas secuencias consenso del gen del citocromo P450 se contaron en cada una de las 16 bibliotecas de EST. Los valores obtenidos se normalizaron sobre la base del número total de EST obtenido por biblioteca.

- 25 **Amplificación y secuenciación de los genes del citocromo P450 del ADN genómico de GSK NOSCAPINE CVS1.**

#### a) Preparación del ADN genómico

- 30 Preparación del ADN: Se recolectaron muestras de hojas (30-50mg) para la extracción del ADN a partir de plantas GSK MORFINA CVS1, GSK MORFINA CVS2 GSK NOSCAPINA CVS1, GSK TEBAÍNA CVS1 cultivadas en invernadero. El ADN se extrajo usando Qiagen BioSprint 96. El ADN extraído se cuantificó usando Hoescht 33258 y se normalizó a 10 ng/ul.

#### 35 b) Amplificación y secuenciación de los genes del citocromo P450 a partir del ADN de GSK

- Los cebadores de NOSCAPINA CVS1 y las combinaciones de cebadores usadas para la amplificación de los respectivos genes del citocromo P450 del ADN genómico extraído se muestran en la Tabla 1. Los cebadores se diseñaron basándose en los respectivos contigs del citocromo P450 ensamblados a partir de EST únicas para el cultivar GSK NOSCAPINA CVS1. Los contigs de *PSCYP1* y *PSCYP2* contenían el marco de lectura abierto tanto de las regiones 5' como las 3' no traducidas. El *PSCYP3* se representó mediante dos contigs que cubrían los extremos 5' y 3' del marco de lectura abierto con 200 bases desde el centro del marco de lectura abierto perdido. Este tramo perdido de secuencia codificante se amplificó y se confirmó por amplificación y secuenciación a partir del ADNc
- 40 (preparado tal como se ha descrito anteriormente) además del ADN genómico para determinar la posición exacta del intrón 1 (Figura 3c). La amplificación se realizó en conjuntos de ADN que comprenden el ADN de al menos cuatro individuos y las combinaciones de cebadores que se muestran en la Tabla 2. Las condiciones para la PCR fueron las siguientes:

Mezcla de reacción:

50

Tampón 5xHF (Finnzymes)	5 µl
dNTPs (cada uno de 20 mM)	0,25 µl
Cebador Fwd (10 µM)	2,5 µl
Cebador Rev (10 µM)	2,5 µl
ADN (10 ng/µl)	5 µl
Phusion Hot Start (Finnzymes)	0,25 µl
H <sub>2</sub> O	9,5 µl

Volumen de reacción: 25 µl

El Phusion Hot Start de Finnzymes se adquirió a través de New England Biolabs, (Bishops Stortford, Reino Unido).

#### **Programa PCR:**

	desnaturalización inicial	98°C	1 min
30 ciclos de:	desnaturalización	98°C	30 segundos
	temperatura de hibridación	Tablas 2 y 3	30 segundos
	extensión	72°C	40 segundos
	extensión final	72°C	10 min
	incubación	4°C	almacenamiento

5

El extremo 5' y parte de la región promotora de *PSCYP3* se amplificaron a partir de ADN genómico a través de una configuración de PCR de largo alcance usando los cebadores PSCYP1\_LA\_R1 y PSCYP3\_LA\_R1:

Mezcla de reacción del PCR de largo alcance:

10

Tampón 5xLongAmp (New England Biolabs)	10 µl
dNTPs (cada uno de 10 mM)	1,5 µl
Cebador Fwd (10 µM)	2 µl
Cebador Rev (10 µM)	2 µl
ADNg (100 ng/µl)	2 µl
Taq LongAmp (New England Biolabs)	2 µl
H <sub>2</sub> O	30,5 µl

Volumen de reacción: 50 µl

#### **Programa PCR de largo alcance:**

15

	desnaturalización inicial	94°C	30 segundos
30 ciclos de:	desnaturalización	94°C	30 segundos
	hibridación y extensión	65°C	13,5min
	extensión final	65°C	10 min
	incubación	4°C	almacenamiento

Los productos resultantes de las diversas PCR se purificaron usando el kit de purificación Agencourt AMPure (Beckman Coulter LTD, Bromley, Reino Unido). 30-50 ng de los respectivos productos de PCR purificados se sometieron a una secuenciación de Sanger usando los cebadores mostrados en la Tabla 2 como cebadores de secuenciación. Debido a que la combinación de cebadores PSCYP1\_F4/R7 dio como resultado una amplificación de un producto más pequeño inespecífico además del amplicón esperado (ver también la Figura 4d), el último se escindió y se purificó del gel usando el kit de extracción QIAEX II Gel (Qiagen, Hilden, Alemania) antes de la secuenciación.

25 Las secuencias de aminoácido de citocromo P450s respectivo previstas a partir de la secuencia de Sanger confirmaron secuencias de marco de lectura abierto cuando se compararon con las secuencias de proteínas depositadas en la base de datos de proteínas no redundantes usando el programa Standard Protein BLAST (blastp).

#### **c) Análisis del ADN genómico de GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1 para la presencia de genes del citocromo P450**

Para investigar si los genes del citocromo P450 estaban presentes en los cuatro cultivares, se realizó una amplificación del ADN genómico (grupos de cuatro individuos por cultivar) de GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1 en una serie de fragmentos solapados usando la combinaciones de cebadores que se muestran en la Tabla 3. Para obtener las secuencias genómicas de longitud completa de GSK NOSCAPINA CVS1, se utilizaron exactamente las mismas condiciones para la PCR descritas anteriormente. =. 5 µl de cada reacción de PCR se resolvieron en agarosa al 1% junto con los estándares de tamaño adecuados.

#### **40 Generación de un mapeo de la población, extracción y análisis del ADN genómico de material de las hojas más extracción y análisis de alcaloides de la paja de la adormidera**

**a) extracción del ADN de plantas F2**

Se recogieron 40-50 mg de tejido foliar, por duplicado, de todas las plantas de adormidera dentro del mapeo de la población y de las plantas parentales de GSK NOSCAPINA CVS1 X GSK TEABÍNA CVS1 F2) en la etapa de crecimiento de 'pequeña roseta' (~ 10 hojas presentes en cada planta).

El tejido foliar (40-50 mg de peso húmedo) se recogió en tubos de muestras de 1,2 ml en un formato de 8x12 (Número de pieza 1760-00, Scientific Specialties Inc, 130 Thurman St, Lodi, CA 95240 USA), cerrados con tapón de tira (Número de pieza 1702-00, Scientific Specialties Inc) y se enviaron a la sede de AGRF (Australian Genome Research Facility) en Adelaida en paquetes con bloques de hielo seco Techni por mensajería durante la noche.

A su recepción, se extrajeron los tapones de tira y se añadió un cordón de carburo de tungsteno de 3 mm a cada tubo (Número de pieza 69997, Qiagen GmbH, Hilden, Germany). Las muestras se colocaron a -80 ° C (Modelo de congelador; Sanyo MDF-U73V) durante un mínimo de dos horas antes de la liofilización durante 18 horas (Christ Model Alpha 2-4 LSC).

Después de la liofilización, los tubos se sellaron con unos tapones de tira nuevos (como los anteriores), y las muestras se pulverizaron mediante molienda de perlas (Modelo «TissueLyser», número de pieza 85300, Qiagen) a 3000 rpm durante 2 ciclos x 60 segundos separados por inversión de placa. La extracción de ADN se realizó usando el sistema Nucleospin Plant II» (Macherey-Nagel, GmbH & Co. KG Neumann-Neander-Straße 6-8, 52355 Düren, Germany).

La lisis de las células se realizó usando el tampón Set PL2 / 3 suministrado. Para la extracción centrífuga se siguió el protocolo del fabricante (modelo de centrífuga 4-K 15; Sigma Laborzentrifugen GmbH, 37520 Osterode am Harz, Alemania).

El ADN recuperado (12/96 muestras, una muestra por columna de plancha) se comprobó para la calidad y cantidad mediante espectroscopia ultravioleta (Modelo Nanodrop-8000; NanoDrop products, 3411 Silverstone Rd, Bancroft Building; Wilmington, DE 19810, USA) a 230, 260 y 280 nM.

**b) Genotipado de muestras de ADN F2 para la presencia o ausencia de genes del citocromo P450**

Las muestras de ADN de un total de 275 plantas F2 fueron genotipadas para la presencia o ausencia de *PSCYP1*, *PSCYP2* y *PSCYP3*, respectivamente, por amplificación de un fragmento corto de cada uno de los genes. Para poder marcar con fluorescencia los fragmentos de PCR resultantes, los cebadores directos llevaban una etiqueta-VIC (Applied Biosystems, UK) en sus extremos 5'-prima. Los análisis de los fragmentos se llevaron cabo con 96-electroforesis capilar en el analizador 3730xl DNA (Applied Biosystems, UK) de acuerdo con las instrucciones del fabricante. Además de los respectivos fragmentos de citocromo P450, se amplificó un control positivo interno en cada ensayo por PCR para distinguir la falta de amplificación debida a la ausencia de los genes del citocromo P450 en las muestras de ADN de la falta de amplificación debida a fallos en el ensayo de PCR. Las muestras donde el ensayo por PCR fracasó fueron excluidas de los análisis de co-segregación de los genes con el rasgo noscapina.

Los siguientes cebadores fueron usados (las secuencias de los cebadores se muestran en la Tabla 1; los cebadores directos estaban marcados en el extremo 5' con VIC):

*PSCYP1*: VIC-PSCYP1\_F3/ PSCYP1\_R2; tamaño del fragmento amplificado: 166 bp

*PSCYP2*: VIC-PSCYP2\_F2/ PSCYP2\_R1; tamaño del fragmento amplificado: 226 bp

*PSCYP3*: VIC-PSCYP3\_F3/ PSCYP3\_R1; tamaño del fragmento amplificado: 638 bp

El fragmento *PSCYP1* se amplificó con las siguientes condiciones de la PCR:

Mezcla de reacción:

Tampón 5xGoTaq (Promega)	2 µl
dNTPs (mezcla de 2,5 mM)	0,5 µl
MgCl <sub>2</sub> (25 mM)	0,6 µl
Cebador directo (10 µM)	0,5 µl
Cebador inverso (10 µM)	0,5 µl
ADNg (5 ng/µl)	2 µl
GoTaq (Promega)	0,2 µl
H <sub>2</sub> O	3,7 µl

Volumen de reacción: 10 µl

##### 5 Programa PCR:

	desnaturalización inicial	94°C	1 min
30 ciclos de:	desnaturalización	94°C	30 segundos
	temperatura de hibridación	62°C	30 segundos
	extensión	72°C	20-30 segundos
	extensión final	72°C	5 min
	incubación	4°C	almacenamiento

Los fragmentos *PSCYP2* y *PSCYP3* se amplificaron con las siguientes condiciones de la PCR:

Mezcla de reacción:

10

Mezcla de PCR 5xTipo-it multiplex (Qiagen)	5 µl
Cebador directo (10 µM)	0,5 µl
Cebador inverso (10 µM)	0,5 µl
ADNg (5 ng/µl)	2 µl
dH <sub>2</sub> O	2 µl

Volumen de reacción: 10 µl

##### Programa PCR:

15

	desnaturalización inicial	95°C	15 min
30 ciclos de:	desnaturalización	95°C	15 segundos
	temperatura de hibridación	60°C	30 segundos
	extensión	72°C	30 segundos
	extensión final	72°C	5 min
	incubación	4°C	almacenamiento

##### c) Análisis de la paja de adormidera

20 Se recogieron las cápsulas de amapola a mano de la población de mapeo una vez que las cápsulas se habían secado hasta aproximadamente 10% de humedad en la planta. La semilla se separó de la cápsula manualmente y después el material de paja (paja de la adormidera) se envió a las instalaciones de extracción de GSK en Port Fairy, Australia.

25 Las muestras de paja de adormidera se molieron en un molino de bolas Retsch Modelo MM04, hasta obtener un polvo fino. Después, se pesaron dos gramos de muestras de paja de adormidera molida con precisión ( $2 \pm 0,003$ g) y se extrajeron en 50 ml de una solución de ácido acético al 10%. La suspensión de extracción se agitó en un agitador orbital a 200 rpm durante un mínimo de 10 minutos y después se filtró para proporcionar un filtrado claro. El filtrado final se pasó a través de un filtro de 0,22 µm antes del análisis.

30 Todas las soluciones se analizaron usando un sistema Waters Acquity UPLC equipado con una columna Waters Acquity BEH C18, 2,1mm x 100mm con un empaquetamiento de 1,7 micras. La fase móvil usó un perfil de gradiente con eluyente A que consistía en ácido trifluoroacético al 0,01% en agua desionizada y eluyente B que consistía en acetonitrilo al 100%. Las condiciones del gradiente de la fase móvil usadas se indican en la Tabla 2, el número de la curva del gradiente se determinó usando un programa informático de cromatografía Waters Empower. El caudal fue

de 0,6 mL por minuto y la columna se mantuvo a 45°C. El volumen de inyección fue de 1 µL y los alcaloides se detectaron usando un detector UV a 285 nm.

La pérdida en el secado (LOD) de la paja se determinó mediante secado en una estufa a 105°C durante 3 horas.

5

#### Programa de flujo de gradiente

TIEMPO (minutos)	% De eluyente A	% De eluyente B	Flujo (ml/min)	Curva No
0,00	95,0	5,0	0,60	INICIAL
0,80	90,0	10,0	0,60	6
3,40	75,0	25,0	0,60	3
3,60	95,0	5,0	0,60	6
4,00	95,0	5,0	0,60	11

Las concentraciones de alcaloide para la morfina, la codeína, la tebaína, la oripavina y la noscapina se determinaron mediante comparación con las soluciones estándar y los resultados se calcularon en base al peso seco.

10

Los tiempos de retención típicos fueron los siguientes:

Compuesto	Tiempo de retención (minutos)
Morfina	1,14
Pseudo morfina	1,26
Codeína	1,69
Oripavina	1,80
10-Hidroxitebaína	2,32
Tebaína	2,53
Noscapina	3,16

#### 15 Silenciamiento génico inducido por virus (VIGS) de *PSCYP3* y *PSCYP3*

##### a) Generación de constructos de silenciamiento

Se usó un sistema de silenciamiento del gen inducido por virus basado en el virus del tabaco cascabeleo (TRV) desarrollado y descrito por Liu et al. (2002) Plant J. 30(4): 415-429 para investigar la función del gen de *PSCYP2* y *PSCYP3*. Los fragmentos de ADN seleccionados para el silenciamiento de *PSCYP2* y *PSCYP3*, respectivamente, se amplificaron por PCR y se clonaron en el vector de silenciamiento pTRV2 (Número de acceso de GenBank . AF406991; Liu et al. (2002) Plant J. 30(4): 415-429). Se unieron a un fragmento de 129 pb de longitud del gen *P. somniferum phytoene desaturase (PsPDS)* para silenciar los genes respectivos del citocromo P450 y *PsPDS* simultáneamente. Las plantas que mostraron el fenotipo de blanqueamiento que resultó del silenciamiento de *PsPDS* (Hileman et al. (2005) Plant J. 44(2): 334-341) se identificaron como plantas infectadas satisfactoriamente con los constructos respectivos de silenciamiento y se seleccionaron para análisis.

