

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 624 776**

51 Int. Cl.:

**C12N 15/62** (2006.01)  
**A01H 5/00** (2006.01)  
**A01H 5/10** (2006.01)  
**C07K 19/00** (2006.01)  
**C12N 15/54** (2006.01)  
**C12N 15/82** (2006.01)  
**C12N 5/10** (2006.01)  
**C12N 9/10** (2006.01)  
**C12P 21/02** (2006.01)  
**C12P 21/08** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **13.06.2008 PCT/CA2008/001139**
- 87 Fecha y número de publicación internacional: **18.12.2008 WO08151440**
- 96 Fecha de presentación y número de la solicitud europea: **13.06.2008 E 08772802 (8)**
- 97 Fecha y número de publicación de la concesión europea: **10.08.2016 EP 2155880**

54 Título: **Producción modificada de glicoproteínas en plantas**

30 Prioridad:

**15.06.2007 US 944344 P**

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

**17.07.2017**

73 Titular/es:

**MEDICAGO INC. (33.0%)**  
**1020, Route de l'Eglise, bureau 600**  
**Sainte-Foy, QC G1V 3V9, CA;**  
**CENTRE NATIONAL DE LA RECHERCHE**  
**SCIENTIFIQUE (33.0%) y**  
**UNIVERSITE DE ROUEN (33.0%)**

72 Inventor/es:

**D'Aoust, Marc-André;**  
**Marquet-Blouin, Estelle;**  
**Bardor, Muriel;**  
**Burel, Carole;**  
**Faye, Loïc;**  
**Lerouge, Patrice;**  
**Vešina, Louis-Philippe;**  
**Gomord, Véronique;**  
**Aquin, Stéphanie;**  
**Rihouey, Christophe;**  
**Paccalet, Thomas y**  
**Sourrouille, Christophe**

74 Agente/Representante:

**CAMPello ESTEBARANZ, Reyes**

ES 2 624 776 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 modificada de glicoproteínas en plantas.

## 5 CAMPO DE LA INVENCION

La presente invención se refiere a métodos para modificar la producción de glicoproteínas en plantas. La presente invención también proporciona plantas con producción modificada de glicoproteínas.

## 10 ANTECEDENTES DE LA INVENCION

Las inmunoglobulinas (IgGs) son proteínas heteromultiméricas complejas con afinidad característica para homólogos antígenicos e específicos de varias naturalezas. En la actualidad, el aislamiento de rutina de líneas celulares que producen IgG, y el advenimiento de tecnologías para evolución dirigida a IgG y la ingeniería molecular han tenido un profundo impacto en su evolución como bioterapéuticos y en el mercado general de las ciencias biológicas. La IgG monoclonal terapéutica (anticuerpos monoclonales, mAb) dominan el mercado actual de nuevos fármacos antiinflamatorios y anticancerosos y cientos de nuevos candidatos están actualmente bajo investigación y desarrollo clínico para aplicaciones mejoradas y novedosas. El mercado anual de mandagamas de mAb de varios gramos (diagnóstico), algunos kilogramos (antitoxina) hasta uno a varios cientos de kilogramos (defensa bioquímica, anticanceroso, antiinfeccioso, antiinflamatorio).

Aunque el cultivo de células CHO es aún su huésped de producción preferido a escala comercial, se admite generalmente para que los mAb alcancen su impacto completo en el mercado de las ciencias biológicas, se han desarrollado sistemas de producción alternativos, ya que las instalaciones requeridas para estos cultivos no se modulan fácilmente en escala, su coste de construcción y mantenimiento son extremadamente elevados y aumentan constantemente, y su validación bajo GMP aún requiere un promedio de tres años tras la construcción. Incluso en las etapas tempranas de desarrollo, la selección de líneas celulares de CHO con rendimientos y productividad aceptables sigue siendo un proceso costoso y largo. Los nuevos sistemas de producción que podrían disminuir los costes ascendentes (rendimientos más elevados, tecnologías e infraestructuras más sencillas), tienen un tiempo de espera más corto, son más flexibles en capacidad, al mismo tiempo que reúnen las características actuales de reproducibilidad, calidad y seguridad de los sistemas de cultivo celular actuales, probablemente tienen un impacto significativo en el desarrollo de mAb y vacunas para el mercado de las ciencias biológicas, en cada una de las fases de desarrollo.

Las plantas son huéspedes adecuados para la producción de mAb y varias otras proteínas que tienen aplicaciones actuales en las ciencias biológicas (véase Koy y Koprowski 2005; Ma y col., 2005; Yusibov y col., 2006 para revisiones recientes). Los MAb se han producido en líneas vegetales transgénicas estables a rendimientos de hasta 200 mg/kg de peso fresco (FW), y mediante una expresión transitoria en índices de hasta 20 mg/kg de FW (Kathuria, 2002). Giritch y col. (2006) indican niveles de expresión de 200-300 mg/kg de peso de hoja para una IgG, con un máximo citado de 500 mg/kg a través del uso de un sistema de expresión transitoria basado en múltiples virus.

El documento WO 03/078637 desvela métodos para optimizar el procesamiento de glicanos en los organismos (y, en particular, las plantas), de manera que pueda obtenerse una glicoproteína que tenga glicanos biantenarios de tipo completo y, por lo tanto, que contenga residuos de galactosa en ambos brazos y que estén desprovistos de (o reducidos en) xilosa y fucosa.

El documento WO 2004/065540 desvela moléculas de ácido nucleico, incluyendo construcciones de fusión, que tienen actividad catalítica y el uso de las mismas en el diseño de glicosilación de células huésped para generar polipéptidos con propiedades terapéuticas mejoradas, incluyendo anticuerpos con un aumento de la unión al receptor Fc y un aumento de la función efectora.

El documento WO 01/29242 desvela métodos para producir un polipéptido heterólogo modificado postraduccionalmente en un sistema huésped vegetal alterando las capacidades de modificación postraduccionales naturales de ese sistema de huésped vegetal. En otro aspecto de este método, la alteración de las modificaciones postraduccionales naturales se hace transformando el sistema huésped vegetal con una o más secuencias de ácido nucleico que codifican una enzima de modificación postraduccionales.

El proceso de N-glicosilación de plantas y mamíferos difiere. Las últimas etapas de N-glicosilación en células de mamíferos añaden  $\beta$ 1,4galactosa,  $\alpha$ 1,6fucosa (beta-1,4galactosa, alfa-1,6fucosa) y residuos de ácido siálico terminal

a los glicanos complejos. Sin embargo, se añaden en las plantas residuos de  $\beta$ 1,3galactosa,  $\alpha$ 1,3-fucosa (beta-1,3galactosa, alfa-1,3fucosa),  $\alpha$ 1,4fucosa y  $\beta$ 1,2xilosa (alfa-1,4fucosa y beta-1,2xilosa). alfa-1,3fucosa y  $\beta$ 1,2xilosa que son constituyentes de glico-epítomos de algunos alérgenos de plantas, estos residuos se consideran potencialmente inmunógenos y es indeseada su aparición en proteínas terapéuticas, incluyendo anticuerpos.

La adición de una secuencia de KDEL al extremo C de un péptido se utiliza típicamente para asegurar la recuperación del péptido a partir del Golgi de nuevo al ER. Este enfoque se ha utilizado para producir un anticuerpo no fucosilado y no xilosilado utilizando agroinfiltración de hojas de tabaco (Sriraman y col., 2004). Sin embargo, el péptido de KDEL añadido es potencialmente inmunógeno, y este enfoque ha limitado la aplicabilidad a la producción de proteínas terapéuticas.

El control de la adición de  $\alpha$ 1,3fucosa (alfa-1,3fucosa) y  $\beta$ 1,2xilosa (beta-1,2xilosa) se ha logrado también modificando la expresión de fucosiltransferasa y xilosiltransferasa. Se han producido mutantes en un musgo que carecen de la capacidad de añadir  $\alpha$ 1,3fucosa y  $\beta$ 1,2xilosa en glicanos complejos de musgo (*Physcomitrella patens*; Koprivova y col., 2004) y *Arabidopsis thaliana* (Strasser y col., 2004). La inhibición parcial de la expresión de fucosiltransferasa y xilosiltransferasa en plantas se ha logrado también por la expresión de ARN interferentes (ARNi) dirigidos a genes  $\alpha$ 1,3fucosiltransferasa y  $\beta$ 1,2xilosiltransferasa en *Lemna minor* (Cox y col., 2006). Sin embargo, la inhibición completa de estas actividades enzimáticas tiene efectos nocivos en varias especies de plantas, ya que interfieren con eventos de desarrollo claves, tal como formación de polen y producción de semillas. La degradación específica inducida por ARNi del ARNm puede no ser estable con el paso del tiempo, y puede no ser aplicable a un espectro grande de plataformas basadas en plantas utilizadas para la producción de productos terapéuticos, ya que se ha indicado que son sensibles a factores ambientales.

El documento 03/078637 desvela el uso de galactosiltransferasa humana para catalizar la adición de  $\beta$ 1,4 galactosa terminal (GalT) en glicanos vegetales. La expresión de GalT y el direccionamiento de su actividad al cis Golgi mediante el uso de una fusión con el dominio transmembrana de una xilosiltransferasa dieron como resultado la adición de galactosa terminal y una disminución en residuos específicos de plantas que tienen N-glicanos (véase también Bakker y col., 2006). El cultivo de estas plantas con plantas que contienen una IgG recombinante dio como resultado un descenso significativo, aunque variable, de los glicanos que contenían fucosa y xilosa.

La unión de un residuo de N-acetilglucosamina unida a beta-1,4 (GlnNAc) a manosa unida a beta para producir una GlcNAc dividida se cataliza por N-acetilglucosaminiltransferasa II (GnT-III; EC 2.4.1.144). Se ha descrito la introducción de esta enzima en plantas (Rouwendaal y col., 2007), y las plantas que expresaban proteínas comprendidas por GnT-III con N-glicanos complejos se dividieron y se portaron dos residuos GlnAc.

## RESUMEN DE LA INVENCIÓN

La presente invención se refiere a métodos para modificar la producción de glicoproteínas en plantas. La presente invención también proporciona plantas con producción de glicoproteínas modificada.

Es un objeto de la invención proporcionar un método mejorado para modificar la producción de glicoproteínas en plantas.

Se proporciona en el presente documento un ácido nucleico que tiene una secuencia nucleotídica (A) que comprende los nucleótidos 1-1077 de la SEQ ID NO: 17 (GNT1-GalT; figura 5d), o que comprende una secuencia nucleotídica que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 1-1077 de la SEQ ID NO: 17, como se determina usando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: complejidad baja; Alineamiento: en pares; Tamaño de palabra: 11, en el que la secuencia nucleotídica codifica una proteína que modifica la glicosilación de una proteína de interés.

También se desvela un ácido nucleico que tiene una secuencia nucleotídica (B) que comprende una primera secuencia de ácido nucleico que comprende los nucleótidos 5-1198 de la SEQ ID NO: 14 (GalT), o que comprende una secuencia que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 5-1198 de la SEQ ID NO: 14, como se determina usando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: complejidad baja; Alineamiento: en pares; Tamaño de palabra: 11, en el que la primera secuencia de ácido nucleico codifica una proteína que modifica la glicosilación de una proteína de interés, y la primera secuencia de ácido nucleico está unida operativamente con una segunda secuencia de ácido nucleico que comprende un promotor 35S, o un promotor de plastocianina.

Se desvela un ácido nucleico que tiene una secuencia nucleotídica que comprende los nucleótidos 1-1641 de la SEQ ID NO: 26 (GNT1-GnT-III), o que comprende una secuencia nucleotídica que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 1-1641 de la SEQ ID NO: 26, como se determina usando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: baja complejidad; Alineamiento: en pares; Tamaño de palabra: 11, en el que la secuencia nucleotídica codifica una proteína que modifica la glicosilación de una proteína de interés.

También se desvela un ácido nucleico que tiene una secuencia nucleotídica que comprende una primera secuencia de ácido nucleico que comprende los nucleótidos 1-1460 de la SEQ ID NO: 16 (GnT-III), o que comprende una secuencia que es de aproximadamente el 80 % al 100 % similar con los nucleótidos 1-1460 de la SEQ ID NO: 16, como se determina usando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: baja complejidad; Alineamiento: en pares; Tamaño de palabra: 11, en el que la primera secuencia de ácido nucleico codifica una proteína que modifica la glicosilación de una proteína de interés, la primera secuencia de ácido nucleico unida operativamente con una segunda secuencia de ácido nucleico que comprende un promotor 35S, o un promotor de plastocianina.

Se proporciona también una planta, una célula vegetal, o una semilla que comprende la secuencia nucleotídica de (A) como se ha descrito anteriormente.

La presente invención también pertenece a una proteína híbrida GNT1-GalT, que comprende un dominio CTS de N-acetilglucosaminil transferasa fusionado a un dominio catalítico de beta-1,4galactosiltransferasa, y que comprende la secuencia SEQ ID NO: 18. La secuencia aminoácida de la SEQ ID NO: 18 puede codificarse por la secuencia nucleotídica SEQ ID NO: 17. Se proporciona también una planta, una célula vegetal, o una semilla que comprende la proteína híbrida como se acaba de describir. Se proporciona también una planta, una célula vegetal, o una semilla que comprende el ácido nucleico que comprende la secuencia nucleotídica SEQ ID NO: 17.

Se desvela una proteína híbrida GNT1-GnT-III, que comprende un dominio CTS de N-acetilglucosaminil transferasa fusionado a un dominio catalítico de N-acetilglucosaminiltransferasa III, y que comprende la secuencia aminoácida SEQ ID NO: 20. La secuencia aminoácida de la SEQ ID NO: 21 puede codificarse por la secuencia nucleotídica SEQ ID NO: 26. Se proporciona también una planta, una célula vegetal o una semilla que comprende la proteína híbrida como se acaba de describir. Se proporciona también una planta, una célula vegetal, o una semilla que comprende el ácido nucleico que comprende la secuencia nucleotídica SEQ ID NO: 26.

De acuerdo con la presente invención se proporciona un método (1) para sintetizar una proteína de interés que comprende expresar dentro de una planta o una porción de una planta, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica una proteína híbrida, GNT1-GalT, que comprende un dominio CTS de N-acetilglucosaminil transferasa (GNT1) fusionado a un dominio catalítico de beta-1,4galactosiltransferasa (GalT), la primera secuencia nucleotídica unida operativamente con una primera región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una segunda región reguladora que está activa en la planta, y expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con N-glicosilación modificada.

La primera secuencia nucleotídica y la segunda secuencia nucleotídica como se ha descrito anteriormente, se expresan de forma transitoria en la planta. Además, la primera región reguladora puede ser un primer promotor específico de tejidos, y la segunda región reguladora es un segundo promotor específico de tejidos. Cada uno del primer y segundo promotores específicos de tejidos puede ser un promotor de plastocianina.

La presente invención también proporciona un método (2) para sintetizar una proteína de interés. El método es como se ha descrito anteriormente (método 1), en el que la proteína de interés puede ser un anticuerpo. Si la proteína de interés es un anticuerpo, entonces la segunda secuencia nucleotídica que codifica la proteína de interés comprende una secuencia nucleotídica 2A, unida operativamente con una región reguladora 2A que está activa en la planta, y una secuencia nucleotídica 2B, unida operativamente con una región reguladora 2B que está activa en la planta, y el producto codificado para cada una de 2A y 2B se combinan para producir el anticuerpo. La región reguladora 2A puede ser un promotor de plastocianina, y la región reguladora 2B puede ser un promotor de plastocianina.

La presente invención también proporciona un método (1) o (2), como se ha descrito anteriormente, en el que una tercera secuencia nucleotídica se expresa dentro de la planta, la tercera secuencia nucleotídica que codifica un supresor de silenciamiento se une operativamente con una tercera región reguladora que está activa en la planta. La



tercera secuencia nucleotídica que codifica un supresor de silenciamiento puede ser, por ejemplo, HcPro, TEV - p1/HC-Pro, BYV-p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 o GVA-p10. El tercer promotor específico de tejidos puede ser un promotor de plastocianina.

5

Se desvela un sistema de expresión vegetal para conducir la expresión de una proteína de interés en una planta, en el que la proteína de interés comprende un patrón de glicosilación modificado. Por ejemplo, la proteína de interés puede comprender N -glicanos fucosilados, xilosilados, o ambos, fucosilados y xilosilados reducidos. Como alternativa, la proteína de interés puede comprender un patrón de glicosilación modificado, en el que la proteína carece de residuos fucosilados, xilosilados o tanto fucosilados como xilosilados, y muestra un aumento de galactosilación. Además, puede añadirse una galactosa terminal que dé como resultado una reducción o eliminación de la fucosilación y la xilosilación de la proteína de interés en comparación con la misma proteína de interés producida en una planta de tipo silvestre.

10

Se desvela un método (3) para sintetizar una proteína de interés con un perfil de N-glicosilación modificado que comprende, co-expresar dentro de una planta, una porción de una planta, o una célula vegetal, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica una proteína híbrida, GNT1-GnT-III, que comprende un dominio CTS de N-acetilglucosaminil transferasa (GNT1) fusionado a un dominio catalítico de N-acetilglucosaminiltransferasa III (GnT-III), la primera secuencia nucleotídica unida operativamente con una primera región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una segunda región reguladora que está activa en la planta, y co-expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con el perfil de N-glicosilación modificado.

20

La primera secuencia nucleotídica y la segunda secuencia nucleotídica como se ha descrito anteriormente (método 3), pueden expresarse de forma transitoria en la planta, o pueden expresarse de forma estable. Además, la primera región reguladora puede ser un primer promotor específico de tejido, y la segunda región reguladora es un segundo promotor específico de tejido. Cada uno de los primeros y segundos promotores específicos de tejido puede ser un promotor de plastocianina.

30

También se desvela un método (4) para sintetizar una proteína de interés con un perfil de N-glicosilación modificado que comprende, co-expresar dentro de una planta, una porción de una planta, o una célula vegetal, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica una proteína híbrida, GNT1-GalT, que comprende un dominio CTS de N-acetilglucosaminil transferasa (GNT1) fusionado a un dominio catalítico de beta-1,4galactosiltransferasa (GalT), la primera secuencia nucleotídica unida operativamente con una primera región reguladora que está activa en la planta, una segunda secuencia nucleotídica que codifica beta-1,4galactosiltransferasa, la segunda secuencia nucleotídica unida operativamente con una segunda región reguladora que está activa en la planta, y una tercera secuencia nucleotídica para codificar la proteína de interés, la tercera secuencia nucleotídica unida operativamente con una tercera región reguladora que está activa en la planta, y co-expresar la primera, segunda y tercera secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con el perfil de N-glicosilación modificado.

35

40

La primera secuencia nucleotídica y la segunda secuencia nucleotídica como se ha descrito anteriormente (4), pueden expresarse de forma transitoria en la planta, o pueden expresarse de forma estable. Además, la primera región reguladora puede ser un primer promotor específico de tejido, y la segunda región reguladora es un segundo promotor específico de tejido. Cada uno del primer y segundo promotores específicos de tejido puede ser un promotor de plastocianina.

45

Se desvela un método (5) para sintetizar una proteína de interés con un perfil de N-glicosilación modificado que comprende, co-expresar dentro de una planta, una porción de una planta, o una célula vegetal, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica una proteína híbrida, GNT1-GnT-III, que comprende un dominio CTS de N-acetilglucosaminil transferasa (GNT1) fusionado a un dominio catalítico de N-acetilglucosaminiltransferasa III (GnT-III), la primera secuencia nucleotídica unida operativamente con una primera región reguladora que está activa en la planta, una segunda secuencia nucleotídica que codifica N-acetilglucosaminiltransferasa III, la segunda secuencia nucleotídica unida operativamente con una segunda región reguladora que está activa en la planta, y una tercera secuencia nucleotídica para codificar la proteína de interés, la tercera secuencia nucleotídica unida operativamente con una tercera región reguladora que está activa en la planta, y co-expresar la primera, segunda y tercera secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con el perfil de N-glicosilación modificado.

50

55

La primera secuencia nucleotídica, la segunda secuencia nucleotídica y la tercera secuencia nucleotídica como se ha descrito en el método (5) anteriormente, pueden expresarse de forma transitoria en la planta, o pueden expresarse de forma estable. Además, la primera, segunda y tercera regiones reguladoras pueden ser promotores específicos de tejidos. Por ejemplo, cada uno de los promotores específicos de tejidos puede ser un promotor de plastocianina.

De acuerdo con los métodos descritos en el presente documento, una proteína de interés puede producirse, por lo tanto, con un alto rendimiento y que carece de glicanos que son conocidos por estar implicados en reacciones de hipersensibilidad, o de otro modo, estar implicados en reacciones alérgicas. Esto se logra por la co-expresión de enzimas glico-diseñadas junto con la proteína de interés, y esto da como resultado la producción de una proteína inmunógena que podría producirse dentro de la planta de tipo silvestre.

Como se describe en el presente documento, puede utilizarse un sistema de expresión simplificado para la producción de una proteína de interés utilizando un sistema de expresión transitoria, sin embargo, los métodos pueden también ser utilizados con sistemas de transformación estable. Por lo tanto, la presente invención no se limita a sistemas de expresión transitoria.

Usando la co-expresión transitoria, el sistema descrito en el presente documento evita los extensos tiempos de producción, y el proceso de selección de líneas transgénicas mutantes elite o glico-diseñadas y su uso posterior como líneas parentales (por ejemplo, como se describe por Bakker, 2005). También se evitan los problemas simultáneos a menudo encontrados con plantas mutantes o glico-diseñadas, en cuanto a productividad, producción de polen, producción de semillas (Bakker y col. 2005) y viabilidad (Boisson y col., 2005). Como se describe en el presente documento, la co-expresión de una proteína de interés con una galactosiltransferasa humana química modificante no tuvo efecto en la cinética de producción o los rendimientos.

El sistema de expresión transitoria descrito en el presente documento produce niveles de expresión que alcanzan 1,5 g de anticuerpo de alta calidad por kilogramo de peso fresco de hoja, excediendo el nivel de acumulación indicado para cualquier anticuerpo en plantas con otros sistemas de expresión, incluyendo sistemas basados en múltiples virus y plantas transgénicas.

El enfoque descrito en el presente documento, por ejemplo, que implica la expresión de GalT o GAIT-GNT1, puede usarse también con plantas transformadas de forma estable. Aunque se muestran muchas ventajas que se han descrito anteriormente, esta invención no se limita a los sistemas de expresión transitoria.

Este resumen de la invención no describe necesariamente todas las características de la invención.

## BREVE DESCRIPCIÓN DE LAS FIGURAS

Estas y otras características de la invención se harán más evidentes a partir de la siguiente descripción en la que se hace referencia a las figuras adjuntas, en las que:

La figura 1A muestra un casete basado en plastocianina ensamblado para la expresión de C5-1. R610 comprende una secuencia nucleotídica que codifica C5-1 LC y C5-1 HC-KDEL; R612 comprende una secuencia nucleotídica que codifica C5-1 LC y C5-1 HC. C5-1 LC: secuencia codificante de cadena ligera de C5-1; C5-1 HC: secuencia codificante de cadena pesada de C5-1. La figura 1B muestra la secuencia nucleotídica para el promotor de plastocianina y 5' UTR (SEQ ID NO: 23), el sitio de inicio de la transcripción se muestra en negrita, y el codón de inicio de la traducción está subrayado. La figura 1C muestra la secuencia nucleotídica para la plastocianina 3' UTRC y el terminador (SEQ ID NO: 24), el codón de terminación está subrayado.

La figura 2 muestra la acumulación del anticuerpo C5-1 en hojas de *Nicotiana benthamiana* infiltradas con R610 y R612 (casetes de expresión basados en plastocianina) con o sin la co-expresión del supresor de silenciamiento HcPro. Los valores presentados corresponden al nivel de acumulación promedio y la desviación estándar obtenida a partir de las 6 mediciones en 3 plantas (jeringa) o 6 mediciones en lotes de infiltración individuales de aproximadamente 12 plantas (250 g).

