

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 709 439**

51 Int. Cl.:

C12P 19/44 (2006.01)
A23G 3/38 (2006.01)
A23G 3/36 (2006.01)
A23G 4/06 (2006.01)
C12P 19/56 (2006.01)
C12N 9/10 (2006.01)
A23L 27/30 (2006.01)
C12N 15/52 (2006.01)
A23L 33/105 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **03.10.2014 PCT/US2014/059081**
- 87 Fecha y número de publicación internacional: **07.05.2015 WO15065650**
- 96 Fecha de presentación y número de la solicitud europea: **03.10.2014 E 14857965 (9)**
- 97 Fecha y número de publicación de la concesión europea: **05.12.2018 EP 3063286**

54 Título: **Producción recombinante de glucósidos de esteviol**

30 Prioridad:

01.11.2013 US 201361898571 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
16.04.2019

73 Titular/es:

**CONAGEN INC. (100.0%)
15 DeAngelo Drive
Bedford, Massachusetts 01730, US**

72 Inventor/es:

**MAO, GUOHONG y
YU, XIAODAN**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 709 439 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

Producción recombinante de glucósidos de esteviol.

En la presente descripción se proporciona un listado de secuencias. Este listado de secuencias consta de las secuencias con núm. de ident.: 1 - 12.

5

Antecedentes de la invención

La presente descripción se refiere generalmente a la biosíntesis de glucósidos de esteviol. En particular, la presente descripción se refiere a un polipéptido recombinante que cataliza la producción de glucósidos de esteviol tales como rebaudiósido D, rebaudiósido E y un nuevo rebaudiósido (rebaudiósidoZ).

10

Los glucósidos de esteviol son productos naturales aislados de las hojas de *Stevia rebaudiana* y se usan ampliamente como edulcorantes de alta intensidad y bajos en calorías. Los glucósidos de esteviol de origen natural tienen la misma estructura de base (esteviol) y difieren en el contenido de residuos de carbohidratos (por ejemplo, residuos de glucosa, ramnosa y xilosa) en las posiciones C13 y C19. Los glucósidos de esteviol con estructura conocida incluyen esteviósido, rebaudiósido A, rebaudiósido B, rebaudiósido C, rebaudiósido D, rebaudiósido E, rebaudiósido F y dulcósido A.

15

Sobre la base del peso seco, el esteviósido, el rebaudiósido A, el rebaudiósido C y el dulcósido A representan 9,1, 3,8, 0,6 y 0,3 % del peso total de los glucósidos de esteviol en las hojas, respectivamente, mientras que los otros glucósidos de esteviol están presentes en cantidades mucho menores. Los extractos de la planta de *Stevia rebaudiana* se encuentran disponibles comercialmente y contienen típicamente esteviósido y rebaudiósido A como compuestos primarios. Los otros glucósidos de esteviol típicamente están presentes en el extracto de estevia como componentes menores. Por ejemplo, la cantidad de rebaudiósido A en preparaciones comerciales puede variar de aproximadamente 20 % a más de 90 % del contenido total de glucósido de esteviol, mientras que la cantidad de rebaudiósido B puede ser de aproximadamente 1-2 %, la cantidad de rebaudiósido C puede ser de aproximadamente 7-15 %, y la cantidad de rebaudiósido D puede ser de aproximadamente 2 % del total de glucósidos de esteviol.

20

25

Como edulcorantes naturales, los diferentes glucósidos de esteviol tienen diferentes grados de dulzor y sabor. La dulzura de los glucósidos de esteviol es significativamente más alta que la de la sacarosa. Por ejemplo, el esteviósido es 100-150 veces más dulce que la sacarosa con sabor amargo, mientras que el rebaudiósido A y E son 250-450 veces más dulce que la sacarosa y el sabor es mucho mejor que el esteviósido. En consecuencia, el perfil de sabor de cualquier extracto de estevia se ve profundamente afectado por el contenido relativo de los glucósidos de esteviol en el extracto, que a su vez puede verse afectado por la fuente de la planta, los factores ambientales (como el contenido del suelo y el clima), y el proceso de extracción. En particular, las variaciones de las condiciones de extracción pueden llevar a composiciones inconsistentes de los glucósidos de esteviol en los extractos de estevia, de modo que el perfil de sabor varía entre diferentes lotes de productos de extracción. El perfil de sabor de los extractos de estevia también puede verse afectado por contaminantes derivados de las plantas (como pigmentos, lípidos, proteínas, compuestos fenólicos y sacáridos) que permanecen en el producto después del proceso de extracción. Estos contaminantes típicamente tienen sabores desagradables indeseables para el uso del extracto de estevia como edulcorante.

30

35

40

La mayoría de los glucósidos de esteviol están formados por varias reacciones de glucosilación del esteviol, que típicamente son catalizadas por las UDP-glucosiltransferasas (UGT) mediante el uso de uridina 5'-difosfoglucosa (UDP-glucosa) como donante del resto de azúcar. En las plantas, las UGT son un grupo de enzimas muy divergentes que transfieren un residuo de glucosa de UDP-glucosa al esteviol. El esteviósido es un intermedio en la biosíntesis de compuestos rebaudiósido. Por ejemplo, la glucosilación del C-3' de la C-13-O-glucosa del esteviósido produce el rebaudiósido A; y la glucosilación del C-2' de la 19-O-glucosa del esteviósido produce rebaudiósido E. Además, la glucosilación del rebaudiósido A (en 19-O-glucosa) o el rebaudiósido E (en C-13-O-glucosa) produce rebaudiósido D. (Figura 1).

45

50

Un enfoque práctico para mejorar la calidad del sabor del extracto de estevia es aumentar el rendimiento de los compuestos de rebaudiósidos mediante la glucosilación adicional del esteviósido. Las UGT con una actividad de glucosilación de 1,2-19-O-glucosa son enzimas importantes para la producción de rebaudiósidos D y E.

55

La sacarosa sintasa (SUS) cataliza la conversión de la UDP a UDP-glucosa en presencia de sacarosa. Por lo tanto, para una reacción de glucosilación que utiliza UDP-glucosa (como las catalizadas por los UGT), la SUS puede usarse para regenerar UDP-glucosa a partir de UDP, lo que mejora la eficiencia de dicha reacción (Figura 2).

60

Por consiguiente, existe la necesidad de glucósidos de esteviol con un perfil de sabor consistente y menos sabor desagradable que los productos comerciales existentes. La entrada F2DT21_HORVV de Uniprot describe una proteína glucosiltransferasa de *Hordeum vulgare* subsp. *vulgare* (cebada domesticada). El documento US 2013/0171328 describe la producción de glucósidos de esteviol en microorganismos. El documento WO 2013/022989 también describe la producción recombinante de glucósidos de esteviol.

Como se describe en la presente descripción, la presente descripción proporciona un polipéptido recombinante útil para preparar glucósidos de esteviol (tales como el rebaudiósido D y el rebaudiósido E). La presente descripción también proporciona un método para producir una composición de glucósido de esteviol (rebaudiósido Z) mediante el uso de dicho polipéptido recombinante.

5

Breve descripción

10

15

20

25

30

35

40

45

50

55

60

65

La presente tecnología se refiere generalmente a polipéptidos recombinantes que tienen actividades de UDP-glucosiltransferasa, en particular, polipéptidos que tienen una actividad de glucosilación de 1,2-19-0-glucosa en compuestos de glucósido de esteviol. La presente invención proporciona el uso de un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. ident.: 6 para producir un glucósido de esteviol. Más particularmente, la presente invención proporciona un método para producir un glucósido de esteviol o la síntesis de un compuesto de glucósido de esteviol, el método comprende incubar un sustrato con un polipéptido recombinante que es una UDP-glucosiltransferasa que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de la glucosiltransferasa HVI como se expone en la sec. con núm. de ident.: 6 en presencia de UDP-glucosa como donante de glucosa. En un aspecto ilustrativo, la secuencia de aminoácidos del polipéptido recombinante para tal uso tiene al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % de identidad con la sec. con núm. de ident.: 6.

En la presente descripción también se describe un ácido nucleico aislado que comprende una secuencia de nucleótidos que codifica un polipéptido recombinante para dicho uso, en donde la secuencia de nucleótidos tiene al menos 80 % de identidad con la sec. con núm. de ident.: 7. También se describe un vector que comprende un ácido nucleico de este tipo y una célula huésped que comprende el vector descrito en la presente descripción. En una modalidad ilustrativa, la célula huésped de la presente tecnología se selecciona del grupo que consiste en bacterias, levaduras, hongos filamentosos, cianobacterias, algas y células vegetales.

En otro aspecto, la presente tecnología también se refiere a un método para producir una composición de glucósido de esteviol, el método comprende incubar un sustrato (tal como esteviósido, rebaudiósido A, rebaudiósido E o una combinación de los mismos) con un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.: 6. En un aspecto ilustrativo, la secuencia de aminoácidos del polipéptido recombinante usado en el método descrito en la presente descripción tiene al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % de identidad con la sec. con núm. de ident.:6.

En otro aspecto, la presente tecnología también se refiere a un método para producir una composición de glucósido de esteviol, el método comprende incubar un sustrato (tal como esteviósido y rebaudiósido E) con un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.:11. En un aspecto ilustrativo, la secuencia de aminoácidos del polipéptido recombinante usado en el método descrito en la presente descripción tiene al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % de identidad con la sec. con núm. de ident.:11.

En una modalidad, el método comprende además incubar una sacarosa sintasa recombinante (tal como una con una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de AtSUS1 como se indica en la sec. con núm. de ident.: 9) con el sustrato y el polipéptido recombinante descrito en la presente descripción. En otra modalidad, el método comprende además incubar una UDP-glucosiltransferasa recombinante (tal como una con una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de UGT76G1 como se indica en la sec. con núm. de ident.: 11) con la sacarosa sintasa recombinante, el sustrato y el polipéptido recombinante descritos en la presente descripción. En otro aspecto, el método descrito en la presente descripción comprende incubar el sustrato con una célula huésped que expresa el polipéptido recombinante.

La presente tecnología también se refiere a un nuevo glucósido de esteviol, denominado rebaudiósido Z, que se caracteriza por un tiempo de retención de aproximadamente 6,68 minutos en una HPLC bajo las condiciones descritas en la presente descripción. La presente tecnología también se refiere a un método para producir rebaudiósido Z descrito en la presente descripción, el método comprende incubar un sustrato con un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.: 6. Como se usa en la presente descripción, los términos "rebaudiósido Z" o "Reb Z" se refieren a una mezcla de compuestos, en particular, una mezcla de rebaudiósido Z1 ("Reb Z1") y rebaudiósido Z2 ("Reb Z2").

En una modalidad, la presente descripción se dirige además a un método para sintetizar rebaudiósido Z a partir de rebaudiósido E. El método incluye: preparar una mezcla de reacción que comprende rebaudiósido E, un sustrato seleccionado del grupo que consiste en sacarosa, difosfato de uridina (UDP) y uridina difosfato-glucosa (UDP-glucosa) y HV1, y sacarosa sintasa, incubar la mezcla de reacción durante un tiempo suficiente para producir rebaudiósido Z, en donde una glucosa se acopla covalentemente al rebaudiósido E para producir rebaudiósido Z, una glucosa se acopla

covalentemente a la C2'-13-O-glucosa del rebaudiósido E para producir rebaudiósido Z1, y una glucosa se acopla covalentemente a C2'-19-O-glucosa del rebaudiósido E para producir el rebaudiósido Z2.

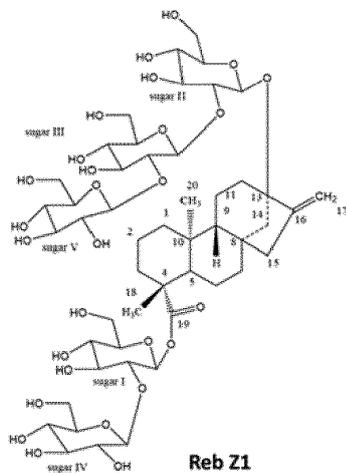
En una modalidad, el compuesto rebaudiósido Z es rebaudiósido Z1 (Reb Z1) que tiene la estructura:

5

10

15

20



Reb Z1

25

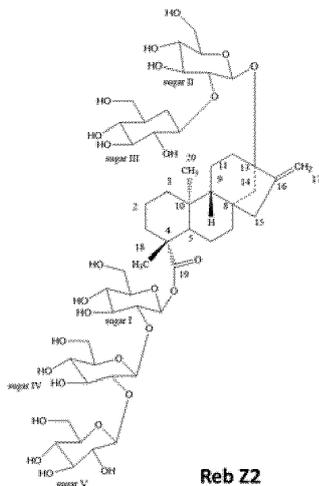
En una modalidad, el compuesto rebaudiósido Z es rebaudiósido Z2 (Reb Z2) que tiene la estructura:

30

35

40

45



Reb Z2

Como se describe en la presente descripción, los polipéptidos recombinantes de la presente tecnología se usan para desarrollar un método biosintético para preparar glucósidos de esteviol que son típicamente de baja abundancia en fuentes naturales, tales como el rebaudiósido D y el rebaudiósido E. En consecuencia, la presente tecnología también proporciona una composición de glucósidos de esteviol producida por el método biosintético descrito en la presente descripción. Dicha composición puede comprender un compuesto de glucósido de esteviol que se selecciona del grupo que consiste en rebaudiósido D, rebaudiósido E, un nuevo rebaudiósido(denominado en la presente descripción "rebaudiósido Z" y "Reb Z") y sus combinaciones. Además, también se proporciona un edulcorante que comprende la composición de glucósido de esteviol que se describe en la presente descripción.

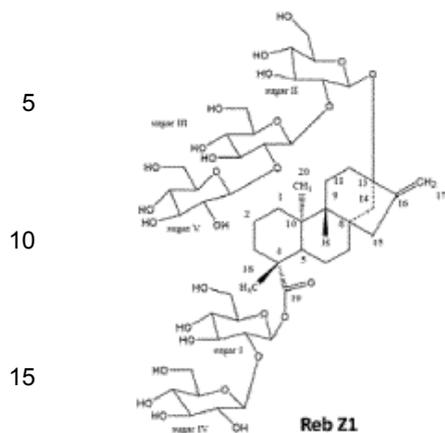
50

55

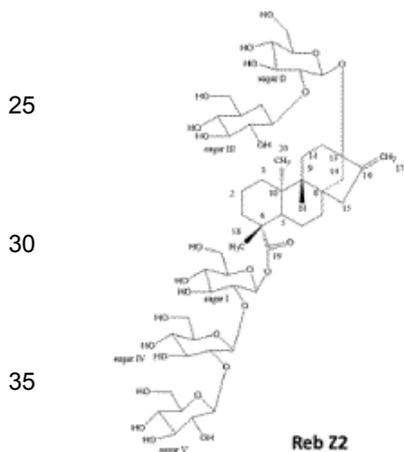
En una modalidad, la presente descripción se dirige a un edulcorante que incluye un compuesto que tiene la estructura química:

60

65



20 En otra modalidad, el edulcorante incluye un compuesto que tiene la estructura química:



40 La presente descripción se refiere además al uso de edulcorantes en productos consumibles tales como bebidas, confitería, productos de panadería, galletas y chicles.

45 Breve descripción de las figuras

La descripción se comprenderá mejor, y las características, aspectos y ventajas distintas de las expuestas anteriormente se harán evidentes cuando se considere la siguiente descripción detallada de la misma. Dicha descripción detallada hace referencia a las siguientes figuras, en donde:

50 La Figura 1 es un esquema que ilustra las rutas de biosíntesis de glucósidos de esteviol a partir de esteviósido. Como se describe en la presente descripción, el polipéptido HV1 recombinante ("HV1") tiene una actividad de glucosilación 1,2-19-O-glucosa que transfiere un segundo resto de azúcar al C-2' de 19-O-glucosa de esteviósido para producir rebaudiósido E ("Reb E"), o de manera similar para producir rebaudiósido D ("Reb D") a partir de rebaudiósido A ("Reb A"). La Figura 1 también muestra que una enzima UGT76G1 recombinante ("UGT76G1", diferente del polipéptido HV1 recombinante) cataliza la reacción que transfiere un resto de azúcar a C-3' de la C-13-O-glucosa de esteviósido para producir rebaudiósido A, o de manera similar para producir rebaudiósido D a partir de rebaudiósido E.

60 La Figura 2 muestra un esquema ilustrativo de un sistema de reacción de acoplamiento de UDP-glucosiltransferasa ("UGT") y sacarosa sintasa ("SUS"). La reacción 1 muestra una reacción catalizada por UGT que convierte el rebaudiósido A ("Reb A") en rebaudiósido D ("Reb D"), que usa UDP-glucosa como donante de glucosa y da como resultado la producción de UDP. La reacción 2 muestra una reacción catalizada por SUS que convierte UDP en UDP-glucosa, que usa sacarosa como donante de glucosa. La reacción 2 también muestra que la reacción catalizada por SUS puede acoplarse a la reacción catalizada por UGT.

65 La Figura 3 muestra la producción *in vitro* de rebaudiósido D ("Reb D") a partir de rebaudiósido A ("Reb A") catalizada por un polipéptido HV1 recombinante (sec. con núm. de ident.: 6) y un AtSUS1 recombinante (sec. con núm. de ident.: 9) en

un sistema de reacción de acoplamiento HV1-AtSUS1 como se describió en la presente descripción. La Figura 3A muestra los estándares de esteviósido ("Ste"), rebaudiósido A ("Reb A") y rebaudiósido D ("Reb D"). Los resultados a las 6, 9, 12 y 24 horas se muestran en las Figuras 3B-E, respectivamente. Los resultados de la reacción sin el AtSUS1 recombinante (es decir, una reacción sin acoplamiento) a las 12 y 24 horas se muestran en las Figuras 3F y 3G, respectivamente.

La Figura 4 muestra la producción *in vitro* de rebaudiósido E ("Reb E") a partir de esteviósido catalizada por un polipéptido HV1 recombinante (sec. con núm. de ident.: 6) y un AtSUS1 recombinante (sec. con núm. de ident.: 9) en un sistema de reacción de acoplamiento HV1-AtSUS1 como se describió en la presente descripción. La Figura 4A muestra los estándares de esteviósido ("Ste"), rebaudiósido A ("Reb A") y rebaudiósido D ("Reb D"). El resultado a las 20 horas se muestra en las Figuras 4B, que incluye un compuesto rebaudiósido Z ("Reb Z").

La Figura 5 muestra la producción *in vitro* de rebaudiósido D ("Reb D") a partir de esteviósido catalizada por una combinación de un polipéptido HV1 recombinante (sec. con núm. de ident.: 6), un UGT76G1 recombinante (sec. con núm. de ident.: 11) y un AtSUS1 recombinante (sec. con núm. de ident.: 9). La Figura 5A muestra los estándares de esteviósido ("Ste"), rebaudiósido A ("Reb A") y rebaudiósido D ("Reb D"). Los resultados a las 6, 9, 12 y 24 horas se muestran en las Figuras 5BE, respectivamente.

La Figura 6 muestra el análisis de SDS-PAGE del polipéptido HV1 recombinante purificado.

La Figura 7 muestra el análisis de SDS-PAGE del polipéptido AtSUS1 recombinante purificado.

La Figura 8 muestra la producción *in vitro* de rebaudiósido Z ("Reb Z") a partir de rebaudiósido E ("Reb E") catalizada por un polipéptido HV1 recombinante (sec. con núm. de ident.: 6) y un AtSUS1 recombinante (sec. con núm. de ident.: 9) en un sistema de reacción de acoplamiento HV1-AtSUS1 como se describió en la presente descripción. La Figura 8A muestra los estándares del rebaudiósido E ("Reb E"). El resultado a las 24 horas se muestra en las Figuras 8B, que incluye un compuesto rebaudiósido Z ("Reb Z").

La Figura 9 muestra el análisis de SDS-PAGE del polipéptido UGT76G1 recombinante purificado.

Figura 10: Estructuras de rebaudiósido Z (incluyendo Reb Z1 y Reb Z2) y rebaudiósido E.

Figura 11: Correlaciones principales de TOCSY y HMBC para Reb Z1 y Reb Z2.

Figura 12: muestra la producción *in vitro* de rebaudiósido D ("Reb D") a partir de rebaudiósido E ("Reb E") catalizada por un UGT76G1 recombinante (sec. con núm. de ident.: 11). La Figura 12A-12B muestra los estándares de rebaudiósido E ("Reb E"), rebaudiósido D ("Reb D"). Los resultados de la reacción sin el AtSUS1 recombinante (Figura 12C) (es decir, una reacción sin acoplamiento) y con el AtSUS1 recombinante (Figura 12D) (es decir, la reacción de acoplamiento UGT-SUS) a las 6 horas se muestran respectivamente.

Si bien la descripción es susceptible de diversas modificaciones y formas alternativas, sus modalidades específicas se mostraron a manera de ejemplo en los dibujos y se describen en la presente descripción a detalle más abajo.

Descripción Detallada

La presente tecnología proporciona el uso de un polipéptido recombinante que tiene actividades de UDP-glucosiltransferasa, más particularmente actividad de glucosilación de 1,2-19-O-glucosa y actividad de glucosilación de 1,2-13-O-glucosa, para sintetizar glucósidos de esteviol. El polipéptido recombinante de la presente tecnología (que también puede denominarse "polipéptido HV1 recombinante" en lo sucesivo) es útil para la biosíntesis de compuestos de glucósido de esteviol. En la presente descripción, la UDP-glucosiltransferasa (UGT) se refiere a una enzima que transfiere un residuo de azúcar de una molécula donadora activada (típicamente UDP-glucosa) a una molécula aceptora. La actividad de la glucosilación 1,2-19-O-glucosa se refiere a una actividad enzimática que transfiere un resto de azúcar al C-2' del resto 19-O-glucosa de esteviósido, rebaudiósido A o rebaudiósido E (Figura 1 y Figura 10). La actividad de la glucosilación 1,2-13-O-glucosa se refiere a una actividad enzimática que transfiere un resto de azúcar al C-2' del resto 13-O-glucosa del rebaudiósido E (Figura 10).

Los nombres de las enzimas UGT usadas en la presente descripción son consistentes con el sistema de nomenclatura adoptado por el Comité de Nomenclatura UGT (Mackenzie y otros, "La superfamilia del gen de la glucosiltransferasa UDP: nomenclatura recomendada actualizada según la divergencia evolutiva," *Pharmacogenetics*, 1997, vol. 7, pp. 255-269), que clasifica los genes UGT por la combinación de un número de familia, una letra que denota una subfamilia y un número para un gen individual. Por ejemplo, el nombre "UGT76G1" se refiere a una enzima UGT codificada por un gen que pertenece a la familia de UGT número 76 (que es de origen vegetal), subfamilia G y gen número 1.

Existe una gran familia de genes UGT en las plantas. Sin embargo, las funciones biológicas de la mayoría de estas UGT siguen desconociéndose.

Definiciones

Como se usa en la presente descripción, las formas en singular "un", "uno/una" y "el/la" incluyen las referencias en plural a menos que el contexto indique claramente lo contrario.

5 En la medida en que el término "incluyen", "tienen" o similares se usan en la descripción o las reivindicaciones, dicho término se destina a ser inclusivo de una manera similar al término "comprenden" tal como se interpreta el término "comprenden" cuando se emplea como una palabra transicional en una reivindicación.

10 La palabra "ilustrativo" se usa en la presente descripción y significa "que sirve como un ejemplo, caso, o ilustración." Cualquier modalidad descrita en la presente descripción como "ilustrativa" no debe interpretarse necesariamente como preferida o ventajosa sobre otras modalidades.

15 El término "complementario" debe recibir su significado común y habitual para una persona con experiencia corriente en la técnica, y se utiliza sin limitación para describir la relación entre las bases de nucleótidos que pueden hibridarse entre sí. Por ejemplo, con respecto al ADN, la adenosina es complementaria a la timina y la citosina es complementaria a la guanina. Por consiguiente, la presente tecnología también incluye fragmentos de ácido nucleico aislados que son complementarios a las secuencias completas tal como se informa en el Listado de Secuencias que se acompaña, así como también aquellas secuencias de ácido nucleico sustancialmente similares.

20 Los términos "ácido nucleico" y "nucleótido" deben recibir sus respectivos significados comunes y habituales para una persona con experiencia corriente en la técnica, y se usan sin limitación para referirse a los desoxirribonucleótidos o ribonucleótidos y sus polímeros, ya sea en forma de cadena simple o doble. A menos que se limite específicamente, el término abarca ácidos nucleicos que contienen análogos conocidos de nucleótidos naturales que tienen propiedades de unión similares al ácido nucleico de referencia y se metabolizan de manera similar a los nucleótidos de origen natural. A menos que se indique de otra forma, una secuencia de ácido nucleico particular implícitamente también abarca las variantes redundantes o modificadas de forma conservadora de estos (por ejemplo, las sustituciones de codones redundantes) y las secuencias complementarias, así como también la secuencia que se indica explícitamente.

