

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 774 190**

51 Int. Cl.:

C12N 9/40 (2006.01)
A61K 38/47 (2006.01)
C12N 15/82 (2006.01)
A61P 3/00 (2006.01)
A61P 43/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **07.09.2011 E 17164713 (4)**

97 Fecha y número de publicación de la concesión europea: **18.12.2019 EP 3272861**

54 Título: **Composiciones de alfa-galactosidasa**

30 Prioridad:

20.01.2011 US 201161434499 P
20.01.2011 US 201161434503 P
02.03.2011 WO PCT/IL2011/000209

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
17.07.2020

73 Titular/es:

PROTALIX LTD. (100.0%)
2 Snunit Street Science Park
2010000 Carmiel, IL

72 Inventor/es:

SHULMAN, AVIDOR;
HANANIA, URI;
KIZHNER, TALI y
SHAALTIEL, YOSEPH

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 774 190 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

Composiciones de alfa-galactosidasa

5 Campo y antecedentes de la invención

La presente invención, en algunas realizaciones de la misma, se refiere a una construcción de expresión de ácido nucleico y, más en particular, pero no de manera exclusiva, a una construcción de ácido nucleico para la expresión de α -galactosidasa en células vegetales, células que expresan la construcción de ácido nucleico y la α -galactosidasa humana, y usos de las mismas.

La enfermedad de Fabry es una enfermedad de almacenamiento lisosómico ligada al cromosoma X provocada por actividad deficiente de la enzima lisosómica α -galactosidasa A (α -Gal A). Los pacientes con enfermedad de Fabry clásica tienen habitualmente actividad de α -Gal A de menos de 1 % y, con frecuencia, demuestran el espectro completo de síntomas, incluyendo dolor grave en las extremidades (acroparestesias), hipohidrosis, cambios corneales y lenticulares, lesiones cutáneas (angioqueratoma), insuficiencia renal, enfermedad cardiovascular, insuficiencia pulmonar, síntomas neurológicos e ictus. En la enfermedad de Fabry atípica, los individuos con actividad enzimática residual demuestran síntomas en el futuro y los síntomas se limitan habitualmente a uno o varios órganos. Las manifestaciones clínicas en portadoras varían en gran medida debido a la inactivación aleatoria del cromosoma X. Aunque los portadores habitualmente permanecen asintomáticos durante toda la vida, muchos demuestran síntomas clínicos tan variables y graves como los de los hombres afectados. De Duve sugirió en primer lugar que el reemplazo de la enzima lisosómica ausente con enzima biológicamente activa exógena podría ser un enfoque viable para el tratamiento de enfermedades de almacenamiento lisosómico [Fed Proc. 23: 1045 (1964)].

Se ha producido α -galactosidasa A recombinante para terapia de reemplazo enzimático en células de insecto (sf9) (véase documento US 7.011.831) en fibroblastos humanos (véase documento US 6.395.884) y en células vegetales (véase documentos US 6.846.968 y WO2008/132743). Se han realizado ensayos clínicos con α -Gal A recombinante (agalsidasa beta [Fabrazyme®]: Genzyme Corporation, Cambridge, Mass; agalsidasa alfa [Replagal®]: Shire Human Genetic Therapies Corporation, Cambridge, Mass) y ambos fármacos han sido aprobados para su uso clínico. Hasta 2009, se habían tratado más de 2000 pacientes con una de las dos formulaciones disponibles y, en general, se ha demostrado la seguridad y eficacia de ambas (Hoffmann, Orphanet J Rare Dis. 2009; 4:21).

La experiencia clínica ha indicado que las composiciones enzimáticas recombinantes disponibles difieren con respecto a la tolerancia del paciente, la inmunogenicidad, el régimen de dosificación y la eficacia clínica (Hoffmann, Orphanet J Rare Dis. 2009; 4:21). Otro inconveniente asociado con las enzimas recombinantes existentes es su costo, que puede suponer una carga económica pesada en los sistemas de salud. El alto costo de estas enzimas recombinantes, cuando se producen en cultivos celulares de mamíferos, biorreactores o células de insectos resulta de un protocolo complejo de purificación y modificación y las cantidades relativamente grandes del producto terapéutico necesarias para los tratamientos existentes. Existe, por lo tanto, una necesidad urgente de reducir el costo de la α -galactosidasa para que esta terapia que salva vidas pueda proporcionarse de manera más asequible a todos los que la necesiten.

Se pueden producir enzimas lisosómicas humanas en plantas transgénicas para resolver problemas de seguridad, infecciones víricas, reacciones inmunitarias, rendimiento de producción y costo. El documento US 2002/0088024 enseña la expresión en células vegetales de α -galactosidasa humana recombinante, usando un péptido señal de direccionamiento al RE de amilasa de arroz y una secuencia codificante de α -galactosidasa humana truncada en la parte C-terminal para expresión y secreción eficaces en el líquido intracelular.

El documento WO 97/10353 enseña la producción de enzimas lisosómicas en plantas, específicamente IDUA y glucocerebrosidasa, usando el promotor MeGA inducido por herida y el marcador FLAG para recuperación. El documento WO 97/10353 no proporciona orientación para construcciones específicas que expresan α -galactosidasa humana recombinante.

Cramer *et al.* (Curr. Topics Trans. 1999 Plants 95-118) revisa métodos de expresión de glucocerebrosidasa e IDUA en plantas de manera similar al documento WO 97/10353, que se encuentra en la endomembrana (IF) de las hojas de tabaco y describe la expresión en planta en general, pero no proporciona orientación para la expresión de α -galactosidasa recombinante humana.

El documento US2006-0204487 enseña métodos para la expresión de enzimas lisosómicas humanas en plantas, empleando diversas estrategias para direccionamiento a los polipéptidos recombinantes (específicamente, glucocerebrosidasa) para glucosilación y remodelación de glucano eficaces. No se desvela clonación ni expresión de α -galactosidasa humana recombinante a partir de construcciones específicas.

El documento WO2007/010533 enseña la administración de células vegetales que expresan moléculas bioactivas recombinantes, por ejemplo, enzimas lisosómicas tales como glucocerebrosidasa humana y α -galactosidasa humana. No se desvela la clonación y expresión de α -galactosidasa humana recombinante.

El documento WO2008/132743 describe la clonación de una secuencia codificante de α -galactosidasa humana en un vector de expresión en plantas y su expresión en células de tabaco, lo que da como resultado una enzima α -galactosidasa biológicamente activa, glucosilada, de longitud completa que demostró captación en fibroblastos. Sin embargo, se necesitan vectores de expresión de α -galactosidasa más avanzados para aumentar aún más la eficacia de la expresión y producción de proteína α -galactosidasa recombinante humana que tiene farmacocinética mejorada en condiciones clínicas.

Sumario de la invención

La invención es como se define en las reivindicaciones. Según un aspecto de algunas realizaciones de la presente invención, se proporciona una construcción de expresión de ácido nucleico, comprendiendo la construcción una secuencia de ácido nucleico que codifica una proteína α -galactosidasa humana, en donde la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo N a un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y en donde la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico.

Según un aspecto de algunas realizaciones de la presente invención, se proporciona una proteína α -galactosidasa humana que tiene un resto de glicina N-terminal, en donde la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo C a un péptido señal de retención en el retículo endoplásmico.

Según algunas realizaciones de la invención, la secuencia de ácido nucleico comprende la SEQ ID NO: 22.

Según algunas realizaciones de la invención, el péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* es como se expone en la SEQ ID NO: 4.

Según algunas realizaciones de la invención, la construcción de ácido nucleico comprende una secuencia de ácido nucleico que codifica el péptido señal de direccionamiento al retículo endoplásmico como se expone en la SEQ ID NO: 5.

Según algunas realizaciones de la invención, el péptido señal de retención en el retículo endoplásmico es como se expone en la SEQ ID NO: 6.

Según algunas realizaciones de la invención, la construcción de ácido nucleico comprende una secuencia de ácido nucleico que codifica el péptido de retención en el retículo endoplásmico como se expone en la SEQ ID NO: 19.

Según algunas realizaciones de la invención, la secuencia de ácido nucleico tiene una secuencia como se expone en la SEQ ID NO: 2

Según algunas realizaciones de la invención, la proteína α -galactosidasa humana tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 1.

Según algunas realizaciones de la invención, el uso codónico de la secuencia de ácido nucleico está optimizado para *Nicotinia tabaccum*.

Según algunos aspectos de algunas realizaciones de la invención, se proporciona una célula aislada que comprende la construcción de ácido nucleico de la invención.

Según algunas realizaciones de la invención, la célula produce de manera recombinante la enzima α -galactosidasa humana.

Según algunas realizaciones de la invención, la α -galactosidasa humana comprende un resto de glicina N-terminal.

La célula es una célula vegetal.

Según algunas realizaciones de la invención, la célula vegetal es una célula de tabaco.

Según algunas realizaciones de la invención, la célula de tabaco es una célula BY-2.

Según algunas realizaciones de la invención, la proteína α -galactosidasa humana se produce de manera recombinante para tener al menos un resto de manosa expuesto.

Según algunas realizaciones de la invención, la proteína α -galactosidasa humana se produce de manera recombinante para tener al menos una β -(1,2) xilosa de núcleo o al menos una α -(1,3) fucosa de núcleo o ambas.

Según algunas realizaciones de la invención, la construcción de ácido nucleico está integrada de manera estable en el genoma de la célula.

Según algunas realizaciones de la invención, la célula es una célula de *Agrobacterium tumefaciens*.

5 Según un aspecto de algunas realizaciones de la presente invención, se proporciona un método para producir una proteína α -galactosidasa humana recombinante, que comprende: proporcionar una célula según la invención; y cultivar la célula para producir la proteína α -galactosidasa humana recombinante; y aislar la proteína α -galactosidasa humana recombinante de la célula.

10 Según algunas realizaciones de la invención, la célula es una célula aislada cultivada en un medio de cultivo celular.

Según algunas realizaciones de la invención, el cultivo se ve afectado en un biorreactor desechable.

15 Según algunas realizaciones de la invención, la proteína α -galactosidasa humana tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 16.

Según un aspecto de algunas realizaciones de la presente invención, se proporciona una composición farmacéutica que comprende, como principio activo, la proteína α -galactosidasa humana de la presente invención y un vehículo farmacéuticamente aceptable.

20 Según un aspecto de algunas realizaciones de la presente invención, se proporciona una composición farmacéutica que comprende, como principio activo, la célula de la presente invención y un vehículo farmacéuticamente aceptable.

25 Según un aspecto de algunas realizaciones de la presente invención, se proporciona una composición farmacéutica que comprende, como principio activo, una población de proteínas α -galactosidasa humanas de la invención, en donde las estructuras de glucano predominantes de dicha población de proteínas α -galactosidasas humanas son manosa 4- β -(1,2) xilosa (M4X); manosa 3- β -(1,2) xilosa- α -(1,3) fucosa [Fc(3)M3X] y manosa 8 (M8) y un vehículo farmacéuticamente aceptable.

30 Según un aspecto de algunas realizaciones de la presente invención, se proporciona un método para tratar enfermedad de Fabry en un sujeto que lo necesite, que comprende administrar al sujeto la composición farmacéutica de la invención.

35 Según un aspecto de algunas realizaciones de la presente invención, se proporciona una composición farmacéutica que comprende la proteína α -galactosidasa humana de la invención, que tiene una estructura de glucano que comprende nueve restos de manosa, en donde tres son restos de manosa expuestos, y un vehículo farmacéuticamente aceptable. La proteína α -galactosidasa humana que tiene una estructura de glucano que comprende nueve restos de manosa, en donde tres son restos de manosa expuestos, puede ser 0,5 %, 0,8 %, 1,0 %, 1,3 %, 2 % o más de la población de proteínas α -galactosidasas humanas de la composición.

40 Según un aspecto de algunas realizaciones de la presente invención, se proporciona un método para tratar enfermedad de Fabry en un sujeto que lo necesite, que comprende administrar al sujeto la composición farmacéutica que comprende la célula de la invención.

45 Según un aspecto de algunas realizaciones de la presente invención, se proporciona el uso de la proteína α -galactosidasa humana de la invención para la fabricación de un medicamento para el tratamiento de la enfermedad de Fabry en un sujeto que lo necesite.

50 A menos que se definan de otra manera, todos los términos técnicos y/o científicos usados en el presente documento tienen el mismo significado que el que entiende habitualmente un experto habitual en la materia a la que pertenece la invención. Aunque pueden usarse métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o el ensayo de realizaciones de la invención, a continuación, se describen métodos y/o materiales ilustrativos. En caso de conflicto, prevalecerá la memoria descriptiva de patente, incluyendo las definiciones. Además, los materiales, métodos y ejemplos son solamente ilustrativos y no se pretende que sean necesariamente limitantes.

55 Breve descripción de los dibujos

60 Algunas realizaciones de la invención se describen en el presente documento, únicamente a modo de ejemplo, haciendo referencia a los dibujos adjuntos. Haciendo referencia específica ahora a los dibujos en detalle, se subraya que los detalles mostrados son a modo de ejemplo y para fines de análisis ilustrativo de realizaciones de la invención. En este sentido, la descripción tomada con los dibujos hace evidente para los expertos en la materia cómo se pueden practicar realizaciones de la invención.

65 En los dibujos:

la FIG. 1 es un alineamiento de la secuencia de aminoácidos de la α -galactosidasa codificada por el ácido nucleico

de la invención (SEQ ID NO: 1, secuencia inferior) y la proteína α -galactosidasa humana nativa (GenBank: X05790, secuencia superior), incluyendo el péptido líder señal nativo (SEQ ID NO: 3, resaltado en rojo). El péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* (SEQ ID NO: 4) está resaltado en verde. El péptido señal de retención en el retículo endoplásmico SEKDEL (SEQ ID NO: 6) se ubica en el extremo C de la proteína expresada en plantas. La secuencia de proteína de enzima α -galactosidasa nativa madura (SEQ ID NO: 7) se resalta en amarillo.

La FIG. 2 es una exploración de un gel de SDS-PAGE que ilustra los niveles de expresión de diversas construcciones de α -galactosidasa humana en plantas. Se separaron extractos de hojas de plantas *N. benthamiana* infiltradas con diferentes construcciones de α -galactosidasa humana en un gel reducido de SDS-PAGE al 12 % y se tiñeron con azul de Coomassie para visualización de la proteína. Carril 1 = productos de expresión de una construcción (SEQ ID NO: 8) que codifica una α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (péptido señal de pectinasa de manzana = SEQ ID NO: 9). Carril 2 = productos de expresión de la SEQ ID NO: 2, que codifica una α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de ABPI (péptido señal de ABPI = SEQ ID NO: 4). Carril 3 = productos de expresión de la SEQ ID NO: 11, que codifica la SEQ ID NO: 12 con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (SEQ ID NO: 9). Carril 4 = control negativo (sin secuencia codificante de α -galactosidasa). La banda prominente en el carril 2 (flecha) corresponde al tamaño del monómero de α -galactosidasa recombinante maduro (SEQ ID NO: 21). La banda teñida con Coomassie indica proteína α -galactosidasa abundante;

la FIG. 3 es un histograma que demuestra los niveles de actividad catalítica en extractos de plantas que expresan diversas construcciones de α -galactosidasa humana. Se infiltraron hojas de plantas *N. benthamiana* con diferentes construcciones de α -galactosidasa humana de la siguiente manera: A = SEQ ID NO: 8, que codifica una proteína α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (péptido señal de pectinasa de manzana = SEQ ID NO: 9). B = SEQ ID NO: 2, que codifica una proteína α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de ABPI (péptido señal de ABPI = SEQ ID NO: 4). C = SEQ ID NO: 13, que codifica una proteína α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de endoquitinasa B (proteína señal de endoquitinasa = SEQ ID NO: 14). D = SEQ ID NO: 11, que codifica una proteína α -galactosidasa (SEQ ID NO: 12) con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (péptido señal de pectinasa de manzana = SEQ ID NO: 9). Se extrajeron hojas de las plantas respectivas en tampón de ensayo de actividad y se determinó la actividad catalítica usando p-nitrofenilalfa-D-galactopiranosido (pNP-alfa-D-Gal, GBB1290, IRIS Biotech, Alemania) como sustrato de hidrólisis. Se supervisó la generación del cromóforo p-nitrofenolato a 405 nm;

la FIG. 4 es un gráfico que compara la estabilidad de dos α -galactosidasas humanas recombinantes de plantas en plasma humano (37 °C) medida por la actividad catalítica residual, α -galactosidasa de hojas de plantas *N. benthamiana* transformadas con la SEQ ID NO: 11, que codifica la SEQ ID NO: 12 con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (cuadrados abiertos) y α -galactosidasa de hojas de plantas *N. benthamiana* transformadas con la SEQ ID NO: 2, que codifica una α -galactosidasa (SEQ ID NO: 1) que tiene un péptido señal de direccionamiento al retículo endoplásmico de ABPI (cuadrados cerrados) purificadas por rotura de la planta, precipitación por adición de sales y cromatografía, e incubadas en plasma humano a 37 °C durante hasta 1 hora, para simular la farmacocinética *in vivo*. Las muestras se ensayaron para determinar la actividad catalítica de α -galactosidasa mediante el ensayo PNP-G a los intervalos indicados. Obsérvese la mayor estabilidad de α -galactosidasa de las hojas de plantas *N. benthamiana* que expresan SEQ ID NO: 1 (cuadrados cerrados);

la FIG. 5 es un gráfico que compara la estabilidad de la actividad catalítica de la α -galactosidasa expresada en células vegetales (cuadrados cerrados) con la de la α -galactosidasa recombinante humana comercial expresada en cultivo de células humanas (Replagal®, triángulos abiertos). Se incubó extracto no purificado de células BY2 que expresaban la SEQ ID NO: 2, que codificaba una α -galactosidasa (SEQ ID NO: 1) que tenía un péptido señal de direccionamiento al retículo endoplásmico de ABPI (cuadrados cerrados) y α -galactosidasa recombinante humana expresada en cultivo de células de mamífero (Replagal®, triángulos abiertos), en plasma a 37 °C durante hasta 1 hora, para simular la farmacocinética *in vivo*. Las muestras se ensayaron para determinar la actividad catalítica de α -galactosidasa mediante el ensayo PNP-G a los intervalos indicados. Obsérvese la estabilidad similar de las enzimas expresadas en células vegetales y expresadas en células humanas;

la FIG. 6 es un gráfico que compara la estabilidad de la actividad catalítica de la α -galactosidasa expresada en plantas (cuadrados cerrados) con la de la α -galactosidasa recombinante humana comercial expresada en cultivo de células humanas (Replagal®, triángulos abiertos). Se incubó extracto no purificado de células BY2 que expresaban la SEQ ID NO: 2, que codificaba una α -galactosidasa (SEQ ID NO: 1) que tenía un péptido señal de direccionamiento al retículo endoplásmico de ABPI (cuadrados cerrados) y α -galactosidasa recombinante humana expresada en cultivo de células de mamífero (Replagal®, triángulos abiertos), en condiciones que simulaban el ambiente lisosómico [ácido (pH 4,6)] a 37 °C durante hasta 11 días. Las muestras se ensayaron para determinar la actividad catalítica de α -galactosidasa mediante el ensayo PNP-G a los intervalos indicados, cuando se incubaron. Obsérvese la estabilidad similar de las enzimas expresadas en plantas y expresadas en mamíferos durante hasta 9 días;

la FIG. 7 es un cromograma que muestra la distribución de estructuras de glucanos en los productos de hidrolizados por glucosilasa de α -galactosidasa expresada en células vegetales. Las muestras de extracto de células BY2 que expresaban SEQ ID NO: 2, que codificaba una α -galactosidasa (SEQ ID NO: 1) que tenía un péptido señal de direccionamiento al retículo endoplásmico de ABPI se redujeron, se alquilaron y se separaron en SDS-PAGE. Las

bandas de proteínas se tomaron para el análisis de glucanos mediante digestión por tripsina seguida de digestión por PNGasa A y PNGasa F. Los glucanos libres resultantes se extrajeron, se purificaron, se marcaron con el reactivo fluorescente antranilamida (2AB), se separaron usando una HPLC de fase normal TSKgel Amide 80 y se detectaron usando un detector de fluorescencia. Obsérvense los picos mayores en 5,9 y 8,9 UG (84,55 minutos y 108,05 minutos) y un pico menor de M9 en 112,35;

la FIG. 8 es una tabla que muestra el perfil de glucanos de la α -galactosidasa expresada en células vegetales, expresado como el área porcentual de los glucanos individuales representados en el cromograma de la figura 7. Obsérvense los perfiles idénticos de cada análisis independiente (A y B).