La figura 3 muestra el análisis de transferencia de proteínas de acumulación de C5-1 en extractos de plantas infiltradas con jeringa y al vacío. La figura 3A muestra la inmunotinción con una IgG de cabra anti-

ratón conjugada con peroxidasa (H+L), en extractos de plantas infiltradas con R612 (para secreción, filas 2) o con R610 (para retención de ER, filas 2). C<sub>1</sub>: Se cargaron 100 ng de IgG1 murina comercial (Sigma M9269), como un control para la movilidad electroforética; C<sub>2</sub>: 12 µg de proteínas totales extraídas de biomasa infiltrada con control (vector vacío). C<sub>3</sub>: 100 ng de IgG1 murina comercial (Sigma M9269) añadidos en 12 µg de proteína total extraída de biomasa infiltrada con control (vector vacío). La figura 3B muestra la inmunotinción de la actividad con una IgG1 humana conjugada de peroxidasa, en extractos de plantas infiltradas con R612 (para secreción, filas 1) o con R610 (para retención de ER, filas 2). C<sub>1</sub>: 2 µg de C5-1 de control purificado a partir de hibridoma (Khoudi y col., 1997); C<sub>2</sub>: 75 µg de proteínas totales extraídas de biomasa infiltrada con control (vector vacío).

La figura 4 muestra un análisis de anticuerpos purificados a partir de plantas infiltradas ya sea con R612 (para secreción, filas 1) o R610 (para retención de ER, filas 2). La figura 4A muestra un análisis SDS-PAGE de extractos en bruto y anticuerpos purificados realizado en condiciones no reductoras. La figura 4B muestra un análisis SDS-PAGE de anticuerpos purificados realizado en condiciones reductoras. La figura 4C muestra la inmunotinción de la actividad de anticuerpos purificados realizada con una IgG1 humana conjugada con peroxidasa. La figura 4D muestra una comparación de contaminantes en 6 lotes de C5-1 purificada a partir de diferentes lotes de infiltración. C: 2,5 µg de IgG1 murina comercial (Sigma M9269), cargada como un control para movilidad electroforética.

La figura 5A muestra una representación de ejemplos de cassetes ensamblados para versiones nativas (R622) e híbridas (R621) de expresión de galactosiltransferasa. GNT1-CTS: dominio CTS de N-acetilglucosaminiltransferasa I; GalT-Cat: dominio catalítico de β1,4galactosiltransferasa humana; GalT (de R622): β1,4galactosiltransferasa humana. La figura 5B muestra la secuencia nucleotídica (SEQ ID NO: 14) para GalT (UDP-Gal:β1,4-galactosiltransferasa 1, β1,4-galactosiltransferasa I), el sitio de inicio ATG está subrayado; el dominio transmembrana está subrayado y en cursiva; la secuencia en negrita corresponde al dominio catalítico de β1,4GalT humana; el epítipo FLAG está en cursiva. La figura 5C muestra la secuencia aminoacídica (SEQ ID NO: 15) para GalT (UDP-Gal:β1,4-galactosiltransferasa 1, β1,4-galactosiltransferasa I). El dominio transmembrana está subrayado y en cursiva; la secuencia en negrita corresponde al dominio catalítico de β1,4GalT humana; el epítipo FLAG está en cursiva. La figura 5D muestra la secuencia nucleotídica (SEQ ID NO: 17) de GNT1GalT, el sitio de inicio ATG está subrayado; el dominio transmembrana (CTS) está subrayado y en cursiva; la secuencia en negrita corresponde al dominio catalítico de β1,4GalT humana; el epítipo FLAG está en cursiva. La figura 5E muestra la secuencia aminoacídica (SEQ ID NO: 18) de GNT1GalT, el dominio transmembrana (CTS) está subrayado y en cursiva; la secuencia en negrita corresponde al dominio catalítico de β1,4GalT humana; el epítipo FLAG está en cursiva. La figura 5F muestra la secuencia nucleotídica del dominio CTS (colocalización citoplasmática, dominio transmembrana, región transmembrana) de N-acetilglucosaminiltransferasa (GNT1; SEQ ID NO: 21). La figura 5G muestra el aminoácido del CTS (SEQ ID NO: 22). La figura 5H muestra la secuencia de ácido nucleico de GntI-Gnt III (SEQ ID NO: 26). La figura 5I muestra la secuencia de ácido nucleico de GntI-Gnt II (SEQ ID NO: 20). La figura 5J muestra la secuencia de ácido nucleico de Gnt III (SEQ ID NO: 16). La figura 5K muestra la secuencia de ácido nucleico de Gnt III (SEQ ID NO: 19).

La figura 6 muestra un perfil de extractos obtenidos a partir de plantas que expresan C5-1 y están teñidos para una proteína, o se someten a análisis Western. El panel superior muestra un gel PAGE teñido con Coomassie. El segundo desde el panel superior muestra afinodetección usando a glutinina de *Erythrina cristagalli* (ECA) que se une específicamente a β1,4galactosa. El tercer panel desde el superior muestra un análisis de transferencia de Western usando anticuerpos anti-α1,3-fucosa. El panel inferior muestra un análisis de transferencia de Western usando anticuerpos específicos de anti-β1,2xilosa. R612: C5-1 expresado en solitario; R612+R622-. C5-1 co-expresado (co-infiltrado) con GalT; R612+R621: C5-1 co-expresado con GNT1-GalT.

La figura 7 muestra espectrometría de masas MALDI-TOF para determinar la N-glicosilación del glicopéptido triptico EEQFNSTFR (SEQ ID NO: 13) de C5-1 aislado a partir de plantas infiltradas con jeringa *Nicotiana benthamiana* con un rango de construcciones. La N-glicosilación del glicopéptido se determinó después de la separación en el análisis por HPLC preparativa. Se obtuvieron resultados similares después de la infiltración al vacío. La figura 7A muestra la espectrometría de masas MALDI-TOF del glicopéptido triptico después de la expresión de R612 (C5-1; véase la figura 1). La figura 7B muestra una espectrometría de masas MALDI-TOF del glicopéptido triptico después de la expresión de C5-1 (R612;

véase la figura 1) junto con GalT nativa (R622, véase la figura 5). La figura 7C muestra la espectrometría de masas MALDI-TOF del glicopéptido triptico después de la expresión de C5-1 (R612, véase la figura 1); con GNT1GalT (R621, véase la figura 5). La figura 7C-inserción corresponde al alargamiento m/z 2650-2800 del espectro para mostrar la ausencia del ión J complejo principal detectado en la planta R612 (flecha). A: GlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub>; B: Man<sub>5</sub>GlcNAc<sub>2</sub>; C: GalGlcNAcMan<sub>3</sub>GlcNAc<sub>2</sub>; D: GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>; E: Man<sub>6</sub>GlcNAc<sub>2</sub>; F: GalGlcNAcMan<sub>3</sub>(Xyl)GlcNAc<sub>2</sub>; G: GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub>; H: GlcNAc<sub>2</sub>Man<sub>3</sub>(Fuc)GlcNAc<sub>2</sub>; I: Man<sub>7</sub>GlcNAc<sub>2</sub>; J: GlcNAc<sub>2</sub>Man<sub>3</sub>(Xyl)(Fuc)GlcNAc<sub>2</sub>; K: GalGlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub>; L: Man<sub>8</sub>GlcNAc<sub>2</sub>; M: Man<sub>9</sub>GlcNAc<sub>2</sub>.

## 10 DESCRIPCIÓN DETALLADA

La presente invención se refiere a métodos para modificar la producción de glicoproteínas en plantas. La presente invención también proporciona plantas con producción de glicoproteínas modificada.

- 15 Se desvela un sistema de expresión de plantas para conducir la expresión de una proteína de interés en una planta. Con el sistema de expresión descrito, puede obtenerse una proteína de interés que comprende un patrón de glicosilación modificado, por ejemplo, con una reducción de N-glicanos fucosilados, xilosilados o tanto fucosilados como xilosilados. Como alternativa, puede obtenerse una proteína de interés que comprende un patrón de glicosilación modificado, en el que la proteína carece de fucosilación, xilosilación, o ambas, y comprende un
- 20 aumento de la galactosilación. Además, como se describe en el presente documento, la modulación de modificaciones de postraduccionales, por ejemplo, la adición de la galactosa terminal, da como resultado una reducción de la fucosilación y xilosilación de la proteína expresada de interés, por ejemplo, la proteína de interés puede comprender menos del 10 % de fucosilación y xilosilación (es decir, menos del 10 % de los residuos N-glicano están fucosilados y xilosilados), o menos del 5 % de fucosilación y xilosilación (es decir menos del 5 % de los
- 25 residuos N-glicano están fucosilados y xilosilados), menos del 1 % de fucosilación y xilosilación (menos del 1 % de los residuos de N-glicano están fucosilados y xilosilados), de aproximadamente el 0,1 a aproximadamente el 2 % de los residuos de N-glicano están fucosilados y xilosilados, de aproximadamente el 0,5 a aproximadamente el 1,5 % de los residuos de N-glicano están fucosilados y xilosilados, o de aproximadamente el 0,7 a aproximadamente el 1,0 % de los residuos de N-glicano están fucosilados y xilosilados, en comparación con la misma proteína de interés
- 30 producida en una planta de tipo silvestre. Por lo tanto, una proteína de interés puede producirse con un alto rendimiento y carecer de glicanos que puedan provocar reacciones de hipersensibilidad, o de otra manera, estar implicada en reacciones alérgicas.

Un ejemplo no limitante de una proteína de interés a expresar incluye una proteína compleja, tal como un anticuerpo.

- 35 La expresión de tal proteína compleja dentro de la planta agroinfiltrada, por ejemplo, *Nicotiana benthamiana*, produjo niveles de proteínas que alcanzaban 1,5 g/kg de FW (aproximadamente el 25 % de TSP). Los niveles promedio de 558 y 757 mg/kg/FW se lograron para las formas secretadas y retenidas por ER de la proteína de interés, respectivamente. En el ejemplo no limitante proporcionado, este nivel de expresión se obtuvo para un anticuerpo, a un nivel de expresión tres veces más elevado que para un anticuerpo producido utilizando un sistema de expresión
- 40 transitorio de múltiples virus (Giritch y col. 2006).

El impacto de la diferencia entre la planta y la N-glicosilación de mamífero típica ha sido un problema principal que rodea el concepto del uso de plantas para la producción de productos terapéuticos. La aparición de glicanos específicos de plantas puede contribuir a acortar la semivida de una proteína hecha con plantas en el torrente

45 sanguíneo, o que los mismos glicanos provoquen reacciones de hipersensibilidad en pacientes.

- La presencia de un núcleo de  $\alpha$ 1,3fucosa y  $\beta$ 1,2xilosa en glicoproteínas hechas en plantas se percibe como un desafío regulador por la industria, ya que también se encuentran en algunos alérgenos vegetales. Además, se documenta ahora que la eliminación de fucosa central, incluso  $\alpha$ 1,6fucosa que puede encontrarse en las IgG de
- 50 células CHO, aumentará la actividad de ADCC. Una característica del sistema descrito en el presente documento es la capacidad para lograr la modulación de modificaciones postraduccionales, por ejemplo, la adición simultánea de galactosa terminal y la reducción o inhibición de la fucosilación y la xilosilación (en comparación con la misma proteína producida en una planta de tipo silvestre). Como alternativa, el grado de fucosilación puede reducirse mientras se aumenta la cantidad de galactosilación, de nuevo en comparación con la misma proteína producida en
- 55 una planta de tipo silvestre.

La modulación en la cantidad de fucosilación, xilosilación o galactosilación puede determinarse utilizando cualquier método adecuado, por ejemplo, utilizando anticuerpos anti- $\alpha$ 1,3fucosa, para detectar la presencia o ausencia de inmunosímbolos específicos de fucosa (fucosilación) y anticuerpos anti- $\beta$ 1,2-xilosa para detectar xilosilación, o la



presencia o ausencia de inmunoseñales específicas de xilosa, por ejemplo, como se muestra en la figura 6. Como alternativa, la espectrometría de masas MALDI-TOF puede utilizarse para determinar el perfil de N-glicosilación de una proteína o una porción de la proteína, como se muestra en la figura 7. Puede también utilizarse otro método para determinar el perfil de N-glicanos de una proteína o porción de la proteína conocida por un experto en la técnica.

Como se describe en más detalle a continuación, se usó un sistema de agroinfiltración basado en vacío y se encontró adecuado para la producción de una proteína de interés, por ejemplo, un anticuerpo en cuanto a cantidad, calidad y reproducibilidad. El paso a una tecnología de infiltración ampliable permite la producción de gramos de este anticuerpo por día dentro de una pequeña unidad piloto, la cual permite el uso de tal sistema de expresión transitoria para la producción de materiales para ensayos clínicos dentro de plazos extremadamente cortos y para el suministro de un producto permitido con un tamaño de mercado en kilogramos por año. Se obtuvieron anticuerpos de alta calidad a partir de hojas infiltradas después de una única etapa cromatográfica de afinidad. Sin embargo, se entenderá que los métodos descritos en el presente documento pueden también aplicarse a plantas que se transforman de forma estable.

El silenciamiento génico postranscripción (PTGS) puede estar involucrado en la limitación de la expresión de transgenes en plantas, y puede utilizarse la coexpresión de un supresor de silenciamiento, por ejemplo, pero sin limitación, del virus Y de patata (HcPro) para contrarrestar la degradación específica de los ARNm transgénicos (Brigneti y col., 1998). Los supresores alternativos de silenciamiento son bien conocidos en la técnica y pueden utilizarse como se describe en el presente documento (Chiba y col., 2006, Virology 346: 7-14), por ejemplo, pero sin limitación, TEV -p1/Hc-Pro (virus del grabado del tabaco-p1/Hc-Pro), BYV -p21, p19 del virus de enanismo arbustivo del tomate (TBSV p19), proteína de la cápside del virus de rizado del tomate (TCV-CP), 2b del virus del mosaico del pepino; CMV-2b), p25 del virus X de la patata (PVX-p25), p11 del virus M de la patata (PVM-p11), p11 del virus S de la patata (PVS-p11), p16 del virus de quemadura del arándano, (BScV-p16), p23 del virus de la tristeza de los cítricos (CTV-p23), p24 del virus-2 asociado con el enrollamiento de hoja de la vid, (GLRaV-2 p24), p10 del virus A de la vid, (GVA-p10), p14 del virus B de la vid (GVB-p14), p10 del virus latente de Heracleum (HLV-p10) o p16 del virus latente común del ajo (GCLV-p16). Por lo tanto, un supresor de silenciamiento, por ejemplo, HcPro, TEV -p1/Hc-Pro, BYV -p21, TBSV p19, TCV-CP, CMV-2b, PVX-p25, PVM-p11, PVS-p11, BScV-p16, CTV-p23, GLRaV-2 p24, GBV-p14, HLV-p10, GCLV-p16 o GVA-p10, puede coexpresarse junto con GalT, GNT1-GalT, GntT-III, GNT1-GntT-III, o una combinación de las mismas, para asegurar además altos niveles de producción de proteína dentro de una planta.

La tinción de Coomassie de productos purificados producidos utilizando expresión transitoria muestra la presencia de varios contaminantes de baja abundancia. Estos fragmentos parecen ser el producto relacionado, y todos los contaminantes por encima de 70 kDa contenían al menos un Fab como se muestra por la transferencia de actividad (figura 3B). La identidad y cantidad de los contaminantes relacionados con el producto presentes en extractos de plantas son similares a los observados en sistemas de producción de células de mamíferos. Por lo tanto, una serie de purificación típicamente utilizada para la purificación de anticuerpos terapéuticos (por ejemplo, intercambio aniónico, afinidad e intercambio catiónico) produce fácilmente la pureza requerida por las agencias reguladoras para una proteína para uso terapéutico.

La presente invención proporciona un método para la síntesis de una proteína de interés dentro de plantas que se caracterizan con un patrón de glicosilación modificado. El método implica co-expresar la proteína de interés junto con una secuencia nucleotídica que codifica beta-1,4galactosiltransferasa (GalT; SEQ ID NO: 14), por ejemplo, pero sin limitación, GalT de mamífero, o GalT humano, sin embargo, puede utilizarse la GalT de otras fuentes. El dominio catalítico de GalT (por ejemplo, nucleótidos 370-1194 de la SEQ ID NO: 14, secuencia en negrita en la figura 5b, o nucleótidos 238-1062 de la SEQ ID NO: 17, secuencia en negrita en la figura 5d) se fusiona a un dominio CTS (es decir, la cola citoplásmica, el dominio de transmembrana, región tallo) de N-acetilglucosaminil transferasa (GNT1; por ejemplo, que comprende los nucleótidos 34-87 de la SEQ ID NO: 17; figura 5d) y que codifica la secuencia aminoacídica que comprende los aminoácidos 12-29 de la SEQ ID NO: 18 (figura 5e), para producir una enzima híbrida GNT1-GalT, y la enzima híbrida puede co-expresarse con la proteína de interés. Se desvela la co-expresión de la proteína de interés junto con una secuencia nucleotídica que codifica N-acetilglucosaminiltransferasa III (GntT-III; SEQ ID NO: 16; figura 5j), por ejemplo, pero sin limitación, GntT-III de mamífero o GntT-III humana, puede utilizarse también GntT-III de otras fuentes. Adicionalmente, una enzima híbrida GNT1-GntT-III (SEQ ID NO: 26; figura 5d), que comprende el CTS de GNT1 fusionado a GntT-III puede también utilizarse y se describe a continuación.

Los métodos alternativos para anclar la secuencia nucleotídica que codifica GalT mGalT, o hGalT o GalT a partir de otras fuentes, incluyen fusionar la GalT a HDEL, KDEL (ambas secuencias de retención de retículo endoplásmico),

el CTS de una proteína implicada en la biosíntesis de N-glicoproteína, por ejemplo, pero sin limitación, CTS de glucosidasa I, CTS de glucosidasa I, CTS de manosidasa I, CTS de manosidasa I, CTS de beta,1,2xililtransferasa, CTS de alfa, 1,2fucosiltransferasa. El dominio catalítico de GalT puede también fusionarse a HDEL, KDEL (ambas secuencias de retención de retículo endoplásmico), el CTS de glucosidasa I, CTS de glucosidasa II, CTS de manosidasa I, CTS de manosidasa II, CTS de beta, 1,2xililtransferasa, CTS de alfa, 1,2fucosiltransferasa.

El uso de una enzima híbrida comprende GNT1-GalT o la secuencia GNT1-GnT-III, sitúa el dominio catalítico de GalT, o GnT-III, en el aparato cis-Golgi donde tienen lugar etapas tempranas en la maduración de N-glicanos complejos. Sin desear quedar ligado por la teoría, la actividad de GalT secuestrante en una fase temprana de maduración de glicanos puede dar como resultado la adición de  $\beta$ 1,4galactosa en glicanos de maduración y da como resultado la inhibición eficiente de la fucosilación y la xilosilación de la proteína que, de otro modo, podría tener lugar dentro de la planta. De forma similar, la actividad de GnT-III secuestrante en una fase temprana de maduración de glicanos puede dar como resultado la adición de residuos de GlcNAc en manosa unida a beta para producir una GlnAc de bi-antena, y da como resultado la inhibición eficiente de la fucosilación y la xilosilación de la proteína que, de otro modo, podría tener lugar dentro de la planta. Por ejemplo, la proteína de interés puede co-expresarse con una enzima híbrida que comprende un dominio CTS fusionado a un dominio catalítico de GalT, por ejemplo, GNT1-GalT (R621; figuras 5a, 5d), o un dominio catalítico de GnT-III, por ejemplo, GNT1-GnT-III (SEQ ID NO: 26). Si se desea una proteína de interés que comprenda niveles reducidos de fucosilación, mientras comprende aún proteínas xilosiladas y galctosiladas, entonces la GalT no modificada (nativa) puede co-expresarse con la proteína de interés. Si se desea una proteína de interés que comprenda glicosilación modificada que comprenda residuos de GlnAc bi-antena, entonces la GnT-III no modificada (nativa), puede co-expresarse con la proteína de interés. Como apreciará un experto en la técnica, una secuencia de ácido nucleico optimizada en plantas puede usarse para producir una enzima de GalT o GnT-III no modificada o nativa.

Por lo tanto, se proporciona una secuencia nucleotídica que comprende los nucleótidos 1-1062 de la SEQ ID NO: 17 (GNT1-GalT), o que comprende una secuencia nucleotídica que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 1-102 de la SEQ ID NO: 17, en la que la secuencia nucleotídica codifica una proteína que modifica la glicosilación de una proteína de interés. La secuencia puede estar optimizada para plantas. Se desvela también una secuencia nucleotídica que comprende una secuencia de ácidos nucleicos que comprende los nucleótidos 1-1224 de la SEQ ID NO: 14 (GalT), o que comprende una secuencia que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 1-1224 de la SEQ ID NO: 14, en la que la secuencia de ácidos nucleicos codifica una proteína que modifica la glicosilación de una proteína de interés. La secuencia puede estar optimizada para plantas. La identidad de secuencias se determina utilizando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: complejidad baja; Alineamiento: en pares; tamaño de palabra: 11.

También se desvela una secuencia aminoacídica como se expone en la SEQ ID NO: 18 (GNT1-GalT; figura 5e), o la SEQ ID NO: 15 (GalT; figura 5c).

Se desvela un ácido nucleico que tiene una secuencia nucleotídica que comprende los nucleótidos 1-1641 de la SEQ ID NO: 26 (GNT1-GnT-III), o que comprende una secuencia nucleotídica que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 1-1641 de la SEQ ID NO: 26, en el que la secuencia de ácidos nucleicos codifica una proteína que modifica la glicosilación de una proteína de interés. La secuencia puede estar optimizada para plantas. Se proporciona también una secuencia nucleotídica que comprende una secuencia de ácidos nucleicos que comprende los nucleótidos 232-1641 de la SEQ ID NO: 26 (GnT-III; o los nucleótidos 1-1460 de la SEQ ID NO: 16), o que comprende una secuencia que muestra de aproximadamente el 80 % al 100 % de identidad con los nucleótidos 232-1641 de la SEQ ID NO: 26 (o los nucleótidos 1-1460 de la SEQ ID NO: 16), en los que la secuencia de ácidos nucleicos codifica una proteína que modifica la glicosilación de una proteína de interés. La secuencia puede estar optimizada para plantas. La identidad de secuencia se determina usando los siguientes parámetros: Programa: blastn; Base de datos: nr; Esperado 10; filtro: baja complejidad; Alineamiento: en pares; tamaño de palabra: 11, en la que la secuencia nucleotídica codifica una proteína que modifica la glicosilación de una proteína de interés.

Se desvela también una secuencia aminoacídica como se expone en la SEQ ID NO: 20 (GNT1-GnT-III, figura 5i) o la SEQ ID NO: 19 (GnT-III, figura 5k).

Por "glicosilación modificada" de una proteína de interés significa que el perfil de N-glicano de la proteína de interés que comprende glicosilación modificada (por ejemplo, como se describe en el presente documento), es diferente del perfil de N-glicano de la proteína de interés producida en una planta de tipo silvestre. La modificación de glicosilación

puede incluir un aumento o una disminución en uno o más de un glicano de la proteína de interés, o la bisección de GlnAc. Por ejemplo, la proteína de interés puede mostrar una reducción de xilosilación, una reducción de fucosilación, o tanto una reducción de xilosilación como de fucosilación. Como alternativa, el perfil de N-glicano de la proteína de interés puede modificarse de una manera tal que la cantidad de galactosilación sea aumentada, y

5 opcionalmente, la cantidad de xilosilación, fucosilación, o ambas, se reduzcan. Adicionalmente, puede producirse la GlnAc dividida en dos y esto puede dar como resultado la reducción de la cantidad de fucosilación y xilosilación de la proteína.