30 El término "aislado" debe recibir su significado común y habitual para una persona con experiencia corriente en la técnica, y cuando se usa en el contexto de un ácido nucleico aislado o un polipéptido aislado, se usa sin limitación para referirse a un ácido nucleico o polipéptido que, de la mano del hombre, existe aparte de su entorno nativo y, por lo tanto, no es un producto de la naturaleza. Un ácido nucleico o polipéptido aislado puede existir en una forma purificada o puede existir en un entorno no nativo tal como, por ejemplo, en una célula huésped transgénica.

35 Los términos "incubar" e "incubación", como se usan en la presente descripción, significan un proceso de mezclar dos o más entidades químicas o biológicas (como un compuesto químico y una enzima) y permitirles interactuar en condiciones favorables para producir una composición de glucósido de esteviol.

40 El término "variante redundante" se refiere a una secuencia de ácido nucleico que tiene una secuencia de residuos que difiere de una secuencia de ácido nucleico de referencia en una o más sustituciones de codones redundantes. Las sustituciones de codones redundantes pueden lograrse mediante la generación de secuencias en las cuales la tercera posición de uno o más (o todos) codones seleccionados se sustituye con residuos con base mezclada y/o desoxiinosina. Una secuencia de ácido nucleico y todas sus variantes degeneradas expresarán el mismo aminoácido o polipéptido.

45 Los términos "polipéptido", "proteína" y "péptido" deben recibir sus respectivos significados ordinarios y habituales para una persona con experiencia ordinaria en la técnica; los tres términos a veces se usan indistintamente, y se usan sin limitación para referirse a un polímero de aminoácidos, o análogos de aminoácidos, independientemente de su tamaño o función. Aunque la "proteína" se usa a menudo en referencia a polipéptidos relativamente grandes, y el "péptido" se usa a menudo en referencia a polipéptidos pequeños, el uso de estos términos en la técnica se solapa y varía. El término "polipéptido" como se usa en la presente descripción se refiere a péptidos, polipéptidos y proteínas, a menos que se indique lo contrario. Los términos "proteína", "polipéptido" y "péptido" se usan indistintamente en la presente descripción cuando se hace referencia a un producto polinucleotídico. Así, los ejemplos de polipéptidos incluyen productos polinucleotídicos, proteínas naturales, homólogos, ortólogos, parálogos, fragmentos y otros equivalentes, variantes y análogos de los anteriores.

55 A los términos "fragmento de polipéptido" y "fragmento", cuando se usan en referencia a un polipéptido de referencia, se les debe otorgar sus significados comunes y habituales a una persona con experiencia corriente en la técnica, y se usan sin limitación para referirse a un polipéptido en el cual se eliminan residuos de aminoácidos en comparación con el propio polipéptido de referencia, pero donde la secuencia de aminoácidos restante es generalmente idéntica a las posiciones correspondientes en el polipéptido de referencia. Dichas eliminaciones pueden ocurrir en el extremo amino o en el extremo carboxilo del polipéptido de referencia, o alternativamente ambos.

60 El término "fragmento funcional" de un polipéptido o proteína se refiere a un fragmento peptídico que es una porción del polipéptido o proteína de longitud completa, y tiene sustancialmente la misma actividad biológica, o desempeña sustancialmente la misma función que el polipéptido o proteína de longitud completa (por ejemplo, al realizar la misma reacción enzimática).

Los términos "polipéptido variante", "secuencia de aminoácidos modificada" o "polipéptido modificado", que se usan indistintamente, se refieren a una secuencia de aminoácidos que es diferente del polipéptido de referencia por uno o más aminoácidos, por ejemplo, por uno o más Sustituciones, supresiones y/o adiciones de aminoácidos. En un aspecto, una variante es una "variante funcional" que conserva parte o toda la capacidad del polipéptido de referencia.

El término "variante funcional" incluye además variantes con sustituciones conservativas. El término "variante sustituida de forma conservativa" se refiere a un péptido que tiene una secuencia de aminoácidos que difiere de un péptido de referencia en una o más sustituciones de aminoácidos conservativas, y mantiene parte o toda la actividad del péptido de referencia. Una "sustitución conservativa de aminoácidos" es una sustitución de un residuo de aminoácido con un residuo funcionalmente similar. Ejemplos de sustituciones conservativas incluyen la sustitución de un residuo no polar (hidrofóbico) tal como isoleucina, valina, leucina o metionina por otro similar, la sustitución de un residuo polar o cargado (hidrofílico) por otro similar tal como de entre arginina y lisina, entre glutamina y asparagina, entre treonina y serina; la sustitución de un residuo básico tal como lisina o arginina por otro; o la sustitución de un residuo ácido, tal como ácido aspártico o ácido glutámico por otro; o la sustitución de un residuo aromático, tal como fenilalanina, tirosina o triptófano por otro. Se espera que tales sustituciones tengan poco o ningún efecto sobre el peso molecular aparente o el punto isoeléctrico de la proteína o polipéptido. La frase "variante sustituida de forma conservativa" también incluye péptidos en donde un residuo se reemplaza con un residuo derivado químicamente, siempre y cuando el péptido resultante mantenga parte o toda la actividad del péptido de referencia como se describe en la presente descripción.

El término "variante", en relación con los polipéptidos de la presente tecnología, incluye además un polipéptido funcionalmente activo que tiene una secuencia de aminoácidos al menos 75 %, al menos 76 %, al menos 77 %, al menos 78 %, al menos 79 %, al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, e incluso 100 % idéntico a la secuencia de aminoácidos de un polipéptido de referencia.

El término "homólogo" en todas sus formas gramaticales y variaciones ortográficas se refiere a la relación entre polinucleótidos o polipéptidos que poseen un "origen evolutivo común", incluidos los polinucleótidos o polipéptidos de superfamilias y polinucleótidos o proteínas homólogas de diferentes especies (Reeck y otros, Cell 50: 667, 1987). Dichos polinucleótidos o polipéptidos tienen homología de secuencia, como se refleja por su similitud de secuencia, ya sea en términos de por ciento de identidad o la presencia de aminoácidos específicos o motivos en posiciones conservadas. Por ejemplo, dos polipéptidos homólogos pueden tener secuencias de aminoácidos que son al menos 75 %, al menos 76 %, al menos 77 %, al menos 78 %, al menos 79 %, al menos 80 %, al menos 81 %, al menos 82 %, al menos 83 %, al menos 84 %, al menos 85 %, al menos 86 %, al menos 87 %, al menos 88 %, al menos 89 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % e incluso 100 % idénticos.

"Por ciento (%) de identidad de secuencia de aminoácidos" con respecto a las secuencias de polipéptido variante de la presente tecnología se define como el por ciento de residuos de aminoácidos en una secuencia candidato que son idénticos con los residuos de aminoácidos del polipéptido de referencia (como tal, por ejemplo, la sec. con núm. de ident.:6), después de alinear las secuencias e introducir interrupciones, si fuera necesario, para obtener el por ciento máximo de identidad de secuencia, y sin considerar cualquiera de las sustituciones conservativas como parte de la secuencia de identidad.

El alineamiento para fines de determinar el por ciento de identidad de la secuencia de aminoácidos puede lograrse de varias maneras que están dentro del conocimiento en la técnica, por ejemplo, usando programas computarizados públicamente disponibles tales como los programas BLAST, BLAST-2, ALIGN-2 o Megalign (DNASTAR). Los expertos en la técnica pueden determinar los parámetros adecuados para medir el alineamiento, que incluye cualquiera de los algoritmos necesarios para lograr el alineamiento máximo sobre la longitud completa de las secuencias que se comparan. Por ejemplo, el % de identidad de secuencia de aminoácidos puede determinarse utilizando el programa de comparación de secuencias NCBI-BLAST2. El programa de comparación de secuencias NCBI-BLAST2 puede descargarse de ncbi.nlm.nih.gov. El NCBI-BLAST2 usa varios parámetros de búsqueda, en donde todos esos parámetros de búsqueda se fijan para valores predeterminados que incluyen, por ejemplo, desenmascarar sí, cadena = todos, eventos esperados = 10, longitud mínima de baja complejidad = 15/5, valor-e multipaso = 0,01, constante de multipaso = 25, caída para alineación con espacios final = 25 y la matriz de puntuación = BLOSUM62. En situaciones donde se emplea NCBI-BLAST2 para las comparaciones de secuencias de aminoácidos, el % de identidad de secuencia de aminoácidos de una secuencia dada de aminoácidos A para, con, o contra una secuencia dada de aminoácidos B (que puede expresarse alternativamente como una secuencia dada de aminoácidos A que tiene o comprende un cierto % de identidad de secuencia de aminoácidos para, con, o contra una secuencia dada de aminoácidos B) se calcula como sigue: 100 veces la fracción X/Y donde X es el número de residuos de aminoácidos acumulados como coincidencias idénticas por el programa de alineamiento de secuencias NCBI-BLAST2 en ese alineamiento del programa de A y B, y donde Y es el número total de residuos de aminoácidos en B. Se apreciará que cuando la longitud de la secuencia de aminoácidos A no es igual a la longitud de la secuencia de aminoácidos B, el % de identidad de secuencia de aminoácidos de A a B no será igual al % de identidad de secuencia de aminoácidos de B a A.

En este sentido, las técnicas para determinar la "similitud" de la secuencia de aminoácidos se conocen bien en la técnica. En general, "similitud" significa la comparación exacta de aminoácido a aminoácido de dos o más polipéptidos en el lugar apropiado, donde los aminoácidos son idénticos o poseen propiedades químicas y/o físicas similares, como la carga o la hidrofobicidad. Un denominado "por ciento de similitud" puede entonces determinarse entre las secuencias polipeptídicas comparadas. Las técnicas para determinar la identidad de la secuencia de aminoácidos y ácidos nucleicos también son bien conocidas en la técnica e incluyen la determinación de la secuencia de nucleótidos del ARNm para ese gen (generalmente a través de un intermedio de ADNc) y la determinación de la secuencia de aminoácidos codificada en el mismo, y comparando esto con una segunda secuencia de aminoácidos. En general, "identidad" se refiere a una correspondencia exacta de nucleótido a nucleótido o de aminoácido a aminoácido de dos secuencias polinucleotídicas o polipeptídicas, respectivamente. Pueden compararse dos o más secuencias de polinucleótidos determinando su "por ciento de identidad", al igual que dos o más secuencias de aminoácidos. Los programas disponibles en el Paquete de Análisis de Secuencias de Wisconsin, Versión 8 (disponible de Genetics Computer Group, Madison, WI), por ejemplo, el programa GAP, son capaces de calcular tanto la identidad entre dos polinucleótidos como la identidad y similitud entre dos secuencias de polipéptidos, respectivamente. Los expertos en la técnica conocen otros programas para calcular la identidad o similitud entre secuencias.

Una posición de aminoácido "correspondiente a" una posición de referencia es una posición que se alinea con una secuencia de referencia, como se identifica al alinear las secuencias de aminoácidos. Dichos alineamientos pueden hacerse a mano o utilizando programas de alineamiento de secuencias bien conocidos, como ClustalW2, Blast 2, etcétera.

A menos que se especifique de cualquier otra manera, el por ciento de identidad de dos secuencias polipeptídicas o polinucleotídicas se refiere al por ciento de residuos de aminoácidos o nucleótidos idénticos en toda la longitud de la más corta de las dos secuencias.

La "secuencia codificante" debe otorgar a un experto en la técnica su significado común y habitual, y se utiliza sin limitación para referirse a una secuencia de ADN que codifica una secuencia de aminoácidos específica.

"Secuencias reguladoras adecuadas" deben recibir su significado ordinario y habitual para un experto en la técnica y se usan sin limitación para referirse a secuencias de nucleótidos ubicadas aguas arriba (secuencias no codificantes en 5'), dentro, o aguas abajo (secuencias no codificantes en 3') de una secuencia codificante, y que influyen en la transcripción, el procesamiento o estabilidad del ARN, o la traducción de la secuencia codificante asociada. Las secuencias reguladoras pueden incluir promotores, secuencias líder de traducción, intrones, y secuencias de reconocimiento de poliadenilación.

Al promotor se le debe otorgar su significado común y habitual para una persona con experiencia corriente en la técnica, y se usa sin limitación para referirse a una secuencia de ADN capaz de controlar la expresión de una secuencia codificante o ARN funcional. En general, una secuencia codificante se ubica hacia el 3' de una secuencia promotora. Los promotores pueden derivarse en su totalidad de un gen nativo o pueden estar compuestos de diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso comprender segmentos sintéticos de ADN. Se entiende por los expertos en la técnica que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos de células o en diferentes etapas del desarrollo o en respuesta a diferentes condiciones ambientales. Los promotores, que hacen que un gen se exprese en la mayoría de los tipos de células en la mayoría de los casos, se conocen comúnmente como "promotores constitutivos." Se reconoce además que, dado que en la mayoría de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, los fragmentos de ADN de diferentes longitudes pueden tener una actividad promotora idéntica.

El término "unido operativamente" se refiere a la asociación de secuencias de ácido nucleico en un solo fragmento de ácido nucleico de manera que la función de una se afecta por la otra. Por ejemplo, un promotor está unido operativamente con una secuencia codificante cuando es capaz de afectar la expresión de esa secuencia codificante (es decir, que la secuencia codificante está bajo el control transcripcional del promotor). Las secuencias codificantes pueden estar unidas operativamente a secuencias reguladoras en orientación sentido o antisentido.

El término "expresión", como se usa en la presente descripción, debe otorgar a un experto en la técnica su significado común y habitual, y se usa sin limitación para referirse a la transcripción y acumulación estable de sentido (ARNm) o ARN antisentido derivado a partir del fragmento de ácido nucleico de la presente tecnología. "Sobreexpresión" se refiere a la producción de un producto génico en organismos transgénicos o recombinantes que excede los niveles de producción en organismos normales o no transformados.

"Transformación" debe recibir su significado ordinario y habitual para una persona con experiencia corriente en la técnica, y se usa sin limitación para referirse a la transferencia de un polinucleótido a una célula objetivo. El polinucleótido transferido puede incorporarse al genoma o al ADN cromosómico de una célula objetivo, lo que resulta en una herencia genéticamente estable, o puede replicarse independientemente del cromosoma del huésped. Los organismos huésped que contienen los fragmentos de ácido nucleico transformados se denominan organismos "transgénicos" o "recombinantes" o "transformados".

Los términos "transformado", "transgénico" y "recombinante", cuando se usan en la presente descripción en relación con las células huésped, deben recibir sus respectivos significados comunes y habituales para una persona con experiencia

corriente en la técnica, y se usan sin limitación para referirse a una célula de un organismo huésped, como una planta o célula microbiana, en la que se ha introducido una molécula de ácido nucleico heteróloga. La molécula de ácido nucleico puede integrarse de manera estable en el genoma de la célula huésped, o la molécula de ácido nucleico puede estar presente como una molécula extracromosómica. Dicha molécula extracromosómica puede ser autoreplicante. Se entiende que las células, tejidos o sujetos transformados abarcan no solo el producto final de un proceso de transformación, sino también su progenie transgénica.

Los términos "recombinante", "heterólogo" y "exógeno", cuando se usan en la presente descripción en relación con polinucleótidos, deben otorgar sus significados comunes y habituales a una persona con experiencia corriente en la técnica, y se usan sin limitación para referirse a un polinucleótido (por ejemplo, una secuencia de ADN o un gen) que se origina en una fuente extraña a la célula huésped en particular o, si es de la misma fuente, se modifica de su forma original. Por lo tanto, un gen heterólogo en una célula huésped incluye un gen que es endógeno respecto de la célula huésped en particular pero se ha modificado, por ejemplo, mediante el uso de mutagénesis sitio dirigida u otras técnicas recombinantes. Los términos incluyen además múltiples copias de origen no natural de una secuencia de ADN de origen natural. Por lo tanto, los términos se refieren a un segmento de ADN que es foráneo o heterólogo a la célula, u homólogo a la célula pero en una posición o forma dentro de la célula huésped en la que el elemento no se encuentra normalmente. De manera similar, los términos "recombinante", "heterólogo" y "exógeno", cuando se usan en la presente descripción en relación con un polipéptido o secuencia de aminoácidos, significa un polipéptido o secuencia de aminoácidos que se origina en una fuente extraña a la célula huésped particular o, si de la misma fuente, se modifica de su forma original. Por lo tanto, los segmentos de ADN recombinante pueden expresarse en una célula huésped para producir un polipéptido recombinante.

Los términos "plásmido", "vector" y "casete" deben recibir sus respectivos significados comunes y habituales para una persona con experiencia corriente en la técnica, y se usan sin limitación para referirse a un elemento cromosómico adicional que a menudo lleva genes que no son parte del metabolismo central de la célula, y generalmente en forma de moléculas de ADN de doble cadena circular. Dichos elementos pueden ser secuencias de replicación autónoma, secuencias de integración al genoma, secuencias de fagos o nucleótidos, lineales o circulares, de un ADN o ARN de cadena simple o doble, derivados de cualquier fuente, en la que se hayan unido o recombinado varias secuencias de nucleótidos en una construcción única que es capaz de introducir un fragmento de un promotor y una secuencia de ADN para un producto génico seleccionado junto con la secuencia 3' no traducible apropiada en una célula. "Casete de transformación" se refiere a un vector específico que contiene un gen extraño y que tiene elementos además del gen extraño que facilita la transformación de una célula huésped en particular. "Casete de expresión" se refiere a un vector específico que contiene un gen extraño y que tiene elementos además del gen extraño que permite una expresión mejorada de ese gen en un huésped extraño.

Las técnicas de recombinación estándar de ADN y de clonación molecular usadas en la presente descripción se conocen bien en la técnica y se describen, por ejemplo, por Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2da ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989 (en adelante, "Maniatis"); y por Silhavy, T.J., Bannan, M.L. y Enquist, L.W. *Experiments with Gene Fusions*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1984; y por Ausubel, F.M. y otros, *Current Protocols in Molecular Biology*, publicado por Greene Publishing and Wiley-Interscience, 1987.

A menos que se defina de cualquier otra manera, todos los términos técnicos y científicos usados en la presente descripción tienen el mismo significado que el comprendido comúnmente por el experto en la técnica a la que pertenece esta descripción. Si bien los métodos y materiales similares o equivalentes a aquellos descritos en la presente descripción pueden usarse en la práctica o en la prueba de la presente descripción, los materiales y métodos preferidos se describen a continuación.

La descripción se comprenderá más completamente después de considerar los siguientes ejemplos no limitantes. Debe entenderse que estos ejemplos, aunque indican modalidades preferidas de la presente tecnología, se dan solo a modo de ilustración.

Polipéptidos recombinantes

En un aspecto, la presente descripción se refiere a un polipéptido recombinante que tiene una secuencia de aminoácidos que tiene al menos un 70 %, al menos un 75 %, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % e incluso 100 % de identidad con la secuencia de aminoácidos expuesta en la sec. con núm. de ident.: 6. Adecuadamente, la secuencia de aminoácidos del polipéptido recombinante tiene al menos un 80 % de identidad con la sec. con núm. de ident.: 6. Más adecuadamente, la secuencia de aminoácidos del polipéptido recombinante tiene al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % e incluso 100 % de identidad con la sec. con núm. de ident.: 6. En un aspecto ilustrativo, la secuencia de aminoácidos del polipéptido recombinante consiste en la sec. con núm. de ident.: 6. En consecuencia, el polipéptido recombinante descrito en la presente descripción incluye fragmentos funcionales de la sec. con núm. de ident.: 6, variantes funcionales de la sec. con núm. de ident.: 6 y otros polipéptidos homólogos que tienen, por ejemplo, al menos un 80 %, al menos un 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al

menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % e incluso 100 % identidad de secuencia a sec. con núm. de ident.: 6.

5 En otro aspecto, la presente descripción se refiere a un ácido nucleico aislado que tiene una secuencia de nucleótidos que codifica el polipéptido recombinante descrito en la presente descripción. Por ejemplo, el ácido nucleico aislado puede incluir una secuencia de nucleótidos que codifica un polipéptido que tiene una secuencia de aminoácidos que tiene al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % e incluso 100 % de identidad con el amino secuencia ácida expuesta en la sec. con núm. de ident.: 6. Adecuadamente, el ácido nucleico aislado incluye una secuencia de nucleótidos que codifica un polipéptido que tiene una secuencia de aminoácidos que es al menos 80 % de identidad con la secuencia de aminoácidos expuesta en la sec. con núm. de ident.: 6. Más adecuadamente, el ácido nucleico aislado incluye una secuencia de nucleótidos que codifica un polipéptido que tiene una secuencia de aminoácidos que tiene al menos 85 %, al menos un 90 %, al menos un 91 %, al menos un 92 %, al menos un 93 %, al menos un 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, e incluso 100 % de identidad de secuencia con la secuencia de aminoácidos establecida en la sec. con núm. de ident.: 6. El ácido nucleico aislado, por lo tanto, incluye aquellos que codifican fragmentos funcionales de la sec. con núm. de ident.: 6, variantes funcionales de la sec. con núm. de ident.: 6 u otros polipéptidos homólogos que tienen, por ejemplo, al menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % e incluso 100 % identidad de secuencia con la sec. con núm. de ident.: 6.

25 En un aspecto, la presente descripción se refiere a un ácido nucleico aislado que tiene una secuencia de nucleótidos que tiene al menos un 70 %, al menos un 75 %, al menos 80 %, al menos 85 %, al menos un 90 %, al menos un 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % e incluso 100 % de identidad con la secuencia de nucleótidos expuesta en sec. con núm. de ident.: 7. Adecuadamente, el ácido nucleico aislado incluye una secuencia de nucleótidos que tiene al menos 80 % de identidad con la secuencia de nucleótidos expuesta en la sec. con núm. de ident.: 7. Más adecuadamente, el ácido nucleico aislado incluye una secuencia de nucleótidos que tiene al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, e incluso 100 % de identidad con la secuencia de ácido nucleico expuesta en la sec. con núm. de ident.: 7.

35 En otro aspecto, la presente tecnología se refiere a un vector que tiene los ácidos nucleicos descritos en la presente descripción, y a una célula huésped que tiene el vector descrito en la presente descripción. En algunos aspectos, la presente descripción se refiere a un vector de expresión que incluye al menos un polinucleótido de la presente tecnología y en el que el vector de expresión, tras la transfección en una célula huésped, es capaz de expresar al menos un polipéptido HV1 recombinante descrito en la presente descripción. En un aspecto, el vector de expresión incluye una secuencia de nucleótidos expuesta en la sec. con núm. de ident.: 7 o una variante de esta.

40 El diseño del vector de expresión depende de factores tales como la elección de la célula huésped a transformar, el nivel de expresión de la proteína deseada y similares. Los vectores de expresión pueden introducirse en la célula huésped para producir así el polipéptido recombinante de la presente tecnología, tal como el polipéptido HV1 recombinante que tiene una secuencia de aminoácidos de la sec. con núm. de ident.: 6 o una variante de la misma.

45 La expresión de proteínas en procariontes se realiza con mayor frecuencia en una célula huésped bacteriana con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de proteínas de fusión o no de fusión. Los vectores de fusión agregan una cantidad de aminoácidos a una proteína codificada en ellos, generalmente hacia el extremo amino de la proteína recombinante. Tales vectores de fusión típicamente sirven para tres propósitos: 1) para aumentar la expresión de proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante; y 3) para ayudar en la purificación de la proteína recombinante actuando como un ligando en la purificación por afinidad. A menudo, se introduce un sitio de escisión proteolítica en la unión del resto de fusión y la proteína recombinante para permitir la separación de la proteína recombinante del resto de fusión después de la purificación de la proteína de fusión. Tales vectores están dentro del alcance de la presente descripción.

55 En un aspecto, el vector de expresión incluye aquellos elementos genéticos para la expresión del polipéptido recombinante en células bacterianas. Los elementos para la transcripción y traducción en la célula bacteriana pueden incluir un promotor, una región codificante para el complejo de proteínas y un terminador de la transcripción.

60 En un aspecto, los vectores de expresión de la presente tecnología incluyen vectores de expresión bacterianos, por ejemplo, ADN de bacteriófago recombinante, ADN plásmido o ADN cósmido, vectores de expresión de levadura, por ejemplo, vectores de expresión de levadura recombinante, vectores para expresión en células de insecto, por ejemplo, vectores de expresión de virus recombinante, por ejemplo, baculovirus, o vectores para la expresión en células vegetales, por ejemplo, vectores de expresión de virus recombinantes tales como virus del mosaico de la coliflor (CaMV), virus del mosaico del tabaco (TMV), o vectores de expresión de plásmidos recombinantes tales como plásmidos Ti.

En un aspecto, el vector incluye un vector de expresión bacteriano. En otro aspecto, el vector de expresión incluye un vector de expresión de alto número de copias; alternativamente, el vector de expresión incluye un vector de expresión de bajo número de copias, por ejemplo, un plásmido Mini-F.