10 Descripción de realizaciones específicas de la invención

La presente invención, en algunas realizaciones de la misma, se refiere a un vector de expresión para la expresión recombinante de α -galactosidasa humana en células vegetales, métodos para su expresión de la misma en células vegetales y la α -galactosidasa humana producida de este modo.

15 Debe entenderse que la invención no se limita necesariamente en su aplicación a los detalles expuestos en la siguiente descripción o ilustrados por los ejemplos. La invención tiene capacidad para otras realizaciones o puede practicarse o llevarse a cabo de diversas maneras.

20 Los presentes inventores han construido un vector de expresión para la expresión recombinante de α -galactosidasa humana en células vegetales, plantas de tabaco transformadas y células vegetales con el vector y han aislado α -galactosidasa humana catalíticamente activa de las plantas y los cultivos celulares. La α -galactosidasa humana recombinante expresada conservó actividad catalítica favorable en condiciones que simulan las encontradas normalmente en administración *in vivo*, lo que sugiere farmacocinética simulada mejorada en comparación con otra α -galactosidasa humana recombinante de origen vegetal y similar a la de α -galactosidasa humana recombinante expresada en mamíferos (Replagal®).

30 Por tanto, según un aspecto de la presente invención, se proporciona una construcción de expresión de ácido nucleico, comprendiendo la construcción una secuencia de ácido nucleico que codifica una proteína α -galactosidasa humana, en donde la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo N a un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y en donde la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico.

35 Como se usa en el presente documento, la expresión "proteína α -galactosidasa humana" se refiere a un polipéptido de α -galactosidasa A humana (α -D-galactósido galactohidrolasa; EC 3.2.1.22; α -Gal A). Según algunas realizaciones de la invención, la proteína α -galactosidasa humana es la proteína α -galactosidasa humana madura, que tiene la secuencia de aminoácidos como se expone en la SEQ ID NO: 7 (α -Gal humano maduro). Se apreciará que la proteína α -galactosidasa humana puede ser una proteína α -galactosidasa humana modificada, que tiene una secuencia de aminoácidos diferente a la de la SEQ ID NO: 7. Un ejemplo no limitante de una proteína α -galactosidasa humana modificada codificada por el vector de expresión de la invención es la SEQ ID NO: 16 (α -Gal). En otro ejemplo no limitante, la proteína α -galactosidasa humana codificada por el vector de expresión puede ser una mezcla de proteínas α -galactosidasas humanas que comprenden tanto proteínas α -galactosidasas que tienen un resto de glicina N-terminal (G), tales como SEQ ID NO: 16 o SEQ ID NO: 21, como sin un resto de glicina N-terminal, tales como, pero sin limitación, SEQ ID NO: 7 y SEQ ID NO: 20.

45 Según otro aspecto de la invención, la proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y la proteína α -galactosidasa humana también se fusiona traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico.

50 Como se usa en el presente documento, la expresión "fusionado traduccionalmente en el extremo N" o "fusionado traduccionalmente en el extremo C" se refiere a la unión covalente del péptido indicado mediante un enlace peptídico con el aminoácido N-terminal o C-terminal de la proteína α -galactosidasa humana madura normalmente como resultado de expresión recombinante.

55 Como se usa en el presente documento, la expresión "péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis*" se refiere a la secuencia de péptido líder de la proteína de unión a auxina de *Arabidopsis thaliana* (n.º de referencia de Uni-Prot P33487), que es capaz de dirigir la proteína expresada al retículo endoplásmico dentro de la célula vegetal. En una realización, el péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* es un polipéptido de 33 aminoácidos como se expone en la SEQ ID NO: 4.

60 Como se usa en el presente documento, la expresión "péptido señal de retención en el retículo endoplásmico" se refiere a una secuencia peptídica que, cuando está presente en el extremo N o C de un polipéptido, provoca que el polipéptido se recupere del aparato de Golgi y se retenga en el retículo endoplásmico (véase Rayon *et al.* Journal of Experimental Botany, Vol. 49, n.º 326, págs. 1463-1472, 1998; y Neumann, *et al.* Annals of Botany, 2003; 92:167-180). En una realización, el péptido señal de retención en el retículo endoplásmico es KDEL (SEQ ID NO: 17) o SEKDEL

(SEQ ID NO: 6). Otras señales adecuadas de retículo endoplásmico de plantas son conocidas en la técnica (véase Pagny *et al*, Journal of Experimental Botany, Vol. 50, págs. 157-64).

Por tanto, según una realización de la presente invención, la proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico ABPI de *Arabidopsis* y la proteína α -galactosidasa humana y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 1.

Como se usa en el presente documento, la expresión "secuencia de ácido nucleico" se refiere a una secuencia de ácido nucleico monocatenaria o bicatenaria que se aísla y se proporciona en forma de una secuencia de ARN, una secuencia polinucleotídica complementaria (ADNc), una secuencia polinucleotídica genómica y/o una secuencia polinucleotídica compuesta (p. ej., una combinación de las anteriores).

Como se usa en el presente documento, la expresión "secuencia polinucleotídica complementaria" se refiere a una secuencia, que resulta de la transcripción inversa de ARN mensajero usando una transcriptasa inversa o cualquier otra ADN polimerasa dependiente de ARN. Dicha secuencia se puede amplificar posteriormente *in vivo* o *in vitro* usando una ADN polimerasa dependiente de ADN.

Como se usa en el presente documento, la expresión "secuencia polinucleotídica genómica" se refiere a una secuencia procedente (aislada) de un cromosoma y, por tanto, representa una parte contigua de un cromosoma.

Como se usa en el presente documento, la expresión "secuencia polinucleotídica compuesta" se refiere a una secuencia, que es al menos parcialmente complementaria y al menos parcialmente genómica. Una secuencia compuesta puede incluir algunas secuencias exonales necesarias para codificar el polipéptido de la presente invención, así como algunas secuencias intrónicas interpuestas entre medias. Las secuencias intrónicas pueden ser de cualquier fuente, incluyendo de otros genes, y normalmente incluirán secuencias señal de corte y empalme conservadas. Dichas secuencias intrónicas pueden incluir además elementos reguladores de la expresión que actúan en cis.

Según algunas realizaciones de la presente invención, las secuencias de ácido nucleico que codifican los polipéptidos de la presente invención están optimizadas para su expresión. Los ejemplos de dichas modificaciones de secuencias incluyen, pero sin limitación, un contenido de G/C alterado para acercarse más estrechamente al hallado normalmente en la especie vegetal de interés y la retirada de codones encontrados de manera poco habitual en la especie vegetal denominada habitualmente optimización de codones. En una realización, el uso de codones de la secuencia de ácido nucleico que codifica la proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico ABPI de *Arabidopsis* y la proteína α -galactosidasa humana y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico está optimizado para *Nicotiana tabacum* o *Nicotiana benthamiana*.

La expresión "optimización de codones" se refiere a la selección de nucleótidos de ADN adecuados para su uso dentro de un gen estructural o fragmento del mismo que aborda el uso codónico dentro de la planta de interés. Por lo tanto, un gen o secuencia de ácido nucleico optimizado se refiere a un gen en el que la secuencia de nucleótidos de un gen nativo o de origen natural se ha modificado para utilizar codones estadísticamente preferidos o estadísticamente favorecidos dentro de la planta. La secuencia de nucleótidos normalmente se examina al nivel de ADN y la región codificante optimizada para expresión en la especie de planta determinada usando cualquier procedimiento adecuado, por ejemplo, como se describe en Sardana *et al*. (1996, Plant Cell Reports 15:677-681). En este método, la desviación típica del uso codónico, una medida del sesgo de uso codónico, se puede calcular encontrando primero la desviación proporcional cuadrática del uso de cada codón del gen nativo en relación con el de genes vegetales con alta expresión, seguido de un cálculo de la desviación cuadrática promedio. La fórmula usada es: $1 \text{ DTUC} = n = 1 N[(X_n - Y_n)/Y_n]^2/N$, donde X_n se refiere a la frecuencia de uso del codón n en genes vegetales con alta expresión, en donde Y_n se refiere a la frecuencia de uso del codón n en el gen de interés y N se refiere al número total de codones en el gen de interés. Una tabla de uso codónico de genes con alta expresión de plantas dicotiledóneas se compila utilizando los datos de Murray *et al*. (1989, Nuc Acids Res. 17:477-498).

Un método de optimización de la secuencia de ácido nucleico de acuerdo con el uso codónico preferido para un tipo de célula vegetal particular se basa en el uso directo, sin realizar ningún cálculo estadístico adicional, de tablas de optimización de codones tales como las proporcionadas en línea en la base de datos de uso codónico a través del banco de ADN NIAS (Instituto Nacional de Ciencias Agrobiológicas) en Japón (protocolo de transferencia de hipertexto://World Wide Web (punto) kazusa (punto) or (punto) jp/codon/). La base de datos de uso codónico contiene tablas de uso codónico para varias especies diferentes, habiéndose determinado cada tabla de uso codónico estadísticamente basándose en los datos presentes en Genbank.

Usando las tablas anteriores para determinar los codones más preferidos o más favorecidos para cada aminoácido en una especie particular (por ejemplo, arroz), una secuencia de nucleótidos de origen natural que codifica una proteína de interés se puede someter a optimización de codones para esa especie vegetal particular. Esto se efectúa reemplazando codones que pueden tener una baja incidencia estadística en el genoma de la especie particular con

5 los codones correspondientes, con respecto a un aminoácido, que están estadísticamente más favorecidos. Sin embargo, se pueden seleccionar uno o más codones menos favorecidos para suprimir sitios de restricción existentes, para crear nuevos en uniones potencialmente útiles (extremos 5' y 3' para añadir péptido señal o casetes de terminación, sitios internos que se podrían usar para cortar y empalmar segmentos entre sí para producir una secuencia de longitud completa correcta) o para eliminar secuencias de nucleótidos que pueden afectar negativamente a la estabilidad o expresión del ARNm.

10 La secuencia de nucleótidos codificante deseada puede contener ya, con antelación a cualquier modificación, varios codones que corresponden a un codón estadísticamente favorecido en una especie vegetal particular. Por lo tanto, la optimización de codones de la secuencia de nucleótidos nativa puede comprender determinar qué codones, dentro de la secuencia de nucleótidos deseada, no están estadísticamente favorecidos con respecto a una planta particular, y modificar estos codones de acuerdo con una tabla de uso codónico de la planta particular para producir un derivado de codones optimizados. Una secuencia de nucleótidos modificada se puede optimizar completa o parcialmente para uso codónico de plantas siempre que la proteína codificada por la secuencia de nucleótidos modificada se produzca a un nivel mayor que la proteína codificada por el gen de origen natural o nativo correspondiente. La construcción de genes sintéticos alterando el uso codónico se describe, por ejemplo, en la Solicitud de Patente PCT 93/07278.

20 Según algunas realizaciones de la invención, la proteína α -galactosidasa humana está codificada por una secuencia de ácido nucleico como se expone en la SEQ ID NO: 18. Según realizaciones adicionales de la invención, el péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* está codificado por una secuencia de ácido nucleico como se expone en la SEQ ID NO: 5. Según más realizaciones adicionales de la invención, el péptido señal de retención en el retículo endoplásmico está codificado por una secuencia de ácido nucleico como se expone en la SEQ ID NO: 19. Según otras realizaciones adicionales de la invención, la proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y la proteína α -galactosidasa humana y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico está codificada por una secuencia de ácido nucleico como se expone en la SEQ ID NO: 2.

30 Como se usa en el presente documento, la expresión "construcción de expresión de ácido nucleico" se refiere a una construcción de ácido nucleico que incluye el ácido nucleico de algunas realizaciones de la invención (p. ej., SEQ ID NO: 2) y al menos un promotor para dirigir la transcripción de ácido nucleico en un célula de planta hospedadora. A continuación en el presente documento se proporcionan detalles adicionales de enfoques de transformación adecuados.

35 Según algunas realizaciones de la invención, se proporciona una construcción de expresión de ácido nucleico que comprende la secuencia de ácido nucleico de la invención y un promotor para dirigir la transcripción de la secuencia de ácido nucleico en una célula hospedadora vegetal.

40 Según algunas realizaciones de la invención, el ácido nucleico está unido operativamente con la secuencia promotora. Una secuencia de ácido nucleico codificante está "unida operativamente" a una secuencia reguladora (p. ej., promotor) si la secuencia reguladora es capaz de ejercer un efecto regulador sobre la secuencia codificante unida a la misma.

45 La construcción de ácido nucleico de la presente invención puede usar cualquier secuencia promotora adecuada. Preferentemente, el promotor es un promotor constitutivo, un promotor específico de tejido o uno inducible.

50 Como se usa en el presente documento, la expresión "expresable en plantas" se refiere a una secuencia promotora, incluyendo cualquier elemento regulador adicional añadido a la misma o contenido en la misma, es al menos capaz de inducir, conferir, activar o potenciar la expresión en una célula, tejido u órgano vegetal, preferentemente una célula, tejido u órgano vegetal monocotiledóneo o dicotiledóneo. Dicho promotor puede ser constitutivo, es decir, capaz de dirigir alto nivel de expresión génica en una pluralidad de tejidos, específico de tejido, es decir, capaz de dirigir la expresión génica en un tejido o tejidos particulares, inducible, es decir, capaz de dirigir la expresión génica bajo un estímulo, o quimérico, es decir, formado por partes de al menos dos promotores diferentes.

55 Se presentan ejemplos de promotores preferidos útiles para los métodos de algunas realizaciones de la invención en la tabla I, II, III y IV.

ES 2 774 190 T3

Tabla I

<i>Promotores constitutivos ilustrativos para su uso en la ejecución de algunas realizaciones de la invención</i>		
<i>Fuente del gen</i>	<i>Patrón de expresión</i>	<i>Referencia</i>
Actina	constitutivo	McElroy <i>et al</i> , Plant Cell, 2: 163-171, 1990
CAMV 35S	constitutivo	Odell <i>et al</i> , Nature, 313: 810-812, 1985
CaMV 19S	constitutivo	Nilsson <i>et al.</i> , Physiol. Plant 100:456-462, 1997
G0S2	constitutivo	de Pater <i>et al</i> , Plant J Nov; 2(6):837-44, 1992
ubiquitina	constitutivo	Christensen <i>et al</i> , Plant Mol. Biol. 18: 675-689, 1992
Ciclofilina de arroz	constitutivo	Bucholz <i>et al</i> , Planta Mol Biol. 25(5):837-43, 1994
Histona H3 de maíz	constitutivo	Lepetit <i>et al</i> , Mol. Gen. Genet. 231:276-285,1992
Actina 2	constitutivo	An <i>et al</i> , Plant J. 10(1); 107-121, 1996

Tabla II

<i>Promotores preferidos en semillas ilustrativos para su uso en la ejecución de algunas realizaciones de la invención</i>		
<i>Fuente del gen</i>	<i>Patrón de expresión</i>	<i>Referencia</i>
Genes específicos de semillas	semilla	Simon, <i>et al.</i> , Plant Mol. Biol. 5. 191, 1985; Scofield, <i>et al.</i> , J. Biol. Chem. 262: 12202, 1987.; Baszczynski, <i>et al.</i> , Plant Mol. Biol. 14: 633, 1990.
Albúmina de nuez de Brasil	semilla	Pearson' <i>et al.</i> , Plant Mol. Biol. 18: 235- 245, 1992.
legumbre	semilla	Ellis, <i>et al</i> . Plant Mol. Biol. 10: 203-214, 1988
Glutelina (arroz)	semilla	Takaiwa, <i>et al.</i> , Mol. Gen. Genet. 208:15-22, 1986; Takaiwa, <i>et al.</i> , FEBS Letts. 221: 43-47, 1987
Zeína	semilla	Matzke <i>et al</i> Plant Mol Biol, 143.323-32 1990
napA	semilla	Stalberg, <i>et al</i> , Planta 199: 515-519, 1996
LMW y HMW de trigo, glutenina-1	endospermo	Mol Gen Genet 216:81-90, 1989; NAR 17:461-2,
SPA de trigo	semilla	Albani <i>et al</i> , Plant Cell, 9: 171- 184, 1997
gliadinas a, b y g de trigo	endospermo	EMB03:1409-15, 1984
Promotor Itrl de cebada	endospermo	
hordeína B1, C, D de cebada	endospermo	TheorAppl Gen 98:1253-62, 1999; Plant J 4:343-55, 1993; Mol Gen Genet 250:750- 60, 1996
DOF de cebada	endospermo	Mena <i>et al</i> , The Plant Journal, 116(1): 53-62, 1998
Biz2	endospermo	Documento EP99106056.7
Promotor sintético	endospermo	Vicente-Carbajosa <i>et al.</i> , Plant J. 13: 629-640, 1998
prolamina de arroz NRP33	endospermo	Wu <i>et al</i> , Plant Cell Physiology 39(8) 885-889, 1998
Globulina de arroz Glb-1	endospermo	Wu <i>et al</i> , Plant Cell Physiology 39(8) 885-889, 1998
OSH1 de arroz	embrión	Sato <i>et al</i> , Proc. Nati. Acad. Sci. USA, 93: 8117-8122
alfa-globulina de arroz REB/OHP-1	endospermo	Nakase <i>et al</i> . Plant Mol. Biol. 33: 513-S22, 1997
ADP-glucosa de arroz PP	endospermo	Trans Res 6:157-68, 1997
familia de genes ESR de maíz	endospermo	Plant J 12:235-46, 1997
gammakafirina de sorgo	endospermo	PMB 32:1029-35, 1996
KNOX	embrión	Postma-Haarsma <i>et al</i> , Plant Mol. Biol. 39:257-71, 1999
oleosina de arroz	Embrión y aleutón	Wu <i>et al</i> , J. Biochem., 123:386, 1998
oleosina de girasol	Semilla (embrión y semilla seca)	Cummins, <i>et al.</i> , Plant Mol. Biol. 19: 873-876, 1992

Tabla III

<i>Promotores específicos de flor ilustrativos para su uso en la ejecución de la invención</i>		
<i>Fuente del gen</i>	<i>Patrón de expresión</i>	<i>Referencia</i>
AtPRP4	flores	www.salus.medium.edu/m mg/tierney/html
chalcona sintasa (chsA)	flores	Van der Meer, <i>et al.</i> , Plant Mol. Biol. 15, 95-109, 1990.
LAT52	antera	Twel <i>et al</i> Mol. Gen Genet. 217:240-245 (1989)
apetala-3	flores	

Tabla IV

<i>Promotores alternativos de arroz para su uso en la ejecución de la invención</i>		
<i>PRO n.º</i>	<i>gen</i>	<i>expresión</i>
PR00001	Metalotioneína Mte	capa de transferencia de embrión + callos
PR00005	posible beta-amilasa	capa de transferencia de embrión
PR00009	Posible celulosa sintasa	Débil en las raíces
PR00012	lipasa (posible)	
PR00014	Transferasa (posible)	
PR00016	peptidil prolil cis-trans isomerasa (posible)	
R00019	desconocido	
PR00020	proteína prp (posible)	
PR00029	nodulina (posible)	
PR00058	Inhibidor de proteinasa Rgpi9	semilla
PR00061	beta expansina EXPB9	Débil en flores jóvenes
PR00063	Proteína estructural	tejidos jóvenes + callos + embrión
PR00069	xilosidasa (posible)	
PR00075	Prolamina 10 Kda	fuerte en endospermo
PR00076	alérgeno RA2	fuerte en endospermo
PR00077	prolamina RP7	fuerte en endospermo
PR00078	CBP80	
PR00079	enzima de ramificación de almidón I	
PR00080	ML2 de tipo metalotioneína	capa de transferencia de embrión + callos
PR00081	posible cafeoil-CoA 3-0 metiltransferasa	brote
PR00087	prolamina RM9	fuerte en endospermo
PR00090	prolamina RP6	fuerte en endospermo
PR00091	prolamina RP5	fuerte en endospermo
PR00092	alérgeno RA5	
PR00095	posible metionina aminopeptidasa	embrión
PR00098	proteína de unión a GTP relacionada con ras	
PR00104	beta expansina EXPB1	
PR00105	Proteína rica en glicina	
PR00108	proteína de tipo metalotioneína (posible)	
PR00110	Raíz fuerte RCc3	
PR00111	proteína de tipo uclacianina 3	centro de discriminación débil/meristemo de brote
PR00116	subunidad no ATPasa 11 de partículas reguladoras de proteasoma 26S	específico de meristemo muy débil
PR00117	posible proteína ribosómica 40S	débil en el endospermo
PR00122	precursor de la proteína de unión a clorofila a/lo (Cab27)	muy débil en brote
PR00123	posible protoclorofilida reductasa	Hojas fuertes

(continuación)

<i>Promotores alternativos de arroz para su uso en la ejecución de la invención</i>		
<i>PRO n.º</i>	<i>gen</i>	<i>expresión</i>
PR00126	metalotioneína RiCMT	centro de discriminación fuerte meristemo de brote
PR00129	GOS2	Constitutivo fuerte
PR00131	GOS9	
PR00133	quitinasa Cht-3	específico de meristemo muy débil
PR00135	alfa-globulina	Fuerte en endospermo
PR00136	alanina aminotransferasa	Débil en el endospermo
PR00138	Ciclina A2	
PR00139	Ciclina D2	
PR00140	Ciclina D3	
PR00141	Ciclofilina 2	Brote y semilla
PR00146	sacarosa sintasa SS1 (cebada)	constitutivo medio
PR00147	inhibidor de tripsina ITR1 (cebada)	débil en el endospermo
PR00149	ubiquitina 2 con intrón	constitutivo fuerte
PR00151	WS118	Embrión y tensión
PR00156	homólogo de HVA22 (posible)	
PR00157	EL2	
PR00169	acuaporina	constitutivo medio en plantas jóvenes
PR00170	Proteína de grupo de alta movilidad	Constitutivo fuerte
PR00171	proteína glucosilada de manera reversible RGP1	constitutivo débil
PR00173	MDH citosólico	brote
PR00175	RAB21	Embrión y tensión
PR00176	CDPK7	
PR00177	Cdc2-1	muy débil en meristemo
PR00197	sacarosa sintasa 3	
PRO0198	OsVP1	
PR00200	OSH1	muy débil en meristemo de plantas jóvenes
PR00208	posible clorofilasa	
PRO0210	OsNRT1	
PRO0211	EXP3	
PR00216	transportador de fosfato OjPT1	
PRO0218	oleosina 18 kd	aleurona + embrión
PRO0219	ubiquitina 2 sin intrón	
PR00220	RFL	
PRO0221	intrón delta de UBI de maíz	no detectado
PR00223	glutelina-1	
PRO0224	fragmento de promotor de prolamina RP6	
PRO0225	4xABRE	
PRO0226	glutelina OSGLUA3	
PRO0227	BLZ-2_corto (cebada)	
PRO0228	BLZ-2_largo (cebada)	

Según una realización específica, el promotor utilizado por la presente invención es un promotor constitutivo fuerte de modo que la sobreexpresión de los insertos de construcción se efectúa después de la transformación. En determinadas realizaciones, el promotor es el promotor de actina 2 de *Arabidopsis* (ACT2) (n.º de referencia: U41998, véase tabla I anterior).