Por "reducción de la xilosilación" o "reducción de la fucosilación" de una proteína de interés, significa que la cantidad

10 de xilosilación, fucosilación, o tanto xilosilación como fucosilación de N-glicanos detectables en la proteína de interés, es menor del 10 % que la cantidad de xilosilación, fucosilación o ambas, que es detectable en la proteína de interés cuando se produce dentro de una planta de tipo silvestre, y donde la proteína de interés se aísla, y donde se determina la xilosilación o la fucosilación, utilizando el mismo método. Por ejemplo, la proteína de interés puede comprender menos del 5 % de residuos de N-glicano que están fucosilados, xilosilados, o ambos, menos del 1 % de

15 los residuos de N-glicano detectables en la proteína de interés pueden estar fucosilados, xilosilados, o ambos, de aproximadamente el 0,1 a aproximadamente el 2 % de los residuos de N-glicano que son detectables en la proteína de interés pueden estar fucosilados y xilosilados, de aproximadamente el 0,5 a aproximadamente el 1,5 % de los residuos de N-glicano están fucosilados y xilosilados, o de aproximadamente el 0,7 a aproximadamente el 1,0 % de los residuos de N-glicano están fucosilados y xilosilados, en comparación con la misma proteína de interés

20 producida en una planta de tipo silvestre.

Como se muestra en las figuras 6 y 7, al utilizar los métodos descritos en el presente documento, puede producirse una proteína de interés que muestra un perfil de glicosilación modificado. Por ejemplo, una proteína de interés con residuos de mucosa o xilosa inmunológicamente indetectables se ha producido cuando la proteína de interés se co-

25 expresa con GNT1-GalT. El análisis MALDI-TOF de un epítipo de una proteína de interés de muestra que una proteína de interés con un patrón de glicosilación modificado puede obtenerse cuando la proteína de interés se co-expresa con GalT o GNT-GalT (véanse las figuras 7A-C, inserción C). Por ejemplo, la figura 7A, muestra el perfil de glicosilación de un epítipo de una proteína de interés expresada dentro de una planta de tipo silvestre. La proteína de interés comprende varios residuos xilosilados y fucosilados (picos H y J, respectivamente, figura 7A). Estos

30 residuos se reducen o están ausentes cuando la proteína de interés se co-expresa con GalT (figura 7B). Además, se observan nuevos residuos xilosilados (pico F), así como un aumento de los residuos galactosilados en la proteína de interés, al co-expresarse con GalT (véase los picos C, F, K, figura 7B). La co-expresión de la proteína de interés con GNT1-GalT da como resultado el perfil de N-glicanos que está caracterizado por que tiene menos del 1 % de residuos xilosilados y fucosilados (véase la figura 7C-inserción), y un aumento de residuos galactosilados (pico K;

35 figura 7C).

Por lo tanto, se desvela un método (Método A) para sintetizar una proteína de interés con N-glicosilación modificada que comprende, proporcionar una planta que comprende una secuencia nucleotídica que codifica una primera

40 secuencia nucleotídica que codifica beta-1,4galactosiltransferasa humana (GalT), la primera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, que hace crecer la planta, y expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia que codifica GalT junto con la segunda secuencia que codifica la proteína de interés, da como

45 resultado la adición de beta-1,4galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación, la xilosilación, o ambas, de los glicanos de la proteína de interés. Además, con este método, el grado de galactosilación de la proteína de interés puede aumentarse en comparación con la cantidad de galactosilación de la proteína de interés producida en una planta de tipo silvestre que no expresa GalT.

Se desvela un método (Método B) para sintetizar una proteína de interés con N-glicosilación modificada que implica expresar dentro de una planta o una porción de una planta, en una manera transitoria, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica beta-1,4galactosiltransferasa humana (GalT; SEQ ID

50 NO: 14; figura 5b), la primera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia

55 nucleotídica unida operativamente con una región reguladora que está activa en la planta, y expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con N-glicosilación modificada. La co-expresión de la primera secuencia que codifica GalT junto con la segunda secuencia que codifica la proteína de interés, da como resultado la adición de beta-1,4-galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación, la xilosilación, o ambas, de los glicanos de la

proteína de interés. Además, el grado de galactosilación de la proteína de interés puede aumentarse en comparación con la cantidad de galactosilación de la proteína de interés producida en una planta de tipo silvestre que no expresa GalT. La etapa de expresión puede implicar co-expresar de forma transitoria la primera y segunda secuencia nucleotídica, o co-expresar de forma estable la primera y segunda secuencia nucleotídica.

5

La presente invención proporciona un método (Método C) para sintetizar una proteína de interés con N-glicosilación modificada que comprende, proporcionar una planta que comprende una primera secuencia nucleotídica híbrida que codifica una primera proteína híbrida que comprende un dominio CTS de N-acetilglucosaminil transferasa (GNT1) fusionado al dominio catalítico de GalT (beta-1,4galactosiltransferasa humana; GNT1-GalT; SEQ ID NO: 17; figura 10 5d), la primera secuencia nucleotídica híbrida unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, que hace crecer la planta, y expresar la primera y segunda secuencias nucleotídicas híbridas, para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia híbrida que codifica GNT1-GalT 15 junto con la segunda secuencia que codifica la proteína de interés da como resultado la adición de beta-1,4 galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación y la xilosilación de los glicanos de la proteína de interés. Por ejemplo, el grado de fucosilación, xilosilación, o ambas, puede reducirse de aproximadamente el 0,5 a aproximadamente el 5 %, o de aproximadamente el 0,5 a aproximadamente el 2 % de la cantidad en comparación con la proteína de interés producida en una planta de tipo 20 silvestre que no expresa GNT1-GalT.

Se desvela un método alternativo adicional (Método D) para sintetizar una proteína de interés con N-glicosilación modificada que implica expresar dentro de una planta, o una porción de una planta, de una manera transitoria, una secuencia nucleotídica que codifica una primera secuencia nucleotídica híbrida que codifica GNT1-GalT, la primera 25 secuencia nucleotídica híbrida unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia híbrida que codifica GNT1-GalT junto con la segunda secuencia que codifica la proteína de 30 interés, da como resultado la adición de beta-1,4 galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación y la xilosilación de los glicanos de la proteína de interés. Por ejemplo, el grado de fucosilación, xilosilación, o ambas, puede reducirse de aproximadamente el 0,5 a aproximadamente el 5 %, o de aproximadamente el 0,5 a aproximadamente el 2 %, de la cantidad en comparación con la proteína de interés producida en una planta de tipo silvestre que no expresa GNT1-GalT. La etapa de expresión puede implicar 35 la co-expresión transitoria de la primera y segunda secuencia nucleotídica, o la co-expresión estable de la primera y segunda secuencia nucleotídica.

Se desvela un método alternativo adicional (Método E) para sintetizar una proteína de interés con N-glicosilación modificada que comprende, proporcionar una planta que comprende una secuencia nucleotídica que codifica una 40 primera secuencia nucleotídica que codifica beta-1,4galactosiltransferasa humana (GalT), la primera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y una tercera secuencia nucleotídica que codifica un supresor del silenciamiento génico, por ejemplo HcPro, la tercera secuencia nucleotídica unida operativamente con una región 45 reguladora que está activa en la planta, hacer crecer la planta, y expresar la primera, segunda y tercera secuencias nucleotídicas para sintetizar una proteína de interés que comprende glicanos con N-glicosilación modificada. La co-expresión de la primera secuencia que codifica GalT junto con la segunda secuencia que codifica la proteína de interés, da como resultado la adición de beta-1,4 galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación y la xilosilación de los glicanos de la proteína de interés, en comparación 50 con la proteína de interés producida usando una planta de tipo silvestre que no expresa GalT. El grado de galactosilación de la proteína de interés puede aumentarse en comparación con la cantidad de galactosilación de la proteína de interés producida en una planta de tipo silvestre que no expresa GalT. La expresión de la tercera secuencia que codifica un supresor de silenciamiento, asegura altos rendimientos de la galactosiltransferasa y la proteína de interés.

55

También se desvela un método para sintetizar una proteína de interés (Método F) con N-glicosilación modificada que implica expresar dentro de una planta, o una porción de una planta, de una manera transitoria, una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica beta-1,4galactosiltransferasa humana (GalT), la primera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la



planta, una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y una tercera secuencia nucleotídica que codifica un supresor del silenciamiento génico, por ejemplo HcPro, la tercera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y expresar la primera, segunda y tercera secuencias nucleotídicas para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia que codifica GalT junto con la segunda secuencia que codifica la proteína de interés, da como resultado la adición de beta-1,4 galactosa en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación y la xilosilación de los glicanos de la proteína de interés. El grado de galactosilación de la proteína de interés puede aumentarse en comparación con la cantidad de galactosilación de la proteína de interés producida en una planta de tipo silvestre que no expresa GalT. La expresión de la tercera secuencia que codifica un supresor de silenciamiento, asegura altos rendimientos de la galactosiltransferasa y la proteína de interés. La etapa de expresión puede implicar la co-expresión transitoria de la primera y segunda secuencia nucleotídica, o la co-expresión estable de la primera y segunda secuencia nucleotídica.

Por lo tanto, se desvela un método (Método G) para sintetizar una proteína de interés con N-glicosilación modificada que comprende, proporcionar una planta que comprende una secuencia nucleotídica que codifica una primera secuencia nucleotídica que codifica N-acetilglucosaminiltransferasa (GnT-III), la primera secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, hacer crecer la planta, y expresar la primera y segunda secuencias nucleotídicas para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia que codifica GnT-III junto con la segunda secuencia que codifica la proteína de interés, da como resultado la adición de residuos de N-acetilglucosamina unida a beta-1,4 (GlnAc) a manosa unida a beta (GlnAc de bisección) en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación, la xilosilación, o ambas, de los glicanos de la proteína de interés, en comparación con la proteína de interés producida en una planta de tipo silvestre que no expresa GnT-III. La secuencia que codifica GnT-III (primera secuencia nucleotídica), la proteína de interés (segunda secuencia nucleotídica), o tanto la primera como la segunda secuencia nucleotídica, pueden expresarse de forma transitoria. Si las secuencias se expresan de forma transitoria, también puede usarse una tercera secuencia nucleotídica que codifica un supresor de silenciamiento génico, por ejemplo HcPro, unida operativamente con una región reguladora activa en la planta, y la primera, segunda y tercera secuencias nucleotídicas se expresan para sintetizar una proteína de interés que comprende glicosilación modificada.

También se desvela un método alternativo (Método H) para sintetizar una proteína de interés con N-glicosilación modificada que comprende, proporcionar una planta que comprende una primera secuencia nucleotídica híbrida que codifica una primera proteína híbrida que codifica un dominio CTS de N-acetilglucosaminiltransferasa (GNT1) fusionado al dominio catalítico de GnT-III (N-acetilglucosaminiltransferasa; GNT1-GnT-III SEQ ID NO: 20), la primera secuencia nucleotídica híbrida unida operativamente con una región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la segunda secuencia nucleotídica unida operativamente con una región reguladora que está activa en la planta, hacer crecer la planta, y expresar la primera y segunda secuencias nucleotídicas híbridas para sintetizar una proteína de interés que comprende N-glicosilación modificada. La co-expresión de la primera secuencia híbrida que codifica GNT1-GnT-III junto con la segunda secuencia que codifica la proteína de interés da como resultado la adición de residuos de N-acetilglucosamina unida a beta-1,4 (GlnAc) a manosa unida a beta (GlnAc de bisección) en glicanos de maduración de la proteína de interés, reduciendo de esta manera la fucosilación y la xilosilación de los glicanos de la proteína de interés. Por ejemplo, el grado de fucosilación, xilosilación, o ambas, puede reducirse de aproximadamente el 0,5 a aproximadamente el 5 %, o de aproximadamente el 0,5 a aproximadamente el 2 %, de la cantidad en comparación con la proteína de interés producida en una planta de tipo silvestre que no expresa GNT1-GnT-III. La secuencia que codifica GNT1-GnT-III (primera secuencia nucleotídica), la proteína de interés (segunda secuencia nucleotídica), o tanto la primera como la segunda secuencia nucleotídica, pueden expresarse de forma transitoria. Si las secuencias se expresan de forma transitoria, también puede usarse una tercera secuencia nucleotídica que codifica un supresor de silenciamiento génico, por ejemplo HcPro, unida operativamente con una región reguladora activa en la planta, y la primera, segunda y tercera secuencias nucleotídicas se expresan para sintetizar la proteína de interés que comprende glicosilación modificada.

Pueden hacerse modificaciones adicionales a la secuencia nucleotídica que codifica la proteína de interés para asegurar un alto rendimiento. Por ejemplo, la segunda secuencia que codifica la proteína de interés puede también fusionarse a una secuencia que codifica una secuencia activa para retener la proteína dentro del retículo endoplásmico (ER), por ejemplo, pero sin limitación, una secuencia de KDEL (Lys-Asp-Glu-Leu), u otras secuencias

de retención de ER conocidas, por ejemplo, HDEL o KKSS.

Además, cuando una proteína de interés compleja se produce, la segunda secuencia nucleotídica, utilizada en cualquiera de los Métodos A-H como se ha descrito anteriormente, puede codificar más de un péptido o dominio de la proteína compleja. Por ejemplo, en el caso donde la proteína de interés es un anticuerpo, la segunda secuencia nucleotídica puede comprender dos segundas secuencias nucleotídicas 2A y 2B, codificando cada una, una porción del anticuerpo, por ejemplo, el nucleótido 2A puede codificar una cadena ligera y una secuencia 2B codifica una cadena pesada del anticuerpo. Los ejemplos no limitantes de tales construcciones se proporcionan en la figura 1, donde la construcción de cada una de R 216 y R610 comprende dos segundas secuencias nucleotídicas 2A, que codifican C5-1 LC (la cadena ligera de C5-1) unida operativamente a una región reguladora activa en una planta, por ejemplo, pero sin limitación, el promotor de plastocianina, y 2B que codifica la cadena pesada de C5-1 (C5-1 HC) unida operativamente a una región reguladora activa en una planta, por ejemplo, pero sin limitación, el promotor de plastocianina que comprende los nucleótidos 556 -999 de la figura 1b o la SEQ ID NO: 23; documento US 7.125.978). Como se muestra en la figura 1, y con referencia a R610, una secuencia de KDEL puede fusionarse a la región C-terminal de uno de los péptidos 2A o 2B, por ejemplo, que no se considerará limitante, la secuencia de KDEL puede fusionarse a la cadena pesada del anticuerpo para asegurar que el anticuerpo se retenga con el ER.

Cada proteína codificada por la segunda secuencia nucleotídica puede ser glicosilada.

La proteína de interés así producida por cualquiera de los Métodos A-H o puede ser recuperada de la planta. Además, la proteína de interés puede purificarse parcialmente o purificarse utilizando técnicas estándares como se conocerá por un experto en la técnica.

En los casos donde las secuencias nucleotídicas se co-expresan dentro de la planta, cada una de las secuencias nucleotídicas deseadas puede ser introducida en la planta utilizando técnicas de transformación estándares, técnicas de transformación transitoria, dos o más de dos plantas, expresando cada una, una o más de las secuencias nucleotídicas deseadas que pueden cruzarse para obtener una planta que co-expresa la combinación requerida de secuencias nucleotídicas, o puede combinarse una combinación de las técnicas anteriores. Por ejemplo, la expresión transitoria puede realizarse usando una planta transformada de forma estable expresando una secuencia que codifica GalT, GNT1-GalT, GalT y GNT1-GalT, Gnt-III, GNT1-Gnt-III, GNT1-Gnt-III y Gnt-III, o una combinación de las mismas.

Por lo tanto, se desvela un método (Método I) para producir una planta que puede utilizarse como una plataforma para la producción de una proteína de interés con N-glicosilación modificada. Este método comprende, proporcionar una planta que expresa una o más de una primera secuencia nucleotídica que codifica GalT, GNT1-GalT, o tanto GalT como GNT1-GalT, y que expresa la una o más secuencias nucleotídicas. Con el fin de producir la proteína de interés, una segunda secuencia nucleotídica que codifica la proteína de interés se introduce dentro de la planta de plataforma utilizando técnicas estándares que implican la transformación estándar o la transformación transitoria, y la segunda secuencia nucleotídica se expresa de manera que la proteína de interés producida comprenda glicanos con N-glicosilación modificada, o la planta que expresa la primera secuencia nucleotídica se cruza con una segunda planta que expresa la segunda secuencia nucleotídica, y de esta manera la proteína de interés producida comprende glicanos con N-glicosilación modificada. La proteína de interés puede extraerse de la planta, y si se desea, la proteína de interés puede aislarse y purificarse utilizando métodos estándares.

Por lo tanto, se desvela un método (Método J) para producir una planta que puede utilizarse como una plataforma para la producción de una proteína de interés con N-glicosilación modificada. Este método comprende proporcionar una planta que expresa una o más de una primera secuencia nucleotídica que codifica Gnt-III, GNT1-Gnt-III o tanto Gnt-III como GNT1-Gnt-III, y que expresa una o más secuencias nucleotídicas. Con el fin de producir la proteína de interés, una segunda secuencia nucleotídica que codifica la proteína de interés se introduce en la planta de plataforma utilizando técnicas estándares que implican la transformación estable o la transformación transitoria, y la segunda secuencia nucleotídica se expresa de manera que la proteína de interés producida comprenda glicanos con N-glicosilación modificada, o la planta que expresa la primera secuencia nucleotídica se cruza con una segunda planta que expresa la segunda secuencia nucleotídica y de esta manera la proteína de interés producida comprende glicanos con N-glicosilación modificada. La proteína de interés puede extraerse de la planta, y si se desea, la proteína de interés puede aislarse y purificarse utilizando métodos estándares.

Por lo tanto, se desvela un método (Método K) para producir una planta que puede utilizarse como una plataforma para la producción de una proteína de interés con N-glicosilación modificada. Este método comprende proporcionar una planta que expresa una o más de una primera secuencia nucleotídica que codifica GalT, GNT1-GalT, GalT y

GNT1-GalT, GnT-III, GNT1-GnT-III, GnT-III y GNT1-GNT-III o una combinación de las mismas, y que expresa una o más secuencias nucleotídicas. Con el fin de producir la proteína de interés, una segunda secuencia nucleotídica que codifica la proteína de interés se introduce en la planta de la plataforma utilizando técnicas estándares que implican la transformación estable, o la transformación transitoria, y la segunda secuencia nucleotídica se expresa de manera que la proteína de interés producida comprenda glicanos con N-glicosilación modificada, o la planta que expresa la primera secuencia nucleotídica se cruza con una segunda planta que expresa la segunda secuencia nucleotídica y de esta manera la proteína de interés producida comprende glicanos con N-glicosilación modificada. La proteína de interés puede extraerse de la planta, y si se desea, la proteína de interés puede aislarse y purificarse utilizando métodos estándares.

10

Las secuencias nucleotídicas que codifican GalT, GNT1-GalT, GnT-III, GNT1-GnT-III, la proteína de interés, o una combinación de las mismas, pueden ser un codón optimizado para aumentar el nivel de expresión dentro de la planta. Por optimización de codón se refiere a la selección de nucleótidos de ADN apropiados para la síntesis de bloques de construcción de oligonucleótidos, y su ensamblaje enzimático posterior, de un gen o fragmento estructural del mismo con el fin de adoptar el uso del codón dentro de las plantas.

15

Con el fin de optimizar la expresión de la secuencia extraña dentro de una planta, la secuencia nucleotídica, que puede ser una secuencia silvestre o sintética, puede utilizarse o alterarse según se requiera de manera que la proteína correspondiente, por ejemplo, GalT, GNT1-GalT, GnT-III, GNT1-GnT-III, la proteína de interés, o una combinación de las mismas, se produzca a un nivel más elevado que se producirá al codificarse por la secuencia nucleotídica no modificada. Por ejemplo, sin que se considere limitante, la secuencia puede ser una secuencia sintética, optimizada para uso del codón dentro de una planta, que comprende al menos aproximadamente un 80 % de identidad con la secuencia de tipo silvestre, como se determina utilizando técnicas de comparación de secuencia, por ejemplo, pero sin limitación, BLAST (disponible a través de GenBank; usando parámetros predeterminados). Se contempla también que los fragmentos o porciones de la secuencia que codifica la proteína de interés, o derivados de la misma, que muestran propiedades biológicas útiles, por ejemplo, pero sin limitación, propiedades antigénicas, pueden expresarse dentro de los tejidos vegetales.

20

Con el fin de maximizar los niveles de expresión y la producción de proteínas transgénicas de GalT, GNT1-GalT, GnT-III, GNT1-GnT-III, y una proteína de interés, la secuencia de ácido nucleico puede ser examinada y la región de codificación modificada para optimizar la expresión del gen en plantas, utilizando un procedimiento similar al descrito por Sardana y col. (Plant Cell Reports 15: 677-681; 1996). Una tabla de uso del codón a partir de genes altamente expresados de plantas dicotiledóneas está disponible a partir de varias fuentes, incluyendo Murray y col. (Nuc Acids Res. 17: 477-498; 1989). Además, la optimización de secuencia puede también incluir la reducción de repeticiones de codón en tándem, la eliminación de sitios de empalme críptico, la reducción de secuencias repetidas (incluyendo repeticiones invertidas) y puede determinarse utilizando, por ejemplo, Leto 1.0 (Entelechon, Alemania).

30

Por lo tanto, se desvela un método (L) para sintetizar una proteína de interés con N-glicosilación modificada, como se describe en uno cualquiera de los métodos anteriores (Método A-K), en el que una o más de una de las secuencias nucleotídicas que codifican GalT, GNT1-GalT, GnT-III, GNT1-GnT-III, la proteína de interés o una combinación de las mismas, se optimizan para la expresión en una planta.

40

Además, la presente invención pertenece a una planta, una célula vegetal, o una semilla que comprende una secuencia nucleotídica que codifica GNT1-GalT unida operativamente con una región reguladora que está activa en la planta. La planta, célula vegetal o semilla puede además comprender una segunda secuencia nucleotídica que codifica una o más de una de una proteína de interés, la segunda secuencia nucleotídica unida operativamente a una o más de una segunda región reguladora activa dentro de la planta. La primera secuencia nucleotídica, la segunda secuencia nucleotídica, o tanto la primera secuencia nucleotídica como la segunda secuencia nucleotídica, pueden ser un codón optimizado para la expresión dentro de la planta, célula vegetal o semilla de planta.

50

Por "unida operativamente" significa que las secuencias particulares interactúan directa o indirectamente para realizar una función pretendida, tal como la mediación o la modulación de la expresión genética. La interacción de secuencias unidas operativamente puede, por ejemplo, mediarse por proteínas que interactúan con las secuencias unidas operativamente. Una región reguladora de transcripción y una secuencia de interés se unen operablemente cuando las secuencias se conectan funcionalmente para permitir la transcripción de la secuencia de interés que se media o modula por la región reguladora de transcripción.

55

Por la expresión "porción de una planta", se refiere a cualquier parte derivada de una planta, incluyendo la planta entera, el tejido obtenido de la planta, por ejemplo, pero sin limitación, las hojas, las hojas y el tallo, las raíces, la

porción aérea, incluyendo las hojas, el tallo y opcionalmente la porción floral de la planta, las células o protoplastos obtenidos de la planta.

