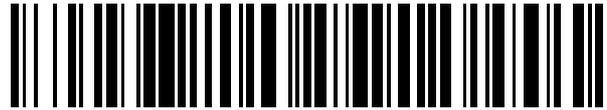


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 562 211**

51 Int. Cl.:

C12N 15/82 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.10.2006 E 06806151 (4)**

97 Fecha y número de publicación de la concesión europea: **02.12.2015 EP 1951879**

54 Título: **Plantas con un contenido de amino azúcares aumentado**

30 Prioridad:

05.10.2005 EP 05090279

11.10.2005 US 725388 P

22.09.2006 EP 06090177

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

03.03.2016

73 Titular/es:

**BAYER INTELLECTUAL PROPERTY GMBH
(100.0%)**

**Alfred Nobel Strasse 10
40789 Monheim am Rhein, DE**

72 Inventor/es:

**FROHBERG, CLAUS y
ESSIGMANN, BERND**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 562 211 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).

DESCRIPCION

Plantas con un contenido de amino azúcares aumentado

5 La presente invención se refiere a células vegetales y plantas que tienen un contenido aumentado de derivados de glucosamina N-acetilados. Además, la presente invención se refiere a células vegetales y plantas que sintetizan glucosaminoglucanos. La presente invención también proporciona procedimientos para la producción de dichas plantas y composiciones que comprenden dichas células vegetales.

10 La glucosamina de amino azúcar, los derivados de glucosamina y los polímeros que comprenden derivados de glucosamina se usan, entre otras cosas, como suplementos alimenticios para la profilaxis de trastornos en las articulaciones en animales y en el hombre. También, en el campo médico, algunos polímeros que contienen derivados de glucosamina se usan para tratar trastornos.

El documento WO 06 032538 describe plantas transgénicas que se han transformado con moléculas de ácido nucleico que codifican para hialuronano sintetas. La síntesis del hialuronano en las plantas en cuestión podría demostrarse de manera unívoca.

15 El documento WO 98 35047 (documento US 6.444.878) describe una ruta metabólica para la síntesis de GlcNAc en células vegetales en donde la glucosamina se convierte mediante una serie de etapas de reacción sucesivas catalizadas enzimáticamente con formación de los metabolitos GlcNAc, N-acetil-glucosamina 6-fosfato, N-acetil-glucosamina 1-fosfato en UDP-GlcNAc. Una ruta metabólica que se ha descrito como una alternativa para plantas comprende la conversión de fructosa 6-fosfato y glutamina en glucosamina 6-fosfato que después se convierte mediante una serie de etapas de reacción sucesivas catalizadas enzimáticamente con la formación de los metabolitos glucosamina 1-fosfato y N-acetilglucosamina 1-fosfato en UDP-GlcNAc. La conversión de fructosa 6-fosfato y glutamina en glucosamina 6-fosfato se cataliza mediante una proteína con la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT, por sus siglas en inglés) (Mayer y col., 1968, Plant Physiol. 43, 1097-1107). Las concentraciones relativamente altas de glucosamina 6-fosfato son tóxicas para la células vegetales (documento WO 98 35047).

25 El documento WO 00 11192 describe la sobreexpresión específica de endospermo de una molécula de ácido nucleico de maíz que codifica una proteína con la actividad enzimática de una GFAT vegetal en plantas de maíz transgénicas con el objetivo de sintetizar un almidón catiónico que tiene moléculas de 2-amidoanhidroglucosa en plantas. La ruta metabólica descrita que, de acuerdo con la descripción del documento WO 00 11192, debería dar como resultado la incorporación de 2-aminoanhidroglucosa en el almidón, comprende entre otras cosas la incorporación de UDP-glucosamina mediante almidón y/o glucógeno sintetas en el almidón. Fue posible demostrar cantidades aumentadas de UDP-glucosamina en la harina de endospermo de las plantas transgénicas en cuestión que sobreexpresan una molécula de ácido nucleico que codifica una proteína que tiene la actividad enzimática de una GFAT vegetal fusionada traduccionalmente con un péptido de señal plastídico. Cuando la proteína que tiene la actividad enzimática de una GFAT se expresó sin péptido de señal, fue posible demostrar una cantidad aumentada de glucosamina 1-fosfato en la harina correspondiente de tejido de endospermo de maíz. No fue posible detectar almidón catiónico o cantidades aumentadas de derivados de N-glucosamina acetilados, tales como, por ejemplo, UDP-GlcNAc o N-acetilglucosamina 6-fosfato en las plantas transgénicas.

40 El amino azúcar beta-D-glucosamina (glucosamina) y/o los derivados de glucosamina son componentes de diversos polímeros (glucosaminoglucanos) que, entre otras cosas, son componentes esenciales del exoesqueleto de los artrópodos, la matriz extracelular de mamíferos o de exopolisacáridos de algunos microorganismos bacterianos. De esta forma, por ejemplo, N-acetil-D-glucos-2-amina (N-acetilglucosamina, GlcNAc) es un derivado de glucosamina acetilado en el átomo de nitrógeno. GlcNAc es, por ejemplo, un bloque de construcción molecular de hialuronano (beta-1,4-[ácido glucurónico beta-1,3-GlcNAc]_n), que es un componente esencial del fluido sinovial.

45 En el campo médico, los productos que contienen hialuronano se usan actualmente para el tratamiento intra-articular de la artrosis y como productos oftálmicos usados para cirugía ocular. El hialuronano reticulado derivatizado se usa para tratar trastornos de articulares (Fong Chong y col., 2005, Appl Microbiol Biotechnol 66, 341-351). Además, el hialuronano es un componente de algunos productos rinológicos que, por ejemplo, en la forma de gotas oculares y nasales, sirven para humectar las membranas mucosas secas. Las soluciones que contienen hialuronano para inyección se usan como analgésicos y antiinflamatorios. Se emplean parches que comprenden hialuronano o hialuronano derivatizado en la curación de heridas. Como productos dérmicos, los implantes de gel que contienen hialuronano se usan para corregir deformaciones de la piel en cirugía plástica. En cirugía cosmética, las preparaciones de hialuronano se encuentran entre los materiales de relleno adecuados para la piel. Mediante la inyección de hialuronano, durante un periodo limitado de tiempo, es posible suavizar las arrugas o aumentar el volumen de los labios.

55 En productos cosméticos, en particular cremas y lociones para la piel, el hialuronano se usa frecuentemente como un humectante por virtud de su alta capacidad de adhesión al agua. Además, las preparaciones que contienen hialuronano se venden como lo que se denomina nutracéuticos (suplementos alimenticios) que también pueden usarse en animales (por ejemplo, perros, caballos) para la profilaxis y alivio de la artrosis.

La catálisis de la síntesis de hialuronano se efectúa por una sola enzima de membrana o asociada a membrana, es decir, hialuronano sintasa (DeAngelis, 1999, CMLS, Cellular and Molecular Life Sciences 56, 670-682). La hialuronano sintasa cataliza la síntesis de hialuronano a partir de los sustratos de ácido UDP-glucurónico (UDP-GlcA) y UDP-N-acetilglucosamina (UDP-GlcNAc).

- 5 El hialuronano usado con fines comerciales actualmente se aísla de tejidos de animal (cresta de gallo) o se prepara mediante fermentación usando cultivos bacterianos.

Los proteoglicanos, una clase de glucoproteínas, son, entre otras cosas, un componente esencial del cartílago y tienen, unidos a la proteína núcleo, glucosaminoglicanos compuestos de unidades de disacárido repetitivas. Las unidades de disacárido repetitivas por su parte están unidas covalentemente a la proteína núcleo a través de una secuencia de unión de carbohidrato característica. Dependiendo de la composición de las unidades de disacárido, se hace una distinción, entre otras cosas, de los glucosaminoglicanos sulfato de heparano/heparina, sulfato de queratano y sulfato de condroitina/dermatano, cuyas unidades de disacárido cada una contiene una molécula que es bien una glucosamina o un derivado de glucosamina. En estas sustancias, los grupos sulfato se introducen en varios átomos o sustituyentes de las unidades de disacárido, de tal forma que las sustancias respectivas mencionadas no son polímeros uniformes sino grupos de polímeros resumidos bajo el término genérico respectivo. En este caso, las moléculas individuales de los grupos de polímero en cuestión pueden diferir tanto en el grado de sulfatación como en la posición de los monómeros que contienen los grupos sulfato.

La síntesis de la cadena de disacárido de la condroitina/dermatano ([beta-1,3]-[ácido glucurónico beta-1,4-N-acetilgalactosamina]_n) se caliza mediante una condroitina sintasa partiendo de UDP-GlcA y UDP-N-acetilgalactosamina, un epímero de UDP-GlcNAc (Kitagawa y col., 2001, J Biol Chem 276(42), 38721-38726). Las moléculas de ácido glucurónico de condroitina se pueden convertir mediante una epimerasa en ácido idurónico. Si más del 10% de las moléculas de ácido glucurónico están presentes como ácido idurónico, el polímero se cita como dermatano. La introducción de los grupos sulfato en diversas posiciones de la cadena de disacárido de la condroitina o del dermatano se cataliza entonces mediante enzimas adicionales, dando como resultado un sulfato de condroitina/dermatano. En el presente documento, el grado de sulfatación puede diferir de molécula a molécula.