5 Un experto en la técnica conocerá las técnicas de biología molecular disponibles para la preparación de vectores de expresión. El polinucleótido usado para la incorporación en el vector de expresión de la presente tecnología, como se describe anteriormente, puede prepararse mediante técnicas rutinarias como la reacción en cadena de la polimerasa (PCR).

10 Se han desarrollado varias técnicas de biología molecular para vincular operativamente el ADN a vectores a través de extremos cohesivos complementarios. En un aspecto, pueden añadirse tratos homopolímeros complementarios a la molécula de ácido nucleico para insertarse en el vector ADN. El vector y la molécula de ácido nucleico luego se unen mediante enlaces de hidrógeno entre las colas homopoliméricas complementarias para formar moléculas de ADN recombinante.

15 En un aspecto alternativo, los enlazadores sintéticos que contienen uno o más sitios de restricción proporcionados se usan para enlazar operativamente el polinucleótido de la presente tecnología al vector de expresión. En un aspecto, el polinucleótido se genera por digestión con endonucleasas de restricción. En un aspecto, la molécula de ácido nucleico se trata con ADN polimerasa del bacteriófago T4 o ADN polimerasa I de *E. coli*, enzimas que eliminan los extremos 3' monocatenarios sobresalientes con sus actividades 3'-5'-exonucleolíticas, y rellenan los extremos 3' con sus actividades de polimerización, generan segmentos de ADN de extremos romos. Los segmentos de extremos romos se incuban luego con un gran exceso molar de moléculas enlazadoras en presencia de una enzima que es capaz de catalizar la ligación de las moléculas de ADN de extremos romos, como la ligasa de ADN del bacteriófago T4. Por lo tanto, el producto de la reacción es un polinucleótido que lleva secuencias enlazadoras poliméricas en sus extremos. Estos polinucleótidos se escinden luego con la enzima de restricción apropiada y se ligan a un vector de expresión que se ha escindido con una enzima que produce extremos compatibles con los del polinucleótido.

20 Alternativamente, puede emplearse un vector que tenga sitios de clonación independientes de la ligación (LIC). El polinucleótido amplificado por PCR requerido puede luego clonarse en el vector LIC sin digestión de restricción o ligación (Aslanidis y de Jong, Nucl. Acid. Res. 18, 6069-6074, (1990), Haun, y otros, Biotechniques 13, 515-518 (1992)).

25 En un aspecto, para aislar y/o modificar el polinucleótido de interés para la inserción en el plásmido elegido, es adecuado usar PCR. Pueden diseñarse cebadores apropiados para usar en la preparación de PCR de la secuencia para aislar la región codificante requerida de la molécula de ácido nucleico, agregar endonucleasas de restricción o sitios LIC, colocar la región codificante en el marco de lectura deseado.

30 En un aspecto, se prepara un polinucleótido para su incorporación en un vector de expresión de la presente tecnología mediante el uso de PCR y cebadores oligonucleotídicos apropiados. La región codificante se amplifica, mientras que los propios cebadores se incorporan al producto de secuencia amplificada. En un aspecto, los cebadores de amplificación contienen sitios de reconocimiento de endonucleasas de restricción, que permiten que el producto de secuencia amplificada se clone en un vector apropiado.

35 En un aspecto, el polinucleótido de sec. con núm. de ident.: 7 o una variante de la misma se obtiene por PCR y se introduce en un vector de expresión usando la digestión y la ligación con endonucleasas de restricción de acuerdo con técnicas que se conocen bien en la técnica.

40 La presente descripción se refiere además a una célula huésped que comprende el vector de expresión descrito en la presente descripción. Los huéspedes adecuados de la presente tecnología incluyen típicamente huésped microbianos o huésped de plantas. Por ejemplo, la célula huésped de la presente tecnología se selecciona del grupo que consiste en bacterias, levaduras, hongos filamentosos, algas cianobacterias y células vegetales.

45 Los huéspedes microbianos pueden incluir cualquier organismo capaz de expresar el polinucleótido (tal como la sec. con núm. de ident.: 7) para producir el polipéptido HV1 recombinante descrito en la presente descripción. Los microorganismos útiles en la presente tecnología incluyen bacterias, tales como las bacterias entéricas (*Escherichia* y *Salmonella* por ejemplo), así como *Bacillus*, *Acinetobacter*, *Actinomycetes* como *Streptomyces*, *Corynebacterium*, *Methanotrophs*, tales como *Methylosinus*, *Methylomonas*, *Rhodococcus* y *Pseudomonas*; Cianobacterias, tales como *Rhodobacter* y *Synechocystis*; levaduras, tales como *Saccharomyces*, *Zygosaccharomyces*, *Kluyveromyces*, *Candida*, *Hansenula*, *Debaryomyces*, *Mucor*, *Pichia* y *Torulopsis*; y hongos filamentosos, tales como *Aspergillus* y *Arthrobotrys*, y algas, y *Escherichia*, *Klebsiella*, *Pantoea*, *Salmonella*, *Corynebacterium*, *Clostridium* y *Clostridium acetobutylicum*, por ejemplo. Preferentemente, el huésped microbiano es una bacteria (como *Escherichia*) o una levadura (como *Saccharomyces*). Los vectores de expresión pueden incorporarse en estos y otros huéspedes microbianos para preparar grandes cantidades comercialmente útiles de glucósidos de esteviol.

50 En un aspecto, el polipéptido recombinante puede expresarse en una célula huésped que es una célula vegetal. Como se usa en la presente descripción, el término "célula vegetal" se entiende que significa cualquier célula derivada de una planta monocotiledónea o dicotiledónea y capaz de constituir tejidos no diferenciados tales como callos, tejidos diferenciados

tales como embriones, porciones de plantas monocotiledóneas, plantas monocotiledóneas o semillas. El término "planta" se entiende que significa cualquier organismo multicelular diferenciado capaz de la fotosíntesis, incluyendo monocotiledóneas y dicotiledóneas. En algunos aspectos, la célula vegetal puede ser una célula vegetal de Arabidopsis, una célula vegetal de tabaco, una célula vegetal de soya, una célula vegetal de petunia o una célula de otro cultivo de semillas oleaginosas que incluye, entre otros, una célula vegetal de canola, una célula vegetal de colza, una célula vegetal de palma, una célula vegetal de girasol, una célula vegetal de algodón, una célula vegetal de maíz, una célula vegetal de cacahuete, una célula vegetal de lino y una célula vegetal de sésamo.

Los huéspedes de plantas útiles pueden incluir cualquier planta que soporte la producción de los polipéptidos recombinantes de la presente tecnología. Las plantas verdes adecuadas para usar como huésped incluyen, pero no se limitan a, soya, colza (*Brassica napus*, *B. campestris*), girasol (*Helianthus annuus*), algodón (*Gossypium hirsutum*), maíz, tabaco (*Nicotiana tabacum*), alfalfa (*Medicago sativa*), trigo (*Triticum sp*), cebada (*Hordeum vulgare*), avena (*Avena sativa*), sorgo (*Sorghum bicolor*), arroz (*Oryza sativa*), Arabidopsis, vegetales crucíferos (brócoli, coliflor, col, parsnips, etc.), melones, zanahorias, apio, perejil, tomates, papas, fresas, cacahuetes, uvas, cultivos de semillas de hierba, remolacha azucarera, caña de azúcar, frijoles, guisantes, centeno, lino, árboles de madera dura, árboles de madera blanda y pastos forrajeros. Las especies de algas incluyen, pero no se limitan a, huésped comercialmente significativos, como la espirulina, *Haemotacoccus* y *Dunaliella*. Las plantas adecuadas para el método de la presente tecnología también incluyen biocombustible, biomasa y cultivos bioenergéticos. Las plantas ilustrativas incluyen Arabidopsis thaliana, arroz (*Oryza sativa*), *Hordeum vulgare*, pasto varilla (*Panicum virgatum*), *Brachypodium spp*, *Brassica spp.* y *Crambe abyssinica*.

En algunos aspectos, la presente descripción incluye células huésped transgénicas o huésped que se han transformado con uno o más de los vectores descritos en la presente descripción.

Alternativamente, las células huésped pueden ser aquellas adecuadas para la producción de biosíntesis, que incluyen organismos de una sola célula, microorganismos, organismos multicelulares, plantas, hongos, bacterias, algas, cultivos cultivados, cultivos no cultivados y/o similares.

Los vectores de expresión pueden introducirse en células huésped microbianas o de plantas mediante técnicas convencionales de transformación o transfección. La transformación de las células apropiadas con un vector de expresión de la presente tecnología se realiza mediante métodos conocidos en la técnica y típicamente depende tanto del tipo de vector como de la célula. Las técnicas adecuadas incluyen coprecipitación con fosfato de calcio o cloruro de calcio, transfección mediada por DEAE-dextrano, lipofección, quimioporación o electroporación.

Las células transformadas con éxito, es decir, aquellas células que contienen el vector de expresión, pueden identificarse mediante técnicas bien conocidas en la técnica. Por ejemplo, las células transfectadas con un vector de expresión de la presente tecnología pueden cultivarse para producir polipéptidos descritos en la presente descripción. Las células pueden examinarse para determinar la presencia del ADN del vector de expresión mediante técnicas bien conocidas en la técnica. Las células huésped pueden contener una sola copia del vector de expresión descrito anteriormente, o alternativamente, copias múltiples del vector de expresión.

En algunos aspectos, la célula transformada es una célula animal, una célula de insecto, una célula vegetal, una célula de algas, una célula fúngica o una célula de levadura. En algunos aspectos, la célula es una célula vegetal seleccionada del grupo que consiste en: célula vegetal de canola, una célula vegetal de colza, una célula vegetal de palma, una célula vegetal de girasol, una célula vegetal de algodón, una célula vegetal de maíz, una célula vegetal de maní, una célula vegetal de lino, una célula vegetal de sésamo, una célula vegetal de soya y una célula vegetal de petunia.

Los expertos en la técnica conocen bien los sistemas de expresión de la célula huésped microbiana y los vectores de expresión que contienen secuencias reguladoras que dirigen un alto nivel de expresión de proteínas exógenas. Cualquiera de estos podría usarse para construir vectores para la expresión del polipéptido recombinante de la presente tecnología en una célula huésped microbiana. Estos vectores podrían introducirse después en microorganismos apropiados mediante transformación para permitir una expresión de alto nivel del polipéptido recombinante de la presente tecnología.

Los vectores o casetes útiles para la transformación de células huésped microbianas adecuadas se conocen bien en la técnica. Típicamente, el vector o casete contiene secuencias que dirigen la transcripción y traducción del polinucleótido relevante, un marcador seleccionable, y secuencias que permiten la replicación autónoma o la integración cromosómica. Los vectores adecuados comprenden una región 5' del polinucleótido que alberga controles del inicio de la transcripción y una región 3' del fragmento de ADN que controla la terminación de la transcripción. Se prefiere que ambas regiones de control se deriven de genes homólogos de la célula huésped transformada, aunque debe entenderse que tales regiones de control no necesitan derivarse de los genes nativos de las especies específicas elegidas como huésped.

Las regiones de control del inicio o los promotores, que son útiles para dirigir la expresión del polipéptido recombinante en la célula huésped microbiana deseada, son numerosos y comunes para los expertos en la técnica. Prácticamente cualquier promotor capaz de conducir estos genes es adecuado para la presente tecnología, incluidos, entre otros, CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI (útil para la expresión en *Saccharomyces*); AOX1 (útil para la expresión en *Pichia*); y lac, trp, IPL, IPR, T7, tac y trc (útiles para la expresión en *Escherichia coli*).

Las regiones de control de terminación también pueden derivarse de diversos genes nativos de los huéspedes microbianos. Opcionalmente, puede incluirse un sitio de terminación para los huéspedes microbianos descritos en la presente descripción.

5 En células vegetales, los vectores de expresión de la presente tecnología pueden incluir una región codificante unida operativamente a promotores capaces de dirigir la expresión del polipéptido recombinante de la presente tecnología en los tejidos deseados en la etapa deseada de desarrollo. Por razones de conveniencia, los polinucleótidos a expresar pueden comprender secuencias promotoras y secuencias líder de traducción derivadas del mismo polinucleótido. Las secuencias 3' no codificantes que codifican las señales de terminación de la transcripción también deben estar presentes.
10 Los vectores de expresión también pueden comprender uno o más intrones para facilitar la expresión de polinucleótidos.

Para las células huésped vegetales, cualquier combinación de cualquier promotor y cualquier terminador capaz de inducir la expresión de una región codificante puede usarse en las secuencias de vectores de la presente tecnología. Algunos ejemplos adecuados de promotores y terminadores incluyen los de los genes de la nopalina sintasa (nos), la octopina sintasa (ocs) y el virus del mosaico de la coliflor (CaMV). Un tipo de promotor de plantas eficiente que puede usarse es un promotor de plantas de alto nivel. Dichos promotores, en un enlace operable con un vector de expresión de la presente tecnología, deberían ser capaces de promover la expresión del vector. Los promotores de plantas de alto nivel que pueden usarse en la presente tecnología incluyen el promotor de la subunidad pequeña de la ribulosa-1,5-bisfosfato carboxilasa (ss), por ejemplo, de la soya (Berry-Lowe y otros, *J. Molecular y App. Gen.*, 1: 483 498 (1982)), y el promotor de la proteína de unión a clorofila a/b. Se conoce que estos dos promotores son inducidos por la luz en células vegetales (ver, por ejemplo, *Genetic Engineering of Plants, una Perspectiva Agrícola*, A. Cashmore, Plenum, NY (1983), páginas 29 38; Coruzzi, G. y otros, *The Journal of Biological Chemistry*, 258: 1399 (1983), y Dunsmuir, P. y otros, *Journal of Molecular and Applied Genetics*, 2: 285 (1983)).

25 La elección del vector plasmídico depende del método que se utilizará para transformar las plantas huésped. El experto en la técnica conoce bien los elementos genéticos que deben estar presentes en el vector plasmídico para transformar, seleccionar y propagar con éxito las células huésped que contienen el polinucleótido quimérico. El experto en la técnica también reconocerá que diferentes eventos de transformación independientes darán lugar a diferentes niveles y patrones de expresión (Jones y otros, *EMBO J.* 4: 2411 2418 (1985); De Almeida y otros, *Mol. Gen. Genetics* 218: 78 86 (1989)), y por lo tanto, deben analizarse múltiples eventos para obtener líneas que muestren el nivel de expresión y el patrón deseados. Dicha selección puede realizarse mediante el análisis Southern de las transferencias de ADN, el análisis Northern de la expresión del ARNm, el análisis Western de la expresión de proteínas o el análisis fenotípico.

35 La introducción del vector de expresión de la presente tecnología en una célula vegetal puede realizarse mediante una variedad de métodos conocidos por los expertos en la técnica, incluida la inserción de una secuencia de ácido nucleico de interés en un plásmido de *Agrobacterium rhizogenes* Ri o *Agrobacterium tumefaciens* Ti. Microinyección, electroporación o precipitación directa. A modo de ejemplo, en algunos aspectos, la expresión transitoria de un polinucleótido de interés puede realizarse mediante métodos de infiltración. En este sentido, una suspensión de *Agrobacterium tumefaciens* que contiene un polinucleótido de interés puede cultivarse y luego inyectarse en una planta colocando la punta de una jeringa contra el lado inferior de una hoja mientras se aplica una contrapresión suave en el otro lado de la hoja. La solución de *Agrobacterium* luego se inyecta en los espacios aéreos dentro de la hoja a través de los estomas. Una vez dentro de la hoja, el *Agrobacterium* transforma el gen de interés en una porción de las células de la planta donde el gen se expresa de forma transitoria.

45 Como otro ejemplo, la transformación de un plásmido de interés en una célula vegetal puede realizarse mediante técnicas de bombardeo con pistola de partículas (es decir, biolística). Con respecto a esto, una suspensión de embriones de plantas puede crecer en cultivo líquido y luego bombardearse con plásmidos o polinucleótidos que están unidos a partículas de oro, en donde las partículas de oro unidas al plásmido o al ácido nucleico de interés pueden ser propulsadas a través de las membranas de tejidos vegetales, como el tejido embrionario. Después del bombardeo, los embriones transformados pueden seleccionarse usando un antibiótico apropiado para generar nuevos cultivos de suspensión embriogénicos transformados, propagados clonalmente.

Las células huésped pueden ser células o líneas celulares no modificadas, o líneas celulares modificadas genéticamente. En algunos aspectos, la célula huésped es una línea celular que se ha modificado para permitir el crecimiento en condiciones deseadas, como a una temperatura más baja.
55

Pueden usarse metodologías estándar de ADN recombinante para obtener un ácido nucleico que codifica un polipéptido recombinante descrito en la presente descripción, incorporar el ácido nucleico en un vector de expresión e introducir el vector en una célula huésped, como las descritas en Sambrook, y otros (edición) *Molecular Cloning; A Laboratory Manual*, tercera edición, Cold Spring Harbor, (2001); y Ausubel, F. M. y otros (eds) *Current Protocols in Molecular Biology*, John Wiley Sons (1995). Un ácido nucleico que codifica un polipéptido puede insertarse en un vector o vectores de expresión de manera que los ácidos nucleicos estén unidos operativamente a secuencias de control de transcripción y traducción (como una secuencia promotora, una secuencia de terminación de la transcripción, etcétera). Las secuencias del vector de expresión y control de expresión se eligen generalmente para ser compatibles con la célula huésped de expresión que se usa.
60
65

La expresión del polipéptido en un huésped descrito en la presente descripción puede mejorarse adicionalmente mediante la optimización de codones. Por ejemplo, la modificación de un codón menos común con un codón más común puede afectar la vida media del ARNm o alterar su estructura mediante la introducción de una estructura secundaria que interfiere con la traducción del mensaje. Puede optimizarse toda o una parte de una región codificante. En algunos casos, la modulación de la expresión deseada se logra optimizando esencialmente todo el gen. En otros casos, la modulación deseada se logrará optimizando parte de la secuencia del gen, pero no la totalidad.

El uso del codón de cualquier secuencia codificante puede ajustarse para lograr una propiedad deseada, por ejemplo, altos niveles de expresión en un tipo de célula específico. El punto de partida para tal optimización puede ser una secuencia codificante con 100 % de codones comunes, o una secuencia codificante que contenga una mezcla de codones comunes y no comunes.

Pueden generarse y probar dos o más secuencias candidatas que difieren en su uso de codones para determinar si poseen la propiedad deseada. Las secuencias candidatas pueden evaluarse usando un ordenador para buscar la presencia de elementos regulatorios, como silenciadores o potenciadores, y para buscar la presencia de regiones de secuencia codificante que podrían convertirse en dichos elementos regulatorios mediante una alteración en el uso del codón. Los criterios adicionales pueden incluir el enriquecimiento de nucleótidos particulares, por ejemplo.

A, C, G o U, preferencia códonica para un aminoácido particular, o la presencia o ausencia de una estructura secundaria o terciaria de ARNm particular. El ajuste a la secuencia candidata puede realizarse en función de una serie de criterios de este tipo.

En ciertos aspectos, la secuencia de ácido nucleico optimizada en codones puede expresar su proteína, a un nivel que es aproximadamente 110 %, aproximadamente 150 %, aproximadamente 200 %, aproximadamente 250 %, aproximadamente 300 %, aproximadamente 350 %, aproximadamente 400 %, aproximadamente 450 %, o aproximadamente el 500 %, de la expresada por una secuencia de ácido nucleico que no ha sido optimizada por codones.

Además del ácido nucleico que codifica el polipéptido recombinante de la presente tecnología, el vector de expresión de la presente tecnología puede llevar adicionalmente secuencias reguladoras que controlan la expresión de la proteína en una célula huésped, como promotores, potenciadores u otros elementos de control de la expresión que controlan la transcripción o traducción de los ácidos nucleicos. Tales secuencias reguladoras son conocidas en la técnica. Los expertos en la técnica apreciarán que el diseño del vector de expresión, incluida la selección de secuencias reguladoras, puede depender de factores tales como la elección de la célula huésped que se va a transformar, el nivel de expresión de la proteína deseada, etcétera. Además, los vectores de expresión recombinantes de la presente tecnología pueden llevar secuencias adicionales, como secuencias que regulan la replicación del vector en las células huésped (por ejemplo, orígenes de replicación) y genes marcadores de selección.

Biosíntesis de los glucósidos de esteviol

Como se describe en la presente descripción, los polipéptidos recombinantes de la presente tecnología tienen actividades de UDP-glucosiltransferasa, que incluyen más particularmente una actividad de glucosilación de 1,2-19-O-glucosa, y son útiles para desarrollar métodos biosintéticos para preparar glucósidos de esteviol que son típicamente de baja abundancia en fuentes naturales, como el rebaudiósido D y el rebaudiósido E. Los polipéptidos recombinantes de la presente tecnología tienen actividades de UDP-glucosiltransferasa, son útiles para desarrollar métodos biosintéticos para preparar nuevos glucósidos de esteviol, como el rebaudiósido Z1 y el rebaudiósido Z2.

En consecuencia, en un aspecto, la presente invención proporciona un método para producir una composición de glucósido de esteviol, incluyendo el método de incubar un sustrato con un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.: 6.

El sustrato puede ser cualquier compuesto natural o sintético capaz de convertirse en un compuesto de glucósido de esteviol en una reacción catalizada por una o más UDP-glucosiltransferasas. Por ejemplo, el sustrato puede ser extracto de estevia natural, esteviol, esteviol-13-O-glucósido, esteviol-19-O-glucósido, esteviol-1,2-biósidio, rubusosido, esteviósido, rebaudiósido A, rebaudiósido G o rebaudiósido E. El sustrato puede ser un compuesto puro o una mezcla de diferentes compuestos. Preferentemente, el sustrato incluye un compuesto que se selecciona del grupo que consiste en rubusosido, esteviósido, rebaudiósido A, rebaudiósido E y sus combinaciones.

El método descrito en la presente descripción también proporciona un sistema de reacción de acoplamiento en el que los péptidos recombinantes descritos en la presente descripción pueden funcionar en combinación con una o más enzimas adicionales para mejorar la eficiencia o modificar el resultado de la biosíntesis global de compuestos de glucósido de esteviol. Por ejemplo, la enzima adicional puede regenerar la UDP-glucosa necesaria para la reacción de glucosilación convirtiendo el UDP producido a partir de la reacción de glucosilación de nuevo a UDP-glucosa (usando, por ejemplo, sacarosa como donante del residuo de glucosa), mejorando así la eficacia de la reacción de glucosilación. En otro ejemplo, el polipéptido recombinante de la presente tecnología puede producir un producto de glucósido de esteviol intermedio (por ejemplo, rebaudiósido E), que se convierte adicionalmente en otro glucósido de esteviol (por ejemplo, rebaudiósido D) en una reacción catalizada por otra UDP-glucosiltransferasa, tal como UGT76G1. En otro ejemplo, el polipéptido

recombinante de la presente tecnología puede producir un producto de glucósido de esteviol intermedio (por ejemplo, rebaudiósido E), que se convierte adicionalmente en otro glucósido de esteviol (por ejemplo, rebaudiósido Z1 y rebaudiósido Z2) en una reacción catalizada por UDP-glucosiltransferasa, como HV1.

5 En consecuencia, en una modalidad, el método de la presente tecnología incluye además incubar una sacarosa sintasa (SUS) recombinante con el sustrato y el polipéptido recombinante descrito en la presente descripción. La sacarosa sintasa recombinante convierte UDP en UDP-glucosa usando la sacarosa como fuente de glucosa. La sacarosa sintasa adecuada incluye aquellos genes SUS derivados de *Arabidopsis thaliana* y *Vigna radiate*, o de cualquier gen que codifique un homólogo funcional de la sacarosa sintasa codificada por la secuencia SUS1 de *Arabidopsis thaliana* y *Vigna radiate*, o los homólogos funcionales de la misma. La sacarosa sintasa adecuada puede ser, por ejemplo, una sacarosa sintasa 1 de *Arabidopsis*; una sacarosa sintasa 3 de *Arabidopsis*; y una sacarosa sintasa de *Vigna radiate*. Una sacarosa sintasa particularmente adecuada puede ser, por ejemplo, sacarosa sintasa 1 de *Arabidopsis*. Por ejemplo, el SUS recombinante incluye una secuencia de aminoácidos que tiene al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, a al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % de identidad con la secuencia de aminoácidos de AtSUS1 expuesta en la sec. con núm. de ident.: 9. Preferentemente, la SUS recombinante de la presente tecnología incluye una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de AtSUS1 expuesta en la sec. con núm. de ident.: 9.