- Por la expresión "materia vegetal", se refiere a cualquier material derivado de una planta. La materia vegetal comprende una planta entera, tejido, células o cualquier fracción de los mismos. Además, la materia vegetal puede comprender componentes vegetales intracelulares, componentes vegetales extracelulares, extractos líquidos o sólidos de plantas, o una combinación de los mismos. Además, la materia vegetal puede comprender plantas, células vegetales, tejido, un extracto líquido, o una combinación de los mismos, de hojas de las plantas, tallos, frutos, raíces o una combinación de los mismos. La materia vegetal puede comprender una planta o porción de la misma la cual no se ha sometido a ninguna etapa de procesamiento. Sin embargo, se contempla también que la materia vegetal puede someterse a etapas de procesamiento mínimas como se define a continuación, o procesamiento más riguroso, incluyendo purificación de proteínas parcial o sustancial utilizando técnicas comúnmente conocidas en la técnica, incluyendo, pero sin limitación, cromatografía, electroforesis y similares.
- Por la expresión "procesamiento mínimo", se refiere a materia vegetal, por ejemplo, una planta o porción de la misma que comprende una proteína de interés que se purifica parcialmente para producir un extracto vegetal, homogenado, fracción del homogenado vegetal o similares. La purificación parcial puede comprender, pero sin limitación, la interrupción de estructuras celulares vegetales, por lo que se crea una composición que comprende componentes vegetales solubles y componentes vegetales insolubles los cuales pueden separarse, por ejemplo, pero sin limitación, por centrifugación, filtración o una combinación de las mismas. En este sentido, las proteínas secretadas dentro del espacio extracelular de la hoja u otros tejidos podrían obtenerse fácilmente utilizando extracción al vacío o por centrifuga, o los tejidos podrían extraerse a presión por el paso a través de rodillos o molienda o similares para comprimir o liberar la proteína libre del espacio extracelular. El procesamiento mínimo podría también implicar la preparación de extractos en bruto de proteínas solubles, ya que estas preparaciones podrían tener una contaminación insignificante de productos vegetales secundarios. Además, el procesamiento mínimo puede implicar la extracción acuosa de proteína soluble a partir de hojas, seguida de la precipitación con cualquier sal adecuada. Otros métodos pueden incluir maceración y extracción de jugos a gran escala con el fin de permitir el uso directo del extracto.
- La materia vegetal, en la forma de material o tejido vegetal puede administrarse por vía oral a un sujeto. La materia vegetal puede administrarse como parte de un complemento dietético, junto con otros alimentos, o encapsularse. La materia o tejido vegetal puede también concentrarse para mejorar o aumentar la palatabilidad, o proporcionarse junto con otros materiales, ingredientes o excipientes farmacéuticos, según se requiera.
- Se contempla que una planta que comprende la proteína de interés puede administrarse a un sujeto, por ejemplo, un animal o ser humano, en una variedad de formas dependiendo de la necesidad y la situación. Por ejemplo, la proteína de interés obtenida a partir de la planta puede extraerse antes de su uso en ya sea una forma en bruto, parcialmente purificada o purificada. Si la proteína va a purificarse, entonces puede producirse en plantas comestibles o no comestibles. Además, si la proteína se administra por vía oral, el tejido vegetal puede cosecharse y suministrarse directamente al sujeto, o el tejido cosechado puede secarse antes de la alimentación, o puede dejarse que un animal se alimente de la planta sin tener lugar la cosecha. También se desvela el uso de los tejidos vegetales cosechados como un complemento alimenticio dentro del alimento para animales. Si el tejido vegetal se suministra a un animal con poco o ningún procesamiento adicional, se prefiere que el tejido vegetal que se administra sea comestible.
- Como se describe en más detalle en los Ejemplos, GalT, GNT1-GalT y la proteína de interés se introdujeron en plantas en una manera transitoria. El análisis inmunológico, utilizando anticuerpos apropiados, demostró que una proteína de PMr 150 kDa se presentó en las células transformadas (figuras 2, 3A y 3B). Además, GalT o GNT1-GalT fue detectable en extractos obtenidos a partir de plantas que expresan cualquier construcción, y la N-glicosilación alterada de una proteína de interés se observó cuando se expresó GNT1-GalT en la planta (figura 6). Por lo tanto, GalT expresado de forma recombinante, o GNT1-GalT, está biológicamente activa *in planta*.
- Se introdujeron en plantas GalT-FLAG o GNT1-GalT-FLAG (es decir, GalT o GNT1-GalT etiquetado en su extremo C con un epítopo FLAG para permitir la inmunodetección de la proteína recombinante e transformantes). Puede utilizarse un análisis de transferencia de Western utilizando anticuerpos anti-FLAG, para demostrar que la proteína apropiada se encuentra presente en las células transformadas.

Un "análogo" o "derivado" incluye cualquier sustitución, supresión o adición a la secuencia nucleotídica que codifica GalT (SEQ ID NO: 14), el dominio catalítico de GalT (nucleótidos 368-1198 de SEQ ID NO: 14; la secuencia de

- codificación en negrita en la figura 5b), o GNT1-GaIT (nucleótidos 248-1077 de SEQ ID NO: 17; secuencia de codificación en negrita que aparece en la figura 5d), siempre y cuando la secuencia codifique una proteína que modifique, o que al fusionarse a un CST de GaIT modifique el perfil de glicosilación de una proteína de interés cuando se expresa en una planta, por ejemplo, con reducción de la fucosilación, xilosilación, o ambas, de glicanos de la proteína de interés, o con el aumento de galactosilación de la proteína de interés en comparación con el perfil de glicosilación de la proteína de interés producida en una planta, en ausencia de GaIT (SEQ ID NO: 14) o GNT1-GaIT (SEQ ID NO: 17) expresadas ectópicamente. Por ejemplo, la proteína codificada por la secuencia puede añadir una galactosa terminal durante la maduración de N-glicanos. Típicamente, los derivados y análogos de las secuencias de ácido nucleico muestran una identidad mayor del 80 % con una secuencia de ácido nucleico, por ejemplo, de aproximadamente el 80-100 % de identidad de secuencia, o de aproximadamente el 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100 % de identidad. La identidad de secuencia puede determinarse mediante el uso del algoritmo BLAST (GenBank: [ncbi.nlm.nih.gov/cgi-bin/BLAST/](http://ncbi.nlm.nih.gov/cgi-bin/BLAST/)), utilizando parámetros predeterminados (Programa: blastn; Base de Datos: nr; Esperado 10: filtro: baja complejidad; Alineamiento: por pares; Longitud de las Palabras: 11). Los análogos o derivados de los mismos, también incluyen las secuencias nucleotídicas que hibridan en condiciones rigurosas de hibridación (véase Maniatis y col., en *Molecular Cloning (A Laboratory Manual)*, Cold Spring Harbor Laboratory, 1982, pág. 387-389, o Ausubel, y col. (ed.), 1989, *Current Protocols in Molecular Biology*, Vol. 1, Green Publishing Associates, Inc., y John Wiley & Sons, Inc., Nueva York, en pág. 2.10.3) con cualquiera de las secuencias de GaIT (SEQ ID NO: 14), GNATI-GaIT (SEQ ID NO: 18) descritas en el presente documento, siempre y cuando la secuencia codifique una proteína que modifique el perfil de glicosilación de una proteína de interés, por ejemplo, reduciendo la fucosilación, xilosilación, o ambas, de glicanos de la proteína de interés, o aumentando la galactosilación de la proteína de interés en comparación con el perfil de glicosilación de la proteína de interés producido en ausencia de GaIT (SEQ ID NO: 14) o GNT1-GaIT (SEQ ID NO: 17). Por ejemplo, la proteína codificada por la secuencia puede añadir una galactosa terminal durante la maduración de N-glicanos. Un ejemplo de una de tales condiciones rigurosas de hibridación puede ser la hibridación con una sonda adecuada, por ejemplo, pero sin limitación, una sonda etiquetada con [gamma-<sup>32</sup>P] dATP durante 16-20 h a 65 °C en SDS al 7 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 0,5 M, pH 7,2. Seguido de un lavado a 65 °C en SDS al 5 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 40 mM, pH 7,2 durante 30 min seguido de un lavado a 65 °C en SDS al 1 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 40 mM, pH 7,2 durante 30 min. El lavado en este tampón puede repetirse para reducir el fondo.
- De manera similar, se desvela una modificación en la secuencia nucleotídica que codifica GnT-in (SEQ ID NO: 16), el dominio catalítico de GnT-III, o GNT1-GnT-III (SEQ ID NO: 20), siempre y cuando la secuencia codifique una proteína que modifique, o al fusionarse con un CST de GnT-III modifique el perfil de glicosilación de una proteína de interés cuando se expresa en una planta, por ejemplo, con la bisección de GlNAc y la reducción de la fucosilación, xilosilación, o ambas, de glicanos de la proteína de interés, en comparación con el perfil de glicosilación de la proteína de interés producido en una planta, en ausencia de GnT-III (SEQ ID NO: 16) o GNT1-GnT-III (SEQ ID NO: 26) expresada ectópicamente. Típicamente, los derivados y análogos de las secuencias de ácido nucleico muestran una identidad mayor del 80 % con una secuencia de ácido nucleico, por ejemplo, aproximadamente del 80-100 % de identidad de la secuencia, o aproximadamente del 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100 % de identidad. La identidad de la secuencia puede determinarse mediante el uso del algoritmo BLAST (GenBank: [ncbi.nlm.nih.gov/cgi-bin/BLAST/](http://ncbi.nlm.nih.gov/cgi-bin/BLAST/)), utilizando parámetros predeterminados (Programa: blastn; Base de Datos: nr; Esperado 10: filtro: baja complejidad; Alineamiento: por pares; Longitud de las Palabras: 11). Los análogos o derivados de los mismos, también incluyen las secuencias nucleotídicas que hibridan en condiciones rigurosas de hibridación (véase Maniatis y col., en *Molecular Cloning (A Laboratory Manual)*, Cold Spring Harbor Laboratory, 1982, pág. 387-389, o Ausubel, y col. (ed.), 1989, *Current Protocols in Molecular Biology*, Vol. 1, Green Publishing Associates, Inc., y John Wiley & Sons, Inc., Nueva York, en pág. 2.10.3) con cualquiera de las secuencias de GnT-III (SEQ ID NO: 16), GNATI-GnT-III (SEQ ID NO: 26) descritas en el presente documento, siempre y cuando la secuencia codifique una proteína que modifique el perfil de glicosilación de una proteína de interés en comparación con el perfil de glicosilación de la proteína de interés producido en ausencia de GnT-III (SEQ ID NO: 16) o GNT1-GnT-III (SEQ ID NO: 26). Un ejemplo de una de tales condiciones rigurosas de hibridación puede ser la hibridación con una sonda adecuada, por ejemplo, pero sin limitación, una sonda etiquetada con [gamma-<sup>32</sup>P]dATP durante 16-20 h a 65 °C en SDS al 7 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 0,5 M, pH 7,2. Seguido de un lavado a 65 °C en SDS al 5 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 40 mM, pH 7,2 durante 30 min seguido de un lavado a 65 °C en SDS al 1 %, EDTA 1 mM, Na<sub>2</sub>HPO<sub>4</sub> 40 mM, pH 7,2 durante 30 min. El lavado en este tampón puede repetirse para reducir el fondo.
- Por "región reguladora", "elemento regulador" o "promotor" se entiende una porción de ácido nucleico típicamente, aunque no siempre, aguas arriba de la región que codifica las proteínas de un gen, la cual puede estar compuesta por ADN o ARN o tanto ADN como ARN. Cuando una región reguladora está activa, y en asociación operativa, o unida operativamente, con un gen de interés, esto puede dar como resultado la expresión de un gen de interés. Un elemento regulador puede ser capaz de mediar la especificidad orgánica, o controlar el desarrollo o la activación



genética temporal. Una "región reguladora" incluye elementos promotores, elementos promotores centrales que exhiben una actividad promotora basal, elementos que son inducibles en respuesta a un estímulo externo, elementos que median la actividad promotora tal como elementos reguladores negativos o mejoradores de la transcripción. La "región reguladora", como se utiliza en el presente documento, también incluye elementos que son

- 5 activos después de la transcripción, por ejemplo, elementos reguladores que modulan la expresión genética tales como mejoradores de la traducción y de la transcripción, represores de la traducción y de la transcripción, secuencias de activación aguas arriba, y determinantes de la inestabilidad de la ARNm. Varios de estos últimos elementos pueden ubicarse próximos a la región de codificación.
- 10 En el contexto de esta divulgación, la expresión "elemento regulador" o "región reguladora" típicamente se refiere a una secuencia de ADN, normalmente, aunque no siempre, aguas arriba (5') a la secuencia de codificación de un gen estructural, el cual controla la expresión de la región de codificación proporcionando el reconocimiento para la ARN polimerasa y/u otros factores requeridos para que la transcripción inicie en un sitio particular. Sin embargo, debe entenderse que otras secuencias de nucleótidos, ubicadas dentro de los intrones, o 3' de la secuencia pueden
- 15 también contribuir a la regulación de expresión de una región de codificación de interés. Un ejemplo de un elemento regulador que proporciona el reconocimiento para la ARN polimerasa u otros factores de transcripción para asegurar el inicio en un sitio particular es un elemento promotor. La mayoría, aunque no todos los elementos promotores eucariotas contienen una caja TATA, una secuencia de ácido nucleico conservada que consta de pares de bases de nucleótidos de adenosina y timidina usualmente situados aproximadamente 25 pares de bases aguas arriba de un sitio de inicio de la transcripción. Un elemento promotor comprende un elemento promotor basal, responsable por el
- 20 inicio de la transcripción, así como otros elementos reguladores (como se ha enumerado anteriormente) que modifican la expresión génica.

- Existen varios tipos de regiones reguladoras, incluyendo aquellas que se regulan por desarrollo, inducibles o
- 25 constitutivas. Una región reguladora que se regula por desarrollo, o controla la expresión diferencial de un gen bajo su control, se activa dentro de ciertos órganos o tejidos de un órgano en momentos específicos durante el desarrollo de ese órgano o tejido. Sin embargo, algunas regiones reguladoras que se regulan por desarrollo pueden activarse preferiblemente dentro de ciertos órganos o tejidos en etapas específicas del desarrollo, también pueden ser activas en una forma regulada por desarrollo, o a un nivel basal en otros órganos o tejidos dentro de la planta también.
- 30 Ejemplos de regiones reguladoras específicas de tejidos, por ejemplo, véase - específicas de una región reguladora, incluyen el promotor de napina, y al promotor de cruciferina (Rask y col., 1998, J. Plant Physiol. 152: 595-599; Bilodeau y col., 1994, Plant Cell 14: 125-130). Un ejemplo de un promotor específico de la hoja incluye el promotor de plastocianina (documento US 7.125.978; SEQ ID NO: 23; figura 1b).

- 35 Una región reguladora inducible es aquella que es capaz de activar directa o indirectamente la transcripción de una o más secuencias o genes de ADN en respuesta a un inductor. En ausencia de un inductor, las secuencias o genes de ADN no serán transcritos. Típicamente, el factor de proteína que se une específicamente a una región reguladora inducible para activar la transcripción puede presentarse en una forma activa, que es después directa o indirectamente convertida en la forma activa por el inductor. Sin embargo, el factor de proteína también puede estar
- 40 ausente. El inductor puede ser un agente químico tal como una proteína, metabolito, regulador de crecimiento, herbicida o compuesto fenólico o un estrés fisiológico impuesto directamente por calor, frío, sal o elementos tóxicos o indirectamente mediante la acción de un patógeno o agente infeccioso tal como un virus. Una célula vegetal que contiene una región reguladora inducible puede ser expuesta a un inductor aplicando externamente el inductor a la célula o planta por ejemplo, por aspersión, riego, calentamiento o métodos similares. Los elementos reguladores
- 45 inducibles pueden derivarse ya sea de genes de la planta o que no son de la planta (por ejemplo, Gatz, C. y Lenk, I.R.P., 1998, Trends Plant Sci. 3, 352-358). Los ejemplos de promotores inducibles potenciales incluyen, pero sin limitación, promotor inducible de tetraciclina (Gatz, C., 1997, Ann. Rev. Plant Physiol. Plant. Mol. Biol. 48, 89-108), el promotor inducible por esteroides (Aoyama, T. y Chua, N. H., 1997, Plant J. 2, 397-404) y el promotor inducible por etanol (Salter, M. G., y col., 1998, Plant Journal 16, 127 - 132; Caddick, M. X. y col., 1998, Nature Biotech, 16, 177
- 50 180), genes IB6 y CK11 inducibles por citoquinina (Brandstatter, I. y Kieber, J. J., 1998, Plant Cell 10, 1009 - 1019; Kakimoto, T., 1996, Science 274, 982 - 985) y el elemento inducible por auxina, DR5 (Ulmasov, T., y col., 1997, Plant Cell 9, 1963 - 1971). Una región reguladora constitutiva dirige la expresión de un gen a través de las diferentes partes de una planta y continuamente durante todo el desarrollo de la planta. Ejemplos de elementos reguladores constitutivos conocidos incluyen promotores asociados con la transcripción 35S de CaMV (Odell y col., 1985, Nature, 313: 810 - 812), la actina 1 del arroz (Zhang y col., 1991, Plant Cell, 3: 1155 - 1165), actina 2 (An y col., 1996, Plant J., 10: 107 - 121) o tms 2 (documento US 5.428.147) y los genes de trifosfato isomerasa 1 (Xu y col., 1994, Plant Physiol. 106: 459 - 467), el gen de la ubiquitina 1 de maíz (Cornejo y col., 1993, Plant Mol. Biol. 29: 637 - 646), los genes 1 y 6 de ubiquitina de *Arabidopsis* (Holtorf y col., 1995, Plant Mol. Biol. 29: 637 - 646), y el gen del factor 4A de inicio de la traducción de tabaco (Mandel y col., 1995 Plant Mol. Biol. 29: 995 - 1004). El término "constitutivo"
- 55

como se usa en el presente documento, no necesariamente indica que un gen bajo control de la región reguladora constitutiva se exprese al mismo nivel en todos los tipos de células, sino que el gen se expresa en un amplio rango de tipos celulares aun cuando a menudo se observa una variación en abundancia.

5 La una o más de las secuencias de nucleótidos de la presente invención puede expresarse en cualquier huésped vegetal adecuado que se transforma por la secuencia de nucleótidos, o construcciones, o vectores de la presente invención. Los ejemplos de huéspedes adecuados incluyen, pero sin limitación, cultivos agrícolas incluyendo alfalfa, colza, *Brassica spp.*, maíz, *Nicotiana spp.*, alfalfa, patata, ginseng, guisante, avena, arroz, soja, trigo, cebada, girasol y algodón.

10

La una o más construcciones genéticas quiméricas de la presente invención pueden comprender además una región 3' no traducida. Una región 3' no traducida se refiere a aquella porción de un gen que comprende un segmento de ADN que contiene una señal de poliadenilación y cualquier otra señal reguladora capaz de efectuar el procesamiento del ARNm o expresión genética. La señal de poliadenilación se caracteriza usualmente efectuando la adición de rastros de ácido poliadenílico al extremo 3' del precursor de ARNm. Las señales de poliadenilación se reconocen comúnmente por la presencia de homología con la forma canónica 5' AATAAA-3', aunque no son infrecuentes las variaciones. Una o más de las construcciones genéticas quiméricas de la presente invención pueden también incluir potenciadores adicionales, ya sea potenciadores de traducción o de transcripción, según se requiera. Estas regiones potenciadoras son bien conocidas por los expertos en la técnica, y pueden incluir el codón de inicio ATG y las secuencias adyacentes. El codón de inicio debe estar en fase con el marco de lectura de la secuencia de codificación para asegurar la traducción de la secuencia completa.

Los ejemplos no limitantes de regiones 3' adecuadas son las regiones 3' no traducidas, transcritas que contienen una señal de poliadenilación de genes del plásmido que inducen tumor (Ti) de *Agrobacterium*, tal como los genes de plantas y de nopalina sintasa (gen Nos) tal como los genes de proteína de almacenamiento de soja y la subunidad pequeña del gen de ribulosa-1,5-bisfosfato carboxilasa (ssRUBISCO).

Para facilitar la identificación de las células de plantas transformadas, las construcciones de esta invención pueden manipularse además para incluir marcadores seleccionables de plantas. Marcadores seleccionables útiles incluyen enzimas que proporcionan resistencia a compuestos químicos tal como un antibiótico, por ejemplo, gentamicina, higromicina, kanamicina, o herbicidas, tal como fosfinotricina, glifosato, clorosulfurón y similares. De forma similar, se pueden utilizar enzimas que proporcionen la producción de un compuesto identificable por medio de cambio de color tal como GUS (beta-glucuronidasa), o luminiscencia, tal como luciferasa o GFP.

35 También se consideran parte de esta invención las plantas transgénicas, células vegetales o semillas que contienen la construcción genética quimérica de la presente invención. Los métodos para regenerar plantas completas a partir de células vegetales se conocen también en la técnica. En general, se cultivan células vegetales transformadas en un medio apropiado, que puede contener agentes selectivos tal como antibióticos, donde se usan marcadores seleccionables para facilitar la identificación de células vegetales transformadas. Una vez se forman los callos, se puede estimular la formación de brotes empleando las hormonas apropiadas de plantas de acuerdo con métodos conocidos y se transfieren los brotes al medio de enraizamiento para regeneración de las plantas. Las plantas pueden utilizarse entonces para establecer generaciones repetitivas, ya sea a partir de semillas o utilizando técnicas de propagación vegetativas. También pueden generarse plantas transgénicas sin utilizar cultivos de tejidos.

45 Los elementos reguladores descritos en la presente invención pueden también combinarse con la región de codificación de interés para la expresión de uno de un rango de organismos huéspedes que son sensibles a transformación, o expresión transitoria. Tales organismos incluyen, pero sin limitación, plantas, tanto monocotiledóneas como dicotiledóneas, por ejemplo, pero sin limitación, maíz, plantas de cereales, trigo, cebada, avena, *Nicotiana spp.*, *Brassica spp.*, soja, judías, guisantes, alfalfa, patata, tomate, ginseng y *Arabidopsis*.

50

Los métodos para la transformación estable y la regeneración de estos organismos se encuentran establecidos en la técnica y se conocen entre los expertos en la técnica. El método para obtener plantas transformadas y regeneradas no es crucial para la presente invención.

55 El término "transformación" se refiere a la transferencia interespecífica de información genética (secuencia nucleotídica) que se manifiesta de forma genotípica, fenotípica, o ambas. La transferencia interespecífica de la información genética proveniente de una construcción quimérica a un huésped puede heredarse y la transferencia de información genética puede considerarse estable, o la transferencia puede ser transitoria y la transferencia de información genética no puede heredarse.

Se desvela un vector adecuado que comprende una construcción quimérica adecuada para utilizarse con sistemas de expresión estables o transitorios. La información genética también puede proporcionarse dentro de una o más construcciones. Por ejemplo, una secuencia nucleotídica que codifica una proteína de interés puede introducirse en una construcción, y una segunda secuencia nucleotídica que codifica una proteína que modifica la glicosilación de la proteína de interés puede introducirse utilizando una construcción separada. Estas secuencias nucleotídicas pueden co-expresarse posteriormente dentro de una planta. Sin embargo, también puede utilizarse una estructura que comprende una secuencia nucleotídica que codifica tanto la proteína de interés como la proteína que modifica el perfil de glicosilación de la proteína de interés. En este caso, la secuencia nucleotídica comprenderá una primera secuencia que comprende una primera secuencia de ácido nucleico que codifica la proteína de interés operativamente unida a un promotor o región reguladora, y una segunda secuencia que comprende una segunda secuencia de ácido nucleico que codifica la proteína que modifica el perfil de glicosilación de la proteína de interés, la segunda secuencia operativamente unida a un promotor o región reguladora.

Por "coexpresado" se entiende que dos o más de dos secuencias de nucleótidos se expresan aproximadamente al mismo tiempo dentro de la planta, y dentro del mismo tejido de la planta. Sin embargo, las secuencias de nucleótidos no necesitan expresarse exactamente al mismo tiempo. Más bien, las dos o más secuencias de nucleótidos se expresan en una forma tal que los productos codificados tengan una oportunidad de interactuar. Por ejemplo, la proteína que modifica la glicosilación de la proteína de interés puede expresarse ya sea antes o durante el período cuando la proteína de interés se expresa de tal manera que la modificación de la glicosilación de la proteína de interés tenga lugar. Las dos o más de dos secuencias de nucleótidos pueden coexpresarse utilizando un sistema de expresión transitorio, donde las dos o más secuencias se introducen dentro de la planta aproximadamente al mismo tiempo bajo condiciones tales que se expresen ambas secuencias. Como alternativa, una planta de plataforma que comprende una de las secuencias de nucleótidos, por ejemplo, la secuencia que codifica la proteína que modifica el perfil de glicosilación de la proteína de interés, puede transformarse, ya sea transitoriamente o en una forma estable, con una secuencia adicional que codifica la proteína de interés. En este caso, la secuencia que codifica la proteína que modifica el perfil de glicosilación de la proteína de interés puede expresarse dentro de un tejido deseado, durante una etapa deseada de desarrollo, o su expresión puede inducirse utilizando un promotor inducible, y la secuencia adicional que codifica la proteína de interés puede expresarse bajo condiciones similares y en el mismo tejido, para asegurar que se coexpresan las secuencias de nucleótidos.

