

DESCRIPCIÓN

TÍTULO

GEN DE LA 2-HIDROXIFLAVANONA C-GLUCÓSIDO DESHIDRATASA Y SU USO

5

CAMPO DE LA INVENCION

[0001]

10 La presente invención se refiere a un polinucleótido que codifica una enzima con actividad de deshidratación para la 2-hidroxi flavanona C-glucósido, y a su uso.

ANTECEDENTES DE LA INVENCION

[0002]

15 Los flavonoides, como metabolitos secundarios típicos de las plantas, tienen como sustancia de partida la fenilalanina y se clasifican en numerosos grupos, como flavanonas, flavonas, isoflavonas, flavonoles y antocianinas, según la estructura del anillo C. Las enzimas implicadas en su biosíntesis, y los genes que las codifican, han sido ampliamente investigados. Tanto las isoflavonas como las flavonas se biosintetizan a partir de flavanonas. En el caso de una isoflavona, la isoflavona sintasa cataliza la hidroxilación en la posición 2 de una flavanona y la transferencia del anillo B (NPL 1), y posteriormente la 2-hidroxi-isoflavona deshidratasa actúa para producir una isoflavona (NPL 2). La isoflavona sintasa pertenece a la familia del citocromo P450, mientras que la 2-hidroxi-isoflavona deshidratasa pertenece a la familia de las carboxiesterasas. Las flavonas están presentes en numerosas plantas como agliconas, C-glucósidos y O-glucósidos (NPL 3). Por ejemplo, plantas de la familia Caryophyllaceae, género *Dianthus*, como el clavel (*Dianthus caryophyllus*) y el nadeshiko (*Dianthus* sp.) producen flavonas O-glucósidos y flavonas C-glucósidos (NPL 4, NPL 5 y NPL 6). Las plantas de genciana, rábano picante, trigo sarraceno y Poaceae también producen flavonas C-glucósidos. Las funciones de las flavonas en las plantas incluyen la protección contra los rayos ultravioleta, la interacción con microorganismos, la protección contra el estrés biológico mediante la producción de fitoalexinas y el azulado de las antocianinas por acción copigmentaria. Por ejemplo, flavonas C-glucósidos como la isovitexina funcionan como copigmentos para los lirios y los recombinantes genéticos de claveles, produciendo el color azul de las flores (NPL 7 y NPL 8).

[0003]

35 Las flavonas se sintetizan a partir de flavanonas por dos vías (NPL 3). Las flavonas sintasas catalizan la reacción de conversión de flavanonas en flavonas. Para las flavonas sintasas, se

conocen genes que codifican la flavona sintasa I, perteneciente a la familia de las dioxigenasas (NPL 9), y la flavona sintasa II, perteneciente a la familia del citocromo P450 (NPL 10). Las flavonas producidas son metabolizadas en flavonas O-glucósidos o flavonas C-glucósidos por la glicosiltransferasa dependiente de UDP-glucosa. Por ejemplo, ya se han aislado el gen de la flavona O-glucosiltransferasa de *Nemophila* (NPL 11) y el gen de la flavona C-glucosiltransferasa (NPL 12, NPL 13) de la genciana y el rábano picante. Otra ruta es a través de la 2-hidroxi-flavanona como intermediario, donde la 2-hidroxi-flavanona se produce primero por la acción catalítica de la flavanona 2-hidroxilasa (NPL 14). Posteriormente, se produce una 2-hidroxi-flavanona C-glucósido por la acción de la glicosiltransferasa en el grupo hidroxilo en la posición 2. Este gen de la glicosiltransferasa ya se ha descrito (NPL 15). La 2-hidroxi-flavanona C-glucósido se convierte en una mezcla de flavonas 6-C-glucósidos y flavonas 8-C-glucósidos (típicamente isovitexina y vitexina) mediante una reacción de deshidratación espontánea *in vitro* (NPL 15, NPL 16).

[0004]

En el organismo, se cree que la reacción de deshidratación es catalizada por la 2-hidroxi-flavanona C-glucósido deshidratasa. Si bien se ha detectado dicha actividad enzimática en hojas de arroz (NPL 17), se desconoce la enzima en sí y su gen relacionado. Al igual que en la biosíntesis de isoflavonas descrita anteriormente, se ha sugerido que las enzimas pertenecientes a la familia de las carboxiesterasas podrían catalizar la misma reacción de deshidratación (NPL 18), pero no se ha observado actividad enzimática, por lo que esta posibilidad sigue siendo una conjetura. Además, la deshidratasa no se ha purificado ni se ha aislado su gen, y tampoco existen ejemplos de su uso industrial.

[LISTA DE CITAS]

[LITERATURA NO PATENTE]

[0005]

[NPL 1] *Plant Physiol.* 1999 121:821-828

[NPL 2] *Plant Physiol* 2005 137:882-91

[NPL 3] *Phytochemistry* 2005 66:2399-2407

[NPL 4] *Z. Naturforsch.* 63c, 161D168 (2008)

[NPL 5] *Chem. Pharm. Bull.* 59(9)1141-1148(2011)

[NPL 6] *Phytochemistry.* 2003 Mayo; 63(1):15-23

[NPL 7] *Plant Physiol Biochem.* 2013 72:116-24

[NPL 8] *Phytochemistry.* 2003; 63:15-23

[NPL 9] *Phytochemistry* 2001 58:43-6

[NPL 10] *Plant Cell Physiol.* 1999 40:1182-6

[NPL 11] *Plant Cell Physiol.* 2018 59:2075-2085

[NPL 12] FEBS Lett. 2015 589:182-7
 [NPL 13] Plant Cell Physiol. 2019 Dic 1; 60(12):2733-2743
 [NPL 14] FEBS Lett. 1998 Jul 17; 431(2):287-90
 [NPL 15] J Biol Chem. 2009 3; 284:17926-34
 5 [NPL 16] Metabolic Engineering 2013 16:11-20
 [NPL 17] J Biol Chem. 2009 3; 284:17926-34
 [NPL 18] Journal of Agricultural and Food Chemistry 2019 67(40), 11262-11276

RESÚMEN

[PROBLEMA TÉCNICO]

[0006]

Un objeto de la presente invención es proporcionar una planta transgénica que tenga una resistencia al estrés mejorada y/o un color de flor modificado, preferiblemente una modificación a un color de flor azul, o su progenie endogámica o exogámica, o sus propágulos, cuerpos
 15 vegetales parciales, tejidos o células, así como un método para producir de manera fácil y eficiente de flavona C-glucósido.

[SOLUCIÓN AL PROBLEMA]

[0007]

Los presentes inventores han descubierto la actividad de la 2-hidroxiflavanona C-glucósido
 20 deshidratasa en pétalos de plantas Caryophyllaceae y han logrado purificar la enzima de pétalos de plantas Caryophyllaceae. Gracias a una investigación exhaustiva, los presentes inventores han logrado obtener un gen que codifica la 2-hidroxiflavanona C-glucósido deshidratasa derivada de plantas Caryophyllaceae, y la invención se ha completado gracias a este hallazgo.

[0008]

Específicamente, la presente invención proporciona lo siguiente.

[1] Un polinucleótido que codifica una enzima con actividad de deshidratación para la 2-hidroxiflavanona C-glucósido, donde el polinucleótido se selecciona del grupo que consiste en los siguientes (A) a (E):

(A) polinucleótidos que consisten en las secuencias de nucleótidos de SEQ ID NO: 3, 5, 7 o
 30 9;

(B) polinucleótidos que se hibridan con polinucleótidos que consisten en secuencias de nucleótidos complementarias a las secuencias de nucleótidos de SEQ ID NO: 3, 5, 7 o 9 en condiciones rigurosas;

(C) polinucleótidos que codifican proteínas que consisten en las secuencias de aminoácidos
 35 de SEQ ID NO: 4, 6, 8 o 10;

(D) polinucleótidos que codifican proteínas que consisten en las secuencias de aminoácidos

de SEQ ID NO: 4, 6, 8 o 10 que tienen una delección, sustitución, inserción y/o adición de uno o más aminoácidos; y

(E) polinucleótidos que codifican proteínas que tienen secuencias de aminoácidos con al menos un 90% de identidad con las secuencias de aminoácidos de SEQ ID NO: 4, 6, 8 o 10.

5 [2] Un polinucleótido según [1], que es un polinucleótido que consiste en la secuencia de nucleótidos de SEQ ID NO: 3, 5, 7 o 9.

[3] Un polinucleótido según [1], que es un polinucleótido que codifica una proteína que consiste en la secuencia de aminoácidos de SEQ ID NO: 4, 6, 8 o 10.

[4] Una proteína codificada por un polinucleótido según cualquiera de [1] a [3].

10 [5] Un vector que comprende un polinucleótido según cualquiera de los puntos [1] a [3].

[6] Una planta transgénica, o su progenie endogámica o exogámica, que comprende un polinucleótido según cualquiera de los puntos [1] a [3].

[7] Propágulos, cuerpos vegetales parciales, tejido o células de una planta transgénica según [6] o su progenie endogámica o exogámica.

15 [8] Flores cortadas de una planta transgénica según [6], o su progenie endogámica o exogámica, o una forma procesada creada a partir de las flores cortadas.

[9] Un método para crear una planta transgénica,
el método que comprende una etapa de introducción en células vegetales de un polinucleótido según cualquiera de [1] a [3].

20 [10] Un huésped no humano que contiene un polinucleótido según cualquiera de [1] a [3].

[11] Un método para producir una enzima con actividad de deshidratación para la 2-hidroxiflavanona C-glucósido, comprendiendo el método:

un paso de cultivo de un huésped no humano según [10], y

una etapa de recolección del huésped no humano de una enzima con actividad de

25 deshidratación para la 2-hidroxiflavanona C-glucósido.

[12] Un método para producir flavona C-glucósido,

en donde el método incluye poner en contacto una enzima producida por el método de acuerdo con [11] con 2-hidroxiflavanona C-glucósido.

[EFECTOS VENTAJAS DE LA INVENCION]

30 **[0009]**

Según la invención, es posible proporcionar una planta transgénica con mayor resistencia al estrés y/o color de flor modificado, y modificación a un color de flor azul, o su progenie endogámica o exogámica, o sus propágulos, cuerpos vegetales parciales, tejidos o células. Según la invención, también es posible proporcionar un huésped no humano (por ejemplo, levadura o

35 bacteria) que exprese una enzima con actividad de deshidratación para la 2-hidroxiflavanona C-glucósido y utilizarla para producir flavona C-glucósido de forma fácil y eficiente.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0010]

La Figura 1 muestra la vía biosintética de flavona C-glucósido en plantas, a través de la 2-hidroxi flavanona como intermediario.

5 La Figura 2 muestra los resultados de cromatografía de líquidos de alta resolución de un extracto de solución de proteína cruda de *E. coli* que expresa Dca32587.1, derivado del clavel, y de una solución de reacción enzimática obtenida mediante la reacción enzimática con 2-hidroxinaringenina 6-C-glucósido.