Durante algún tiempo, el sulfato de condroitina se ha considerado como un compuesto potencial activo para el tratamiento de la artrosis (Clegg y col., 2006, The New England Journal of Medicine 354(8), 795-808).

La síntesis de la cadena de disacárido de heparina/heparano (heparosano) ([alfa-1,4]-[ácido glucurónico beta-1,4-glucosamina]_n o [alfa-1,4]-[ácido idurónico alfa-1,4-glucosamina]_n), se cataliza mediante la heparina/heparosano sintasa de UDP-GlcA y UDP-GlcNAc (DeAngelis y White, 2004, J- Bacteriology 186(24), 8529-8532). Las moléculas de ácido glucurónico de heparina/heparosano se pueden convertir mediante una epimerasa en ácido idurónico. La introducción de los grupos sulfato en diversas posiciones de la cadena de disacárido de heparosano se cataliza entonces por enzimas adicionales, dando lugar a sulfato de heparina o sulfato de heparano. El sulfato de heparina tiene una sustitución considerablemente más alta por grupos sulfato que el sulfato de heparano. El sulfato de heparina tiene aproximadamente un 90% de moléculas de ácido idurónico, mientras que en el caso del sulfato de heparano predomina la fracción de moléculas de ácido glucurónico (Gallagher y col., 1992, Int. J. Biochem 24, 553-560). Como en el caso del sulfato de condroitina/dermatano, en el caso del sulfato de heparina/heparano, también, el grado de sulfatación puede diferir de molécula a molécula.

El sulfato de heparina se utiliza, en otras cosas, como anticoagulante, por ejemplo, para la profilaxis y tratamiento de la trombosis.

El sulfato de condroitina/dermatano y el sulfato de heparina/heparano se producen actualmente mediante el aislamiento de tejidos animales. El sulfato de condroitina se aísla principalmente a partir de cartílago bovino o de tiburón, y el sulfato de heparina/heparano se aísla a partir del intestino porcino o de pulmones de bovino. Ya que las cadenas de disacárido del sulfato de condroitina/dermatano o el sulfato de heparina/heparano no tienen un patrón de sulfatación uniforme, es difícil obtener un producto específico uniforme. Por consiguiente, los productos siempre son una mezcla de moléculas con diversos grados de sulfatación.

El glucosaminoglicano quitina ([beta-1,4-GlcNAc]_n) es uno de los principales componentes de la pared celular de los hongos y del exoesqueleto de los insectos, miriápodos, arácnidos y crustáceos y es un polímero que es insoluble en agua. La enzima quitina sintasa cataliza la síntesis de quitina mediante la unión de UDP-GlcNAc (Merzendorfer y Zimoch, 2003, J. Experimental Biology 206, 4393-4412).

Como fuente de materia prima para aislar quitina, se usan hasta la fecha principalmente crustáceos (camarones, cangrejos) y hongos, tales como, por ejemplo, *Aspergillus spec.*, *Penicillium Spec.* *Mucor spec.* El documento WO 03 031435 describe, por ejemplo, un procedimiento para preparar GlcNAc mediante fermentación de levaduras. Dependiendo del procedimiento mediante el cual se aísla la quitina a partir de la fuente de materia prima en cuestión, la quitina contiene además de GlcNAc su forma desacetilada de glucosamina como un bloque de construcción. Si más de 50% de los bloques de construcción son GlcNAc, el polímero se cita como quitina, mientras que los polímeros que comprenden más del 50% de glucosamina se citan como quitosano. Actualmente, la glucosamina o derivados de la misma, tales como, por ejemplo, GlcNAc, se producen mediante la degradación de

quitina. La quitina puede bien desacetilarse en primer lugar, dando como resultado la formación de quitosano, o degradarse directamente, dando como resultado la formación de GlcNAc.

La quitina puede desacetilarse enzimáticamente con la ayuda de quitina desacetilasas (Kafetzopoulos y col., 1993, Pro. Natl. Acad. Sci. 90, 2564-2568) o mediante desacetilación química.

- 5 La degradación de quitina o de quitosano también puede tener lugar tanto enzimáticamente (por ejemplo usando quitinasas, gluconasas, beta-N-acetilglucosaminidasas) y mediante hidrólisis química.

La degradación de quitosano o la desacetilación de GlcNAc dan como resultado la formación de glucosamina.

- 10 Una desventaja sustancial de todos los procedimientos para preparar amino azúcares mediante la degradación de quitina consiste en el hecho de que, debido a la hidrólisis incompleta y/o a la desacetilación incompleta, lo que se obtiene no es un producto uniforme sino una mezcla de varios monómeros y oligómeros.

Un procedimiento alternativo para preparar glucosamina con la ayuda de microorganismos recombinantes, en particular *Escherichia coli*, que no requiere la degradación de quitina, se describe en el documento US 2002/0160459.

- 15 Durante algún tiempo, la glucosamina y las sustancias que contienen glucosamina, también, se han considerado como compuestos activos potenciales para el tratamiento de la artrosis (Clegg y col., 2006, The New England Journal of Medicine 354(8), 795-808). La glucosamina o las sustancias que contienen glucosamina también están presentes en muchos suplementos alimenticios. Los alimentos enriquecidos con GlcNAc se describen, por ejemplo, en el documento US 2006/0003965.

- 20 Como ya se ha descrito, los glucosaminoglucanos, tales como, por ejemplo, sulfato de condroitina, sulfato de heparina/heparano o quitina se aíslan actualmente a partir de tejidos animales. Además de las sustancias deseadas en cada caso, estos tejidos contienen también otros glucosaminoglucanos. La separación de los glucosaminoglucanos individuales, en caso de ser posible una separación completa, es difícil y complicada. Además, la presencia potencial, en tejidos animales, de microorganismos patógenos y/o de otras sustancias, tales como, por ejemplo, el patógeno BSE o el patógeno de la gripe aviar, que podría causar enfermedades en el hombre, representa un problema cuando se utilizan glucosaminoglucanos aislados a partir de tejido animal. El uso de preparaciones medicinales contaminadas con proteínas animales puede, en el paciente, dar como resultado reacciones inmunológicas indeseadas del organismos (para preparaciones de hialuronano, véase, por ejemplo, el documento US 4.141.973), en particular si el paciente es alérgico a proteínas animales.

- 30 Un problema adicional durante el aislamiento de los glucosaminoglucanos a partir de tejidos animales consiste en el hecho de que el peso molecular de los glucosaminoglucanos se reduce frecuentemente durante la purificación, ya que los tejidos animales también contienen enzimas que degradarán el glucosaminoglucano.

La glucosamina o los derivados de la misma aislados a partir de crustáceos contienen frecuentemente sustancias (proteínas) que pueden desencadenar una reacción alérgica en el hombre. La glucosamina o los derivados obtenidos a partir de hongos pueden contener micotoxinas.

- 35 Las cantidades (rendimiento) de glucosaminoglucanos que pueden obtenerse con calidad y pureza satisfactorias a partir de tejidos de animales son bajas (por ejemplo, hialuronano de crestas de gallo: 0.079% p/p, documentos EP 0144019, US 4.782.046), lo que significa que tienen que procesarse grandes cantidades de tejidos animales.

- 40 La producción de glucosaminoglucanos con la ayuda de fermentación bacteriana se asocia con altos costes, ya que las bacterias tienen que fermentarse en contenedores estériles sellados en condiciones de cultivo controladas complejas (para hialuronano, véase, por ejemplo, el documento US 4.897.349). Además la cantidad de glucosaminoglucanos que pueden producirse mediante fermentación de cepas bacterianas está limitada por las instalaciones de producción existentes. En este caso, también se ha tenido en cuenta que, debido a las limitaciones físicas, no es posible construir fermentadores para volúmenes de cultivo relativamente grandes. En este contexto, En este contexto, puede hacerse mención en particular a la mezcla homogénea, necesaria para la producción eficaz de sustancias de alimentación (por ejemplo, fuentes de nutrientes esenciales para bacterias, reactivos para regular el pH, oxígeno) con el medio de cultivo, el cual, como mucho, puede asegurarse en grandes fermentadores solamente con un alto gasto técnico.

Además, las sustancias preparadas a partir de materias primas animales no son aceptables para determinados estilos de vida, tales como, por ejemplo, el veganismo o preparación de alimentos kosher.

- 50 Las plantas no producen glucosaminoglucanos de manera natural, tales como, por ejemplo, hialuronano, quitina, sulfato de heparano/heparina, sulfato de queratano o sulfato de condroitina/dermatano.