Las construcciones pueden introducirse en las células vegetales utilizando plásmidos Ti, plásmidos Ri, vectores virales de la planta, transformación directa de ADN, microinyección, electroporación, etc. Para la revisión de tales técnicas véase por ejemplo, Weissbach y Weissbach, *Methods for Plant Molecular Biology*, Academy Press, Nueva York VIII, págs. 421 - 463 (1988); Geier y Corey, *Plant Molecular Biology*, 2ª edición, (1988); y Miki y Iyer, *Fundamentals of Gene Transfer in Plants*. En *Plant Metabolism*, 2ª edición, DT. Dennis, DH Turpin, DD Lefebvre, DB Layzell (eds), Addison Wesley, Longmans Ltd, Londres, págs. 561 - 579 (1997). Otros métodos incluyen la incorporación directa de ADN, el uso de liposomas, electroporación, por ejemplo, utilizando protoplastos, microinyección, microproyectiles o bigotes e infiltración al vacío. Véase por ejemplo, Bilang, y col. (*Gene* 100: 247 - 250 (1991), Scheid y col. (*Mol. Gen. Genet.* 228: 104 - 112, 1991), Guerche y col. (*Plant Science* 52: 111 - 116, 1987), Neuhauser y col. (*Theor. Appl. Genet.* 75: 30 - 36, 1987), Klein y col., *Nature* 327: 70 - 73 (1987); Howell y col. (*Science* 208: 1265, 1980), Horsch y col. (*Science* 227: 1229 - 1231, 1985), DeBlock y col., *Plant Physiology* 91: 694 - 701, 1989), *Methods for Plant Molecular Biology* (Weissbach y Weissbach, eds. Academic Press Inc., 1988), *Methods in Plant Molecular Biology* (Schuler y Zielinski eds., Academic Press Inc., 1989), Liu y Lomonosoff (*J. Virol. Meth.* 105: 343 - 340, 2002), las Pat. de Estados Unidos n.º 4.945.050; 5.036.006; y 5.100.792, las Solicitudes de Patente de Estados Unidos n.º de Serie 08/438.666, presentada el 10 de mayo de 1995, y 07/951.715, presentada el 25 de septiembre de 1992.

Como se describe a continuación, se usan métodos de expresión transitoria para expresar las construcciones de la presente invención (véase Liu y Lomonosoff, 2002, *Journal of Virological Methods*, 105: 343 - 348). Como alternativa, puede utilizarse un método de expresión transitoria con base en vacío, como se describe por Kapila y col. 1997. Estos métodos pueden incluir, por ejemplo, pero sin limitación, un método de Agroinoculación o Agroinfiltración, sin embargo, también se pueden utilizar otros métodos transitorios como se observó anteriormente. Ya sea con la Agroinoculación o con la Agroinfiltración, una mezcla de Agrobacterias que incluye el ácido nucleico deseado ingresa a los espacios intercelulares de un tejido, por ejemplo, las hojas, la porción aérea de la planta (incluyendo tallo, hojas y flores), otra porción de la planta (tallo, raíz, flor) o la planta completa. Después de cruzar la epidermis, las *Agrobacteria* infectan y transfieren copias de ADN-t dentro de las células. El ADN-t se transcribe en forma episomal y se traduce el ARNm, conduciendo a la producción de la proteína de interés en células infectadas, sin embargo, el paso del ADN-t dentro del núcleo es transitorio.

Por "gen de interés", "secuencia nucleotídica de interés", o "región codificante de interés", se refiere a que cualquier gen, secuencia nucleotídica, o región de codificación que se va a expresar dentro de un organismo huésped, por ejemplo, una planta. Estas expresiones se utilizan de forma intercambiable. Tal secuencia nucleotídica de interés puede incluir, pero sin limitación, un gen o región de codificación cuyo producto es una proteína de interés. Ejemplos de una proteína de interés incluyen, por ejemplo, pero sin limitación, una enzima industrial, un complemento proteínico, un nutracéutico, un producto de valor añadido, o un fragmento de los mismos para su uso en forraje, alimento, o en ambos, una proteína farmacéuticamente activa, por ejemplo, pero sin limitación, factores de crecimiento, reguladores de crecimiento, anticuerpos, antígenos, y fragmentos de los mismos, o sus derivados útiles para inmunización o vacunación o similares. Proteínas de interés adicionales pueden incluir, pero sin limitación, interleucinas, por ejemplo, una o más de una de IL-1 a IL-24, IL-26 y IL-27, citocinas, eritropoyetina (EPO), insulina, G-CSF, GM-CSF, hPG-CSF, M-CSF o combinaciones de las mismas, interferones, por ejemplo, interferón-alfa, interferón-beta, interferón-gama, factores de coagulación sanguínea, por ejemplo, Factor VIII, Factor IX, o tPA hGH, receptores, agonistas de receptores, anticuerpos, neuropolipéptidos, insulina, vacunas, factores de crecimiento, por ejemplo, pero sin limitación, factor del crecimiento epidérmico, factor del crecimiento queratinocítico, factor de crecimiento de transformación, reguladores de crecimiento, antígenos, autoantígenos, fragmentos de los mismos, o combinaciones de los mismos.

Si la secuencia nucleotídica de interés codifica un producto que es directa o indirectamente tóxico para la planta, entonces, usando el método de la presente invención, tal toxicidad puede reducirse por toda la planta expresando de forma selectiva el gen de interés dentro de un tejido deseado o en una etapa deseada del desarrollo de la planta. Además, el periodo limitado de expresión resultante de la expresión transitoria puede reducir el efecto al producir un producto tóxico en la planta.

La región de codificación de interés o la secuencia nucleotídica de interés puede expresarse en cualquier planta huésped adecuada, y a sea transformada o que comprenda las secuencias nucleotídicas, o moléculas de ácidos nucleicos, o construcciones genéticas, o vectores de la presente invención. Los ejemplos de huéspedes adecuados incluyen, pero sin limitación, *Arabidopsis*, cultivos agrícolas que incluyen, por ejemplo, colza, *Brassica* spp., maíz, *Nicotiana* spp., alfalfa, patata, ginseng, guisante, avena, arroz, soja, trigo, cebada, girasol y algodón.

Como se describe en más detalle en los siguientes ejemplos, la síntesis de una proteína de interés, por ejemplo, pero sin limitación, un anticuerpo, C5-1, con una N-glicosilación modificada, se produjo en una planta co-expresando GalT o GNT1-GalT.

Lista de secuencias:

| Secuencia             | SEQ ID NO: | Secuencia                          | SEQ ID NO: |
|-----------------------|------------|------------------------------------|------------|
| XmaI-pPlas.c          | 1          | GNT1-GalT (secuencia nucleotídica) | 17         |
| SacI-ATG-pPlas.r      | 2          | aminoácido GNT1-GalT               | 18         |
| SacI-PlasTer.c        | 3          | aminoácido GnT-III                 | 19         |
| EcoRI-PlasTer.r       | 4          | aminoácido GNT1-GnT-III            | 20         |
| Plasto-443c           | 5          | Dominio CTS de GNT1 (nucleótido)   | 21         |
| Plas+LC-C51.r         | 6          | Dominio CTS de GNT1 (aminoácido)   | 22         |
| LC-C51.c              | 7          | Promotor de plastocianina y 5'UTR  | 23         |
| LC-C51XhoSac.r        | 8          | Plastocianina 3' UTR y terminador  | 24         |
| Plas+HC-C51.r         | 9          | FgalTSpe                           | 25         |
| HC-C51.c              | 10         | GNT1-GnT-III (nucleótido)          | 26         |
| HC-C51XhoSac.r        | 11         | FgalT                              | 27         |
| HC-C51KDEL(SacI).r    | 12         | RgalTFlagStu                       | 28         |
| Glucopéptido triptico | 13         | FGNT                               | 29         |
| GalT I (nucleótido)   | 14         | RGBTSpe                            | 30         |
| aminoácido GalT I     | 15         |                                    |            |
| GnT-III (nucleótido)  | 16         |                                    |            |

## Ejemplos

### Ejemplo 1: Ensamblaje de los casetes de expresión R610, R612 (figura 1a), R621 y R622 (figura 5a)

Todas las manipulaciones se realizaron usando protocolos de biología molecular general de Sambrook y Russel

(2001).

R610, R612 (figura 1a)

5 Los cebadores de oligonucleótidos usados se presentan a continuación:

**XmaI-pPlas.c:** SEQ ID NO: 1 5'-AGTTCCCGGGCTGGTATATTTATATGTTGTC-3' SEQ ID NO: 1  
**SacI-ATG-pPlas.r:** SEQ ID NO: 2 5'-AATAGAGCTCCATTTTCTCTCAAGATGATTAATTAATTAATTAGTC-3' SEQ ID NO: 2  
**SacI-PlasTer.c:** SEQ ID NO: 3 5'-AATAGAGCTCGTTAAATGCTTCTTCGTCTCCTATTTATAATATGG-3' SEQ ID NO: 3  
**EcoRI-PlasTer.r:** SEQ ID NO: 4 5'-TTACGAATTCCTTCCTAATTGGTGTACTATCATTTATCAAAGGGGA-3' SEQ ID NO: 4  
**Plasto-443c:** SEQ ID NO: 5 5'-GTATTAGTAATTAGAATTTGGTGTGTC-3' SEQ ID NO: 5  
**Plas+LC-C51.r:** SEQ ID NO: 6 5'-ATCTGAGGTGTGAAAACCATTTTCTCTCAAGATG-3' SEQ ID NO: 6  
**LC-C51.c:** SEQ ID NO: 7 5'-ATGGTTTTACACCTCAGATACTTGG-3' SEQ ID NO: 7  
**LC-C51XhoSac.r:** SEQ ID NO: 8 5'-ATATGAGCTCCTCGAGCTAACACTCATTCTGTTGAAGC-3' SEQ ID NO: 8  
**Plas+HC-C51.r:** SEQ ID NO: 9 5'-CAAGGTCCACACCCAAGCCATTTTCTCTCAAGATG-3' SEQ ID NO: 9  
**HC-C51.c:** SEQ ID NO: 10 5'-ATGGCTTGGGTGTGGACCTTGC-3' SEQ ID NO: 10  
**HC-C51XhoSac.r:** SEQ ID NO: 11 5'-ATAAGAGCTCCTCGAGTCATTACCAGGAGAGTGGG-3' SEQ ID NO: 11  
**HC-C51KDEL(SacI).r:** SEQ ID NO: 12 5'-ATAAGAGCTCTCAAAGTTCATCTTTTACCAGGAGAGTGGG-3' SEQ ID NO: 12

25 La primera etapa de clonación consistió en ensamblar un plásmido receptor que contenía elementos reguladores aguas arriba y aguas abajo del gen de plastocianina de alfalfa. El promotor de plastocianina (Patente de Estados Unidos 7.125.978) y las secuencias 5'UTR se amplificaron a partir de ADN genómico de alfalfa utilizando los cebadores de oligonucleótidos XmaI-pPlas.c (SEQ ID NO: 1) y SacI-ATG-pPlas.r (SEQ ID NO: 2). El producto resultante de la amplificación se digirió con XmaI y SacI y se ligó con pCAMBIA2300, previamente digerido con las mismas enzimas, para crear pCAMBIA-PromoPlasto. De manera similar, las secuencias de 3'UTR y el terminador, del gen de plastocianina (figura 1c; nucleótidos 1-399 de la SEQ ID NO: 24) se amplificaron a partir de ADN genómico de alfalfa utilizando los siguientes cebadores: SacI-PlasTer.c (SEQ ID NO: 3) y EcoRI-PlasTer.r (SEQ ID NO: 4) y el producto se digirió con SacI y EcoRI antes de insertarse en los mismos sitios de pCAMBIA-PromoPlasto para crear pCAMBIAPlasto.

Los plásmidos R 610 y R 612 se prepararon para contener secuencias de cadena ligera C5-1 y de cadena pesada C5-1 bajo el promotor de plastocianina de alfalfa como construcciones en tándem; R 610 se diseñó para permitir la retención en el ER de la IgG ensamblada y comprendió la secuencia de KDEL fusionada al extremo de la cadena pesada de C5-1, y R 612 se diseñó para permitir la secreción.

El ensamblaje de casetes de expresión de C5-1 se realizó utilizando un método de ligación mediado por PCR descrito por Darveau y col. (1995). Para ensamblar las secuencias de codificación de cadena ligera aguas abajo del promotor de plastocianina, una primera etapa consistió en amplificar los primeros 443 pares de bases (pb) del promotor de plastocianina de alfalfa (nucleótidos 556-999 de la figura 1b o SEQ ID NO: 23) descritos por D'Aoust y col. (Patente de Estados Unidos 7.125.978) aguas abajo del ATG inicial mediante PCR usando pCAMBIAPlasto como plantilla y los siguientes cebadores:

50 Plasto-443c (SEQ ID NO: 5) y Plas+LC-C51.r (SEQ ID NO: 6; el solapamiento está subrayado, anteriormente).

De forma paralela, la secuencia de codificación de cadena ligera se amplificó por PCR a partir del plásmido pGA643-kappa (Khoudi y col., 1999) con los siguientes cebadores:

55 LC-C51.C (SEQ ID NO: 7) y LC-C51XhoSac.r. (SEQ ID NO: 8; el solapamiento está subrayado).

Los dos productos de amplificación obtenidos se mezclaron y utilizaron como plantilla en una tercera reacción por PCR utilizando los cebadores Plasto-443c (SEQ ID NO: 5) y LC-C51XhoSac.r (SEQ ID NO: 8). El solapamiento entre los cebadores Plas+LC-C51.r (SEQ ID NO: 6) y LC-C51.C (SEQ ID NO: 7) utilizados en las primeras reacciones

condujo al ensamblaje de los productos de la amplificación durante la tercera reacción. El producto ensamblado resultante de la tercera reacción por PCR se digirió con *DraIII* y *SacI* y se ligó con pCAMBIAPlasto digerido con *DraIII* y *SacI* para generar el plásmido R540.

- 5 La secuencia codificante de cadena pesada se fusionó con el elemento regulador aguas arriba de plastocianina mediante la amplificación de los 443 pb aguas arriba del ATG inicial de plastocianina, los nucleótidos 556-999 de la figura 1b; SEQ ID NO: 23, mediante PCR utilizando pCAMBIAPlasto como plantilla con los siguientes cebadores:

Plasto-443c (SEQ ID NO: 5) y Plas+HC-C51.r (SEQ ID NO: 9; el solapamiento subrayado anterior).

10

El producto de estas reacciones se mezcló y ensambló en una tercera reacción por PCR utilizando los cebadores Plasto-443c (SEQ ID NO: 5) y HC-C51XhoSac.r (SEQ ID NO: 11). El fragmento resultante se digirió con *DraIII* y *SacI* y se ligó en pCAMBIAPlasto entre los sitios de *DraIII* y *SacI*. El plásmido resultante se denominó R541.

- 15 Se añadió una etiqueta KDEL en el extremo C de la secuencia codificante de cadena pesada mediante amplificación por PCR con los cebadores Plasto-443c (SEQ ID NO: 8) y HC-C51KDEL (*SacI*).r (SEQ ID NO: 12) utilizando el plásmido R541 como una plantilla. El fragmento resultante se digirió con *DraIII* y *SacI* se clonó en los mismos sitios de pCAMBIAPlasto, creando el plásmido R550.

- 20 El ensamblaje de los casetes de expresión de cadena ligera y pesada en el mismo plásmido binario se realizó como se indica a continuación: R541 y R550 se digirieron con *EcoRI*, se redujeron, se digirieron con *HindIII* y se ligaron en los sitios *HindIII* y *SmaI* de R540 para crear R610 (con KDEL) y R612 (sin KDEL; véase la figura 1).

R621 y R622 (figura 5a) - Los cebadores oligonucleotídicos usados se presentan a continuación

25

**FgaIT** SEQ ID NO: 27 5'-GACTCTAGAGCGGGAAGATGAGGCTTCGGGAGCCGCTC-3' SEQ ID NO: 27

**RgaITFlagStu** SEQ ID NO: 28

5'-AAGGCCTACGCTACTTGTCATCGTCATCTTTGTAGTCGCAC

GGTGTCCCGAAGTCCAC-3' SEQ ID NO: 28

30

**FGNT** SEQ ID NO: 29 5'-ATCGAAATCGCACGATGAGAGGGAACAAGTTTTGC-3' SEQ ID NO: 29

**RGNTSpe** SEQ ID NO: 30 5'-CGGGATCCACTAGTCTGACGCTTCATTTGTTCTTC-3' SEQ ID NO: 30

35

**FgaITSpe** SEQ ID NO: 25 5'-GGACTAGTGCCTGTCGCTGCCCGCCTGC-3' SEQ ID NO: 25

- Los plásmidos para la expresión de GalT y GNT1GalT se ensamblaron a partir de pBLTI121 (Pagny y col., 2003). El gen de  $\beta(1,4)$ -galactosiltransferasa (hGalT) (galactosa UDP:  $\beta$ -N-acetilglucosaminida:  $\beta(1,4)$ -galactosiltransferasa; EC 2.4.1.22) se aisló de pUC19-hGalT (Watzel y col., 1991) con digestión de *EcoRI*. Después del tratamiento de Klenow, el fragmento de hGalT de 1, 2 kb se clonó en los sitios *SmaI*, dando como resultado el plásmido pBLTI221hGalT. Después, una etiqueta indicadora se fusionó con el extremo C-terminal de la región codificante mediante PCR utilizando los cebadores FgaIT (SEQ ID NO: 27) y RgaITFlagStu (SEQ ID NO: 28) para amplificación. Posteriormente, R622 se produjo clonando este fragmento *XbaI*-*StuI* en el vector binario pBLTI121. Los primeros 77 a. a. de N-acetilglucosaminiltransferasa I (GNT1) que corresponden al dominio transmembrana se amplificaron por PCR utilizando el ADNc de *N. tabacum* que codifica N-GNT1 como plantilla (Strasser y col., 1999) y FGNT (SEQ ID NO: 29) y RGNTSpe (SEQ ID NO: 30) como cebadores. El producto de amplificación se clonó en primer lugar en el vector pGEM-T, y el plásmido resultante se asimiló con *Apal* y *BamHI*, y se ligó en pBLTI221, produciendo un plásmido denominado pBLTI221-GNTI. El dominio catalítico de hGalT se obtuvo mediante amplificación por PCR en pBLTI221hGalT usando los cebadores FgaITSpe (SEQ ID NO: 25) y RgaITFlagStu (SEQ ID NO: 28), creando los sitios *SpeI* y *StuI* en el extremo 5' y 3', respectivamente. Después, el fragmento *SpeI*/*StuI* hGalT se clonó en pBLTI221-GNTI utilizando los mismos sitios (*SpeI* y *StuI*), creando pBLTI221-GNTIGalT. Finalmente, pBLTI221-GNTIGalT se asimiló con *XbaI* y *StuI* aislando la secuencia codificante de GNT1GalT (figura 5d; SEQ ID NO: 17), y R621 se produjo clonando este fragmento en el vector binario pBLTI121.

- 55 Todos los clones se secuenciaron para confirmar la integridad de las construcciones. Los plásmidos se usaron para transformar *Agrobacterium tumefaciens* (AGL1; ATCC, Manassas, VA 20108, Estados Unidos) por electroporación (Höfgen y Willmitzer, 1988) usando un aparato Gene Pulser II (Biorad, Hercules, CA, Estados Unidos) como para la



transformación de *E. coli* (W.S. Dower, Electroporation of bacteria, En "Genetic Engineering", Volumen 12, Plenum Press, Nueva York, 1990, J.K. Setlow eds.). La integridad de todas las cepas de *A. tumefaciens* se confirmó mediante un mapeo de restricción.

5 Se preparó una construcción HcPro como se describe en Hamilton y col. (2002).

#### Ejemplo 2: Preparación de biomasa vegetal, inoculación, agroinfiltración y cosecha

Se cultivaron plantas de *Nicotiana benthamiana* a partir de semillas en superficies llenas con un sustrato de musgo esfagnáceo comercial. Se dejó que las plantas crecieran en el invernadero bajo un fotoperiodo de 16/8 y un régimen de temperatura de 25 °C durante el día/20 °C durante la noche. Después del sembrado, se seleccionaron plántulas individuales, se transplantaron en macetas y se dejó que crecieran en el invernadero durante tres semanas adicionales en las mismas condiciones ambientales.

15 Las cepas de *Agrobacterium* R612, R610, R621, R622 o 35SHcPro crecieron en un medio YEB complementado con ácido 2-[N-morfolino]etanosulfónico 10 mM (MES), acetosiringona 20 µM, 50 µg/ml de cloranfenicol y 25 µg/ml de carbenicilina a pH 5,6 hasta que alcanzaron una OD<sub>600</sub> entre 0,6 y 1,6. Las suspensiones de *Agrobacterium* se centrifugaron antes de utilizarse y se suspendieron de nuevo en un medio de infiltración (MgCl<sub>2</sub> 10 mM y MES 10 mM a pH 5,6).

20 La infiltración con jeringa se realizó como se describe por Liu y Lomonosoff (2002, Journal of Virological Methods, 105: 343-348)).

Para la infiltración al vacío, las suspensiones de *A. tumefaciens* se centrifugaron, se suspendieron de nuevo en el medio de infiltración y se almacenaron durante la noche a 4 °C. El día de la infiltración, los lotes de cultivo se diluyeron en 2,5 volúmenes de cultivo y se dejó que se calentaran antes de utilizarse. Se pusieron al revés plantas completas de *Nicotiana benthamiana* en la suspensión bacteriana en un depósito de acero inoxidable cerrado herméticamente, con un vacío de 20-40 Torr durante 2 min. Tras la infiltración con jeringa o al vacío, las plantas se devolvieron al invernadero durante un periodo de incubación de 4-5 días hasta su cultivo.

#### Muestreo de Hojas y Extracción Total de Proteínas

Después de la incubación, se cosechó la parte aérea de las plantas, se congeló a -80 °C, se trituró y se separó en sub-muestras de 1,5 o 7,5 g. Las proteínas solubles totales se extrajeron al homogeneizar (Polytron) cada sub-muestra de material congelado-triturado de planta en 3 volúmenes de Tris 50 mM frío a pH 7,4, NaCl 0,15 M, Triton X-100 al 0,1 %, fluoruro de fenilmetansulfonilo 1 mM y quitostatina 10 µM. Después de la homogenización, las suspensiones se centrifugaron a 20.000 g durante 20 min a 4 °C y los extractos en bruto depurados (sobrenadante) se conservaron para su análisis. El contenido de proteína total de los extractos en bruto depurados se determinó mediante el ensayo de Bradford (Bio-Rad, Hercules, CA) utilizando albúmina sérica bovina como estándar de referencia.

#### Ejemplo 3: Análisis de proteínas, inmunotransferencia y ELISA

C5-1 es una IgG murina anti-humana, por lo tanto, la detección y cuantificación puede realizarse a través de su afinidad característica con IgG (marcas de actividad) o mediante su inmunoreactividad con IgG anti-ratón.