20 La sacarosa sintasa recombinante de la presente tecnología puede obtenerse mediante la expresión de un ácido nucleico que tiene una secuencia de nucleótidos que codifica una secuencia de aminoácidos de interés (por ejemplo, una que tiene al menos 80 % de identidad con la secuencia de aminoácidos expuesta en la sec. con núm. de ident.: 9) en una célula huésped como se describió anteriormente. Por ejemplo, un vector que incluye una secuencia de nucleótidos expuesta en la sec. con núm. de ident.: 10 puede introducirse en un huésped microbiano (como *E. coli*) mediante técnicas de transformación convencionales para producir la sacarosa sintasa recombinante.

En otra modalidad, el método de la presente tecnología incluye además incubar una UDP-glucosiltransferasa recombinante con la sacarosa sintasa recombinante, el sustrato y el polipéptido recombinante descrito en la presente descripción. La UDP-glucosiltransferasa recombinante puede catalizar una reacción de glucosilación diferente a la catalizada por el polipéptido recombinante de la presente tecnología. Por ejemplo, la UDP-glucosiltransferasa recombinante puede catalizar la reacción que transfiere un resto de azúcar al C-3' de la C-13-O-glucosa de esteviósido para producir rebaudiósido A (o de manera similar para producir rebaudiósido D a partir de rebaudiósido E), mientras que el polipéptido recombinante de la presente tecnología transfiere un segundo resto de azúcar al C-2' de 19-O-glucosa de esteviósido para producir rebaudiósido E (o de manera similar para producir rebaudiósido D a partir de rebaudiósido A).

35 La UDP-glucosiltransferasa adecuada incluye cualquier UGT conocida en la técnica que capaz de catalizar una o más reacciones en la biosíntesis de compuestos de glucósido de esteviol, tales como UGT85C2, UGT74G1, UGT76G1, o los homólogos funcionales de los mismos. Por ejemplo, la UDP-glucosiltransferasa como se describe en la presente descripción puede incluir una secuencia de aminoácidos que tiene al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 91 %, al menos 92 %, al menos 93 %, al menos 94 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o incluso 100 % de identidad con la secuencia de aminoácidos de UGT76G1 establecido en la sec. con núm. de ident.: 11. Preferentemente, la UDP-glucosiltransferasa incluye una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de UGT76G1 expuesta en la sec. con núm. de ident.: 11.

45 La UDP-glucosiltransferasa recombinante puede obtenerse expresando un ácido nucleico que tiene una secuencia de nucleótidos que codifica una secuencia de aminoácidos de interés (por ejemplo, una que tiene al menos 80 % de identidad con la secuencia de aminoácidos expuesta en la sec. con núm. de ident.: 11) en una célula huésped como se describió anteriormente. Por ejemplo, un vector que incluye una secuencia de nucleótidos expuesta en la sec. con núm. de ident.: 12 puede introducirse en un huésped microbiano (como *E. coli*) mediante técnicas de transformación convencionales para producir la UDP-glucosiltransferasa recombinante de la presente tecnología.

Tanto la producción *in vitro* como *in vivo* de compuestos de glucósidos de esteviol son abarcadas por la presente tecnología.

55 Por ejemplo, la incubación puede ser un proceso *in vitro*, en el que se permite que un sustrato interactúe con un polipéptido recombinante de la presente tecnología. Preferentemente, en un proceso *in vitro*, el polipéptido recombinante se purifica antes de incubarse con el sustrato. En los métodos de la presente tecnología se incluyen técnicas de purificación de polipéptidos convencionales, tales como centrifugación, lisado celular y cromatografía. Por ejemplo, el ácido nucleico que codifica el polipéptido recombinante de la presente tecnología puede clonarse en un vector de expresión con una etiqueta de histidina, de modo que el polipéptido recombinante expresado pueda purificarse mediante cromatografía de afinidad en columna.

65 El método *in vitro* de la presente tecnología incluye cualquier sistema tampón que sea adecuado para la producción de glucósidos de esteviol utilizando uno o más polipéptidos recombinantes de la presente tecnología. Típicamente, el sistema de tampón es una solución acuosa, como el tampón Tris, el tampón HEPES, el tampón MOPS, el tampón fosfato, con un

pH de aproximadamente 6,0 a aproximadamente 8,0. Más adecuadamente, el pH es de aproximadamente 6,5 a aproximadamente 7,5. Aún más adecuadamente, el pH es de aproximadamente 7,0 a aproximadamente 7,5.

5 Normalmente, en el método *in vitro* de la presente tecnología, el sustrato está presente en el tampón a una concentración de aproximadamente 0,2 mg/ml a aproximadamente 5 mg/ml, preferentemente de aproximadamente 0,5 mg/ml a aproximadamente 2 mg/ml, más preferentemente de aproximadamente 0,7 mg/ml a aproximadamente 1,5 mg/ml.

10 Normalmente, en el método *in vitro* de la presente tecnología, se incluye UDP-Glucosa en el tampón a una concentración de aproximadamente 0,2 mM a aproximadamente 5 mM, preferentemente de aproximadamente 0,5 mM a aproximadamente 2 mM, más preferentemente de aproximadamente 0,7 mM a alrededor de 1,5 mM. En un aspecto, cuando se incluye una sacarosa sintasa recombinante en la reacción, también se incluye sacarosa en el tampón a una concentración de aproximadamente 100 mM a aproximadamente 500 mM, preferentemente de aproximadamente 200 mM a aproximadamente 400 mM, con mayor preferencia de aproximadamente 250 mM a aproximadamente 350 mM.

15 Típicamente, en el método *in vitro* de la presente tecnología, la relación en peso del polipéptido recombinante al sustrato, sobre una base en peso seco, es de aproximadamente 1:100 a aproximadamente 1:5, preferentemente de aproximadamente 1:50 a aproximadamente 1:10, más preferentemente de aproximadamente 1:25 a aproximadamente 1:15.

20 Típicamente, la temperatura de reacción del método *in vitro* es de aproximadamente 20 °C a aproximadamente 40, °C, adecuadamente de 25, °C a aproximadamente 37, °C, más adecuadamente de 28, °C a aproximadamente 32, °C.

25 La presente descripción también proporciona una composición de glucósido de esteviol producida por el método biosintético descrito en la presente descripción. La naturaleza exacta de la composición de glucósido de esteviol producida usando el método descrito en la presente descripción, como los tipos de especies moleculares y su contenido porcentual en el producto final, depende del sustrato utilizado, las condiciones de incubación y las actividades enzimáticas incluidas en el sistema de reacción. Por ejemplo, cuando se usa esteviósido como sustrato, la composición de glucósido de esteviol producida puede incluir rebaudiósido A, rebaudiósido D, rebaudiósido E, rebaudiósido Z1 y rebaudiósido Z2 y sus combinaciones.

30 La presente tecnología proporciona un método para convertir las especies predominantes de glucósidos de esteviol (es decir, esteviósido y rebaudiósido A) en el extracto de estevia natural en rebaudiósido D y rebaudiósido E, que de otro modo son de baja abundancia en el extracto natural. En consecuencia, la presente tecnología también proporciona un método para enriquecer el contenido de uno o más glucósidos de esteviol específicos (como el rebaudiósido D y el rebaudiósido E), el método incluye la incubación de un sustrato (como un extracto de estevia natural) con un polipéptido recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.: 6. Por ejemplo, cuando se usa el extracto de estevia natural como sustrato, la composición de glucósido de esteviol producida puede enriquecerse con rebaudiósido D y/o rebaudiósido E, que son de baja abundancia en el extracto de estevia natural.

40 Un experto en la técnica reconocerá que la composición de glucósido de esteviol producida por el método descrito en la presente descripción puede purificarse adicionalmente y mezclarse con otros glucósidos, sabores o edulcorantes de esteviol para obtener un sabor o una composición edulcorante deseada. Por ejemplo, una composición enriquecida con rebaudiósido D producida como se describe en la presente descripción puede mezclarse con un extracto de estevia natural que contiene rebaudiósido A como el glucósido de esteviol predominante, o con otros productos de glucósido de esteviol sintéticos o naturales para obtener una composición edulcorante deseada. Alternativamente, un glucósido de esteviol sustancialmente purificado (por ejemplo, rebaudiósido D) obtenido a partir de la composición de glucósido de esteviol descrita en la presente descripción puede combinarse con otros edulcorantes, como sacarosa, maltodextrina, aspartamo, sucralosa, neotamo, acesulfamo de potasio y sacarina. La cantidad de glucósido de esteviol con relación a otros edulcorantes puede ajustarse para obtener un sabor deseado, como se conoce en la técnica. La composición de glucósido de esteviol descrita en la presente descripción (incluyendo rebaudiósido D, rebaudiósido E, rebaudiósido Z1, rebaudiósido Z2 o una combinación de los mismos) puede incluirse en productos alimenticios (como bebidas, refrescos, helados, productos lácteos, confitería, cereales, chicles, productos horneados, etcétera), suplementos dietéticos, nutrición médica, así como también productos farmacéuticos.

55 **EJEMPLOS**

EJEMPLO 1: Selección de genes UGT candidatos

60 Se utilizaron análisis filogenéticos y BLAST de proteínas para identificar 7 genes candidatos que pertenecen a la subfamilia UGT91 para la actividad de glucosilación de 1,2-19-O-glucosa (Tabla 1).

Tabla 1: Lista de genes candidatos UGT

65

Nombre	Descripción	Acceso	Identidad de secuencia
BD1	PREDICHO: Tipo UDP-glucosiltransferasa 91C1 [Brachypodium distachyon]	XP_003560664.1	sec. con núm. de ident.: 1
BD2	PREDICHO: Tipo UDP-glucosiltransferasa 91C1 [Brachypodium distachyon]	XP_003560669.1	sec. con núm. de ident.: 2
BD3	PREDICHO: PROTEÍNA DE BAJA CALIDAD: Tipo UDP-glucosiltransferasa 91C1 [Brachypodium distachyon]	XP_003581636.1	sec. con núm. de ident.: 3
BD4	PREDICHO: Tipo UDP-glucosiltransferasa 91C1 [Brachypodium distachyon]	XP_003580515.1	sec. con núm. de ident.: 4
BD5	PREDICHO: PROTEÍNA DE BAJA CALIDAD: Tipo UDP-glucosiltransferasa 91B1 [Brachypodium distachyon]	XP_003559500.1	sec. con núm. de ident.: 5
HV1	proteína predicha [Hordeum vulgare subsp. vulgare]	BAJ98242.1	sec. con núm. de ident.: 6
HV2	proteína predicha [Hordeum vulgare subsp. vulgare]	BAJ93155.1	sec. con núm. de ident.: 8

Ejemplo 2: Cribado de la actividad enzimática de los genes UGT candidatos

Los fragmentos de ADN de longitud completa de todos los genes UGT candidatos se sintetizaron comercialmente. Casi todos los codones del ADNc se cambiaron a los preferidos para *E. coli* (Genscript, NJ). El ADN sintetizado se clonó en un vector de expresión bacteriano pETite N-His SUMO Kan Vector (Lucigen).

Cada construcción de expresión se transformó en *E. coli* BL21 (DE3), que subsecuentemente se cultivó en medio LB que contenía 50 µg/ml de kanamicina a 37 °C hasta alcanzar una OD₆₀₀ de 0,8-1,0. La expresión de la proteína se indujo mediante la adición de 1 mM de isopropil β-D-1-tiogalactopiranosido (IPTG) y el cultivo se cultivó a 16 °C durante 22 h. Las células se colectaron por centrifugación (3.000 xg; 10 min; 4 °C). Los sedimentos celulares se colectaron y se usaron inmediatamente o se almacenaron a -80 °C.

Los sedimentos celulares típicamente se resuspendieron en tampón de lisis (tampón de fosfato de potasio 50 mM, pH 7,2, lisozima 25 ml/ml, DNasa I 5 ml/ml, imidazol 20 mM, NaCl 500 mM, glicerol al 10 % y 0,4% Triton X-100). Las células se rompieron por sonicación a 4 °C, y el debris celular se clarificó por centrifugación (18,000 x g; 30 min). El sobrenadante se cargó en un equilibrado (tampón de equilibrio: Tampón de fosfato de potasio 50 mM, pH 7,2, imidazol 20 mM, NaCl 500 mM, glicerol al 10 %) en la columna de afinidad Ni-NTA (Qiagen). Después de cargar la muestra de proteína, la columna se lavó con tampón de equilibrio para eliminar las proteínas contaminantes no unidas. Los polipéptidos UGT recombinantes etiquetados con His se eluyeron mediante un tampón de equilibrio que contenía imidazol 250 mM.

Los polipéptidos recombinantes UGT candidatos purificados se analizaron para determinar la actividad de la glucosilación 1,2-19-O-glucosa utilizando esteviósido o Reb A como sustrato (Figura 1). Normalmente, el polipéptido recombinante (10 µg) se probó en un sistema de reacción *in vitro* de 200 µl. El sistema de reacción contiene un tampón de fosfato de potasio 50 mM, pH 7,2, MgCl₂ 3 mM, 1 mg/ml de esteviósido o rebaudiósido A, UDP-glucosa 1 mM. La reacción se realizó a 30 °C y se terminó agregando 200 µL de 1-butanol. Las muestras se extrajeron tres veces con 200 µL de 1-butanol. La fracción combinada se secó y se disolvió en 70 µl de metanol al 80 % para un análisis de cromatografía líquida de alto rendimiento (HPLC). El extracto de estevia (Blue California, CA), que contenía 95 % de esteviósido, se usó como sustrato de esteviósido. El rebaudiósido A (pureza 99 %) también fue suministrado por Blue California.

La reacción de glucosilación catalizada por UGT se acopló a una reacción generadora de UDP-glucosa catalizada por una sacarosa sintasa (tal como AtSUS1 de la sec. con núm. de ident.: 9). En este método, la UDP-glucosa se generó a partir de sacarosa y UDP (Figura 2), de manera que puede omitirse la adición extra de UDP-glucosa. Secuencia AtSUS1 (Bieniawska y otros, Plant J. 2007, 49: 810-828) se sintetizó e insertó en un vector de expresión bacteriano. La proteína AtSUS1 recombinante se expresó y purificó por cromatografía de afinidad. El polipéptido AtSUS1 recombinante purificado se analizó mediante SDS-PAGE (peso molecular: 106,3 kD, Figura 7).

En consecuencia, las actividades de los polipéptidos UGT recombinantes se probaron sin el acoplamiento AtSUS1 (tampón fosfato potásico 50 mM, pH 7,2, MgCl₂ 3 mM, esteviósido o rebaudiósido A 1 mM, UDP 1 mM) o con acoplamiento AtSUS (tampón fosfato potásico 50 mM, pH 7,2, MgCl₂ 3 mM, 1 mg/ml de esteviósido o rebaudiósido A, UDP 1 mM y sacarosa 285 mM). Típicamente, se usaron 10 µg de AtSUS1 para una reacción *in vitro* de 200 µl. La reacción *in vitro* se incubó a 30 °C y se detuvo por extracción utilizando 1-butanol.

El análisis de HPLC se realizó después con el uso de un sistema Dionex UPLC ultimate 3000 (Sunnyvale, CA), que incluye una bomba cuaternaria, un compartimento de columna de temperatura controlada, un muestreador automático y un detector de absorbancia de UV. Se usó Phenomenex Luna NH2 con columna protectora para la caracterización de

glucósidos de esteviol. El acetonitrilo en agua se usó para la elución isocrática en el análisis de HPLC. Los productos de rebaudiósido D, rebaudiósido E, rebaudiósido Z se identificaron mediante análisis de RMN.

5 El polipéptido recombinante (sec. con núm. de ident.: 6) codificado por la sec. con núm. de ident.: 7 mostró una actividad de glucosilación de 1,2-19-O-glucosa, y se sometió a análisis adicional. El gen se derivó de *Hordeum vulgare subsp. vulgare* (abreviado como "HV1" en la presente descripción). El polipéptido HV1 recombinante purificado se analizó mediante SDS-PAGE (Figura 6). Como se muestra en la Figura 6, la proteína HV1 recombinante (peso molecular: 61,4 kD) se purificó por cromatografía de afinidad. Los polipéptidos codificados por otros genes candidatos (Tabla 1) no mostraron ninguna actividad detectable en los ensayos descritos en la presente descripción, aunque comparten una identidad de secuencia de aproximadamente 62-74 % con el polipéptido HV1 recombinante.

15 Como se describe en la presente descripción, el polipéptido recombinante de HV1 transfirió un resto de azúcar al rebaudiósido A para producir rebaudiósido D en todas las condiciones de reacción con o sin AtSUS 1. El rebaudiósido A se convirtió completamente en rebaudiósido D mediante el polipéptido HV1 recombinante en un sistema de reacción de acoplamiento UGT-SUS (Figura 3B-E). Sin embargo, solo el rebaudiósido A parcial se convirtió en rebaudiósido D después de 24 horas por el polipéptido HV1 recombinante solo sin estar acoplado a AtSUS1 (Figura 3F-G). Por lo tanto, el polipéptido HV1 recombinante mostró una actividad de glucosilación 1,2-19-O-glucosa para producir rebaudiósido D a partir de rebaudiósido A y AtSUS 1 mejoró la eficiencia de conversión en el sistema de acoplamiento UGT-SUS.

20 Además, el polipéptido HV1 recombinante acoplado con AtSUS1 (sec. con núm. de ident.: 9) convirtió el esteviósido en rebaudiósido E *in vitro* (Figura 4). Se produjo un compuesto inesperado ("Reb Z") con un tiempo de retención de HPLC de 6.68 minutos (ver Figura 4) que era distintivo de los rebaudiósidos D y E. Este compuesto representa un nuevo glucósido de esteviol, y se denomina "rebaudiósido Z" ("Reb Z"). Para confirmar la conversión de Reb E en Reb Z, se incubó el sustrato Reb E (0,5 mg/ml) con el polipéptido HV1 recombinante (20 µg) y AtSUS1 (20 µg) en un sistema de reacción de acoplamiento UGT-SUS (200 µL) bajo Condiciones similares a las utilizadas en los ejemplos anteriores. Como se muestra en la Figura 8, Reb Z se produjo mediante la combinación del polipéptido HV1 recombinante y AtSUS1. Estos resultados indicaron que HV1 puede transferir el resto de glucosa a Reb E para formar Reb Z.

30 EJEMPLO 3: Biosíntesis de glucósidos de esteviol usando el polipéptido HV1 recombinante

Como se muestra en la Figura 1, el rebaudiósido D también puede formarse por glucosilación del C-3' de la C-13-O-glucosa del rebaudiósido E. Por lo tanto, el rebaudiósido D se puede producir por diferentes rutas biosintéticas (por ejemplo, a través del rebaudiósido A versus rebaudiósido E), en dependencia del orden en que se producen las reacciones de glucosilación. Por ejemplo, la glucosilación en el C-3' de la C-13-O-glucosa del esteviósido puede ocurrir primero para producir el rebaudiósido A intermedio, seguido de la glucosilación en el C-2' de la 19-O-glucosa del rebaudiósido A para producir rebaudiósido D. Hasta ahora, la UGT76G1 (sec. con núm. de ident.: 11) de la estevia se ha identificado como una enzima que transfiere un residuo de azúcar al C-3' de la C-13-O-glucosa del esteviósido para formar el rebaudiósido A.

40 El ADNc de la UGT76G1 optimizado por codones se insertó en un vector de expresión bacteriano, y la proteína UGT76G1 recombinante se expresó y purificó mediante cromatografía de afinidad. El polipéptido UGT76G1 recombinante purificado se analizó mediante SDS-PAGE (peso molecular: 65,4 kD, Figura 9). El sustrato rebaudiósido E se incubó con la UGT76G1 recombinante, con o sin AtSUS1, en condiciones similares a las usadas en los ejemplos anteriores. Los productos fueron analizados por HPLC. Como se muestra en la Figura 12, el rebaudiósido D fue producido por la UGT76G1 recombinante. La adición de AtSUS recombinante en la reacción mejoró la eficiencia de conversión en el sistema de acoplamiento UGT-SUS. Por lo tanto, el polipéptido UGT76G1 recombinante mostró una actividad de glucosilación de 1,3-13-O-glucosa para producir Reb D a partir de Reb E.

50 En consecuencia, la actividad catalítica del polipéptido HV1 recombinante para la biosíntesis de glucósidos de esteviol (por ejemplo, la producción de rebaudiósido D) se determinó además en combinación con la UGT76G1. El sustrato de esteviósido se incubó con el polipéptido HV1 recombinante (10 µg), UGT76G1 (10 µg) y AtSUS1 (10 µg) en un sistema de reacción de acoplamiento UGT-SUS (200 µL) en condiciones similares a las utilizadas en los ejemplos anteriores. Los productos se analizaron por HPLC. Como se muestra en la Figura 5, el rebaudiósido D se produjo mediante la combinación del polipéptido HV1 recombinante, UGT76G1 y AtSUS1. Por lo tanto, el polipéptido HV1 recombinante, que mostró al menos una actividad de glucosilación de 1,2-19-O-glucosa, puede usarse en combinación con otras enzimas UGT (tal como UGT76G1) para la compleja biosíntesis de múltiples etapas de glucósidos de esteviol.

Ejemplo 4: Análisis de la estructura de Reb Z por NRM.

60 El material usado para la caracterización del rebaudiósido Z (Reb Z) se produjo usando la conversión enzimática del rebaudiósido E y se purificó por HPLC.

65 Los datos HRMS se generaron con un instrumento LTQ Orbitrap Discovery HRMS, con su resolución establecida en 30k; datos escaneados de *m/z* 150 a 1500 en modo de electropulverización de iones positivos. El voltaje de la aguja se ajustó a 4 kV; las otras condiciones de la fuente fueron gas de impulsión = 25, gas auxiliar = 0, gas de barrido = 5 (todos los flujos de gas en unidades arbitrarias), voltaje capilar = 30 V, temperatura capilar = 300 °C y voltaje de la lente del tubo =

75. La muestra se diluyó con acetonitrilo:metanol:agua 2:2:1 (igual que el eluyente de infusión) y se inyectaron 50 microlitros.

5 Los espectros de NMR se adquirieron en instrumentos Bruker Avance DRX 500 MHz o Varian INOVA 600 MHz usando secuencias de pulsos estándar. Los espectros de NMR ID (^1H y ^{13}C) y 2D (COSY, TOCSY, HMQC y HMBC) se realizaron en $\text{C}_5\text{D}_5\text{N}$.

10 El Compuesto Reb Z, que se muestra como una mezcla de Reb Z1 y Reb Z2, se muestra en la Figura 10. La fórmula molecular del compuesto Reb Z se dedujo como $\text{C}_{50}\text{H}_{80}\text{O}_{28}$ en base a su espectro de masas de alta resolución positiva (HR), que mostró iones aductos correspondientes a $[\text{M}+\text{Na}]^+$ en m/z 1151.4713; esta composición fue soportada por los datos de espectro NMR de ^{13}C . Los datos del espectro NMR ^1H del Reb Z mostraron la presencia de una mezcla de dos compuestos (Reb Z1 y Reb Z2) en la relación entre 60:40 y 70:30. Por lo tanto, los datos del espectro NMR de ^1H y ^{13}C de Reb Z mostraron un conjunto de picos para cada protón y carbono presentes en su estructura. La hidrólisis ácida de Reb Z con 5 % de H_2SO_4 produjo D-glucosa que se identificó por comparación directa con una muestra auténtica por TLC. La hidrólisis enzimática de Reb Z proporcionó una aglicona, que se identificó como esteviol por comparación entre ^1H NMR y co-TLC con un compuesto estándar. Los valores de NRM de ^1H y ^{13}C para el compuesto Reb Z se asignaron en base a los datos TOCSY, HMQC y HMBC. Las altas constantes de acoplamiento observadas para los cinco protones anoméricos de los restos de glucosa, sugirieron su orientación β según lo informado para los glucósidos de esteviol.

20 Tabla 2. Datos del espectro NMR de ^1H y ^{13}C (cambios químicos y constantes de acoplamiento) para rebaudiósido Z ("Reb Z") y rebaudiósido E ^{a-c}.