Para la síntesis de glucosaminoglucanos, es necesario, entre otras cosas, que haya suficientes cantidades de derivados de glucosamina acetilados (en particular UDP-GLcNAc) y/o UDP-GlcA disponibles como sustratos para las enzimas respectivas implicadas en la síntesis. No existe información con respecto a las cantidades de glucosaminas

semi-acetiladas presentes en células vegetales. El documento WO 2005 035710 describe un procedimiento que permite que el contenido de glucosamina del material vegetal aumente mediante secado. El contenido de glucosamina más alto en material vegetal fresco y húmedo se determinó para achicoria con 10 mg de glucosamina en 1 kilo de peso fresco, que, para un peso molecular de 178 para la glucosamina, corresponde a aproximadamente 56 nmol de glucosamina en 1 gramo de peso fresco del material vegetal. El documento WO 2005 035710 no contiene información concerniente al contenido de derivados de glucosamina N-acetilados en plantas.

Además, a partir de la técnica anterior descrita antes, es evidente que las rutas metabólicas de la glucosamina en las plantas aún no se han dilucidado completamente. En el documento WO 00 11192, fue posible generar plantas mediante transformación con una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT vegetal, teniendo dichas plantas un contenido elevado de derivados de glucosamina (UDP-glucosamina o glucosamina 1-fosfato); sin embargo, no se encontraron cantidades aumentadas de derivados de glucosamina N-acetilados.

Por consiguiente, es un objeto de la presente invención proporcionar fuentes alternativas de derivados de glucosamina N-acetilada y procedimientos para preparar dichas fuentes alternativas para derivados de glucosamina N-acetilada.

Un primer aspecto de la presente invención se refiere a células vegetales o plantas que tienen un contenido de derivados de glucosamina N-acetilada de por lo menos 2,50 μmol por gramo de peso fresco, preferentemente de al menos 5,00 μmol por gramo de peso fresco, particularmente preferentemente de al menos 10,00 μmol por gramo de peso fresco, de manera muy particularmente preferente de al menos 15,00 μmol por gramo de peso fresco, de manera especialmente preferente de al menos 20,00 μmol por gramo de peso fresco.

Preferentemente, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención tienen un contenido de derivados de glucosamina N-acetilada de como máximo 250 μmol por gramo de peso fresco, preferentemente de como máximo 200 μmol por gramo de peso fresco, particularmente preferentemente de como máximo 150 μmol por gramo de peso fresco, de manera muy particularmente preferente de como máximo 100 μmol por gramo de peso fresco, de manera especialmente preferente de como máximo 50 μmol por gramo de peso fresco.

En comparación con la técnica anterior, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención ofrecen la ventaja de que contienen cantidades mayores de derivados de glucosamina N-acetilada. En comparación con la producción de derivados de glucosamina N-acetilada mediante la fermentación de microorganismos o el aislamiento de glucosaminas N-acetiladas de fuentes de materia prima de animales u hongos, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la presente invención ofrecen la ventaja de que las células vegetales de acuerdo con la invención y las plantas de acuerdo con la invención pueden propagarse infinitamente de manera vegetativa o sexual, y que producen continuamente derivados de glucosamina N-acetilada. Además, en comparación con las plantas conocidas, las plantas de acuerdo con la invención ofrecen la ventaja de que son más adecuadas para la preparación de glucosaminoglucanos, tales como, por ejemplo, condroitina, hialuronano, quitina, heparosano, ya que contienen una cantidad más alta de sustratos para las enzimas implicadas en la catálisis de los glucosaminoglucanos mencionados (glucosaminoglucano sintasas).

Los derivados de glucosamina N-acetilada pueden detectarse usando procedimientos conocidos para los expertos en la técnica (Morgan y Elson (1934, Biochem J. 28(3), 988-995). En el contexto de la presente invención, para determinar el contenido de derivados de glucosamina N-acetilada, se prefiere hacer uso del procedimiento descrito en el punto 4 de procedimientos generales.

En el contexto de la presente invención, se entiende que la expresión "derivados de glucosamina N-acetilada" significa todos los derivados de glucosamina (2-amino-2-desoxiglucosa), que también incluye epímeros, tales como, por ejemplo, galactosamina (2-amino-2-desoxigalactosa) o manosamina (2-amino-2-desoximanosa), que se miden usando el procedimiento descrito en el punto 4 de procedimientos generales. Los derivados de glucosamina N-acetilada son preferentemente fosfato de N-acetilglucosamina (N-acetilglucosamina 1-fosfato y/o N-acetilglucosamina 6-fosfato), N-acetilglucosamina y/o UDP-N-acetilglucosamina.

Preferentemente las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención tienen un contenido aumentado de fosfato de glucosamina (glucosamina 1-fosfato y/o glucosamina 6-fosfato) además de un contenido aumentado de derivados de glucosamina N-acetilada.

Las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención pueden prepararse, por ejemplo, introduciendo moléculas de ácido nucleico exógenas que codifican una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT) de isoforma II (GFAT-2) o que codifican una proteína que tiene la actividad de una GFAT bacteriana.

En una realización preferida de la presente invención, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención son por lo tanto células vegetales modificadas genéticamente y plantas modificadas genéticamente, respectivamente.

Sorprendentemente, se ha descubierto que las células vegetales o las plantas que contienen una molécula de ácido nucleico que codifica una proteína que tiene una actividad de una GFAT-2 o una proteína que tiene la actividad de una GFAT bacteriana contienen considerablemente más derivados de glucosamina N-acetilada que en células vegetales o plantas que contienen una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma I (GFAT-1). Como ya se mencionó, no fue posible detectar cantidades aumentadas de derivados de glucosamina acetilada en plantas que contienen una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT vegetal (documento WO 00 11192).

Por consiguiente, la presente invención también proporciona células vegetales modificadas genéticamente o plantas modificadas genéticamente que contienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato aminotransferasa (GFAT), en donde la molécula de ácido nucleico exógena codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de la isoforma II (GFAT-2) o una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa bacteriana (GFAT bacteriana).

La modificación genética de una célula vegetal de acuerdo con la invención o una planta de acuerdo con la invención puede ser cualquier modificación genética adecuada para integrar una molécula de ácido nucleico exógena en una célula vegetal o planta.

Preferentemente, la molécula de ácido nucleico exógena se integra en el genoma; de manera particularmente preferente, la molécula de ácido nucleico exógena está establemente integrada en el genoma en las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención.

Hay disponible un gran número de técnicas para (de manera estable) integrar las moléculas de ácido nucleico en la célula vegetal hospedadora para producir células vegetales de acuerdo con la invención o plantas de acuerdo con la invención. Estas técnicas incluyen la transformación de las células vegetales con ADN-T usando *Agrobacterium tumefaciens* o *Agrobacterium rhizogenes* como medios de transformación, fusión de protoplastos, inyección, electroporación de ADN, introducción de ADN mediante el procedimiento biolístico, y también opciones adicionales (revisar en "Transgenic Plants", Leandro ed. Humana Press 2004, ISBN 1-59259-827-7).

El uso de transformación mediada por agrobacterium de células vegetales se ha sometido a estudios profundos y se ha descrito exhaustivamente en el documento EP 120516 y en Hoekema, IN: The Binary Plant Vector System Offsetdrukkerij Kanters B.V. Alblasterdam (1985), Capítulo V; Fraley y col., Crit. Rev. Plant Sci. 4, 1-46 y en An y col., EMBO J. 4, (1985), 277-287. Para la transformación de patatas, véase, por ejemplo Rocha-Sosa y col., EMBO J. 8, (1989), 29-33, para la transformación de plantas de tomate véase, por ejemplo, el documento US 5.565.347.

También se ha descrito la transformación de plantas monocotiledóneas usando vectores basados en la transformación con *Agrobacterium* (Chan y col., Plant Mol. Biol. 22, (1993), 491-506; Hiei y col., Plant J. 6, (1994) 271-282; Deng y col., Science in China 33, (1990), 28-34; Wilmink y col., Plant Cell Reports 11, (1992), 76-80; May y col., Bio/Technology 13, (1995), 486-492; Conner y Domisse, Int. J. Plant Sci. 153 (1992), 550-555; Ritchie y col., Transgenic Res. 2, (1993), 252-265). Los sistemas alternativos para transformar plantas monocotiledóneas es la transformación usando el procedimiento biolístico (Wan y Lemaux, Plant Physiol. 104, (1994), 37-48; Vasil y col., Bio/Technology 11 (1993), 1553-1558; Ritala y col., Plant Mol. Biol. 24, (1994), 317-325; Spencer y col., Theor. Appl. Genet. 79, (1990), 625-631), la transformación con protoplastos, la electroporación de células parcialmente permeabilizadas, o la introducción de ADN usando fibras de vidrio. En particular, se ha descrito la transformación de maíz varias veces en la bibliografía (consúltese, por ejemplo, los documentos WO95/06128, EP0513849, EP0465875, EP0292435; Fromm y col., Biotechnology 8, (1990), 833-844; Gordon-Kamm y col., Plant Cell 2, (1990), 603-618; Kozziel y col., Biotechnology 11 (1993), 194-200; Moroc y col., Theor. Appl. Genet. 80, (1990), 721-726). También se ha descrito la transformación de otras hierbas, tales, como por ejemplo pasto varilla (*Panicum virgatum*) (Richards y col., 2001, Plant Cell Reporters 20, 48-54).