Las proteínas de extractos totales en bruto o anticuerpos purificados se separaron mediante SDS-PAGE y se tiñeron con Azul de Coomassie R-250 o G-250 o se electrotransferieron a membranas de difluoruro de polivinileno (Roche Diagnostics Corporation, Indianapolis, IN) para inmunodetección. Antes de la inmunotransferencia, las membranas se bloquearon con leche desnatada al 5 % y Tween-20 al 0,1 % en una solución salina tamponada con Tris (TBS-T) durante 16-18 h a 4 °C.

La inmunotransferencia se realizó mediante incubación con los siguientes anticuerpos: un anticuerpo IgG (H+L) anti-ratón de cabra conjugado con peroxidasa (Jackson ImmunoResearch, West Grove, PA, Cat. n.º 115-035-146) (0,04 µg/ml en leche desnatada al 2 % en TBS-T), un anticuerpo IgG humano conjugado con peroxidasa (Gamunex® Bayer Corp., Elkhart, IN) (0,2 µg/ml en leche desnatada al 2 % en TBS-T) o un anticuerpo IgG anti-ratón de cabra policlonal (específico de cadena pesada) (Sigma-Aldrich, St-Louis, MO) (0,25 µg/ml en leche desnatada al 2 % en TBS-T). Un anticuerpo IgG anti-cabra de asno conjugado con peroxidasa (Jackson ImmunoResearch) (0,04 µg/ml en leche desnatada al 2 % en TBS-T) se utilizó como anticuerpo secundario para membranas tratadas

con el anticuerpo específico de cadena pesada. Se detectaron los complejos inmunorreactivos mediante quimioluminiscencia utilizando luminol como el sustrato (Roche Diagnostics Corporation). La conjugación de la enzima de peroxidasa de rábano picante del anticuerpo IgG humano se realizó usando el equipo de conjugación de peroxidasa activada EZ-Link Plus® (Pierce, Rockford, IL).

5

#### Ensayo cuantitativo de ELISA

- Placas con múltiples pocillos (Immulon 2H B, ThermoLab System, Franklin, MA) se cubrieron con 2,5 µg/ml de anticuerpos anti-ratón de cabra específicos para cadena pesada de IgG1 (Sigma M8770) en tampón de carbonato
- 10 50 mM (pH 9,0) a 4 °C durante 16-18 h. Las placas multipocillo se bloquearon después a través de 1 h de incubación en caseína al 1 % en una solución salina tamponada con fosfato (PBS) (Pierce Biotechnology, Rockford, IL) a 37 °C. Una curva estándar se generó con diluciones de un control de IgG1 de ratón purificado (Sigma M9269). Al realizar los inmunoensayos, todas las diluciones (control y las muestras) se realizaron en un extracto vegetal obtenido de un tejido vegetal infiltrado e incubado con inóculo de control, de manera que se eliminase cualquier efecto de matriz.
- 15 Las placas se incubaron con muestras de proteínas y diluciones con curvas estándar durante 1 h a 37 °C. Después de tres lavados con Tween-20 al 0,1 % en PBS (PBS-T), las placas se incubaron con un anticuerpo IgG (H+L) anti-ratón de cabra conjugado con peroxidasa (0,04 µg/ml en una solución de bloqueo) (Jackson ImmunoResearch 115-035-146) durante 1 h a 37 °C. Los lavados con PBS-T se repitieron y las placas se incubaron con un sustrato de peroxidasa de 3,3', 5, 5'-Tetrametilbencidina (TMB) Sure Blue (KPL, Gaithersburg, MD). La reacción se detuvo
- 20 añadiendo HCl 1 N y la absorbancia se leyó a 450 nm. Cada muestra se analizó por triplicado y las concentraciones se interpolaron en la porción lineal de la curva estándar.

#### Ejemplo 4: Purificación de IgG

- 25 La purificación de C5-1 a partir del material de las hojas implicó recoger hojas congeladas de *N. benthamiana* (100-150 g), añadir fosfato sódico 20 mM, NaCl 150 mM y meta-bisulfito sódico 2 mM con un pH de 5,8-6,0, y mezclar con ayuda de un mezclador comercial durante 2-3 min a temperatura ambiente. Se retiraron las fibras insolubles mediante una filtración basta de Miracloth™ (Calbiochem, San Diego, CA) y se añadió fluoruro de fenilmetansulfonilo 10 mM (PMSF) al filtrado. El extracto se ajustó a un pH 4,8 ± 0,1 con HCl 1 M y se depuró mediante centrifugación a
- 30 18.000 g durante 15 min a 2-8 °C. El sobrenadante depurado se ajustó a un pH de 8,0 ± 0,1 con TRIS 2 M, se depuró de nuevo mediante centrifugación a 18000 g durante 15 min a 2-8 °C, y se filtró en membranas secuenciales de 0,8 y 0,2 µm (Pall Corporation, Canadá). El material filtrado se concentró mediante filtración de flujo tangencial utilizando una membrana de ultrafiltración de corte con peso molecular de 100 kDa de 0,2 pies² de área eficaz (GE Healthcare Biosciences, Canadá) para reducir el volumen del material depurado de 5 a 10 veces. Después, la
- 35 muestra concentrada se aplicó a una columna de 5 mm x 5 cm (volumen de columna de 1 ml) de proteína recombinante G-Sepharose Fast Flow (Sigma-Aldrich, St-Louis, MO, Cat. n.º P4691). La columna se lavó con 5 volúmenes de columna de TRIS-HCl 20 mM, NaCl 150 mM con un pH de 7,5. El anticuerpo se eluyó con glicina 100 mM con pH de 2,9-3,0, e inmediatamente se llevó a un pH neutro mediante recolección en tubos con volúmenes calculados de TRIS-HCl 1 M a pH 7,5. Las fracciones agrupadas de anticuerpos eluidos se centrifugaron a 21000 g
- 40 durante 15 min a 2-8 °C y se almacenaron a -80 °C hasta su análisis. Después de la purificación, la columna de afinidad se limpió y se almacenó de acuerdo con las indicaciones del fabricante. El mismo medio cromatográfico pudo volverse a utilizar para varias purificaciones sin una modificación significativa del rendimiento de la purificación (hasta 10 ciclos probados).

#### 45 Ejemplo 5: Análisis de N-Glicosilación

- Las muestras que comprendían C5-1 (50 µg) se realizaron en SDS/PAGE al 15 %. Las cadenas pesadas y ligeras se revelaron con azul de Coomassie y la banda de proteínas correspondiente a la cadena pesada se extrajo y se cortó en pequeños fragmentos. Los fragmentos se lavaron 3 veces con 600 µl de una solución de N H<sub>4</sub>HCO<sub>3</sub>
- 50 0,1 M/CH<sub>3</sub>CN (1/1) durante 15 minutos cada vez y se secaron.

- La reducción de puentes de disulfuro se produjo mediante la incubación de los fragmentos de gel en 600 µl de una solución de DTT 0,1 M en NH<sub>4</sub>HCO<sub>3</sub> 0,1 M, a 56 °C durante 45 minutos. La alquilación se realizó añadiendo 600 µl de una solución de yodoacetamida 55 mM en N H<sub>4</sub>HCO<sub>3</sub> 0,1 M, a temperatura ambiente durante 30 minutos. Los
- 55 sobrenadantes se desecharon y los fragmentos de poliacrilamida se lavaron una vez más en NH<sub>4</sub>HCO<sub>3</sub> 0,1 M/CH<sub>3</sub>CN (1/1).

Después, las proteínas se digirieron con 7,5 µg de tripsina (Promega) en 600 µl de N H<sub>4</sub>HCO<sub>3</sub> 0,05 M, a 37 °C

durante 16 h. Se añadieron doscientos µl de CH<sub>3</sub>CN y se recogió el sobrenadante. Después, los fragmentos de gel se lavaron con 200 µl de NH<sub>4</sub>HCO<sub>3</sub> 0,1 M, después con 200 µl de CH<sub>3</sub>CN de nuevo y, finalmente, con 200 µl de ácido fórmico al 5 %. Todos los sobrenadantes se agruparon y se liofilizaron.

- 5 La separación peptídica mediante HPLC se realizó en una columna de fase inversa C18 (4,5 x 250 mm) con un gradiente lineal de CH<sub>3</sub>CN en TFA al 0,1%. Las fracciones se recogieron, se liofilizaron y se analizaron por MALDI-TOF-MS en un instrumento Voyager DE-Pro MALDI-TOF (Applied Biosystems, Estados Unidos) equipado con un láser de nitrógeno de 337 nm. Se realizaron los espectros de masas en el modo de extracción retardada del reflector usando ácido alfa-ciano-4-hidroxicinámico (Sigma-Aldrich) como matriz.

10

#### Ejemplo 6: Cuantificación de expresión transitoria de IgG en hojas agroinfiltradas de *Nicotiana benthamiana*

Para probar si los casetes de expresión a base de plastocianina fuertes podían provocar una alta acumulación de una IgG completamente ensamblada, las secuencias codificantes de la cadena ligera y pesada de C5-1, una IgG anti-humana murina (Khouli y col. 1997) se ensamblaron en construcciones en tándem aguas abajo del promotor de plastocianina y secuencias sin traducir 5', y se flanquearon con las secuencias no traducidas 3' de plastocianina y de terminación de transcripción en el mismo segmento de ADN-T de un plásmido binario pCambia como se representa en la figura 1.

15

- 20 Tanto en el casete de expresión R612 como R610 (véase el Ejemplo 1), las secuencias de codificación de cadena pesada y ligera contenían el péptido de señal nativo de C5-1 (Khouli y col. 1999), pero en R610, la secuencia de codificación de un péptido KDEL se añadió en el extremo C-terminal de la cadena pesada para restringir el movimiento de la IgG ensamblada al aparato de Golgi.

- 25 Tras las etapas de clonación y la transferencia de plásmidos en *Agrobacterium tumefaciens* (AGL1), cada hoja de las tres plantas de *Nicotiana benthamiana* se infiltraron con jeringa con cepas de *Agrobacterium* transformadas con los plásmidos R612 o R610, y se incubaron en condiciones de invernadero durante 6 días antes de su análisis, como se describe en el Ejemplo 2. Después del periodo de incubación, las hojas de cada planta (aproximadamente 20 g de biomasa) se congelaron, se trituraron, y el polvo congelado se mezcló para producir una muestra homogénea de la cual se tomaron 2 sub-muestras de 1,5 gramos para su extracción (de cada planta; véase el Ejemplo 3). El contenido en C5-1 se cuantificó en extractos totales de proteínas de cada muestra mediante un ensayo inmunoabsorbente ligado a enzimas (ELISA) utilizando una cadena pesada de IgG1 anti-ratón de cabra policlonal para la captura y una IgG anti-ratón de cabra conjugada con peroxidasa (H+L) para la detección (véase el Ejemplo 3).

35

Como se muestra en la figura 2, la agroinfiltración de R612 condujo a la acumulación de 106 mg del anticuerpo por kg de peso fresco, mientras que la forma retenida en ER del anticuerpo (R610) alcanzó 211 mg/kg FW en las mismas condiciones.

- 40 Debido a que el silenciamiento genético postranscripcional (PTGS) ha demostrado limitar la expresión de transgenes en plantas agroinfiltradas de *Nicotiana benthamiana* y que la co-expresión de un supresor de silenciamiento del virus de la patata Y (HcPro) contrarrestó la degradación específica de ARNm de los transgenes (Brigneti y col., 1998), la co-infiltración de una construcción HcPro (Hamilton y col., 2002) se probó para determinar su eficiencia con la expresión en aumento de C5-1. La co-expresión de R612 y R610 con HcPro aumentó los niveles de acumulación de anticuerpos 5,3 y 3,6 veces, respectivamente, en comparación con los observados en ausencia de HcPro. En presencia de HcPro, la expresión de C5-1 controlada con plastocianina alcanzó valores promedio de 558 mg/kg de FW con R612, y 757 mg/kg de FW con R610 (figura 2). Los niveles máximos de expresión de C5-1 excedieron 1,5 g/kg de FW (25 % de las proteínas solubles totales) en algunos extractos de hojas infiltradas con R612 y R610.

45

- 50 Para evaluar la escalabilidad de un sistema de expresión de agroinfiltración, la acumulación de C5-1 se cuantificó seguida de un procedimiento de infiltración al vacío adaptado de Kapila y col. (1997). En esta serie de experimentos, la parte aérea de las plantas completas se infiltraron al vacío con R612 + HcPro o R610 + HcPro y se devolvieron al invernadero durante 6 días antes de la cosecha. En un esfuerzo para proporcionar datos que sean representativos de un sistema de producción a gran escala, se congelaron lotes de aproximadamente 250 g de hojas/pecíolos de varias plantas, se molieron en una muestra homogénea, y se recogieron 3 sub-muestras de 7,5 gramos de material por lote para su análisis. Como se muestra a través de la cuantificación por ELISA, los niveles de acumulación promedio de C5-1 alcanzaron 238 y 328 mg/kg de FW para las infiltraciones de R612 y R610 respectivamente (figura 2).

55

Ejemplo 7: Caracterización del anticuerpo producido

Se usaron análisis de transferencia de proteínas (véase el Ejemplo 3) para revelar el nivel de ensamblaje y fragmentación de la IgG C5-1 en plantas que producen las formas secretadas de (R612) y de retención en ER (R610) de la proteína, después de tanto experimentos de infiltración de jeringa como al vacío. Una transferencia de Western sondeada con IgG H+L anti-ratón de cabra conjugada con peroxidasa se utilizó en primer lugar para destacar la presencia de un máximo de fragmentos de anticuerpos, independientemente de su origen en la molécula de C5-1. Como se muestra en la figura 3A, todos los extractos de proteínas contenían fragmentos de tamaño molecular similar y en abundancia relativa similar, sin importar la estrategia de direccionamiento subcelular o el método de infiltración utilizado. En cada caso, se reveló una banda mayor ( $\geq 85\%$ ) correspondiente al anticuerpo completo a aproximadamente 150 kDa, con dos bandas menores a aproximadamente 135 kDa y aproximadamente 100 kDa, mostrando que el anticuerpo se acumuló principalmente en su forma completamente ensamblada ( $H_2L_2$ ). De manera interesante, los fragmentos de movilidad electroforética similar también estaban presentes en la IgG1 de control purificada a partir de una línea celular tumoral murina (MOPC-21; Sigma #M9269), lo que sugería que las fragmentaciones producidas en las líneas celulares de plantas y animales eran similares y, quizá resultantes de actividades proteolíticas comunes. También se obtuvieron resultados similares con un anticuerpo anti-ratón específico de cadena pesada para la detección.

Para ensayar la identidad de los fragmentos de anticuerpos presentes en los extractos, se utilizó una transferencia de actividad, en la que las proteínas transferidas se prueban con una IgG1 humana conjugada con peroxidasa, el antígeno de C5-1. La identidad de un anticuerpo completamente ensamblado con una banda de aproximadamente 150 kDa puede observarse en la figura 3B. Además, el patrón de fragmentación observado en las transferencias de Western, con excepción de una banda de 100 kDa (véase la figura 3A) es visible en la transferencia de actividad (figura 3B). Sin desear quedar ligado por la teoría, este resultado sugiere que el fragmento de 100 kDa no contiene las regiones Fab del anticuerpo de C5-1, y puede consistir, al menos en parte, en dímeros de cadena pesada, un intermedio del ensamblaje de anticuerpos.

Ejemplo 8: Purificación de anticuerpos y caracterización del producto purificado

El anticuerpo se purificó a partir de la biomasa, utilizando una única etapa cromatográfica de afinidad de Proteína G y el producto obtenido se analizó mediante SDS-PAGE (véase el Ejemplo 4). El gel teñido con azul de Coomassie que se presenta en la figura 4a muestra una banda principal a 150 kDa en la fracción del eluato de la Proteína G. Esta banda representa más del 85 % del producto purificado en la forma secretada y retenida en ER, y el contenido en los contaminantes es idéntico para ambas formas (figura 4A, filas 4 y 5). Un análisis de transferencia Western, probado con una IgG anti-ratón policlonal, ha mostrado el origen murino de IgG del contaminante principal en las fracciones purificadas de C5-1 (datos no mostrados). En condiciones de reducción, se detectaron dos productos principales a aproximadamente 26 kDa y aproximadamente 55 kDa, lo que corresponde al peso molecular de las cadenas ligera y pesada, respectivamente (figura 4B, fila 2). La cadena pesada del anticuerpo retenido en ER mostró una movilidad electroforética más elevada que la cadena pesada del anticuerpo apoplásico (figura 4B, fila 3), lo que se interpreta como los resultados combinados de aminoácidos KDEL adicionales presentes en el extremo C-terminal y las diferencias en la N-glicosilación debido a la retención en ER. La figura 4C muestra que los anticuerpos purificados (150 kDa) unidos con IgG1 humana, como ocurrió con los fragmentos contaminantes de 75, 90, 100, y 120 kDa, destacando la presencia de al menos un segmento Fab en estos fragmentos. La presencia de Fab en el fragmento de 100 kDa contrastó con el resultado obtenido del análisis de extractos en bruto, donde la banda de 100 kDa no se unió con la IgG humana. Existen hipótesis de que la cantidad de fragmentos que contienen Fab y que migran a 100 kDa en el extracto en bruto fue demasiado baja para la detección con esta transferencia de actividad o que el fragmento que migra a 100 kDa consistía en dos moléculas diferentes, una de dímeros de cadena pesada (sin Fab) y la otra con regiones de unión a antígeno.

La capacidad de reproducción de este sistema para la producción de anticuerpos se evaluó con una comparación paralela de los productos purificados de 2 lotes de infiltración diferentes y 3 lotes de purificación distintos de cada lote. El análisis de SDS-PAGE teñido con azul de Coomassie de los lotes de purificación mostró la presencia de bandas idénticas en todos los lotes, y en abundancia relativa altamente similar (figura 4D).

Ejemplo 9: Modificación de N-glicosilación de anticuerpos mediante la co-expresión de una galactosiltransferasa humana

Para investigar si la co-expresión transitoria puede usarse para controlar la glicosilación de proteínas emergentes durante la expresión transitoria, se prepararon cassetes de expresión de plastocianina que comprendían

- $\beta$ 1,4galactosiltransferasa humana nativa (GalT). R622 comprendía GalT (figura 5B), y R621 comprendía el dominio catalítico de GalT fusionado con el dominio CTS de N-acetilglucosaminil transferasa (GNT1; GNT1GalT (figura 5A). El dominio CTS de N-acetilglucosaminil transferasa (GNT1) se seleccionó como anclaje de membrana para el dominio catalítico de GalT humano, ya que GNT1 actúa en una etapa temprana de la síntesis N-glicanos complejos en el ER y el aparato de cis-Golgi (Saint-Jore-Dupas y col., 2006). Sin desear quedar ligado por la teoría, el secuestro de la actividad de GalT en una etapa temprana de la maduración de la proteína puede dar como resultado la adición de  $\beta$ 1,4galactosa en glicanos de maduración y la inhibición eficiente de la fucosilación y xilosilación del núcleo. Estas construcciones se co-infiltraron en plantas con C5-1.
- 10 Las plantas de *Nicotiana benthamiana* se infiltraron con R612 (véase el Ejemplo 2) (forma secretada de C5-1), R612 + R621 (GNT1GalT) o R612 + R622 (GalT) en presencia de HcPro. La figura 6 muestra un análisis inmunológico de C5-1 purificado a partir de estas muestras de biomasa.
- La galactosilación del anticuerpo se estimó mediante afinodetección con la aglutinina de *Erythrina cristagalli* (ECA) que se une específicamente a  $\beta$ 1,4galactosa. Como se esperaba, no se detectó galactosa cuando C5-1 se expresó en solitario (R612; figura 6). La galactosilación se observó en C5-1 purificado a partir de co-infiltraciones con R512+R622 (GalT) pero no a partir de co-infiltraciones con R612+R621 (GNT1GalT, Figura 6). La transferencia de Western realizada utilizando anticuerpos anti- $\beta$ 1,3fucosa revelaron la fucosilación del N-glicano en el control C5-1 expresado sin galactosiltransferasa. La fucosilación del N-glicano no se detectó en el anticuerpo co-expresado con GNT1GalT, sin importar el método de agroinfiltración, mientras que la co-expresión con el GalT nativa no condujo a una reducción detectable en la fucosilación del anticuerpo (figura 6). Se obtuvieron resultados similares con los anticuerpos específicos anti- $\beta$ 1,2xilosa, que mostraron la ausencia total de inmunosignales específicas de xilosa en C5-1 co-expresado con GNT1GalT y su presencia cuando C5-1 se co-expresó con GalT (figura 6).
- 25 Los geles teñidos con Coomassie para estimaciones visuales directas de IgG completamente ensambleda, y las transferencias de Western y de actividad se realizaron en los mismos extractos. En base a estos datos, el sistema de expresión de anticuerpos que se describe alcanza rendimientos de 1,5 g/kg de peso fresco con más del 85 % del producto que consiste en IgG tetramérica de tamaño completo de aproximadamente 150 kDa en extractos en bruto.
- 30 La adición de un péptido KDEL en el extremo C-terminal de la cadena pesada se ha utilizado previamente para aumentar la acumulación de anticuerpos (2-10X) mediando en la recuperación del anticuerpo del Golgi de vuelta al ER (Schillberg y col., 2003). Con el sistema de expresión que se describe en el presente documento, la adición de un péptido KDEL a la cadena pesada de C5-1 duplicó el rendimiento de C5-1 cuando el supresor de HcPro de silenciamiento no estaba en uso. La diferencia entre el rendimiento de C5-1 en presencia o ausencia de KDEL se redujo de forma significativa cuando se usó HcPro para reducir el silenciamiento. La retención en ER no afectó a la calidad del producto, debido a que los fragmentos observados en los extractos en bruto de plantas que producen las formas de retención en ER y secretadas del anticuerpo tenían un tamaño idéntico y abundancia relativa.
- C5-1 se separó en un análisis por HPLC preparativa y el perfil de N-glicano del glicopéptido tríptico EEQFNSTFR (SEQ ID NO: 13) de C5-1 se analizó mediante espectrometría de masas de MALDI-TOF. Como se muestra en la figura 7A, cuando C5-1 se expresó en solitario, su población de N-glicano estuvo principalmente representada por formas complejas, incluyendo iones que consistían en oligosacáridos fucosilados y xilosilados, como se observa para la expresión estable (Bardor y col., 2003). La presencia de glicanos específicos de ER no maduros, tales como Man-8 y Man-9, pueden estar relacionados con una fracción de proteínas "en camino", como se ha indicado previamente para anticuerpos derivados de plantas producidos por expresión transitoria (Sriraman, y col., 2004).
- La co-expresión del anticuerpo C5-1 con GalT nativa dio como resultado una modificación significativa de la estructura de N-glicano, aunque los iones correspondientes a los N-glicanos de tipo de manosa elevada siguieron siendo abundantes. El complejo fucosilado principal y el N-glicano xilosilado (J) desaparecieron y se detectaron nuevos oligosacáridos parcialmente galactosilados, algunos de los cuales contenían maduraciones específicas de plantas y extensiones de galactosa, tales como GalGlcNAcMan<sub>3</sub>(Xyl)GlcNAc<sub>2</sub> (figura 7B). Esto demostró que la co-expresión de C5-1 con  $\beta$ 1,4galactosiltransferasa humana dio como resultado un glico-diseño eficiente del anticuerpo derivado de la planta.
- 50 La co-expresión de C5-1 con GNT1GalT produjo una preparación de C5-1 purificada, en la que la población de N-glicano era significativamente distinta de la de GalT/C5-1. Como se muestra en la figura 7C, los híbridos galactosilados y no galactosilados (GalGlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> (K), y GalGlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> (G), estaban presentes junto con N-glicanos de oligomanosa. Sin desear quedar ligado por la teoría, según las hipótesis de Bakker y col. (2006), GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> (G) y Man<sub>5</sub>GlcNAc<sub>2</sub> (B) pueden ser fragmentos derivados de la degradación del

GalGlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> híbrido mediante glicosidasas endógenas, en lugar de intermediarios de la formación de N-glicanos maduros. El efecto del anclaje de membrana de GNT1 fue notable; el C5-1 purificado a partir de plantas en las que la enzima sintética GNT1GalT se co-expresó de forma transitoria con C5-1 no contenía residuos ( $\leq 99\%$ ) de glicanos que albergasen  $\alpha 1,3$ fucosa o  $\beta 1,2$ xilosa específicas de plantas, lo que demuestra que la modificación completa de la fucosilación y xilosilación se logró durante la co-expresión transitoria.