10 La Fig. 3 muestra los resultados de cromatografía de líquidos de alta resolución de un extracto de solución de proteína cruda de *E. coli* que expresa DN21819, derivado de *Silene* y *Osteospermum* KAH9622207.1, y de una solución de reacción enzimática obtenida mediante la reacción enzimática con 2-hidroxinaringenina 6-C-glucósido. También se muestra una solución de proteína derivada de Dca32587.1, derivada de clavel, como control.

15 La Figura 4 muestra los resultados de cromatografía de líquidos de alta resolución para un extracto de solución de proteína cruda de *E. coli* que expresa ADNc de longitud completa para la deshidrogenasa derivada de una variedad nadeshiko (SEQ ID NO: 9), y para una solución de reacción enzimática obtenida por reacción enzimática en 2-hidroxinaringenina 6-C-glucósido.

[0011]

20 Las flores y hojas de las plantas contienen flavonas. Las flavonas son un tipo típico de metabolito secundario vegetal, comúnmente conocido como flavonoides, como cetonas cíclicas de derivados de flavanos que se encuentran principalmente como glucósidos en las plantas. La flavona, en sentido estricto, se refiere a la 2,3-didehidroflavan-4-ona, un compuesto con fórmula química $C_{15}H_{10}O_2$ y peso molecular 222,24. Sin embargo, en un sentido más amplio, las flavonas son una categoría de flavonoides. Un flavonoide se clasifica como "flavona" si su estructura

25 básica es flavónica y carece del grupo hidroxilo en la posición 3.

[0012]

En las plantas, las flavonas también se distribuyen como glucósidos, además de su forma libre, de modo que se producen principalmente flavonas O-glucósidos y flavonas C-glucósidos, pero se sabe que los flavonas C-glucósidos, en particular, exhiben una potente acción

30 copigmentaria, con un color azul exhibido por interacción intermolecular en la copresencia de antocianinas (un grupo de pigmentos abundante en plantas), y especialmente antocianinas de tipo delphinidina como delphinidina, malvidina y petunidina. La acción copigmentaria incluye no solo un efecto de profundidad de color que induce la producción de color azul, sino también un efecto de color profundo o un efecto de aumento de la estabilidad del color. Las flavonas tienen la

35 función de impartir a las plantas diversas resistencias al estrés biológico o no biológico, como las fitoalexinas, o contribuir al establecimiento de la interacción con microorganismos.

[0013]

En el presente documento, «flavona C-glucósido» se refiere a un glucósido de flavona en sentido amplio, es decir, un derivado que se incluye en la definición de flavonas, donde un aglicón está directamente unido al carbono anomérico de una aldosa. Las flavonas C-glucósidos incluyen, entre otros, luteolina C-glucósido, tricetina C-glucósido, apigenina C-glucósido y acetina C-glucósido. Las flavonas C-glucósidos también incluyen glucósidos de derivados de apigenina, luteolina, tricetina y acetina. Se conocen dos rutas para la biosíntesis de flavonas C-glucósidos en plantas. De estas, flavona 6-C-glucósido y flavona 8-C-glucósido se producen por la acción de la flavanona 2-hidroxilasa (F₂H), la flavona C-glicosilasa (2HCGT) y deshidratasa (2HDH) en una vía que utiliza la 2-hidroxiflavanona como intermediario (Fig. 1). Por lo tanto, se conjetura que, al incorporar el gen 2HDH, uno de los genes esenciales de esta vía, en una planta hospedadora, podría ser posible la acumulación de flavona C-glucósido en células vegetales.

[0014]

Los presentes inventores han logrado obtener un polinucleótido que codifica una enzima con actividad de deshidratación para 2-hidroxiflavanona C-glucósido a partir de pétalos de plantas Caryophyllaceae. La presente invención se refiere a un nuevo polinucleótido que codifica una enzima con actividad de deshidratación para 2-hidroxiflavanona C-glucósido.

[0015]

Las enzimas que tienen actividad de deshidratación para 2-hidroxiflavanona C-glucósido, que se utilizarán para la invención, se derivarán típicamente de plantas Caryophyllaceae, ejemplos de plantas Caryophyllaceae incluyen clavel (*Dianthus caryophyllus*), *Silene littorea*, *Heliosperma pusillum* y especies de jardinería nadeshiko (*Dianthus hybrida*), sin limitación a estas, sin embargo.

[0016]

Un polinucleótido de la invención se selecciona preferiblemente del grupo que consiste en los siguientes (A) a (E):

(A) polinucleótidos que consisten en las secuencias de nucleótidos de SEQ ID NO: 3, 5, 7 o 9;

(B) polinucleótidos que se hibridan con polinucleótidos que consisten en secuencias de nucleótidos complementarias a las secuencias de nucleótidos de SEQ ID NO: 3, 5, 7 o 9 en condiciones rigurosas;

(C) polinucleótidos que codifican proteínas que consisten en las secuencias de aminoácidos de SEQ ID NO: 4, 6, 8 o 10;

(D) polinucleótidos que codifican proteínas que consisten en las secuencias de aminoácidos de SEQ ID NO: 4, 6, 8 o 10 que tienen una delección, sustitución, inserción y/o adición de uno o más aminoácidos; y

(E) polinucleótidos que codifican proteínas que tienen secuencias de aminoácidos con al menos un 90% de identidad con las secuencias de aminoácidos de SEQ ID NO: 4, 6, 8 o 10.

[0017]

SEQ ID NO: 3 es la secuencia de nucleótidos de un gen que codifica una enzima (Dca32587.1) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada del clavel, SEQ ID NO: 5 es la secuencia de nucleótidos de un gen que codifica una enzima (DN21819) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada de *Silene littorea*, SEQ ID NO: 7 es una enzima (KAH9622207.1) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada de *Heliosperma pusillum*, y SEQ ID NO: 9 es la secuencia de nucleótidos de un gen que codifica una enzima que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada de la variedad nadeshiko Miite.

[0018]

El término “polinucleótido” como se utiliza en este documento se refiere a ADN o ARN.

También en el presente documento, el término "condiciones rigurosas" se refiere a las condiciones que permiten la unión específica entre un polinucleótido u oligonucleótido y el ADN genómico de forma selectiva y detectable. Las condiciones rigurosas se definen por una combinación adecuada de concentración de sal, disolvente orgánico (por ejemplo, formamida), temperatura y otras condiciones conocidas. Específicamente, la rigurosidad aumenta al reducir la concentración de sal, aumentar la concentración de disolvente orgánico o elevar la temperatura de hibridación. La rigurosidad también se ve afectada por las condiciones de enjuague tras la hibridación. Las condiciones de enjuague se definen por la concentración de sal y la temperatura, y la rigurosidad del enjuague aumenta al reducir la concentración de sal y elevar la temperatura. Por lo tanto, el término "condiciones rigurosas" se refiere a condiciones tales que la hibridación específica solo se produce entre secuencias de nucleótidos con alta identidad, como un grado de "identidad" entre las secuencias de nucleótidos de aproximadamente el 80 % o superior, preferiblemente de aproximadamente el 90 % o superior, más preferiblemente de aproximadamente el 95 % o superior, incluso más preferiblemente del 97 % o superior y, lo más preferiblemente, del 98 % o superior, en promedio. Las “condiciones rigurosas” pueden ser, por ejemplo, una temperatura de 60 °C a 68 °C, una concentración de sodio de 150 a 900 mM y preferiblemente de 600 a 900 mM, y un pH de 6 a 8, con ejemplos específicos que incluyen hibridación en condiciones de 5 x SSC (750 mM NaCl, 75 mM citrato trisódico), 1 % SDS, 5 x solución Denhardt, 50 % formaldehído, 42 °C y enjuague en condiciones de 0,1x SSC (15 mM NaCl, 1,5 mM citrato trisódico), 0,1 % SDS, 55 °C.

[0019]

La hibridación puede llevarse a cabo mediante un método conocido en la materia o un

método similar, como el descrito en Current Protocols in Molecular Biology (editado por Frederick M. Ausubel et al., 1987). Cuando se utilice una biblioteca comercial, la hibridación puede realizarse según el método descrito en las instrucciones de uso adjuntas. El gen seleccionado mediante hibridación puede ser de origen natural, ya sea vegetal o no. El gen
5 seleccionado mediante hibridación puede ser ADNc, ADN genómico o ADN sintetizado químicamente.

[0020]

SEQ ID NO: 4 es la secuencia de aminoácidos de una enzima (Dca32587.1) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada del clavel, SEQ ID
10 NO: 6 es la secuencia de aminoácidos de una enzima (DN21819) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada de *Silene littorea*, SEQ ID NO: 8 es una enzima (KAH9622207.1) que tiene actividad de deshidratación para 2-hidroxi flavanona C-glucósido derivada de *Heliosperma pusillum*, y SEQ ID NO: 10 es la secuencia de aminoácidos de una enzima que tiene actividad de deshidratación para 2-hidroxi flavanona C-
15 glucósido derivada de la variedad nadeshiko Miite.

[0021]

La frase "secuencia de aminoácidos con una delección, sustitución, inserción o adición de uno o más aminoácidos" se refiere a una secuencia de aminoácidos con una delección, sustitución, inserción o adición de 1 a 20, preferiblemente de 1 a 5 y, más preferiblemente, de 1 a 3
20 aminoácidos arbitrarios. La mutagénesis dirigida es un método útil de ingeniería genética, ya que permite la introducción de mutaciones específicas en sitios específicos, y puede llevarse a cabo mediante el método descrito en Molecular Cloning: A Laboratory Manual, 2.^a ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York, 1989. Al expresar el ADN mutante mediante un sistema de expresión adecuado, es posible obtener una proteína consistente en una
25 secuencia de aminoácidos con una delección, sustitución, inserción o adición de uno o más aminoácidos. Un polinucleótido puede obtenerse mediante un método que es públicamente conocido por aquellos expertos en la materia, tal como un método de síntesis química que utiliza el método de fosforamidita, o un método de amplificación de ácidos nucleicos que utiliza una muestra de ácido nucleico vegetal como plantilla y cebadores diseñados en base a la secuencia de
30 nucleótidos del gen objetivo.

[0022]

A lo largo de la presente descripción, el término "identidad" significa, para secuencias polipeptídicas (o secuencias de aminoácidos) o secuencias polinucleotídicas (o secuencias nucleotídicas), la cantidad (número) de residuos de aminoácidos o nucleótidos que las componen
35 que pueden determinarse como idénticos entre dos cadenas, en el sentido de acuerdo mutuo entre ellos, es decir, el grado de correlación de secuencia entre dos secuencias polipeptídicas o dos

secuencias polinucleotídicas, y esta "identidad" puede calcularse fácilmente. Se conocen numerosos métodos para medir la identidad entre dos secuencias de polinucleótidos o secuencias de polipéptidos, y el término "identidad" es bien conocido por los expertos en la materia (por ejemplo, véase Lesk, A.M. (Ed.), Computational Molecular Biology, Oxford University Press, Nueva York, (1988); Smith, D.W. (Ed.), Biocomputing: Informatics and Genome Projects, Academic Press, Nueva York, (1993); Griffin, A.M. y Griffin, H.G. (Ed.), Computer Analysis of Sequence Data: Part I, Human Press, Nueva Jersey, (1994); von Heinje, G., Sequence Analysis in Molecular Biology, Academic Press, Nueva York (1987); Gribskov, M. y Devereux, J. (Ed.), Sequence Analysis Primer, M-Stockton Press, Nueva York, (1991) y en otros lugares).