5

La construcción de ácido nucleico de algunas realizaciones de la invención puede incluir además un marcador seleccionable adecuado y/o un origen de replicación. Según algunas realizaciones de la invención, la construcción de

ácido nucleico utilizada es un vector lanzadera, que puede tanto propagarse en *E. coli* (en donde la construcción comprende un marcador seleccionable adecuado y origen de replicación) como ser compatible con la propagación en células. La construcción según la presente invención puede ser, por ejemplo, un plásmido, un bácido, un fagémido, un cósmido, un fago, un virus o un cromosoma artificial.

La construcción de ácido nucleico de algunas realizaciones de la invención se puede utilizar para transformar células vegetales de manera estable o transitoria. En la transformación estable, la molécula de ácido nucleico se integra en el genoma vegetal y como tal representa un rasgo estable y hereditario. En la transformación transitoria, el polinucleótido exógeno es expresado por la célula transformada pero no está integrada en el genoma y como tal representa un rasgo transitorio.

Por tanto, según algunos aspectos de la presente invención, se proporciona una célula aislada que comprende la construcción de ácido nucleico de la invención.

Como se usa en el presente documento, la expresión "célula aislada" se refiere a una célula al menos parcialmente separada de una planta. En algunas realizaciones, la célula aislada es una célula vegetal de una planta completa. La célula aislada es una célula vegetal, por ejemplo, una célula vegetal en cultivo.

El término "planta" como se usa en el presente documento abarca plantas completas, ancestros y descendencia de las plantas y partes de plantas, incluyendo semillas, brotes, tallos, raíces (incluyendo tubérculos) y células, tejidos y órganos vegetales. La planta puede estar en cualquier forma, incluyendo cultivos en suspensión, embriones, regiones meristemáticas, tejido de callo, hojas, gametofitos, esporofitos, polen y microesporas. Las plantas que son particularmente útiles en los métodos de la invención incluyen todas las plantas que pertenecen a la superfamilia Viridiplantae, en particular, plantas monocotiledóneas y dicotiledóneas, incluyendo una legumbre forrajera o de pienso, planta ornamental, cultivo alimentario, árbol o arbusto seleccionados de la lista que comprende *Acacia* spp., *Acer* spp., *Actinidia* spp., *Aesculus* spp., *Agathis australis*, *Albizia amara*, *Alsophila tricolor*, *Andropogon* spp., *Arachis* spp., *Areca catechu*, *Astelia fragrans*, *Astragalus cicer*, *Baikiaea plurijuga*, *Betula* spp., *Brassica* spp., *Bruguiera gymnorhiza*, *Burkea africana*, *Butea frondosa*, *Cadaba farinosa*, *Calliandra* spp., *Camellia sinensis*, *Canna indica*, *Capsicum* spp., *Cassia* spp., *Centroema pubescens*, *Chacoomeles* spp., *Cinnamomum cassia*, *Coffea arabica*, *Colophospermum mopane*, *Coronillia varia*, *Cotoneaster serotina*, *Crataegus* spp., *Cucumis* spp., *Cupressus* spp., *Cyathea dealbata*, *Cydonia oblonga*, *Cryptomeria japonica*, *Cymbopogon* spp., *Cynthea dealbata*, *Cydonia oblonga*, *Dalbergia monetaria*, *Davallia divaricata*, *Desmodium* spp., *Dicksonia squarosa*, *Dibeteropogon amplexans*, *Dioclea* spp., *Dolichos* spp., *Dorycnium rectum*, *Echinochloa pyramidalis*, *Ehrafia* spp., *Eleusine coracana*, *Eragrostis* spp., *Erythrina* spp., *Eucalyptus* spp., *Euclea schimperii*, *Eulalia villosa*, *Pagopyrum* spp., *Feijoa sellowiana*, *Fragaria* spp., *Flemingia* spp., *Freyinetia banksii*, *Geranium thunbergii*, *Ginkgo biloba*, *Glycine javanica*, *Gliricidia* spp., *Gossypium hirsutum*, *Grevillea* spp., *Guibourtia coleosperma*, *Hedysarum* spp., *Hemaphysa altissima*, *Heteropogon contoffus*, *Hordeum vulgare*, *Hyparrhenia rufa*, *Hypericum erectum*, *Hypeffhelia disolute*, *Indigo incamata*, *Iris* spp., *Leptarrhenia pyrolifolia*, *Lespedeza* spp., *Lettuca* spp., *Leucaena leucocephala*, *Loudetia simplex*, *Lotonus bainesii*, *Lotus* spp., *Macrotyloma axillare*, *Malus* spp., *Manihot esculenta*, *Medicago saliva*, *Metasequoia glyptostroboides*, *Musa sapientum*, *Nicotianum* spp., *Onobrychis* spp., *Ornithopus* spp., *Oryza* spp., *Peltophorum africanum*, *Pennisetum* spp., *Persea gratissima*, *Petunia* spp., *Phaseolus* spp., *Phoenix canariensis*, *Phormium cookianum*, *Photinia* spp., *Picea glauca*, *Pinus* spp., *Pisum sativum*, *Podocarpus totara*, *Pogonarthria fleckii*, *Pogonaffhria squarosa*, *Populus* spp., *Prosopis cineraria*, *Pseudotsuga menziesii*, *Pterolobium stellatum*, *Pyrus communis*, *Quercus* spp., *Rhaphiolepis umbellata*, *Rhopalostylis sapida*, *Rhus natalensis*, *Ribes grossularia*, *Ribes* spp., *Robinia pseudoacacia*, *Rosa* spp., *Rubus* spp., *Salix* spp., *Schyzachyrium sanguineum*, *Sciadopitys veficillata*, *Sequoia sempervirens*, *Sequoiadendron giganteum*, *Sorghum bicolor*, *Spinacia* spp., *Sporobolus fimbriatus*, *Stiburus alopecuroides*, *Stylosanthes humilis*, *Tadehagi* spp., *Taxodium distichum*, *Themeda triandra*, *Trifolium* spp., *Triticum* spp., *Tsuga heterophylla*, *Vaccinium* spp., *Vicia* spp., *Vitis vinifera*, *Watsonia pyramidata*, *Zantedeschia aethiopica*, *Zea mays*, amaranto, alcachofa, espárrago, brócoli, coles de Bruselas, repollo, canola, zanahoria, coliflor, apio, berza, lino, col común, lenteja, colza de semillas oleaginosas, oca, cebolla, patata, arroz, soja, paja, remolacha azucarera, caña de azúcar, girasol, tomate, té de calabaza, maíz, trigo, cebada, centeno, avena, cacahuete, guisante, lenteja y alfalfa, algodón, colza, canola, pimiento, girasol, tabaco, berenjena, eucalipto, un árbol, una planta ornamental, una hierba perenne y un cultivo forrajero. Como alternativa, se pueden usar algas y otras no Viridiplantae para los métodos de la presente invención.

Según algunas realizaciones de la invención, la planta o célula vegetal usada por el método de la invención es una planta de cultivo o célula de una planta de cultivo tal como arroz, maíz, trigo, cebada, cacahuete, patata, sésamo, olivo, aceite de palma, plátano, soja, girasol, canola, caña de azúcar, alfalfa, mijo, leguminosas (judía, guisante), lino, altramu, colza, tabaco, álamo y algodón.

Según realizaciones adicionales, las células vegetales incluyen células de tabaco, célula de raíz transformada de *Agrobacterium rhizogenes*, célula de apio, célula de jengibre, célula de rábano picante y células de zanahoria. En una realización, las células de tabaco son de una línea celular de tabaco, tal como, pero sin limitación, células de *Nicotiana tabacum* L. cv amarillo brillante (BY-2) (Nagata *et al*, Int Rev Cytol 1992; 132: 1-30). Las células vegetales pueden cultivarse según cualquier tipo de método de cultivo adecuado, incluyendo, pero sin limitación, cultivo en una superficie sólida (tal como un recipiente o una placa de cultivo de plástico, por ejemplo) o en suspensión. Se observará que algunas células, tales como las células BY-2 y de zanahoria se pueden cultivar y crecen en suspensión. Se conocen

en la técnica dispositivos y métodos adecuados para cultivar células vegetales en suspensión, por ejemplo, como se describe en las solicitudes de patente internacional PCT WO98/13469, WO2005/080544 y WO2007/010533. En otra realización más, las células son células de plantas de tabaco completas o tejidos vegetales, incluyendo, pero sin limitación, *Nicotiana benthamiana*.

Existen diversos métodos para introducir genes extraños en plantas tanto monocotiledóneas como dicotiledóneas (Potrykus, I., Annu. Rev. Plant. Physiol., Plant. Mol. Biol. (1991) 42:205-225; Shimamoto *et al.*, Nature (1989) 338:274-276).

Los principales métodos para provocar integración estable de ADN exógeno en ADN genómico de planta incluyen dos enfoques principales:

(i) Transferencia génica mediada por *Agrobacterium*: Klee *et al.* (1987) Annu. Rev. Plant Physiol. 38:467-486; Klee y Rogers en Cell Culture and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant Nuclear Genes, eds. Schell, J. y Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) págs. 2-25; Gatenby, en Plant Biotechnology, eds. Kung, S. y Arntzen, C. J., Butterworth Publishers, Boston, Mass. (1989) págs. 93-112.

(ii) Captación directa de ADN: Paszkowski *et al.*, en Cell Culture and Somatic Cell Genetics of Plants, Vol. 6, Molecular Biology of Plant Nuclear Genes eds. Schell, J. y Vasil, L. K., Academic Publishers, San Diego, Calif. (1989) págs. 52-68; incluyendo métodos para captación directa de ADN en protoplastos, Toriyama, K. *et al.* (1988) Bio/Technology 6:1072-1074. Captación de ADN inducida por descarga eléctrica breve de células vegetales: Zhang *et al.* Plant Cell Rep. (1988)7:379-384. Fromm *et al.* Nature (1986) 319:791-793. Inyección de ADN en células o tejidos vegetales mediante bombardeo de partículas, Klein *et al.* Bio/Technology (1988) 6:559-563; McCabe *et al.* Bio/Technology (1988) 6:923-926; Sanford, Physiol. Plant. (1990) 79:206-209; mediante el uso de sistemas de micropipetas: Neuhaus *et al.*, Theor. Appl. Genet. (1987) 75:30-36; Neuhaus y Spangenberg, Physiol. Plant. (1990)79:213-217; transformación con fibras de vidrio o filamentos de carburo de silicio de cultivos celulares, embriones o tejido calloso, patente de los Estados Unidos n.º 5.464.765 o mediante la incubación directa de ADN con polen germinante, DeWet *et al.* en Experimental Manipulation of Ovule Tissue, eds. Chapman, G. P. y Mantell, S. H. y Daniels, W. Longman, Londres, (1985) págs. 197-209; y Ohta, Proc. Natl. Acad. Sci. USA (1986) 83:715-719.

El sistema de *Agrobacterium* incluye el uso de vectores plasmídicos que contienen segmentos de ADN definidos que se integran en el ADN genómico de la planta. Los métodos de inoculación del tejido vegetal varían según la especie de la planta y el sistema de administración de *Agrobacterium*. Un enfoque ampliamente usado es el procedimiento del disco de hoja que puede realizarse con cualquier explante tisular que proporcione una buena fuente para el inicio de la diferenciación de la planta completa. Véase, p. ej., Horsch *et al.* en Plant Molecular Biology Manual A5, Kluwer Academic Publishers, Dordrecht (1988) págs. 1-9. Otro enfoque emplea el sistema de suministro de *Agrobacterium* en combinación con infiltración al vacío. El sistema de *Agrobacterium* es especialmente viable en la creación de plantas dicotiledóneas transgénicas.

Existen diversos métodos de transferencia directa de ADN a células vegetales. En electroporación, los protoplastos se exponen brevemente a un campo eléctrico fuerte. En microinyección, el ADN se inyecta de manera mecánica directamente en las células usando micropipetas muy pequeñas. En bombardeo de micropartículas, el ADN se adsorbe en microproyectiles, tales como cristales de sulfato de magnesio o partículas de tungsteno, y los microproyectiles se aceleran físicamente en células o tejidos vegetales.

Después de la transformación estable, se realiza propagación de la planta. El método más común de propagación de plantas es mediante semillas. La regeneración mediante propagación por semillas, sin embargo, tiene la deficiencia de que, debido a la heterocigosidad, existe una falta de uniformidad en el cultivo, ya que las semillas son producidas por plantas según las variaciones genéticas regidas por las normas mendelianas. Básicamente, cada semilla es genéticamente diferente y cada una crecerá con sus propios rasgos específicos. Por lo tanto, se prefiere que la planta transformada se produzca de manera que la planta regenerada tenga los rasgos y características idénticos de la planta transgénica original. Por lo tanto, se prefiere que la planta transformada se regenere mediante micropropagación, lo que proporciona una reproducción rápida, uniforme, de las plantas transformadas.

La micropropagación es un proceso de cultivo de plantas de nueva generación a partir de un solo trozo de tejido que se ha escindido de una planta o cultivar original seleccionado. Este proceso permite la reproducción en masa de plantas que tienen el tejido preferido que expresa la proteína de fusión. Las plantas de nueva generación que se producen son genéticamente idénticas a, y tienen todas las características de, la planta original. La micropropagación permite la producción en masa de material vegetal de calidad en un periodo de tiempo corto y ofrece una multiplicación rápida de cultivares seleccionados en la conservación de las características de la planta original transgénica o transformada. Las ventajas de la clonación de plantas son la velocidad de multiplicación de plantas y la calidad y uniformidad de plantas producidas.

La micropropagación es un procedimiento de múltiples etapas que requiere alteración del medio de cultivo o de las condiciones de crecimiento entre las etapas. Por tanto, el proceso de micropropagación implica cuatro etapas básicas: Etapa uno, cultivo de tejido inicial; etapa dos, multiplicación de cultivo tisular; etapa tres, diferenciación y formación de

plantas; y etapa cuatro, cultivo y endurecimiento en invernadero. Durante la etapa uno, cultivo tisular inicial, el cultivo tisular se establece y se certifica que está exento de contaminantes. Durante la etapa dos, el cultivo tisular inicial se multiplica hasta que se produce un número suficiente de muestras tisulares para cumplir los objetivos de producción. Durante la etapa tres, las muestras tisulares cultivadas en la etapa dos se dividen y cultivan en plántulas individuales. En la etapa cuatro, las plántulas transformadas se transfieren a un invernadero para su endurecimiento, donde la tolerancia de las plantas a la luz aumenta gradualmente de modo que pueda cultivarse en el entorno natural.

Según algunas realizaciones de la invención, las plantas transgénicas se generan mediante transformación transitoria de células de la hoja, células meristemáticas o la planta completa.

Puede efectuarse transformación transitoria mediante cualquiera de los métodos de transferencia directa de ADN descritos anteriormente o mediante infección vírica usando virus vegetales modificados.

Los virus que se han mostrado que son útiles para la transformación de hospedadores vegetales incluyen CaMV, virus del mosaico del tabaco (TMV), virus del mosaico del bromo (BMV) y virus del mosaico común de la judía (BV o BCMV). Se describe transformación de plantas usando virus vegetales en la patente de los Estados Unidos n.º 4.855.237 (virus del mosaico dorado de la judía; BGV), documento EP-A 67.553 (TMV), solicitud publicada japonesa n.º 63-14693 (TMV), documento EPA 194.809 (BV), documento EPA 278.667 (BV); y Gluzman, Y. *et al.*, *Communications in Molecular Biology: Viral Vectors*, Cold Spring Harbor Laboratory, Nueva York, págs. 172-189 (1988). Se describen partículas de pseudovirus para su uso en la expresión de ADN extraño en muchos hospedadores, incluyendo plantas, en el documento WO 87/06261.

Según algunas realizaciones de la invención, el virus usado para transformaciones transitorias es avirulento y, por tanto, es incapaz de provocar síntomas graves, tales como tasa de crecimiento reducida, mosaico, manchas anulares, enrollamiento de las hojas, amarilleo, picado, formación de viruela, formación de tumores y erosión. Un virus avirulento adecuado puede ser un virus avirulento de origen natural o un virus atenuado artificialmente. Puede efectuarse atenuación de virus usando métodos bien conocidos en la técnica incluyendo, pero sin limitación, calentamiento subletal, tratamiento químico o mediante técnicas de mutagénesis dirigida tal como se describe, por ejemplo, en Kurihara y Watanabe (*Molecular Plant Pathology* 4: 259-269, 2003), Galon *et al.* (1992), Atreya *et al.* (1992) y Huet *et al.* (1994).

Se pueden obtener cepas de virus adecuadas de fuentes disponibles tales como, por ejemplo, la Colección Americana de Cultivos Tipo (ATCC) o mediante aislamiento de plantas infectadas. Se puede efectuar aislamiento de virus a partir de tejidos vegetales infectados mediante técnicas bien conocidas en este campo, tal como se describe, por ejemplo, en Foster y Tatlor, Eds. "Plant Virology Protocols: From Virus Isolation to Transgenic Resistance (Methods in Molecular Biology (Humana Pr), Vol 81)", Humana Press, 1998. En resumen, se muelen tejidos de una planta infectada que se cree que contienen una alta concentración de un virus adecuado, preferentemente hojas jóvenes y pétalos de flores, en una solución de tampón (p. ej., solución de tampón de fosfato) para producir una savia infectada con virus que puede usarse en inoculaciones posteriores.

La construcción de virus de ARN de plantas para la introducción y expresión de secuencias de ácido nucleico no víricas en plantas se demuestra por las referencias anteriores, así como por Dawson, W. O. *et al.*, *Virology* (1989) 172:285-292; Takamatsu *et al.* *EMBO J.* (1987) 6:307-311; French *et al.* *Science* (1986) 231:1294-1297; Takamatsu *et al.* *FEBS Letters* (1990) 269:73-76; y patente de los Estados Unidos n.º 5.316.931.