La presente invención se ha descrito con respecto a una o más realizaciones. Sin embargo, será evidente para los expertos en la técnica que pueden hacerse varias variaciones y modificaciones sin apartarse del alcance de la invención, según se define en las reivindicaciones.

10

#### Referencias:

- Bakker, H. *et al.* An antibody produced in tobacco expressing a hybrid beta-1,4-galactosyltransferase is essentially devoid of plant carbohydrate epitopes. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7577-7582 (2006).
- Boisson, M., *et al.* A rabidopsis glucosidase mutants reveal a critical role of N-glycan trimming in seed development. *EMBO J.* 20, 1010-1019 (2001).
- Brigneti, G. *et al.* Viral pathogenicity determinants are suppressors of transgene silencing in *Nicotiana benthamiana*. *EMBO J.* 17, 6739-6746 (1998).
- Chua, Y.L., Watson, L.A., & Gray, J.C. The transcriptional enhancer of the pea plastocyanin gene associates with the nuclear matrix and regulates gene expression through histone acetylation. *Plant Cell* 15, 1468-1479 (2003).
- Cox, K.M., *et al.* Glycan optimization of a human monoclonal antibody in the aquatic plant *Lemna minor*. *Nat. Biotechnol.* 24, 1591-1597 (2006).
- D'Aoust, M. A. *et al.* Efficient and reliable production of pharmaceuticals in alfalfa. *Molecular Farming*, pp 1-12. Rainer Fischer and Stefan Schillberg (eds.), Wiley-VCH, Weinheim, Germany (2004).
- Darveau, A., Pelletier, A. & Perreault, J. PCR-mediated synthesis of chimeric molecules. *Methods Neurosc.* 26, 77-85 (1995).
- Elmayan, T. & Vaucheret, H. Expression of single copies of a strongly expressed 35S transgene can be silenced post-transcriptionally. *Plant J.* 9, 787-797 (1996).
- Fischer, R. *et al.* Towards molecular farming in the future: transient protein expression in plants. *Biotechnol. Appl. Biochem.* 30, 113-116 (1999).
- Fischer, R., Drossard, J., Commandeur, U., Schillberg, S. & Emans, N. Towards molecular farming in the future: moving from diagnostic protein and antibody production in microbes to plants. *Biotechnol. Appl. Biochem.* 30, 101-108 (1999).
- Giritch, A. *et al.* High-yield expression of full-size IgG antibodies in plants coinfected with noncompeting viral vectors. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14701-14706 (2006).
- Gomord, V., Chamberlain, P., Jefferis, R. & Faye, L. Biopharmaceutical production in plants: problems, solutions and opportunities. *Trends Biotechnol.* 23, 559-565 (2005).
- Hamilton, A., Voinnet, O., Chappell, L. & Baulcombe, D. Two classes of short interfering RNA in RNA silencing. *EMBO J.* 21, 4671-4679 (2002).
- Hiatt, A., Cafferkey, R. & Bowdich, K. Production of antibodies in transgenic plants. *Nature* 342, 76-78 (1989).
- Hiatt, A. & Pauly, M. Monoclonal antibodies from plants: a new speed record. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14645-14646 (2006).
- Hibino, Y., Ohzeki, H., Sugano, N. & Hiraga, K. Transcription modulation by a rat nuclear scaffold protein, P130, and a rat highly repetitive DNA component or various types of animal and plant matrix or scaffold attachment regions. *Biochem. Biophys. Res. Commun.* 279, 282-287 (2000).
- Höfgen, R. & Willmitzer, L. Storage of competent cells for *Agrobacterium* transformation. *Nucleic Acid Res.* 16, 9877 (1988).
- Hull, A. K. *et al.* Human-derived, plant-produced monoclonal antibody for the treatment of anthrax. *Vaccine* 23, 2082-2086 (2005).
- Kapila, J., Derycke, R., Van Montagu, M. & Angenon, G. An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci.* 122, 101-108 (1997).
- Kathuria, S.R. *et al.* Functional recombinant antibodies against human chorionic gonadotropin expressed in plants. *Curr. Sci.* 82, 1452-1456 (2002).
- Ko, K. *et al.* Function and glycosylation of plant-derived antiviral monoclonal antibody. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8013-8018 (2003).
- Ko, K. & Koprowski, H. Plant biopharming of monoclonal antibodies. *Virus Res.* 111, 93-100 (2005).
- Koprivova, A. *et al.* Targeted knockouts of phycomitrella lacking plant-specific immunogenic N-glycans.

Plant Biotechnol. J. 2, 517-523 (2004).

Khoudi, H. et al. Production of a diagnostic monoclonal antibody in perennial alfalfa plants. Biotechnol. Bioeng. 64, 135-143 (1999).

5 Liu, L. & Lomonosoff, G.P. Agroinfection as a rapid method for propagating Cowpea mosaic virus-based constructs. J. Virol. Methods 105, 343-348 (2002).

Ma, J. K-C., Drake, P.M.W., Chargelegue, D., Obregon, P. & Prada, A. Antibody processing and engineering in plants, and new strategies for vaccine production. Vaccine 23, 1814-1818 (2005).

Misaki, R., Fujiyama, K. & Seki, T. Expression of human CMP-N-acetylneuraminic acid synthetase and

10 CMP-sialic acid transporter in tobacco suspension-cultured cell. Biochem. Biophys. Res. Com. 339, 1184-1189 (2004).

Paccalet et al. Engineering of a sialic acid synthesis pathway in transgenic plants by expression of bacterial Neu5Ac-synthesizing enzymes. Plant Biotech. J. 5, (2007).

Pagny, S. et al. Protein recycling from the Golgi to the endoplasmic reticulum is very active in plants but has a minor contribution to calreticulin retention. Plant Cell 12, 739-752 (2000).

15 Pagny, S. et al. Structural requirements for Arabidopsis  $\beta$ 1,2-xylosyltransferase activity and targeting to the Golgi. Plant J. 33, 189-203 (2003).

Peterson, E., Owens, S.M. & Henry, R. L. Monoclonal antibody form and function: manufacturing the right antibodies for treating drug abuse. AAPS J. 8, E383-E390 (2006).

Petrucelli, S. et al. A KDEL-tagged monoclonal antibody is efficiently retained in the endoplasmic reticulum

20 in leaves, but is both partially secreted and sorted to protein storage vacuoles in seeds. Plant Biotechnol. J. 4, 511-527 (2006).

Pwee, K-H. & Gray, J.C. The pea plastocyanin promoter directs cell-specific but not full light-regulated expression in transgenic tobacco plants. Plant J. 3, 437-449 (1993).

Rodriguez, M. et al. Transient expression in tobacco leaves of an aglycosylated recombinant antibody

25 against the epidermal growth factor receptor. Biotechnol. Bioeng. 89, 188-194 (2005).

Rouwendaal, G. J.A. et al. Efficient introduction of a bisecting GlNAc residue in tobacco N-glycans by expression of the gene encoding human N-acetylglucosaminyltransferase II. Glycobiology 17, 334-344 (2007)

Saint-Jore-Dupas, C. et al. Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. Plant Cell 18, 3182-3200 (2006).

30 Sambrook J., and Russell, D.W., Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York. (2001).

Sandhu, J.S., Webster, C.I., & Gray, J.C. A/T-rich sequences act as quantitative enhancers of gene expression in transgenic tobacco and potato plants. Plant Mol. Biol. 37, 885-896 (1998).

35 Schillberg, S., Fischer, R. & Emans, N. Molecular farming of recombinant antibodies in plants. Cell. Mol. Life Sci. 60, 433-445 (2003).

Sharp, J. M. & Doran, P. M. Characterization of monoclonal antibody fragments produced by plant cells. Biotechnol. Bioeng. 73, 338-346 (2001).

Sriraman, R. et al. Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core- $\alpha$ (1-3)-fucose residues. Plant Biotechnol. J. 2, 279-287 (2004).

40 Strasser, R., Altmann, F., Mach, L., G16ssl, J. & Steinkellner, H. Generation of Arabidopsis thaliana plants with complex N-glycans lacking  $\beta$ 1,2 linked xylose and core  $\alpha$ 1,3-linked fucose. FEBS Lett. 561, 132-136 (2004).

Szittyá, G. et al. Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. EMBO J. 22, 633-640 (2003).

45 Verch, T., Yusibov, V. & Koprowski, H. Expression and assembly of a full-length monoclonal antibody in plants using a plant virus vector. J. Immunol. Methods 220, 69-75 (1998).

Watzel, G., Bachofner, R. & Berger, E.G.. Immunocytochemical localization of the Golgi apparatus using protein-specific antibodies to galactosyltransferase. Eur. J. Cell. Biol. 56, 451-458 (1991).

50 Yusibov, V., Rabindran, S., Commandeur, U. Wyman, R.M. & Fischer, R. The potential of plant virus vectors for vaccine production. Drugs R.D. 7, 203-217 (2006).



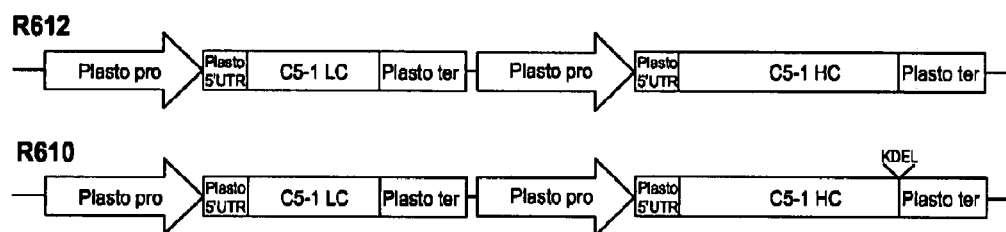
## REIVINDICACIONES

1. Un método para sintetizar una proteína de interés con una reducción de fucosilación y xilosilación que comprende, introducir en una planta, una porción de una planta, o una célula vegetal, una secuencia nucleotídica que comprende del 80 % al 100 % de identidad con una secuencia nucleotídica definida por la SEQ ID NO: 17 y que codifica una proteína híbrida que comprende una cola citoplásmica, un dominio transmembrana, una región tallo (dominio CTS) de N-acetilglucosaminil transferasa (GNT1) fusionada a un dominio catalítico de bet a-1,4galactosiltransferasa (GalT), la primera secuencia nucleotídica unida operativamente con una primera región reguladora que está activa en la planta, y una segunda secuencia nucleotídica para codificar la proteína de interés, la secuencia nucleotídica unida operativamente con una segunda región reguladora que está activa en la planta, y co-expresar de forma transitoria la primera y segunda secuencias nucleotídicas para sintetizar la proteína de interés que comprende menos del 10 % de fucosilación y xilosilación en comparación con la misma proteína de interés producida en una planta de tipo silvestre.
2. El método de la reivindicación 1, en el que la primera región reguladora es un promotor de plastocianina, o un promotor 35S, y la segunda región reguladora es un promotor de plastocianina, o un promotor 35S.
3. El método de la reivindicación 1, en el que una tercera secuencia nucleotídica se expresa dentro de la planta, la tercera secuencia nucleotídica que codifica un supresor de silenciamiento, la tercera secuencia nucleotídica unida operativamente con una tercera región reguladora que está activa en la planta.
4. El método de la reivindicación 3, donde la proteína de interés se sintetiza en una cantidad de hasta 1,5 g por kilogramo de peso fresco de planta.
5. El método de la reivindicación 3, en el que la tercera región reguladora es un promotor de plastocianina, o un promotor 35S.
6. El método de la reivindicación 2, en el que una tercera secuencia nucleotídica se expresa en la planta, la tercera secuencia nucleotídica que codifica un supresor del silenciamiento, la tercera secuencia nucleotídica unida operativamente con una tercera región reguladora que está activa en la planta.
7. El método de la reivindicación 6, en el que la tercera región reguladora es un promotor de plastocianina, o un promotor 35S.
8. El método de la reivindicación 1, en el que la proteína de interés es un anticuerpo, un antígeno o una vacuna.
9. El método de la reivindicación 8, en el que la segunda secuencia nucleotídica que codifica la proteína de interés comprende una secuencia nucleotídica 2A, unida operativamente con una región reguladora 2A que está activa en la planta, y una secuencia nucleotídica 2B, unida operativamente con una región reguladora 2B que está activa en la planta, y el producto codificado por cada una de 2A y 2B se combinan para producir un anticuerpo.
10. El método de la reivindicación 9, en el que la región reguladora 2A es un promotor de plastocianina, y la región reguladora 2B es un promotor de plastocianina.
11. Un ácido nucleico que comprende los nucleótidos 1 a 1062 de la SEQ ID NO: 17 (GNT1-GalT), en el que la secuencia nucleotídica codifica una proteína que modifica la glicosilación de una proteína de interés.
12. Una proteína híbrida GNT1-GalT, que comprende un dominio CTS de N-acetilglucosaminil transferasa fusionado a un dominio catalítico de bet a-1,4galactosiltransferasa, comprendiendo la proteína híbrida los aminoácidos 1 a 354 de la secuencia SEQ ID NO: 18.
13. Una planta, célula vegetal, o una semilla que comprende la secuencia nucleotídica de la reivindicación 12.
14. Una planta, célula vegetal, o una semilla que comprende la proteína híbrida de la reivindicación 12.
15. El método de la reivindicación 3, en el que la proteína de interés es un anticuerpo y se sintetiza en una

cantidad de hasta 1,5 g por kilogramo de peso fresco de hoja.

16. El método de la reivindicación 1, en el que la proteína de interés sintetizada carece de fucosilación, xilosilación, o ambas.

5



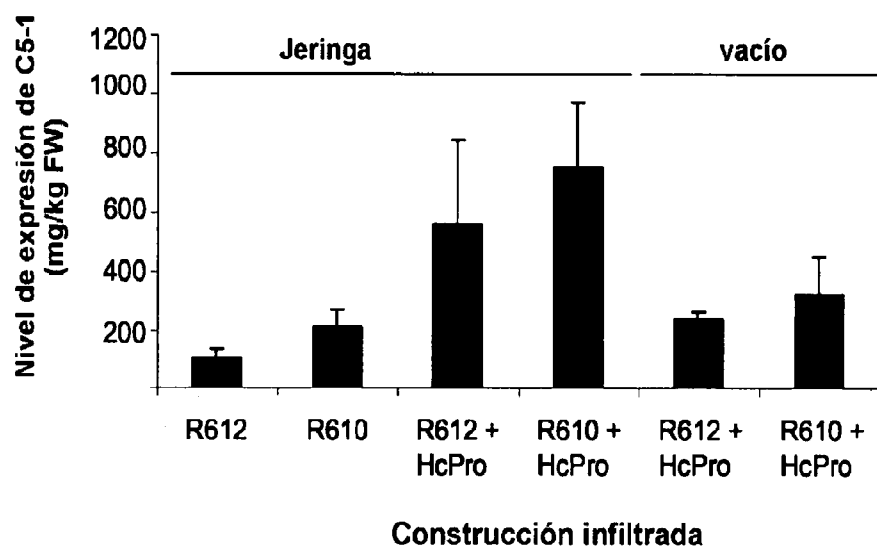
**Figura 1A**

AGAGGTACCC CGGGCTGGTA TATTTATATG TTGTCAAATA ACTCAAAAAC  
 CATAAAAGTT TAAGTTAGCA AGTGTGTACA TTTTACTTG AACAAAAATA  
 TTCACCTACT ACTGTTATAA ATCATTATTA AACATTAGAG TAAAGAAATA  
 TGGATGATAA GAACAAGAGT AGTGATATTT TGACAACAAT TTTGTTGCAA  
 CATTTGAGAA AATTTTGTTG TTCTCTCTTT TCATTGGTCA AAAACAATAG  
 AGAGAGAAAA AGGAAGAGGG AGAATAAAAA CATAATGTGA GTATGAGAGA  
 GAAAGTTGTA CAAAAGTTGT ACCAAAATAG TTGTACAAAT ATCATTGAGG  
 AATTTGACAA AAGCTACACA AATAAGGGTT AATTGCTGTA AATAAATAAG  
 GATGACGCAT TAGAGAGATG TACCATTAGA GAATTTTTGG CAAGTCATTA  
 AAAAGAAAGA ATAAATTATT TTTAAAATTA AAAGTTGAGT CATTTGATTA  
 AACATGTGAT TATTTAATGA ATTGATGAAA GAGTTGGATT AAAGTTGTAT  
 TAGTAATTAG AATTTGGTGT CAAATTTAAT TTGACATTTG ATCTTTTCCT  
 ATATATTGCC CCATAGAGTC AGTTAACTCA TTTTATATTT TCATAGATCA  
 AATAAGAGAA ATAACGGTAT ATTAATCCCT CCAAAAAAAA AAAACGGTAT  
 ATTTACTAAA AAATCTAAGC CACGTAGGAG GATAACAGGA TCCCCGTAGG  
 AGGATAACAT CCAATCCAAC CAATCACAAC AATCCTGATG AGATAACCCA  
 CTTTAAGCCC ACGCATCTGT GGCACATCTA CATTATCTAA ATCACACATT  
 CTTCCACACA TCTGAGCCAC ACAAAAACCA ATCCACATCT TTATCACCCA  
 TTCTATAAAA AATCACACTT TGTGAGTCTA CACTTTGATT CCCTTCAAAC  
 ACATACAAAG AGAAGAGACT AATTAATTAA TTAATCATCT TGAGAGAAAA  
TG

**Figura 1B**

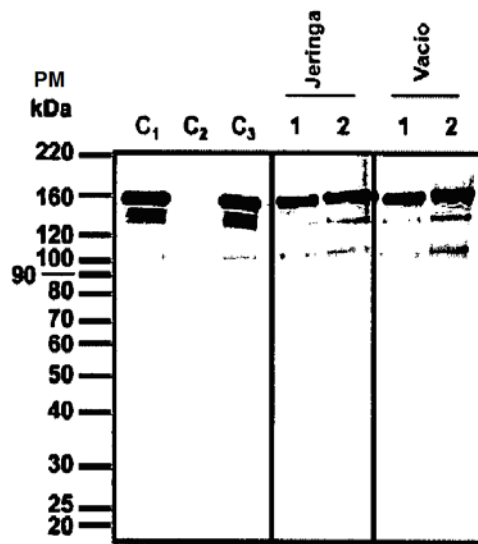
TAAGTTAAAA TGCTTCTTCG TCTCCTATTT ATAATATGGT TTGTTATTGT  
 TAATTTTGTT CTTGTAGAAG AGCTTAATTA ATCGTTGTTG TTATGAAATA  
 CTATTTGTAT GAGATGAACT GGTGTAATGT AATTCATTTA CATAAGTGGA  
 GTCAGAATCA GAATGTTTCC TCCATAACTA ACTAGACATG AAGACCTGCC  
 GCGTACAATT GTCTTATATT TGAACAACTA AAATTGAACA TCTTTTGCCA  
 CAACTTTATA AGTGGTTAAT ATAGCTCAA TATATGGTCA AGTTCAATAG  
 ATTAATAATG GAAATATCAG TTATCGAAAT TCATTAACAA TCAACTTAAC  
 GTTATTAACT ACTAATTTTA TATCATCCCC TTTGATAAAT GATAGTACA

**Figura 1C**

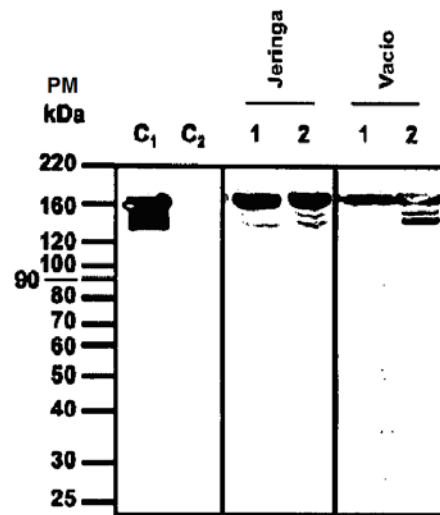


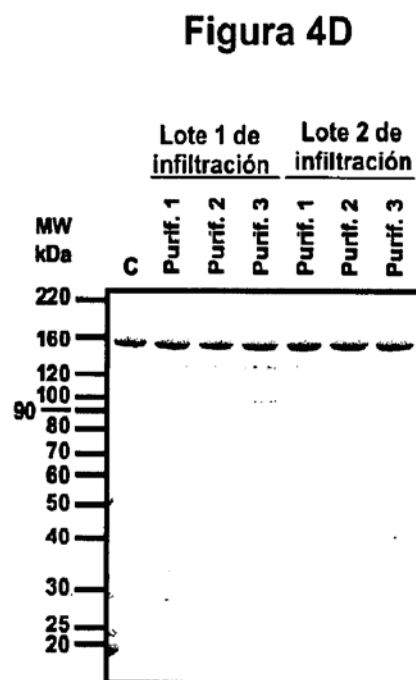
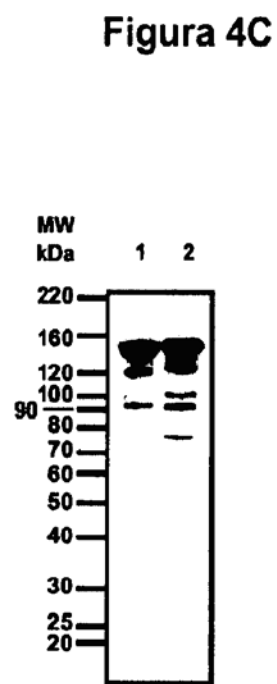
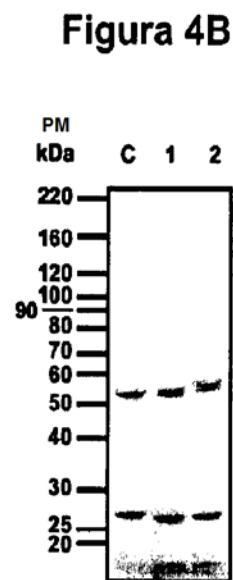
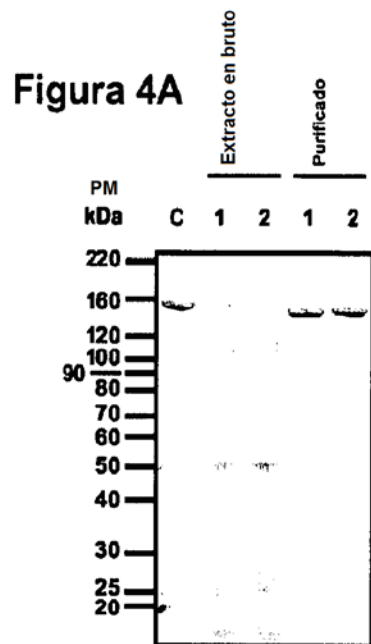
**Figura 2**