También ya se han descrito las transformaciones exitosas de otras especies de cereales, por ejemplo para la cebada (Wan y Lemaux, citado anteriormente; Ritala y col., citado anteriormente; Krens y col., Nature 296, (1982), 72-74) y para el trigo (Nehra y col., Plant J. 5, (1994), 285-297; Becker y col., 1994, Plant Journal 5, 299-307). Todos los procedimientos anteriores son adecuados en el contexto de la presente invención.

Las células vegetales modificadas genéticamente y las plantas modificadas genéticamente que tienen una molécula de ácido nucleico exógena pueden distinguirse de las células vegetales de tipo silvestre y de las plantas de tipo silvestre, respectivamente, que no tienen la molécula de ácido nucleico exógena, entre otras cosas, por el hecho de que contienen una molécula de ácido nucleico exógena que no es natural en células vegetales de tipo silvestre y plantas de tipo silvestre, respectivamente. Dicha integración de la molécula de ácido nucleico exógena en una célula vegetal o planta puede detectarse usando procedimientos conocidos para un experto en la técnica, tales como, por ejemplo, análisis de transferencia de Southern o mediante PCR.

En el contexto de la presente invención, se entiende que la expresión "molécula de ácido nucleico integrada de manera estable" significa la integración de una molécula de ácido nucleico en el genoma de la planta. Una molécula de ácido nucleico integrada de manera estable se caracteriza por que, durante la replicación del sitio de integración

correspondiente, se multiplica junto con las secuencias de ácido nucleico del hospedador que colinda con el sitio de integración, de tal forma que el sitio de integración en la hebra de ADN hija replicada está rodeada por las mismas secuencias de ácido nucleico que la hebra madre de lectura que sirve como matriz para la replicación.

5 La integración estable de una molécula de ácido nucleico en el genoma de una célula vegetal o una planta puede demostrarse mediante procedimientos genéticos y/o procedimientos de biología molecular. Una integración estable de una molécula de ácido nucleico en el genoma de una célula vegetal o el genoma de una planta está caracterizada por que en la descendencia que ha heredado la molécula de ácido nucleico, la molécula de ácido nucleico integrada de forma estable existe en el mismo entorno genómico que en la generación progenitora. La presencia de una integración estable de una secuencia de ácido nucleico en el genoma de la célula vegetal o en el genoma de una planta puede demostrarse usando procedimientos conocidos por los expertos en la técnica, entre otros con la ayuda del análisis de transferencia de Southern, o el análisis RFLP (polimorfismo de longitud de fragmento de restricción) (Nam y col., 1989, *The Plant Cell* 1, 699-705; Leister y Dean, 1993, *The Plant Journal* 4 (4), 745-750), con procedimientos basados en la PCR, tales como, por ejemplo, el análisis de diferencias en longitud en el fragmento amplificado (AFLP, Polimorfismo de Longitud de Fragmento Amplificado) (Castiglioni y col., 1998, *Genetics* 149, 2039-2056; Meksem y col., 2001, *Molecular Genetics and Genomics* 265, 207-214; Meyer y col., 1998, *Molecular and General Genetics* 259, 150-160) o usando fragmentos divididos usando endonucleasas de restricción (CAPS, Secuencias Polimórficas Amplificadas Divididas) (Konieczny y Ausubel, 1993, *The Plant Journal* 4, 403-410; Jarvis y col., 1994, *Plant Molecular Biology* 24, 685-687; Bachem y col., 1996, *The Plant Journal* 9 (5), 745-753).

20 En el contexto de la presente invención se entiende que el término "genoma" significa el material genético completo presente en una célula vegetal. Se sabe por el experto en la técnica, que además del núcleo, otros componentes (por ejemplo, plástidos, mitocondrias) también contienen material genético.

25 Una materia objeto adicional de la presente invención se refiere a células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención que expresan una molécula de ácido nucleico exógena que codifica la proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o codifican una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa bacteriana (GFAT bacteriana).

30 En el contexto de la presente invención, el término "expresar" o "expresión" se entiende que significa la presencia de transcritos (ARNm) codificados por una molécula de ácido nucleico exógena y/o la presencia de proteínas que tienen la actividad de un GFAT-2 o un GFAT bacteriana.

Puede demostrarse la expresión, por ejemplo, mediante la detección de transcritos específicos (ARNm) o moléculas de ácido nucleico exógenas mediante análisis de transferencia de Northern o RT-PCR.

35 Puede determinarse si las células vegetales o las plantas contienen proteínas que tienen la actividad de una GFAT-2 o proteínas que tienen la actividad de una GFAT bacteriana, por ejemplo, mediante procedimientos inmunológicos, tales como análisis por transferencia de Western, ELISA (Ensayo Inmunoabsorbente Acoplado a Enzimas) o RIA (ensayo radioinmunológico). El experto en la técnica está familiarizado con los procedimientos para preparar anticuerpos que reaccionan específicamente a una determinada proteína, es decir, que se unen específicamente a una determinada proteína (véase, por ejemplo, Lottspeich y Zorbas (Eds.), 1998, *Bioanalytik, Spektrum akad. Verlag, Heidelberg, Berlín, ISBN 3-8274-0041-4*). Algunas compañías (por ejemplo Eurogentec, Bélgica) ofrecen la preparación de los anticuerpos como servicio por encargo.

40 En una realización preferida adicional de la presente invención, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención tienen una actividad de una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o de codificar una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa bacteriana (GFAT bacteriana).

45 La actividad de las proteínas que tienen la actividad de una GFAT-2 o de las proteínas que tienen la actividad de una GFAT bacteriana en extractos de células vegetales de acuerdo con la invención o de plantas de acuerdo con la invención se puede detectar usando procedimientos conocidos por los expertos en la técnica, tales como, por ejemplo, los descritos en Samac y otros (2004, *Applied Biochemistry and Biotechnology* 113-116, Humana Press, Editor Ashok Mulehandani, 1167-1182, ISSN 0273-2289). Un procedimiento preferido para determinar la cantidad de actividad de una proteína con la actividad de una GFAT se proporciona en el punto 8 de Procedimientos Generales.

50 En el contexto de la presente invención, el término "glutamina:fructosa 6-fosfato amidotransferasa (GFAT)" (E.C. 2.6.1.16), en la bibliografía experta también se cita como glucosamina sintasa, que se entiende que significa una proteína que sintetiza, a partir de los materiales de partida de glutamina y fructosa 6-fosfato (Fruc-6-P), glucosamina 6-fosfato (GlcN-6-P). Esta catálisis tiene lugar de acuerdo con el siguiente esquema de reacción:

55
$$\text{glutamina} + \text{Fruc-6-P} \rightarrow \text{GlcN-6-P} + \text{glutamato}$$

En el contexto de la presente invención, el término "glutamina:fructosa 6-fosfato amidotransferasa (GFAT)" se usa como término genérico que incluye todas las isoformas conocidas.

Un artículo de revisión de Milewski (2002, *Biochimica et Biophysica Acta* 1597, 173-193) describe características estructurales de proteínas que tienen la actividad de una GFAT. La secuencia de aminoácidos de todas las proteínas conocidas que tienen la actividad de GFAT contiene regiones con secuencias de aminoácido conservadas. La secuencia de aminoácidos de proteínas que tienen la actividad de una GFAT tiene un dominio de unión de glutamina N-terminal y un dominio de unión de fructosa 6-fosfato C-terminal que están separados por una secuencia de 40 a 90 aminoácidos no conservados. Ambos dominios son activos aún si están presentes en moléculas de aminoácido separadas. Los análisis de la estructura cristalina de un fragmento que comprende el dominio de unión de glutamina N-terminal de la proteína que tiene la actividad de una GFAT de *Escherichia coli* demostraron que el centro activo de este dominio está situado en el extremo N terminal y el aminoácido Cys1 está implicado en la hidrólisis de la glutamina. Los aminoácidos Arg73 y Asp123 interactúan con los grupos carboxilo y amino de la glutamina. Esta interacción está soportada por los aminoácidos Thr76 y His77. La formación de enlaces de hidrógeno con el grupo amino de la glutamina se atribuye a los aminoácidos Gly99 y Trp74. Los aminoácidos Asn98 y Gly99 estabilizan el bolsillo de cuatro caras del centro activo. Los aminoácidos 25 a 29 y 73-80 forman bucles flexibles que, después de la unión al sustrato de glutamina, contribuyen mediante un cambio conformacional de la proteína para la reacción catalizada mediante una proteína que tiene la actividad de una GFAT. El análisis de la estructura cristalina del dominio de unión de fructosa 6-fosfato C-terminal de la proteína tiene la actividad de una GFAT de *Escherichia coli* demostró que este dominio está construido por dos dominios topológicamente idénticos (aminoácidos 241 a 424 y 425 a 592) seguido de un dominio presente en el extremo C-terminal como un bucle irregular (aminoácidos 593 a 608), pero que solo tiene un centro activo. Los aminoácidos Ser303, Ser347, Gln348, Ser349 y Thr352 están implicados en la unión del sustrato, mientras que los aminoácidos Glu488, His504 y Lys603 están directamente implicados en la catálisis de la reacción de la proteína que tiene la actividad de una GFAT.