[0023]

Además, los valores numéricos de "identidad" utilizados en la presente descripción, a menos que se especifique lo contrario, pueden ser los calculados mediante un programa de búsqueda de identidad conocido por los expertos en la materia, pero preferiblemente se calculan mediante el programa ClustalW de MacVector Application (versión 9.5, Oxford Molecular Ltd., Oxford, Inglaterra). Según la invención, el grado de "identidad" entre secuencias de aminoácidos es, por ejemplo, aproximadamente del 65 % o superior, aproximadamente del 70 % o superior, aproximadamente del 75 % o superior, aproximadamente del 80 % o superior, aproximadamente del 85 % o superior, aproximadamente del 90 % o superior, preferiblemente aproximadamente del 95 % o superior, aún más preferiblemente aproximadamente del 97 % o superior, aún más preferiblemente aproximadamente del 98 % o superior y, lo más preferiblemente, aproximadamente del 99 % o superior.

[0024]

El polinucleótido (ácido nucleico, gen) de la invención "codifica" una proteína de interés. El término "codifica" también incluye la codificación de una secuencia estructural (exón), que es una sección continua de la proteína de interés, y la codificación mediante una secuencia intermedia (intrón). Asimismo, el término "codifica" incluye la codificación de una secuencia estructural (exón), que es una sección continua de la proteína de interés, y la codificación mediante una secuencia intermedia (intrón).

[0025]

Un gen con una secuencia de nucleótidos natural puede obtenerse, por ejemplo, mediante análisis con un secuenciador de ADN. El ADN que codifica una enzima con una secuencia de aminoácidos modificada puede sintetizarse mediante mutagénesis específica de sitio común o PCR, basándose en el ADN con la secuencia de nucleótidos natural. Por ejemplo, un fragmento de ADN que se va a modificar puede obtenerse mediante tratamiento con enzimas de restricción de ADNc natural o ADN genómico y utilizarse como plantilla para mutagénesis específica de sitio o PCR utilizando cebadores con la mutación deseada, para obtener un fragmento de ADN

con la modificación deseada. El fragmento de ADN con la mutación puede entonces unirse con un fragmento de ADN que codifica otra porción de la enzima objetivo.

Como alternativa, para obtener ADN que codifica una enzima que consiste en una secuencia de aminoácidos acortada, es posible cortar ADN que codifica una secuencia de aminoácidos más larga que la secuencia de aminoácidos objetivo, tal como la secuencia de aminoácidos de longitud completa, con una enzima de restricción deseada, y si el fragmento de ADN obtenido no codifica la secuencia de aminoácidos objetivo completa, entonces se puede sintetizar un fragmento de ADN que consiste en la secuencia de la porción faltante y vincularlo a éste.

[0026]

Al expresar el polinucleótido obtenido utilizando un sistema de expresión genética en *Escherichia coli* o levadura y medir la actividad enzimática, es posible confirmar que el polinucleótido obtenido codifica una proteína con la actividad deseada.

[0027]

La presente invención se refiere a un vector (recombinante), y en particular a un vector de expresión, que comprende el polinucleótido mencionado anteriormente. El vector de la invención también comprende una región de control de la expresión, como un promotor, un terminador y un origen de replicación, que dependen del tipo de planta hospedadora en la que se introduce. Entre los ejemplos de promotores que expresan constitutivamente polinucleótidos en células vegetales se incluyen el promotor 35S del virus del mosaico de la coliflor, el promotor El235S, que tiene dos regiones potenciadoras del promotor 35S unidas entre sí, y los promotores de los genes rd29A, rbcS y mac-1. Para la expresión génica específica de tejido, se puede utilizar un promotor para un gen expresado específicamente en ese tejido.

[0028]

El vector puede crearse mediante un método común que utiliza una enzima de restricción y una ligasa. La transformación de una planta hospedadora mediante el vector de expresión también puede llevarse a cabo mediante un método común.

[0029]

La invención se refiere además a una planta transgénica que contiene el polinucleótido o vector antes mencionado, o a su progenie endogámica o exogámica.

[0030]

Con el nivel actual de tecnología para la creación de plantas transgénicas, es posible utilizar técnicas para introducir un polinucleótido en una planta y expresarlo de forma constitutiva o tisular específica. La transferencia de ADN a la planta puede realizarse mediante un método conocido por los expertos en la materia, como el método de *Agrobacterium*, el método del vector binario, el método de electroporación, el método PEG o el método de la pistola de partículas.

[0031]

Las plantas que se utilizarán como hospedadoras para la invención no están particularmente restringidas y pueden ser plantas pertenecientes al género *Rosaceae Rosa*, *Solanaceae Petunia*, *Compositae Chrysanthemum*, *Caryophyllaceae Dianthus* o *Liliaceae Lilium*, entre las cuales se prefiere especialmente el cultivar de rosa de *Rosaceae Rosa* (nombre científico: *Rosa hybrida*). El término "planta de rosa", como se utiliza en este documento, es un cultivar de rosa de *Rosaceae Rosa* (nombre científico: *Rosa hybrida*), que es su clasificación taxonómica. Las rosas se clasifican principalmente como rosas Hybrid Tea, Floribunda y Polyantha según su forma de árbol y tamaño de flor, con el pigmento principal (antocianina) en los pétalos de todas las líneas siendo de dos tipos, el tipo cianidina y el tipo pelargonidina. El tipo de planta de rosa utilizada como hospedadora para la invención no está particularmente restringido, y cualquiera de estas variedades o líneas son adecuadas. Algunos ejemplos de variedades de rosas que se pueden utilizar como hospedadoras incluyen Ocean Song, Noblesse, Rita Perfumera, Cool Water, Fame, Topless y Peach Avalanche.

[0032]

Al introducir un polinucleótido de la invención en células vegetales, la flavona C-glucósido se acumula en las células vegetales y actúa como una fitoalexina, mejorando así la resistencia al estrés y exhibiendo una acción copigmentaria en copresencia de antocianinas de tipo delfinidina, y permitiendo la creación de una planta transgénica con un color de flor modificado al color de flor azul, o su progenie endogámica o exogámica.

[0033]

La invención también se refiere a flores cortadas de la planta transgénica obtenida o su progenie endogámica o exogámica, o a los propágulos, partes del cuerpo vegetal, tejido o células, o a una forma procesada creada a partir de las flores cortadas (especialmente flores cortadas procesadas). Las flores cortadas procesadas a las que se hace referencia aquí incluyen flores prensadas formadas a partir de flores cortadas, flores preservadas, flores secas o productos sellados con resina, sin limitarse a estos.

[0034]

La invención también se refiere a un huésped no humano que comprende el polinucleótido o vector mencionado anteriormente. El huésped no humano utilizado puede ser un procariota o un eucariota. Un procariota puede ser un huésped común, como una bacteria del género *Escherichia*, como *Escherichia coli*, o un microorganismo del género *Bacillus*, como *Bacillus subtilis*. Los eucariotas utilizados pueden ser organismos eucariotas inferiores, como microorganismos eucariotas, por ejemplo, hongos como levaduras u hongos filamentosos. Entre las levaduras se incluyen microorganismos del género *Saccharomyces*, como *Saccharomyces cerevisiae*, y entre los hongos filamentosos, microorganismos del género *Aspergillus*, como

Aspergillus oryzae, *Aspergillus niger* y microorganismos del género *Penicillium*. También pueden utilizarse como huésped no humano células animales o vegetales, pudiendo utilizarse como huésped células animales de líneas celulares de ratones, hámsteres, monos o seres humanos, y células de insectos, como células de gusanos de seda o los propios gusanos de seda.

[0035]

La invención se refiere además a un método para producir una enzima con actividad de deshidratación para 2-hidroxi flavanona C-glucósido. Este método comprende la incorporación de un polinucleótido de la invención en un huésped no humano, su cultivo y la extracción de dicho huésped de una enzima con actividad de deshidratación para 2-hidroxi flavanona C-glucósido. El método se refiere además a un método para producir flavona C-glucósido, así como a un método para poner en contacto la enzima producida mediante este método con 2-hidroxi flavanona C-glucósido. Este método permite una producción fácil y eficiente de flavonas C-glucósidos.

[0036]

La presente invención se explicará ahora con mayor detalle mediante ejemplos.

EJEMPLOS

[0037]

[Ejemplo 1: Detección de flavonas C-glucósidos en pétalos de nadeshiko]

Se ha informado previamente que los pétalos de nadeshiko contienen flavonas O-glucósidos y flavona C-glucósidos. Se realizó un análisis de pigmentos para confirmar si se detectó la flavona C-glucósido objetivo en los pétalos de las variedades comerciales nadeshiko (NP) y violeta grande (BNG) analizadas aquí.

[0038]

Tras congelar las muestras de pétalos de nadeshiko, se secaron durante la noche con un secador por congelación al vacío VirTis Sentry 2.0 (SP SCIENTIFIC) y se trituraron suavemente con una espátula. A la muestra triturada se le añadieron 4 ml de ácido trifluoroacético (TFA) al 0,1 % con acetonitrilo al 50 % por cada 0,5 g de peso fresco. Tras tratarla con ultrasonidos durante 20 minutos, se centrifugó a 3.600 rpm, 4 °C, durante 10 minutos y se recuperó el sobrenadante. El sobrenadante obtenido se filtró con un filtro de 0,45 µm (filtro COSMO NICE (acuoso), 0,45 µm, 13 mm). Se secó una porción de 200 µL y, tras la adición de β-glucosidasa y naringinasa, se trató durante la noche a 30 °C, y se añadieron 200 µL de acetonitrilo al 90 % con 0,1 % de TFA para suspender la reacción. A continuación, se trató durante 2 minutos con ultrasonidos, tras lo cual se centrifugó a 15.000 rpm, 4 °C, durante 5 minutos. El sobrenadante obtenido se filtró con un filtro de 0,45 µm (Milex-LH, 0,45 µm, Millipore) y se preparó para cromatografía de líquidos de alta resolución. Las condiciones de análisis fueron las siguientes.

< Condiciones de análisis >

Equipo: Sistema HPLC Prominence (producto de Shimadzu Corp.)