Posición	Reb Z		Rebaudiósido E	
	^1H NMR	^{13}C NMR	^1H NMR	^{13}C NMR
1	0,74 t (12,7), 1,65 m	41,2	0,73 t (13,2), 1,68 m	41,0
2	1,45 m, 2,12 m	20,6	1,46 m, 2,13 m	20,6
3	1,13 m,	38,2	1,12 m, 2,78 d (12,8)	38,2
	2,922 d (13,2)/2,79 d (12,8)			
4	---	44,9/44,8	---	44,8
5	0,99 m	58,1	0,97 d (11,8)	57,9
6	1,87 m, 2,14 m	22,6	1,85 m, 2,09 m	22,6
7	1,24 m, 1,68 m	42,2	1,27 m, 1,63 m	42,1
8	---	43,2/43,1	---	43,0
9	0,88 br s	54,5	0,88 br s	54,5
10	---	40,2	---	40,2
11	1,68m	21,1	1,65m	21,1
12	1,92 m, 2,28 m	37,8	1,96 m, 2,16 m	37,8
13	---	86,6	---	86,6
14	1,72 m, 2,48 d (10,8)	44,9/44,8	1,74 d (11,4), 2,54 d (11,0)	44,8
15	1,92 m, 2,14 m	48,5	2,04 m, 2,12 m	48,5
16	---	155,2/155,0	---	154,9
17	5,09/5,12 s, 5,68/5,74 s	105,4/105,2	5,09 s, 5,76 s	105,4
18	1,43/1,49 s	29,9	1,43 s	29,8
19	---	176,3/176,1	---	176,2
20	1,09 s	17,3	1,10s	17,2

ES 2 709 439 T3

	1'	6,30 d (7,9)/6,35 d (7,8)	93,9/93,6	6,30 d (7,9)	93,9
5	2'	4,38 m	81,8	4,38 m	81,7
	3'	4,27 m	78,5	4,26 m	78,4
	4'	4,24 m	71,8	4,22 m	71,8
10	5'	3,94 m	79,6	3,92 m	79,5
	6'	4,33 m, 4,46 m	62,6	4,33 m, 4,43 m	62,6
	1"	5,12 d (7,4)/	98,4/98,2	5,16 d (7,5)	98,4
		5,14 d (7,4)/			
15	2"	4,18 m	84,9	4,17 m	84,9
	3"	4,29 m	78,6	4,32 m	78,5
	4"	4,20 m	72,1	4,22 m	72,1
20	5"	3,74 m	78,5	3,72 m	78,2
	6"	4,32 m, 4,38 m	62,8	4,26 m, 4,35 m	62,9
	1'''	5,33 d (7,8)/	106,7/104,5	5,32 d (7,5)	107,2
25		5,46 d (7,8)/			
	2'''	4,14 t (8,4)	85,7/77,7	4,15 t (8,4)	77,7
	3'''	4,25 m	78,7	4,26 m	78,6
30	4'''	4,34 m	72,1/71,5	4,36 m	72,3
	5'''	3,88 m	79,1	3,96 m	79,0
	6'''	4,43 m, 4,56 m	63,1	4,46 m, 4,56 m	63,2
35	1''''	5,48 d (7,9)/	106,3/104,4	5,48 d (7,9)	106,2
		5,40 d (7,6)			
	2''''	4,04 t (7,9)	85,6/77,3	4,06 t (7,9)	76,8
40	3''''	4,22 m	78,8	4,25 m	78,7
	4''''	4,38 m	71,4/71,0	4,31 m	71,2
	5''''	3,96 m	79,1	4,02 m	79,1
45	6''''	4,38 m, 4,57 m	63,4	4,42 m, 4,54 m	63,4
	1'''''	5,29 d (7,5)/	107,1		
		5,34 d (7,5)/			
	2'''''	4,02 m	77,0		
50	3'''''	4,21 m	78,6		
	4'''''	4,25 m	71,6/71,2		
	5'''''	3,98 m	79,1		
55	6'''''	4,34 m, 4,48 m	63,3		
	^a asignaciones hechas sobre la base de las correlaciones TOCSY, HMQC y HMBC; los valores de desplazamiento químico están en δ (ppm); ^c constantes de acoplamiento están en Hz,				

60

Con base en los resultados de los datos del espectro de NRM y los experimentos de hidrólisis de Reb Z, y una comparación cercana de los valores de NRM de ¹H y ¹³C de Reb Z con rebaudiósido E sugirió que la mezcla de dos compuestos producidos por la conversión enzimática se dedujo como 13-[(2-O-β-D-glucopiranosil-2-O-β-D-glucopiranosil-β-D-glucopiranosil)oxi]ácido *ent*-kaur-16-en-19-oico-2-O-β-D-glucopiranosilo-β-D-glucopiranosil éster (Reb Z1) o 13-[(2-O-β-

65

D-glucopiranosil-β-D-glucopiranosil) oxi]ácido *ent*-kaur-16-en-19-oico-[(2-O-β-D-glucopiranosil-2-O-β-D-glucopiranosil-β-D-glucopiranosilo)éster (Reb Z2).

5 Hidrólisis ácida del compuesto Reb Z. A una solución del compuesto Reb Z (5 mg) en MeOH (10 ml) se agregaron 3 ml de H₂SO₄ al 5 % y la mezcla se calentó a reflujo durante 24 horas. La mezcla de reacción se neutralizó luego con carbonato de sodio saturado y se extrajo con acetato de etilo (EtOAc) (2 x 25 ml) para dar una fracción acuosa que contenía azúcares y una fracción de EtOAc que contenía la parte de aglicona. La fase acuosa se concentró y se comparó con los azúcares estándar usando los sistemas TLC EtOAc/n-butanol/agua (2:7:1) y CH₂Cl₂/MeOH/agua (10:6:1); los azúcares fueron identificados como D-glucosa.

10 Hidrólisis enzimática del compuesto Reb Z. El Compuesto Reb Z (1 mg) se disolvió en 10 ml de tampón de acetato de sodio 0,1 M, pH 4,5 y se añadió pectinasa cruda de *Aspergillus niger* (50 uL, Sigma-Aldrich, P2736). La mezcla se agitó a 50 °C durante 96 horas. El producto precipitó durante la reacción y se filtró y luego se cristalizó. El producto resultante obtenido de la hidrólisis de 1 se identificó como esteviol por comparación de su co-TLC con el compuesto estándar y los datos espectrales de ¹H NMR (Figura 11).

20 Se produjo una mezcla de dos compuestos llamada Reb Z mediante la bioconversión de rebaudiósido E usando una metodología enzimática y las estructuras se caracterizaron como 13-[(2-O-β-D-glucopiranosil-2-O-β-D-glucopiranosina-β-D-glucopiranosil)oxi] ácido *ent*-kaur-16-en-19-oico-2-O-β-D-glucopiranosil-β-D-glucopiranosil éster (Reb Z1), o 13-[(2-ácido O-β-D-glucopiranosil-β-D-glucopiranosil)oxi] ácido *ent*-kaur-16-en-19-oico-[(2-O-β-D-glucopiranosil-2-O-β-D-glucopiranosilo-β-D-glucopiranosil) éster (Reb Z2), sobre la base de la extensa NMR 1D y 2D, así como también de los datos de espectros de masas de alta resolución y los estudios de hidrólisis.

25 Por lo tanto, la actividad de glucosilación de 1,2-19-O-glucosa del polipéptido HV1 recombinante se confirmó por su capacidad para transferir un segundo resto de azúcar al C-2' de una 19-O-glucosa de esteviósido para producir rebaudiósido E. El polipéptido recombinante HV1 también tiene la actividad de transferir el tercer resto de glucosa al C-2' de una 13-O-glucosa o el C-2' de una 19-O-glucosa de rebaudiósido E para producir Reb Z1 o Reb Z2.

30 Listado de secuencias

<110> Conagen, Inc.
Mao, Guohong
Yu, Xiaodan

35 <120> PRODUCCIÓN RECOMBINANTE DE GLUCÓSIDOS DE ESTEVIOL

<130> 32559-17

<150> US 61/898,571

40 <151> 2013-11-01

<160> 12

<170> PatentIn versión 3.5

45

<210> 1

<211> 470

<212> PRT

<213> *Brachypodium distachyon*

50

<400> 1

55

60

65

ES 2 709 439 T3

5 Met Asp Asp Ala Gly Tyr Ser Ser Ala Tyr Ser Pro Leu His Val Val
1 5 10 15

10 Ile Cys Pro Trp Leu Ala Phe Gly His Gln Leu Pro Cys Leu Asp Leu
20 25 30

15 Ala Glu Arg Leu Ala Ser Arg Gly His Arg Val Ser Phe Val Ser Thr
35 40 45

20 Pro Arg Ile Ile Ala Arg Leu Pro Pro Val Arg Pro Thr Ala Ala Gln
50 55 60

25 Leu Ile Asn Phe Val Ala Leu Pro Leu Pro Ser Val Asp Gly Leu Pro
65 70 75 80

30 Glu Gly Ala Glu Ser Thr Asn Asp Val Pro Phe Asp Lys Phe Glu Leu
85 90 95

35 His Arg Lys Ala Phe Asp Gly Leu Ala Leu Pro Phe Ser Glu Phe Leu
100 105 110

40 Gly Ala Ala Cys Ala Lys Gly Gln Gly His Lys Pro Asp Trp Ile Leu
115 120 125

45 Val Asp Ile Phe His His Trp Ala Ala Ala Ala Val Glu His Lys
130 135 140

50 Val Pro Cys Ala Met Leu Leu Leu Gly Ala Ala Ser Phe Ile Ala Ser
145 150 155 160

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

Gly Ala Gly Gln Leu Phe Glu His Ala Ala Ser Gly Val Gln Val Gln
 165 170 175

Glu Arg Pro Ser Ser Thr Glu Pro Pro Lys Phe Glu Ile Glu Met Arg
 180 185 190

Glu Leu Ile Ile Thr Gln Arg Ala Ser Gly Met Ser Ile Ala Glu Arg
 195 200 205

Val Ser Leu Thr Leu Gln Arg Ser Asn Leu Ala Ala Met Arg Ser Cys
 210 215 220

Val Glu Trp Glu Pro Glu Ser Val Pro Leu Val Ala Ser Leu Gly Val
 225 230 235 240

Gly Gly Lys Pro Val Val Pro Leu Gly Leu Leu Pro Pro Ser Pro Glu
 245 250 255

Gly Gly Arg Gly Val Cys Lys Asp Gly Lys Lys Asp Ala Thr Val Lys
 260 265 270

Trp Leu Asp Val Gln Pro Ala Lys Ser Val Val Tyr Val Ala Met Gly
 275 280 285

Thr Glu Val Pro Leu Pro Ala Glu Gln Val His Glu Leu Ala Phe Gly
 290 295 300

Ile Glu Leu Ala Gly Thr Arg Phe Leu Trp Ala Leu Arg Lys Pro Ser
 305 310 315 320

Gly Gly Ala Pro Asp Ala Asp Ile Leu Pro Pro Gly Phe Glu Asp Arg
 325 330 335

Thr Ala Gly Arg Gly Leu Val Arg Thr Gly Trp Val Pro Gln Met Ser
 340 345 350

Ile Leu Gly His Asp Ala Val Gly Ala Phe Leu Thr His Cys Gly Trp
 355 360 365

Asn Ser Ile Ile Glu Gly Leu Leu Phe Gly His Pro Leu Val Met Leu
 370 375 380

Pro Ile Leu Gly Asp Gln Gly Pro Asn Ala Arg Leu Met Glu Gly Lys
 385 390 395 400

Lys Val Gly Val Gln Val Gln Arg Asp Gly Asn Asp Gly Ser Phe Asn
 405 410 415

ES 2 709 439 T3

5 Arg Glu Gly Val Ala Met Ala Val Arg Ala Val Met Val Glu Glu Glu
 420 425 430
 Ser Lys Lys Ile Phe Lys Ala Asn Ala Lys Lys Met Gln Glu Ile Val
 435 440 445
 10 Ala Asp Thr Glu Arg His Glu Arg Tyr Ile Asp Gly Phe Ile Gln Gln
 450 455 460
 15 Leu Arg Ser Tyr Lys Glu
 465 470
 20 <210> 2
 <211> 460
 <212> PRT
 <213> Brachypodium distachyon
 25 <400> 2
 Met Asp Asn Gly Ser Ser Ser Ser Ser Ser Ser Ser Pro Leu His Val Val
 1 5 10 15
 30 Ile Cys Pro Trp Leu Ala Phe Gly His Gln Leu Pro Cys Leu Asp Leu
 20 25 30
 35 Ala Glu Arg Leu Ala Leu Arg Gly His Arg Val Ser Phe Val Ser Thr
 35 40 45
 40 Pro Arg Ile Ile Ala Arg Leu Pro Pro Val Arg Pro Val Ala Ala Ser
 50 55 60
 45 Leu Val Asp Leu Val Ala Leu Pro Leu Pro Arg Val Asp Gly Leu Pro
 65 70 75 80
 50 Glu Gly Ala Glu Ser Thr Asn Asp Val Pro Tyr Glu Lys Phe Glu Leu
 85 90 95
 55 His Arg Lys Ala Phe Asp Gly Leu Ala Val Pro Phe Ser Glu Phe Leu
 100 105 110
 60 Arg Ala Ala Cys Ala Glu Glu Gly Lys Lys Pro Asp Trp Ile Ile Val
 115 120 125
 Asp Thr Phe His His Trp Ala Ala Ala Ala Ala Ile Glu His Lys Val
 130 135 140
 65 Pro Cys Ala Met Leu Met Leu Gly Ala Ala Gly Leu Ile Val Ala Trp
 145 150 155 160

5 Ala Thr Gln Pro Ser Lys His Val Thr Ser Glu Gln Gln Glu Gln Ser
165 170 175

Ala Ala Glu Pro Pro Arg Phe Glu Thr Glu Arg Arg Lys Leu Ala Thr
180 185 190

10 Thr Gln Arg Ala Ser Gly Met Ser Ile Ala Glu Arg Cys Ser Val Thr
195 200 205

15 Leu Glu Arg Cys Asn Leu Val Ala Met Arg Ser Cys Leu Glu Trp Glu
210 215 220

20 Pro Glu Ser Ile Pro Leu Ala Thr Thr Ile Gly Gly Lys Gln Leu Val
225 230 235 240

Pro Leu Gly Leu Leu Pro Pro Ser Pro Glu Gly Gly Arg Gly Val Ser
245 250 255

25 Lys Glu Asp Ala Thr Val Arg Trp Leu Asp Ala Gln Pro Thr Lys Ser
260 265 270

30 Val Val Tyr Val Ala Leu Gly Ser Glu Val Pro Leu Gly Ala Lys Glu
275 280 285

35 Val His Glu Leu Ala Leu Gly Leu Glu Leu Ala Gly Thr Arg Phe Leu
290 295 300

Trp Ser Leu Arg Lys Pro Ser Gly Val Ser Asp Ala Asp Ile Leu Pro
305 310 315 320

40 Ser Gly Phe Glu Glu Arg Thr Arg Gly Arg Gly Leu Val Thr Met Gly
325 330 335

45 Trp Val Pro Gln Ile Ser Val Leu Ala His Gly Ala Val Gly Ala Phe
340 345 350

50 Leu Thr His Cys Gly Trp Asn Ser Ile Ile Glu Gly Leu Gln Phe Gly
355 360 365

55 His Pro Leu Val Met Leu Pro Ile Phe Gly Asp Gln Gly Pro Asn Ala
370 375 380

Arg Met Met Glu Gly Arg Lys Val Gly Val Gln Val Pro Arg Asp Glu
385 390 395 400

60 Ser Asn Gly Ser Phe Asp Arg Glu Gly Val Ala Thr Thr Val Arg Ala

ES 2 709 439 T3

5 405 410 415
 Val Ala Val Glu Glu Glu Gly Asn Arg Ile Phe Thr Ala Asn Ala Lys
 420 425 430
 10 Lys Met Gln Glu Ile Val Ala Asp Lys Gly Cys His Asp Lys Tyr Val
 435 440 445
 15 Asp Lys Phe Ile Gln Lys Leu Arg Ser Tyr Met Glu
 450 455 460
 <210> 3
 <211> 466
 <212> PRT
 20 <213> Brachypodium distachyon
 <220>
 <221> CARACT_MISC
 <222> (407)..(407)
 25 <223> Desconocido
 <400> 3
 30 Met Asp Ala Ala Gly Ser Ser Ser Ser Ser Pro Leu Arg Ile Ala Ile
 1 5 10 15
 35 Val Pro Trp Leu Ala Phe Gly His Leu Leu Pro Tyr Leu Glu Leu Ala
 20 25 30
 40 Glu Arg Leu Ala Ala Arg Gly His Arg Val Ser Phe Val Ser Thr Pro
 35 40 45
 45 Arg Asn Leu Ala Arg Leu Pro Pro Leu Arg Pro Ala Ala Ala Pro Arg
 50 55 60
 50 Val Asp Leu Val Ala Leu Pro Leu Pro Arg Val Glu Gly Leu Pro Asp
 65 70 75 80
 55 Gly Ala Glu Ser Thr Asn Asp Val Pro Asp Asp Glu Arg Glu Pro Leu
 85 90 95
 60 Trp Lys Ala Phe Asp Gly Leu Ala Ala Pro Phe Ala Gly Phe Leu Thr
 100 105 110
 65 Ala Ala Cys Ala Asp Glu Gly Thr Arg Pro His Trp Ile Ile Ala Asp
 115 120 125
 70 Ser Phe His His Trp Ala Ala Ala Ala Leu Glu His Lys Val Pro
 130 135 140

ES 2 709 439 T3

5 Cys Ala Met Leu Leu Pro Thr Ala Ala Leu Ile Ala Ala Ser Ala Gly
 145 150 155 160
 Ala Gly Arg Pro Ser Pro Glu Glu His Ala Glu Gln Gln Pro Gln Pro
 165 170 175
 10 Arg Tyr Glu Gln Glu Gly Arg Ala Thr Leu Leu Thr Asp Gly Asp Met
 180 185 190
 Ser Gly Met Ser Ile Met Gln Arg Ser Val Leu Thr Leu Glu Arg Cys
 195 200 205
 15 Lys Leu Thr Ala Ile Arg Ser Cys Val Glu Trp Glu Pro Glu Cys Leu
 210 215 220
 20 Pro Leu Val Ser Glu Phe Ile Gly Lys Pro Val Val Pro Leu Gly Leu
 225 230 235 240
 25 Leu Pro Pro Ser Pro Asp Gly Gly Arg Arg Ala Ala Asn Thr Asn Gly
 245 250 255
 30 Glu Asp Ala Thr Ile Arg Trp Leu Asp Ala Gln Pro Pro Asn Ser Val
 260 265 270
 Val Tyr Val Ala Leu Gly Ser Glu Val Pro Leu Pro Val Glu Gln Thr
 275 280 285
 35 His Glu Leu Ala Leu Gly Leu Glu Leu Ser Lys Thr Arg Phe Leu Trp
 290 295 300
 40 Ala Leu Arg Lys Pro Ser Gly Val Leu Asp Ala Glu Met Leu Pro Met
 305 310 315 320
 Gly Phe Gln Glu Arg Ile His Gly His Gly Leu Val Thr Thr Gly Trp
 325 330 335
 45 Val Pro Gln Met Ser Ile Leu Ala His Gly Ala Val Gly Ser Phe Leu
 340 345 350
 50 Thr His Cys Gly Arg Asn Ser Leu Ile Glu Gly Leu Leu Phe Gly His
 355 360 365
 55 Pro Leu Ile Met Leu Pro Ile Phe Gly Asp Gln Gly Pro Asn Ala Arg
 370 375 380
 60 Leu Met Glu Gly Lys Lys Val Gly Leu Gln Val Ala Arg Asn Glu Asn
 385 390 395 400
 65

ES 2 709 439 T3

Asp Gly Ser Phe Asp Arg Xaa Gly Val Ala Ser Ala Val Arg Ser Val
 405 410 415
 5
 Met Leu Glu Glu Asp Ala Arg Lys Ser Phe Val Ala Asn Ala Leu Glu
 420 425 430
 10
 Met Gln Lys Ile Val Ala Asp Lys Glu Arg His Glu Arg Tyr Ile Asp
 435 440 445
 15
 Glu Phe Ile His Gln Leu Arg Ser Tyr Thr Ser Ala Pro Thr Ser Asp
 450 455 460
 20
 Arg Asn
 465
 <210> 4
 <211> 460
 <212> PRT
 <213> Brachypodium distachyon
 <400> 4
 30
 Met Asp Ala Ala Gly Ser Ser Pro Pro Leu Arg Ile Val Ile Val Pro
 1 5 10 15
 35
 Trp Leu Ala Phe Gly His Met Leu Pro Tyr Leu Glu Leu Ala Glu Arg
 20 25 30
 40
 Leu Ala Ala Arg Gly His Arg Val Ser Tyr Val Ser Thr Pro Arg Asn
 35 40 45
 45
 Leu Ala Arg Leu Pro Pro Leu Arg Pro Ala Ala Ala Pro Arg Val Asp
 50 55 60
 45
 Leu Val Ala Leu Pro Leu Pro Arg Val Glu Gly Leu Pro Asp Gly Ala
 65 70 75 80
 50
 Glu Ser Thr Asn Asp Val Pro Asp Asp Glu Arg Glu Pro Leu Trp Lys
 85 90 95
 55
 Ala Phe Asp Gly Leu Ala Ala Pro Phe Arg Ser Val Pro Arg Gln Arg
 100 105 110
 55
 Cys Ala Arg Asp Asp Thr Arg Pro His Trp Ile Leu Ala Asp Cys Phe
 115 120 125
 60
 His His Trp Ala Val Asp Ala Ala Leu Asp His Lys Val Pro Cys Ala
 130 135 140
 65

ES 2 709 439 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

Met Phe Leu Pro Thr Ala Ala Val Ile Ala Thr Met Pro Gln Arg Gln
 145 150 155 160

Pro Asp His Ala Ala Ser Ala Pro Ala Glu His Ala Val Pro Arg His
 165 170 175

Glu Ile Glu Ala Thr Ala Pro Leu Leu Ser Asp Gln Gly Val Ser Gly
 180 185 190

Met Ser Ile Val Gln Arg Tyr Leu Leu Thr Lys Glu Arg Cys Thr Val
 195 200 205

Gly Ala Ile Arg Ser Cys Val Glu Trp Glu Pro Asp Ser Tyr Pro Leu
 210 215 220

Ala Ala Thr Ile Leu Gly Met Pro Val Val Pro Leu Gly Leu Leu Pro
 225 230 235 240

Pro Ser Pro Asp Gly Gly Arg Arg Ala Pro Asp Gly Ser Glu His Ala
 245 250 255

Thr Val Arg Trp Leu Asp Ala Gln Pro Pro Ser Ser Val Val Tyr Val
 260 265 270

Ala Leu Gly Ser Glu Val Pro Leu Pro Val Asp His Val His Glu Leu
 275 280 285

Ala Leu Gly Leu Glu Leu Ala Gly Thr Arg Phe Leu Trp Ala Leu Arg
 290 295 300

Lys Pro Asn Gly Val Pro Asp Ala Asp Met Leu Pro Ala Gly Phe Gln
 305 310 315 320

Asp Arg Thr Arg Gly His Gly Leu Val Thr Thr Gly Trp Val Pro Gln
 325 330 335

Met Ser Ile Leu Ala His Gly Ala Val Gly Ala Phe Leu Thr His Cys
 340 345 350

Gly Arg Asn Ser Leu Ile Glu Gly Leu Leu Leu Gly His Pro Leu Val
 355 360 365

Met Leu Pro Ile Phe Gly Asp Gln Gly Pro Asn Ala Arg Ala Met Glu
 370 375 380

Arg Lys Lys Val Gly Leu Gln Val Lys Arg Asp Asp Asn Asp Gly Ser

Ile Phe His His Trp Ala Ala Ala Asp Thr Leu Glu Tyr Lys Val Pro
 130 135 140
 5
 Ile Gly Ala Ala Asn Val Ile Ala Thr Trp His Gly Arg Leu Val Lys
 145 150 155 160
 10
 His Thr Thr Met Ser Lys Gln Glu Gln Pro Ala Ser Lys Leu Pro Arg
 165 170 175
 15
 Phe Glu Ile Glu Arg Arg Gln Leu Ser Thr Thr Gln Arg Ala Ser Gly
 180 185 190
 20
 Met Ser Ile Ala Glu Arg Ile Ser Leu Thr Leu Arg Arg Cys Asn Leu
 195 200 205
 25
 Val Val Met Arg Thr Arg Leu Glu Trp Glu Pro Glu Ser Val Pro Leu
 210 215 220
 30
 Ala Ala Ser Leu Gly Gly Lys Pro Val Ile Ser Leu Gly Leu Leu Pro
 225 230 235 240
 35
 Pro Leu Arg Lys Gly Cys Cys Gly Val Thr Glu Asp Gly Lys Asn Ile
 245 250 255
 40
 Met Cys Trp Leu Asp Leu Gln Pro Ala Lys Ser Val Val Tyr Val Ala
 260 265 270
 45
 Leu Gly Thr Glu Val Pro Leu Pro Val Glu Gln Met His Glu Leu Ala
 275 280 285
 50
 Leu Arg Leu Glu Leu Ala Gly Met Gln Phe Leu Trp Ala Leu Arg Lys
 290 295 300
 55
 Pro Arg Gly Val His Glu Ala Glu Ile Leu Pro Leu Gly Phe Glu Glu
 305 310 315 320
 60
 Arg Met Xaa Gly Leu Val Thr Thr Gly Leu Ala Pro Gln Ile Asn Ile
 325 330 335
 65
 Leu Ala His Gly Ala Val Gly Thr Phe Leu Thr His Cys Gly Trp Ser
 340 345 350
 70
 Leu Thr Ile Glu Gly Leu Leu Phe Gly His Pro Leu Ile Met Leu Pro
 355 360 365
 75
 Met Phe Gly Asp Gln Gly Pro Asn Ala Arg Leu Met Glu Gly Arg Lys
 370 375 380