Generación del constructo pTRV2: PDS: Un fragmento de 622 pb (Figura 11) de *PsPDS* se amplificó a partir del ADNc preparad a partir de GSK NOSCIPINA CVS1 como se ha descrito anteriormente usando los cebadores ps\_pds\_F y ps\_pds\_R4 (Tabla 4). La secuencia del cebador directo se basó en un contig de 412 pb de longitud derivado de las bibliotecas EST que compartía 99% de identidad en su extremo 3' con la secuencia codificante parcial de *P. somniferum phytoene desaturase* (No de acceso de GenBank . DQ116056). La secuencia del cebador inverso se diseñó basándose en la secuencia DQ116056. Las condiciones para la PCR fueron idénticas a las descritas anteriormente para la amplificación de los genes de citocromo P450 a partir de la secuencia genómica excepto que la etapa de hibridación se llevó a cabo a 70°C y el tiempo de extensión se aumentó a 60 segundos.

La digestión con Sau3AI de los fragmentos de PCR produjo entre otros dos fragmentos (280 pb y 129 pb de longitud) que presentaban unos extremos de unión compatibles con BamHI en ambos extremos 5' y 3'. El fragmento de 129 pb de longitud (tramo subrayado en Figura 11) se clonó en el sitio BamHI del vector pTRV2. Debido a que se usó Sau3AI para producir los extremos de unión compatibles con BamHI, el sitio BamHI en el extremo 5' del inserto PDS se abolió en el constructo pYL156-PDS. Sin embargo, el sitio de reconocimiento de BamHI en su extremo 3' se mantuvo intacto debido a la naturaleza de la secuencia del inserto PDS.

Un constructo de secuencia confirmada pTRV2-PDS, con el fragmento de 129 pb en orientación sentido, se usó



posteriormente para la generación de los constructos de silenciamiento *PSCYP2* y *PSCYP3*, y sirvió como control en los experimentos VIGS.

- 5 Generación de constructos de silenciamiento para *PSCYP2* y *PSCYP3* (pTRV2-PDS-*PSCYP2* y pTRV2-PDS-*PSCYP3*): Los fragmentos de ADN seleccionados para el silenciamiento de *PSCYP2* y *PSCYP3* se amplificaron a partir del ADNc de GSK NOSCAPINA CVS1 preparado como se ha descrito anteriormente con el uso de las secuencias de cebador mostradas en la Tabla 4. Se añadieron sitios de restricción adicionales (cebadores directos: XhoI y HindIII para los cebadores directos; sitio KpnI para los cebadores inversos) a los cebadores específicos del gen para facilitar la clonación. Las condiciones de amplificación fueron como las descritas anteriormente para la  
10 amplificación del fragmento *PDS* excepto que las temperaturas de hibridación fueron 60,9 °C para *PSCYP2* y 66°C para *PSCYP3* y el tiempo de extensión fue 30 segundos.

- Las secuencias seleccionadas para silenciar *PSCYP2* (Figura 12) y *PSCYP3* (Figura 12), respectivamente, fueron clonadas en pTV00 (Ratcliff et al. (2001) Plant J. 25(2): 237-245) usando HindIII y KpnI y subclonadas en TRV2-PDS  
15 usando BamHI y KpnI. Los constructos de secuencia confirmada pTRV2-PDS-*PSCYP2* y pTRV2-PDS-*PSCYP3* se usaron en los experimentos VIGS.

#### **b) Transformación de los constructos en *Agrobacterium tumefaciens***

- 20 La propagación de los constructos de silenciamiento se llevó a cabo con la cepa DH5α de *E. coli* y, posteriormente, los constructos de silenciamiento respectivos, así como pTRV1 (nº de acceso de GenBank. AF406990; Liu et al. (2002) Plant J. 30(4): 415-429) se transformaron independientemente en *Agrobacterium tumefaciens* electrocompetente (cepa GV3101).

#### **25 c) Infiltración de plantas**

- Los cultivos líquidos durante una noche de *A. tumefaciens* que contienen cada constructo de silenciamiento se usaron para inocular medio Luria-Bertani (LB) que contiene MES 10 mM, acetosiringona 20 µM y kanamicina a 50 µg/ml. Los cultivos se mantuvieron a 28°C durante 24 horas, se recogieron por centrifugación a 3000 g durante 20  
30 minutos y se resuspendieron en una solución de infiltración (MES 10 mM, acetosiringona 200 µM, MgCl<sub>2</sub> 10 mM) a una DO600 de 2,5. *A. tumefaciens* que alberga los constructos respectivos (pTRV2-PDS-*PSCYP2*, pTRV2-PDS-*PSCYP3* o, como control, pTRV2-PDS) se mezclaron cada uno a 1:1 (v/v) con *A. tumefaciens* que contiene pTRV1, y se incubaron durante dos horas a 22°C antes de la infiltración. Se infiltraron plántulas maduras de GSK NOSCAPINA CVS1 de dos semanas de edad cultivadas en condiciones de invernadero estándar (22°C, fotoperíodo  
35 de 16 horas), con primeras hojas emergentes, tal como se describe en Hagel and Facchini (2010) Nat. Chem. Biol. 6: 273-275.

#### **d) Análisis de látex y cápsula de plantas silenciadas**

- 40 Se analizó el látex foliar de las plantas de adormidera infiltradas que mostraban for-blanqueamiento como marcador visual de una infección y silenciamiento exitosos cuando aparecieron los primeros brotes de las flores (plantas de 7~ semanas de edad). Las plantas que presentaban un grado similar de blanqueamiento fotográfico de las hojas se seleccionaron para análisis.
- 45 Se recogió látex de pecíolos cortados, con una sola gota dispersa en 500 µl de ácido acético al 10%. Esto se diluyó 10 veces en ácido acético al 1% para proporcionar una solución de alcaloide en ácido acético al 2% para un análisis posterior. Las cápsulas de las mismas plantas usadas para el análisis de látex se recolectaron manualmente del invernadero de cultivo y las cápsulas separadas se molieron en un molinillo de bolas de Retsch, modelo MM04 hasta conseguir un polvo fino. Se pesaron diez mg de las muestras de paja de adormidera molida con precisión (10 ± 0,1  
50 mg) y se extrajeron en 0,5 ml de una solución de ácido acético al 10% con agitación suave durante 1h a temperatura ambiente. Después, las muestras se clarificaron mediante centrifugación y una submuestra de 50 µl se diluyó 10 veces en ácido acético al 1% para dar una solución de alcaloide en ácido acético al 2% para un análisis posterior. Todas las soluciones se analizaron usando un sistema Waters Acquity UPLC (Waters Ltd., Elstree, Reino Unido). equipado con una columna Waters Acquity BEH C18, 2.1mm x 100mm con un empaquetamiento de 1,7 micras. La  
55 fase móvil usó un perfil de gradiente de eluyente A que consistía en 10 mM de bicarbonato de amonio de pH 10.2 y eluyente B metanol. Las condiciones de gradiente de fase móvil utilizadas se muestran en la Tabla 1 con un gradiente lineal. El caudal fue de 0,5 ml por minuto y la columna se mantuvo a 60°C. El volumen de inyección fue de 2 µl y los picos eluidos se ionizaron en modo APCI positivo y se detectaron con una precisión de masa de 3 ppm utilizando un Thermo LTQ-Orbitrap. Las pruebas fueron controladas por el software Thermo Xcalibur.

60

**Programa de flujo de gradiente:**

TIEMPO (minutos)	% De eluyente A	% De eluyente B	Flujo (ml/min)
0,0	98,0	2,0	0,50
0,2	98,0	2,0	0,50
0,5	60,0	40	0,50
4,0	20,0	80,0	0,50
4,5	20,0	0,0	0,50

- Todos los análisis de datos se llevaron a cabo en R. Los picos putativos de alcaloides se cuantificaron por sus áreas de iones pseudomoleculares utilizando unos scripts personalizados. Las listas de picos se compilaron y cualquier diferencia significativa entre las muestras se identificó utilizando un test ANOVA de una vía con valores de p ajustados usando la corrección de Bonferroni para el número de picos únicos en el conjunto de datos. Para cualquier comparación de picos con valores de p ajustados <0.05, la prueba de HSD de Tukey se utilizó para identificar picos que fueron significativamente diferentes entre cualquier muestra dada y el control. Los alcaloides se identificaron comparando la masa exacta y los valores de tiempo de retención con los de los estándares. Donde los estándares no estaban disponibles, masas exactas neutras se usaron para generar los de las fórmulas moleculares dentro de las restricciones elementales de C = 1:100, H = 1:200, O = 0:200, N = 0:3 y exactitud de masa <20 ppm. El hit con el menor error de ppm dentro de estas restricciones se usó para asignar una fórmula putativa.

**15 Ejemplo 1****Ensamblaje de la secuencia de longitud completa de ADNc de *PSCYP1* de EST y confirmación mediante secuenciación a partir del ADN genómico.**

- 20 El marco de lectura abierto de longitud completa de *PSCYP1* (Figura 1a) se ensambló a partir de las EST derivadas de la plataforma de secuenciación 454 usando el programa de ensamblaje de secuencias CAP3. La secuencia de longitud completa de ADNc se confirmó mediante amplificación directa del ADNc de la longitud completa del ADN genómico de GSK NOSCAPINA CVS1.

**25 Ejemplo 2****El *PSCYP1* se expresa exclusivamente en el cultivar de *Papaver somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

- 30 La Figura 2a muestra la distribución normalizada de las EST asociadas con la secuencia consenso de *PSCYP1* a lo largo de cada una de las 16 bibliotecas de EST preparadas a partir de dos órganos (cápsulas y tallos) de dos etapas de desarrollo (cosecha temprana y tardía) de cada uno de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBÁINA CVS1. Las EST correspondientes a *PSCYP1* se encontraron exclusivamente en bibliotecas derivadas del cultivar GSK NOSCAPINA CVS1 que produce noscapina (Figura 2a). La expresión de *PSCYP1* fue más fuerte en los tejidos de los tallos inmediatamente después de la floración.

**Ejemplo 3**

- 40 **Amplificación por PCR de *PSCYP1* a partir del ADN genómico de los cuatro cultivares de *Papaver somniferum* GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBÁINA CVS1.**

- Las amplificaciones por PCR de fragmentos de *PSCYP1* se realizaron sobre el ADN genómico de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBÁINA CVS1 usando las combinaciones de cebadores que se muestran en la Tabla 2 y 3.

- La Figura 5 muestra la amplificación por PCR de *PSCYP1* a partir del ADN genómico de los cuatro cultivares de *Papaver somniferum* GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBÁINA CVS1;

- 50 La amplificación del ADN genómico produjo la secuencia de genes que se muestra en la Figura 3a.

**Ejemplo 4**

**La proteína putativa codificada por *PSCYP1* muestra una similitud más elevada a un citocromo P450 de *Coptis japonica* y *Thalictrum flavum*.**

Los homólogos más cercanos a la proteína putativa codificada por el marco de lectura abierto de *PSCYP1* (Figura 4a) son un citocromo P450 de *Coptis japonica* (nº de acceso de GenBank: BAF98472.1, 46% idéntico a nivel de aminoácido). El homólogo más cercano con una asignación a la subfamilia del citocromo P450 es CYP82C4 de *Arabidopsis lyrata* (nº de acceso de GenBank: XP\_002869304.1, 44% idéntico a nivel del aminoácido).

**Ejemplo 5**

**El *PSCYP1* sólo está presente en el genoma del cultivar de *P. somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

La región transcrita cubierta por los EST contenía la secuencia codificante completa de *PSCYP1* (incluyendo las regiones no traducidas 5' y 3'), que se usó para el diseño del cebador (Tabla 1) para amplificar el gen *PSCYP1* de ADN genómico en unas series de fragmentos solapados para la secuenciación. Al ensayar un subconjunto de combinaciones del cebador (Tabla 3) en las muestras de ADN genómico para los cuatro cultivares se encontró que los fragmentos *PSCYP1* sólo se podían amplificar a partir del ADN genómico del cultivar GSK NOSCAPINA CVS1 que produce noscapina pero no desde el ADN genómico de los cultivares que producen predominantemente morfina (GSK MORFINA CVS1, GSK MORFINA CVS2) o tebaína GSK TEBAINA CVS1 (Figura 5). Las amplificaciones por PCR se realizaron en conjuntos de ADN genómico que comprenden ADN de cuatro individuos por cultivar. Este descubrimiento explica por qué el *PSCYP1* sólo se expresa en el cultivar de GSK NOSCAPINA CVS1 y no está presente en el transcriptoma de los otros tres cultivares.

**Ejemplo 6**

**Ensamblaje de la secuencia de longitud completa de ADNc de *PSCYP2* de EST y confirmación por secuenciación a partir del ADN genómico.**

El marco de lectura abierto de longitud completa de *PSCYP2* (Figura 1b) se ensambló a partir de las EST derivadas de la plataforma de secuenciación 454 usando el programa de ensamblaje de secuencias CAP3. La secuencia de longitud completa de ADNc se confirmó mediante amplificación directa del ADNc de longitud completa a partir del ADN genómico de GSK NOSCAPINA CVS1.

**Ejemplo 7**

**El *PSCYP2* se expresa exclusivamente en el cultivar de *Papaver somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

La Figura 2b muestra la distribución normalizada de las EST asociadas con la secuencia consenso de *PSCYP2* a lo largo de cada una de las 16 bibliotecas de EST preparadas a partir de dos órganos (cápsulas y tallos) de dos etapas de desarrollo (cosecha temprana y tardía) de cada uno de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1. Las EST correspondientes a *PSCYP2* se encontraron exclusivamente en bibliotecas derivadas del cultivar GSK NOSCAPINA CVS1 que produce noscapina (Figura 2b). La expresión de *PSCYP2* fue más fuerte en los tejidos de los tallos inmediatamente después de la floración.

**Ejemplo 8**

**Amplificación por PCR de *PSCYP2* a partir del ADN genómico de los cuatro cultivares de *Papaver somniferum* GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1.**

Las amplificaciones por PCR de fragmentos de *PSCYP2* se realizaron sobre el ADN genómico de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1 usando las combinaciones de cebadores que se muestran en la Tabla 2 y 3. La Figura 6 muestra la amplificación por PCR de *PSCYP2* a partir del ADN genómico de los cuatro cultivares de *Papaver somniferum* GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1;

La amplificación del ADN genómico produjo la secuencia de genes que se muestra en la Figura 3b.

#### Ejemplo 9

- 5 **La proteína putativa codificada por *PSCYP2* muestra una similitud más elevada a un citocromo P450 de *Coptis japonica* y *Thalictrum flavum*.**

Los homólogos más cercanos a la proteína putativa codificada por el marco de lectura abierto de *PSCYP2* (Figura 4b) son citocromos P450s anotados como estilopina sintasa de *Argemone mexicana* (nº de acceso de GenBank: ABR14721, identidades: 366/475 (78%)) y de *Papaver somniferum* (nº de acceso de GenBank: ADB89214, identidades = 373/491 (76%)). La comparación de secuencias se realizó usando un algoritmo de NCBI 'blastp' (método: ajuste composicional de la matriz).