Detector: SPD-M20A (250 a 450 nm)

Columna: Shim-pack FC-ODS 150 x 4,6 mm, 3 µm (Shimadzu GLC)

5 Eluyente A: Solución acuosa de TFA al 0,1 %

Eluyente B: TFA al 0,1 % con acetonitrilo al 90 %

Velocidad de flujo: 0,6 ml/min

[0039]

10 La elución se realizó utilizando la solución A (solución acuosa de TFA al 0,1 %) y la solución B (solución acuosa de acetonitrilo al 90 % que contenía TFA al 0,1 %). La elución se realizó con un gradiente de concentración lineal de 20 minutos de una mezcla 9:1 a una mezcla 8:2, un gradiente de concentración lineal de 15 minutos de una mezcla 8:2 a una mezcla 2:8 y un gradiente de concentración lineal de 5 minutos de una mezcla 2:8 a una mezcla 0:10, seguido de una elución con una mezcla 0:10 durante 1 minuto. La velocidad de flujo fue de 0,6 ml/min.

15 Como resultado del análisis por el método descrito anteriormente, se detectaron isovitexina e isoorientina de NP, mientras que la isovitexina sola como flavona C-glucósido se detectó de BNG.

[Tabla 1]

Variedad	Flavona C-glucósido (contenido en mg por 1 g de peso fresco)	
	Isovitexina	Isoorientina
NP	0,961	0
BNG	0,097	0,874

20 **[0040]**

[Ejemplo 2: Medición de la actividad de la 2-hidroxi flavanona C-glucósido deshidratasa en el extracto enzimático crudo de pétalos de nadeshiko]

Se pulverizó aproximadamente 1 g de pétalos de nadeshiko (NP y BNG) en un mortero en presencia de nitrógeno líquido. Se añadieron 15 ml de tampón de extracción A (tampón Tris-HCl

25 0,1 M, pH 7,5; ditioneitol (DTT) 1 mM; cóctel de inhibidores de proteasas diluido al 1/100 (uso general, concentración 100 veces mayor), Nacalai Tesque, Inc.) y se extrajo la proteína en hielo. Se centrifugó a 10.000 rpm durante 10 minutos a 4 °C y se recogió el sobrenadante. Se midió el volumen del sobrenadante obtenido, se añadió sulfato de amonio hasta una saturación del 35 % y, a continuación, se agitó lentamente la mezcla a 4 °C durante 1,5 horas. Se centrifugó a 10.000

30 rpm durante 10 minutos a 4 °C y se recogió el sobrenadante. Se midió el volumen del sobrenadante obtenido, se añadió sulfato de amonio hasta saturarlo al 70 % y, a continuación, la mezcla se agitó lentamente a 4 °C durante 3 horas. Tras la separación por centrifugación a

10.000 rpm y 4 °C durante 10 minutos, se descartó el sobrenadante y el precipitado se suspendió en 400 µL de tampón de extracción B (tampón Tris-HCl 20 mM a pH 7,5, DTT 1 mM, cóctel de inhibidores de proteasas diluido al 1/100 (uso general, concentración 100 veces mayor)). La mezcla se centrifugó a 10.000 rpm, 4°C durante 10 minutos, el sobrenadante obtenido se desalinizó con columnas MicroSpin G-25 (Cytiva) según el método recomendado por el fabricante y se recuperó la fracción proteica.

[0041]

La reacción enzimática se llevó a cabo durante 30 minutos a 30 °C con una mezcla de reacción de 50 µl que comprendía 2-hidroxinaringenina 6-C-glucósido (que se puede preparar mediante el método descrito en Metabolic Engineering 2013 16:11-20, por ejemplo) a una concentración final de 200 µM, tampón de fosfato de potasio 100 mM pH 7,5, 10 % de sacarosa, 1 mM de DTT y 32,5 µl de una solución de enzima cruda. Tras detener la reacción mediante la adición de 50 µl de metanol al 90 %, la mezcla se centrifugó a 15.000 rpm, 4 °C, durante 5 minutos. El sobrenadante obtenido se filtró con un filtro de 0,45 µm (Millex-LH, 0,45 µm, Millipore) y se preparó para cromatografía de líquidos de alta resolución. Las condiciones de análisis fueron las siguientes.

< Condiciones de análisis >

Equipo: Cromatógrafo de líquidos de alta resolución LC-2030 (Shimadzu Corp.)

Columna: Shim-pack FC-ODS 150 x 4,6 mm, 3 µm (Shimadzu GLC)

Eluyente A: Solución acuosa de ácido fórmico al 0,1 %

Eluyente B: Metanol con ácido fórmico al 0,1 %

Velocidad de flujo: 1,0 ml/min

[0042]

Se utilizaron las soluciones A (solución acuosa de ácido fórmico al 0,1 %) y B (metanol con ácido fórmico al 0,1 %) para la elución. La elución se realizó con un gradiente de concentración lineal de 12 minutos, desde una mezcla 9:1 hasta una mezcla 4:6 de ambas, manteniéndose durante 8 minutos con la mezcla 4:6, seguido de un gradiente de concentración lineal de 1 minuto, desde una mezcla 4:6 hasta una mezcla 0:10 de ambas, y manteniéndose durante 1 minuto con la mezcla 0:10. La velocidad de flujo fue de 1 ml/min. En estas condiciones, el sustrato 2-hidroxinaringenina 6-C-glucósido eluye aproximadamente a los 8,7 minutos, la vitexina a los 10,6 minutos y la isovitexina a los 11,1 minutos. Como resultado del análisis mediante el método descrito anteriormente, se detectaron nuevos picos de compuestos de reacción en ambas soluciones de reacción de NP y BNG, que coincidían con el tiempo de retención y el espectro de absorción de la muestra de isovitexina. Esto demostró la presencia de una enzima en los pétalos de nadeshiko que cataliza la reacción de deshidratación de 2-hidroxinaringenina 6-C-glucósido a isovitexina.

[0043]

Ejemplo 3: Aislamiento de genes candidatos que codifican proteínas que catalizan la reacción de deshidratación de 2-hidroxi-flavanona C-glucósido.

5 <Aislamiento de ARN total>

Utilizando un kit RNeasy Plant Mini (Qiagen), se aisló el ARN total de pétalos de nadeshiko (variedad: Miite Raspberry Rose, Peach Latte) y hojas (variedad: Miite Raspberry Rose), según el protocolo recomendado por el fabricante. Los pétalos se clasificaron en una escala de 5 niveles según su crecimiento. La variedad Miite Peach Latte se utilizó solo en el

10 Nivel 1: Pétalos completamente rodeados por el cáliz

Nivel 2: Puntas de los pétalos ligeramente expuestas del cáliz

Nivel 3: Puntas de los pétalos expuestas a aproximadamente 1 mm del cáliz

Nivel 4: Aproximadamente 1/3 de los pétalos expuestos del cáliz

15 Nivel 5: Al menos la mitad de los pétalos expuestos del cáliz

[0044]

< Análisis de la expresión del ADNc derivado de nadeshiko >

Se preparó una biblioteca de ARN-Seq a partir del ARN total preparado como se describió anteriormente, con diferentes etiquetas respectivas, utilizando un kit de preparación de

20 bibliotecas de ARN específicas de cadena SureSelect (Agilent Technologies) de acuerdo con el protocolo recomendado por el fabricante. Se obtuvieron secuencias de extremos emparejados de 150 bases de longitud utilizando un secuenciador de nueva generación NextSeq (Illumina). Las lecturas de todas las muestras se combinaron y ensamblaron utilizando un Trinity v2.12.0, y se obtuvieron 415.080 secuencias contig. Las secuencias contig que no habían sido mapeadas por

25 RSEM 1.3.0 se eliminaron, y se obtuvieron 94.257 secuencias contig como productos de transcripción válidos. Cuando se encontraron variantes en cada gen, se seleccionó la secuencia más larga, y finalmente se obtuvieron 46.267 secuencias contig. Se mapearon los pares de lecturas para cada muestra, y se calculó el valor FPKM para determinar el nivel de expresión.

[0045]

30 < Estimación de la función genética >

Las funciones genéticas de las secuencias contig obtenidas se estimaron basándose en una búsqueda DIAMOND en la base de datos NCBI NR y una búsqueda BLAST en Araport11.

[0046]

< Adquisición de ADNc de longitud completa del gen candidato >

35 De las secuencias de contig obtenidas, se realizó una búsqueda BLAST utilizando la secuencia de aminoácidos del gen de la 2-hidroxi-isoflavanona deshidratasa derivada de

Glycyrrhiza echinata (GeHIDM, Accession NCBI: AB154414.1) o una búsqueda por palabra clave utilizando "Deshidratasa", y se seleccionaron 9 tipos diferentes de genes (DN226, DN601, DN1155, DN4559, DN5671, DN9033, DN1313, DN1658, DN1010) entre los genes candidatos obtenidos con altos niveles de expresión, expresados principalmente en los pétalos. Los clones de ADNc de longitud completa se obtuvieron mediante síntesis génica artificial basada en la secuencia de ADNc de longitud completa ensamblada. Dado que la secuencia de DN601 obtenida por ensamblaje era de longitud incompleta, se buscó un gen correspondiente al gen en una base de datos de clavel (<http://carnation.kazusa.or.jp/>), para obtener Dca57486.1.

[0047]

[Ejemplo 4: Medición de la actividad enzimática en *E. coli* mediante una proteína con actividad de deshidratación para 2-hidroxiflavanona C-glucósido]

<Construcción del vector de expresión de *E. coli*>

Los 9 genes candidatos descritos anteriormente se utilizaron como candidatos para preparar vectores de expresión de *E. coli* que contienen sus longitudes completas utilizando pET15b (Novagen), de acuerdo con el protocolo recomendado por el fabricante.

[0048]

<Expresión de *E. coli*>

Cada vector de expresión de *E. coli* preparado se introdujo en la cepa BL21(DE3) de *E. coli* (New England Biolabs Japan Inc.) según el protocolo recomendado por el fabricante para obtener transformantes de *E. coli*. La *E. coli* se cultivó utilizando un sistema de autoinducción Overnight Express 1 (Novagen), según el protocolo recomendado por el fabricante. Los transformantes de *E. coli* se cultivaron a 37 °C (aproximadamente 4 horas) con 2 ml de solución de cultivo preparada, hasta alcanzar un valor DO600 de 0,5. La solución de *E. coli* se añadió como solución de precultivo a 50 ml de solución de cultivo, y el cultivo principal se realizó durante la noche a 25 °C. La solución de *E. coli* que se había cultivado mediante cultivo principal durante la noche se centrifugó a 3000 rpm, 4 °C, 15 minutos, y las células recolectadas se suspendieron en un tampón sónico (tampón Tris-HCl 20 mM a pH 7,5, DTT 10 mM, fluoruro de amidinofenilmetanosulfonilo HCl (APMSF) 10 µM). Después de agregar 5 ml de tampón sónico a 1 g de *E. coli*, la *E. coli* suspendida se interrumpió mediante tratamiento ultrasónico y luego se centrifugó a 15.000 rpm, 4 °C, 10 minutos, y se recuperó el sobrenadante. El sobrenadante se usó como extracto de solución de proteína cruda de *E. coli* que expresa el gen candidato.