ES 2 709 439 T3

5 Val Gly Val Gln Val Pro Arg Asn Glu Gly Asp Gly Ser Phe Asp Arg
 385 390 395 400
 10 Glu Gly Val Ala Thr Thr Val Arg Ala Val Thr Val Glu Glu Glu Gly
 405 410 415
 15 Lys Arg Ile Phe Thr Ser Asn Ala Lys Lys Met Gln Glu Ile Lys Ala
 420 425 430
 20 Asp Thr Glu Cys Gln Glu Arg Tyr Ile Asn Gly Phe Ile Lys Lys Leu
 435 440 445
 25 Arg Ser Tyr Lys Glu Pro Asp Leu Ala His Pro Ile Cys Pro Val Thr
 450 455 460
 30 Ser Ser Pro Thr His Tyr Glu
 465 470
 35 <210> 6
 <211> 459
 <212> PRT
 <213> Hordeum vulgare
 40 <400> 6
 45 Met Asp Gly Asn Ser Ser Ser Ser Pro Leu His Val Val Ile Cys Pro
 1 5 10 15
 50 Trp Leu Ala Leu Gly His Leu Leu Pro Cys Leu Asp Ile Ala Glu Arg
 20 25 30
 55 Leu Ala Ser Arg Gly His Arg Val Ser Phe Val Ser Thr Pro Arg Asn
 35 40 45
 60 Ile Ala Arg Leu Pro Pro Leu Arg Pro Ala Val Ala Pro Leu Val Asp
 50 55 60
 65 Phe Val Ala Leu Pro Leu Pro His Val Asp Gly Leu Pro Glu Gly Ala
 65 70 75 80
 70 Glu Ser Thr Asn Asp Val Pro Tyr Asp Lys Phe Glu Leu His Arg Lys
 85 90 95
 75 Ala Phe Asp Gly Leu Ala Ala Pro Phe Ser Glu Phe Leu Arg Ala Ala
 100 105 110
 80 Cys Ala Glu Gly Ala Gly Ser Arg Pro Asp Trp Leu Ile Val Asp Thr
 115 120 125
 85

ES 2 709 439 T3

5 Phe His His Trp Ala Ala Ala Ala Ala Val Glu Asn Lys Val Pro Cys
 130 135 140
 Val Met Leu Leu Leu Gly Ala Ala Thr Val Ile Ala Gly Phe Ala Arg
 145 150 155 160
 10 Gly Val Ser Glu His Ala Ala Ala Ala Val Gly Lys Glu Arg Pro Ala
 165 170 175
 Ala Glu Ala Pro Ser Phe Glu Thr Glu Arg Arg Lys Leu Met Thr Thr
 180 185 190
 15 Gln Asn Ala Ser Gly Met Thr Val Ala Glu Arg Tyr Phe Leu Thr Leu
 195 200 205
 20 Met Arg Ser Asp Leu Val Ala Ile Arg Ser Cys Ala Glu Trp Glu Pro
 210 215 220
 25 Glu Ser Val Ala Ala Leu Thr Thr Leu Ala Gly Lys Pro Val Val Pro
 225 230 235 240
 30 Leu Gly Leu Leu Pro Pro Ser Pro Glu Gly Gly Arg Gly Val Ser Lys
 245 250 255
 Glu Asp Ala Ala Val Arg Trp Leu Asp Ala Gln Pro Ala Lys Ser Val
 260 265 270
 35 Val Tyr Val Ala Leu Gly Ser Glu Val Pro Leu Arg Ala Glu Gln Val
 275 280 285
 40 His Glu Leu Ala Leu Gly Leu Glu Leu Ser Gly Ala Arg Phe Leu Trp
 290 295 300
 Ala Leu Arg Lys Pro Thr Asp Ala Pro Asp Ala Ala Val Leu Pro Pro
 305 310 315 320
 45 Gly Phe Glu Glu Arg Thr Arg Gly Arg Gly Leu Val Val Thr Gly Trp
 325 330 335
 50 Val Pro Gln Ile Gly Val Leu Ala His Gly Ala Val Ala Ala Phe Leu
 340 345 350
 55 Thr His Cys Gly Trp Asn Ser Thr Ile Glu Gly Leu Leu Phe Gly His
 355 360 365
 Pro Leu Ile Met Leu Pro Ile Ser Ser Asp Gln Gly Pro Asn Ala Arg

ES 2 709 439 T3

5 ggcgttctgg ctcattggtgc ggtggetgog tttctgaccc actgtggctg gaactctacg 1080
 atogaaggcc tgctgttogg tcatccgctg attatgctgc cgatcagctc tgatcagggt 1140
 ccgaatgocg gcctgatgga aggcctgaaa gtcggtatgc aagtgcocgog tgatgaatca 1200
 gacggctcgt ttctgtcgcga agatgttgcc gcaacogtcc gcgccgtggc agttgaagaa 1260
 10 gacggctcgc gcgtcttcac ggctaacgog aaaaagatgc aagaaattgt ggccgatggc 1320
 gcatgccacg aacgttgat tgacggtttt atccagcaac tgogcagtta caaggogtga 1380

15 <210> 8
 <211> 488
 <212> PRT
 <213> Hordeum vulgare

20 <400> 8

25 Met Asp Ala Gly Ser Ser Thr Ala Gly Pro Leu Arg Ile Val Ile Cys
 1 5 10 15
 Pro Trp Leu Ala Phe Gly His Leu Leu Pro Tyr Leu Glu Leu Ala Glu
 20 25 30
 30 Arg Leu Ala Ser Arg Gly His Arg Val Ala Phe Val Ser Thr Pro Arg
 35 40 45
 Asn Leu Ala Arg Leu Pro Pro Pro Ala Ser Pro Cys Ser Val Asp Leu
 50 55 60
 35 Val Ala Leu Gln Leu Pro Arg Val Asp Gly Leu Pro Glu Gly Ala Glu
 65 70 75 80
 40 Ser Thr Asn Asp Val Pro Asp Glu Met Arg Glu Leu His Trp Lys Ala
 85 90 95
 45 Phe Asp Gly Leu Ala Ala Pro Phe Ala Asp Phe Leu Ala Ala Ala Cys
 100 105 110
 Ala Asp Asp Gly Arg Arg Pro His Trp Ile Ile Ala Asp Cys Phe His
 115 120 125
 50 His Trp Ala Ala Ala Ala Ala Leu Asp His Lys Val Pro Cys Ala Val
 130 135 140
 55 Leu Leu Pro Thr Ala Ala Met Leu Ala Ala Ala Pro Arg Gln Gln Pro
 145 150 155 160
 60 Leu Gly Ser Lys Pro Val Glu Ala Ala Ala Ala Ala Ser Val Leu Gly
 165 170 175

65

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60
 65

Gln Ala Ala Ala Ala Val Arg Leu Ala Val Pro Ser Tyr Glu Arg Asp
 180 185 190

Asp Val Ala Pro Ala Tyr Ala Asp Asp Cys Ala Ser Gly Met Ser Ile
 195 200 205

Ala Gln Arg Trp Phe Leu Ala Lys Glu Arg Cys Thr Val Leu Ala Ile
 210 215 220

Arg Ser Cys Val Glu Trp Glu Pro Glu Thr Phe Pro Leu Val Glu Ala
 225 230 235 240

Leu Leu Gly Lys Pro Val Val Pro Leu Gly Leu Leu Pro Pro Ser Ala
 245 250 255

Asp Gly Gly Arg Arg Arg Ala Ala Gly Ser Ser Glu Asp His Val Thr
 260 265 270

Leu Arg Trp Leu Glu Glu Gln Pro Pro Asp Ser Val Val Tyr Ile Ala
 275 280 285

Leu Gly Ser Glu Val Pro Leu Ser Ile Glu Gln Val His Glu Leu Ala
 290 295 300

Leu Gly Leu Glu Leu Ala Gly Thr Arg Phe Leu Trp Ala Leu Arg Lys
 305 310 315 320

Pro Ala Gly Ala Val Val Gly Asn Asn Asp Asp Asp Thr Leu Pro Pro
 325 330 335

Gly Phe Arg Asp Cys Thr Arg Gly His Gly Leu Val Thr Met Gly Trp
 340 345 350

Val Pro Gln Ile Ser Ile Leu Ala His Ala Ala Val Gly Ala Phe Leu
 355 360 365

Thr His Cys Gly Arg Asn Ser Leu Ile Glu Gly Leu Leu Phe Gly His
 370 375 380

Pro Leu Val Met Leu Pro Ile Phe Gly Asp Gln Gly Pro Asn Ala Arg
 385 390 395 400

Gln Met Glu Ala Lys Lys Val Gly Leu Gln Val Ala Arg Asp Asp Asp
 405 410 415

Asp Gly Ser Phe Asp Arg His Gly Val Ala Thr Ala Val Arg Ala Val
 420 425 430

ES 2 709 439 T3

5 Met Val Asp Gly Glu Ala Arg Arg Gly Phe Val Ala Asn Ala Ile Lys
435 440 445

10 Met Gln Ala Ile Val Ala Asp Lys Glu Arg His Glu Arg Tyr Ile Asp
450 455 460

15 Gly Phe Val Gln Gln Leu Arg Ser Tyr Leu Gln Ala Thr Asp Asp Leu
465 470 475 480

20 Thr Thr Ala Thr Pro Asn Ser Ser
485

<210> 9
<211> 808
<212> PRT
<213> Arabidopsis thaliana

<400> 9

25 Met Ala Asn Ala Glu Arg Met Ile Thr Arg Val His Ser Gln Arg Glu
1 5 10 15

30 Arg Leu Asn Glu Thr Leu Val Ser Glu Arg Asn Glu Val Leu Ala Leu
20 25 30

35 Leu Ser Arg Val Glu Ala Lys Gly Lys Gly Ile Leu Gln Gln Asn Gln
35 40 45

40 Ile Ile Ala Glu Phe Glu Ala Leu Pro Glu Gln Thr Arg Lys Lys Leu
50 55 60

45 Glu Gly Gly Pro Phe Phe Asp Leu Leu Lys Ser Thr Gln Glu Ala Ile
65 70 75 80

50 Val Leu Pro Pro Trp Val Ala Leu Ala Val Arg Pro Arg Pro Gly Val
85 90 95

55 Trp Glu Tyr Leu Arg Val Asn Leu His Ala Leu Val Val Glu Glu Leu
100 105 110

60 Gln Pro Ala Glu Phe Leu His Phe Lys Glu Glu Leu Val Asp Gly Val
115 120 125

65 Lys Asn Gly Asn Phe Thr Leu Glu Leu Asp Phe Glu Pro Phe Asn Ala
130 135 140

70 Ser Ile Pro Arg Pro Thr Leu His Lys Tyr Ile Gly Asn Gly Val Asp
145 150 155 160

ES 2 709 439 T3

5 Phe Leu Asn Arg His Leu Ser Ala Lys Leu Phe His Asp Lys Glu Ser
165 170 175

10 Leu Leu Pro Leu Leu Lys Phe Leu Arg Leu His Ser His Gln Gly Lys
180 185 190

15 Asn Leu Met Leu Ser Glu Lys Ile Gln Asn Leu Asn Thr Leu Gln His
195 200 205

20 Thr Leu Arg Lys Ala Glu Glu Tyr Leu Ala Glu Leu Lys Ser Glu Thr
210 215 220

25 Leu Tyr Glu Glu Phe Glu Ala Lys Phe Glu Glu Ile Gly Leu Glu Arg
225 230 235 240

30 Gly Trp Gly Asp Asn Ala Glu Arg Val Leu Asp Met Ile Arg Leu Leu
245 250 255

35 Leu Asp Leu Leu Glu Ala Pro Asp Pro Cys Thr Leu Glu Thr Phe Leu
260 265 270

40 Gly Arg Val Pro Met Val Phe Asn Val Val Ile Leu Ser Pro His Gly
275 280 285

45 Tyr Phe Ala Gln Asp Asn Val Leu Gly Tyr Pro Asp Thr Gly Gly Gln
290 295 300

50 Val Val Tyr Ile Leu Asp Gln Val Arg Ala Leu Glu Ile Glu Met Leu
305 310 315 320

55 Gln Arg Ile Lys Gln Gln Gly Leu Asn Ile Lys Pro Arg Ile Leu Ile
325 330 335

60 Leu Thr Arg Leu Leu Pro Asp Ala Val Gly Thr Thr Cys Gly Glu Arg
340 345 350

65 Leu Glu Arg Val Tyr Asp Ser Glu Tyr Cys Asp Ile Leu Arg Val Pro
355 360 365

70 Phe Arg Thr Glu Lys Gly Ile Val Arg Lys Trp Ile Ser Arg Phe Glu
370 375 380

75 Val Trp Pro Tyr Leu Glu Thr Tyr Thr Glu Asp Ala Ala Val Glu Leu
385 390 395 400

80 Ser Lys Glu Leu Asn Gly Lys Pro Asp Leu Ile Ile Gly Asn Tyr Ser

ES 2 709 439 T3

5 405 410 415
 Asp Gly Asn Leu Val Ala Ser Leu Leu Ala His Lys Leu Gly Val Thr
 420 425 430
 10 Gln Cys Thr Ile Ala His Ala Leu Glu Lys Thr Lys Tyr Pro Asp Ser
 435 440 445
 15 Asp Ile Tyr Trp Lys Lys Leu Asp Asp Lys Tyr His Phe Ser Cys Gln
 450 455 460
 20 Phe Thr Ala Asp Ile Phe Ala Met Asn His Thr Asp Phe Ile Ile Thr
 465 470 475 480
 25 Ser Thr Phe Gln Glu Ile Ala Gly Ser Lys Glu Thr Val Gly Gln Tyr
 485 490 495
 30 Glu Ser His Thr Ala Phe Thr Leu Pro Gly Leu Tyr Arg Val Val His
 500 505 510
 35 Gly Ile Asp Val Phe Asp Pro Lys Phe Asn Ile Val Ser Pro Gly Ala
 515 520 525
 40 Asp Met Ser Ile Tyr Phe Pro Tyr Thr Glu Glu Lys Arg Arg Leu Thr
 530 535 540
 45 Lys Phe His Ser Glu Ile Glu Glu Leu Leu Tyr Ser Asp Val Glu Asn
 545 550 555 560
 50 Lys Glu His Leu Cys Val Leu Lys Asp Lys Lys Lys Pro Ile Leu Phe
 565 570 575
 55 Thr Met Ala Arg Leu Asp Arg Val Lys Asn Leu Ser Gly Leu Val Glu
 580 585 590
 60 Trp Tyr Gly Lys Asn Thr Arg Leu Arg Glu Leu Ala Asn Leu Val Val
 595 600 605
 65 Val Gly Gly Asp Arg Arg Lys Glu Ser Lys Asp Asn Glu Glu Lys Ala
 610 615 620
 70 Glu Met Lys Lys Met Tyr Asp Leu Ile Glu Glu Tyr Lys Leu Asn Gly
 625 630 635 640
 75 Gln Phe Arg Trp Ile Ser Ser Gln Met Asp Arg Val Arg Asn Gly Glu
 645 650 655

ES 2 709 439 T3

5 Leu Tyr Arg Tyr Ile Cys Asp Thr Lys Gly Ala Phe Val Gln Pro Ala
 660 665 670
 10 Leu Tyr Glu Ala Phe Gly Leu Thr Val Val Glu Ala Met Thr Cys Gly
 675 680 685
 15 Leu Pro Thr Phe Ala Thr Cys Lys Gly Gly Pro Ala Glu Ile Ile Val
 690 695 700
 20 His Gly Lys Ser Gly Phe His Ile Asp Pro Tyr His Gly Asp Gln Ala
 705 710 715 720
 25 Ala Asp Thr Leu Ala Asp Phe Phe Thr Lys Cys Lys Glu Asp Pro Ser
 725 730 735
 30 His Trp Asp Glu Ile Ser Lys Gly Gly Leu Gln Arg Ile Glu Glu Lys
 740 745 750
 35 Tyr Thr Trp Gln Ile Tyr Ser Gln Arg Leu Leu Thr Leu Thr Gly Val
 755 760 765
 40 Tyr Gly Phe Trp Lys His Val Ser Asn Leu Asp Arg Leu Glu Ala Arg
 770 775 780
 45 Arg Tyr Leu Glu Met Phe Tyr Ala Leu Lys Tyr Arg Pro Leu Ala Gln
 785 790 795 800
 50 Ala Val Pro Leu Ala Gln Asp Asp
 805

40 <210> 10
 <211> 2427
 <212> ADN
 <213> Arabidopsis thaliana

45 <400> 10
 atggcaaacg ctgaacgtat gataacgogc gtccacagcc aacgtgagcg tttgaacgaa 60
 acgcttgttt ctgagagaaa cgaagtcctt gccttgcttt ccagggttga agccaaaggt 120
 50 aaaggtatatt tacaacaaaa ccagatcatt gctgaattog aagctttgcc tgaacaaacc 180
 cggaagaaac ttgaaggtgg tcctttcttt gaccttctca aatocactca ggaagcaatt 240
 gtgttgccac catgggttgc tctagctgtg aggccaaggc ctggtgtttg ggaatactta 300
 55 cgagtcaatc tccatgctct tgtcgttgaa gaactccaac ctgotgagtt tottoatttc 360
 aaggaagaac tcgttgatgg agttaagaat ggtaatttca ctcttgagct tgatttcgag 420
 ccattcaatg cgtctatccc togtccaaca ctccacaaat acattggaaa tgggtgtgac 480
 60 ttcttaacc gtcatttatc ggctaagctc ttocatgaca aggagagttt gcttcattg 540

ES 2 709 439 T3

	cttaagttcc ttcgtcttca cagccaccag ggcaagaacc tgatggtgag cgagaagatt	600
5	cagaacctca acactctgca acacaccttg aggaaagcag aagagtatct agcagagctt	660
	aagtccgaaa cactgtatga agagtttgag gccaaagttg aggagattgg tcttgagagg	720
	ggatggggag acaatgcaga gcgtgtcctt gacatgatac gtcttctttt ggaccttctt	780
10	gaggcgctg atccttgcaac tcttgagact tttcttggaag gactaccaat ggtgttcaac	840
	gttgtgatcc tctctccaca tggttacttt gctcaggaca atgttcttgg ttacctgac	900
	actggtggac aggttgttta cattcttgat caagttcgtg ctctggagat agagatgctt	960
15	caacgtatta agcaacaagg actcaacatt aaaccaagga ttctcattct aactcgactt	1020
	ctacctgatg cggtaggaac tacatgcggt gaacgtctcg agagagttta tgattctgag	1080
20	tactgtgata ttcttctgtg gcccttcaga acagagaagg gtattgttctg caaatggatc	1140
	tcaaggttcg aagtctggcc atatctagag acttacaccg aggatgctgc ggttgagcta	1200
	tcgaaagaat tgaatggcaa gcctgacctt atcattggta actacagtga tggaaatctt	1260
25	gttgcttctt tattggctca caaacttggg gtcaactcagt gtaccattgc tcatgctctt	1320
	gagaaaacaa agtaccggga ttctgatatc tactggaaga agcttgacga caagtaccat	1380
	ttctcatgcc agttcactgc ggatattttc gcaatgaacc aactgattt catcatcact	1440
30	agtactttcc aagaaattgc tggaaagcaa gaaactgttg gccagtatga aagccacaca	1500
	gcctttactc ttcccggatt gtatcgagtt gttcacggga ttgatgtgtt tgatcccaag	1560
35	ttcaacattg tctctcctgg tgctgatatg agcatctact tcccttacac agaggagaag	1620
	cgtagattga ctaagttoca ctctgagatc gaggagctcc tctacagoga tgttgagaac	1680
	aaagagcact tatgtgtgct caaggacaag aagaagcoga ttctcttcac aatggctagg	1740
40	cttgatcgtg tcaagaactt gtcaggtcct gttgagtggg acgggaagaa caccgcttg	1800
	cgtgagctag ctaacttggg tgttgttggg ggagacagga ggaagagtc aaaggacaat	1860
	gaagagaaag cagagatgaa gaaaatgtat gatctcattg aggaatacaa gctaaacggt	1920
45	cagttcaggt ggatctcctc tcagatggac cgggtaagga acggtgagct gtaccggtac	1980
	atctgtgaca ccaagggtgc ttttgtccaa cctgcattat atgaagcctt tgggttaact	2040
	gttgtggagg ctatgacttg tggtttacog actttcgcca cttgcaaagg tgggtccagct	2100
50	gagatcattg tgcacggtaa atcgggtttc cacattgacc cttaccatgg tgatcaggct	2160
	gctgatactc ttgctgattt cttcaccaag tgtaaggagg atccatctca ctgggatgag	2220
	atctcaaaag gagggttca gaggattgag gagaaataca cttggcaaat ctattcacag	2280
55	aggctcttga cattgactgg tgtgtatgga ttctggaagc atgtctcga ccttgaccgt	2340
	cttgaggctc gccgttaact tgaatgttc tatgcattga agtatgccc attggctcag	2400
60	gctgttctc ttcacacaaga tgattga	2427

65

ES 2 709 439 T3

<210> 11
 <211> 458
 <212> PRT
 <213> Stevia rebaudiana

5
 <400> 11

10 Met Glu Asn Lys Thr Glu Thr Thr Val Arg Arg Arg Arg Arg Ile Ile
 1 5 10 15

15 Leu Phe Pro Val Pro Phe Gln Gly His Ile Asn Pro Ile Leu Gln Leu
 20 25 30

20 Ala Asn Val Leu Tyr Ser Lys Gly Phe Ser Ile Thr Ile Phe His Thr
 35 40 45

25 Asn Phe Asn Lys Pro Lys Thr Ser Asn Tyr Pro His Phe Thr Phe Arg
 50 55 60

30 Phe Ile Leu Asp Asn Asp Pro Gln Asp Glu Arg Ile Ser Asn Leu Pro
 65 70 75 80

35 Thr His Gly Pro Leu Ala Gly Met Arg Ile Pro Ile Ile Asn Glu His
 85 90 95

40 Gly Ala Asp Glu Leu Arg Arg Glu Leu Glu Leu Leu Met Leu Ala Ser
 100 105 110

45 Glu Glu Asp Glu Glu Val Ser Cys Leu Ile Thr Asp Ala Leu Trp Tyr
 115 120 125

50 Phe Ala Gln Ser Val Ala Asp Ser Leu Asn Leu Arg Arg Leu Val Leu
 130 135 140

55 Met Thr Ser Ser Leu Phe Asn Phe His Ala His Val Ser Leu Pro Gln
 145 150 155 160

60 Phe Asp Glu Leu Gly Tyr Leu Asp Pro Asp Asp Lys Thr Arg Leu Glu
 165 170 175

65 Glu Gln Ala Ser Gly Phe Pro Met Leu Lys Val Lys Asp Ile Lys Ser
 180 185 190

70 Ala Tyr Ser Asn Trp Gln Ile Leu Lys Glu Ile Leu Gly Lys Met Ile
 195 200 205

75 Lys Gln Thr Lys Ala Ser Ser Gly Val Ile Trp Asn Ser Phe Lys Glu

ES 2 709 439 T3

	210		215		220														
5	Leu	Glu	Glu	Ser	Glu	Leu	Glu	Thr	Val	Ile	Arg	Glu	Ile	Pro	Ala	Pro			
	225					230					235					240			
10	Ser	Phe	Leu	Ile	Pro	Leu	Pro	Lys	His	Leu	Thr	Ala	Ser	Ser	Ser	Ser			
				245						250						255			
15	Leu	Leu	Asp	His	Asp	Arg	Thr	Val	Phe	Gln	Trp	Leu	Asp	Gln	Gln	Pro			
				260					265						270				
20	Pro	Ser	Ser	Val	Leu	Tyr	Val	Ser	Phe	Gly	Ser	Thr	Ser	Glu	Val	Asp			
			275					280						285					
25	Glu	Lys	Asp	Phe	Leu	Glu	Ile	Ala	Arg	Gly	Leu	Val	Asp	Ser	Lys	Gln			
		290					295					300							
30	Ser	Phe	Leu	Trp	Val	Val	Arg	Pro	Gly	Phe	Val	Lys	Gly	Ser	Thr	Trp			
	305					310					315					320			
35	Val	Glu	Pro	Leu	Pro	Asp	Gly	Phe	Leu	Gly	Glu	Arg	Gly	Arg	Ile	Val			
					325					330					335				
40	Lys	Trp	Val	Pro	Gln	Gln	Glu	Val	Leu	Ala	His	Gly	Ala	Ile	Gly	Ala			
				340					345						350				
45	Phe	Trp	Thr	His	Ser	Gly	Trp	Asn	Ser	Thr	Leu	Glu	Ser	Val	Cys	Glu			
			355					360						365					
50	Gly	Val	Pro	Met	Ile	Phe	Ser	Asp	Phe	Gly	Leu	Asp	Gln	Pro	Leu	Asn			
		370					375					380							
55	Ala	Arg	Tyr	Met	Ser	Asp	Val	Leu	Lys	Val	Gly	Val	Tyr	Leu	Glu	Asn			
	385					390					395					400			
60	Gly	Trp	Glu	Arg	Gly	Glu	Ile	Ala	Asn	Ala	Ile	Arg	Arg	Val	Met	Val			
					405					410					415				
65	Asp	Glu	Glu	Gly	Glu	Tyr	Ile	Arg	Gln	Asn	Ala	Arg	Val	Leu	Lys	Gln			
				420					425					430					
70	Lys	Ala	Asp	Val	Ser	Leu	Met	Lys	Gly	Gly	Ser	Ser	Tyr	Glu	Ser	Leu			
			435					440						445					
75	Glu	Ser	Leu	Val	Ser	Tyr	Ile	Ser	Ser	Leu									
		450					455												