**Figura 3A**



**Figura 3B**







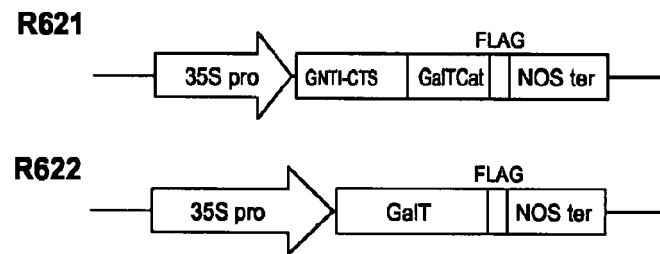


Figura 5A

ATGAGGCTTCGGGAGCCGCTCCTGAGCGGCAGCGCCGCGATGCCAGGCGCGTCCCTACAGCGGGC  
 CTGCCGCTGCTCGTGGCGCTCTGCGCTCTGCACCTTGGCGTCACCTCGTTTACTACCTGGCTGGC  
 CGCGACCTGAGCCGCTGCCCAACTGGTCGGAGTCTCCACACCGCTGCAGGGCGGCTCGAACAGT  
 GCCGCCGCCATCGGGCAGTCCCGGGGAGCTCCGGACCGGAGGGGCCCGGCCGCTCCTCT  
 AGGCGCCTCCTCCAGCCGCGCCCGGGTGGCGACTCCAGCCAGTCTGGATTCTGGCCCTGGCCC  
 CGCTAGCAACTTGACCTCGGTCCAGTGCCCCACACCACCGCACTGTGCTGCCCGCCTGCCCTGA  
 GGAGTCCCCGCTGCTTGTGGGCCCCATGCTGATTGAGTTTAAACATGCCTGTGGACCTGGAGCTCGT  
 GGCAAAGCAGAACCCTAAATGTGAAGATGGGCGGCCGCTATGCCCCAGGGACTGCGTCTCTCCTC  
 ACAAGGTGGCCATCATCATTCCATTCCGCAACCGGCAGGAGCACCTCAAGTACTGGCTATATTATT  
 GCACCCAGTCCCTGCAGCGCCAGCAGCTGGACTATGGCATCTATGTTATCAACCAGGCGGGAGACAC  
 TATATTCAATCGTGCTAAGCTCCTCAATGTTGGCTTTCAAGAAGCCTTGAAGGACTATGACTACACCT  
 GCTTTGTGTTTAGTGACGTGGACCTCATTCCAATGAATGACCATAATGCGTACAGGTGTTTTTCACAG  
 CCACGGCACATTTCCGTTGCAATGGATAAGTTTGGATTACGCTACCTTATGTTAGTATTTTGGAGG  
 TGTCTCTGCTCTAAGTAAACAACAGTTTCTAACCATCAATGGATTTCCTAATAATTATTGGGGCTGGG  
 GAGGAGAAGATGATGACATTTTAAACAGATTAGTTTTAGAGGCATGTCTATATCTCGCCCAATGC  
 TGTGGTCCGGAGGTGTCGCATGATCCGCCACTCAAGAGACAAGAAAAATGAACCCATCCTCAGAG  
 GTTTGACCGAATTGCACACACAAAGGAGACAATGCTCTCTGATGGTTTGAACCTCACTACCTACCAG  
 GGGCTGGATGTACAGAGATACCCATTGTATACCCAAATCACAGTGGACATCGGGACACCGAGCGAT  
 TACAAGGATGACGATGACAAGATCGATTAG

Figura 5B

MRLREPLLSSGAAMPGASLQ**RACRL**LVA**VCALHLGVTLVYYLAG**RDLSRLPQLVGVSTPLQGGNSAAAI  
 GQSSGELRTGGARPPPPLGASSQPRPGDSSPVVDSGPGPASNLTSVPVPHTTALSLPACPEESPLLVG  
 PMLIEFNMPVDLELVAKQNPVVKMGGRYAPRDCVSPHKVAIIIPFRNRQEHLKYWLYLHPVLQRQQLD  
 YGIYVINQAGDTIFNRAKLLNVGFQALKDYDYTCFVSDVDLIPMNDHNAYRCFSQPRHISVAMDKFGF  
 SLPYVQYFGVVSALSKQQLTINGFPNNYWGWWGGEDDIFNRLVFRGMSISRPNAVVGRCRMRHSRD  
 KKNENPNQRFDRIAHTKETMLSDGLNSLTYQGLDVQRYPLYTQITVDIGTPSDYKDDDDKID\*

Figura 5C

ATGAGAGGGTACAAGTTTTGCTGTGATTCCGGTACCTCCTCATCTTGGCTGCTGTGCGCCTTCATCTAC  
ATACAGATGCGGCTTTTTGCGACACAGTCAGAATATGCAGATCGCCTTGCTGCTGCAATTGAAGCAGAA  
AATCACTGTACAAGTCAGACCAGATTGCTTATTGACCAGATTAGCCAGCAGCAAGGAAGAATAGTTGCT  
CTTGAAGAACAAATGAAGCGTCAGACTAGTGCCTGTGCGTGGCCGCTGCCCTGAGGAGTCCCCGC  
TGCTTGTGGGCCCCATGCTGATTGAGTTTAACATGCCTGTGGACCTGGAGCTCGTGGCAAAGCAGAA  
CCCAAATGTGAAGATGGGCGGCCGCTATGCCCCAGGGACTGCGTCTCTCCTCACAAAGGTGGCCAT  
CATCATTCCATTCCGCAACCGGCAGGAGCACCTCAAGTACTGGCTATATTATTTGCACCCAGTCCTG  
CAGCGCCAGCAGCTGGACTATGGCATCTATGTTATCAACCAGGCGGGAGACACTATATTCAATCGTG  
CTAAGCTCCTCAATGTTGGCTTTCAAGAAGCCTTGAAGGACTATGACTACACCTGCTTTGTGTTTAGT  
GACGTGGACCTCATTCCAATGAATGACCATAATGCGTACAGGTGTTTTTCACAGCCACGGGCACATTT  
CCGTTGCAATGGATAAGTTTGGATTGAGCCTACCTTATGTTTCAAGTATTTGGAGGTGTCTCTGCTCTA  
AGTAAACAACAGTTTCTAACCATCAATGGATTTCCTAATAATTATTGGGGCTGGGGAGGAGAAGATG  
ATGACATTTTTAACAGATTAGTTTTAGAGGCATGTCTATATCTCGCCCAAATGCTGTGGTGGGAGG  
TGTCGCATGATCCGCCACTCAAGAGACAAGAAAAATGAACCCAATCCTCAGAGGTTTGACCGAATTG  
CACACACAAAGGAGACAATGCTCTCTGATGGTTTGAACCTCACTCACCTACCAGGGGGCTGGATGTACA  
GAGATACCCATTGTATACCCAAATCACAGTGGACATCGGGACACCGAGCGATTACAAGGATGACGAT  
GACAAGATCGATTAG

Figura 5D

MRGYKFCCDFRYLLILA AVAFIYIQMRLFATQSEYADRLAAAEAEHCTSQTRLLIDQISQQQGRIVALEEQ  
MKRQTSALSLPACPEESPLLVGPMLEFNMPVDLELVAKQNPVNMKGGRYAPRDCVSPHKVAIIPFRNR  
QEHLKYWLYLHPVLQRQQLDYGIYINQAGDTIFNRAKLLNVGFQEALKDYDYTCFVFSVDVLPMDH  
NAYRCFSQPRHISVAMDKFGFSLPYVQYFGGVSALSKQQLTINGFPNNYWGWWGEDDDIFNRLVFRG  
MSISRPNAVVGRCRMIRHSRDKKNEPNPQRFDRIAHTKETMLSDGLNSLTYQGLDVQRYPLYTQITVDIG  
TPSDYKDDDDKID\*

Figura 5E

TACCTCCTCATCTTGGCTGCTGTCGCCTTCATCTACATACAGATGCGGCTTTTT

## **Figura 5F**

YLLILAAVAFIYIQMRLF

## **Figura 5G**

Secuencia nucleotídica GNTI-GNTIII

|            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------|
| ATGAGAGGGA | ACAAGTTTTG | CTGTGATTTT | CGGTACCTCC | TCATCTTGGC | 50   |
| TGCTGTCGCC | TTCATCTACA | CACAGATGCG | GCTTTTTGCG | ACACAGTCAG | 100  |
| AATATGCAGA | TCGCCTTGCT | GCTGCAATTG | AAGCAGAAAA | TCATTGTACA | 150  |
| AGCCAGACCA | GATTGCTTAT | TGACCAGATT | AGCCTGCAGC | AAGGAAGAAT | 200  |
| AGTTGCTCTT | GAAGAACAAA | TGAAGCGTCA | GACTAGTTCG | CCCCTGCTGC | 250  |
| AGCCGCTGCC | GCCCAGCAAG | GCGGCCGAGG | AGCTCCACCG | GGTGGACTTG | 300  |
| GTGCTGCCCC | AGGACACCAC | CGAGTATTTT | GTGCGCACCA | AGGCCGCGCG | 350  |
| CGTCTGCTTC | AAACCCGGCA | CCAAGATGCT | GGAGAGGCCG | CCCCCGGGAC | 400  |
| GGCCGGAGGA | GAAGCCTGAG | GGGGCCAACG | GCTCCTCGGC | CCGGCGGCCA | 450  |
| CCCCGGTACC | TCCTGAGCGC | CCGGGAGCGC | ACGGGGGGCC | GAGGCGCCCC | 500  |
| GCGCAAGTGG | GTGGAGTGCG | TGTGCCTGCC | CGGCTGGCAC | GGACCCAGCT | 550  |
| GCGGCGTGCC | CACTGTGGTG | CAGTACTCCA | ACCTGCCAC  | CAAGGAGCGG | 600  |
| CTGGTGCCCA | GGGAGGTGCC | GCGCCGCGTC | ATCAACGCCA | TCAACGTCAA | 650  |
| CCACGAGTTC | GACCTGCTGG | ACGTGCGCTT | CCACGAGCTG | GGCGACGTGG | 700  |
| TGGACGCCTT | TGTGGTGTGC | GAGTCCAAC  | TCACGGCTTA | TGGGGAGCCG | 750  |
| CGGCCGCTCA | AGTTCCGGGA | GATGCTGACC | AATGGCACCT | TCGAGTACAT | 800  |
| CCGCCACAAG | GTGCTCTATG | TCTTCCTGGA | CCACTTCCCG | CCCGGCGGCC | 850  |
| GGCAGGACGG | CTGGATCGCC | GACGACTACC | TGCGCACCTT | CCTCACCCAG | 900  |
| GACGGCGTCT | CGCGGCTGCG | CAACCTGCGG | CCCGACGACG | TCTTCATCAT | 950  |
| TGACGATGCG | GACGAGATCC | CGGCCCGTGA | CGGCGTCCTT | TTCCTCAAGC | 1000 |
| TCTACGATGG | CTGGACCGAG | CCCTTCGCCT | TCCACATGCG | CAAGTCGCTC | 1050 |
| TACGGCTTCT | TCTGGAAGCA | GCCGGGCACC | CTGGAGGTGG | TGTCAGGCTG | 1100 |
| CACGGTGAC  | ATGCTGCAGG | CAGTGTATGG | GCTGGACGGC | ATCCGCCTGC | 1150 |
| GCCGCCGCCA | GTACTACACC | ATGCCCAACT | TCAGACAGTA | TGAGAACCGC | 1200 |
| ACCGGCCACA | TCCTGGTGCA | GTGGTCGCTG | GGCAGCCCCC | TGCACTTCGC | 1250 |
| CGGCTGGCAC | TGCTCCTGGT | GCTTCACGCC | CGAGGGCATC | TACTTCAAGC | 1300 |
| TCGTGTCCGC | CCAGAATGGC | GACTTCCCAC | GCTGGGGTGA | CTACGAGGAC | 1350 |
| AAGCGGGACC | TGAACATCAT | CCGCGGCCTG | ATCCGCACCG | GGGGCTGGTT | 1400 |
| CGACGGCACG | CAGCAGGAGT | ACCCGCCTGC | AGACCCAGC  | GAGCACATGT | 1450 |
| ATGCGCCCAA | GTACCTGCTG | AAGAACTACG | ACCGGTTCCA | CTACCTGCTG | 1500 |
| GACAACCCCT | ACCAGGAGCC | CAGGAGCACG | GCGGCGGGCG | GGTGGCGCCA | 1550 |
| CAGGGGTCCC | GAGGGAAGGC | CGCCCGCCCG | GGGCAAAC   | GACGAGGCGG | 1600 |
| AAGTCGAACA | AAAACATCAT | TCAGAAGAGG | ATCTGAATTA | G          | 1641 |

Figura 5H

**Secuencia peptídica GNTIGNTIII**

MRGNKFCCDFRYLLILAAVAFIYTMRLFATQSEYADRLAAAIEAENHCTSQTRLLIDQ  
ISLQQGRIVALEEQMKRQTSSPLLQPLPPSKAAEELHRVDLVLPEDTTEYFVRTKAGGV  
CFKPGTKMLERPPPGRPPEEKPEGANGSSARRPPRYLLSARERTGGRGARRKWVECVCLP  
GWHGPSCGVPTVVQYSNLPTKERLVPREVP RRVINAINVNHEFDLLDVRFHGELGDVVDA  
FVVCESNFTAYGEP RPLK FREMLTNGTFEYIRHKVLYVFLDHFPPGGRQDGWIADDYLR  
TFLTQDGVSRRLNLRPDDVFIIDDADEIPARDGVLFLKLYDGWTEPFAFHRKSLYGFF  
WKQPGTLEV VSGCTVDMLQAVYGLDGIRLRRRQYYTMPNFRQYENRTGHILVQWSLGSP  
LHFAGWHCSWCFTPEGIYFKLVSAQNGDFPRWGDYEDKRDNLNIRGLIRTGGWFDGTQQ  
EYPPADPSEHMYAPKYLLKNYDRFHYLLDNPYQEPRSTAAGGWRHRGPEGRPPARGKLD  
EAEVEQKLI SEEDLN\*

**Figura 5I**

**Secuencia peptídica GNTIII**

TSSPLLQPLPPSKAAEELHRVDLVLPEDTTEYFVRTKAGGVCFKPGTKMLERPPPGRPE  
EKPEGANGSSARRPPRYLLSARERTGGRGARRKWVECVCLPGWHGPSCGVPTVVQYSNL  
PTKERLVPREVP RRVINAINVNHEFDLLDVRFHGELGDVVDAFVVCESNFTAYGEP RPLK  
FREMLTNGTFEYIRHKVLYVFLDHFPPGGRQDGWIADDYLR TFLTQDGVSRRLNLRPDD  
VFIIDDADEIPARDGVLFLKLYDGWTEPFAFHRKSLYGFFWKQPGTLEV VSGCTVDML  
QAVYGLDGIRLRRRQYYTMPNFRQYENRTGHILVQWSLGSP LHFAGWHCSWCFTPEGIY  
FKLVSAQNGDFPRWGDYEDKRDNLNIRGLIRTGGWFDGTQQEYPPADPSEHMYAPKYLL  
KNYDRFHYLLDNPYQEPRSTAAGGWRHRGPEGRPPARGKLD EAEVEQKLI SEEDLN\*

**Figura 5K**

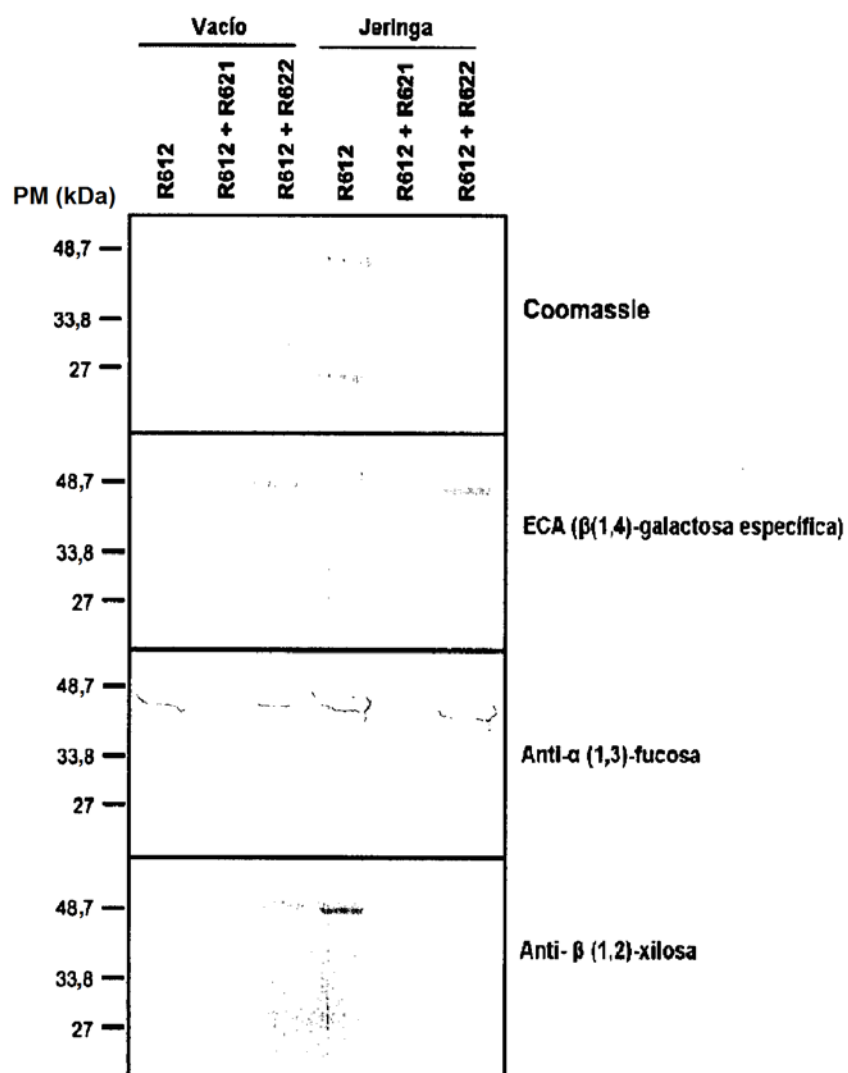
## GNT III

```

—
ACTAGTTCGC CCCTGCTGCA GCCGCTGCCG CCCAGCAAGG CGGCCGAGGA 50
GCTCCACCGG GTGGACTTGG TGCTGCCCCA GGACACCACC GAGTATTTTCG 100
TGCGCACCAA GGCCGGCGGC GTCTGCTTCA AACCCGGCAC CAAGATGCTG 150
GAGAGGCCGC CCCC GGACG GCCGGAGGAG AAGCCTGAGG GGGCCAACGG 200
CTCCTCGGCC CGGCGGCCAC CCCGGTACCT CCTGAGCGCC CGGGAGCGCA 250
CGGGGGGCCG AGGCGCCCCG CGCAAGTGGG TGGAGTGCGT GTGCCTGCCC 300
GGCTGGCACG GACCCAGCTG CGGCGTGCCC ACTGTGGTGC AGTACTCCAA 350
CCTGCCCCACC AAGGAGCGGC TGGTGCCCCA GGAGGTGCCG CGCCGCGTCA 400
TCAACGCCAT CAACGTCAAC CACGAGTTTC ACCTGCTGGA CGTGCGCTTC 450
CACGAGCTGG GCGACGTGGT GGACGCCTTT GTGGTGTGCG AGTCCAACCT 500
CACGGCTTAT GGGGAGCCGC GGCCGCTCAA GTTCCGGGAG ATGCTGACCA 550
ATGGCACCTT CGAGTACATC CGCCACAAGG TGCTCTATGT CTTCTGGAC 600
CACTTCCCGC CCGGCGGCCG GCAGGACGGC TGGATCGCCG ACGACTACCT 650
GCGCACCTTC CTCACCCAGG ACGGCGTCTC GCGGCTGCGC AACCTGCGGC 700
CCGACGACGT CTTTCATCATT GACGATGCGG ACGAGATCCC GGCCCCGTGAC 750
GGCGTCCTTT TCCTCAAGCT CTACGATGGC TGGACCGAGC CCTTCGCCTT 800
CCACATGCGC AAGTCGCTCT ACGGCTTCTT CTGGAAGCAG CCGGGCACCC 850
TGGAGGTGGT GTCAGGCTGC ACGGTGGACA TGCTGCAGGC AGTGTATGGG 900
CTGGACGGCA TCCGCCTGCG CCGCCGCCAG TACTACACCA TGCCCAACTT 1000
CAGACAGTAT GAGAACCGCA CCGGCCACAT CCTGGTGCAG TGGTCGCTGG 1050
GCAGCCCCCT GCACTTCGCC GGCTGGCACT GCTCCTGGTG CTTACGCCC 1100
GAGGGCATCT ACTTCAAGCT CGTGTCCGCC CAGAATGGCG ACTTCCCACG 1150
CTGGGGTGAC TACGAGGACA AGCGGGACCT GAACTACATC CGCGGCCTGA 1200
TCCGCACCGG GGGCTGGTTC GACGGCACGC AGCAGGAGTA CCCGCCTGCA 1250
GACCCCAGCG AGCACATGTA TGCGCCCAAG TACCTGCTGA AGAACTACGA 1300
CCGGTTCCAC TACCTGCTGG ACAACCCCTA CCAGGAGCCC AGGAGCACGG 1350
CGGCGGGCGG GTGGCGCCAC AGGGGTCCCG AGGGAAGGCC GCGCGCCCGG 1400
GGCAAACCTG ACGAGGCGGA AGTCGAACAA AAATCATCT CAGAAGAGGA 1450
TCTGAATTAG 1460

```

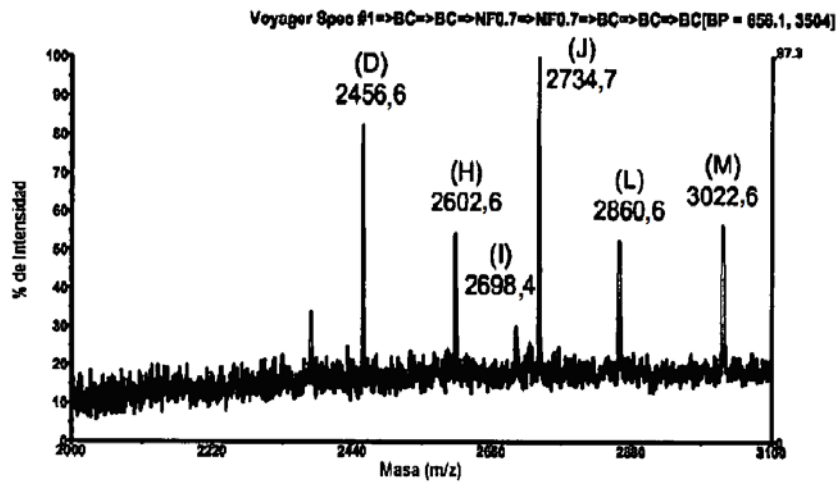
Figura 5J



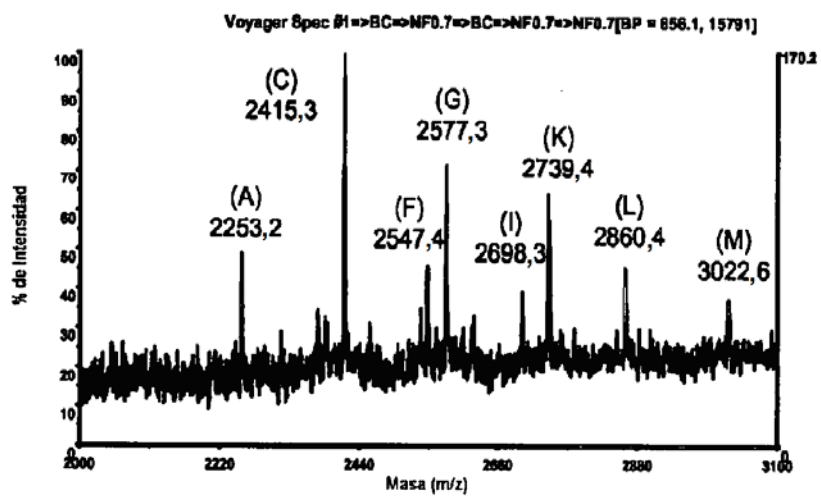
**Figura 6**

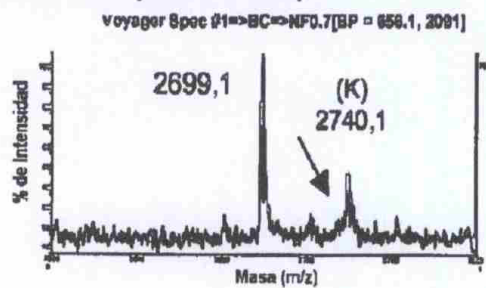


**Figura 7A**



**Figura 7B**



**Figura 7C (inserción)****Figura 7C**