En particular en organismos animales, fue posible demostrar dos diferentes isoformas de proteínas que tienen la actividad de una GFAT (citada en la bibliografía como GFAT-1 y GFAT-2, respectivamente). Hu y col., (2004), *J. Biol. Chem.* 279(29), 29988-29993) describen diferencias de las isoformas de proteínas respectivas que tienen la actividad de una GFAT. Además de las diferencias en la expresión específica de tejido de las isoformas en cuestión que tienen la actividad de una GFAT-1 y una GFAT-2, fue posible demostrar que ambas isoformas están reguladas por fosforilación mediante una proteína cinasa dependiente de AMPc. La actividad de una proteína que tiene la actividad enzimática de una GFAT-1 se inhibe mediante la fosforilación del resto de serina conservado (Serina 205 en la GFAT-1 de ratón, n.º de referencia de GenBank: AF334736.1) de la secuencia de aminoácidos en cuestión, mientras que la actividad de una proteína que tiene la actividad de una GFAT-2 aumenta mediante la fosforilación de un resto de serina conservado (serina 202 en la GFAT-2 de ratón, n.º de referencia de GenBank: NM_013529) de una secuencia de aminoácido en cuestión. Tanto las proteínas que tienen la actividad de una GFAT-1 como las proteínas que tienen la actividad de una GFAT-2 se inhiben en una forma dependiente de la concentración a través de UDP-GlcNAc; sin embargo, para una proteína que tiene la actividad de una GFAT-2, la inhibición mediante UDP-GlcNAc es inferior (reducción máxima de actividad a través de UDP-GlcNAc de aproximadamente 15%) en comparación con una proteína que tiene la actividad de una GFAT-1 (reducción máxima de la actividad de UDP-GlcNAc en aproximadamente un 51% u 80%, respectivamente). Existen indicaciones de que la inhibición de una proteína que tiene la actividad de una GFAT-1 en organismos animales se basa en el hecho de que a concentraciones de UDP-GlcNAc elevadas existe una glucosilación de O-glucosa-N-acetilglucosamina de las proteínas en cuestión. Todavía no se ha determinado si la regulación de la actividad de las proteínas a través de O-glucosilación también toma parte en las células vegetales (Huber y Hardin, 2004, *Current Opinion in Plant Biotechnology* 7, 318-322).

Las proteínas que tienen la actividad de una GFAT bacteriana se distinguen por el hecho de que no se inhiben mediante UDP-GlcNAc (Komfeld, 1967, *J. Biol. Chem.* 242(13), 3135-3141).

Las proteínas que tienen la actividad de una GFAT-1, las proteínas que tienen la actividad de GFAT-2, e incluso las proteínas que tienen la actividad de una GFAT bacteriana se inhiben por el producto de glucosamina 6-fosfato formado en su reacción (Broschat y col., 2002, *J. Biol. Chem.* 277(17), 14764-14770; Deng y col., 2005, *Metabolic Engineering* 7, 2001-214).

En el contexto de la presente invención, se entiende que la expresión "proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de la isoforma I (GFAT-1)" significa una proteína que tiene la actividad de una GFAT y cuya actividad se inhibe mediante la fosforilación por una proteína cinasa dependiente de AMPc.

En el contexto de la presente invención, se entiende que la expresión "proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2)" significa una proteína que tiene la actividad de una GFAT y que se activa mediante fosforilación por una proteína cinasa dependiente de AMPc.

En el contexto de la presente invención, se entiende que la expresión "proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa bacteriana (GFAT bacteriana)" significa una proteína que tiene la actividad de una GFAT y cuya actividad no se inhibe por UDP-GlcNAc. Como alternativa, las "proteínas que tienen la actividad de una GFAT bacteriana" también pueden citarse como "proteínas que tienen la actividad de una GFAT no eucariótica".

En el contexto de la presente invención, se entiende que la expresión “molécula de ácido nucleico exógena” significa una molécula que no existe de manera natural en las células vegetales de tipo silvestre correspondientes o que no existen de manera natural en la configuración espacial específica en células vegetales de tipo silvestre o que se localiza en un sitio en el genoma de la célula vegetal de tipo silvestre donde no existe de forma natural.

- 5 Preferentemente, la molécula de ácido nucleico exógena es una molécula recombinante que consta de varios elementos (moléculas de ácido nucleico) cuya combinación o configuración espacial específica no existe de manera natural en las células vegetales.

En el contexto de la presente invención, se entiende que la expresión “molécula de ácido nucleico recombinante” significa una molécula de ácido nucleico que tiene varias moléculas de ácido nucleico que no están presentes de manera natural en una combinación del modo que están presentes en la molécula de ácido nucleico recombinante. Por lo tanto, las moléculas de ácido nucleico recombinante pueden, además de las moléculas de ácido nucleico exógenas que codifican una proteína, tener, por ejemplo, secuencias de ácido nucleico adicionales que no están presentes de manera natural en combinación con las moléculas de ácido nucleico que codifican la proteína. En el presente documento, las secuencias de ácido nucleico adicionales mencionadas, que están presentes en una molécula de ácido nucleico recombinante en combinación con una molécula de ácido nucleico que codifica una proteína, pueden ser cualquier secuencia. Estas pueden, por ejemplo, representar secuencias de ácido nucleico genómico y/o vegetal.

Las secuencias de ácido nucleico adicionales mencionadas son preferentemente secuencias reguladoras (promotoras, señales de terminación, potenciadores, intrones), de manera particularmente preferente secuencias reguladoras activas en tejido vegetal, de manera muy particularmente preferente secuencias reguladoras específicas de tejido activas en tejido vegetal.

Los procedimientos para generar moléculas de ácido nucleico recombinantes son conocidos por los expertos en la técnica e incluyen procedimientos de ingeniería genética, tales como, por ejemplo, el enlace de las moléculas de ácido nucleico mediante ligamiento, recombinación genética o la síntesis de nuevas moléculas de ácido nucleico (véase, por ejemplo, Sambrook y col., *Molecular Cloning, A Laboratory Manual*, 3ª edición (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. ISBN: 0879695773, Ausubel y col., *Short Protocols in Molecular Biology*, John Wiley & Sons; 5ª edición (2002), ISBN: 0471250929).

La presente invención proporciona preferentemente células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención en las que las moléculas de ácido nucleico exógenas que codifican la proteína tienen la actividad de una GFAT-2 o codifican una proteína que tiene la actividad de una GFAT bacteriana están enlazadas a los elementos reguladores que inician la transcripción en las células vegetales (promotores). Estos pueden ser promotores homólogos o heterólogos. Los promotores pueden ser constitutivos, específicos de tejido o promotores específicos del desarrollo o promotores regulados por factores externos (por ejemplo, después de la aplicación de sustancias químicas, mediante la acción de factores abióticos, tales como calor y/o frío, sequedad, enfermedad, etc.).

En general, cualquier promotor activo en células vegetales es adecuado para la expresión de una molécula de ácido nucleico exógena. Los promotores adecuados son, por ejemplo, el promotor de ARN 35S de un virus del mosaico de la coliflor o el promotor de ubiquitina de maíz o el promotor YLCV de *Cestrum* (Virus de Hoja Rizada Amarilla; documento WO 01 73087; Stavolone y col., 2003, *Plant Mol. Biol.* 53, 703-713) para una expresión constitutiva, el promotor del gen patatín B33 (Rocha-Sosa y col., *E;BO J.* 8 (1989), 23-29) para una expresión específica de tubérculo en patatas o un promotor específico de frutos para tomate, tal como, por ejemplo, el promotor de poligalacturonasa de tomate (Montgomery y col., 1993, *Plant Cell* 5, 1049-1062) o el promotor E8 de tomate (Metha y col., 2002, *Nature Biotechnol.* 20(6), 613-618) o el promotor de oxidasa ACC de melocotón (Moon y Callahan, 2004, *J. Experimental Botany* 55 (402), 1519-1528), o un promotor que asegura la expresión solamente en tejidos fotosintéticamente activos, por ejemplo el promotor ST-LS1 (Stockhaus y col., *Proc. Natl. Acad. Sci. USA* 84 (1987), 7943-7947; Stockhaus y col., *EMBO J.* 8 (1989), 2445-2451) o para una expresión específica de endospermo del promotor HMWG de trigo, el promotor USP, el promotor de faseolina, promotores de genes de zeína de maíz (Pedersen y col., *Cell* 29 (1982), 1015-1026; Quatroccio y col., *Plant Mol. Biol.* 15 (1990), 81-93), un promotor de glutelina (Leisy y col., *Plant Mol. Biol.* 14 (1990), 41-50; Zheng y col., *Plant J.* 4 (1993), 357-366; Yoshihara y col., *FEBS Lett.* 383 (1996), 213-218), un promotor de globulina (Nakase y col., 1996, *Gene* 170(2), 223-226), un promotor de prolamina (Qu y Takaiwa, 2004, *Plant Biotechnology Journal* 2(2), 113-125) o un promotor de encogido-1 (Werr y col., *EMBO J.* 4 (1985), 1373-1380). Sin embargo, también es posible usar promotores que solamente son activos en un punto en el tiempo determinado por factores externos (véase, por ejemplo, el documento WO 9307279). De particular interés en este caso pueden ser los promotores de proteínas de choque por calor que permiten una inducción simple. Además es posible usar promotores específicos de semillas, tales como, por ejemplo, el promotor USP de *Vicia faba* que asegura la expresión específica de semilla en *Vicia faba* y otras plantas (Fiedler y col., *Plant Mol. Biol.* 22 (1993), 669-679; Bäumllein y col., *Mol. Gen. Genet.* 225 (1991), 459-467).