#### Ejemplo 10

- 15 **El *PSCYP2* sólo está presente en el genoma del cultivar de *P. somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

La región transcrita cubierta por las EST contenía la secuencia codificante completa de *PSCYP2* (incluyendo las regiones no traducidas 5' y 3'), que se usó para el diseño del cebador (Tabla 1) para amplificar el gen *PSCYP2* a partir de ADN genómico en unas series de fragmentos solapados para la secuenciación. Al ensayar un subconjunto de combinaciones del cebador (Tabla 3) en las muestras de ADN genómico para los cuatro cultivares se encontró que los fragmentos *PSCYP2* sólo se podían amplificar a partir del ADN genómico del cultivar GSK NOSCAPINA CVS1 que produce noscapina pero no desde el ADN genómico de los cultivares que producen predominantemente morfina (GSK MORFINA CVS1, GSK MORFINA CVS2) o tebaína GSK TEBAINA CVS1 (Figura 6). Las amplificaciones por PCR se realizaron en conjuntos de ADN genómico que comprenden ADN de cuatro individuos por cultivar. Este descubrimiento explica por qué el *PSCYP2* sólo se expresa en el cultivar de GSK NOSCAPINA CVS1 y no está presente en el transcriptoma de los otros tres cultivares.

#### Ejemplo 11

**Ensamblaje de la secuencia de longitud completa de ADNc de *PSCYP3* a partir de EST y por secuenciación a partir del ADNc y del ADN genómico.**

35 Dos posibles marcos de lectura abiertos de longitud completa de *PSCYP3* (Figura 1c y 1c) se ensamblaron parcialmente a partir de las EST derivadas de la plataforma de secuenciación 454 usando el programa de ensamblaje de secuencias CAP3. Las EST cubrieron las áreas 5' y 3' de la secuencia con un tramo de 200 bases perdidas. El tramo perdido de bases se obtuvo por amplificación directa y secuenciación a partir del ADNc de la GSK NOSCAPINA CVS1. Las secuencias de longitud completa se confirmaron además por amplificación y secuenciación directa de *PSCYP3* a partir del ADN genómico de la GSK NOSCAPINA CVS1. Se identificaron dos posibles codones de iniciación ATG. Dado que estaban en la estructura y adyacente uno al otro las secuencias de cadena completa del marco de lectura abierto que se muestran en la Figura 1c y 1d, respectivamente, se diferencian sólo por un codón ATG en el 5'-terminal.

#### Ejemplo 12

**El *PSCYP3* se expresa exclusivamente en el cultivar de *Papaver somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

50 La Figura 2c muestra la distribución normalizada de las EST asociadas con la secuencia consenso de *PSCYP3* a lo largo de cada una de las 16 bibliotecas de EST preparadas a partir de dos órganos (cápsulas y tallos) de dos etapas de desarrollo (cosecha temprana y tardía) de cada uno de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAINA CVS1. Las EST correspondientes a *PSCYP3* se encontraron exclusivamente en bibliotecas derivadas del cultivar GSK NOSCAPINA CVS1 que produce noscapina (Figura 2c). La expresión de *PSCYP3* fue más fuerte en los tejidos de los tallos inmediatamente después de la floración.

#### Ejemplo 13

- 60 **Amplificación por PCR de *PSCYP3* a partir del ADN genómico de los cuatro cultivares de *Papaver***

**somniferum GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAIÑA CVS1.**

Las amplificaciones por PCR de fragmentos de *PSCYP3* se realizaron sobre el ADN genómico de los cuatro cultivares de adormidera GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAIÑA CVS1 usando las combinaciones de cebadores que se muestran en la Tabla 2 y 3. La Figura 7 muestra la amplificación por PCR de *PSCYP3* a partir del ADN genómico de los cuatro cultivares de *Papaver somniferum* GSK MORFINA CVS1, GSK MORFINA CVS2, GSK NOSCAPINA CVS1 y GSK TEBAIÑA CVS1;

La amplificación a partir del ADN genómico produjo la secuencia de genes que se muestra en la Figura 3c.

**10 Ejemplo 14**

**La proteína putativa codificada por *PSCYP3* muestra una similitud de secuencia más elevada con la protopina 6-hidroxilasa de *Eschscholzia californica*.**

15 El homólogo más cercano a las proteínas putativas codificadas por los dos marcos de lectura abiertos posibles de *PSCYP3* (Figura 1c y 1d) es un citocromo P450s anotado como protopina 6-hidroxilasa de *Eschscholzia californica* (nº de acceso de GenBank: BAK20464, identidades: 228/522 (44%)) y un citocromo P450 de *Coptis japonica* (nº de acceso de GenBank: BAF98472, identidades = 230/539 (43%)). La comparación de secuencias se realizó usando un algoritmo de NCBI 'blastp' (método: ajuste composicional de la matriz).

**20 Ejemplo 15**

**El *PSCYP3* sólo está presente en el genoma del cultivar de *P. somniferum* GSK NOSCAPINA CVS1 que produce noscapina.**

25 La región transcrita cubierta por los EST contenía la secuencia codificante parcial de *PSCYP3* (incluyendo las regiones no traducidas 5' y 3'), que se usó para el diseño del cebador (Tabla 1) para amplificar el gen *PSCYP3* a partir del ADN genómico en una serie de fragmentos solapados para la secuenciación. Al ensayar un subconjunto de combinaciones del cebador en las muestras de ADN genómico para los cuatro cultivares se encontró que los fragmentos *PSCYP3* sólo se podían amplificar a partir del ADN genómico del cultivar GSK NOSCAPINA CVS1 que produce noscapina pero no desde el ADN genómico de los cultivares que producen predominantemente morfina (GSK MORFINA CVS1, GSK MORFINA CVS2) o tebaína (GSK TEBAIÑA CVS1) (Figura 7). Las amplificaciones por PCR se realizaron en conjuntos de ADN genómico que comprenden ADN de cuatro individuos por cultivar usando las combinaciones de cebador en la Tabla 3. Este descubrimiento explica por qué el *PSCYP3* sólo se expresa en el cultivar de GSK NOSCAPINA CVS1 y está ausente en el transcriptoma de los otros tres cultivares.

**Ejemplo 16**

**Análisis de segregación de *PSCYP1* y producción de noscapina en un mapeo de población F2 derivado del cruce entre GSK NOSCAPINA CVS1 y GSK TEBAIÑA CVS1.**

El cultivar GSK NOSCAPINA CVS1, que produce noscapina se polinizó por polinización cruzada con GSK TEBAIÑA CVS1 que produce unas cantidades insignificantes de noscapina. Las plantas F1 resultantes se cultivaron hasta su madurez y se recolectaron las semillas de F2. Doscientos setenta y cinco individuos F2 de 10 cruces de GSK NOSCAPINA CVS1 y GSK TEBAIÑA CVS1 se cultivaron hasta su madurez en el campo. El material de las hojas de cada individuo se recolectó y se usó para la extracción y el análisis del ADN. Las cápsulas maduras de cada individuo se recolectaron para la extracción y el análisis del alcaloide.

Las Figuras 8a-c muestran los resultados del análisis del mapeo de la población F2. Los genes *PSCYP1*, *PSCYP2* y *PSCYP3* se unen y segregan con la producción de noscapina en el mapeo de la población F2. Los datos demuestran que en el mapeo de la población GSK NOSCAPINA CVS1, los niveles están presentes en 61 de los 275 individuos de plantas F2. Los genes *PSCYP1*, *PSCYP2* y *PSCYP3* se detectaron en todas las plantas que contienen noscapina lo que de este modo confirma que los genes *PSCYP1*, *PSCYP2* y *PSCYP3* y la producción de noscapina están vinculados. Además, todas las plantas en las cuales no se detectaron los genes *PSCYP1*, *PSCYP2* y *PSCYP3* carecían de noscapina (El ensayo de genotipado de *PSCYP3* falló en 16 muestras tal como indica el fracaso del control positivo interno incluido en el ensayo; ya que estas muestras se excluyeron del análisis de segregación de *PSCYP3* con el rasgo noscapina). Estos datos son estadísticamente muy relevantes y confirman que los genes *PSCYP1*, *PSCYP2* y *PSCYP3* son necesarios para la producción de niveles de noscapina de GSK NOSCAPINA CVS1.

60

**Ejemplo 17****Acumulados de tetrahydrocolumbamina putativa en plantas con *PSCYP2*-silenciado**

- 5 El silenciamiento de genes inducido por virus conduce a la acumulación de tetrahydrocolumbamina putativa tanto en el látex como en las cápsulas maduras de las plantas con ***PSCYP2*-silenciado** pero no en las plantas con *PSCYP3*-silenciado, las plantas control con *PDS*-silenciado o las plantas no infectadas de GSK NOSCAPINA CVS1 (Figure 14). Los datos sugieren que *PSCYP2* codifica para una enzima que forma puentes metilenodioxo que convierte tetrahydrocolumbamina en canadina lo que de este modo conduce a la formación del puente metilenodioxo presente
- 10 en los C-3a/C-9a' del resto isoquinolina de noscapina.

**Ejemplo 18****Acumulados de secoberbinas putativas en plantas con *PSCYP3*-silenciado**

- 15 El silenciamiento de genes inducido por virus conduce a la acumulación de alcaloides putativos de secoberbina tanto en el látex como en las cápsulas maduras de las plantas con *PSCYP3*-silenciado pero no en las plantas con ***PSCYP2*-silenciado**, las plantas control con *PDS*-silenciado o las plantas no infectadas de GSK NOSCAPINA CVS1 (Figura 15 y 16). La masa, la fórmula molecular asignada ( $C_{21}H_{23}NO_6$ ) y el patrón de fragmentación de la
- 20 secoberbina putativa descritos que se acumulan en la Figura 15 es consistente con desmetoxihidroxi-macrantaldéhid o desmetoximacrantoridina. Ambas secoberbinas carecen de un grupo metoxi en el carbono del resto isoquinolina que es equivalente al C-4' de noscapina. La masa, la fórmula molecular asignada ( $C_{21}H_{25}NO_6$ ) y el patrón de fragmentación del segundo compuesto encontrado que se acumulan en las plantas con *PSCYP3*-silenciado (Figura 16) es consistente con las dos secoberbinas, desmetoxinarcotinodiol y narcotolinol,
- 25 respectivamente. El compuesto anterior carece del grupo metoxi en el carbono que es equivalente al C-4' de noscapina. Conjuntamente los datos sugieren que la proteína codificada por *PSCYP3*-hidroliza el resto isoquinolina de las secoberbinas en una posición equivalente al C-4' de noscapina lo que permite de este modo la formación del grupo metoxi presente en noscapina en esta posición mediante la O-metilación posterior. Los derivados metoxilados respectivos (metoxilados en el carbono equivalente al C-4' de noscapina) de las secoberbinas putativas que se
- 30 acumulan en las plantas con *PSCYP3*-silenciado se han encontrado en varias especies de *Papaver* que producen noscapina (Sariyar and Phillipson (1977) Phytochem. 16: 2009-2013; Sariyar and Shamma (1986) Phytochem. 25: 2403-2406, Sariyar (2002) Pure Appl. Chem. 74: 557-574). Han estado implicados, en suelos estructurales, en la conversión biosintética de protoberberinas en ftalideisoquinolinas como noscapina (Sariyar and Shamma (1986) Phytochem. 25: 2403-2406, Sariyar and Phillipson (1977) Phytochem. 16: 2009-2013).
- 35

## REIVINDICACIONES

1. Una molécula aislada de ácido nucleico que codifica un polipéptido del citocromo P450 donde dicha molécula de ácido nucleico comprende o consiste en una secuencia de nucleótido seleccionada de entre el grupo que consiste en:
  - i) una secuencia de nucleótidos representada por la secuencia en la Figura 1c, 1d ó 3c;
  - ii) una secuencia de nucleótidos donde dicha secuencia está degenerada como resultado del código genético de la secuencia de nucleótidos definida en (i);
  - 10 iii) una molécula de ácido nucleico cuya cadena complementaria se hibrida con secuencias de nucleótidos con al menos 90% de identidad bajo condiciones de hibridación muy rigurosas a la secuencia en la Figura 1c, 1d ó 3c donde dicha molécula de ácido nucleico codifica un polipéptido del citocromo P450;
  - iv) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos como se representa en la Figura 4c;
  - 15 v) una secuencia de nucleótidos que codifica un polipéptido que comprende una secuencia de aminoácidos donde dicha secuencia de aminoácidos se modifica por adición, delección o sustitución de al menos un residuo de aminoácido tal como se representa en iv) anterior y que presenta al menos un 85% de secuencia idéntica a la secuencia de longitud completa representada en la Figura 4c que ha retenido la actividad del citocromo P450.
- 20 2. Un polipéptido aislado seleccionado del grupo que consiste en:
  - i) un polipéptido que comprende o que consiste en una secuencia de aminoácidos como se representa en la Figura 4c; o
  - ii) un polipéptido modificado donde los polipéptidos variantes presentan al menos un 85% de identidad con la
  - 25 secuencia de aminoácidos de longitud completa representada en la Figura 4c y que retiene la actividad del citocromo P450.
3. El polipéptido aislado de acuerdo con la reivindicación 2 donde dicho polipéptido comprende o consiste en una secuencia de aminoácidos que es un 85% idéntica a la secuencia de aminoácidos de longitud
- 30 completa en la Figura 4c y que ha retenido la actividad del citocromo P450.
4. Un vector que comprende una molécula de ácido nucleico que codifica un polipéptido del citocromo P450 de acuerdo con la reivindicación 1 donde dicha molécula de ácido nucleico está unida operativamente a una molécula de ácido nucleico que comprende una secuencia promotora.
- 35 5. Una célula transgénica transformada o transfectada con una molécula de ácido nucleico o un vector de acuerdo con la reivindicación 1 o 4.
6. Una célula transgénica de acuerdo con la reivindicación 5 donde la célula es una célula vegetal donde
- 40 dicha célula vegetal es opcionalmente de la familia *Papaveraceae*.
7. Una planta que comprende una célula vegetal de acuerdo con la reivindicación 6 opcionalmente dicha planta es de la familia *Papaveraceae*.
- 45 8. Una célula transgénica de acuerdo con la reivindicación 5, donde dicha célula es una célula microbiana.
9. Molécula de ácido nucleico de acuerdo con la reivindicación 1 que comprende un casete de transcripción seleccionado de entre el grupo que consiste en
- 50 i) un casete que está adaptado de modo que las moléculas de ácido ribonucleico sentido y antisentido se transcriban a partir de dicho casete, donde dichas moléculas de ácido nucleico sentido y antisentido están adaptadas para reconocer al menos una parte o la totalidad de su longitud para formar un ARNsi o ARNhc.
- ii) un casete tal como se define en (i) anterior que está provisto de al menos dos promotores adaptados para
- 55 transcribir cadenas sentido y antisentido de dicha molécula de ácido ribonucleico.
- iii) un casete tal como se define en (i) anterior que comprende una molécula de ácido nucleico donde dicha molécula comprende una primera parte unida a una segunda parte, donde dichas primera y segunda parte son complementarias de al menos parte de su secuencia y además donde la transcripción de dicha molécula de ácido nucleico produce una molécula de ácido ribonucleico que forma una región bicatenaria por apareamiento de bases
- 60 complementarias de dichas primera y segunda parte, formando así un ARNhc.

iv) un casete de acuerdo con una cualquiera de i) a iv) que comprende una molécula de ácido nucleico que es parte de un vector adaptado para la expresión en una célula vegetal.