Cuando el virus es un virus de ADN, se pueden hacer modificaciones adecuadas al virus en sí mismo. Como alternativa, el virus se puede clonar primero en un plásmido bacteriano para facilitar la construcción del vector vírico deseado con el ADN extraño. Después, el virus puede escindirse del plásmido. Si el virus es un virus de ADN, puede unirse a un origen bacteriano de replicación al ADN vírico, que después es replicado por las bacterias. La transcripción y traducción de este ADN producirá la proteína de cubierta que encapsulará el ADN vírico. Si el virus es un virus de ARN, el virus generalmente se clona como un ADNc y se inserta en un plásmido. El plásmido se usa después para realizar todas las construcciones. El virus de ARN se produce después al transcribir la secuencia vírica del plásmido y la traducción de los genes víricos para producir la o las proteínas de cubierta que encapsidan el ARN vírico.

En una realización, se proporciona un polinucleótido vírico de planta en el que la secuencia codificante de proteína de cubierta nativa se ha suprimido de un polinucleótido vírico, se ha insertado una secuencia codificante de proteínas de la cubierta vírica de plantas no nativas y un promotor no nativo, preferentemente el promotor subgenómico de la secuencia codificante de proteínas de la cubierta no nativa, con capacidad de expresión en el hospedador vegetal, empaquetado del polinucleótido vírico de la planta recombinante y que garantiza una infección sistémica del hospedador por el polinucleótido vírico de la planta recombinante. Como alternativa, el gen de proteína de cubierta puede inactivarse mediante inserción de la secuencia polinucleotídica no nativa dentro de él, de modo que se produce una proteína. El polinucleótido vírico de la planta recombinante puede contener uno o más promotores subgenómicos no nativos adicionales. Cada promotor subgenómico no nativo es capaz de transcribir o expresar genes adyacentes o secuencias polinucleotídicas en el hospedador vegetal e incapaz de recombinarse entre sí y con promotores subgenómicos nativos. Pueden insertarse secuencias polinucleotídicas no nativas (extrañas) adyacentes al promotor subgenómico vírico o los promotores subgenómicos víricos de planta nativa y una no nativa si se incluye más de una

secuencia polinucleotídica. Las secuencias polinucleotídicas no nativas se transcriben o expresan en la planta hospedadora bajo control del promotor subgenómico para producir los productos deseados.

5 En una segunda realización, se proporciona un polinucleótido vírico de planta recombinante como en la primera realización, excepto que la secuencia codificante de proteína de cubierta nativa se coloca adyacente a uno de los promotores subgenómicos de proteína de cubierta no nativa en lugar de una secuencia codificante de proteína de cubierta no nativa.

10 En una tercera realización, se proporciona un polinucleótido vírico de planta recombinante en el que el gen de la proteína de cubierta nativa está adyacente a su promotor subgenómico y se han insertado uno o más promotores subgenómicos no nativos en el polinucleótido vírico. Los promotores subgenómicos no nativos insertados son capaces de transcribir o expresar genes adyacentes en una planta hospedadora y son incapaces de recombinarse entre sí y con promotores subgenómicos nativos. Pueden insertarse secuencias polinucleotídicas no nativas adyacentes a los promotores víricos de plantas subgenómicos no nativos de modo que las secuencias se transcriban o expresen en la
15 planta hospedadora bajo el control de los promotores subgenómicos para producir el producto deseado.

20 En una cuarta realización, se proporciona un polinucleótido vírico de planta recombinante como en la tercera realización, excepto que la secuencia codificante de proteína de cubierta nativa se reemplaza por una secuencia codificante de proteína de cubierta no nativa.

25 Los vectores víricos están encapsidados por las proteínas de cubierta codificadas por el polinucleótido vírico de la planta recombinante para producir un virus vegetal recombinante. El polinucleótido vírico de planta recombinante o virus de planta recombinante se usa para infectar plantas hospedadoras adecuadas. El polinucleótido vírico de la planta recombinante tiene capacidad de replicación en el hospedador, diseminación sistémica en el hospedador y transcripción o expresión de gen o genes extraños (polinucleótido exógeno) en el hospedador para producir la proteína deseada.

30 Se pueden encontrar técnicas para inoculación de virus en plantas en Foster y Taylor, eds. "Plant Virology Protocols: From Virus Isolation to Transgenic Resistance (Methods in Molecular Biology (Humana Pr), Vol 81)", Humana Press, 1998; Maramorosh y Koprowski, eds. "Methods in Virology" 7 vols, Academic Press, Nueva York 1967-1984; Hill, S.A. "Methods in Plant Virology", Blackwell, Oxford, 1984; Walkey, D.G.A. "Applied Plant Virology", Wiley, Nueva York, 1985; y Kado y Agrawa, eds. "Principles and Techniques in Plant Virology", Van Nostrand-Reinhold, Nueva York.

35 Además de lo anterior, el polinucleótido de la presente invención también puede introducirse en un genoma de cloroplasto permitiendo de ese modo la expresión en cloroplasto.

40 Se conoce una técnica para introducir secuencias de ácido nucleico exógenas en el genoma de los cloroplastos. Esta técnica implica los siguientes procedimientos. En primer lugar, las células vegetales se tratan químicamente para reducir el número de cloroplastos por célula a aproximadamente uno. Después, el polinucleótido exógeno se introduce mediante bombardeo de partículas en las células con el objetivo de introducir al menos una molécula polinucleotídica exógena en los cloroplastos. Los polinucleótidos exógenos se seleccionan de modo que se puedan integrar en el genoma del cloroplasto mediante recombinación homóloga que se efectúa fácilmente mediante enzimas inherentes al cloroplasto. Con este fin, la secuencia de ácido nucleico incluye, además de un gen de interés, al menos un tramo de polinucleótidos que procede del genoma del cloroplasto. Además, el polinucleótido exógeno incluye un marcador
45 seleccionable, que actúa mediante procedimientos de selección secuencial para determinar que todas o sustancialmente todas las copias de los genomas de cloroplasto después de dicha selección incluirán el polinucleótido exógeno. Se encuentran detalles adicionales relacionados con esta técnica en las patentes de los Estados Unidos n.º 4.945.050; y 5.693.507 a las que se hace referencia en el presente documento. Por tanto, el sistema de expresión de proteínas del cloroplasto puede producir un polipéptido y este puede integrarse en la membrana interna del cloroplasto.
50

Según algunas realizaciones de la invención, el método comprende además cultivar la célula vegetal que expresa el ácido nucleico. La célula vegetal se puede cultivar como parte de una planta completa o, como alternativa, en cultivo de células vegetales. En una realización, la célula vegetal se cultiva en cultivo de suspensión axénica, en un biorreactor. Se describen en detalle biorreactores adecuados para cultivar las suspensiones de células vegetales de la invención y métodos para hacer crecer las células vegetales en suspensión en, entre otros, patentes de los Estados Unidos n.º 6.391.638 y 7.951.557 y documentos PCT WO98/13469, WO2005/080544, WO2007/010533, WO2008/135991 y WO2008/132743, todos los cuales se mencionan en el presente documento.
55

60 Por tanto, la invención abarca plantas o cultivos de plantas que expresan las secuencias de ácido nucleico, para producir la proteína α -galactosidasa humana recombinante de la invención. Una vez expresada dentro de la célula vegetal o la planta completa, el nivel de la proteína α -galactosidasa humana codificada por la secuencia de ácido nucleico puede determinarse mediante métodos bien conocidos en la técnica, tales como, ensayos de actividad, transferencia de Western que usa anticuerpos capaces de unirse específicamente a la proteína α -galactosidasa humana, ensayo inmunoabsorbente ligado a enzimas (ELISA), radioinmunoensayos (RIA), inmunohistoquímica, inmunocitoquímica, inmunofluorescencia y similares.
65

Se conocen bien en la técnica métodos para determinar el nivel en la planta del ARN transcrito a partir de la secuencia de ácido nucleico e incluyen, por ejemplo, análisis de transferencia de Northern, análisis de reacción en cadena de la polimerasa de transcripción inversa (RT-PCR) (incluyendo RT-PCR cuantitativa, semicuantitativa o en tiempo real) e hibridación de ARN-*in situ*.

El análisis de glucanos de la proteína α -galactosidasa humana recombinante expresada de la presente invención indica que el polipéptido está glucosilado en la célula vegetal, dando como resultado una proteína α -galactosidasa humana recombinante que tiene restos de glucanos específicos de planta y manosa expuestos (véase el ejemplo IV a continuación). Por tanto, según algunas realizaciones de la invención, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene al menos uno, opcionalmente al menos dos, opcionalmente al menos tres u opcionalmente al menos cuatro o más restos de β -(1,2) xilosa de núcleo. En otras realizaciones más, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene al menos uno, opcionalmente al menos dos, opcionalmente al menos tres u opcionalmente al menos cuatro o más restos de α -(1,3) fucosa de núcleo. En otras realizaciones, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene al menos uno, opcionalmente al menos dos, opcionalmente al menos tres u opcionalmente al menos cuatro o más β (1-2)N acetilglucosaminas terminales. En otras realizaciones, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene al menos uno, opcionalmente al menos dos, opcionalmente al menos tres u opcionalmente al menos cuatro o más restos de manosa expuestos. En una realización particular, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene nueve restos de manosa (M9), siendo al menos uno, opcionalmente al menos dos, opcionalmente al menos tres de ellos restos de manosa expuestos. En una realización, las células que expresan el vector de expresión de la invención producen una proteína α -galactosidasa humana que tiene al menos un resto de manosa expuesto, al menos un resto de β -(1,2) xilosa de núcleo y al menos un resto de α -(1,3) fucosa.

Adicionalmente, según algunas realizaciones de la invención, se proporciona una composición que comprende una pluralidad de moléculas de proteína α -galactosidasa humana producidas por células que expresan el vector de expresión de la invención, en donde las estructuras de glucano más predominantes de la pluralidad de moléculas de proteína α -galactosidasa humana son manosa 3- β -(1,2) xilosa- α -(1,3) fucosa [Fc(3)M3X] y/o manosa 4- β -(1,2) xilosa (M4X); y manosa 8 (M8). En algunas realizaciones de la invención, las estructuras de glucano manosa 3- β -(1,2) xilosa- α -(1,3) fucosa [Fc(3)M3X] y/o manosa 4- β -(1,2) xilosa (M4X) comprenden aproximadamente 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 % o aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 % del perfil de moléculas de proteína α -galactosidasa humana en la pluralidad, y la estructura de glucano manosa 8 (M8) comprende 20 %, aproximadamente 25 %, aproximadamente 30 %, aproximadamente 35 % o aproximadamente 40 %, aproximadamente 45 %, aproximadamente 50 %, aproximadamente 55 %, aproximadamente 60 %, aproximadamente 70 %, aproximadamente 80 % del perfil de las moléculas de proteína α -galactosidasa humana en la pluralidad.

En más realizaciones adicionales de la presente invención, la composición comprende una pluralidad de moléculas de proteína α -galactosidasa humana producidas por células que expresan el vector de expresión de la invención, en donde dicha pluralidad de moléculas de proteína α -galactosidasa humana comprende al menos aproximadamente 0,5 %, al menos aproximadamente 0,6 %, aproximadamente 0,7 %, aproximadamente 0,8 %, aproximadamente 0,9 %, aproximadamente 1,0 %, aproximadamente 1,1 %, aproximadamente 1,2 %, aproximadamente 1,3 %, aproximadamente 1,4 %, aproximadamente 1,5 %, aproximadamente 1,75 %, aproximadamente 2,0 %, aproximadamente 2,5 %, aproximadamente 3,0 %, aproximadamente 3,5 %, aproximadamente 4,0 %, aproximadamente 4,5 %, aproximadamente 5,0 % o más moléculas de proteína α -galactosidasa humana que tienen estructuras de glucano de nueve restos de manosa, siendo al menos uno, opcionalmente al menos dos, opcionalmente al menos tres de ellos restos de manosa expuestos.

Se mostró que la proteína α -galactosidasa humana producida por células que expresan el vector de expresión de la invención tiene actividad catalítica de α -galactosidasa idéntica a la de la enzima α -galactosidasa humana nativa. La exposición de la α -galactosidasa recombinante de la invención a condiciones que simulan un ambiente *in vivo* de la enzima cuando se administra para terapia de reemplazo enzimático indica además que la proteína α -galactosidasa humana producida por células que expresan el vector de expresión de la invención podría poseer farmacocinética similar a la de α -galactosidasa humana recombinante expresada en células de mamífero (Replagal®) (véase ejemplo 3 y figs. 2 y 3 en el presente documento).

Por tanto, la proteína α -galactosidasa humana producida por células que expresan el vector de expresión de la invención puede usarse para producir una composición farmacéutica. La composición farmacéutica se puede usar para el tratamiento o la prevención de enfermedad de Fabry en un sujeto que lo necesite. En algunas realizaciones, pueden usarse células que expresan la proteína α -galactosidasa humana para el tratamiento o la prevención de enfermedad de Fabry en un sujeto que lo necesite. Se conocen en la técnica métodos para la preparación y la administración de células vegetales que expresan de manera recombinante proteínas humanas, por ejemplo, documento PCT WO2007/010533 de Shaaltiel *et al.*

Por tanto, según otro aspecto de la presente invención, se proporciona una composición farmacéutica que incluye, como principio activo de la misma, una proteína α -galactosidasa humana fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico y un vehículo farmacéuticamente aceptable. En algunas realizaciones de la presente invención, la proteína α -galactosidasa humana tiene un resto de glicina N-terminal. En algunas realizaciones de la presente invención, la proteína α -galactosidasa humana tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 16. En algunas realizaciones, la proteína α -galactosidasa humana fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico tiene un aminoácido como se expone en la SEQ ID NO: 21.

Como se usa en el presente documento, una "composición farmacéutica" se refiere a una preparación de uno o más de los principios activos descritos en el presente documento con otros componentes químicos tales como vehículos y excipientes fisiológicamente adecuados. El fin de una composición farmacéutica es facilitar la administración de un compuesto a un organismo.

En el presente documento, la expresión "principio activo" se refiere a la proteína recombinante α -galactosidasa humana responsable del efecto biológico.

En lo sucesivo en el presente documento, las expresiones "vehículo fisiológicamente aceptable" y "vehículo farmacéuticamente aceptable" que pueden usarse indistintamente se refieren a un vehículo o un diluyente que no provoca irritación significativa a un organismo y no anula la actividad biológica y propiedades del compuesto administrado. Se incluye un adyuvante en estas expresiones.

En el presente documento, la expresión "excipiente" se refiere a una sustancia inerte añadida a una composición farmacéutica para facilitar adicionalmente la administración de un principio activo. Los ejemplos, sin limitación, de excipientes incluyen carbonato de calcio, fosfato de calcio, diversos azúcares y tipos de almidón, derivados de celulosa, gelatina, aceites vegetales y polietilenglicoles.

Pueden encontrarse técnicas para formulación y administración de fármacos en "Remington's Pharmaceutical Sciences", Mack Publishing Co., Easton, PA, última edición, que se incorpora en el presente documento por referencia.

Las vías adecuadas de administración pueden, por ejemplo, incluir suministro oral, rectal, transmucoso, especialmente transnasal, intestinal o parenteral, incluyendo inyecciones intramuscular, subcutánea e intramedular así como inyecciones intratecal, intraventricular directa, intracardiaca, p. ej., en la cavidad ventricular derecha o izquierda, en la arteria coronaria común, intravenosa, intraperitoneal, intranasal o intraocular.

Los enfoques convencionales para suministro de fármacos al sistema nervioso central (SNC) incluyen: estrategias neuroquirúrgicas (p. ej., inyección intracerebral o infusión intracerebroventricular); manipulación molecular del agente (p. ej., producción de una proteína de fusión quimérica que comprende un péptido de transporte que tiene una afinidad por una molécula de superficie celular endotelial en combinación con un agente que en sí mismo es incapaz de cruzar la BHE) en un intento de aprovechar una de las rutas de transporte endógenas de la BHE; estrategias farmacológicas diseñadas para aumentar la solubilidad en lípidos de un agente (p. ej., conjugación de agentes hidrosolubles con vehículos lipídicos o de colesterol); y la alteración transitoria de la integridad de la BHE mediante alteración hiperosmótica (resultante de una solución de manitol en la arteria carótida o el uso de un agente biológicamente activo tal como un péptido de angiotensina). Sin embargo, cada una de estas estrategias tiene limitaciones, tales como los riesgos inherentes asociados con un procedimiento quirúrgico invasivo, una limitación de tamaño impuesta por una limitación inherente a los sistemas de transporte endógeno, efectos secundarios biológicos potencialmente indeseables asociados con la administración sistémica de una molécula quimérica compuesta por un motivo portador que podría estar activo fuera del SNC y el posible riesgo de daño cerebral dentro de las regiones del cerebro donde se altera la BHE, lo que lo convierte en un método de suministro subóptimo.

Como alternativa, se puede administrar la composición farmacéutica de una manera local en lugar de sistémica, por ejemplo, mediante inyección de la composición farmacéutica directamente en una región tisular de un paciente.

El término "tejido" se refiere a parte de un organismo que consiste en células diseñadas para realizar una función o funciones. Los ejemplos incluyen, pero sin limitación, tejido cerebral, retina, tejido cutáneo, tejido hepático, tejido pancreático, hueso, cartílago, tejido conjuntivo, tejido sanguíneo, tejido muscular, tejido cardíaco, tejido cerebral, tejido vascular, tejido renal, tejido pulmonar, tejido gonadal, tejido hematopoyético.

Las composiciones farmacéuticas de algunas realizaciones de la invención pueden fabricarse mediante procesos bien conocidos en la técnica, p. ej., por medio de procesos convencionales de mezclado, disolución, granulación, preparación de grageas, levigación, emulsión, encapsulación, inmovilización o liofilización.

Las composiciones farmacéuticas para su uso de acuerdo con algunas realizaciones de la invención pueden formularse por tanto de manera convencional usando uno o más vehículos fisiológicamente aceptables que comprenden excipientes y adyuvantes, que facilitan el procesamiento de los principios activos en preparaciones que pueden usarse de forma farmacéutica. La formulación apropiada depende de la vía de administración elegida.

Para inyección, los principios activos de la composición farmacéutica pueden formularse en soluciones acuosas, preferentemente, en tampones fisiológicamente compatibles, tales como solución de Hank, solución de Ringer o tampón salino fisiológico. Para administración transmucosa, se usan en la formulación agentes penetrantes adecuados para la barrera que va a permearse. Dichos penetrantes son conocidos en general en la técnica.

Para administración oral, los principios activos pueden formularse opcionalmente mediante la administración de las células completas que producen una proteína α -galactosidasa humana según la presente invención. Los principios activos también se pueden formular opcionalmente combinando los principios activos y/o las células con vehículos farmacéuticamente aceptables bien conocidos en la técnica, produciendo composiciones farmacéuticas. Dichos vehículos permiten que las células y/o la composición farmacéutica se formulen como comprimidos, píldoras, grageas, cápsulas, líquidos, geles, jarabes, pastas, suspensiones y similares, para ingesta oral por un paciente. Cuando se usan para administración oral, se pueden usar células enteras o fracciones de células nuevas o procesadas (p. ej., liofilizadas, congeladas, secas, etc).

Pueden prepararse preparaciones farmacológicas para uso oral usando un excipiente sólido, opcionalmente moliendo la mezcla resultante y procesando la mezcla de gránulos, después de añadir adyuvantes adecuados si se desea, para obtener comprimidos o núcleos de grageas. Son excipientes adecuados, en particular, cargas tales como azúcares, incluyendo lactosa, sacarosa, manitol o sorbitol; preparaciones de celulosa, tales como, por ejemplo, almidón de maíz, almidón de trigo, almidón de arroz, almidón de patata, gelatina, goma de tragacanto, metilcelulosa, hidroxipropilmetilcelulosa, carbometilcelulosa de sodio; y/o polímeros fisiológicamente aceptables tales como polivinilpirrolidona (PVP). Si se desea, pueden añadirse agentes disgregantes, tales como polivinilpirrolidona reticulada, agar o ácido algínico o una sal de los mismos, tal como alginato de sodio.

Los núcleos de grageas se proporcionan con recubrimientos adecuados. Para este fin, pueden usarse soluciones concentradas de azúcar que opcionalmente pueden contener goma arábiga, talco, polivinilpirrolidona, gel de carbopol, polietilenglicol, dióxido de titanio, soluciones de laca y disolventes orgánicos o mezclas de disolventes adecuados. Se pueden añadir colorantes o pigmentos a los comprimidos o recubrimientos de grageas para identificación o para caracterizar diferentes combinaciones de dosis de compuestos activos.

Las composiciones farmacéuticas que pueden usarse por vía oral incluyen capsulas de ajuste por presión hechas de gelatina, así como cápsulas selladas, blandas, hechas de gelatina y un plastificante, tal como glicerol o sorbitol. Las cápsulas de ajuste por presión pueden contener los principios activos en mezcla con carga tal como lactosa, aglutinantes tales como almidones, lubricantes tales como talco o estearato de magnesio y, opcionalmente, estabilizantes. En las cápsulas blandas, los principios activos pueden disolverse o suspenderse en líquidos adecuados, tales como aceites grasos, parafina líquida o polietilenglicoles líquidos. Además, se pueden añadir estabilizantes. Todas las formulaciones para administración oral deben estar en dosis adecuadas para la vía de administración elegida.