[0049]

< Medición de la actividad enzimática >

La reacción enzimática se llevó a cabo durante 30 minutos a 30 °C con una mezcla de

reacción de 50 µl que comprendía 2-hidroxinarigenina 6-C-glucósido a una concentración final de 40 µM, tampón de fosfato de potasio 100 mM pH 7,5, 10 % de sacarosa, 14 µM de β-mercaptoetanol y 33,6 µl de solución de enzima cruda. Tras detener la reacción mediante la adición de 50 µl de metanol al 90 %, la mezcla se centrifugó a 15.000 rpm, 4 °C durante 5 minutos, y el sobrenadante obtenido se filtró con un filtro de 0,45 µm (Milex-LH, 0,45 µm, Millipore), preparado para cromatografía de líquidos de alta resolución, y se analizó mediante el mismo método que en el Ejemplo 2. Como resultado, no se detectó ningún pico del compuesto de reacción en ninguna de las mezclas de reacción, ni se observó actividad de deshidratasa para 2-hidroxinarigenina 6-C-glucósido. Por lo tanto, no fue posible obtener genes objetivo basándose en la similitud con secuencias de aminoácidos descritas previamente ni en la estimación de la función génica.

[Tabla 2]

Gen candidato	Actividad para2HNCG
DN226	-
DN1155	-
DN4559	-
DN5671	-
DN9033	-
DN1313	-
DN1658	-
DN1010	-
Dca57486.1	-

[0050]

[Ejemplo 5: Medición de la actividad deshidratasa de 2-hidroxiflavanona C-glucósido]

La reacción enzimática se llevó a cabo durante 1 a 3 horas a 30 °C con una mezcla de reacción de 20 µl compuesta por 2-hidroxiflavanona 6-C-glucósido a una concentración final de 200 µM, tampón de fosfato de potasio 20 mM, pH 7,0, y de 1 a 10 µl de solución enzimática cruda. Tras añadir 20 µl de acetonitrilo para suspender la reacción, se añadieron 160 µl de acetonitrilo al 50 %, la mezcla se centrifugó a 10.000 rpm durante 15 minutos a 25 °C, se filtró el sobrenadante y, a continuación, el producto de reacción se detectó mediante HPLC (LC-2030C de Shimadzu Corp.) en las siguientes condiciones. Columna: Shim-pack FC-ODS de 4,6 x 150 mm, disolvente A: ácido fórmico al 0,1 %, disolvente B: metanol con ácido fórmico al 0,1 %. La fase móvil presenta un gradiente de concentración lineal desde A:B = 90:10 a los 0 minutos hasta A:B = 40:60 a los 12 minutos, un gradiente de concentración lineal de A:B = 40:60 hasta los 20

minutos y A:B = 0:100 durante 21 minutos, y tras mantener la temperatura durante 1 minuto, un gradiente de concentración lineal hasta A:B = 90:10. La velocidad de flujo fue de 1 ml/min y la temperatura de la columna, de 40 °C. El espectro de absorción se midió mediante la absorbancia a 280 nm y 330 nm utilizando un arreglo de fotodiodos (SPD-M20A). En estas condiciones, el sustrato de 2-hidroxi flavanona 6-C-glucósido eluye aproximadamente a los 8,6 minutos, la vitexina eluye aproximadamente a los 10,6 minutos y la isovitexina eluye aproximadamente a los 11,1 minutos.

[0051]

[Ejemplo 6: Construcción de una base de datos para genes expresados en pétalos]

La base de datos construida a partir de los pétalos de nadeshiko se muestra en el Ejemplo 3. Además, los datos de RNAseq (NCBI Accession: PRJNA299760 ID:299760) para los pétalos de *Silene littorea* de la familia Caryophyllaceae se utilizaron para el ensamblaje utilizando Trinity v2.12.0 para crear secuencias contig, para obtener una biblioteca de transcripción para *Silene littorea*.

[0052]

[Ejemplo 7: Purificación enzimática de la variedad nadeshiko NP]

Tras congelar 25 g de pétalos preaperturados de la variedad nadeshiko NP con nitrógeno líquido, se trituraron con un mortero. Se añadieron 200 ml de tampón de fosfato de potasio 20 mM a pH 7,0, cloruro de sodio 100 mM y DTT 1 mM, y se realizó una extracción durante 40 minutos con agitador, enfriando con hielo. El extracto se centrifugó con una centrífuga refrigerada a 4°C a 6000 rpm durante 15 minutos, y el sobrenadante se recogió y filtró con un Miracloth. Tras disolver el sulfato de amonio en el filtrado hasta una saturación del 35%, se centrifugó de la misma manera. A continuación, se disolvió el sulfato de amonio en el sobrenadante hasta una saturación del 65% y la mezcla se centrifugó de nuevo. El precipitado se disolvió en una cantidad mínima de tampón de fosfato de potasio 20 mM a pH 7,0 y se realizó diálisis durante 16 horas contra tampón de fosfato de potasio 20 mM a pH 7,0 que contenía DTT 0,1 mM. El precipitado se eliminó por separación centrífuga y se adsorbió en 10 ml de Q-Sepharose Fast Flow (Cytiva). Después de lavar la columna con 35 ml de tampón de fosfato de potasio 20 mM a pH 6,2, se lavó adicionalmente con tampón de fosfato de potasio 20 mM que contenía cloruro de sodio 50 mM a pH 7,0. A continuación, la proteína se eluyó con tampón de fosfato de potasio 20 mM que contenía cloruro de sodio 300 mM a pH 7,0, y se recuperó aproximadamente 1 ml de cada una. La actividad de cada fracción se midió como en el Ejemplo 5. Se recogieron las fracciones con fuerte actividad para la conversión del sustrato a isovitexina y se disolvió sulfato de amonio a 1 M.

[0053]

La solución se cargó en una columna con 10 ml de Butyl Sepharose Fast Flow (Cytiva), se

equilibró con tampón de fosfato de potasio 20 mM a pH 7,0 y sulfato de amonio 1 M, y se lavó con 30 ml del mismo tampón. A continuación, se eluyó con tampón de fosfato de potasio 20 mM a pH 7,0 y sulfato de amonio 0,6 M, y se recuperó aproximadamente 1 ml de cada una. La actividad de cada fracción se midió como se describió anteriormente, y las fracciones con mayor actividad de conversión del sustrato a isovitexina se recogieron y concentraron a aproximadamente 0,6 ml utilizando Amicon Ultra10K (Merck-Millipore). El concentrado se suministró para cromatografía de filtración en gel con SuperdexTM75 10/300GL (Cytiva). Los disolventes utilizados fueron fosfato de potasio 50 mM, pH 7,0 y cloruro de sodio 150 mM, con una velocidad de flujo de 0,8 ml/min y 1 ml/fracción (utilizando AKTA puro, Cytiva).

[0054]

La fracción del pico activo se diluyó 5 veces con agua y se cargó en 1 ml de Hi Trap DEAE FF (Cytiva), equilibrado con tampón de fosfato de potasio 20 mM a pH 7,0. Con un flujo de 1 ml/min, se realizó un lavado con tampón de fosfato de potasio 20 mM a pH 7,0 durante 10 minutos, seguido de una elución con un gradiente de concentración lineal de 30 minutos hasta 20 mM de tampón de fosfato de potasio a pH 7,0 + 300 mM de cloruro de sodio. Se recogieron 0,3 ml de cada muestra (utilizando AKTA puro).

[0055]

Las fracciones con actividad intensa se diluyeron tres veces con agua y se adsorbieron en 0,5 g de 40 µM hidroxiapatita cerámica CHT Tipo 1 (BioRad), equilibrada con fosfato de potasio 5 mM, pH 7,0. La mezcla se lavó con 5 ml del mismo tampón, se lavó con 5 ml de fosfato de potasio 20 mM, pH 7,0 y 5 ml de fosfato de potasio 40 mM, pH 7,0, y se eluyó con fosfato de potasio 120 mM, pH 7,0. Se recuperaron aproximadamente 400 µl de cada una. La fracción activa se diluyó cinco veces con agua y se adsorbió en 1 ml de Hi Trap Q FF (Cytiva). Utilizando una velocidad de flujo de 1 ml/min, se realizó un lavado con el mismo tampón durante 5 minutos a 0,25 ml/tubo, y seguido de una elución con un gradiente de concentración lineal de 10 minutos hasta 20 mM de fosfato de potasio, pH 7,0 + 1 M de cloruro de sodio (utilizando AKTA puro).

[0056]

Se midió la actividad de cada tubo y se realizó el siguiente análisis.

Para obtener fragmentos peptídicos de las proteínas de cada tubo, se realizó una digestión enzimática mediante el método de surfactante de transferencia de fase (PTS) (véase, por ejemplo, J Proteome Res. 2008 7:731-740). La solución de muestra de péptidos fragmentados se midió mediante espectrometría de masas utilizando un sistema nanoLC-MS con un cromatógrafo de nanoflujo UltiMate 3000 Nano (Thermo Fisher Scientific) y un espectrómetro de masas de alta precisión LTQ Orbitrap Fusion Lumos (Thermo Fisher Scientific) conectados en línea. Como condiciones para el nanoLC, la muestra de péptido se inyectó y se adsorbió en una columna de

fase reversa C18 a una velocidad de flujo de 10 µl/min, y se separó a una velocidad de flujo de 300 nl/min con un gradiente lineal usando el disolvente A (solución acuosa de acetonitrilo al 3%/ácido fórmico al 0,1%) y el disolvente B (solución acuosa de acetonitrilo al 90%/ácido fórmico al 0,1%). Luego, el péptido se pulverizó desde una fuente de iones de nanoelectrospray (Thermo Fisher Scientific) y el ion precursor generado se introdujo en un espectrómetro de masas para la medición por MS. El ion precursor se fragmentó por Disociación Inducida por Colisión (CID) para producir el ion producto, que se midió por MS/MS. El valor de energía de colisión normalizada para CID se estableció en 32. El espectro MS y el espectro MS/MS se obtuvieron con la configuración para el cambio automático de MS a MS/MS, en el modo de adquisición dependiente de los datos. El escaneo MS se obtuvo con Orbitrap a una resolución de media anchura de 300.000, y el escaneo MS/MS posterior se obtuvo con una trampa de iones. Cada medición se realizó en modo de control automático de ganancia (AGC), y el valor de AGC se ajustó a 1×10^6 para MS y 1×10^4 para MS/MS.

[0057]

Se utilizó Proteome Discoverer versión 2.5 (PD) para el análisis del espectro MS/MS obtenido mediante LC-MS. Los datos del escaneo CID se compararon con la base de datos de secuencias de proteínas de nadeshiko RNAseq mencionada en el Ejemplo 2, mediante el algoritmo de búsqueda SEQUEST-HT. La tripsina se estableció como proteasa, y la tolerancia de masa fue de 10 ppm para el ion precursor y 0,02 Da para los iones fragmento. La oxidación de la metionina (M) (+15,995 Da) se estableció como modificación variable, y la carbamidometilación de la cisteína (C) (+57,021 Da) se estableció como modificación fija. La tasa de falsos descubrimientos (FDR) del péptido se calculó con PD, basándose en una FDR del 1,0 %. Se observaron numerosas secuencias que se correlacionaban con la abundancia relativa en función del nivel de actividad y la intensidad del espectro de masas, como DN48015_c0_g1_i5.