10. Un proceso para la modificación de un alcaloide opiáceo que comprende:

- i) proporcionar una célula vegetal transgénica de acuerdo con la reivindicación 6;
- ii) cultivar dicha célula vegetal para producir una planta transgénica; y opcionalmente
- iii) cosechar dicha planta transgénica, o parte de la misma, donde opcionalmente dicho material vegetal cosechado es paja seca y dicho alcaloide opiáceo se extrae.

11. Un proceso para la modificación de un alcaloide opiáceo que comprende:

- i) proporcionar una célula microbiana transgénica de acuerdo con la reivindicación 8 que expresa un citocromo P450 de acuerdo con la invención en cultivo con al menos un alcaloide opiáceo;
- ii) cultivar la célula microbiana en unas condiciones que modifican uno o más alcaloides opiáceos; y opcionalmente
- iii) aislar dicho alcaloide opiáceo de la célula microbiana o cultivo celular.

12. El uso de un gen codificado por una molécula de ácido nucleico representada por la secuencia de ácido nucleico en la Figura 1d o una molécula de ácido nucleico que se hibrida en condiciones de hibridación astringentes con la secuencia de nucleótidos en la Figura 1d y codifica un polipéptido con actividad del citocromo P450 como medio para identificar la presencia o ausencia de un gen que codifica dicho citocromo P450 o como medio para identificar un locus donde dicho locus está asociado con una expresión o actividad alterada de dicho citocromo P450 en una planta *Papaveraceae*

13. Un procedimiento para determinar la presencia o ausencia de un gen de citocromo P450 en una variedad *Papaveraceae* que comprende:

- i) obtener una muestra de una planta *Papaveraceae*;
- ii) extraer el ADN genómico de la planta; y
- iii) analizar el ADN genómico para la presencia de un gen que comprende o consiste en una secuencia de nucleótidos tal como se representa en la Figura 3c.

14. Una planta que presenta una expresión alterada de un polipéptido del citocromo P450 codificado por una molécula de ácido nucleico de acuerdo con la reivindicación 1 mediante un procedimiento que comprende las etapas de:

- i) mutagénesis de semillas de tipo silvestre de una planta que expresa dicho polipéptido citocromo P450;
- ii) cultivo de la semilla en i) para producir las primeras y posteriores generaciones de plantas;
- iii) obtención la semilla de la planta de la primera generación y las siguientes generaciones de plantas;
- iv) determinar si la semilla de dicha primera y siguientes generaciones de plantas presenta una secuencia de nucleótidos alterada y/o una expresión alterada de dicho polipéptido de citocromo P450.

15. Un procedimiento para el análisis de una planta que comprende:

- i) obtener una muestra de la planta de acuerdo con la reivindicación 14 y analizar la secuencia de ácido nucleico de una molécula de ácido nucleico seleccionada de entre el grupo que consiste en:
  - a) una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se representa en la Figura 3c;
  - b) una molécula de ácido nucleico que se hibrida con la molécula de ácido nucleico en a) bajo condiciones de hibridación astringentes y que codifica un polipéptido con actividad polipéptido citocromo P450; y opcionalmente

ii) comparar la secuencia de nucleótidos de la molécula de ácido nucleico en dicha muestra con una secuencia de nucleótidos de una molécula de ácido nucleico de la planta original de tipo silvestre, opcionalmente dicho procedimiento comprende también las etapas de:

iii) extraer el ácido nucleico de dichas plantas mutadas;

iv) amplificar una parte de dicha molécula de ácido nucleico mediante una reacción en cadena de la polimerasa;

v) formar una preparación que comprende el ácido nucleico amplificado y el ácido nucleico extraído de la semilla de



tipo silvestre para formar ácido nucleico heterodúplex;

vi) incubar dicha preparación con una nucleasa monocatenaria que corta en una región de ácido nucleico heterodúplex para identificar el desapareamiento en dicho heterodúplex; y

5

vii) determinar el sitio del desapareamiento en dicho heterodúplex de ácido nucleico.

16. Una planta donde dicha planta comprende un vector viral que incluye todo o parte de un gen que comprende una molécula de ácido nucleico de acuerdo con la reivindicación 1.

10

17. Un vector viral que comprende toda o parte de una molécula de ácido nucleico de acuerdo la reivindicación 1.

Figura 1a

ATGGAGTTATTTCATAAAGTTACCATTTATCCAACCAATTCCCTTTCAGTATTATTCTTGTTACTACAGT  
 TTCGATTGTTCTATTATACAGTGTCTTCTTCTGGGTTACTGATAAGAAAAAGAAGAGGAAGAAAGCAC  
 CAAATGCTGCAGGGGCATGGCCGTTAATAGGTCATCTCCGTCTATTGATGAACGACAAGGAACCGTTG  
 TATAGAGCACTAGGGAGCATGGCTGATAAGTACGGACCTGCATTCAACATCCGATTAGGTAACCAAGA  
 AGTTCTTGTTGTGAGTAACTGGGAGATGGTAAAACAGTGTGTTGGTAATCAAATGATAAGCTATTTT  
 CGAATCGTCAAACCTACATTAGCTGCAAAATACATGCTTAATCAAACAACCTCTAGCGGATTTCGCACCA  
 TATGGACCATATTGGAGAGAGCTACGAAAGATAATGGTGCAGCAATTACTCTCTAAACAATCTTTAGA  
 ATCGTGGAAACATCTGAAAATCAAAGAGATGGATGCTTCATTTAGTAAACTTAACGAGTTATGCAACA  
 ACAACGGTACTGGAACAGCTACCTAATTAGGATGGACGAATGGTTTGCTGAGTTGACGTTCAACGTG  
 ATCGCAAGAAATGTCTTTGGCTACCAAAGTGGCGGAAGGTCGACAGCGCTTACGAACGGAGATACGGA  
 ATCAAAGGGCGAGAGGTACAAGAAAACATTGGAAGAAGCACTTCATCTTATGTCAATTTTTGACGTTT  
 CAGACATATTTCCAAGTCTAGAGTGGGTAGATCGGTTAAGAGGCCTTATAAGGAATATGAAACGCTTT  
 GGAGATGAGCTAAATTCATTTGCAGGGTGTCTTATTGAAGAGCACCGCCAAAAGAGATTACAATCCGT  
 ATCTAAAAGTGATAAAGGAGTTGGTGATGAACAAGACTTCGTTGATGTTCTCTTATCGGTTGCTGAAA  
 AATCGCAACTTCCTGGAGATGACCCGTGATTTGGTCATCAAGTCTATGATTCTGGAAATCGTATCAGGT  
 GGGAGTGAGACCACATCGTCAACCTTAACCTGGGCCCTCTGTCTGTTACTGAACCATCCGCATGTGTT  
 AAAGAAGGCAAAAGAGGAATTAGATACGCACGTAGGAAAAGATAGGCATGTAGAAGAGTCAGATACCC  
 CTAAGCTCGTGATTAATGCAATTATCAAAGAATCAATGCGATTGTATCCAAACGGGGCAATGCTT  
 GATCGGTTGGCGTTAGAAGAGTGCGAAGTTGGTGGATTTTCATGTACCGGCCGGGGGACGCTTATTTGT  
 CAATGTTTGGAAGATTCAGAGAGATCCGAGTGTGTTGGGAGAATCCTCTGGAGTTTAAACCAGAGAGGT  
 GGTTTTTGTAGTAATGGTGAAAAGATGGATGTGGATTACAAAGGTCACAATCATGAATTCATACCATTT  
 GGGATAGGTCGGAGGATGTGCGCTGGTATGCTTTGGGCATCGGAGGTGATTCATTTGGTGCTGCCCCG  
 TCTTATTCATGGGTTTGATATGAAAGCAGCAAGTGCCAAATGGGAAAGTAGATATGGCAGAAATGGCAG  
 GCATGGTGATTTGTTTAAAGAAGACACCTCTTGAAGTTATGGTCAATCCTCGAGAGTAG

Figura 1b

ATGATCATGAGTAACTTATGGATTCTTACGCTCATTTCTACCATATTAGCAGTCTTTGCTGCTGTGTT  
 AATCATTTTCAGGAGAAGAATATCAGCATCCACAACGGAATGGCCTGTTGGCCCAAAATTACCAATC  
 ATAGGTAACTTGCACATTCTTGGAGGCACCTGCTCTCCATGTCGTCTTACATAAACTTGCTGAAGTTTA  
 CGGCAGTGTAATGACGATATGGATTGGTAGTTGGAAACCTGTTATTATTTTCCGACTTTGATCGAGCC  
 TGGGAAGTTCTTGTAAACAAATCGTCAGATTATTCAGCTCGTGAAATGCCTGAGATCACTAAAATCGG  
 CACTGCAATTTGGAGAACAATTTCAAGTTCTGATTCTGGCCCTTTGGGCCACTCTTCGAAAAGGTCTT  
 CAGAGTGTAGCATTATCGCCTCAGCATTTAGCATCGCAACTGCACACCAAGAGAGAGATATAATAAA  
 GTTGATCAAAAATTTGAAAGACGAAGCAGTTCGGAATGGTTAAACCACTTGATCATCTCAAGAAAGCA  
 ACTGTAAGATTAAATCAGTCGGTTAATCTATGGTCAGGATTTTGATGACGATAAGTATGTTGAAGATAT  
 GCATGACGTGATCGAGTTTTGATCGTATTAGTGGTTATGCTCAACTTGCTGAGGTATTCTATTATGCT  
 AAATATCTACCAGGTCATAAGAGAGCTGTAAGTGGCGCCGAAGAAGCAAAAAGAAGAGTAATAGCTCT  
 GGTGCGTCTTTCTCAGTCAAACCTGCTACTAACACTTACTTGCAATTTCTCAAATCGCAACTGTAT  
 CCTGAAGAGGTTATCATATTGCTATATTGGAAGCTTATCTTTAGGTGTTGATAGCACTTCTTCAAC  
 ACTGCTGGGCACTCGCATCTTAATACGCGAACCATCTGTTCAAGAGAACTTTATCAAGAGCTTAAAG  
 AATTTACAGCCAAATAACAATCGCACAAATGCTGAAAGTCGAAGACGTCAACAAATTACATATTTCAAG  
 CTGTTGTTAAAGAAACAATGAGGATGAAACCAATTGCACCACTGGCGATTCTCATAAAGCTTGIAAA  
 GACACTTCATTGATGGGCAAGAAAGTTGATAAGGGAACCTAAAGTTATGTTAACATCATGCTTTACATC  
 ATACTGAAAAGGTTTGAAAGAACCTTACAAATTCATACCAGAGAGGTTTCTGCAGAAGCACGATAAG  
 GCGATGGAACAATCACTATTACCATTTAGTGCAGGTAGAGAATTTGGCAGGAATGGAATTAGGAAAAC  
 TTCAGTTTAGTTTTCTCTTCTGCTAATCTTGTTAATGCTTTTAAATGGTCTTGTTGTCTGATGGAGTG  
 CTTCTGATATGAGTGATTTACTGGGGTTGTTCTGTTATGAAAACCCCACTCGAAGCACGTATAGTTC  
 CTCGTTTGTAG

Figura 1c

ATGATGAACAAGTTATTATTTCTCCAACGGATTACTGATTCTCCTTCGACCACCATTATCAGTACTTT  
 TATTGTTACAATAATATCCATTGTTTTCTCTACACTGTCTTGTTGATAAGGACGACTAAGAATAAGC  
 AGAAGATAGCAGCACCAAAAGCATCGGGGGCGTGGCCGTTTCATAGGTCATCTCAAACCTATTCATGAAA  
 CAAGATACTCAGTTTTACAGAACTCTAGGAACCATGTCTGATAAATACGGGTCGGTGTTTCACACTTCG  
 ATTAGGAAACCAAGCAATCCTAGTTGTGAGCAACTGGGAGATGGTAAAAGAATGTTTCACAACAAACG  
 ACAAGTCATTCTCGAATCGTCCAAGTACGTTAAGCACTAAATACATGCTGAATGACACTAATTCTGTC  
 GTGTTTTACCTTACGGAACGTATTGGAGAGAAAATGCGGAAGATATTGGTGCAAAAACCTACTGATCTC  
 TAACCAAAGATCAGAGGCATTGAAAAATCTGAAAACGAAAGAAATCGACAACCTCGTTTGTAAGCTTA  
 ATGATTTATGCAACAACGATGTCAGTGGAGGAGGCACAAAAGTTAGGATGGACGAATGGTTGGCTGAC  
 ATGATGTTCAACATTATTGCTAGGATTACATTTGGTTACCAAAGCGGAGGAGGCGATGCACCTGGCGC  
 TTCTACAACATCCAAGAATGTCGAGAGATACAAGAAAACGTTGGACGAGATGTTTGTTGTTTTAGCGA  
 CGAGGTTTGACGTTTCAGATATATTTCCATCTCTGGAGTTTATAGACCGATTGAGAGGTCCTGTAAAG  
 GATATGAAAATCTTGGGAGACGAATTAAACTCCATTGCTGGATGTTTTATTGAAGAACATCGTCAAAA  
 GAGACGAGAATCATTATCCTCATTGTTATCTTTGTCAAATGAATCCGTTGGTGATGAACAAGATTTCA  
 TTGATGTTCTCTTGTCATAATGGATCAGTCACGGCTTCCCGGAGATGACCCAGATTTTATTATCAAA  
 ATTATGATCCTGGAAGCTTTTGACAGGTGGGACGGACAGTTTAAGTGCAACCTTAACCTGGGTCTCTC  
 TCTACTGCTGAACCAACCAACGTTAAAGAGGGCAAGGGAGGAAATAGATAGGCATGTGGAAAACG  
 GTAAGCAAGTGGAAGTGTCTGATATTCCGAAGCTCGGATACATTGATGCAATAATCAAAGAGACGATG  
 AGATTGTATCCAGTCGGAGCATTAAGCGAACGATACACGACTGAAGAATGCGAGGTTGGTCGGTTTAA  
 CGTACCCGCTGGCACACGCTTACTGGTGAATATATGGAAGATCCACAGAGACCAAGTGTGTGGGAGA  
 ATCCATCAGATTTTCAACCAGAGAGGTTTTTGTGCAGCGATAAGGTGGGTGTGGATTTATATGGCCAG  
 AATTATGAGCTGATACCATTTGGGGCCGGTAGGAGGGTATGTCCGGCTATAGTTTCATCACTGCAGAC  
 GATGCATTATGCGTTGGCGCGTCTTATTCAAGGATATGAAATGAAATCAGCCAGCCTCGATGGGAAGG  
 TGAATATGGAAGAAATGATAGCCATGTCGTGCCACAAGATGAGCCCTCTTGAAGTTATTATCAGTCCT  
 CGGGAGCCGAGGCGGAGTTAA