ES 2 709 439 T3

<210> 12
 <211> 1377
 <212> ADN
 <213> Stevia rebaudiana

5

<400> 12

10 atggagaata agacagaaac aacogtaaga oggaggogga ggattatctt gttccctgta 60
 ccatttcagg gccatattaa tcogatoctc caattagcaa acgtcctcta ctccaagggg 120
 ttttcaataa caatcttcca tactaacttt aacaagccta aaacgagtaa ttatcctcac 180
 15 ttacattca ggttcattct agacaacgac cctcaggatg agogtatctc aaatttacct 240
 acgcatggcc ccttggcagg tatgcaata ccaataatca atgagcatgg agccgatgaa 300
 ctccgtcgcg agttagagct tctcatgctc gcaagtgagg aagacgagga agtttcgtgc 360
 20 ctaataactg atgogctttg gtacttggcc caatcagtcg cagactcaact gaatctaagc 420
 cgtttggtcc ttatgacaag ttcattattc aactttcacg cacatgtatc actgccgcaa 480
 25 tttgacgagt tgggttacct ggaccoggat gacaaaacgc gattggagga acaagcgtcg 540
 ggcttcccca tgctgaaagt caaagatatt aagagcgtt atagtaattg gcaaattctg 600
 aaagaaattc toggaaaaat gataaagcaa accaaagcgt cctctggagt aatctggaac 660
 30 tccttcaagg agttagagga atctgaactt gaaacggta tcagagaaat ccccgctccc 720
 tcgttcttaa ttccactacc caagcacctt actgcaagta gcagttccct cctagatcat 780
 35 gaccgaaccg tgtttcagtg gctggatcag caacccccgt cgtcagttct atatgtaagc 840
 tttgggagta cttcggaagt ggatgaaaag gacttcttag agattgcgcg agggctcgtg 900
 gatagcaaac agagcttccct gtgggtagtg agaccgggat tcgttaaggg ctcgacgtgg 960
 40 gtcgagccgt tgccagatgg ttttctaggg gagagagggg gaatcgtgaa atgggtcca 1020
 cagcaagagg ttttggctca oggagctata ggggcctttt ggaccactc tggttggaat 1080
 45 tctactcttg aaagtgtctg tgaagcgtt ccaatgatat tttctgattt tgggcttgac 1140
 cagcctctaa acgctcgcta tatgtctgat gtgttgaagg ttggcgtgta cctggagaat 1200
 ggttgggaaa ggggggaaat tgccaacgcc atacgcggg taatggtgga ogaggaaggt 1260
 50 gagtacatac gtcagaacgc tcgggtttta aaacaaaaag oggacgtcag ccttatgaag 1320
 ggaggtagct cctatgaatc cctagaatcc ttggtaagct atatatcttc gttataa 1377

55

60

65

Reivindicaciones

- 5 1. Un método para producir una composición de glucósido de esteviol o la síntesis de un compuesto de glucósido de esteviol, el método comprende incubar un sustrato con un polipéptido recombinante que es una UDP-glucosiltransferasa que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de la glucosiltransferasa de HVI como se expone en la sec. con núm. de ident. :6 en presencia de UDP-glucosa como donante de glucosa.
- 10 2. Un método de acuerdo con la reivindicación 1, en donde dicha secuencia de aminoácidos tiene:
 - (a) al menos 90 % de identidad con la sec. con núm. de ident.:6;
 - (b) al menos 95 % de identidad con la sec. con núm. de ident.: 6; o
 - (c) consiste en la secuencia de aminoácidos de la glucosiltransferasa HVI de la sec. con núm. de ident.:6.
- 15 3. Un método como se reivindica en la reivindicación 1 o la reivindicación 2, en donde dicho polipéptido recombinante se expresa por una célula huésped que se incuba con el sustrato.
- 20 4. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 3 para producir una composición de glucósido de esteviol, en donde el sustrato se selecciona del grupo que consiste en esteviósido, rebaudiósido A, rebaudiósido E y sus combinaciones.
- 25 5. Un método como se reivindica en la reivindicación 4, en donde el rebaudiósido E se convierte en rebaudiósido Z mediante dicha UDP-glucosiltransferasa, opcionalmente, dicha composición de glucósido de esteviol comprende además un compuesto de glucósido de esteviol seleccionado de rebaudiósido D, rebaudiósido E y sus combinaciones.
- 30 6. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 5, para producir una composición de glucósido de esteviol que comprende además proporcionar un edulcorante que comprende la composición de glucósido de esteviol.
- 35 7. Un método para sintetizar un compuesto de glucósido de esteviol como se reivindicó en cualquiera de las reivindicaciones 1 a 3, en donde se proporciona un compuesto seleccionado de rubusosido, esteviósido, rebaudiósido A o rebaudiósido E se proporciona como un sustrato para dicha UDP-glucosiltransferasa.
- 40 8. Un método como se reivindica en cualquiera de las reivindicaciones 1 a 7, que comprende además incubar una sacarosa sintasa recombinante con dicho sustrato y el polipéptido recombinante; opcionalmente, en donde la sacarosa sintasa recombinante comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de la sacarosa sintasa 1 de *Arabidopsis thaliana* como se expone en la sec. con núm. de ident.:9.
- 45 9. Un método como se reivindica en la reivindicación 8, que comprende además incubar una segunda UDP-glucosiltransferasa recombinante con la sacarosa sintasa, dicho sustrato y dicho polipéptido recombinante; opcionalmente, en donde también se emplea una UDP-glucosiltransferasa recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la secuencia de aminoácidos de la UGT76G1 como se expone en la sec. con núm. de ident.:11
- 50 10. Un método como se reivindica en la reivindicación 7 para sintetizar rebaudiósido Z1 y rebaudiósido Z2 a partir de rebaudiósido E, el método comprende:

50 proporcionar una mezcla de reacción que comprende rebaudiósido E y

 - (A) una UDP-glucosiltransferasa recombinante que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.:6 y la UDP-glucosa como un sustrato, o
 - (B) sacarosa, UDP y UDP-glucosa como sustratos y una UDP-glucosiltransferasa que comprende una secuencia de aminoácidos que tiene al menos 80 % de identidad con la sec. con núm. de ident.:6 junto con una sacarosa sintasa; e

55 incubar la mezcla de reacción de manera que una glucosa se acopla covalentemente al C2' de la 13-O-glucosa del rebaudiósido E para producir rebaudiósido Z1 y una glucosa se acopla covalentemente al C2' de la 19-O-glucosa del rebaudiósido E para producir rebaudiósido Z2.
- 60 11. Un método como se reivindica en la reivindicación 10, en donde se emplea una sacarosa sintasa seleccionada del grupo que consiste en la sacarosa sintasa 1 de *Arabidopsis*, una sacarosa sintasa 3 de *Arabidopsis*, una sacarosa sintasa de *Vigna radiate* y homólogos funcionales de las mismas.

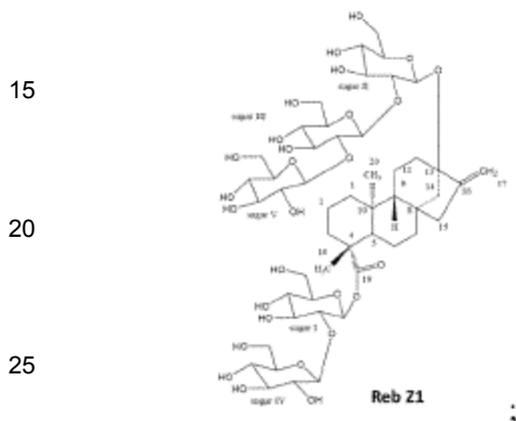
65

12. Un método como se reivindica en la reivindicación 11, en donde dicha sacarosa sintasa tiene al menos 80 % de identidad con la secuencia de aminoácidos de la sacarosa sintasa 1 de *Arabidopsis thaliana* como se expone en la sec. con núm. de ident.: 9.

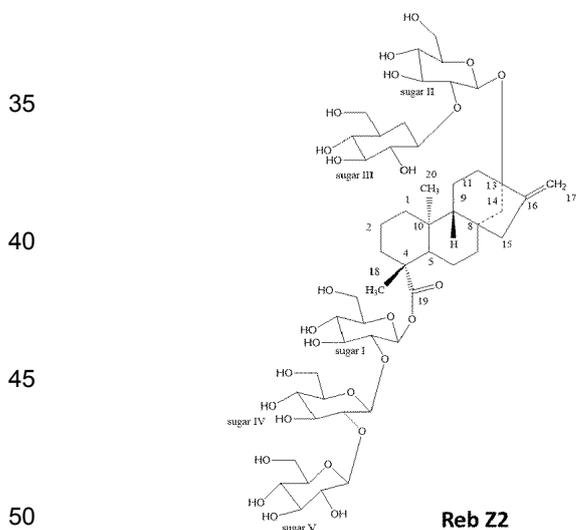
5 13. Un método como se reivindica en cualquiera de las reivindicaciones 10 a 12, que comprende además producir rebaudiósido Z como una mezcla de rebaudiósido Z1 y rebaudiósido Z2 a partir de la mezcla de reacción.

14. Un compuesto rebaudiósido Z que es

10 (a) rebaudiósido Z1 que tiene la estructura:



30 (b) rebaudiósido Z2 que tiene la estructura:



y mezclas de estos.

55 15. Un edulcorante que comprende rebaudiósido Z1 y/o rebaudiósido Z2 como se reivindica en la reivindicación 14, que comprende opcionalmente además al menos un relleno, un agente de carga o un agente antiaglomerante.

60 16. Un producto consumible que comprende una cantidad edulcorante de un edulcorante como se reivindica en la reivindicación 15, opcionalmente en donde el producto consumible se selecciona del grupo que consiste en bebidas, productos de confitería, productos de panadería, galletas y gomas de mascar.