El uso de promotores presentes en el genoma de virus que infectan a algas también es adecuado para la expresión de secuencias de ácido nucleicos en plantas (Mitra y col., 1994, *Biochem. Biophys Res Commun* 204(1), 187-194; Mitra y Higgins, 1994, *Plant Mol Biol* 26(1), 85-93, Van Etten y col., 2002, *Arch Virol* 147, 1479-1516).

En el contexto de la presente invención, debe entenderse que la expresión “específico de tejido” significa la limitación sustancial de una manifestación (por ejemplo, el inicio de la transcripción) para un determinado tejido.

En el contexto de la presente invención, se entiende que los términos “tubérculo, fruta o célula de endospermo” significan todas las células presentes en un tubérculo, una fruta, y un endospermo de una semilla, respectivamente.

5 En el contexto de la presente invención, se entiende que la expresión “promotor homólogo” significa un promotor que está naturalmente presente en las células vegetales o plantas usadas para la preparación de células vegetales modificadas genéticamente de acuerdo con la invención y plantas modificadas genéticamente de acuerdo con la invención, respectivamente, (homólogos con respecto a la célula vegetal o la planta) o que significa un promotor que regula la regulación de la expresión de un gen en el organismo, a partir del cual se aisló la respectiva molécula de ácido nucleico exógena codifica una proteína (homóloga con respecto a la molécula de ácido nucleico que se va a expresar).

15 En el contexto de la presente invención, se entiende que la expresión “promotor heterólogo” significa un promotor que no está presente de manera natural en las células vegetales o plantas que se utilizan para la preparación de células vegetales modificadas genéticamente de acuerdo con la invención y en plantas modificadas genéticamente de acuerdo con la invención, respectivamente (heterólogo con respecto a la célula vegetal o planta) o que significa un promotor que, en el organismo en el cual se aisló la molécula de ácido nucleico exógena respectiva que codifica una proteína, no naturalmente presente para regular la expresión de la molécula de ácido nucleico exógena (heteróloga con respecto a la molécula de ácido nucleico que se va a expresar).

20 También puede estar presente una secuencia de terminación (señal de poliadenilación), que sirve para agregar una cola de poli-A al transcrito. Se cree que la cola de poli-A actúa en la estabilización de los transcritos. Dichos elementos se describen en la bibliografía (consúltese Gielen y col., EMBO J. 8 (1989), 23-29) y pueden intercambiarse según se desee.

25 También es posible que haya secuencias de intrón presentes entre el promotor y la región codificante de la molécula de ácido nucleico exógena. Dichas secuencias de intrón pueden dar lugar a la estabilidad de la expresión y a una expresión aumentada en plantas (Callis y col., 1987, Genes Devel. 1, 1183-1200; Luehrsen, y Walbot, 1991, Mol. Gen. Genet. 225, 81-93; Rethmeier y col., 1997, Plant Journal 12(4), 895-899; Rose y Beliakoff, 2000, Plant Physiol. 122 (2), 535-542; Vasil y col., 1989, Plant Physiol. 91, 1575-1579; XU y col., 2003, Science en China Series C Vol.46 No.6, 561-569). Las secuencias de intrón adecuadas son, por ejemplo, el primer intrón del gen sh1 de maíz, el primer intrón del gen 1 de poli-ubiquitina de maíz, el primer intrón del gen EPSPS de arroz o uno de los primeros dos intrones del gen PAT1 de *Arabidopsis*.

De acuerdo con la invención, la molécula de ácido nucleico exógena que codifica una proteína que tiene actividad enzimática de una GFAT-2 puede originarse a partir de cualquier organismo eucariótico; preferentemente, dicha molécula de ácido nucleico se origina a partir de animales, de manera particularmente preferente de mamíferos y de manera muy particularmente preferente de ratones.

35 De acuerdo con la invención, la molécula de ácido nucleico exógena que codifica una proteína que tiene actividad enzimática de una GFAT bacteriana se puede originar a partir de cualquier organismo no eucariótico o a partir de un genoma de virus; preferentemente, dicha molécula de ácido nucleico se origina a partir de bacterias o virus; de manera particularmente preferente, la molécula de ácido nucleico se origina a partir de *Escherichia coli*. Ya que las secuencias de aminoácido que codifican proteínas virales tienen la actividad de una GFAT tienen una identidad considerablemente mayor con secuencias de aminoácido que codifican proteínas que tienen la actividad de una GFAT bacteriana y una identidad considerablemente inferior con proteínas que tienen la actividad de una GFAT-1 o una GFAT-2, las proteínas virales que tienen la actividad de una GFAT se clasifican con las proteínas bacterianas que tienen la actividad de una GFAT (Landstein y col., 1998, Virology 250, 388-396).

45 Con respecto a los virus, la molécula de ácido nucleico exógena que codifica una proteína que tiene actividad enzimática de una GFAT se origina preferentemente a partir de un virus que infecta a algas, preferentemente un virus que infecta algas del género *Chlorella*, de manera particularmente preferente de un virus *Paramecium bursaria Chlorella* y de manera muy particularmente preferente de un virus *Paramecium bursaria Chlorella*, de una cepa H1.

50 En lugar de una molécula de ácido nucleico de origen natural que codifica una proteína que tiene la actividad de una GFAT-2, o que codifica una proteína que tiene la afinidad de una GFAT bacteriana, una molécula de ácido nucleico introducida en las células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención también puede haberse generado mediante mutagénesis, en donde dicha molécula de ácido nucleico exógena mutagenizada se caracteriza por que codifica una proteína que tiene la actividad de una GFAT-2 o una proteína que tiene la actividad de una GFAT bacteriana que tiene inhibición reducida mediante metabolitos (por ejemplo del metabolismo de glucosamina). De una forma ilustrativa, la preparación de dichas moléculas de ácido nucleico mutagenizadas se describe en Deng y col., (2005, Metabolic Engineering 7, 201-214; documento WO 04 003175) para una proteína que tiene la actividad de una GFAT bacteriana de *Escherichia coli*. Se describen mutantes de una proteína que tienen la actividad de una GFAT-2 de ratón, por ejemplo, en Hu y col., (2004, J. Biol. Chem. 279(29), 29988-29993).

Las moléculas de ácido nucleico que codifican una proteína con la actividad de una GFAT son conocidas por los expertos en la técnica y se describen en la bibliografía. Por lo tanto, las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT bacteriana se describen, por ejemplo, para *Escherichia coli* (Dutka-Malen, 1988, Biochemie 70 (2), 287-290; n.º de ref. EMBL: L10328.1), *Bacillus subtilis* (n.º de ref. EMBL: U21932), *Haemophilus influenzae* (n.º de ref. EMBL AB006424.1, BAA33071). Las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT bacteriana también se describen para virus, tales como, por ejemplo, como el virus k2 de *Chlorella* (n.º de ref. EMBL: AB107976.1).

Las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT-2 se describen entre otros de insectos, por ejemplo, *Drosophila melanogaster* (n.º de ref. NCBI NM_143360.2), de vertebrados, por ejemplo para *Homo sapiens* (n.º de ref. NCBI BC000012.2, Oki y col., 1999, Genomics 57 (2), 227-34) o *Mus musculus* (n.º de ref. EMBL: AB016780.1).