Figura 1d

ATGAACAAGTTATTATTTCTCCAACGGATTACTGATTCTCCTTCGACCACCATTATCAGTACTTTTAT  
 TGTTACAATAATATCCATTGTTTTTCTCTACACTGTCTTGTTGATAAGGACGACTAAGAATAAGCAGA  
 AGATAGCAGCACCAAAAGCATCGGGGGCGTGGCCGTTCATAGGTCATCTCAAACCTATTCATGAAACAA  
 GATACTCAGTTTTACAGAACTCTAGGAACCATGTCTGATAAATACGGGTCGGTGTTCACACTTCGATT  
 AGGAAACCAAGCAATCCTAGTTGTGAGCAACTGGGAGATGGTAAAAGAATGTTTCACAACAAACGACA  
 AGTCATTCTCGAATCGTCCAAGTACGTTAAGCACTAAATACATGCTGAATGACACTAATTCTGTCTG  
 TTTTCACCTTACGGAACGTATTGGAGAGAAATGCGGAAGATATTGGTGCAAAAACCTACTGATCTCTAA  
 CCAAAGATCAGAGGCATTGAAAAATCTGAAAACGAAAGAAATCGACAACCTCGTTTTGTAAAGCTTAATG  
 ATTTATGCAACAACGATGTCAGTGGAGGAGGCACAAAAGTTAGGATGGACGAATGGTTGGCTGACATG  
 ATGTTCAACATTATTGCTAGGATTACATTTGGTTACCAAAGCGGAGGAGGCGATGCACCTGGCGCTTC  
 TACAACATCCAAGAATGTCGAGAGATACAAGAAAACGTTGGACGAGATGTTTGTGTTTTAGCGACGA  
 GGTTCAGTTTTAGATATATTTCCATCTCTGGAGTTTATAGACCGATTGAGAGGTCTTGTAAGGAT  
 ATGAAAATCTTGGGAGACGAATTAACTCCATTGCTGGATGTTTTATTGAAGAACATCGTCAAAAGAG  
 ACGAGAATCATTATCCTCATTGTTATCTTTGTCAAATGAATCCGTTGGTGATGAACAAGATTTTCATTG  
 ATGTTCTCTTGTCAATAATGGATCAGTCACGGCTTCCCGGAGATGACCCAGATTTTATTATCAAAATT  
 ATGATCCTGGAAGCTTTTGCAGGTGGGACGGACAGTTTAAAGTGCAACCTTAACCTGGGTCTCTCTCT  
 ACTGCTGAACCACCCAAACGTGTTAAAGAGGGCAAGGGAGGAAATAGATAGGCATGTGGAAAACGGTA  
 AGCAAGTGGAAGTGTCTGATATTCGAAGCTCGGATACATTGATGCAATAATCAAAGAGACGATGAGA  
 TTGTATCCAGTCGGAGCATTAAAGCGAACGATACACGACTGAAGAATGCGAGGTGGTCGGTTTAACGT  
 ACCCGCTGGCACACGCTTACTGGTGAATATATGGAAGATCCACAGAGACCCAAGTGTGTGGGAGAATC  
 CATCAGATTTTCAACCAGAGAGGTTTTTGTGCAGCGATAAGGTGGGTGTGGATTTATATGGCCAGAAT  
 TATGAGCTGATACATTTGGGGCCGGTAGGAGGGTATGTCCGGCTATAGTTTCATCACTGCAGACGAT  
 GCATTATGCGTTGGCGCGTCTTATTCAAGGATATGAAATGAAATCAGCCAGCCTCGATGGGAAGGTGA  
 ATATGGAAGAAATGATAGCCATGTCGTGCCACAAGATGAGCCCTCTTGAAGTTATTATCAGTCCTCGG  
 GAGCCGAGGCGGAGTTAA

Figura 2a

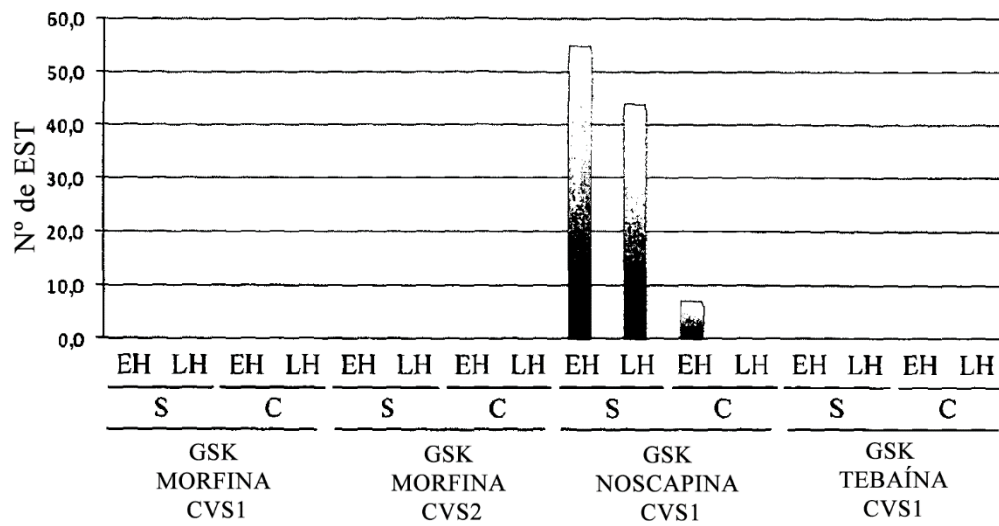


Figura 2b

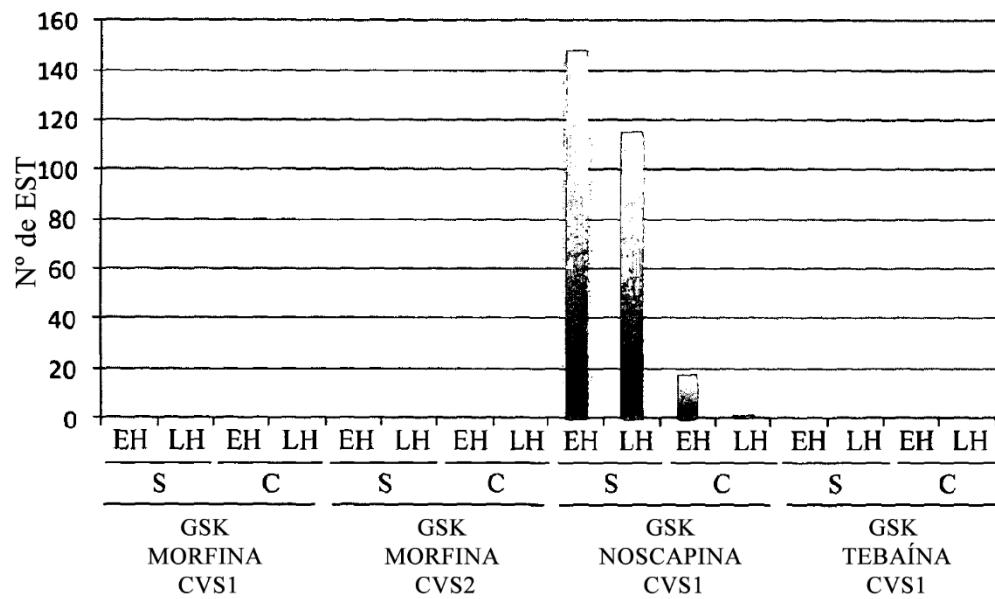


Figura 2c

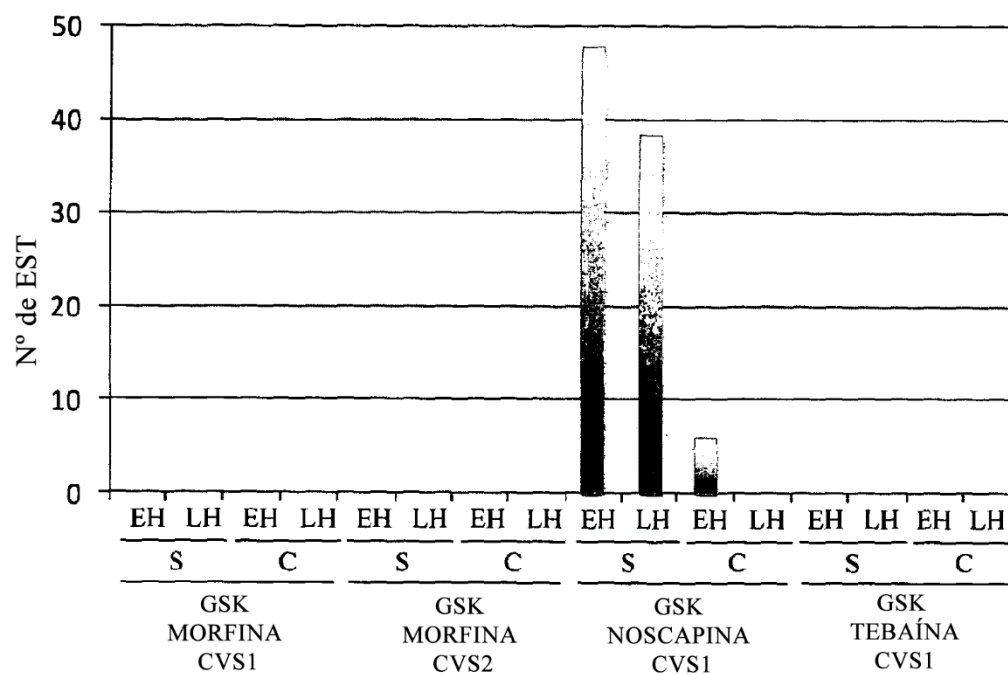


Figura 3a

Inicio	Fin	Característica
1	130	región 5' no traducida
131	800	exón 1
801	881	intrón 1
882	1216	exón 2
1059	1298	intrón 2
1299	1916	exón 3
1917	1921	región 3' no traducida

CTTGAGTCATGCCTTGATATGCTCATATTTTAGTTTGTCATATTCACTATAACTATAAATTTCAATAC  
AATTTCTAAAACATCATCATTCATCAAGAGAGATACAAATACCTTGATATCCTTTTATCATCAATGGAG  
TTATTCAATAAGTTACCATTATCCAACCAATTCCTTTTCAGTATTATTCTTGTACTACAGTTTCGAT  
TGTTCTATTATACAGTGTCTTCTTCTGGGTACTGATAAGAAAAAGAAGAGGAAGAAAGCACCAAATG  
CTGCAGGGGCATGGCCGTTAATAGGTCATCTCCGTCATTGATGAACGACAAGGAACCGTTGTATAGA  
GCACTAGGGAGCATGGCTGATAAGTACGGACCTGCATTCAACATCCGATTAGGTAACCAAGAAGTTCT  
TGTTGTGAGTAACTGGGAGATGGTAAACAGTGTTTTGGTAATCAAATGATAAGCTATTTTCGAATC  
GTCAAACTACATTAGCTGCAAATACATGCTTAATCAAACAACTTCTAGCGGATTGCGACCATATGGA  
CCATATTGGAGAGAGCTACGAAAGATAATGGTGCAGCAATTAATCTCTAAACAATCTTTAGAAATCGTG  
GAAACATCTGAAAATCAAAGAGATGGATGCTTCATTTAGTAAACTTAACGAGTTATGCAACAACAACG  
GTACTGGAACAGCTACCCTAATTAGGATGGACGAATGGTTTGCTGAGTTGACGTTCAACGTGATCGCA  
AGAAATGTCTTTGGCTACCAAAGTGGCGGAAGGTCGACAGCGCTTACGAACGTAATATGATCATACT  
CCCTCAATCTGTATCAATTTAAGGAAATCATTTTGGTCTTGTATTAACTGAATTTTCTATTAGGAG  
ATACGGAATCAAAGGGCGAGAGGTACAAGAAAACATTGGAAGAAGCACTTCATCTTATGTCAATTTTT  
GCAGTTTCAGACATATTTCCAAGTCTAGAGTGGGTAGATCGGTTAAGAGGCCTTATAAGGAATATGAA  
ACGCTTTGGAGATGAGCTAAATTCAATTGCAAGGTGTCTTATTGAAGAGCACCGCCAAAAGAGATTAC  
AATCCCTATCTAAAAGTGATAAAGGAGTTGGTGATGAACAAGACTTCGTTGATGTTCTCTTATCGGTT  
GCTGAAAAATCGCAACTTCTGGAGATGACCCTGATTTGGTCATCAAGTCTATGATTCTGGTTAGGCT  
ATTGATACCAAGTCTATTGCAATTTTGGTTTATGTGCTTGTCTAACTTCGTTTACTGCATATGGAT  
GTGCAGGAAATCGTATCAGGTGGGAGTGAGACCACATCGTCAACCTTAACCTGGGCCCTCTGTCTGTT  
ACTGAACCATCCGCATGTGTTAAGAAGGCAAAAGAGGAATTAGATACGCACGTAGGAAAAGATAGGC  
ATGTAGAAGAGTCAGATACCCCTAAGCTCGTGATCAATTAAAGCAATTATCAAAGAATCAATGCGATTG  
TATCCAAACGGGGCAATGCTTGATCGGTTGGCGTTAGAAGAGTGCGAAGTTGGTGGATTTTATGTACC  
GGCCGGGGGACGCTTATTTGTCAATGTTTGAAGATTTCAGAGAGATCCGAGTGTGTTGGGAGAATCCTC  
TGGAGTTTAAACCAGAGAGGTGGTTTTTGTAGTAATGGTGAAAAGATGGATGTGGATTACAAAGGTCAC  
AATCATGAATTCATACCATTTGGGATAGGTGCGGAGGATGTGCGCTGGTATGCTTTGGGCATCGGAGGT  
GATTCATTTGGTGCTGCCCCGTCTTATTCATGGGTTTGATATGAAAGCAGCAAGTGCCAATGGGAAAG  
TAGATATGGCAGAAATGGCAGGCATGGTGATTTGTTTTAAGAAGACACCTCTTGAAGTTATGGTCAAT  
CCTCGAGAGTAGATGTT

Figura 3b

Inicio	Fin	Característica
1	69	región 5' no traducida
70	1530	ORF
1531	1688	región 3' no traducida

GATGAAATTCCTTTATGCAAAGAGTCAATCTGACTCAAGCTAGCTAGAATATATACCAATCATAAAAGA  
AATGATCATGAGTAACCTTATGGATTCTTACGCTCATTTCTACCATATTAGCAGTCTTTGCTGCTGTGT  
TAATCATTTTCAGGAGAAGAATATCAGCATCCACAACGGAATGGCCTGTTGGCCCAAAACATTACCA  
ATCATAGGTAACCTTGCACATTCTTGGAGGCACTGCTCTCCATGTGCTCTTACATAAACTTGCTGAAGT  
TTACGGCAGTGTAATGACGATATGGATTGGTAGTTGGAAACCTGTTATTATTGTTTCCGACTTTGATC  
GAGCCTGGGAAGTTCTTGTAAACAAATCGTCAGATTATTAGCTCGTGAAATGCCTGAGATCACTAAA  
ATCGGCACTGCAAATGGAGAACAATTTCAAGTTCTGATTCTGGTCCGTTTGGGCCACTCTCGAAA