65

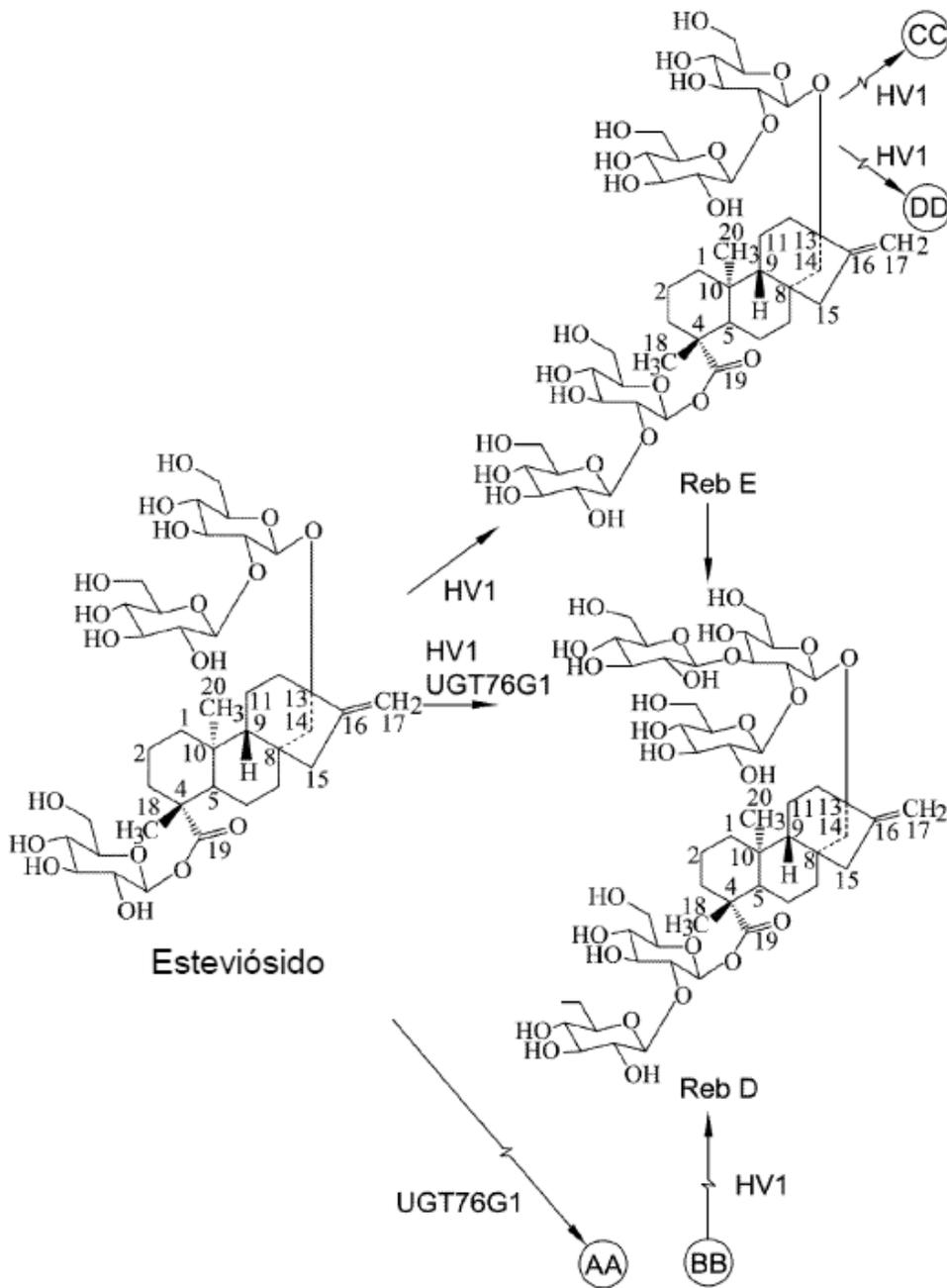


Figura 1A

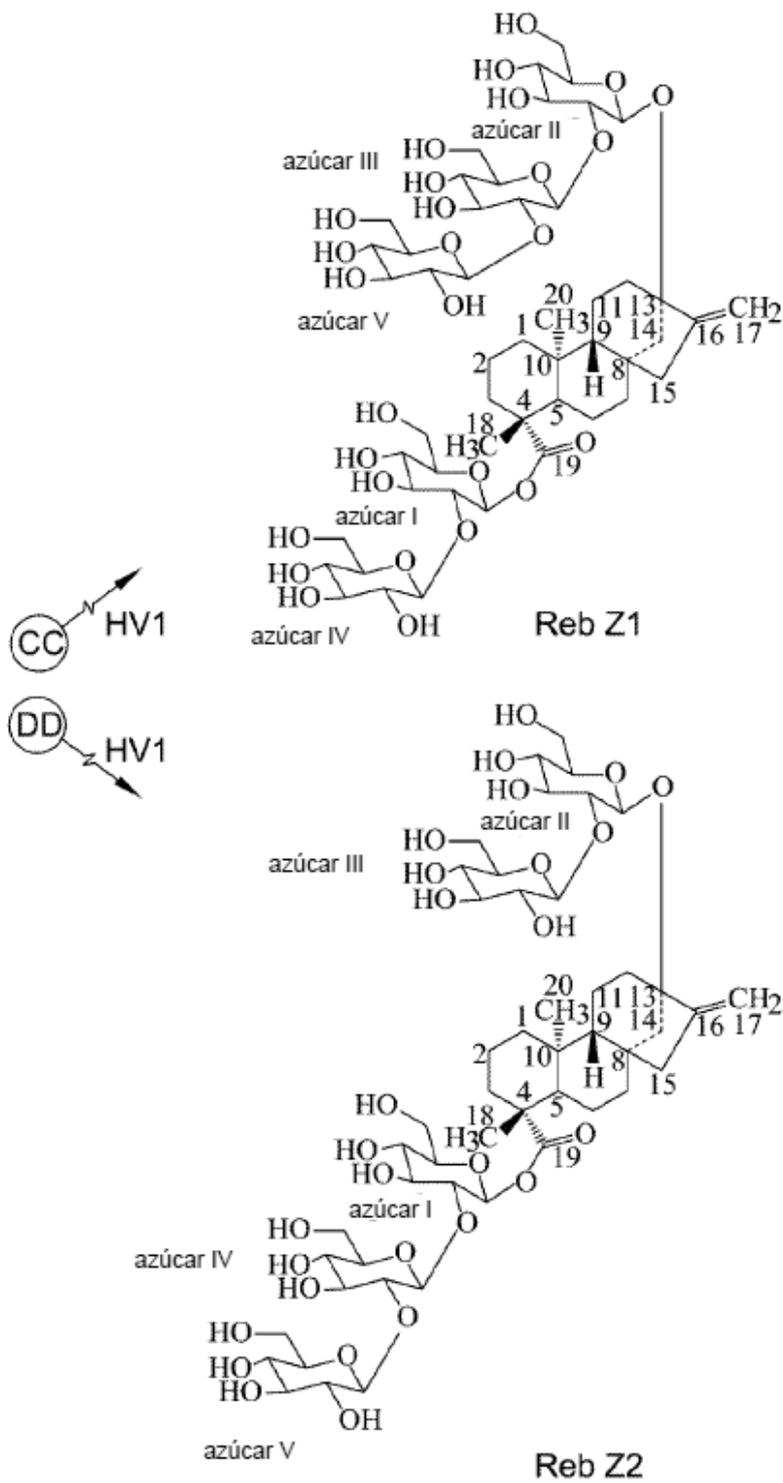


Figura 1B

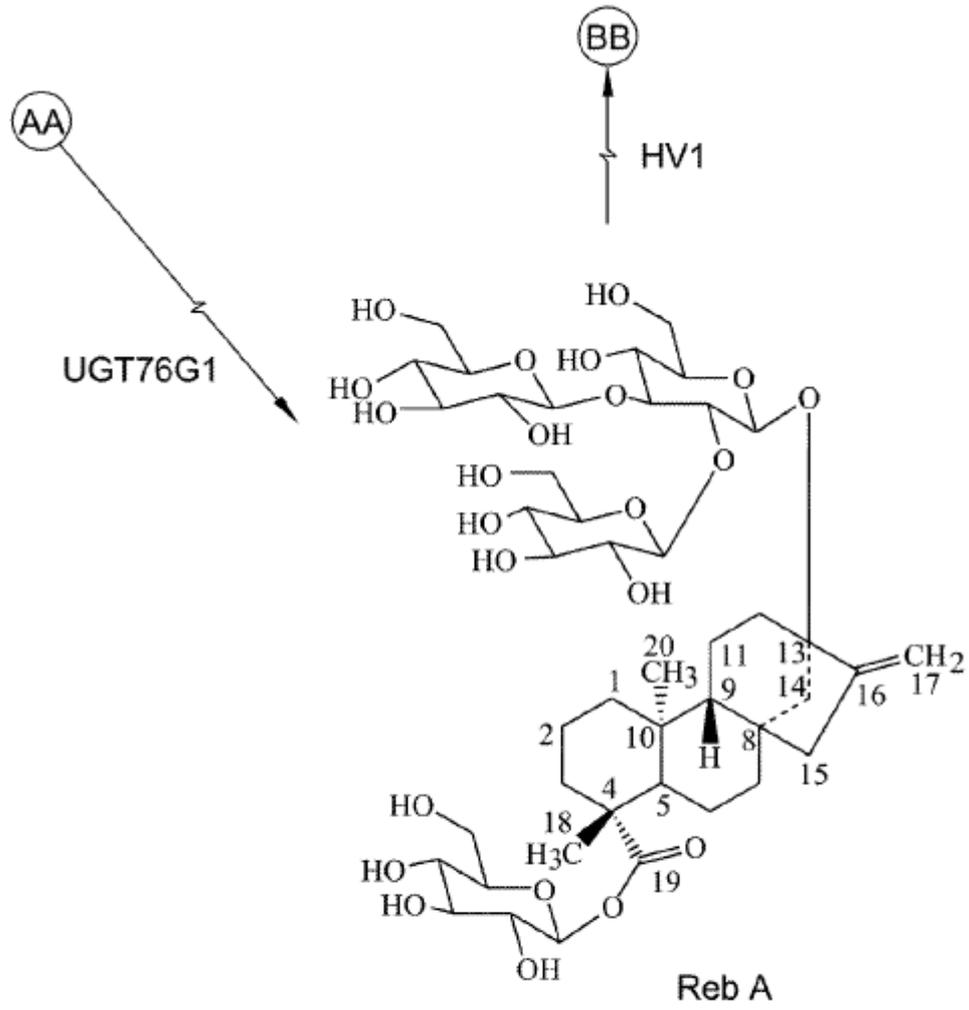


Figura 1C

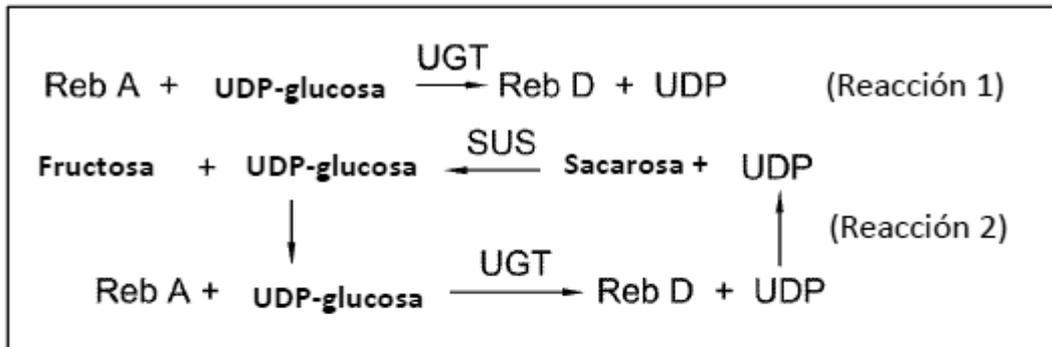


Figura 2

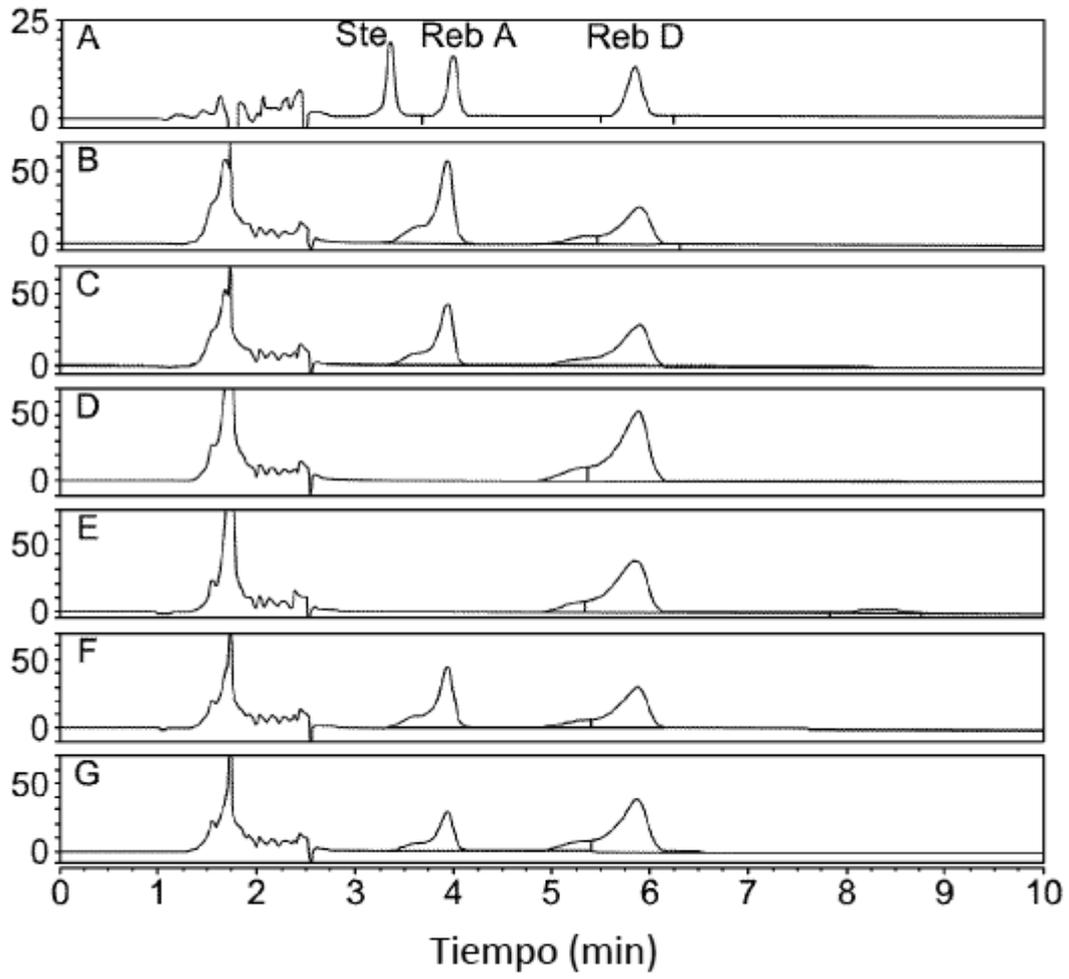


Figura 3

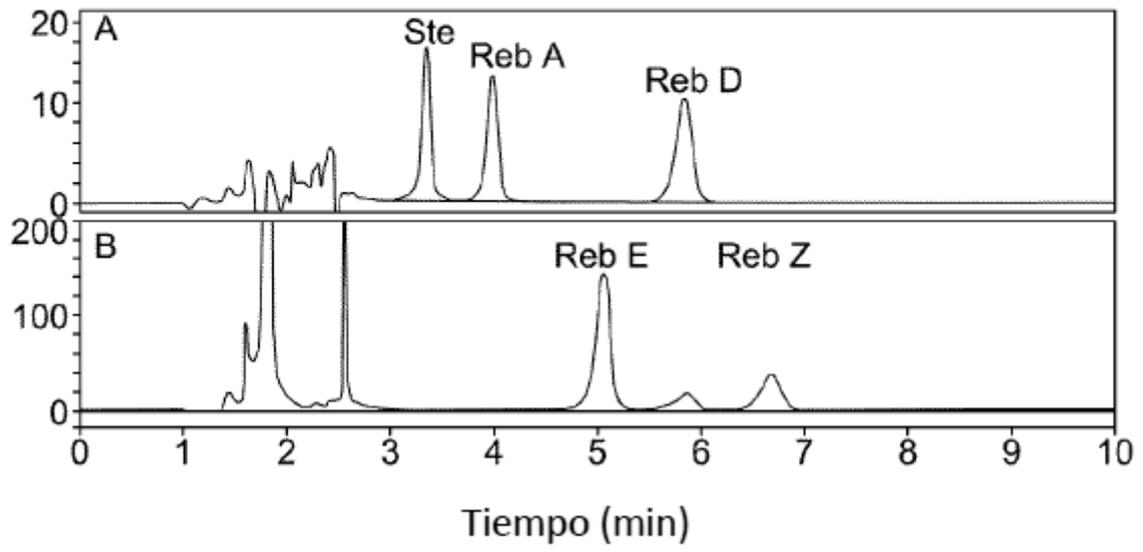


Figura 4

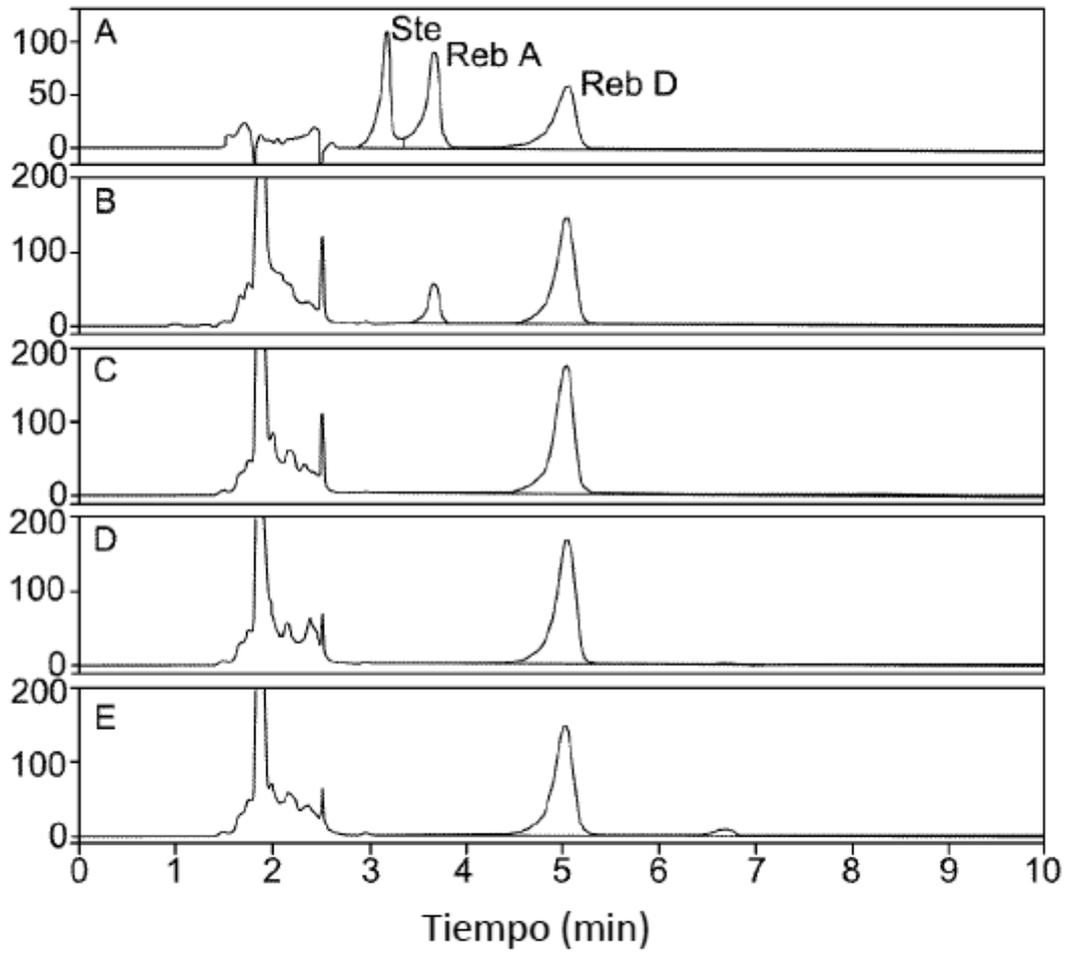


Figura 5

Figura 6

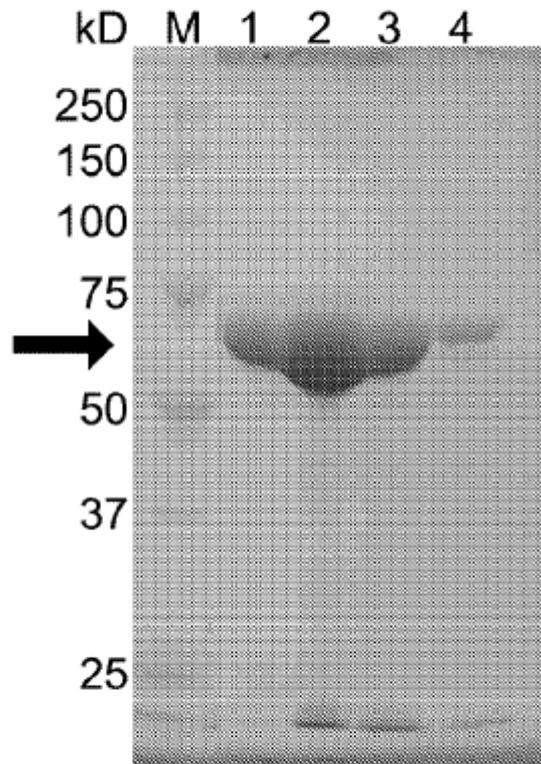
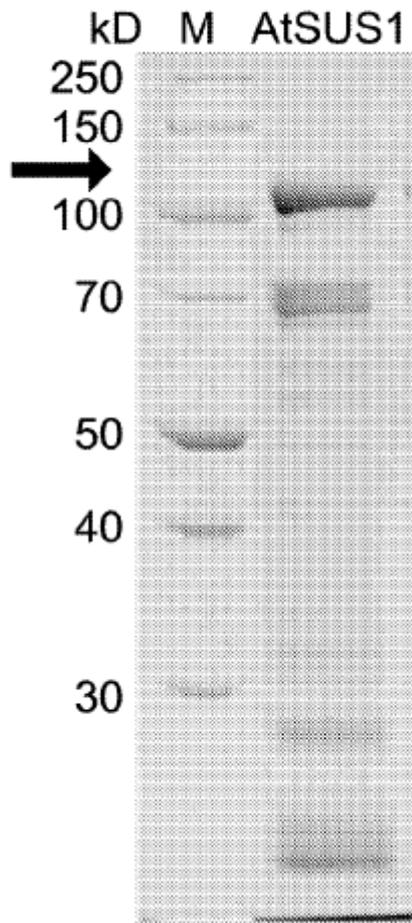


Figura 7



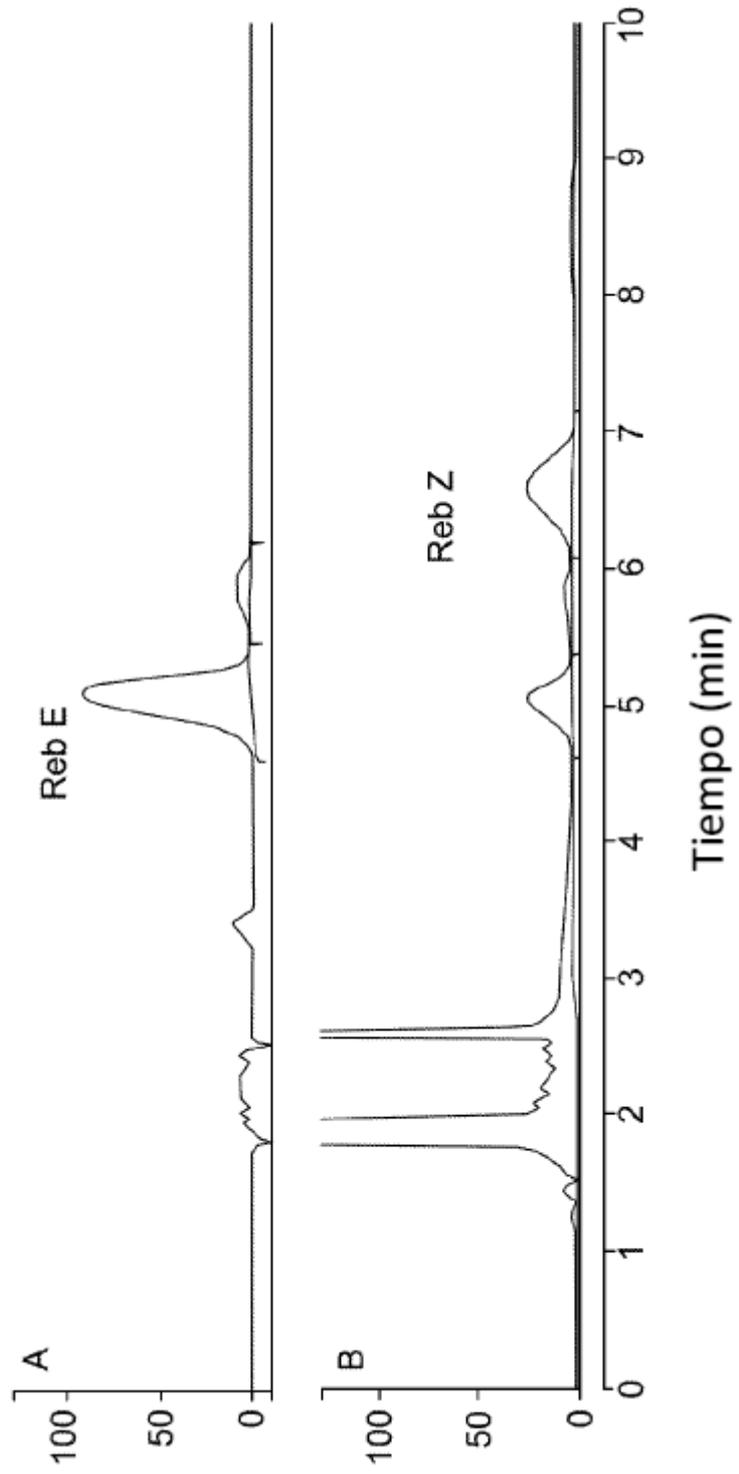
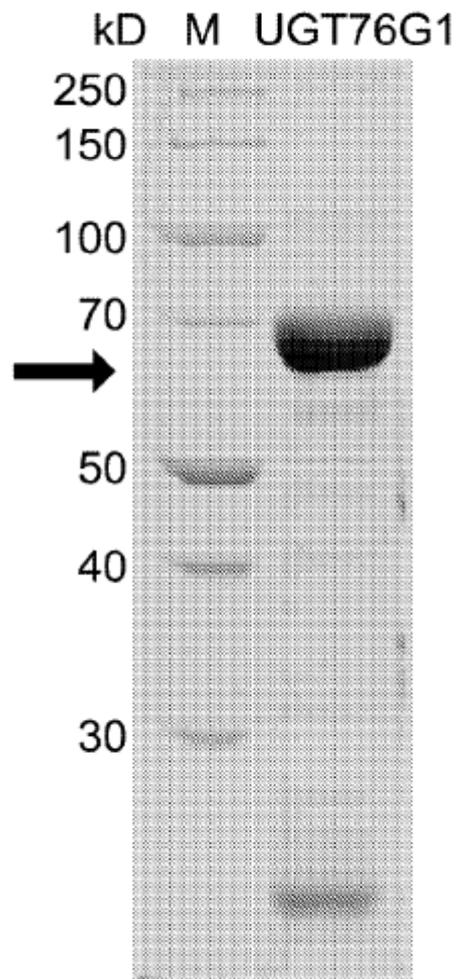


Figura 8

Figura 9



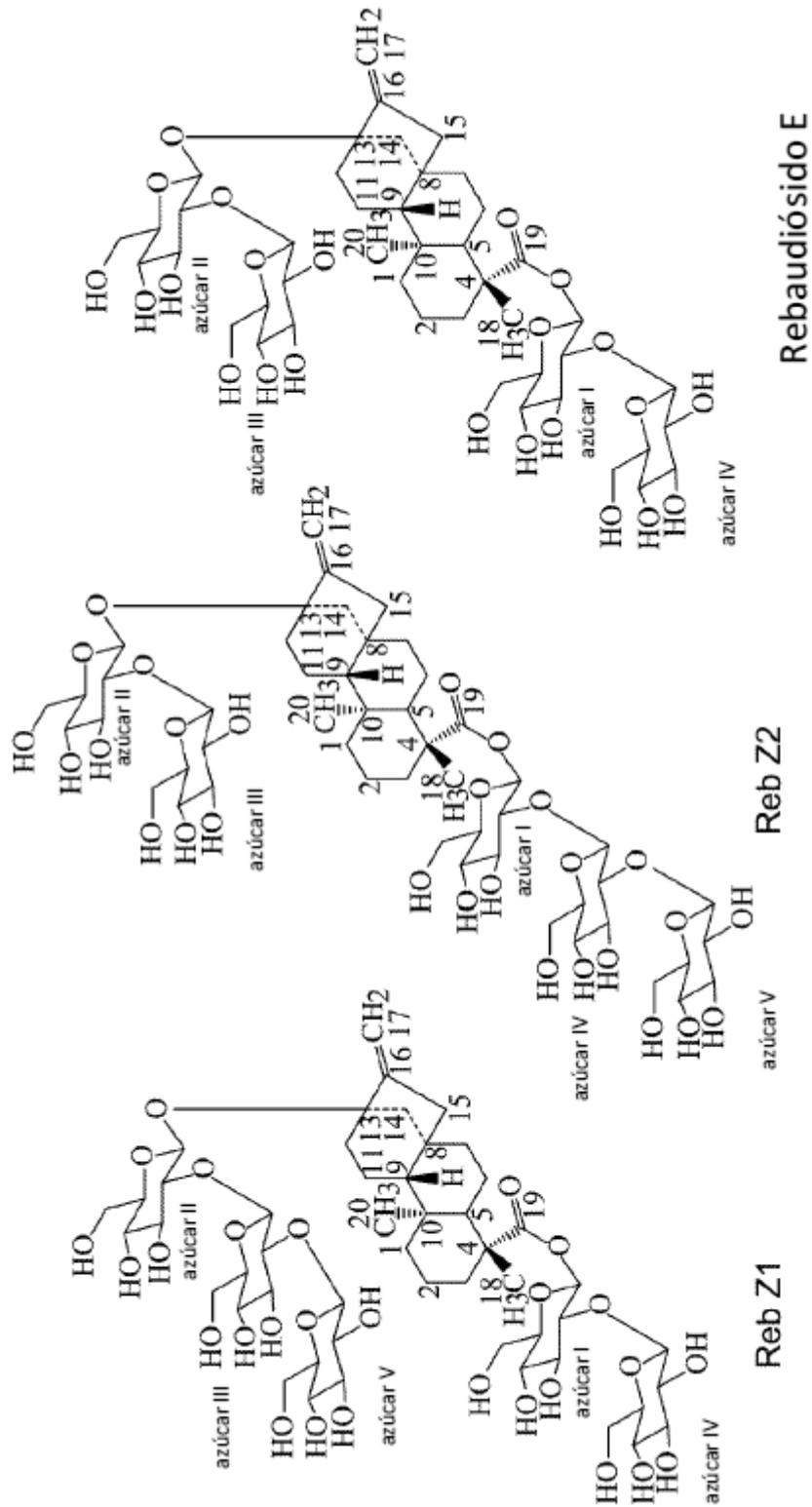


Figura 10

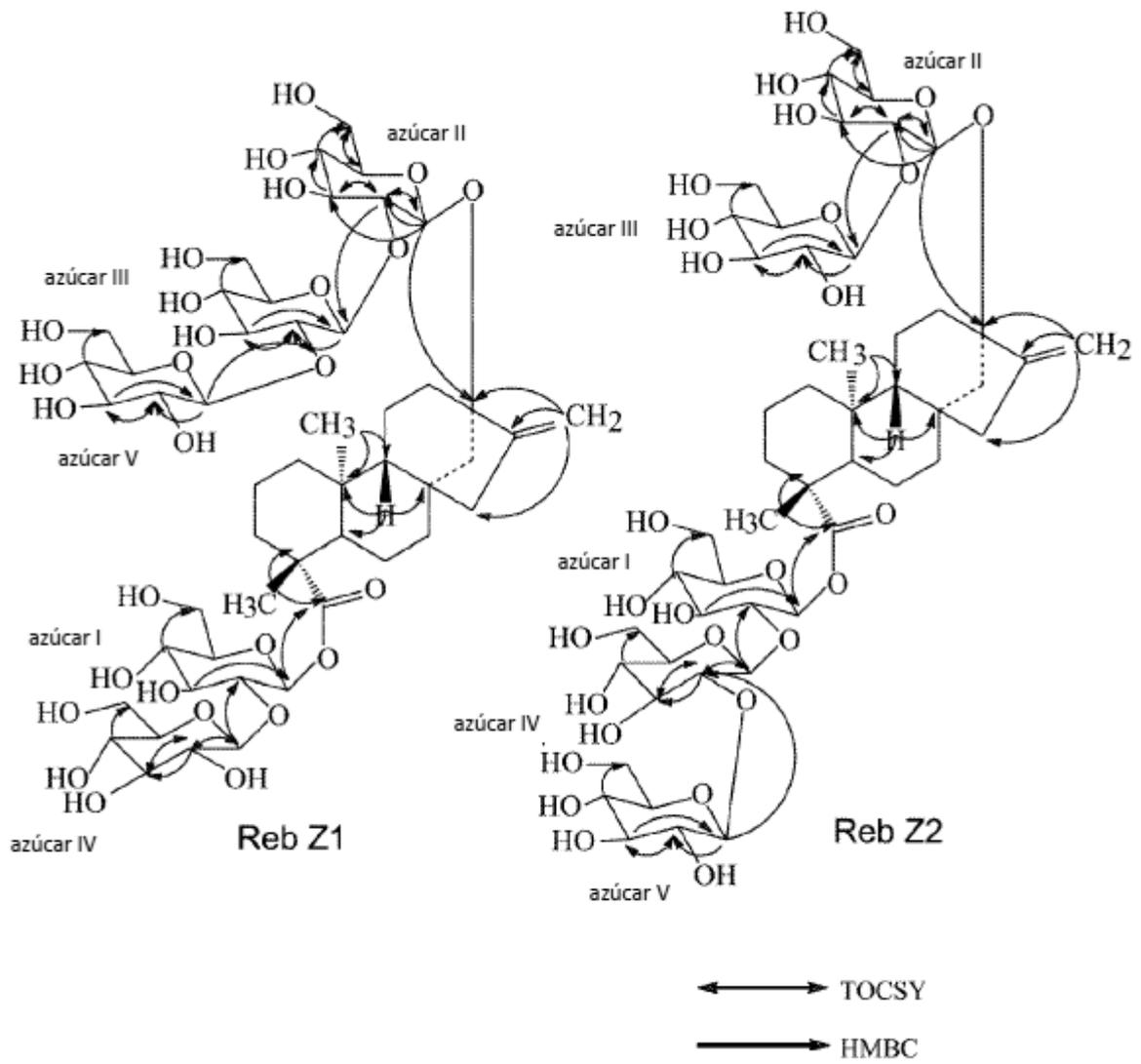


Figura 11

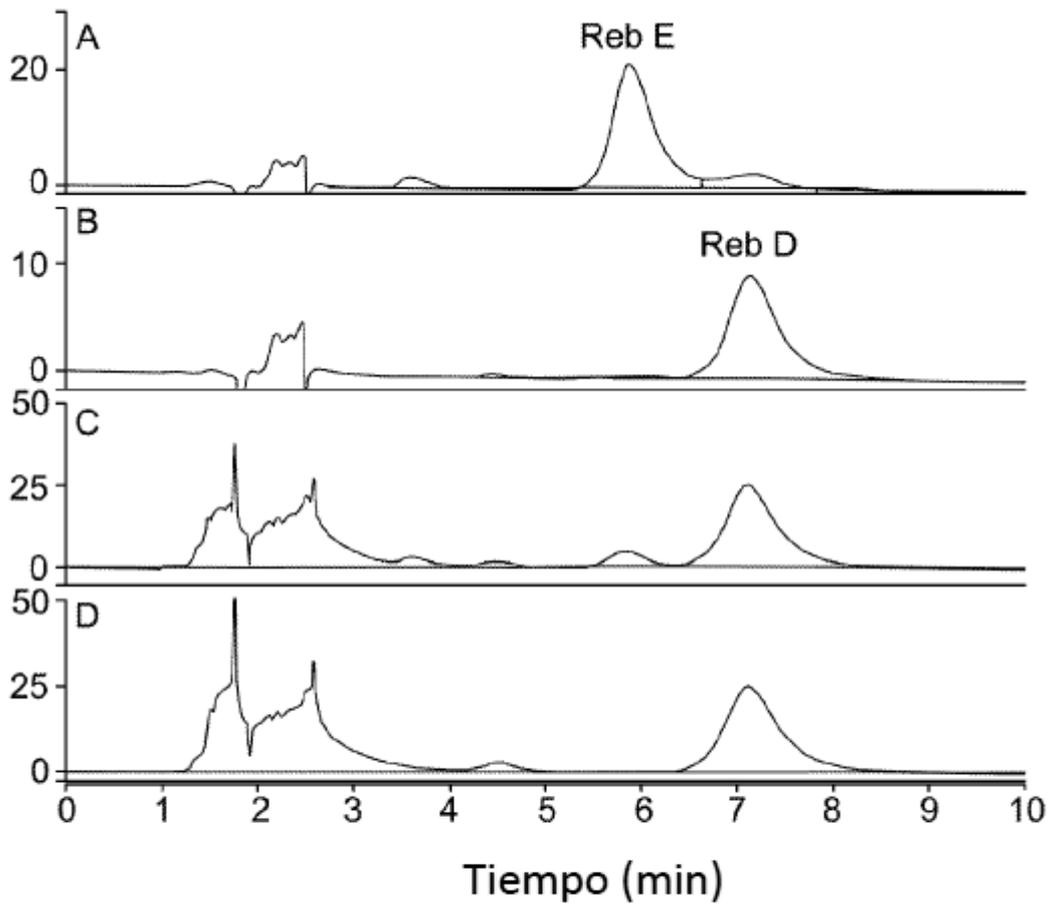


Figura 12