Para administración bucal, las composiciones puede tomar la forma de comprimidos o pastillas para chupar formuladas de forma convencional.

Para administración mediante inhalación nasal, los principios activos para su uso según algunas realizaciones de la invención se suministran convenientemente en forma de una presentación de pulverización en aerosol desde un envase presurizado o un nebulizador con el uso de un propulsor adecuado, p. ej., diclorodifluorometano, triclorofluorometano, diclorotetrafluoroetano o dióxido de carbono. En el caso de un aerosol presurizado, la unidad de dosificación se puede determinar proporcionando una válvula para suministrar una cantidad medida. Pueden formularse cápsulas y cartuchos de, p. ej., gelatina para su uso en un dosificador que contienen una mezcla de polvo del compuesto y una base de polvo adecuada tal como lactosa o almidón.

La composición farmacéutica descrita en el presente documento puede formularse para administración parenteral, p. ej., mediante inyección de embolada o infusión continua. Pueden presentarse formulaciones para inyección en forma de dosificación unitaria, p. ej., en ampollas o en recipientes multidosis con, opcionalmente, un conservante añadido. Las composiciones pueden ser suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos y pueden contener agentes de formulación tales como agentes de suspensión, estabilizantes y/o de dispersión.

Las composiciones farmacéuticas para administración parenteral incluyen soluciones acuosas de la preparación activa en forma hidrosoluble. Adicionalmente, pueden prepararse suspensiones de los principios activos como suspensiones para inyección de base oleosa o acuosa adecuadas. Los disolventes o vehículos lipófilos adecuados incluyen aceites grasos tales como aceite de sésamo o ésteres de ácidos grasos sintéticos tales como oleato de etilo, triglicéridos o liposomas. Las suspensiones acuosas de inyección pueden contener sustancias, que aumentan la viscosidad de la suspensión, tales como carboximetilcelulosa sódica, sorbitol o dextrano. Opcionalmente, la suspensión también puede contener estabilizantes o agentes adecuados que aumenten la solubilidad de los principios activos para permitir la preparación de soluciones muy concentradas.

Como alternativa, el principio activo puede estar en forma de polvo para su constitución con un vehículo adecuado, p.

ej., solución a base de agua apirógena, estéril, antes de su uso.

La composición farmacéutica de algunas realizaciones de la invención también puede formularse en composiciones rectales tales como supositorios o enemas de retención, usando, p. ej., bases de supositorio convencionales tales como manteca de cacao u otros glicéridos.

Las composiciones farmacéuticas adecuadas para su uso en el contexto de algunas realizaciones de la invención incluyen composiciones en donde los principios activos están contenidos en una cantidad eficaz para lograr el fin pretendido. De manera más específica, una cantidad terapéuticamente eficaz significa una cantidad de principios activos (α -galactosidasa) eficaces para prevenir, aliviar o mejorar síntomas de un trastorno (p. ej., enfermedad de Fabry) o prolongar la supervivencia del sujeto que se trate.

La determinación de una cantidad terapéuticamente eficaz está dentro de la capacidad de los expertos en la materia, especialmente a la luz de la divulgación detallada proporcionada en el presente documento.

Para cualquier preparación usada en los métodos de la invención, la cantidad o dosis terapéuticamente eficaz puede estimarse inicialmente a partir de ensayos de cultivo celular e *in vitro*. Por ejemplo, una dosis puede formularse en modelos animales para conseguir una concentración o un título deseado. Dicha información se puede usar para determinar de manera más precisa las dosis útiles en seres humanos.

La toxicidad y eficacia terapéutica de los principios activos descritos en el presente documento pueden determinarse mediante procedimientos farmacéuticos convencionales *in vitro*, en cultivos celulares o animales experimentales. Los datos obtenidos de estos ensayos *in vitro* y de cultivo celular y estudios animales pueden usarse en la formulación de un intervalo de dosificación para su uso en el ser humano. La dosificación puede variar dependiendo de la forma de dosificación empleada y la vía de administración utilizada. La formulación exacta, la vía de administración y la dosificación pueden ser elegidas por el médico individual a la vista del estado del paciente. (Véase, p. ej., Fingl, *et al.*, 1975, en "The Pharmacological Basis of Therapeutics", C. 1 p. 1).

La cantidad y el intervalo de dosificación pueden ajustarse individualmente para proporcionar niveles séricos y celulares del principio activo que sean suficientes para inducir o suprimir el efecto biológico (concentración eficaz mínima, CEM). La CEM variará para cada preparación, pero puede estimarse a partir de datos *in vitro*. Las dosis necesarias para lograr la CEM dependerán de las características individuales y la vía de administración. Pueden usarse ensayos de detección para determinar las concentraciones en plasma.

Dependiendo de la gravedad y sensibilidad de la afección para tratar, la dosificación puede ser de una única administración o una pluralidad de administraciones, durando el ciclo de tratamiento de varios días a varias semanas o hasta que se efectúe la cura o se logre disminución de la patología.

La cantidad de una composición que se va a administrar dependerá, por supuesto, del sujeto que se trate, la gravedad de la afección, la forma de administración, el criterio del médico que la receta, etc.

Las composiciones de algunas realizaciones de la invención pueden presentarse, si se desea, en un envase o dispositivo dosificador, tal como un kit aprobado por la FDA, que puede contener una o más formas farmacéuticas unitarias que contienen el principio activo. El envase puede, por ejemplo, comprender papel metálico o de plástico, tal como un envase de tipo blíster. El envase o dispositivo dosificador puede estar acompañado de instrucciones para la administración. El envase o dosificador también puede estar acompañado de una nota asociada con el recipiente en una forma prescrita por una agencia gubernamental que regule la fabricación, el uso o la venta de productos farmacéuticos, reflejando dicha nota la aprobación por parte de la agencia de la forma de las composiciones o la administración humana o veterinaria. Dicha nota, por ejemplo, puede ser etiquetado aprobado por la administración de alimentos y fármacos de los Estados Unidos para fármacos con receta o de un prospecto de producto aprobado. También pueden prepararse composiciones que comprenden una preparación de la invención formulada en un vehículo farmacéutico compatible, colocarse en un recipiente adecuado y marcarse para el tratamiento de una afección indicada, como se ha detallado anteriormente de manera adicional.

La composición farmacéutica se puede usar para el tratamiento o la prevención de enfermedad de Fabry en un sujeto que lo necesite. Por tanto, según otro aspecto de la presente invención, se proporciona un método para tratar o prevenir enfermedad de Fabry en un sujeto que lo necesite, comprendiendo el método administrar al sujeto una cantidad eficaz de una composición farmacéutica que incluye, como principio activo de la misma, una proteína α -galactosidasa humana fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico y un vehículo farmacéuticamente aceptable. En algunas realizaciones de la presente invención, la proteína α -galactosidasa humana tiene un resto de glicina N-terminal. En otras realizaciones, la proteína α -galactosidasa humana tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 16.

El término "tratar" se refiere a inhibir, prevenir o detener el desarrollo de una patología (enfermedad, trastorno o afección) y/o provocar la reducción, remisión o regresión de una patología. Los expertos en la materia entenderán que pueden usarse diversas metodologías y ensayos para evaluar el desarrollo de una patología y, de manera similar, se

pueden usar diversas metodologías y ensayos para evaluar la reducción, remisión o regresión de una patología.

Como se usa en el presente documento, el término "prevenir" se refiere a evitar que una enfermedad, un trastorno o una afección se produzca en un sujeto que pueda estar en riesgo de padecer la enfermedad, pero al que todavía no se ha diagnosticado que tenga la enfermedad.

Como se usa en el presente documento, el término "sujeto" incluye mamíferos, preferiblemente seres humanos a cualquier edad que padecen la patología. Preferentemente, este término abarca a los individuos que están en riesgo de desarrollar la patología.

Como se usa en el presente documento, la expresión "régimen de tratamiento" se refiere a un plan de tratamiento que especifica el tipo de tratamiento, dosificación, programa y/o duración de un tratamiento proporcionado a un sujeto que lo necesite (p. ej., un sujeto al que se ha diagnosticado una patología). El régimen de tratamiento seleccionado puede ser uno agresivo y se espera que dé como resultado el mejor resultado clínico (p. ej., curación completa de la patología) o uno más moderado que pueda aliviar los síntomas de la patología, pero dé como resultado curación incompleta de la patología. Se apreciará que en determinados casos el régimen de tratamiento más agresivo puede estar asociado con cierta incomodidad para el sujeto o efectos secundarios adversos (p. ej., un daño a células o tejidos sanos). El tipo de tratamiento puede incluir una intervención quirúrgica (p. ej., extirpación de la lesión, células, tejido u órgano enfermos, una terapia de reemplazo celular, una administración de un fármaco terapéutico (p. ej., agonistas de receptores, antagonistas, hormonas, agentes quimioterapéuticos) en modo local o sistémico, una exposición a la radioterapia usando una fuente externa (p. ej., haz externo) y/o una fuente interna (p. ej., braquiterapia) y/o cualquier combinación de las mismas. La dosificación, el programa y la duración del tratamiento pueden variar, dependiendo de la gravedad de la patología y el tipo de tratamiento seleccionado, y los expertos en la materia son capaces de ajustar el tipo de tratamiento con la dosis, el programa y la duración del tratamiento.

Se espera que durante la vida de una patente que surja a partir de esta solicitud, se desarrollen muchos vectores, elementos promotores, células vegetales y vehículos relevantes y se entiende que el alcance de los términos proporcionados en el presente documento incluye todas esas nuevas tecnologías *a priori*.

Como se usa en el presente documento, el término "aproximadamente" se refiere a $\pm 10\%$.

Las expresiones "comprende", "que comprende", "incluye", "que incluye", "que tiene" y sus conjugados significan "que incluye pero sin limitación".

La expresión "que consiste en" significa "que incluye y limitado a".

La expresión "que consiste esencialmente en" significa que la composición, el método y la estructura pueden incluir ingredientes, etapas y/o partes adicionales, pero solamente si los ingredientes, las etapas y/o las partes adicionales no alteran materialmente las características básicas y novedosas de la composición, el método o la estructura reivindicados.

Como se usa en el presente documento, la forma singular "un", "uno/a" y "el/la" incluyen referencias plurales a menos que el contexto indique claramente otra cosa. Por ejemplo, la expresión "un compuesto" o "al menos un compuesto" puede incluir una pluralidad de compuestos, incluyendo mezclas de los mismos.

A lo largo de la presente solicitud, se pueden presentar diversas realizaciones de la presente invención en un formato de intervalo. Debería entenderse que la descripción en formato de intervalo es únicamente por conveniencia y brevedad y no debería interpretarse como una limitación inflexible en el alcance de la invención. En consecuencia, debería considerarse que la descripción de un intervalo ha desvelado específicamente todos los posibles subintervalos así como valores numéricos individuales en ese intervalo. Por ejemplo, debería considerarse que la descripción de un intervalo tal como de 1 a 6 tiene subintervalos específicamente desvelados tales como de 1 a 3, de 1 a 4, de 1 a 5, de 2 a 4, de 2 a 6, de 3 a 6, etc., así como números individuales dentro de ese intervalo, por ejemplo, 1, 2, 3, 4, 5 y 6. Esto es aplicable independientemente de la amplitud del intervalo.

Siempre que se indique un intervalo numérico en el presente documento, se entiende que incluye cualquier número citado (fraccional o integral) dentro del intervalo indicado. Las expresiones "que varía/varía entre" un primer número indicado y un segundo número indicado y "que varía/varía de" un primer número indicado "a" un segundo número indicado se usan en el presente documento indistintamente y se entiende que incluyen los primer y segundo números indicados y todos los números fraccionales e integrales entre ellos.

Como se usa en el presente documento, el término "método" se refiere a formas, medios, técnicas y procedimientos para realizar una tarea dada, incluyendo, pero sin limitación, las formas, los medios, las técnicas y los procedimientos conocidos o fácilmente desarrollados a partir de formas, medios, técnicas y procedimientos conocidos por profesionales de las técnicas química, farmacológica, biológica, bioquímica y médica.

Como se usa en el presente documento, el término "tratar" incluye anular, sustancialmente inhibir, ralentizar o invertir

la progresión de una afección, aliviar sustancialmente síntomas clínicos o estéticos de una afección o prevenir sustancialmente la aparición de síntomas clínicos o estéticos de una afección.

Se apreciará que determinadas características de la invención, que se describen, para mayor claridad, en el contexto de realizaciones separadas, también pueden proporcionarse en combinación en una única realización. Por el contrario, diversas características de la invención, que se describen, para mayor brevedad, en el contexto de una única realización, también se pueden proporcionar por separado o en cualquier subcombinación adecuada o según sea adecuado en cualquier otra realización descrita de la invención. Determinadas características descritas en el contexto de diversas realizaciones no deben considerarse características esenciales de esas realizaciones, a menos que la realización sea inoperante sin esos elementos.

Diversas realizaciones y aspectos de la presente invención como se han delineado anteriormente en el presente documento y como se reivindica en la sección de reivindicaciones posterior encuentran apoyo experimental en los siguientes ejemplos.

Ejemplos

A continuación, se hace referencia a los siguientes ejemplos, que junto con las descripciones anteriores ilustran algunas realizaciones de la invención de una manera no limitante.

En general, la nomenclatura usada en el presente documento y los procedimientos de laboratorio utilizados en la presente invención incluyen técnicas moleculares, bioquímicas, microbiológicas y de ADN recombinante. Dichas técnicas se explican exhaustivamente en la bibliografía. Véase, por ejemplo, "Molecular Cloning: A laboratory Manual" Sambrook *et al.*, (1989); "Current Protocols in Molecular Biology" volúmenes I-III Ausubel, R. M., ed. (1994); Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley and Sons, Baltimore, Maryland (1989); Perbal, "A Practical Guide to Molecular Cloning", John Wiley & Sons, Nueva York (1988); Watson *et al.*, "Recombinant DNA", Scientific American Books, Nueva York; Birren *et al.* (eds) "Genome Analysis: A Laboratory Manual Series", vol. 1-4, Cold Spring Harbor Laboratory Press, Nueva York (1998); metodologías expuestas en las patentes de los Estados Unidos n.º 4.666.828; 4.683.202; 4.801.531; 5.192.659 y 5.272.057; "Cell Biology: A Laboratory Handbook", volúmenes I-III Cellis, J. E., ed. (1994); "Culture of Animal Cells - A Manual of Basic Technique" de Freshney, Wiley-Liss, N. Y. (1994), Tercera Edición; "Current Protocols in Immunology" volúmenes I-III Coligan J. E., ed. (1994); Stites *et al.* (eds), "Basic and Clinical Immunology" (8ª edición), Appleton y Lange, Norwalk, CT (1994); Mishell y Shiigi (eds), "Selected Methods in Cellular Immunology", W. H. Freeman and Co., Nueva York (1980); se describen exhaustivamente en la bibliografía de patentes y científica inmunoensayos disponibles, véase, por ejemplo, patentes de los Estados Unidos n.º 3.791.932; 3.839.153; 3.850.752; 3.850.578; 3.853.987; 3.867.517; 3.879.262; 3.901.654; 3.935.074; 3.984.533; 3.996.345; 4.034.074; 4.098.876; 4.879.219; 5.011.771 y 5.281.521; "Oligonucleotide Synthesis" Gait, M. J., ed. (1984); "Nucleic Acid Hybridization" Hames, B. D. y Higgins S. J., eds. (1985); "Transcription and Translation" Hames, B. D. y Higgins S. J., eds. (1984); "Animal Cell Culture" Freshney, R. I., ed. (1986); "Immobilized Cells and Enzymes" IRL Press, (1986); "A Practical Guide to Molecular Cloning" Perbal, B., (1984) y "Methods in Enzymology" vol. 1, 2, 317, Academic Press; "PCR Protocols: A Guide To Methods And Applications", Academic Press, San Diego, CA (1990); Marshak *et al.*, "Strategies for Protein Purification and Characterization - A Laboratory Course Manual" CSHL Press (1996).

Se proporcionan otras referencias generales a lo largo de este documento. Se cree que los procedimientos en el mismo son bien conocidos en la técnica y se proporcionan por conveniencia del lector.

Ejemplo i: construcción de la construcción de expresión de α -galactosidasa de planta

El ADNc que codificaba la proteína α -galactosidasa humana (EC 3.2.1-22 GenBank: X05790) fue optimizado y sintetizado por GENEART AG (Regensburg, Alemania). El uso codónico sin el péptido líder (péptido señal de direccionamiento al retículo endoplásmico) se adaptó al sesgo codónico de genes de *Nicotiana tabacum*. Durante el proceso de optimización se evitaron los siguientes motivos de secuencia de acción en cis: cajas TATA internas, sitios chi y sitios de entrada ribosómicos, tramos de secuencias ricas en AT o ricas en GC, elementos de inestabilidad de ARN ("motivos asesinos"), secuencias repetidas y estructuras secundarias de ARN, sitios donantes (crípticos) y aceptores de corte y empalme, puntos de ramificación. Además, se han evitado regiones de contenido de GC muy alto (> 80 %) o muy bajo (<30 %).

La secuencia de nucleótidos del péptido líder de α -galactosidasa humana nativa (péptido señal de direccionamiento al retículo endoplásmico) (figura 1, primeros 31 aminoácidos en rojo) de proteína α -galactosidasa humana de longitud completa (GenBank: X05790) se reemplazó con una secuencia de nucleótidos que codificaba el péptido señal de direccionamiento al retículo endoplásmico de 33 aminoácidos (péptido líder) de la proteína ABPI de *Arabidopsis* (marcado en verde en la figura 1, SEQ ID NO: 4). Este péptido señal proporciona direccionamiento eficaz de α -galactosidasa a la ruta secretora y se escinde del polipéptido, por peptidasa señal, una vez que la proteína se ha translocado al retículo endoplásmico. Se añadió una secuencia de nucleótidos (SEQ ID NO: 19) que codificaba la señal de retención en el retículo endoplásmico SEKDEL (SEQ ID NO: 6) a la secuencia de ADNc en el extremo 3', lo que permite la recuperación de la proteína expresada del aparato de Golgi, manteniendo eficazmente la proteína en el retículo endoplásmico. La SEQ ID NO: 2 representa la secuencia codificante completa de la α -galactosidasa humana

recombinante expresada (SEQ ID NO: 1), incluyendo el péptido señal de direccionamiento al retículo endoplásmico de ABPI N-terminal, proteína α -galactosidasa humana y la señal de retención en el retículo endoplásmico SEKDEL.

Ejemplo ii: expresión de α -galactosidasa humana recombinante en plantas

Sistema de expresión transitoria en *N. benthamiana*

El uso de vectores víricos de plantas se eligió en este caso como una alternativa a las plantas transgénicas, permitiendo la expresión transitoria, rápida, de alto nivel de proteínas en plantas completas maduras.

La proteína de interés se expresó a partir de un fuerte promotor de proteína de cubierta vírica subgenómica. El sistema se basa en la amplificación transitoria (por agroinfección) de vectores víricos suministrados a una planta por *Agrobacterium*, en el que el promotor funcional de la planta y el ADNc que codifica un replicón vírico se transfieren como ADN-T de *Agrobacterium* a células vegetales. El ADN-T se transcribe en la planta por el promotor de la planta para generar ARN vírico biológicamente activo que inicia la autorreplicación.

Para la expresión transitoria se empleó un sistema de recombinación de 3 vectores basado en el sistema desarrollado previamente como se ha descrito (Gleba *et al.*, Vaccine 23 2042-2048, 2005), se insertó ADNc de α -galactosidasa en uno de los vectores y los otros dos vectores que contenían genes para la construcción del replicón vírico completo (RdRp e integrasa), generando de este modo un ARN vírico biológicamente activo capaz de iniciar la autorreplicación.

Transfección de plantas completas-

Se germinaron plantas de *N. benthamiana* y se cultivaron en suelo de mezcla comercial (Givaat Ada, IL) complementado con fertilizante granular de liberación lenta (Scott Marysville, OH) en un régimen de luz de día largo (16 h de luz/8 h de oscuridad) a 24 °C-25 °C.