AGGTCTTCAGAGTGTAGCATTATCGCCTCAGCATTTAGCATCGCAAACCTGCACACCAAGAGAGAGATA  
 TAATAAAGTTGATCAAAAATTTGAAAGACGAAGCAGCTTCTGGAATGGTTAAACCACTTGATCATCTC  
 AAGAAAGCAACTGTAAGATTAATCAGTCGGTTAATCTATGGTCAGGATTTTGATGACGATAAGTATGT  
 TGAAGATATGCATGACGTGATCGAGTTTTTGTTCGTATTAGTGGTTATGCTCAACTTGCTGAGGTAT  
 TCTATTATGCTAAATATCTACCAGGTCATAAGAGAGCTGTAACCTGGCGCCGAAGAAGCAAAAAGAAGA  
 GTAATAGCTCTGGTGCCTCTTTCTTCAGTCAAACCTGCTACTAACACTTACTTGCATTTTCTCAA  
 ATCGCAACTGTATCCTGAAGAGGTTATCATATTGCTATATTTCGAAGCTTATCTTTTAGGTGTTGATA  
 GCACTTCTCAACCACTGCATGGGCACTCGCATTCTTAATACGCGAACCATCTGTTCAAGAGAACTT  
 TATCAAGAGCTTAAGAATTTACAGCCAATAACAATCGCACAATGCTGAAAGTCGAAGACGTCACAA  
 ATTACCATATTTACAAGCTGTTGTTAAAGAAACAATGAGGATGAAACCAATTGCACCACTGGCGATTC  
 CTCATAAAGCTTGTAAGACACTTCATTGATGGGCAAGAAAGTTGATAAGGGAACTAAAGTTATGGTT  
 AACATTGCTTTTACATCATACTGAAAGGTTTGGAAAGAACCTTACAAATTCATACCAGAGAGGTT  
 TCTGCAGAAGCAGGATAAGGCGATGGAACAATCACTATTACCATTTAGTGCAGGTATGAGAATTTGTG  
 CAGGAATGGAATTAGGAAACTTCAGTTTACTTTTTCTCTTGCTAATCTTGTTAATGCTTTTAAATGG  
 TCTTGTGTGCTGATGGAGTGCTTCTGATAGTGATTTACTGGGGTTTGTCTGTTTCATGAAAAC  
 CCCACTCGAAGCAGTATAGTTCTCGTTTGTAGTGATGGAATTTTATCTCATGTTGTTGTTTCTCT  
 TCATGTTTACTATTTCTGACTCGTTTGGTTTGGTGTAAAAAATAAGATCTAACTTCCAAATATCAT  
 TAATGTTTACACAAATCGAAATCAATCACTATGTTATGAAAATTAGTGTTCGCG

Figura 3c

Inicio	Fin	Característica
1	783	secuencia promotora
784	905/908	región 5' no traducida
906/909	1581	exón 1
1582	1694	intrón 1
1695	2050	exón 2
2051	2170	intrón 2
2171	2791	exón 3
2792	2918	región 3' no traducida

AAGTGTGCCACTAATCTACTGCTAGTGCTACTGCTCACTGACACTTACACATATGATTGATTTATGGC  
 TAAACAGGATGACCACTAAATTTATTTTGGAAAGCGGAGTGAATTAATTAAGTGGCACATTTTCATG  
 AGAATTATGATGGCATGCATTTAGATGAACAAGATACACCAATGTAGTGACTGAACAAGATGCTCG  
 ATCCTAACCCACCTGCAACTTTAGCTAACTTTAATAATTACATGTCTTATCTTTTATTGAATCAT  
 TTTATCTATCAATGGATGCTGATCAATAATATCATATATCTTTGCTTTTTCTTCAATCATTTAGATGA  
 ACAAAAACACAATAAGTGTAGTGGTTGTTCAATAACCCACCTTCAACTCATCTTCCCTTTAATAAC  
 AAATATCTTTGCTTTTTCTCCAATCATTTACTTGAACAACCAACACTAGTAAGTGTAGTGGTTTCTCA  
 TAACCCACCTGCAATTTTGGCTTACCTTTAATAACATATATCTTTGATTTTCTTCGATCATTTTAGC  
 TACCAATGGATGCTGATCCAAAAAGTTATGGCAAAAAGAGACAACGTGATCGAACACGAGCCTCTCGT  
 GCACCACAGCATCAAGGTTTGTGGAATTAACCGCTTGTAAAAAATGGAGTGCCTGATCATAATGAGG  
 TATTGCTAAGATATAGTATCAACTTTAGTGAAGTGGGCCAACAAAACCTCACGAGTTGTTGAAAATTGG  
 AGATTATATTTATAAGATAAAAGGGTCACTCCCTACACAACGACTTGCACTGCAAGTGAAAAAGAAAA  
 AAAACAACAACCTCAATCTAGCTAGAGTCGTGAAAAAGTTTGTGCGACTGTATTTAGTTAATTAT  
 AAAATTTCAATGAAGTCGTTAATGATGAACAAGTTATTATTTCTCCAACGGATTACTGATTCTCCTTC  
 GACCACCATATCAGTACTTTTATTGTTACAATAATATCCATTGTTTTCTTACACTGTCTTGTGTA  
 TAAGGACGACTAAGAATAAGCAGAAGATAGCAGCACCAAAAGCATCGGGGGCGTGGCCGTTATAGGT  
 CATCTCAAACATTTATGAAACAAGATACTCAGTTTTACAGAACTCTAGGAACCATGTCTGATAAATA  
 CGGGTCGGTGTTCACACTTCGATTAGGAAACCAAGCAATCCTAGTTGTGAGCAACTGGGAGATGGTAA  
 AAGAATGTTTACAACAAACGACAAGTCATTCTCGAATCGTCCAAGTACGTTAAGCACTAAATACATG  
 CTGAATGACACTAATTTCTGCTGTTTTTCACTTACGGAACGTATTGGAGAGAAATGCGGAAGATATT  
 GGTGCAAAAACACTGATCTCTAACCAGATCAGAGGCATTGAAAAATCTGAAAACGAAAGAAATCG  
 ACAACTCGTTTGTAAAGCTAATGATTTATGCAACAACGATGTCAGTGGAGGAGGCACAAAAGTTAGG



ATGGACGAATGGTTGGCTGACATGATGTTCAACATTATTGCTAGGATTACATTTGGTTACCAAAGCGG  
 AGGAGGCGATGCACCTGGTATGTGATCATCAAATTTTCGTTAAAACCAAATTAACCTTGTAATATATCT  
 TATGTTTACATGTTATATTGATCACTTTTGACACGTTCTGATCATTTTCACAAATCGAATTAGGCCGCTT  
 CTACAACATCCAAGAATGTCGAGAGATACAAGAAAACGTTGGACGAGATGTTTGTGTTTTAGCGACG  
 AGGTTTGCAGTTTCAGATATATTTCCATCTCTGGAGTTTATAGACCGATTGAGAGGTCTGTAAAGGA  
 TATGAAAATCTTGGGAGACGAATTAACTCCATTGCTGGATGTTTATTGAAGAACATCGTCAAAAGA  
 GACGAGAATCATTATCCTCATTGTTATCTTTGTCAAATGAATCCGTTGGTGATGAACAAGATTTTCATT  
 GATGTTCTCTTGTCAATAATGGATCAGTCACGGCTTCCCGGAGATGACCCAGATTTTATTATCAAAAT  
 TATGATCCTGGTAACATATATTACAACAGTATTTCTTTAAGTTATGGATTAATGGATGTCGTAACCAT  
 GAATAITTTTTCTGATCTGGATAAATGTAATCCGGAACCTAATATATGAATATTGTTGACGACGGAAGCT  
 TTTGCAGGTGGGACGGACAGTTTAAGTGCAACCTTAACCTGGGTCTCTCTCTACTGCTGAACCAACCC  
 AAACGTGTTAAAGAGGGCAAGGGAGGAAATAGATAGGCATGTGGAAAACGGTAAGCAAGTGGAAGTGT  
 CTGATATTCCGAAGCTCGGATACATTGATGCAATAATCAAAGAGACGATGAGATTGTATCCAGTCCGA  
 GCATTAAGCGAACGATACACGACTGAAGAATGCGAGGTTGGTCGGTTTAACGTACCCGCTGGCACACG  
 CTTACTGGTGAATATATGGAAGATCCACAGAGACCCAAGTGTGTGGGAGAATCCATCAGATTTTCAAC  
 CAGAGAGTTTTTTGTGCAGCGATAAGGTGGGTGTGGATTTATATGGCCAGAATTATGAGCTGATACCA  
 TTTGGGGCCGGTAGGAGGTATGTCCGGCTATAGTTTCATCACTGCAGACGATGCATTATGCGTTGGC  
 GCGTCTTATTCAAGGATATGAAATGAAATCAGCCAGCCTCGATGGGAAGGTGAATATGGAAGAAATGA  
 TAGCCATGTCGTGCCACAAGATGAGCCCTCTTGAAGTTATTATCAGTCCTCGGGAGCCGAGGCGGAGT  
 TAAATCTTATGTTCCAATTTTACATTAGCATCTTTGATTATGAAATGTATTGCTCTTAAGTTTCTTTT  
 TTGTTTTTTATATTTTTTAAGCTTGTATGTGATCATCAGCGAAAATGATGATGACAGAATCGT

Figura 4a

MELFIKLPFIQPIPFISIILVTTVSIVLLYSVFFWVTDKKKKRKKAPNAAGAWPLIGHLRLLMNDKEPL  
 YRALGSMADKYGPAFNIRLGNQEVLVVSNWEMVKQCFGNQNDKLFNQRQTTLAAKYMLNQTTSSGFAP  
 YGPYWRELKIMVQQLLSKQSLKSWKHLKIKEMDASFSLNELCENNNGTGATLIRMDWFAELTFNV  
 IARNVFGYQSGGRSTALTNGDTESKGERYKKTLEELHLMSIFAVSDIFPSLEWVDRLRGLIRNMKRF  
 GDELNSIAGCLIEHRQKRLQSVSKSDKGVGDEQDFVDVLLSVAEKSQLPGDDPDLVIKSMILEIVSG  
 GSETTSSTLTWALCLLNHPHVLKKAKEELDTHVGKDRHVEESDTPKLVYINAI IKESMRLYPNGAML  
 DRLALEECEVGGFHVPAGGRLFVNVWKIQRDPSVWENPLEFKPERWFLSNGEKMDVDYKGNHEFIPF  
 GIGRRMCAGMLWASEVIHLVLPRLIHGFDMKAASANGKVDMAEMAGMVICFKKTPLEVMVNPRE.

Figura 4b

MIMSNLWILTILISTILAVFAAVLIIFRRRISASTTEWPVGPKTLPIIGNLHILGGTALHVVHLKLAEV  
 YGSVMTIIGSWKPVIIVSDFDRAWEVLVNKSSDYSAEMPEITKIGTANWRTISSSDSGPFWATLRK  
 GLQSVALSPOHLASQTAHQERDI IKLIKNLKDEAASGMVKPLDHLKKATVRLISRLIYGQDFDDDKYV  
 EDMHDVIEFLIRISGYAQLAEVFYAKYLPGHKRAVTGAEAKRRVIALVRPFLQSNPATNTYLHFLK  
 SLYPEEVII FAIFEAYLLGVDSTSSTTAWALAFIREPSVQEKLYQELKNFTANNNRMLKVEDVNK  
 LPYLQAVVKETMRMKPIAPLAI PHKACKDTSIMGKKVDKGTKVMVNIHALHHTKVKWKEPYKFIPERF  
 LQKHDKAMEQSLLPFSAGMRICAGMELGKLQFSFSIANLVNAFKWSCVSDGVL PDMSDLLGFVLFMKT  
 PLEARIVPRL.

Figura 4c

MMNKLLFLQRITDSPSTTIISTFIVTIISIVFLYTVLLIRTTKNKQKIAAPKASGAWPFIGHLKLFMK  
QDTQFYRTLGTMSDKYGSVFTLRLGNQAILVVSNWEMVKECFTTNDKSFSNRPSTLSTKYMLNDTNSV  
VFSPYGTYWREMRKILVQKLLISNQRSEALKNLKTKKEIDNSFVKLNDLCNNDVSGGGTKVRMDEWLAD  
MMFNIIARITFGYQSGGGDAPGASTTSKNVERYKKTLD MFVVLATRFVSDIFPSLEFIDRLRGLVK  
DMKILGDELNSIAGCFIEEHRQKRRESLSSLLSLSNESVGDEQDFIDVLLSIMDQSRPGLDPPDFI IK  
IMILEAFAGGTDSL SATLTWVLSLLL NHPNVLKRAREEIDRHVENGKQVEVSDIPKLG YIDAI IKETM  
RLYPVGALSERYTTEECEVGRFNPAGTRLLVNIWKIHRDPSVWENPSDFQPERFLCSDKVGVDLYGQ  
NYELIPFGAGRRVCPAIVSSLQTMHYALARLIQGYEMKSASLDGKVNMEEMIAM SCHKMSPLEVIISP  
REPRRS.