[0058]

[Ejemplo 8: Purificación enzimática de la variedad nadeshiko BNG]

Se utilizó un mortero para pulverizar 130 g de pétalos de nadeshiko en presencia de nitrógeno líquido. Se añadieron 600 ml de tampón de fosfato de potasio 20 mM a pH 7,0, 100 ml de cloruro de sodio y 400 µM de DTT. La mezcla se agitó con un agitador durante 2 horas en hielo para extraer la proteína. La mezcla se centrifugó a 8.000 rpm durante 20 minutos y el sobrenadante se filtró con un paño Miracloth y se recogió. Tras añadir 150 ml de agua al precipitado, este se suspendió y se centrifugó a 8.000 rpm durante 20 minutos. Tras esto, el sobrenadante se filtró con un paño Miracloth y se recogió, y se añadió a la solución previamente recogida.

[0059]

Tras añadir sulfato de amonio hasta saturación del 35% para su disolución, la solución se

centrifugó a 8.000 rpm durante 20 minutos y se recogió el sobrenadante. Se añadió sulfato de amonio a la mezcla hasta saturación del 70% y se descartó el sobrenadante. El precipitado se disolvió en aproximadamente 40 ml de tampón de fosfato de potasio 20 mM a pH 7,0 y se dializó contra tampón de fosfato de potasio 10 mM a pH 7,0 + DTT 0,1 mM. Tras la diálisis, se realizó una separación centrífuga a 6.000 rpm durante 10 minutos y el sobrenadante se adsorbió en dos plataformas de Q-Sepharose Fast Flow de 20 ml, a la mitad cada una. Cada una se lavó con 20 ml de tampón de fosfato de potasio a pH 7,0, 20 mM de tampón de fosfato de potasio a pH 6,2, 50 ml de tampón de fosfato de potasio a pH 7,0 y 50 mM de cloruro de sodio. Posteriormente, se eluyó con 30 ml de tampón de fosfato de potasio a pH 7,0 y 200 mM de cloruro de sodio, y posteriormente con 30 ml de tampón de fosfato de potasio a pH 7,0 y 500 mM de cloruro de sodio. Se recogieron aproximadamente 3 ml de eluido de cada una. Se midió la actividad de cada fracción, y la fracción con actividad observada se dializó frente a tampón de fosfato de potasio 10 mM a pH 7,0.

[0060]

La fracción dializada se adsorbió en 18 ml de DEAE Sepharose Fast Flow, equilibrada con el mismo tampón. El producto se lavó con 20 ml de tampón de fosfato de potasio 10 mM a pH 7,0, 20 ml de tampón de fosfato de potasio 10 mM a pH 7,0, 20 mM de cloruro de sodio, 20 ml de tampón de fosfato de potasio 10 mM a pH 7,0, 100 mM de cloruro de sodio, 20 ml de tampón de fosfato de potasio 10 mM a pH 7,0 y 200 mM de cloruro de sodio, en ese orden. La fracción con actividad observada se adsorbió en 6 g de hidroxiapatita cerámica CHT, Tipo 1, equilibrada con 10 mM de fosfato de potasio 40 μ M a pH 7,0. Tras lavar con 30 ml de fosfato de potasio 40 mM, pH 7,0, se eluyó con fosfato de potasio 120 mM, pH 7,0. La fracción con actividad detectada se diluyó 5 veces con agua y se adsorbió en HiTrap Sepharose Q Fast Flow. La proteína se eluyó utilizando un tampón de fosfato de potasio 20 mM a pH 7,0 como base, con un gradiente de concentración lineal de 0 mM a 1 M (total: 75 ml, velocidad de flujo: 2 ml/min), y se recuperaron a 1,5 ml cada una. La fracción activa se separó en una primera mitad (con una adsorción relativamente débil en HiTrap Sepharose Q, como fracción A) y una segunda mitad (con una adsorción relativamente fuerte en HiTrap Sepharose Q, como fracción B), y cada una se fraccionó de la siguiente manera.

[0061]

Tras disolver el sulfato de amonio en cada una de las fracciones a 1 M, se adsorbió en 6 ml de Butyl Sepharose Fast Flow, equilibrada con tampón de fosfato de potasio 20 mM a pH 7,0 + sulfato de amonio 1 M. Tras lavar con aproximadamente 25 ml de la misma solución, se eluyó con tampón de fosfato de potasio 20 mM a pH 7,0 + sulfato de amonio 0,6 M. Tras recoger 2 ml de cada una, las fracciones activas se concentraron a 0,5 ml con Amicon Ultra 30K (Merck Millipore). A continuación, se realizó una cromatografía de filtración en gel del mismo modo

que en el Ejemplo 7, y cada fracción activa con la actividad más intensa se suministró para su posterior purificación.

[0062]

Después de una dilución de 5 veces con agua, se realizó una cromatografía en columna utilizando 1 ml de Hi Trap DEAE FF (Cytiva), como se describe en el Ejemplo 7. La actividad debida a la fracción A se eluyó en el mismo tiempo de retención que en el Ejemplo 7. Sin embargo, la actividad debida a la fracción B tuvo un pico de actividad en una ubicación con un tiempo de retención largo, lo que sugiere la posibilidad de que esto se debiera a dos especies de moléculas de enzima con diferentes puntos isoeléctricos.

[0063]

Las fracciones activas derivadas de las fracciones A y B se adsorbieron en 0,5 g de 40 μ M hidroxiapatita cerámica CHT, Tipo 1, que se había equilibrado con tampón de fosfato de potasio 20 mM a pH 7,0. Después de lavar con 3 ml de tampón de fosfato de potasio 50 mM a pH 7,0, se eluyó con 3 ml de tampón de fosfato de potasio 100 mM a pH 7,0 y luego se eluyó con 3 ml de tampón de fosfato de potasio 140 mM a pH 7,0. Se recogieron aproximadamente 0,4 ml de eluido de cada uno. La actividad se eluyó con tampón de fosfato de potasio 140 mM a pH 7,0 para la muestra derivada de la fracción A, y con tampones de fosfato de potasio 100 mM y 140 mM a pH 7,0 para la muestra derivada de la fracción B. Las proteínas de cada fracción se analizaron por MS para identificar las secuencias contig nadeshiko correspondientes. Se pensó que las proteínas con actividad paralela e intensidad de MS se derivaban de las secuencias de nucleótidos DN48015_c0_g1_i5 y DN44046_c2_g1_i1 preparadas en el Ejemplo 3.

[0064]

Tras el fraccionamiento de las muestras de las fracciones A y B con hidroxiapatita, la fracción con mayor actividad se fraccionó mediante electroforesis en gel isoeléctrico con Isoelectric-Focusing Novex (Thermo Fisher Science) a pH 3-7. Se tiñó una porción de la banda con un kit de tinción de plata (FujiFilm-Wako Pure Chemical Industries), se recogieron dos geles correspondientes a las bandas de la fracción A y dos de la fracción B, y se realizó una extracción a 4 °C durante 48 horas con 0,4 ml de tampón de fosfato de potasio 20 mM a pH 7,0 + 0,1 % de Triton X-100. Tras la recuperación, la mezcla se concentró con Amicon Ultra 30K y se midió la actividad. El análisis de MS de la proteína se realizó de la misma manera que se describió anteriormente. En este análisis, también se detectó DN44046_c2_g1_i1 como una secuencia contig que exhibía una intensidad de MS correlacionada con la actividad.

[0065]

[Ejemplo 9: Identificación de genes candidatos]

La secuencia de nucleótidos de DN44046_c2_g1_i1 es la siguiente.
TCAGTTCAGCTCACTTCAGTTCAGTTCAGCACAGCTATATTATTTAATAGAACTTCAT

TCAGTTCAGCTCAGCTCAATTCAGTTCAGTTCAGCTCATTTTCAGTTCAGTTCAGCTGT
AAAAAGTCCGAAAGAACAGCCATTTTTTCGGGACAGGAGGCAGAAACCCAGAATCAT
TCCTAAATGAATTCAGCACTGTAGTTCTTCTAACTTCCTATAATACGTCATATGAAA
GTCAACACAAGATCTTTGCTCAGAAGAAGGGGGTGATAGTGAAGGCGACTAAAGAT
5 CAAATCACAGCTATGACAGCAGAAATCGAACAAGAGAAAGAGAAGAACAAATCAA
CGACGAC (SEQ ID NO: 1)

[0066]

La secuencia de nucleótidos de DN48015_c0_g1_i5 es la siguiente.

TCAACAGGTAATATATTTGTCGACATTACGCTCGCTAATCTCTATATTTTTGTCAAAA
10 ACTTGCTTTAGATGTAGCTAATTGAAGGTACCTATAGGATTAGTGTGACCCTGTAAG
CAAATCGGTTTGTGCGACATTTTTGACGCTTGTCTATGTTTCAGTGTATAAAAAATTTTC
TTGTGACTCAATATGTGTGTGTAATTTTATAGGGTGGCATGAACACTCCGTTTATTAG
TTCAAGGGCAACAAAGATCATAATGGTAGTAGACGGTCGAGGCCGCTCGAGATGC
CGGCCTTAAACGATAGCATAGATGGCACAACCGAAAACCTATTACAAGCAAGTCAGC
15 ACGGAACTAAAGCCGGGTATGGTGTGTTGTCGTTCCCTCCCGGCCATCCTATGGTTCTA
ATGTCTTCTCAGGACGAAAACTCGAACTATTTGCTACGAAGTGAATGCCAAAGG
AAATCAAAAGTTTTCCATCGCAGGGAAAGTC (SEQ ID NO: 2)

[0067]

Dado que estas secuencias de nucleótidos eran cortas y, por lo tanto, era improbable que
20 tuvieran la longitud completa, se realizó una búsqueda de genes correspondientes en una base de
datos de claveles (<http://carnation.kazusa.or.jp/>). El gen de clavel Dca32587.1 mostró una
identidad del 96 % con DN44046_c2_g1_i1 y del 99 % con DN48015_c0_g1_i5. En otras
palabras, esto sugirió que DN44046_c2_g1_i1 y DN48015_c0_g1_i5 son fragmentos del mismo
gen.

[0068]

La secuencia de nucleótidos del marco de lectura abierto del gen del clavel Dca32587.1
consta de 1161 pb, incluido el codón terminal.