Las agrobacterias se transformaron con el sistema de vector de replicón basado en pICH20866-alfa-GAL usando electroporación (2500 V, 5 ms) [den Dulk-Ra, A. y Hooykaas, P. J. (1995) Métodos Mol. Biol. 55:63-72], Las plantas se infiltraron con *Agrobacterium* que contenían los 3 plásmidos ICON por infiltración al vacío con métodos convencionales conocidos en la técnica. En resumen, se infiltraron plantas *N. benthamiana* de 5-6 semanas de edad sumergiendo todos los órganos aéreos de la planta en una suspensión bacteriana y se colocaron en una cámara de vacío. Se aplicó un vacío negativo (-) de 80 kPa (0,8 bar) durante 1 minuto, seguido de un retorno rápido a la presión atmosférica. Las plantas se devolvieron al invernadero durante 5-7 días adicionales en las mismas condiciones de crecimiento.

Extracto de planta de tabaco:

Se recogieron muestras de hojas de *Nicotiana benthamiana* 5 días después de la infiltración y se extrajeron en tampón Laemmli para SDS-PAGE o en tampón de ensayo de actividad (ácido cítrico 20 mM, fosfato sódico 30 mM, BSA al 0,1 % y etanol al 0,67 %, pH 4,6) para ensayo de actividad catalítica de la proteína expresada en planta.

Purificación de extracto de planta de tabaco:

Se purificó proteína α -galactosidasa humana a partir de extractos vegetales mediante una precipitación diferencial con sulfato de amonio en dos etapas ("precipitación por adición de sales": 1ª etapa 0,57 M; 2ª etapa 2,27 M), seguido de cromatografía de interacción hidrófoba (resina de fenilo 650 M) y cromatografía de intercambio catiónico.

Expresión estable en células de *N. tobaccum* BY2

La transformación mediada por *Agrobacterium* se usa ampliamente para introducir genes extraños en un genoma de células vegetales. Esta técnica se basa en la capacidad natural del *Agrobacterium* para transformar células vegetales mediante la transferencia de un segmento de ADN plasmídico, el ADN transferido (ADN-T), en el genoma de la célula hospedadora. Usando este enfoque, una molécula de ADN-T, que consiste en un gen extraño y sus elementos reguladores, se introduce aleatoriamente en el genoma de la planta. El sitio de integración, así como el número de copias de las inserciones génicas, no está controlado, por tanto, el proceso de transformación da como resultado un 'grupo' de células transgénicas compuestas de células con diversos niveles de expresión del transgén. El 'grupo' transgénico se usa posteriormente para aislamiento de clones. El aislamiento de clones da como resultado el establecimiento de muchas líneas celulares individuales, a partir de las cuales se selecciona después el clon con el mayor nivel de expresión del gen extraño.

El cultivo en suspensión de BY2 se cocultivó, durante 48 horas, con la cepa de *Agrobacterium tumefaciens* EHA105 que porta el vector que alberga el gen de α -GAL y el gen de selección de neomicina fosfotransferasa (NPTII).

Posteriormente, las células se mantuvieron en medios complementados con 50 mg/l de kanamicina y 250 mg/l de cefotaxima. El gen de NPTII confiere resistencia a la kanamicina, por tanto, solo las células BY2 positivas para NPTII

sobreviven en este medio de selección. La cefotaxima se usó para matar selectivamente la agrobacteria, siendo las células vegetales resistentes a este antibiótico. Una vez que se estableció una suspensión celular transgénica de crecimiento estable, esta se usó para explorar y aislar líneas celulares individuales. Para permitir la selección de líneas celulares individuales, se extendieron alícuotas de suspensión celular muy diluida en medio BY-2 sólido. Las células se cultivaron después hasta que se desarrollaron callos pequeños. Cada callo se resuspendió después en cultivo líquido. Después se tomaron muestras de las células y las muestras se evaluaron para determinar los niveles de α -GAL. Las líneas que expresaron niveles de actividad de α -GAL relativamente altos se volvieron a analizar y se compararon para determinar los niveles de α -GAL que terminaban con la selección final de líneas de expresión de α -GAL candidatas.

Extracto de células de tabaco:

Se homogeneizaron 100 mg de células BY2 transformadas que expresaban la proteína α -galactosidasa humana durante 5 minutos en 200 μ l de tampón de fosfato 20 mM 0,5 M que contenía EDTA 20 mM, DTT 2 mM y PMSF 2 mM, pH=7,2. El homogeneizado se centrifugó y se recogió el sobrenadante que contenía la proteína α -galactosidasa humana recombinante [extracto en bruto], se determinó la actividad catalítica de α -galactosidasa en bruto mediante ensayo de PNP-G.

Ensayo de PNP-G

Se determinó la actividad de α -galactosidasa galactósido hidrolasa usando p-nitrofenilalfa-D-galactopiranosido (pNP-alfa-D-Gal, GBB1290, IRIS Biotech, Alemania) como sustrato de hidrólisis. El tampón de ensayo contenía ácido cítrico 20 mM, fosfato de sodio 30 mM, BSA al 0,1 % y etanol al 0,67 % a pH 4,6. Se incubaron muestras (50 μ l) con 150 μ l de tampón de ensayo. El sustrato (pNP-alfa-D-Gal) se añadió (30 μ l) a una concentración final de p-NP-alfa-D-Gal 8 mM. La mezcla de reacción se incubó a 37 °C durante 90 minutos. Después de 90 minutos, la reacción se interrumpió mediante la adición de 100 μ l de carbonato de sodio (2 M) a cada pocillo, para terminar la reacción y permitir la generación del cromóforo p-nitrofenolato. Los resultados se calcularon a partir de la absorbancia a 405 nm.

Modificación de la α -galactosidasa humana expresada en plantas: La secuenciación del producto proteico de las plantas transformadas reveló que, en algunos casos, tras la escisión de la señal de ABPI, la proteína α -galactosidasa humana recombinante retiene el último aminoácido glicina del péptido señal en el extremo N de la proteína α -galactosidasa madura. Por tanto, la proteína α -galactosidasa humana recombinante puede estar sin la glicina del péptido señal en el extremo N (p. ej., SEQ ID NO: 20, secuencia codificante como en la SEQ ID NO: 22) o incluir una glicina en el extremo N (p. ej., SEQ ID NO: 21, secuencia codificante como en la SEQ ID NO: 23).

Análisis de glucanos de la α -galactosidasa humana expresada en plantas: Se redujeron muestras de producto de proteína recombinante de α -galactosidasa humana de las plantas transformadas, se alquilaron con DTT y yodoacetamida y se separaron en un SDS-PAGE al 12,5 %. Se escindieron bandas correspondientes al peso molecular correcto, se extrajeron las proteínas y se sometieron a análisis de glucanos que consiste en digestión con tripsina seguida de digestión con PNGasa A y PNGasa F. La digestión con PNGasa A libera todos los glucanos ligados a N y la PNGasa F libera todos los glucanos, excepto los que contienen fucosa de núcleo α 1-3.

Los glucanos libres se extrajeron, se purificaron y después se marcaron con el reactivo fluorescente antranilamida (2AB) seguido de la eliminación del exceso de 2AB.

Los glucanos se separaron usando una HPLC de fase normal TSKgel Amide 80 y se detectaron usando un detector de fluorescencia. Un hidrolizado de dextrano actuó como escalera para el cálculo de los valores de unidades de glucosa (UG). El perfil de glucanos se construye calculando las áreas de pico relativas a partir del cromatograma de la digestión con PNGasa A. La asignación de los glucanos se establece mediante el cálculo de los valores de UG de los picos hallados tanto en digestiones de endoglucósidos como según diversas digestiones de exoglucósidos adicionales.

Resultados:

Comparando entre diferentes construcciones de α -galactosidasa, se descubrió que la adición de un péptido señal de direccionamiento al retículo endoplásmico de ABPI (SEQ ID NO: 4) al polipéptido de α -galactosidasa recombinante humano mejoraba la expresión de la proteína recombinante en plantas *N. benthamiana* con expresión transitoria (figura 2, carril 2), en comparación con secuencias de α -galactosidasa expresadas con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (péptido señal de direccionamiento al RE de pectinasa de manzana codificado por la SEQ ID NO: 10) (α -galactosidasa-péptido señal de RE de pectinasa de manzana-proteína codificada por la SEQ ID NO: 8 o la SEQ ID NO: 11), o la proteína α -galactosidasa expresada con un péptido señal de direccionamiento al retículo endoplásmico de endoquitinasa B (proteína señal de endoquitinasa, SEQ ID NO: 14, codificada por la SEQ ID NO: 15). El ensayo de la actividad catalítica (véase anteriormente) confirmó que el aumento de proteína visto en la figura 2 se refleja en el aumento simultáneo de la actividad enzimática (figura 3, columna B) por kg de hojas, en comparación con secuencias de α -galactosidasa expresadas con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana o proteína α -galactosidasa expresada con un péptido señal de

direccionamiento al retículo endoplásmico de endoquitinasa B (proteína señal de endoquitinasa, SEQ ID NO: 14, codificada por la SEQ ID NO: 15).

El análisis de glucanos del producto de proteína α -galactosidasa recombinante humana de las células de tabaco transformadas que expresan una proteína α -galactosidasa humana con la adición de un péptido señal de direccionamiento al retículo endoplásmico de ABPI (p. ej., SEQ ID NO: 1, codificada por la SEQ ID NO: 2) reveló un perfil de glucano que comprende estructuras de glucano principales de estructuras de glucano con alto contenido en manosa (p. ej., de 3 a 6, 7, 8 o 9 restos de manosa, con manosa expuesta terminal) y glucanos específicos de plantas característicos, tales como β -(1,2) xilosa y α -(1,3) fucosa (figura 7). Las estructuras de glucano predominantes que representan un alto porcentaje del perfil de glucano incluían aproximadamente 30 % que tenían manosa 3- β -(1,2) xilosa- α -(1,3) fucosa [Fc(3)M3X] y/o manosa 4- α -(1,2) xilosa (M4X); y otro aproximadamente 30 % de glucanos de manosa 8 (M8) (figura 8). La figura 8 es una tabla ilustrada que muestra la distribución de las estructuras de glucano dentro de las muestras de proteínas de las células transformadas. A-Gal A y A-Gal-B son dos análisis representativos del producto de proteína α -galactosidasa recombinante humana de las células vegetales transformadas.

Las estructuras de glucano pueden clasificarse adicionalmente según los restos de azúcar presentes en los glucanos complejos. La tabla V a continuación detalla la distribución de estructuras de glucano en el producto de proteína α -galactosidasa humana de plantas transformadas, según azúcares específicos:

Tabla V-Contenido de sacárido de glucanos

TIPO DE GLUCANO	% DE GLUCANOS TOTALES
Glucanos que contiene fucosa	~ 35
Glucanos que contienen xilosa	~ 50
Manosa alta	~ 60
Glucanos que contienen N-acetilglucosamina	~ 10

Por tanto, el producto de proteína α -galactosidasa humana expresado en cultivo de células de tabaco a partir de la construcción que comprende la SEQ ID NO: 2 se caracterizó por un porcentaje significativo de estructuras de glucanos específicos de plantas (que tienen componentes de fucosa y xilosa) y alta glucosilación de manosa alta.

Ejemplo iii-estabilidad de α -galactosidasa recombinante humana expresada en planta en condiciones in vivo

Para aproximar la farmacocinética de la α -galactosidasa recombinante humana expresada en planta, se expusieron preparaciones de enzimas recombinantes a condiciones *in vivo* simuladas y actividad catalítica supervisada.

Estabilidad en plasma:

La estabilidad de la actividad enzimática en plasma se evaluó añadiendo extracto vegetal recombinante o α -galactosidasa purificada de extracto vegetal a plasma humano a una concentración final de 1 μ g/ml (p. ej., 10 μ l de extracto o α -galactosidasa purificada a 190 μ l de plasma humano), incubando a 37 °C y siguiendo la actividad catalítica de α -galactosidasa en muestras a intervalos establecidos durante una hora.

La comparación entre α -galactosidasa de plantas *N. benthamiana* transformadas con la SEQ ID NO: 11, que codifica una α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (SEQ ID NO: 12) y α -galactosidasa de plantas *N. benthamiana* transformadas con la SEQ ID NO: 2, que codifica α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de ABPI (SEQ ID NO: 1) (figura 4) indica una clara ventaja de la α -galactosidasa expresada por *N. benthamiana* de la construcción que tiene un péptido señal de direccionamiento al retículo endoplásmico de ABPI (figura 4, cuadrados cerrados o sólidos), en comparación con α -galactosidasa recombinante expresada a partir de la construcción con un líder de señal de direccionamiento al retículo endoplásmico de pectinasa de manzana (figura 4, cuadrados abiertos).

La comparación entre α -galactosidasa expresada en células de plantas de *N. tabacuum* (BY2) transformadas con la SEQ ID NO: 2, que codifica una proteína α -galactosidasa con un péptido señal de direccionamiento al retículo endoplásmico de ABPI (SEQ ID NO: 1), dando como resultado una proteína α -galactosidasa recombinante madura como se expone en la SEQ ID NO: 21 con α -galactosidasa expresada en células de mamífero (REPLAGAL™) reveló estabilidad similar de la actividad enzimática cuando se incubó durante una hora a 37 °C en plasma (figura 5) y cuando se incubó durante 9 días en condiciones similares a las lisosómicas (pH=4,6 y 37 °C, figura 6). Esta similitud en condiciones *in vivo* simuladas sugiere que la α -galactosidasa recombinante expresada en células vegetales tendría farmacocinética comparable a la de la enzima expresada en células de mamíferos, clínicamente aprobada (REPLAGAL™).

Además, la cita o identificación de cualquier referencia en esta solicitud no debería interpretarse como una admisión de que dicha referencia está disponible como técnica anterior a la presente invención. En la medida en que se usan encabezamientos de sección, no deberían interpretarse como necesariamente limitantes.

ES 2 774 190 T3

LISTADO DE SECUENCIAS

- 5 <110> Protalix Ltd. Shulman, Avidor Hanania, Uri Kizhner, Tali Shaaltiel, Yosef
- <120> CONSTRUCCIÓN DE ÁCIDO NUCLEICO PARA LA EXPRESIÓN DE ALFA-GALACTOSIDASA EN PLANTAS Y CÉLULAS VEGETALES
- 10 <130> 51336
- <150> PCT/IL2011/000209
<151> 02/03/2011
- 15 <150> US 61/434.499
<151> 20/01/2011
- <150> US 61/434.503
<151> 20/01/2011
- 20 <160> 23
- <170> PatentIn versión 3.5
- 25 <210> 1
<211> 437
<212> PRT
<213> Secuencia artificial
- 30 <220>
<223> Alfa-Gal recombinante fusionada con ABPI y un péptido señal del RE (ABPI+AGAL+SEKDEL)
- <400> 1

```

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

Val Phe Ser Val Ala Leu Leu Leu Phe Tyr Phe Ser Glu Thr Ser Leu
          20           25           30

Gly Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His
          35           40           45

Trp Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser
          50           55           60

Cys Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser
          65           70           75           80

Glu Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys
          85           90           95

Trp Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro
          100          105          110
    
```

ES 2 774 190 T3

Gln Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser
 115 120 125

Lys Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys
 130 135 140

Ala Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr
 145 150 155 160

Phe Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys
 165 170 175

Asp Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala
 180 185 190

Leu Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu
 195 200 205

Tyr Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr
 210 215 220

Cys Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser
 225 230 235 240

Ile Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val
 245 250 255

Asp Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile
 260 265 270

Gly Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu
 275 280 285

Trp Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His
 290 295 300

Ile Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala
 305 310 315 320

Ile Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly
 325 330 335

Asp Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala
 340 345 350

Val Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr

ES 2 774 190 T3

	355		360		365														
	Ile	Ala	Val	Ala	Ser	Leu	Gly	Lys	Gly	Val	Ala	Cys	Asn	Pro	Ala	Cys			
		370					375					380							
	Phe	Ile	Thr	Gln	Leu	Leu	Pro	Val	Lys	Arg	Lys	Leu	Gly	Phe	Tyr	Glu			
	385					390					395					400			
	Trp	Thr	Ser	Arg	Leu	Arg	Ser	His	Ile	Asn	Pro	Thr	Gly	Thr	Val	Leu			
					405					410					415				
	Leu	Gln	Leu	Glu	Asn	Thr	Met	Gln	Met	Ser	Leu	Lys	Asp	Leu	Leu	Ser			
				420					425					430					
	Glu	Lys	Asp	Glu	Leu														
				435															

- <210> 2
- <211> 1311
- 5 <212> ADN
- <213> Secuencia artificial
- <220>
- 10 <223> Secuencia codificante de Alfa-Gal recombinante fusionada con ABPI y un péptido señal del RE (ABPI+AGAL+SEKDEL)
- <400> 2

ES 2 774 190 T3

<213> Secuencia artificial

<220>

<223> Péptido señal de direccionamiento al retículo endoplásmico de ABPI procedente de *Arabidopsis*

5

<400> 4

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

Val Phe Ser Val Ala Leu Leu Leu Phe Tyr Phe Ser Glu Thr Ser Leu
20 25 30

Gly

10

<210> 5

<211> 99

<212> ADN

<213> Secuencia artificial

15

<220>

<223> Secuencia codificante de péptido señal de direccionamiento al RE de ABPI

<400> 5

atgattgtgc tttctgtggg atctgcttct tttctccaa ttgtggtggt gttctctgtg 60

20

gctcttcttc ttttctactt ctctgagact tctcttggc 99

<210> 6

<211> 6

<212> PRT

25

<213> Secuencia artificial

<220>

<223> Péptido señal de retención en el retículo endoplásmico

30

<400> 6

Ser Glu Lys Asp Glu Leu
1 5

35

<210> 7

<211> 398

<212> PRT

<213> *Homo sapiens*

40

<220>

<221> misc_feature

<223> Secuencia de proteína alfa galactosidasa madura humana

<400> 7

ES 2 774 190 T3

Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp
130 135 140

Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu
145 150 155 160

Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr
165 170 175

Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys
180 185 190

Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile
195 200 205

Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp
210 215 220

Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile Gly
225 230 235 240

Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp
245 250 255

Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile
260 265 270

Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala Ile
275 280 285

Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp
290 295 300

Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val
305 310 315 320

Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile
325 330 335

Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe
340 345 350

Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp
355 360 365

Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu

ES 2 774 190 T3

370

375

380

Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu
385 390 395

5 <210> 8
 <211> 1290
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Secuencia codificante de alfa-galactosidasa fusionada con péptido señal del RE de pectinasa de manzana y SEKDEL

<400> 8

atggctctta agactcaatt gctctggtct ttcgttggtg tgttcgtggt ctccctttca 60
 acaactagtt gctccggact tgataatgga cttgctagga ctccaactat gggatggctt 120
 cattgggaac gtttcatgtg caatctcgat tgtcaagagg aaccagattc ttgcatttcc 180
 gagaagcttt tcatggaat ggctgagctt atggtttctg agggatggaa ggatgctgga 240
 tatgagtacc tctgcatcga tgattgctgg atggctccac aaagagattc tgaggggaagg 300
 cttcaagctg atccacaaag gttcccacat ggaattaggc aactcgctaa ctacgttcac 360
 tctaagggac ttaagcttgg aatctacgct gatgtgggaa acaagacttg tgctggattt 420
 ccaggatctt tcggttacta cgatatcgat gctcagactt ttgctgattg gggagtggat 480
 cttcttaagt tcgatggatg ctactgtgat tctcttgaga acctcgctga tggttataag 540
 cacatgtctc tcgctcttaa tagaactgga cgttccattg tgtattcttg cgagtggcca 600
 ctttcatatg ggccattcca gaagccaaac tacactgaga ttaggcaata ctgcaacat 660
 tggaggaact tcgctgatat tgatgattcc tggaaagtcta tcaagtccat ccttgattgg 720
 acttcattca atcaagaacg tattgtggat gttgctggac ctggtggatg gaatgatcca 780
 gatatgctcg tgattggaaa ctttgactt tcttggaaacc agcaagttac tcaaatggct 840
 ctctgggcta ttatggctgc tccactcttc atgtctaacg atcttaggca catttctcca 900
 caagctaagg ctttgcctca ggataaggat gtgattgcta tcaaccagga tccacttgg 960
 aagcaaggat accaacttag gcagggtgat aatthtgagg tttgggagag gccactttct 1020
 ggacttgctt gggctggtgc tatgattaac aggcaagaga ttggaggacc aaggtcttac 1080
 actattgctg tggcttctct tggaaagggt gttgcttgta atccagcttg cttcattact 1140
 cagcttttgc cagtgaagag gaagcttgg ttttacgagt ggacttctag gcttaggtca 1200
 cacattaacc caactggaac tgtgcttctt cagctcgaga aactatgca gatgtctctt 1260
 aaggatttgc tctccgagaa ggatgagctt 1290