Figura 4d

MNKLLFLQRITDSPSTTIISTFIVTIISIVFLYTVLLIRTTKNKQKIAAPKASGAWPFIGHLKLFMKQ  
DTQFYRTLGTMSDKYGSVFTLRLGNQAILVVSNWEMVKECFTTNDKSFSNRPSTLSTKYMLNDTNSV  
FSPYGTYWREMRKILVQKLLISNQRSEALKNLKTKKEIDNSFVKLNDLCNNDVSGGGTKVRMDEWLAD  
MFNIIARITFGYQSGGGDAPGASTTSKNVERYKKTLD MFVVLATRFVSDIFPSLEFIDRLRGLVKD  
MKILGDELNSIAGCFIEEHRQKRRESLSSLLSLSNESVGDEQDFIDVLLSIMDQSRPGLDPPDFI IKI  
MILEAFAGGTDSL SATLTWVLSLLL NHPNVLKRAREEIDRHVENGKQVEVSDIPKLG YIDAI IKETMR  
LYPVGALSERYTTEECEVGRFNPAGTRLLVNIWKIHRDPSVWENPSDFQPERFLCSDKVGVDLYGQN  
YELIPFGAGRRVCPAIVSSLQTMHYALARLIQGYEMKSASLDGKVNMEEMIAM SCHKMSPLEVIISPR  
EPRRS

Figura 5

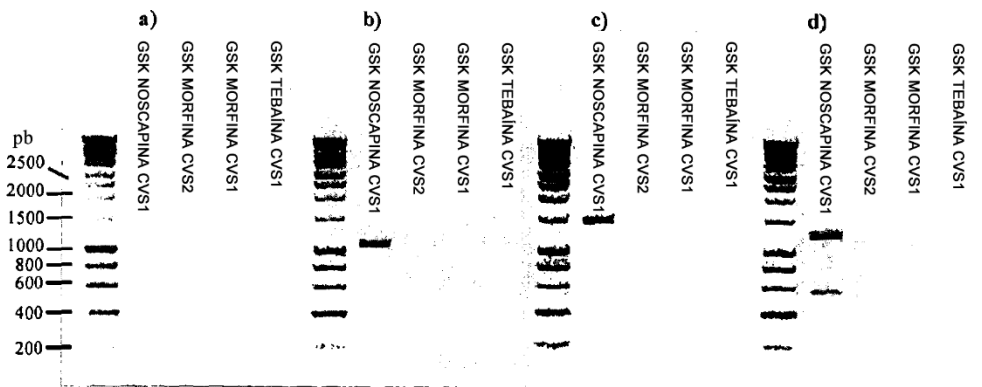


Figura 6

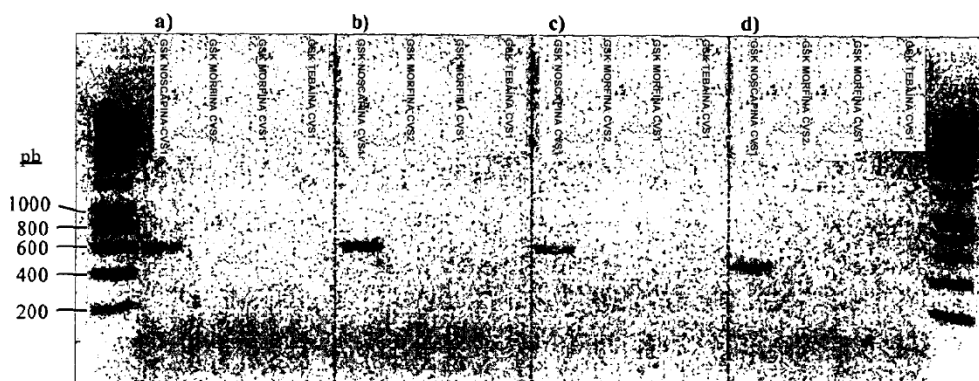


Figura 7

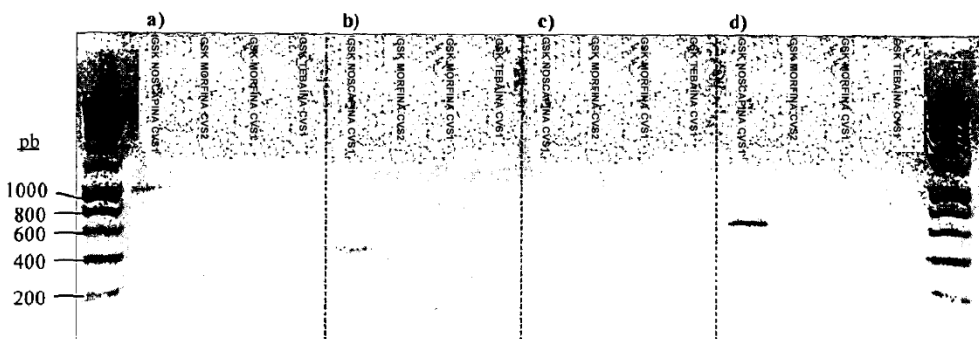


Figura 8a

GSK NOSCAPINA CVS1 x GSK TEBAÍNA CVS1				
SUM	Noscapina +		Noscapina -	
	61		214	
275	<i>PSCYP1</i> +	<i>PSCYP1</i> -	<i>PSCYP1</i> +	<i>PSCYP1</i> -
	61	0	129	85
SUM	<i>PSCYP1</i> +		<i>PSCYP1</i> -	
	190		85	
275	Noscapina +	Noscapina -	Noscapina +	Noscapina -
	61	129	0	85

Figura 8b

GSK NOSCAPINA CVS1 x GSK TEBAÍNA CVS1				
SUM	Noscapina +		Noscapina -	
	61		214	
275	<i>PSCYP2</i> +	<i>PSCYP2</i> -	<i>PSCYP2</i> +	<i>PSCYP2</i> -
	61	0	129	85
SUM	<i>PSCYP2</i> +		<i>PSCYP2</i> -	
	190		85	
275	Noscapina +	Noscapina -	Noscapina +	Noscapina -
	61	129	0	85

Figura 8c

GSK NOSCAPINA CVS1 x GSK TEBAÍNA CVS1				
SUM	Noscapina +		Noscapina -	
	59		200	
259	<i>PSCYP3</i> +	<i>PSCYP3</i> -	<i>PSCYP3</i> +	<i>PSCYP3</i> -
	59	0	122	78
SUM	<i>PSCYP3</i> +		<i>PSCYP3</i> -	
	181		78	
259	Noscapina +	Noscapina -	Noscapina +	Noscapina -
	59	122	0	78

el ensayo de genotipado de *PSCYP3* falló en 16 muestras, éstas se excluyeron del análisis

Figura 9

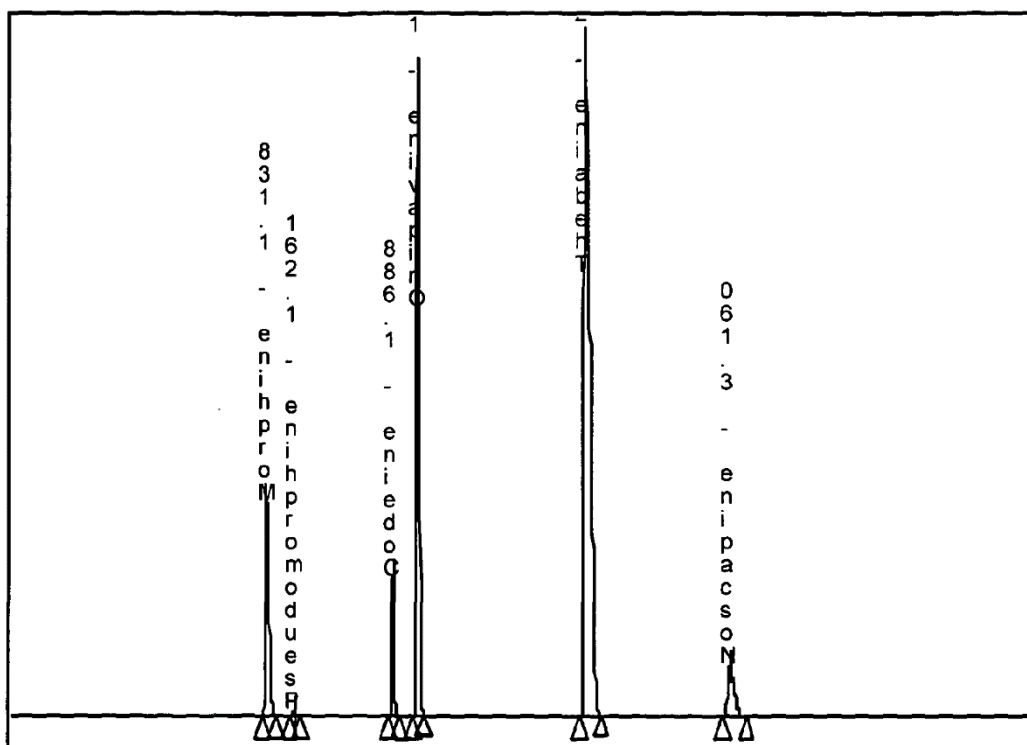


Figura 10

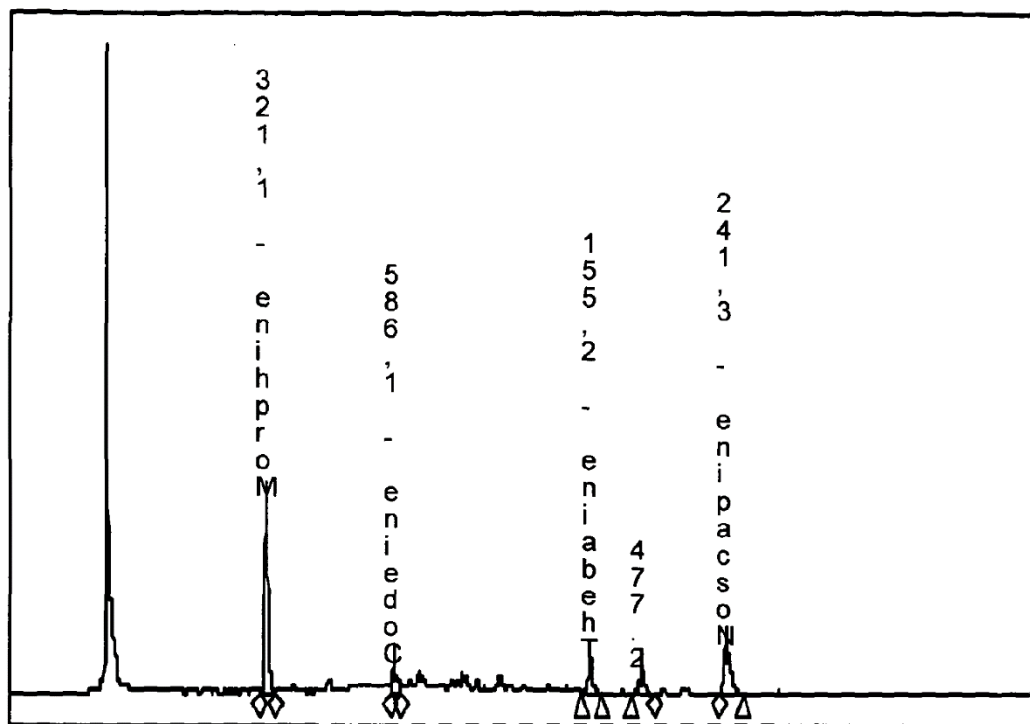


Figura 11

GAGGTGTTCAATTGCCATGTCAAAGGCATTAAACTTCATAAACCCAGATGAGCTTTCGATGCAGTGTCAT  
 TTTGATAGCTTTGAACCGTTTCCTTCAGGAAAAGCATGGTTCCAAGATGGCCTTTTTAGATGGTAATC  
 CTCCCGAGAGACTTTGCAAGCCGGTCGTGGATCATATAGAGTCACTTGGCGGTGAAGTCCGTCTCAAT  
 TCCAGGATTAAAAAGATTGAGCTTAAAAAAGATGGTACTGTGAAACGTCTAATGCTCACCAACGGTGA  
 TGCAATAGAAGGAGATGCTTATGTCATTGCAACCCAGTGGACATCCTAAAGCTGCTTATACCCGAGG  
 AGTGGAAAGAAGTTGGGTACTTTAAAAGATTGGATAAATTAGTTGGAGTTCCTGTGATTAAACGTCCAT  
 ATATGGTTTTGACAGGAAATTGAAAAATACATATGATCATCTTCTCTTCAGCAGAAGTCCCTCTTAAG  
CGTATACGCTGACATGTCAGTGACATGCAAGGAATATTATGACCCAAACAAATCCATGCTTGAGTTGG  
TATTTGCACCCGCTGAGGAATGGATCTCGCGCAGTGACTCTGAAATTATTGAAGCTACTATGCAGGAG  
 CTTGCGAAAC

Figura 12

ATGATCATGAGTAACTTATGGATTCTTACGCTCATTTCTACCATATTAGCAGTCTTTGCTGCTGTGTT  
 AATCATTTTCAGGAGAAGAATATCAGCATCCACAACGGAATGGCCTGTTGG

Figura 13

TAGGAGGGTATGTCCGGCTATAGTTTCATCACTGCAGACGATGCATTATGCGTTGGCGCGTCTTATTC  
 AAGGATATGAAATGAAATCAGCCAGCCTCGATGGGAAGGTGAATATGGAAGAAATGATAGCCATGTCTG  
 TGCCACAAGATGAGCCCTCTTGAAGTTATTATCAGTCCTCGGGAGCCGAGGCGGAGTTAA