En una realización preferida, la presente invención se refiere a células vegetales modificadas genéticamente de acuerdo con la invención y a plantas modificadas genéticamente de acuerdo con la invención, en donde la molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o que codifica una proteína que tiene la actividad de una GFAT bacteriana se seleccionan del grupo que consiste en:

a) moléculas de ácido nucleico que codifican una proteína que tiene la secuencia de aminoácidos proporcionada en SEC ID N° 7 (GFAT-2) o una proteína que tiene la secuencia de aminoácidos proporcionada en SEC ID N° 9 (GFAT bacteriana);

b) moléculas de ácido nucleico que codifican una proteína, cuya secuencia es al menos un 60%, preferentemente al menos un 70%, más preferentemente al menos un 80%, de manera particularmente preferente al menos un 90%, de manera muy particularmente preferente al menos un 95% y preferentemente al menos un 98% idéntica a la secuencia de aminoácidos mostrada como SEC ID N° 7 (GFAT-2) o como SEC ID N° 9 (GFAT bacteriana);

c) moléculas de ácido nucleico que comprenden la secuencia de nucleótidos mostrada como SEC ID N° 6 (GFAT-2) o como SEC ID N° 8 (GFAT bacteriana) o como SEC ID N° 10 (GFAT bacteriana) o una secuencia complementaria a las mismas;

d) moléculas de ácido nucleico que son al menos un 60%, preferentemente al menos un 70%, más preferentemente al menos un 80%, de manera particularmente preferente al menos un 90%, de manera muy particularmente preferente al menos un 95%, y lo más preferentemente al menos un 98% idénticas a las secuencias de ácido nucleico mostradas en los puntos a) o c);

e) moléculas de ácido nucleico que hibridan en condiciones rigurosas con al menos una cepa de las secuencias de ácido nucleico descritas en el punto a) o c);

f) moléculas de ácido nucleico cuyas secuencias de nucleótido difieren de las moléculas de ácido nucleico mencionadas en a) o c) debido a la degeneración del código genético; y

g) moléculas de ácido nucleico que son fragmentos, variantes alélicas y/o derivados de moléculas de ácido nucleico mencionadas en los puntos a), b), c), d), e) o f).

En el contexto de la presente invención, el término "hibridación" significa una hibridación en condiciones de hibridación convencionales, preferentemente en condiciones rigurosas, tal como se describen, por ejemplo, en Sambrook y col., (Molecular Cloning, A Laboratory Manual, 3ª edición. (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). ISBN: 0879695773) o Ausubel y col., (Short Protocols in Molecular Biology, John Wiley & Sons; 5ª edición (2002), ISBN: 0471250929). Con particular preferencia, "hibridación" significa una hibridación en las siguientes condiciones:

Tampón de hibridación:

2xSSC; solución 10xDenhardt (Fikoll 400+PEG+BSA; relación 1:1:1); 0.1% de SDS; 5 mM de EDTA; 50 mM de Na₂HPO₄;
250 µg/ml de ADN de esperma de arenque; 50 µg/ml ARNt; o
25 M de tampón de fosfato de sodio pH 7.1; 1 mM de EDTA; 7% de SDS

Temperatura de hibridación:

T = 65 a 68°C

Tampón de lavado: 0.1xSSC; 0.1% SDS

Temperatura de lavado: T = 65 a 68°C.

Las moléculas de ácido nucleico que hibridan con moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT-2 o codifican una proteína que tiene la actividad de una GFAT bacteriana se pueden originar a partir de cualquier organismo; por consiguiente, se pueden originar a partir de bacterias, hongos, animales, plantas o virus.

Las moléculas de ácido nucleico que hibridan con moléculas de ácido nucleico que codifican la proteína que tiene la actividad de una GFAT-2, preferentemente se originan en animales, de manera particularmente preferente de mamíferos, y de manera muy particularmente preferente de ratón.

Las moléculas de ácido nucleico que hibridan las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de GFAT bacteriana preferentemente se originan a partir de bacterias o virus, preferentemente de *Escherichia coli*.

5 Las moléculas de ácido nucleico que hibridan con las moléculas mencionadas pueden aislarse, por ejemplo, a partir de bibliotecas genómicas o de ADNc. Dichas moléculas de ácido nucleico pueden identificarse y aislarse usando las moléculas de ácido nucleico mencionadas o partes de estas moléculas o los complementos inversos de estas moléculas, por ejemplo, mediante hibridación, de acuerdo con procedimientos convencionales (véase, por ejemplo, Sambrook y col., *Molecular Cloning, A Laboratory Manual*, 3ª edición (2001) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) ISBN: 0879695773) o Ausubel y col., (*Short Protocols in Molecular Biology*, John Wiley & Sons; 5ª edición (2002), ISBN: 0471250929) o mediante amplificación usando PCR.

10 Como muestra de hibridación para aislar una secuencia de ácido nucleico que codifica una proteína con la actividad de una GFAT-2, es posible usar, por ejemplo, moléculas de ácido nucleico que tienen exacta o esencialmente las secuencias de ácido nucleico descritas como SEC ID N° 6 o fragmentos de estas secuencias de ácido nucleico. Como muestra de hibridación para aislar una secuencia de ácido nucleico que codifica una proteína que tiene la afinidad de una GFAT bacteriana, es posible usar, por ejemplo, moléculas de ácido nucleico que tienen exacta o esencialmente las secuencias de ácido nucleico descritas en SEC ID N° 8 o fragmentos de estas secuencias del ácido nucleico.

15 Los fragmentos utilizados como muestras de hibridación también pueden ser fragmentos sintéticos u oligonucleótidos que se prepararon usando técnicas de síntesis convencionales, cuya secuencia es esencialmente idéntica a la molécula de ácido nucleico descrita en el contexto de la presente invención. Una vez que los genes que hibridan con las secuencias de ácido nucleico descritas en el contexto de la presente invención se identifican y se aíslan, la secuencia deberá determinarse y las propiedades de las proteínas codificadas por esta secuencia deberán analizarse para determinar si son proteínas que tienen la actividad de una GFAT-2 o la actividad de una GFAT bacteriana. Los procedimientos de cómo determinar si una proteína tiene la actividad de una proteína que tiene la actividad de una GFAT-2 o que tiene la actividad de una GFAT bacteriana son conocidos por los expertos en la técnica y se describen, entre otros en la bibliografía (GFAT bacteriana: por ejemplo Deng y col., 2005, *Metabolic Engineering* 7, 201-214), Kornfeld, 1967, *J. Biol. Chem.* 242(13), 3135-3141; GFAT-2: por ejemplo Hu y col., 2004, *J. Biol. Chem.* 279 (29), 29988-29993). Las moléculas que hibridan con moléculas de ácido nucleico descritas en el contexto de la presente invención comprenden en particular fragmentos, derivados, y variantes alélicas de las moléculas de ácido nucleico mencionadas. En el contexto de la presente invención, el término "derivado" significa que las secuencias de estas moléculas difieren en una o más posiciones de las secuencias de las moléculas de ácido nucleico previamente descritas y son altamente idénticas a esas secuencias. Las diferencias con las moléculas de ácido nucleico descritas anteriormente pueden, por ejemplo, deberse a la eliminación (en particular la eliminación de las regiones N y/o C-terminales), adición, sustitución, inserción, o recombinación.

20 En el contexto de la presente invención, el término "identidad" significa una identidad de secuencia sobre la longitud completa de la región codificante de una molécula de ácido nucleico o la longitud completa de una secuencia de aminoácido que codifica una proteína de al menos un 60%, en particular una identidad de al menos un 70%, preferentemente de al menos un 80%, de manera particularmente preferente de al menos un 90%, de manera muy particularmente preferente de al menos un 95%, y más preferentemente de al menos un 98%. En el contexto de la presente invención, se entiende que el término "identidad" significa el número de aminoácidos/nucleótidos idénticos (identidad) con otras proteínas/ácidos nucleicos, expresados como porcentaje.

25 Preferentemente, la identidad con respecto a una proteína con la actividad de una GFAT-2 se determina mediante las comparaciones con las secuencias de aminoácidos proporcionadas como SEC ID N° 7 y la identidad con respecto a una molécula de ácido nucleico que codifica una proteína con la actividad de una GFAT-2 según se determina mediante comparaciones de la secuencia de ácido nucleico proporcionada como SEC ID N° 6 con otras proteínas/ácidos nucleicos con la ayuda de programas informáticos. Preferentemente, la identidad con respecto a una proteína que tiene la actividad de una GAFT bacteriana se determina mediante comparaciones de la secuencia de aminoácido proporcionada como SEC ID N°:9 y la identidad con respecto a una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT bacteriana se determina mediante comparaciones de la secuencia de ácido nucleico proporcionadas como SEC ID N° 8 o SEC ID N° 10 con otras proteínas/ácidos nucleicos con la ayuda de programas informáticos. Si las secuencias que se van a comparar entre sí son de longitud diferente, la identidad se determinará mediante la determinación de la identidad en porcentaje del número de aminoácidos/nucleótidos con las secuencias más cortas que comparten con las secuencias más largas. Preferentemente, la identidad se determina usando el programa informático conocido y públicamente disponible ClustalW (Thompson y col., *Nucleic Acids Research* 22 (1994), 4673-4680). ClustalW está públicamente disponible a través de Julie Thompson (Thompson@EMBL-Heidelberg.DE) y Toby Gibson (Gibson@EMBL-Heidelberg.DE), European Molecular Biology Laboratory, Meyerhofstrasse 1, D 69117 Heidelberg, Alemania. ClustalW también puede descargarse de varias páginas de Internet, entre otras de IGBMC (Institut de Génétique et de Biologie Moléculaire et Cellulaire, B.P.163, 67404 Illkirch Cedex, France; <ftp://ftp-igbmc.u-strasbg.fr/pub/>) y del EBI (<ftp://ftp.ebi.ac.uk/pub/software/>), y todas las páginas de Internet en espejo de EBI (European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, R.U.).