ATGGGTGAAAATGGAATAAAAAATGAGAGTAACACTTGTTATGTGTTCAATGATGA
AGATTTTAGTACAATATTTAAGACTGAAGAAGGGAACTAAGTGTTCTTAAGAAATT
30 TGGTGGTTTTTTGAGAGGAATTGAGAAGTATAGAATGTTGATCTTTGAGTCAAATCC
TCTGTCTTTTTTGACACCTAGTTACTATGATTCTGATATCATCTTCTTTGTGCGGACATG
GAAAAGGAACAATAAGTTTAGTATTTGAAAATAGGAGAGAAAGTTATAACATTGAA
AAGGGACATCTAATGACGATACCAGCAGGAGTAACATTCTACATTATCAACACCGA
CGACAAAAACAAATTGGTGTGTTGTCCATGTTTCTCATTTCATTACAAATACCGGTGA
35 CTTTGTGCCATTTTTTCGGGACAGGAGGCAGAAACCCAGAATCATTCTAAATGCATT
CAGCACTGAAGTTCTTCAAACCTTCCTATAATACGTCATATGACAGTCTACAGAAGAT

CTTTGCTCAGAAGAAGGGGGTAATAGTGAAGGCGACTAAAGATCAAATCACAGCTA
 TGACAGCTGAAAACGAACAAGAGAAAGAGAAGAACAAATCAACGACGACACAATT
 TGGACCGTGTCAAATCTATAGTTATTTGAGGGAGTCTAATGAATATGGGAAACTTTT
 TATGGTTGATTATAATGAATATCAGAACTTCAAGACTTTAACATGTTTCATCTCCTTT
 5 AGCAACCTTCAACAGGGTGGCATGAACACTCCGTTTATTAGTTCAAGGGCAACAAA
 GATCATAATGGTAGTAGACGGTCGAGGCCGCTCGAGATGCCGGCCTTAAACGATA
 GCATAGATGGCACAACCGAAAACCTATTACAAGCAAGTCAGCGCGGAACCTAAAGCCG
 GGAATGGTGTGTTTGTCTGTTCCCTCCCGGCCATCCTATGGTTCTAATGTCTTCTCAGGACG
 AAAA ACTCGAACTATTTGCTACGAAGTGAATGCCAAAGGAAATCAAAAAGTTTTCC
 10 CTCGCAGGGAAAGTGAACCTTGATGAAGAACATACAAAAAGAGGCAAAGGTTTTGTG
 ATTTGGAACATTAGCTGAAGAGGTTGACCTTGTGCTCAACAGTCAACAGCAAAATTT
 CTTCATAAAAGCTCCTACTTTTAAGTAA (SEQ ID NO: 3)

[0069]

La secuencia de aminoácidos codificada consta de 386 residuos y es la siguiente.

15 MGENGIKNESNTCYVFNDEDFSTIFKTEEGKLSVLKKFGGFLRGIEKYRMLIFESNPLSFL
 TPSYYDSDIIFVGHGKGTISLVFENRRESYNIEKGHLM TIPAGVT FYIINTDDKNKL VLS
 MFLISITNTGDFVPFFGTGGRNPESFLNAFSTEV LQTSYNTSYDSLQKIFAQKKGVIVKAT
 KDQITAMTAENEQEKEKNKSTTTQFGPCQIYSYLRESNEYGKL FMVDYNEYQKLQDFN
 MFISFSNLQQGGMNTPFISSRATKIIMVVDGRGRLEMPALNDSIDGT TENYYKQVSAELK
 20 PGMVFVVP PGHPMVL MSSQDEKLETICYEVNAKGNQKFS LAGKVNLMKNIQKEAKVLS
 FGTLAEEVDLVLNSQQQNFFIKAPTFK (SEQ ID NO: 4)

[0070]

De las secuencias contig derivadas de *Silene* mencionadas en el Ejemplo 2, la secuencia que exhibió la mayor identidad con el gen fue DN21819, cuya secuencia de nucleótidos del marco de lectura abierto fue la siguiente.

25 ATGGCTTTGTCAAATGGAAAGGGAGAGAATGGGAAGCCATATGTGTTCAAAGATGA
 AGATTTTAGTACATTGTTTGAGACCCAAGAAGGGAGTTTACGAGTTCTTCAGAAATT
 TAATCAAAATTCTGATGTTTTAAGAGGAATTCAGAATTATAGGATGTTGATTTTCGA
 AGCGAATCCGTTGAGTTTTCTAACTCCTAGTCATTGGGATTCTGATATCATTTTCTTT
 30 GTTGGTCAAGGAAAAGGAACTATCAGTCTGGTTTTTGAAAAGAAGAGAGAAAGCTT
 CAACATTGAAAAGGGACATCTCATGACAATACCAGCAGGAGTTACATTCTACATTAT
 CAACAGAGATGATACCGATAAATTAGTCCTCTCAATGTTTGTCAGTTGCGATTGCTAA
 TCCTGGTGATTTTGTGCCATTTTTCACGGCAGGAGGCGGTCAAGACTCTGAATCATT
 CCTATCTGCATTCAGCACTGAGGTTCTTGAAGCAGCTTACAATGAATCAAATGAAAC
 35 ACTAAGCAAGATCTTTTCGCAACAGGGGAAGGGGGTGATAGTGAAGGCATCCAAAG
 AGCAAATTGAAGCTCTGACAGCTGAAGAAGTGACGACGGAAAGCAATTTCGGGCCT

TGCCAAATCCTTAGTTATCTTCGAGAGTCGAGTGAATGTGGGAGACTTTTCATGGTT
GATTACTGTCAATATAAGAACTCCAAGACTTCAATATGTTTCATCTCCTTTTGCAAC
CTTGAACCGGGATGCATGAATAGTCCATTTATTAGCTCAAGGGCAACAAAGATCAT
AATGGTGGTAAAGGGGCGAGGCCGAGTTGAGATGGCAGCCTCTCGCCCTGGCGAAA
5 AGGGCTCAATAGAAATTCAGTACATGAAAATCAGCTCGGAATTGAAACCGGGTGTG
GTGTTTGTATTCTCCTCCAGGGCACCTATGGTCCTAATGGCTTCTCAGGATTCAGTTT
TAGAAACTATTTGCTATGAAGTCAATGCTAAAGGAAATCAGAAATTTTCAATTGCAG
GGAAAGTGAACGTGATGAAGAACATTCAGAAAGAAGCAAAAGTACTTACATTTGGA
AAATCAGCTGAAGAAGTTGACAAGGTGCTCGACAGTTCGCAGTCGGATTCTTTCTTG
10 AAAGCTACTGGTAACAAGTAA (SEQ ID NO: 5)

[0071]

La secuencia de aminoácidos correspondiente es la siguiente.

MALSNGKGENGKPYVFKDEDFSTLFETQEGSLRVLQKFNQNSDVLRLGIQNYRMLIFEAN
PLSFLTPSHWDSDIIFVVGQKGKGTISLVFEKKRESFNIEKGHLM TIPAGVT FYIINRDDTDK
15 LVLSMFVSSIANPGDFVPFFTAGGGQDSESFLSAFSTEVLEAAYNESNETLSKIFSQQGKG
VIVKASKEQIEALTAEEVTTESNFGPCQILSYLRESSECGRLFMVDYCQYKKLQDFNMFI
SFCNLEPGCMNSPFISSRATKIIMVVKGRGRVEMAASRPGEKGSIEIQYMKISSELKPGVV
FVIPPGHMVL MASQDSVLETICYEVNAKGNQKFSIAGKVNVMKNIQKEAKVLTFGKSA
EEVDKVLDSQSFLKATGNK (SEQ ID NO: 6)

La identidad con SEQ ID NO: 4 es del 69%.

[0072]

La secuencia de aminoácidos que tiene la mayor identidad con Dca32587.1 entre las de la
base de datos de proteínas depositada en NCBI es KAH9622207.1 de *Heliosperma pusillum* de
la familia Caryophyllaceae, cuya secuencia de nucleótidos del marco de lectura abierto es la
25 siguiente.

ATGGCTTTGTAAATGGAAATGGAGAGTGTGGGAAACCATATGTGTTTAAAGATGA
AGATTTTAGTACATTGTTTAAGACTGAAGAAGGAAGTTTAAGTGTTCTTCAGAAATT
TAATGAAAGTTCTGATGTTTTAAGAGGGATTGAGAATTATAGGATGTTAATCTTTGA
AGCAAATCCATTGAGTTTCTTAACCTAGTCATTGGGATTCTGATATCATTTTCTTT
30 GTTGGTCAAGGAAAAGGAACAATCAGCCTGGTTTTTGAAAAAAGAGAGAAAGCTT
TAACATTGAAAAGGGACATCTAATGGTTATACCAGCTGGAGTTACATTCTACATTAT
CAACAGAGATGACACTCATAAATTAGTCCTCTCAATGTTTGTTAGTTCTATTGCTAAT
CCTGGTGATTTTGCGCCATTTTTCGGTGCAGGAGGTCAAGACCCTGAATCATTCCTA
TTTGCATTACAGCACTGAGGTTCTTGAAGCAGCTTACAATGCATCAATTAACACTA
35 CAGGGCACCTTTTCGGGGAAGGGAGTGATAGTGAAGGCATCCAAAGAGCAAATTA
AGCTCTGGCAGCTGAAGAAGTGACGGAAAGCAATTTTGGGCCATGCCAAATCCTTA

GTTATCTTAGAGAGTCGAATGAATGCGGGAGACTCTTCATGGTTGATTACTGTGAAT
ATAAGAACTCCAAGACTTCAATATGTTTCATCTCCTTCTGCAACCTTGAACCGGGAT
GCATGAATAGTCCATTTATTAGTTCAAGGGCAACAAAGATCATAATGGTGGTAAAG
GGGCGAGGTTCGAGTGGAGATGGCAGTTTCTCGCAATGGCGAAAAGGGCTCAACAGG
5 AACTCAATATAAGAAAATCAGCTCTGAATTGAAGCCGGGCGTAGTGTGTTATTCC
TCCAGGGCACCTATGGTCCTTATGGCTTCTCAGGACGAAGTTCTAGAACTATTTG
CTATGAAGTCAATGCTAAAGGAAATCAGAAGTTTTCACTTGCAGGTAAAGTGAAGTT
GATGAAGAACATTCAGAAAGAAGCAAAAGTACTGTCATTTGGAACATCAGCTGAAG
AAGTAGACAGTATGCTTGACAGTCAGCAGTCGGATTTCTTTTAAAAGCTACTAGTA
10 ACAAATTACAGGACTAG (SEQ ID NO: 7)

[0073]

La secuencia de aminoácidos codificada es la siguiente.

MALLNGNGECGKPYVFKDEDFSTLFKTEEGSLSVLQKFNESSDVLRGIIENYRMLIFEANP
LSFLTPSHWSDIIFVVGQKGKTISLVFEKKRESFNIEKGHLMVIPAGVTFYIINRDDTHKL
15 VL SMFVSSIANPGDFAPFFGAGGQDPESFLFAFSTEVLEAAYNASIKTLQGTFSGKGIVIVK
ASKEQIKALAAEEVTESNFGPCQILSYLRESNECGRLFMVDYCEYKKLQDFNMFISFCNL
EPGCMNSPFISSRATKIIMVVKGRGRVEMAVSRNGEKGSTGTQYKKISSELKPGVVFVIP
PGHPMVL MASQDEVLETICYEVNAKGNQKFSLAGKVNLMKNIQKEAKVLSFGTSAEEV
DSMLDSQQSDFFLKATSNKLQD (SEQ ID NO: 8)

Las identidades para SEQ ID NO: 4 y 6 fueron 72% y 88%, respectivamente.