Figura 14

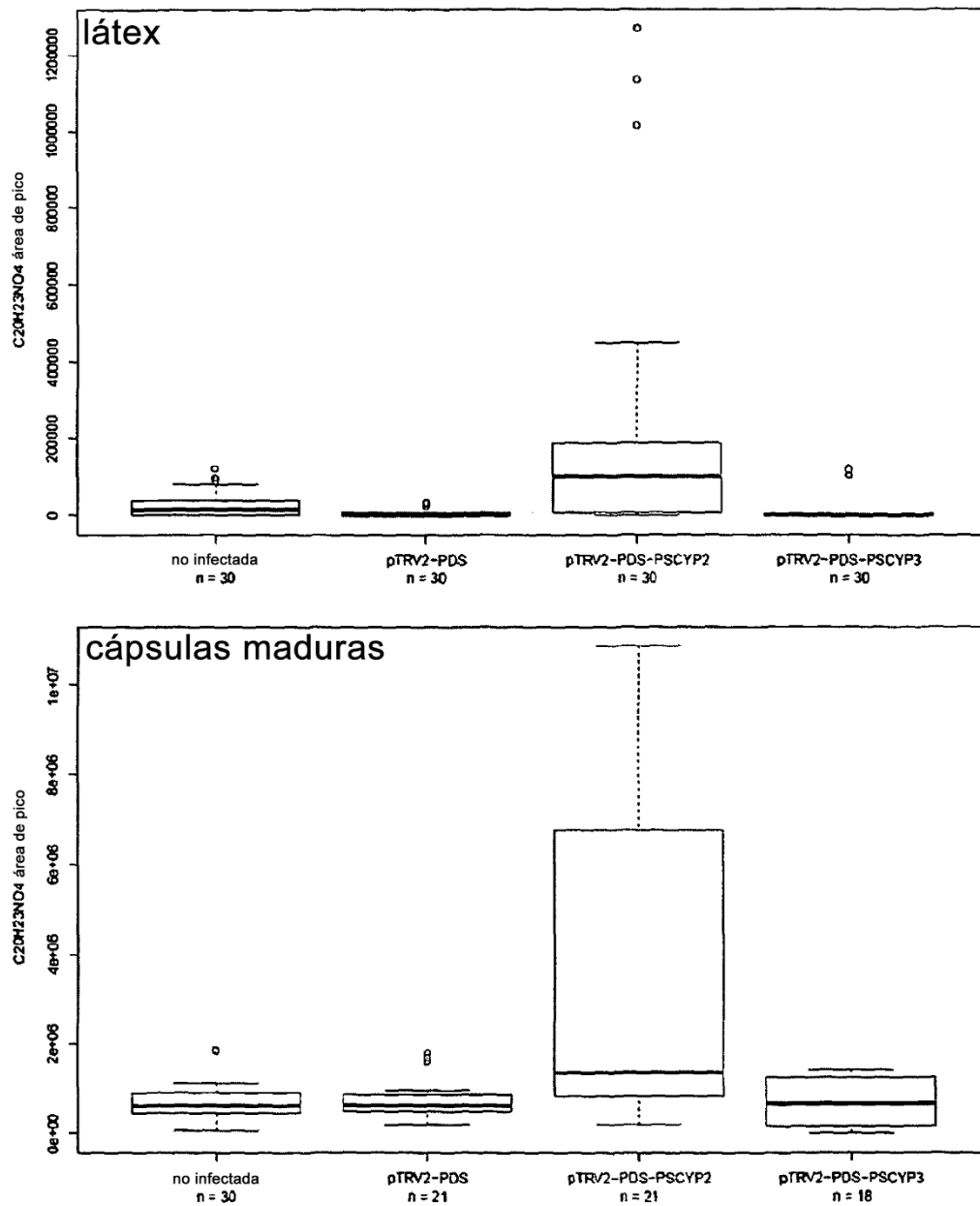




Figura 15

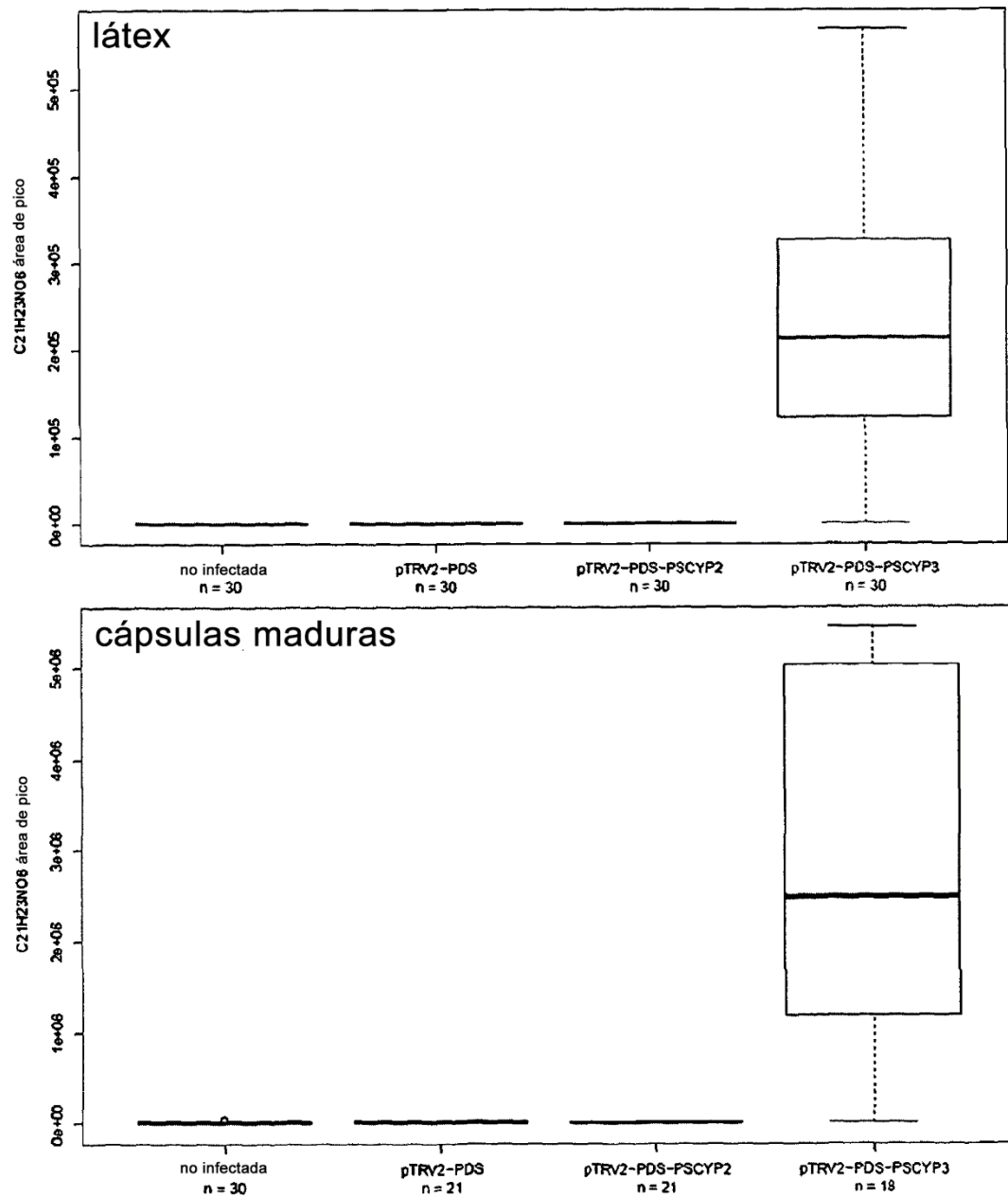
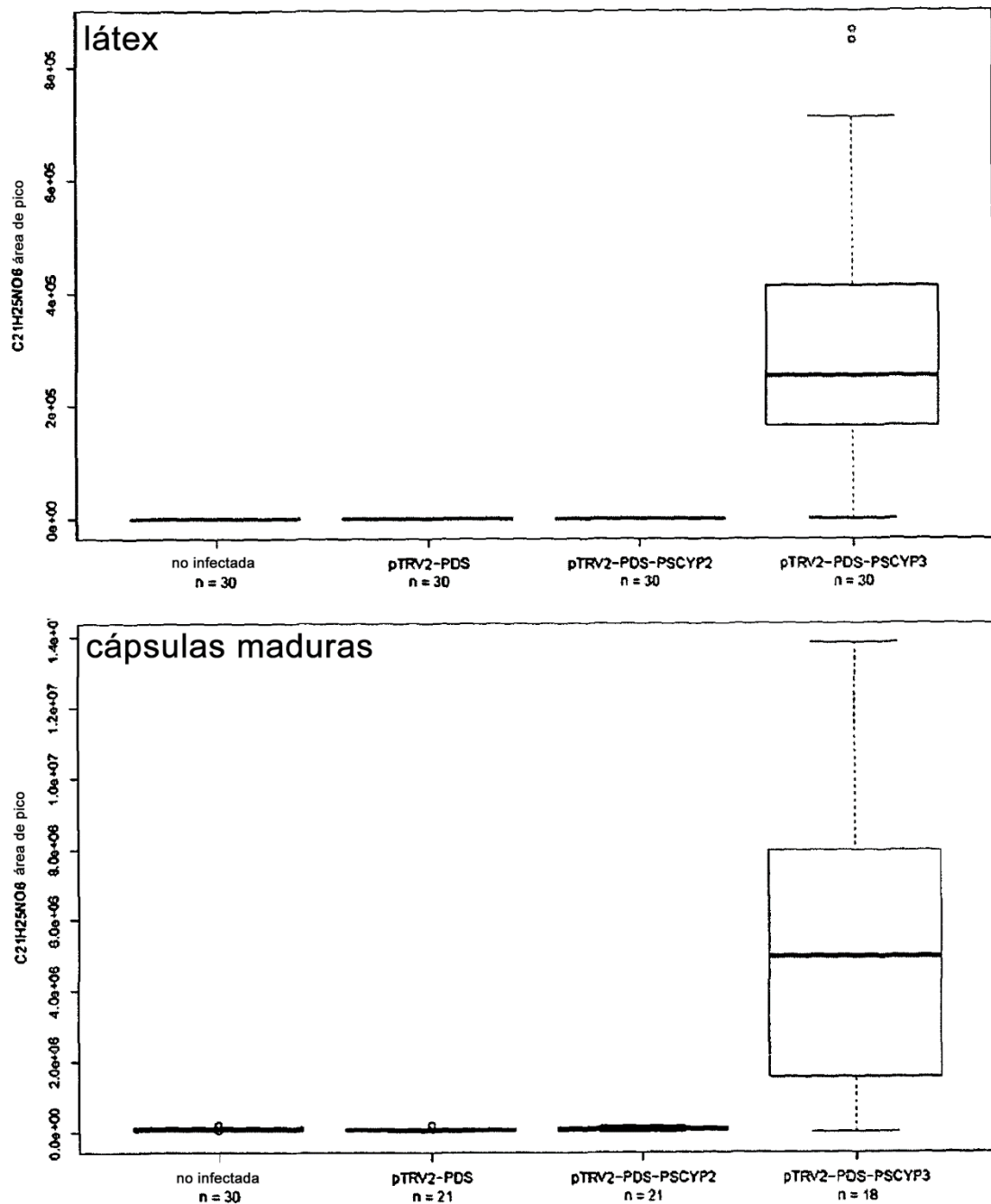


Figura 16



**Tabla 1: Secuencias de cebadores directos e inversos usados para amplificar los genes del citocromo P450 a partir de ADN genómico o ADNc**

gen del citocromo P450	Nombre del cebador	Secuencias de oligonucleótidos (5'-a 3'-)
<i>PSCYP1</i>	PSCYP1_F1	CTTGAGTCATGCCTTGATATGC
	PSCYP1_F2	TTGATGAACGACAAGGAACCG
	PSCYP1_F3	GCTACGAAAGATAATGGTGCAGC
	PSCYP1_F4	TCGACAGCGCTTACGAACG
	PSCYP1_F8	GAACCATTAAACACTTGAGTCATGC
	PSCYP1_LA_R1	GCATTTGGTGCTTTCTTCTCTCTTTTCTTATCAGTA
	PSCYP1_R1	AGCAAACCATTCGTCCATCC
	PSCYP1_R3	TGCAATTGAATTTAGCTCATCT
	PSCYP1_R5	ATTCATGATTGTGACCTTTGTAATCC
	PSCYP1_R7	TACGACAGGTTGCTAGCTTGG
<i>PSCYP2</i>	PSCYP2_F1	CAAAGAGTCAATCTGACTCAAGCTAGC
	PSCYP2_F2	TGAAATGCCTGAGATCACTAAATCG
	PSCYP2_F3	TCAAACCCTGCTACTAACACTTACTTGC
	PSCYP2_F4	TGTAAAGACACTTCATTGATGGGC
	PSCYP2_R1	GAGATGATCAAGTGGTTTAACCATTC
	PSCYP2_R2	CGAGTGCCCATGCAGTGG
	PSCYP2_R3	CACTCCATCAGACACACAAGACC
	PSCYP2_R4	GTAACATTAATGATATTTGGAAGTTTAGATC
	PSCYP2_R5	TTCGATTTGTGTAACATTAATGATATTTGG
<i>PSCYP3</i>	PSCYP3_F1	GTTATCTTTGTCAAATGAATCCGTTGG
	PSCYP3_F2	AATAATGGATCAGTCACGGCTTCC
	PSCYP3_F3	ATGTGGAACGTAAGCAAGTGG
	PSCYP3_F4	AATCCATCAGATTTTCAACCAGAGAGG
	PSCYP3_R1	ACGATTCTGTCATCATCTTTTCGC
	PSCYP3_R2	AGTCGTGTATCGTTCGCTTAATGC
	PSCYP3_LA_F2	GGCTTCCCGGAGATGACCCAGATTTTAT
	PSCYP3_LA_F3	TTGTTATTTTCATGACTATTACCACCAGCTTCCTCTTA
	PSCYP3_LA_F4	AGTGGAGGAGGCACAAAAGTTAGGATGGAC
	PSCYP3_LA_F5	CCATGTCTGATAAATACGGGTCGGTGTC
	PSCYP3_LA_F6	TTGTTGATAAGGACGACTAAGAATAAGCAGAAGATA
	PSCYP3_LA_R1	CATGCCTATCTATTTCTCCCTTGCCCTC
	PSCYP3_LA_R2	TGTCAGCCAACCATTCGTCCATCCTAAC
	PSCYP3_LA_R3	TGTTGATCACGTTGTCTCTTTTGGCATAA
	PSCYP3_LA_R4	TAACAATAAAAGTACTGATAATGGTGGTGAAGGAGAA
	PSCYP3_LA_R5	ATAATGGTGGTGAAGGAGAATCAGTAATC

**Tabla 2: Combinaciones de cebadores usadas para amplificar y secuencia de Sanger de los genes del citocromo P450 a partir de ADN genómico**

gen del citocromo P450	Combinación de cebador	Temperatura de hibridación [°C]	Tiempo de extensión [s]	Cebadores de secuenciación usados para la secuenciación de Sanger del producto de PCR purificado
<i>PSCYP1</i>	PSCYP1_F8/R3	68,5	60	PSCYP1_F3, PSCYP1_F8, PSCYP1_R3
	PSCYP1_F2/R5	69,3	60	PSCYP1_F2, PSCYP1_F4, PSCYP1_F5, PSCYP1_R2, PSCYP1_R4, PSCYP1_R5
	PSCYP1_F4/R7	69,8	60	PSCYP1_F4, PSCYP1_F6, PSCYP1_R4, PSCYP1_R7
<i>PSCYP2</i>	PSCYP2_F1/R5	61,7	60	PSCYP2_F1, PSCYP2_F2, PSCYP2_F3, PSCYP2_F4, PSCYP2_R1, PSCYP2_R2, PSCYP2_R5
<i>PSCYP3</i>	PSCYP3_F2/R1	66	60	PSCYP3_F2, PSCYP3_F4, PSCYP3_R1, PSCYP3_R2
	PSCYP1_LA_R1/ PSCYP_LA_R1	See Long Amp PCR	See Long Amp PCR	PSCYP3_LA_F2, PSCYP3_LA_F3, PSCYP3_LA_F4, PSCYP3_LA_F5, PSCYP3_LA_F6, PSCYP3_LA_R1, PSCYP3_LA_R2, PSCYP3_LA_R3, PSCYP3_LA_R4, PSCYP3_LA_R5

**Tabla 3: Combinaciones de cebadores usadas para amplificar los genes del citocromo P450 a partir de ADN genómico**

gen del citocromo P450	Combinación de cebador	Temperatura de hibridación [°C]	Tiempo de extensión [s]	Tamaño esperado del fragmento [pb]	Figura
<i>PSCYP1</i>	PSCYP1_F1/R3	66	40	1051	Figura 5a
	PSCYP1_F8/R3	68,5	60	1064	Figura 5b
	PSCYP1_F2/R5	69,3	60	1400	Figura 5c
	PSCYP1_F4/R7	69,8	60	~1200	Figura 5d
<i>PSCYP2</i>	PSCYP2_F1/R1	61	60	596	Figura 6a
	PSCYP2_F2/R2	61	60	596	Figura 6b
	PSCYP2_F3/R3	61	60	603	Figura 6c
	PSCYP2_F4/R4	61	60	475	Figura 6d
<i>PSCYP3</i>	PSCYP3_F1/R1	66	60	994	Figura 7a
	PSCYP3_F2/R2	66	60	418	Figura 7b
	PSCYP3_F3/R2	66	60	122	Figura 7c
	PSCYP3_F3/R1	66	60	638	Figura 7d

**Tabla 4: Cebadores usados para amplificar las secuencias seleccionadas para el silenciamiento del gen inducido por virus**

Gen diana que se silencia	Nombre del cebador	Secuencias de oligonucleótidos (5'-a 3'-) (en mayúsculas: secuencia específica del gen; en minúsculas: secuencia añadida; subrayado: sitios de restricción)
<i>PS FITOENO DESATURASA</i>	ps_pds_F	GAGGTGTTTCATTGCCATGTCAA
	ps_pds_R4	GTTCGCAAGCTCCTGCATAGT
<i>PSCYP2</i>	VIGS_PSCYP2_F	aaactcgagaagcttATGATCATGAGTAACTTATGGA
	VIGS_PSCYP2_R	aaaggtaccCCAACAGGCCATTCCGTTG
<i>PSCYP3</i>	VIGS_PSCYP3_F	aaactcgagaagcttTAGGAGGGTATGTCCGGC
	VIGS_PSCYP3_R	aaaggtaccTTAACTCCGCCTCGGCTCC