15 <210> 9
 <211> 26
 <212> PRT

ES 2 774 190 T3

<213> Secuencia artificial

<220>

<223> Secuencia de aminoácidos del péptido señal direccionamiento al RE de pectinasa de manzana

5 <400> 9

Met Ala Leu Lys Thr Gln Leu Leu Trp Ser Phe Val Val Val Phe Val
1 5 10 15

Val Ser Phe Ser Thr Thr Ser Cys Ser Gly
20 25

10 <210> 10
<211> 78
<212> ADN
<213> Secuencia artificial

15 <220>

<223> Secuencia codificante de señal de direccionamiento al RE procedente de pectinasa de manzana

<400> 10

atggctctta agactcaatt gctctggtct ttcgttggtg tgttcgtggt ctccttttca 60

20 acaactagtt gctccgga 78

<210> 11
<211> 1296
<212> ADN

25 <213> Secuencia artificial

<220>

<223> Secuencia codificante de alfa GAL de planta fusionada con péptido líder de pectinasa

30 <400> 11

ES 2 774 190 T3

atggctctta agactcaatt gctctggtct ttcggtggtg tgttcgtggt ctccttttca	60
acaactagtt gctccggaga attccttgat aatggattgg ccaggacacc aactatggga	120
tggctgcatt gggagcgttt tatgtgtaac cttgactgcc aagaagagcc tgattcttgc	180
atctcagaga agctctttat ggaaatggct gagctgatgg tttctgaagg ctggaaagac	240
gcaggttatg agtatttgtg catcgacgat tgctggatgg ctccacagag agatagtgaa	300
gggagacttc aagcagatcc tcagaggttt ccacatggta ttagacagct cgccaattat	360
gttcaactcca agggccttaa gttggggata tacgctgatg tgggaaacaa aacctgtgct	420
ggatttcctg gttccttcgg ctattacgat attgacgcac aaacctttgc tgattgggga	480
gttgacctcc ttaaattcga tggctgttac tgtgattcac ttgagaatct cgccgacggt	540
tataagcaca tgtctctcgc tttgaacagg actggaagga gcatagtcta ctctgtgaa	600
tggccacttt acatgtggcc tttccagaag ccaaactaca ctgagattag gcagttattgc	660
aacctactggc gtaattttgc cgatatcgac gattcttggga agtccatcaa gagcattctg	720
gattggacat ccttcaatca ggaaaggatt gtggatggtg ctggacctgg tggatggaat	780
gatcctgaca tgttgggtgat agggaacttt ggtctctcat ggaatcagca agttacccaa	840
atggcactct gggctattat ggctgcacca ctcttcatgt ctaacgacct caggcatatc	900
tcacctcaag caaaggcatt gctgcaagac aaagacgtca ttgccattaa ccaggaccct	960
cttgggaagc aaggttatca acttcgtcaa ggcgataact ttgaagtgtg ggagagacca	1020
ttgtctggac ttgcttgggc tgttgccatg attaataggc aagaaatcgg tggcccaaga	1080
agttatacaa tcgctgttgc aagtcttggg aaaggtgtcg cttgtaatcc tgctgtttc	1140
attactcagc tcctgcctgt gaaaagaaag ttggggttct atgaatggac ttcaagactc	1200
aggagccata ttaaccctac cggaacagtt cttctccaac tggagaacac catgcaaatg	1260
tccctcaaag atcttagtga gaaggacgag ctctga	1296

<210> 12

<211> 431

<212> PRT

<213> Secuencia artificial

<220>

<223> Alfa GAL de planta fusionada con péptido líder de pectinasa, secuencia de aminoácidos

<400> 12

5

10

ES 2 774 190 T3

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

ES 2 774 190 T3

115					120					125					
Gly	Ile	Tyr	Ala	Asp	Val	Gly	Asn	Lys	Thr	Cys	Ala	Gly	Phe	Pro	Gly
130						135					140				
Ser	Phe	Gly	Tyr	Tyr	Asp	Ile	Asp	Ala	Gln	Thr	Phe	Ala	Asp	Trp	Gly
145					150					155					160
Val	Asp	Leu	Leu	Lys	Phe	Asp	Gly	Cys	Tyr	Cys	Asp	Ser	Leu	Glu	Asn
				165					170					175	
Leu	Ala	Asp	Gly	Tyr	Lys	His	Met	Ser	Leu	Ala	Leu	Asn	Arg	Thr	Gly
			180					185					190		
Arg	Ser	Ile	Val	Tyr	Ser	Cys	Glu	Trp	Pro	Leu	Tyr	Met	Trp	Pro	Phe
		195					200					205			
Gln	Lys	Pro	Asn	Tyr	Thr	Glu	Ile	Arg	Gln	Tyr	Cys	Asn	His	Trp	Arg
	210					215					220				
Asn	Phe	Ala	Asp	Ile	Asp	Asp	Ser	Trp	Lys	Ser	Ile	Lys	Ser	Ile	Leu
225					230					235					240
Asp	Trp	Thr	Ser	Phe	Asn	Gln	Glu	Arg	Ile	Val	Asp	Val	Ala	Gly	Pro
				245					250					255	
Gly	Gly	Trp	Asn	Asp	Pro	Asp	Met	Leu	Val	Ile	Gly	Asn	Phe	Gly	Leu
			260					265					270		
Ser	Trp	Asn	Gln	Gln	Val	Thr	Gln	Met	Ala	Leu	Trp	Ala	Ile	Met	Ala
		275					280					285			
Ala	Pro	Leu	Phe	Met	Ser	Asn	Asp	Leu	Arg	His	Ile	Ser	Pro	Gln	Ala
	290					295					300				
Lys	Ala	Leu	Leu	Gln	Asp	Lys	Asp	Val	Ile	Ala	Ile	Asn	Gln	Asp	Pro
305					310					315					320
Leu	Gly	Lys	Gln	Gly	Tyr	Gln	Leu	Arg	Gln	Gly	Asp	Asn	Phe	Glu	Val
				325					330					335	
Trp	Glu	Arg	Pro	Leu	Ser	Gly	Leu	Ala	Trp	Ala	Val	Ala	Met	Ile	Asn
			340					345					350		
Arg	Gln	Glu	Ile	Gly	Gly	Pro	Arg	Ser	Tyr	Thr	Ile	Ala	Val	Ala	Ser
		355					360					365			

ES 2 774 190 T3

Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile Thr Gln Leu
370 375 380

Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp Thr Ser Arg Leu
385 390 395 400

Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu Gln Leu Glu Asn
405 410 415

Thr Met Gln Met Ser Leu Lys Asp Leu Ser Glu Lys Asp Glu Leu
420 425 430

- <210> 13
- <211> 1278
- <212> ADN
- <213> Secuencia artificial

- <220>
- <223> Secuencia codificante optimizada de proteína de fusión del péptido señal de quitinasa y alfa GAL II
procedente de planta

- <400> 13

ES 2 774 190 T3

```

atgaagacta atctttttct ctttctcatc ttttcaactc tcctatcatt atcctcggcc      60
gaattccttg ataatggact tgctaggact ccaactatgg gatggcttca ttgggaacgt      120
ttcatgtgca atctcgattg tcaagaggaa ccagattcct gcatttccga gaagcttttc      180
atggaaatgg ctgagcttat ggtttctgag ggatggaagg atgctggata tgagtacctc      240
tgcacgcgatg attgctggat ggctccacaa agagattctg agggaaggct tcaagctgat      300
ccacaaaggt tcccacatgg aattaggcaa ctcgctaact acgttcactc taagggactt      360
aagcttgaa tctacgctga tgtgggaaac aagacttggt ctggatttcc aggatctttc      420
ggttactacg atatcgatgc tcagactttt gctgattggg gagtggatct tcttaagttc      480
gatggatgct actgtgattc tcttgagaac ctcgctgatg gttataagca catgtctctc      540
gctcttaata gaactggacg ttccattgtg tattcttgcg agtggccact ttacatgtgg      600
ccattccaga agccaaacta cactgagatt aggcaatact gcaaccattg gaggaacttc      660
gctgatattg atgattcctg gaagtctatc aagtccatcc ttgattggac ttcattcaat      720
caagaacgta ttgtggatgt tgctggacct ggtggatgga atgatccaga tatgctcgtg      780
attgaaaact ttggactttc ttggaaccag caagttactc aatggctct ctgggctatt      840
atggctgctc cactcttcat gtctaacgat cttaggcaca tttctccaca agctaaggct      900
ttgctccagg ataaggatgt gattgctatc aaccaggatc cacttgaaa gcaaggatac      960
caacttaggc agggtgataa ttttgaggtt tgggagaggc cactttctgg acttgcttgg      1020
gctgttgcta tgattaacag gcaagagatt ggaggaccaa ggtcttacac tattgctgtg      1080
gcttctcttg gaaaggggtg tgcttghtaat ccagcttgct tcattactca gcttttgcca      1140
gtgaagagga agcttggatt ttacgagtgg acttctaggc ttaggtcaca cattaacca      1200
actggaactg tgcttcttca gctcgagaac actatgcaga tgtctcttaa ggatttctc      1260
tccgagaagg atgagctt                                     1278

```

- 5 <210> 14
- <211> 22
- <212> PRT
- <213> Secuencia artificial
- 10 <220>
- <223> Secuencia básica de aminoácidos de péptido señal de direccionamiento al RE de endoquitinasa B
- <400> 14

```

Met Lys Thr Asn Leu Phe Leu Phe Leu Ile Phe Ser Leu Leu Leu Ser
1           5           10           15

Leu Ser Ser Ala Glu Phe
20

```

- 15 <210> 15
- <211> 66
- <212> ADN
- <213> Secuencia artificial

ES 2 774 190 T3

<220>

<223> Secuencia básica codificante de péptido señal de direccionamiento al RE de endoquitinasa B

<400> 15

5

atgaagacta atctttttct ctttctcacc ttttcaacttc tcctatcatt atcctcggcc 60

gaattc 66

<210> 16

<211> 399

10

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia de proteína alfa-Gal madura recombinante con G N' terminal

15

<400> 16

Gly Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His
1 5 10 15

Trp Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser
20 25 30

Cys Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser
35 40 45

ES 2 774 190 T3

Glu Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys
 50 55 60

Trp Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro
 65 70 75 80

Gln Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser
 85 90 95

Lys Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys
 100 105 110

Ala Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr
 115 120 125

Phe Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys
 130 135 140

Asp Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala
 145 150 155 160

Leu Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu
 165 170 175

Tyr Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr
 180 185 190

Cys Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser
 195 200 205

Ile Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val
 210 215 220

Asp Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile
 225 230 235 240

Gly Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu
 245 250 255

Trp Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His
 260 265 270

Ile Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala
 275 280 285

Ile Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly
 290 295 300

ES 2 774 190 T3

Asp Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala
305 310 315 320

Val Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr
325 330 335

Ile Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys
340 345 350

Phe Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu
355 360 365

Trp Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu
370 375 380

Leu Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu
385 390 395

<210> 17
<211> 4
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido señal de retención en el retículo endoplásmico

<400> 17

Lys Asp Glu Leu
1

<210> 18
<211> 1194
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia codificante optimizada para alfa-Gal maduro sin péptidos señal de ABPI o SEKDEL

<400> 18

cttgataatg gacttgctag gactccaact atgggatggc ttcattggga acgtttcatg	60
tgcaatctcg attgtcaaga ggaaccagat tcttgcattt ccgagaagct tttcatggaa	120
atggctgagc ttatggtttc tgagggatgg aaggatgctg gatatgagta cctctgcatc	180
gatgattgct ggatggctcc acaaagagat tctgagggaa ggcttcaagc tgatccacaa	240
aggttccac atggaattag gcaactcgct aactacgttc actctaaggg acttaagctt	300
ggaatctacg ctgatgtggg aaacaagact tgtgctggat ttccaggatc tttcggttac	360
tacgatatcg atgctcagac ttttgctgat tggggagtgg atcttcttaa gttc gatgga	420

ES 2 774 190 T3

```

tgctactgtg attctcttga gaacctcgct gatggttata agcacatgtc tctcgctctt      480
aatagaactg gacgttccat tgtgtattct tgcgagtggc cactttacat gtggccattc      540
cagaagccaa actacactga gattaggcaa tactgcaacc attggaggaa cttcgctgat      600
attgatgatt cctggaagtc tatcaagtcc atccttgatt ggacttcatt caatcaagaa      660
cgtattgtgg atgttgctgg acctggtgga tggaatgatc cagatatgct cgtgattgga      720
aactttggac tttcttgaa ccagcaagtt actcaaatgg ctctctgggc tattatggct      780
gctccactct tcatgtctaa cgatcttagg cacatttctc cacaagctaa ggctttgctc      840
caggataagg atgtgattgc tatcaaccag gatccacttg gaaagcaagg ataccaactt      900
aggcagggtg ataattttga ggtttgggag aggccacttt ctggacttgc ttgggctggt      960
gctatgatta acaggcaaga gattggagga ccaaggtctt aactattgc tgtggcttct     1020
cttgaaagg gtggtgcttg taatccagct tgcttcatta ctcagctttt gccagtgaag     1080
aggaagcttg gattttacga gtggacttct aggcttaggt cacacattaa cccaactgga     1140
actgtgcttc ttcagctcga gaacactatg cagatgtctc ttaaggattt gctc         1194

```

5 <210> 19
 <211> 18
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Secuencia codificante de señal de retención en el RE SEKDEL

<400> 19
 tccgagaagg atgagctt 18

15 <210> 20
 <211> 404
 <212> PRT
 <213> Secuencia artificial

20 <220>
 <223> Proteína madura A-Gal recombinante fusionada con SEKDEL sin G N' terminal

<400> 20

```

Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp
1           5           10           15

```

```

Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys
          20           25           30

```

```

Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu
          35           40           45

```

```

Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp

```

ES 2 774 190 T3

50					55					60					
Met	Ala	Pro	Gln	Arg	Asp	Ser	Glu	Gly	Arg	Leu	Gln	Ala	Asp	Pro	Gln
65					70					75					80
Arg	Phe	Pro	His	Gly	Ile	Arg	Gln	Leu	Ala	Asn	Tyr	Val	His	Ser	Lys
				85					90					95	
Gly	Leu	Lys	Leu	Gly	Ile	Tyr	Ala	Asp	Val	Gly	Asn	Lys	Thr	Cys	Ala
			100					105					110		
Gly	Phe	Pro	Gly	Ser	Phe	Gly	Tyr	Tyr	Asp	Ile	Asp	Ala	Gln	Thr	Phe
		115					120					125			
Ala	Asp	Trp	Gly	Val	Asp	Leu	Leu	Lys	Phe	Asp	Gly	Cys	Tyr	Cys	Asp
	130					135					140				
Ser	Leu	Glu	Asn	Leu	Ala	Asp	Gly	Tyr	Lys	His	Met	Ser	Leu	Ala	Leu
145					150					155					160
Asn	Arg	Thr	Gly	Arg	Ser	Ile	Val	Tyr	Ser	Cys	Glu	Trp	Pro	Leu	Tyr
				165					170					175	
Met	Trp	Pro	Phe	Gln	Lys	Pro	Asn	Tyr	Thr	Glu	Ile	Arg	Gln	Tyr	Cys
			180					185					190		
Asn	His	Trp	Arg	Asn	Phe	Ala	Asp	Ile	Asp	Asp	Ser	Trp	Lys	Ser	Ile
		195					200					205			
Lys	Ser	Ile	Leu	Asp	Trp	Thr	Ser	Phe	Asn	Gln	Glu	Arg	Ile	Val	Asp
	210					215					220				
Val	Ala	Gly	Pro	Gly	Gly	Trp	Asn	Asp	Pro	Asp	Met	Leu	Val	Ile	Gly
225					230					235					240
Asn	Phe	Gly	Leu	Ser	Trp	Asn	Gln	Gln	Val	Thr	Gln	Met	Ala	Leu	Trp
				245					250					255	
Ala	Ile	Met	Ala	Ala	Pro	Leu	Phe	Met	Ser	Asn	Asp	Leu	Arg	His	Ile
			260					265					270		
Ser	Pro	Gln	Ala	Lys	Ala	Leu	Leu	Gln	Asp	Lys	Asp	Val	Ile	Ala	Ile
		275					280					285			
Asn	Gln	Asp	Pro	Leu	Gly	Lys	Gln	Gly	Tyr	Gln	Leu	Arg	Gln	Gly	Asp
290						295					300				

ES 2 774 190 T3

Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val
305 310 315 320

Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile
325 330 335

Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe
340 345 350

Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp
355 360 365

Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu
370 375 380

Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu Ser Glu
385 390 395 400

Lys Asp Glu Leu

<210> 21

<211> 405

<212> PRT

<213> Secuencia artificial

<220>

<223> Proteina madura A-Gal recombinante con G N' terminal y fusionada con SEKDEL

<400> 21

Gly Leu Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His
1 5 10 15

Trp Glu Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser
20 25 30

Cys Ile Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser
35 40 45

Glu Gly Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys
50 55 60

Trp Met Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro
65 70 75 80

Gln Arg Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser
85 90 95

5

10

ES 2 774 190 T3

Lys Gly Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys
 100 105 110

Ala Gly Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr
 115 120 125

Phe Ala Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys
 130 135 140

Asp Ser Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala
 145 150 155 160

Leu Asn Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu
 165 170 175

Tyr Met Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr
 180 185 190

Cys Asn His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser
 195 200 205

Ile Lys Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val
 210 215 220

Asp Val Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Val Ile
 225 230 235 240

Gly Asn Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu
 245 250 255

Trp Ala Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His
 260 265 270

Ile Ser Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Val Ile Ala
 275 280 285

Ile Asn Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly
 290 295 300

Asp Asn Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala
 305 310 315 320

Val Ala Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr
 325 330 335

Ile Ala Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys
 340 345 350

ES 2 774 190 T3

Phe Ile Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu
 355 360 365

Trp Thr Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu
 370 375 380

Leu Gln Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu Ser
 385 390 395 400

Glu Lys Asp Glu Leu
 405

<210> 22
 <211> 1212
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia codificante optimizada de proteína A-Gal madura fusionada con SEKDEL sin ABPI

<400> 22

cttgataatg gacttgctag gactccaact atgggatggc ttcattggga acgtttcatg 60
 tgcaatctcg attgtcaaga ggaaccagat tcttgcatth cggagaagct tttcatggaa 120
 atggctgagc ttatggtttc tgagggatgg aaggatgctg gatatgagta cctctgcatc 180
 gatgattgct ggatggctcc acaaagagat tctgagggaa ggcttcaagc tgatccacaa 240
 aggttcccac atggaattag gcaactcgct aactacgttc actctaaggg acttaagctt 300
 ggaatctacg ctgatgtggg aaacaagact tgtgctggat ttccaggatc tttcggttac 360
 tacgatatcg atgctcagac ttttgctgat tggggagtgg atcttcttaa gttcgatgga 420
 tgctactgtg atttcttga gaacctcgct gatggttata agcacatgtc tctcgctctt 480
 aatagaactg gacgttccat tgtgtattct tgcgagtggc cactttacat gtggccattc 540
 cagaagccaa actacactga gattaggcaa tactgcaacc attggaggaa cttcgctgat 600
 attgatgatt cctggaagtc tatcaagtcc atccttgatt ggacttcatt caatcaagaa 660
 cgtattgtgg atgttgctgg acctggtgga tggaatgac cagatatgct cgtgattgga 720
 aactttggac tttcttgaa ccagcaagtt actcaaatgg ctctctgggc tattatggct 780
 gctccactct tcatgtctaa cgatcttagg cacatttctc cacaagctaa ggctttgctc 840
 caggataagg atgtgattgc tatcaaccag gatccacttg gaaagcaagg ataccaactt 900
 aggcagggtg ataattttga ggtttgggag aggccacttt ctggacttgc ttgggctggt 960
 gctatgatta acaggcaaga gattggagga ccaaggtctt aactatttgc tgtggcttct 1020
 cttggaaagg gtgttgcttg taatccagct tgcttcatta ctgagctttt gccagtgaag 1080

ES 2 774 190 T3

aggaagcttg gattttacga gtggacttct aggcttaggt cacacattaa cccaactgga 1140
 actgtgcttc ttcagctoga gaacactatg cagatgtctc ttaaggattt gctctccgag 1200
 aaggatgagc tt 1212

5 <210> 23
 <211> 1215
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Secuencia codificante optimizada de proteína A-Gal madura con G N' terminal y SEKDEL