[0074]

[Ejemplo 10: Medición de la actividad del gen candidato]

La secuencia de Dca32587.1 se expresó mediante un método común (por ejemplo, NPL 11) utilizando el vector de expresión pET15b (Novagen) de *E. coli*, y se midió su actividad. Como se muestra en la Figura 2, *Escherichia coli* que expresa Dca32587.1 mostró actividad de conversión específica de 2-hidroxinaringenina 6-C-glucósido en isovitexina, por lo que se demostró que Dca32587.1 es un gen que codifica la actividad deshidratasa deseada.

De manera similar, cuando DN21819 de *Silene* y KAH9622207.1 de *Heliosperma pusillum* se expresaron usando pET15b y se midió la actividad, exhibieron actividad de conversión específica de 2-hidroxinaringenina 6-C-glucósido a isovitexina como se muestra en la Figura 3, lo que sugiere que estos genes codifican la actividad deshidratasa objetivo.

[0075]

[Ejemplo 11: Adquisición del gen deshidratasa de longitud completa de nadeshiko]

Como se muestra en el Ejemplo 9, la secuencia de nucleótidos del gen de la deshidratasa de nadeshiko se obtuvo únicamente en fragmentos (SEQ ID NO: 1, SEQ ID NO: 2), por lo que la secuencia de nucleótidos completa se obtuvo mediante un método común. En concreto, se buscó

el gen de la deshidratasa derivado del clavel (SEQ ID NO: 3) en la base de datos mencionada en el Ejemplo 6, obteniéndose un total de cinco fragmentos del gen de la deshidratasa de nadeshiko. Dado que no se obtuvo la secuencia completa ni siquiera al ligar los fragmentos, se realizó una PCR utilizando ADNc sintetizado a partir del ARNm de nadeshiko como molde para obtener una

5 secuencia de nucleótidos que codifica la región de traducción completa. La secuencia es la siguiente.

ATGGGTGAAAATGGAGTCAAAATTGTGAGTAACACTTGTTATGTGTTCAATGATGAA
GATTTTAGTACAATGTTTAAGACTGAAGAAGGGAACTAAGTGTCTTAAGAAATTT
GGTGGTTTTTTGAGAGGAATTGAGAAGTATAGAATGTTGATCTTTGAATCAAATCCT
10 CTGTCTTTTTTGACACCTAGTCACTATGATTCTGATATCATCTTCTTTGTTGGACATG
GAAAAGGAACAATAAGTTTAGTATTTGAAAATAGGAGAGAAAGTTACAACATTGAA
AAGGGACATCTAATGACGATACCAGCAGGTGTAACATTTTACATTATCAACACCGA
CGACAAAACAAATTGATGTTGTCCATGTTTGTCAATTCATTACAAATACCGGTGA
CTTTGTGCCATTTTTTTGGGACAGGAGGTGGAAACCCAGAATCATTCTAAATGCATT
15 CAGCACTGAAGTGCTTCAAACCTCCTATAATACGTCATATGACAGTCTACAGAAGAT
CTTTGCTCAGAAGAAGGGGGTGATAGTGAAGGCGACTAAAGATCAAATCACAGCTA
TGACAGCAGAAATCGAACAAGAGAAAGAGAAGAACAAATCAACGACGACACAATT
TGGACCGTGTCAAATCTATAGTTATTTGAGGGAGTCTAATGAATATGGGAACTTTT
TATGGTTGATTATAATGAATATCAGAACTTCAAGACTTTAACATGTTTCATCTCCTTT
20 AGCAACCTTCAACAGGGTGGCATGAACACTCCGTTTATTAGTTCAAGGGCAACAAA
GATCATAATGGTAGTAGACGGTCGAGGCCACCTCGAGATGGCGGCCTTAAATGATA
GCATAGATGGCACAACCGAAAACCTATTACAAGCAAGTCAGCACGGAACCTAAAGCCG
GGTATGGTGTGTTGTCGTTCCCTCCCGGCCATCCTATGGTTCTAATGTCTTCTCAGGACG
AAAAACTCGAACTATTTGCTACGAAGTGAATGCCAAAGGAAATCAAAAGTTTTCC
25 CTCGCAGGGAAAGTGAACCTTGATGAAGAACATACAAAAAGAGGCAAAGGTTTTGTC
ATTTGGAACATTAGCTGAAGAGGTTGACCTTGTGCTCAACGGTCAACAGCAAAATTT
CTTCATAAAAGCTCCTACTTTTTTAA (SEQ ID NO: 9)

[0076]

La secuencia de aminoácidos codificada por la secuencia de nucleótidos es la siguiente.

MGENGVKIVSNTCYVFNDEDFSTMFKTEEGKLSVLKKFGGFLRGIEKYRMLIFESNPLSF
30 LTPSHYDSDIIFVGHGKGTISLVFENRRRESYNIEKGHLMTIPAGVTFYIINTDDKNKLML
SMFVISITNTGDFVPPFGTGGGNPESFLNAFSTEVLQTSYNTSYDSLQKIFAQKKGIVVKA
TKDQITAMTAEIEQEKEKNKSTTTQFGPCQIYSYLRESNEYGKLFMVDYNEYQKLQDFN
MFISFSNLQQGGMNTPFISSRATKIIMVVDGRGHLEMAALNDSIDGTTENYYKQVSTELK
35 PGMVFVPPGHPMVLMSQDEKLETICYEVNAKGNQKFSLAGKVNLMKNIQKEAKVLS
FGTLAEVVDLVLNGQQQNFFIKAPTF*(SEQ ID NO: 10)

Las identidades para SEQ ID NO: 4, 6 y 8 son 96%, 69% y 71%, respectivamente.

[0077]

[Ejemplo 12: Medición de la actividad del gen de la deshidrogenasa derivado de la variedad nadeshiko Miite]

5 El gen de la deshidratasa de longitud completa de nadeshiko (SEQ ID NO: 9) obtenido en el Ejemplo 11 se expresó como se describió anteriormente utilizando el vector de expresión de *E. coli* pET15b (Novagen) y se midió su actividad. Dado que se observó actividad de conversión específica de 2-hidroxinaringenina 6-C-glucósido a isovitexina, como se muestra en la Figura 4, el gen Miite también demostró ser un gen que codifica la actividad de la deshidratasa objetivo (el tiempo de retención se acortó durante el análisis y, por lo tanto, se desvió del estándar.).

[0078]

[Ejemplo 13: Adquisición de la línea de rosas 7472-2-1]

15 Se construyó un vector binario pSPB7472 que contenía los genes de flavonoide 3',5'-hidroxilasa de campánula, la antocianina 3',5'-metil transferasa de torenia, la flavanona 2-hidroxilasa de regaliz, la flavona C-glucosa transferasa de trigo sarraceno y la deshidratasa de Lotus japonicus, y el vector se introdujo en Ocean Song (OS) como hospedadora de rosa, para obtener la línea de rosa 7472-2-1 que almacena 2-hidroxinaringenina 6-C-glucósido en los pétalos.

[0079]

20 [Ejemplo 14: Construcción de un vector binario]

Basado en un método común conocido por aquellos expertos en la técnica (véanse, por ejemplo, Transgenic Res. 1995, 4, 288-290 y Plant Cell Physiol. 1996, 37, 49-59), se fabricaron los vectores binarios pSPB8491 y pSPB8493 utilizando pBinPLUS como estructura base.

· pSPB8491: El2 35S-2HDH de clavel (SEQ ID NO: 3)-terminador nos (para la expresión de 2HDH de clavel))

· pSPB8493: AtFLS-GUS-terminador nos (gen GUS ligado a la región promotora del gen de flavonol sintasa de Arabidopsis, utilizado como control negativo).

[0080]

[Ejemplo 15: Expresión transitoria de pétalos de rosa con clavel 2HDH]

30 Los plásmidos pSPB8491 y pSPB8493 se introdujeron en Agrobacterium Agl0 y se obtuvieron colonias en una placa L con kanamicina (50 mg/l). Tras cultivar durante la noche a 28 °C con 4 ml de medio L (con kanamicina (50 mg/l)), se inocularon 2 ml en 30 ml de medio L (con kanamicina (50 mg/l)) y se cultivó durante 5 1/2 horas a 28 °C, tras lo cual se recogieron las células. Tras lavar con 10 ml de tampón MES (MES-NaOH 10 mM, pH 5,6, acetosiringona 100 μM, MgCl₂ 10 mM), se suspendió en 5 ml de tampón MES.

[0081]

Se prepararon seis pétalos de yema de OS/7472-2-1 obtenidos en el Ejemplo 13 (almacenamiento de 2-hidroxinaringenina 6-C-glucósido) cortándolos en cuadrados de aproximadamente 5 mm por la cara interna. Se les hizo una muesca con una cuchilla y pinzas puntiagudas. Se colocaron en la solución de Agrobacterium y se dejaron reposar durante la noche en un ambiente oscuro a temperatura ambiente. A la mañana siguiente, se lavaron ligeramente los pétalos con agua y se colocó un pétalo cada uno en una placa de titulación de 24 pocillos. Tras añadir 0,5 ml de solución tampón MES a cada uno, la mezcla se dejó reposar durante 3 noches en una cámara de clima artificial (25°, 16 horas, 8 horas de oscuridad).

[0082]

Tras retirar el tampón MES, se añadieron 0,5 ml de acetonitrilo y se extrajeron los flavonoides de los pétalos mediante agitación (durante una noche). La vitexina y la isovitexina se cuantificaron mediante HPLC (columna: Shim-Pack FC-ODS 4,6 x 150 mm, fase móvil A: solución acuosa de ácido fórmico al 0,1 %, fase móvil B: solución de formiato de metanol al 0,1 %, velocidad de flujo: 1 ml/min, gradiente: 0 min → 10 % fase móvil B, 12 min → 60 % fase móvil B, 20 min → 60 % fase móvil B, 21 min → 100 % fase móvil B, 22 min, 100 % fase móvil B, 23 min → 10 % fase móvil B). Aunque la 2-hidroxinaringenina 6-C-glucósido se convierte espontáneamente en vitexina e isovitexina, dado que la isovitexina se produce específicamente cuando la 2HDH actúa sobre la 2-hidroxinaringenina 6-C-glucósido, se espera que la proporción de isovitexina aumente cuando la 2HDH está funcionando.

Los valores numéricos de isovitexina/vitexina en las seis secciones de pétalos se muestran a continuación.

[Tabla 3]

	pSPB8491 (Clavel 2HDH)	pSPB8493 (AtFLS-GUS)
1	2.42	2.22
2	2.29	2.23
3	2.38	2.11
4	2.28	2.15
5	2.29	2.19
6	2.33	2.19

(Cuando se realizó una prueba t en ambos grupos, el valor P fue 0,000492202, que fue menor que 0,01, lo que representa una diferencia significativa).

[0083]

Se conjeturó que, en los pétalos infectados con el gen 2HDH derivado del clavel, la enzima

del mismo gen producía isovitexina. Esto demostró que el gen funciona no solo en *E. coli*, sino también en plantas.

[Listado de secuencias]