<400> 23
 ggccttgata atggacttgc taggactcca actatgggat ggcttcattg ggaacgtttc 60
 atgtgcaatc tcgattgtca agaggaacca gattcttgca tttccgagaa gcttttcatg 120
 gaaatggctg agcttatggg ttctgagggg tggaggatg ctggatatga gtacctctgc 180
 atc gatgatt gctggatggc tccacaaaga gattctgagg gaaggcttca agctgatcca 240
 caaaggttcc cacatggaat taggcaactc gctaactacg ttcactctaa gggacttaag 300
 cttggaatct acgctgatgt gggaaacaag acttgtgctg gatttccagg atctttcggg 360
 tactacgata tcgatgctca gacttttgct gattggggag tggatcttct taagttcgat 420
 ggatgctact gtgattctct tgagaacctc gctgatgggt ataagcacat gtctctcgct 480
 cttaatagaa ctggacgttc cattgtgtat tcttgogagt ggccacttta catgtggcca 540
 ttccagaagc caaactacac tgagattagg caatactgca accattggag gaacttcgct 600
 gatattgatg attcctggaa gtctatcaag tccatccttg attggacttc attcaatcaa 660
 gaacgtattg tggatgttgc tggacctggt ggatggaatg atccagatat gctcgtgatt 720
 ggaaactttg gactttcttg gaaccagcaa gttactcaaa tggctctctg ggctattatg 780
 gctgctccac tcttcatgtc taacgatctt aggcacattt ctccacaagc taaggctttg 840
 ctccaggata aggatgtgat tgctatcaac caggatccac ttggaaagca aggataccaa 900
 cttaggcagg gtgataatth tgaggtttgg gagaggccac tttctggact tgcttgggct 960
 gttgctatga ttaacaggca agagattgga ggaccaaggt cttacactat tgctgtggct 1020
 tctcttgaa aggggtgttc ttgtaatcca gcttgcttca ttactcagct tttgccagtg 1080
 aagaggaagc ttggatttta cgagtggact tctaggctta ggtcacacat taaccaact 1140
 ggaactgtgc ttcttcagct cgagaacact atgcagatgt ctcttaagga tttgctctcc 1200
 gagaaggatg agctt 1215

REIVINDICACIONES

- 5 1. Una proteína α -galactosidasa humana que comprende una glicina como el resto N-terminal en donde dicha proteína α -galactosidasa humana se fusiona traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico y en donde la proteína α -galactosidasa humana se puede obtener expresando la proteína en una célula vegetal que comprende una construcción de ácido nucleico que comprende una secuencia de ácido nucleico que codifica una proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico.
- 10 2. La proteína α -galactosidasa humana de la reivindicación 1, que comprende la secuencia de aminoácidos de la proteína α -galactosidasa humana como se expone en la SEQ ID NO: 21.
- 15 3. La proteína α -galactosidasa humana de la reivindicación 1 o 2, en donde dicha señal de retención en el retículo endoplásmico tiene una secuencia de aminoácidos como se expone en la SEQ ID NO: 6.
- 20 4. La proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-3, en donde dicha secuencia de ácido nucleico que codifica dicha proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico tiene codones optimizados para su expresión en células de planta de tabaco.
- 25 5. La proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-3, en donde dicha secuencia de ácido nucleico que codifica dicha proteína α -galactosidasa humana fusionada traduccionalmente en el extremo N con un péptido señal de direccionamiento al retículo endoplásmico de ABPI de *Arabidopsis* y fusionada traduccionalmente en el extremo C con un péptido señal de retención en el retículo endoplásmico comprende la SEQ ID NO: 2.
- 30 6. La proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-3 en donde dicha proteína α -galactosidasa humana es una proteína α -galactosidasa humana purificada.
- 35 7. La proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-6, en donde dicha proteína α -galactosidasa humana es catalíticamente activa según lo determinado mediante ensayo de p-nitrofenilalfa-D-galactopiranosido.
- 40 8. Una pluralidad de proteínas α -galactosidasa humanas según una cualquiera de las reivindicaciones 1-6, en donde al menos 0,5 % de dicha pluralidad de proteínas α -galactosidasa humanas tiene una estructura de glucano que comprende nueve restos de manosa, en donde tres son restos de manosa expuestos.
- 45 9. La pluralidad de proteínas α -galactosidasa humanas de la reivindicación 8, en donde las estructuras de glucano predominantes de dicha pluralidad de proteínas α -galactosidasas humanas son manosa 4- β -(1,2) xilosa (M4X); manosa 3- β -(1,2) xilosa- α -(1,3) fucosa [Fc(3)M3X] y manosa 8 (M8) y un vehículo farmacéuticamente aceptable.
- 50 10. Una célula vegetal aislada que comprende la proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-7 o la pluralidad de proteínas α -galactosidasa humanas de la reivindicación 8 o 9.
- 55 11. La célula vegetal aislada de la reivindicación 10, en donde dicha célula vegetal es una célula de una línea celular de tabaco.
- 60 12. La célula vegetal aislada de la reivindicación 10, en donde dicha célula vegetal es una célula BY-2.
13. Una composición farmacéutica que comprende, como principio activo, la proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-7 o la pluralidad de proteínas α -galactosidasa humanas de la reivindicación 7 o la reivindicación 8 y un vehículo farmacéuticamente aceptable.
14. Una composición farmacéutica que comprende, como principio activo, la célula vegetal de una cualquiera de las reivindicaciones 10-12 y un vehículo farmacéuticamente aceptable.
15. La proteína α -galactosidasa humana de una cualquiera de las reivindicaciones 1-7, la pluralidad de proteínas α -galactosidasa humanas de la reivindicación 7 o la reivindicación 8, la célula vegetal aislada de una cualquiera de las reivindicaciones 10-12 o la composición farmacéutica de la reivindicación 13 o 14 para su uso en el tratamiento de la enfermedad de Fabry en un sujeto que lo necesite.

[REDACTED] LDNGLARTPTMGWLHWE
[REDACTED] LDNGLARTPTMGWLHWE
1 50

RFMCNLDCQEEPDCI SEKLFMEMAELMVSEGWKDAGYEYLCIDDCWMAP
RFMCNLDCQEEPDCI SEKLFMEMAELMVSEGWKDAGYEYLCIDDCWMAP
51 100

QRDSEGRLQADPQRFPHGIRQLANYVHSGKLGKIYADVGNKTCAGFFGS
QRDSEGRLQADPQRFPHGIRQLANYVHSGKLGKIYADVGNKTCAGFFGS
101 150

FGYYDIDAQTFADWGVDLLKFDGCYCDLENLADGYKHMSLAINRTGRSI
FGYYDIDAQTFADWGVDLLKFDGCYCDLENLADGYKHMSLAINRTGRSI
151 200

VYSCWPLYMWPFOKPNYTEIROQCNHWRNFADIDDSWKSIXSILDWTSF
VYSCWPLYMWPFOKPNYTEIROQCNHWRNFADIDDSWKSIXSILDWTSF
201 250

NQERIVDVAGPGGWNDPMLVIGNFGLSWNQQVTQMALWAIMAAPLFMSN
NQERIVDVAGPGGWNDPMLVIGNFGLSWNQQVTQMALWAIMAAPLFMSN
251 300

DLRHISPOAKALIQDKDVIAINQDPLGKQGYQLRQGDNFVWERPLSGLA
DLRHISPOAKALIQDKDVIAINQDPLGKQGYQLRQGDNFVWERPLSGLA
301 350

WAVAMINRQEI GGPRSYTIAVASLGKGVACNPACFITQLLPVKKRKLGFYE
WAVAMINRQEI GGPRSYTIAVASLGKGVACNPACFITQLLPVKKRKLGFYE
351 400

WTSRLRSHINPTGTVLLQLENTMQMSLKDLL
WTSRLRSHINPTGTVLLQLENTMQMSLKDLLSEKDEL
401 427

FIG. 1

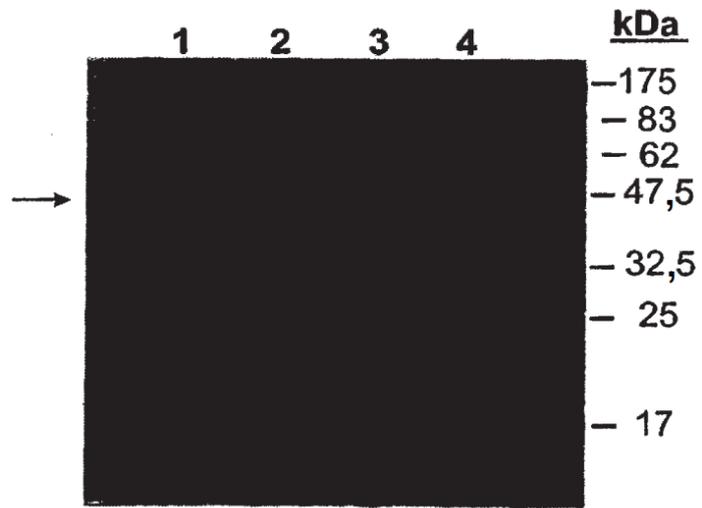


FIG. 2

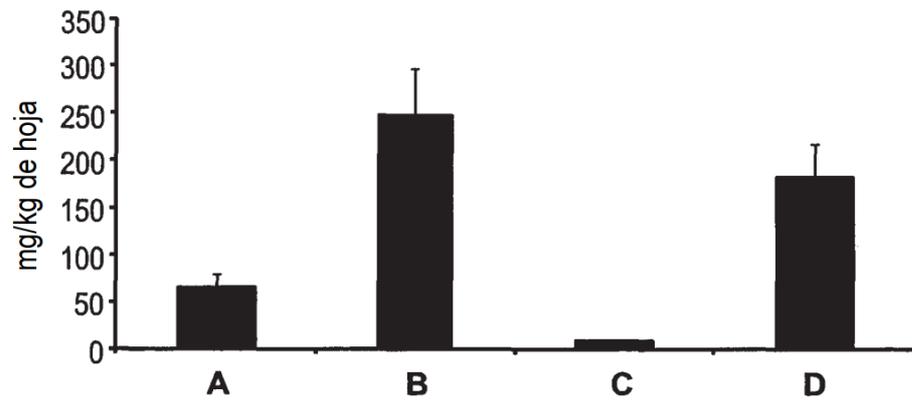


FIG. 3

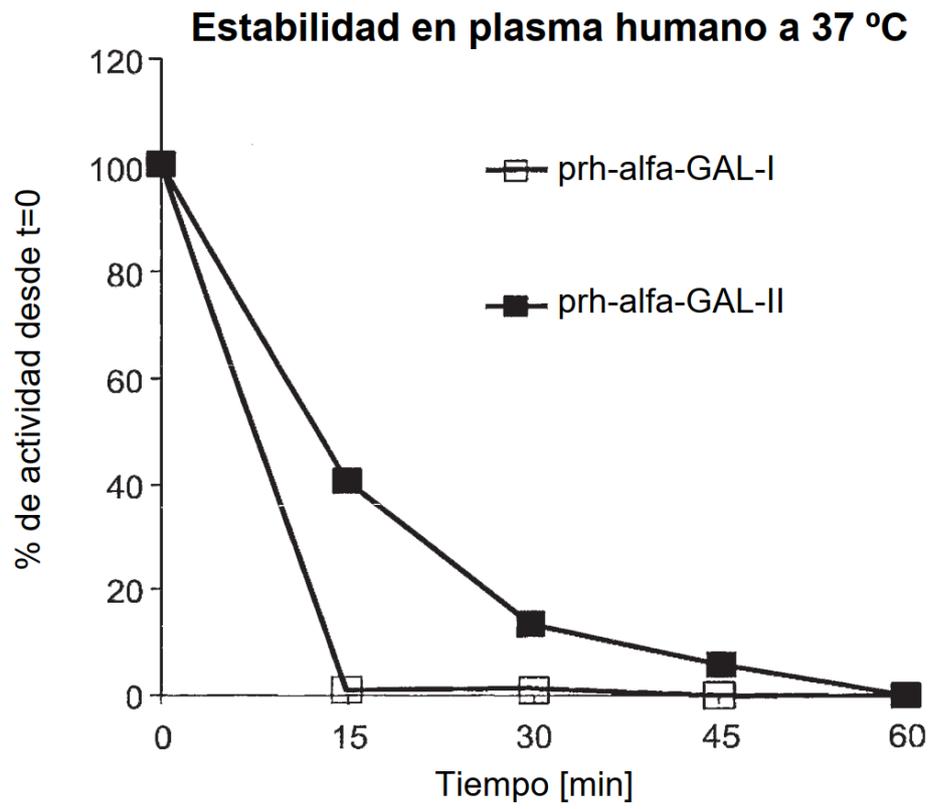


FIG. 4

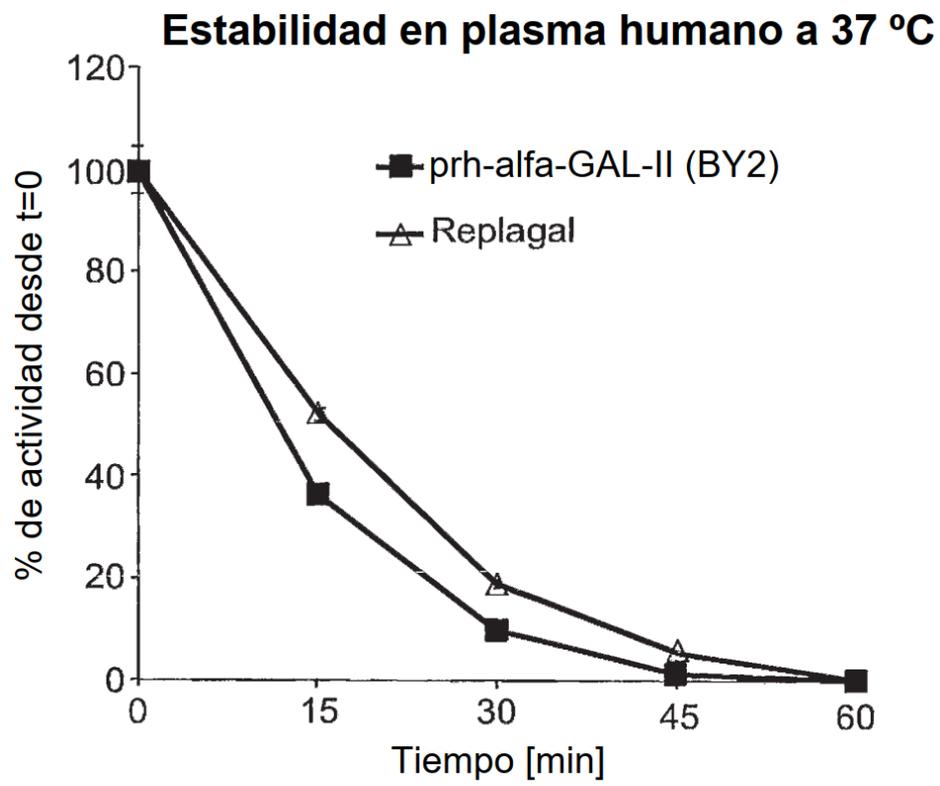


FIG. 5

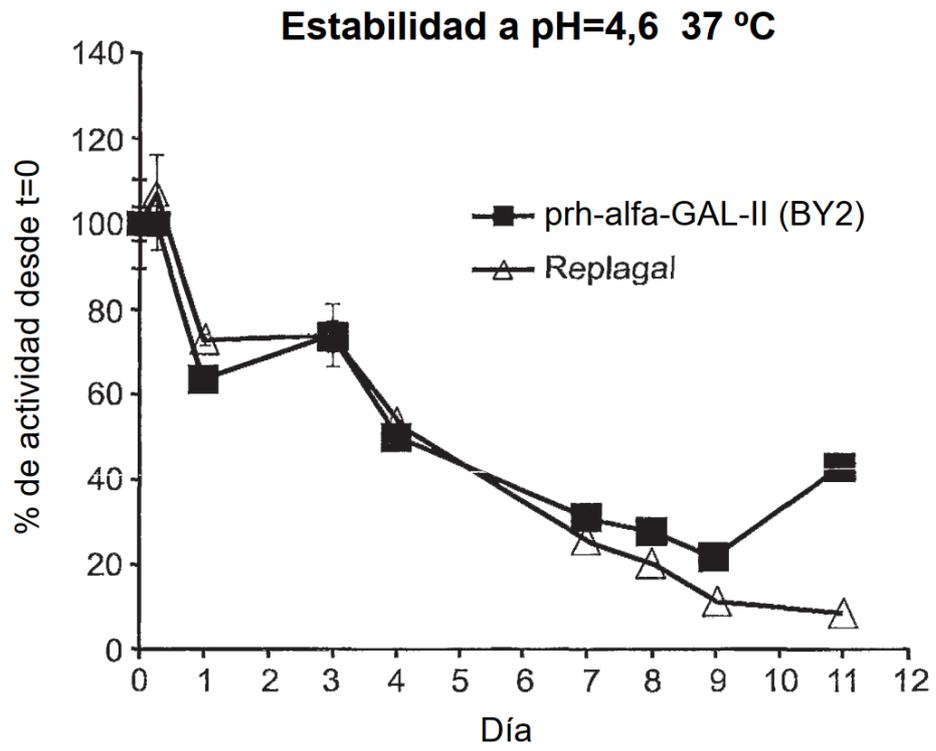


FIG. 6

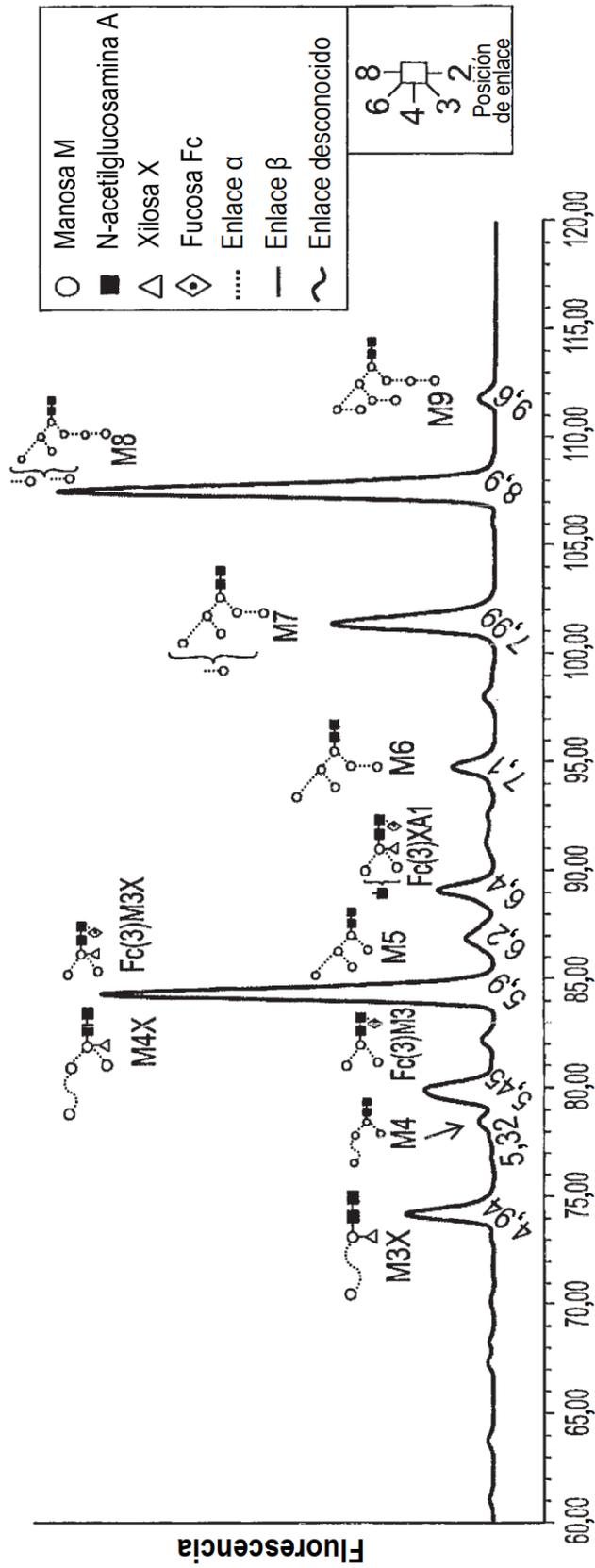


FIG. 7

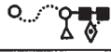
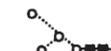
	Nombre	Estructura	Tiempo de retención	UG	A-GAL lote A	A-GAL lote B
1	FcM2X		71,83	4,76	ND	ND
2	M3X		74,32	4,9	6,41	6,51
3	M4		78,5	5,31	1,16	1,19
4	FcM3		79,86	5,32	7,17	6,72
5	FcM3X o M4X		84,55	5,9	26,94	27,59
6	M5		87,126	6,19	3,24	2,44
8	M6		95,083	7,06	3,41	3,34
9	M7		101,81	7,88	14,04	14,27
10	M8		108,05	8,83	31,64	32,12
11	M9		112,347	9,48	1,2	1,3

FIG. 8