Preferentemente, se hace uso de un programa informático ClustalW versión 1.8 para determinar la identidad entre proteínas descritas en el contexto de la presente invención y otras proteínas. En el presente documento, los parámetros que se tienen que establecer son los siguientes: KTUPLE=1, TOPDIAG=5, WINDOW=5, PAIRGAP=3, GAOPEN=10, GAPEXTEND=0.05, GAPDIST=8, MAXDIV=40, MATRIX=GONNET, ENDGAPS(OFF), NOPGAP, NOHGAP.

Preferentemente, se hace uso del programa informático ClustalW de la versión 1.8 para determinar la identidad por ejemplo entre la secuencia de nucleótido de las moléculas de ácido nucleico descritas en el contexto de la presente invención y la secuencia de nucleótidos de otras moléculas de ácido nucleico. En el presente documento, los parámetros tienen que establecerse como sigue: KTUPLE=2, TOPDIAGS=4, PAIRGAP=5, DNAMATRIX:IUB, GAOPEN=10, GAPEXT=5, MAXDIV=40, TRANSITIONS: no ponderadas.

La identidad además significa que existe una equivalencia funcional y/o estructural entre las moléculas de ácido nucleico en cuestión o las proteínas codificadas por esta. Las moléculas de ácido nucleico que son homólogas a las moléculas descritas anteriormente y representan derivados de estas moléculas son generalmente variaciones de estas moléculas que representan modificaciones que tienen la misma función biológica, es decir, codifican una proteína que tiene la actividad de una GFAT-2 o la actividad de una GFAT bacteriana. Estas pueden ser variaciones de origen natural, por ejemplo, secuencias de otras especies, o mutaciones, en donde estas mutaciones pueden haber ocurrido en una forma natural o se introdujeron mediante mutagénesis sistemática. Además, las variaciones pueden ser secuencias producidas sintéticamente. Las variantes alélicas pueden ser variantes de origen natural o variantes producidas sintéticamente o variantes generadas mediante técnicas de ADN recombinante. Una forma especial de derivados son, por ejemplo, moléculas de ácido nucleico que difieren de las moléculas de ácido nucleico descritas en el contexto de la presente invención como un resultado de la degeneración del código genético.

En una realización preferida adicional, la presente invención se refiere a células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención, en donde las moléculas de ácido nucleico codifican una proteína que tiene la actividad de una GFAT-2, o codifican una proteína que tiene la actividad de una GFAT bacteriana se caracterizan en que los codones de las moléculas de ácido nucleico son diferentes de los codones de las moléculas de ácido nucleico que codifican la proteína que tiene la actividad de una GFAT-2 o la proteína tiene la actividad de una GFAT bacteriana del organismo progenitor. Particularmente, preferentemente, los codones de las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT-2 o codifican una proteína que tiene la actividad de una GFAT se cambian de tal forma que se adaptan a la frecuencia de uso de los codones de la célula vegetal o de la planta en cuyos genomas se integran o se integrarán.

Como resultado de la degeneración del código genético, los aminoácidos se pueden estar codificados por uno o más codones. En los diferentes organismos, los codones que codifican un aminoácido se utilizan en diferentes frecuencias. La adaptación de los codones de una secuencia de ácido nucleico de codificación a la frecuencia de su uso en la célula vegetal o en la planta en cuyo genoma la secuencia se va a expresar se integrará para contribuir a una cantidad aumentada de proteína traducida y/o la estabilidad de un ARNm en cuestión en las células vegetales o plantas particulares. La frecuencia del uso de codones en las células vegetales o plantas en cuestión se puede determinar mediante el experto en la técnica examinando tantas secuencias de ácido nucleico de codificación como sea posible del organismo en cuestión para la frecuencia con la cual ciertos codones se utilizan para codificar un cierto aminoácido. La frecuencia de uso de codones de ciertos organismos es conocida por los expertos en la técnica y se puede determinar de una forma simple y rápida usando programas informáticos. Los programas informáticos son públicamente accesibles y se proporcionan gratuitamente en Internet (por ejemplo <http://gcu.schoedl.de/>; <http://www.kazusa.or.jp/codon/>; <http://www.entelechon.com/eng/cutanalysis.html>). La adaptación de los codones de una secuencia de ácido nucleico de codificación a la frecuencia de su uso en la célula vegetal o en la planta en cuyo genoma la secuencia se va a expresar se integrará llevando a cabo la mutagénesis *in vitro*, o preferentemente mediante una nueva síntesis de la secuencia de gen. Los procedimientos para la nueva síntesis de las secuencias de ácido nucleico son conocidos por los expertos en la técnica. Una de la nueva síntesis se puede llevar a cabo, por ejemplo, inicialmente sintetizando los oligonucleótidos de ácido nucleico individuales, hibridar estos oligonucleótidos complementarios de los mismos, de tal forma que se produce una cepa de estructura de doble cadena de ADN y después se liga a oligonucleótidos de estructura de cadena doble individuales de tal forma que se obtienen las secuencias de ácido nucleico deseadas. La nueva síntesis de secuencias de ácido nucleico que incluyen la adaptación de la frecuencia con la cual los codones se utilizan en un cierto organismo objetivo también se puede subcontratar a compañías que ofrecen este servicio (por ejemplo, Entelechon GmbH, Regensburg, Alemania).

Todas las moléculas de ácido nucleico mencionadas son adecuadas para la producción de células vegetales de acuerdo con la invención o plantas de acuerdo con la invención.

Las células vegetales modificadas genéticamente de acuerdo con la invención o las plantas modificadas genéticamente de acuerdo con la invención pueden, en principio, ser células vegetales y plantas, respectivamente, de cualquier especie de planta, es decir, tanto plantas monocotiledóneas como dicotiledóneas. Son preferentemente plantas de cultivo, es decir, plantas cultivadas por el hombre con el propósito de alimentar a seres humanos o animales o para producir biomasa y/o para preparar sustancias con fines técnicos, industriales. Las células vegetales

modificadas genéticamente de acuerdo con la invención o las plantas modificadas genéticamente de acuerdo con la invención son de manera particularmente preferente maíz, arroz, trigo, centeno, avena, cebada, mandioca, patata, tomate, pasto varilla (*Panicum virgatum*), sagú, frijol de oro, guisante, sorgo, zanahoria, berenjena, rábano, colza, soja, cacahuate, pepino, calabazas, melones, puerro, ajo, col, espinaca, batatas, espárragos, calabazas, lechuga, alcachofa, maíz dulce, chirivía, salsifí negro, alcachofa de Jerusalén, remolacha, brócoli, col, cebolla, nabo, diente de león, fresa, manzana, ciruela, melocotón, uvas, coliflor, apio, ají, colinabo, y ruibarbo. Preferentemente son plantas de maíz, arroz, trigo, centeno, avena o cebada, de manera muy particularmente preferente plantas de arroz, tomate o patata.

En el contexto de la presente invención, se entiende que la expresión “planta de patata” o “patata” significa especies de plantas del género *Solanum*, particularmente especies productoras de tubérculos del género *Solanum* y en particular *Solanum tuberosum*.

En el contexto de la presente invención, se entiende que la expresión “planta de tomate” o “tomate” significa especies de plantas del género *Lycopersicon*, en particular *Lycopersicon esculentum*.

En el contexto de la presente invención, se entiende que la expresión “planta de arroz” significa especies de plantas del género *Oryza*, en particular especies de plantas del género *Oryza* agrícolamente cultivadas con fines comerciales, de manera particularmente preferente *Oryza sativa*.

Como ya se discutió, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención son adecuadas para la producción de glucosaminoglucanos, tales como, por ejemplo, condroitina, hialuronano, quitina, heparina (heparosano) ya que contienen una cantidad superior de sustratos para las enzimas implicadas en la catálisis de los glucosaminoglucanos mencionados.

Por consiguiente, la presente invención se refiere además a células vegetales o plantas que sintetizan glucosaminoglucano, preferentemente al menos 500 µg de glucosaminoglucano por gramo de peso fresco, más particularmente al menos 1500 µg de glucosaminoglucano por gramo de peso fresco, de manera particularmente preferente al menos 3500 µg de glucosaminoglucano por gramo de peso fresco, de manera muy particularmente preferente al menos 4000 µg de glucosaminoglucano por gramo de peso fresco, y especialmente preferentemente al menos 5500 µg de glucosaminoglucano por gramo de peso fresco. En este contexto, el glucosaminoglucano es preferentemente condroitina, hialuronano, quitina o heparina (heparosano), de manera particularmente preferente hialuronano.

Las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención preferentemente tienen un contenido de glucosaminoglucano de como máximo 25.000 µmol por gramo de peso fresco, preferentemente como máximo 20.000 µmol por gramo de peso fresco, de manera particularmente preferente como máximo 10.000 µmol por gramo de peso fresco, especialmente preferentemente como máximo 6.500 µmol por gramo de peso fresco.

Las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención que sintetizan glucosaminoglucano se pueden producir, por ejemplo, mediante la introducción de moléculas de ácido nucleico exógenas que codifican una proteína que tienen la actividad de una GFAT, y codifican una proteína que tiene la actividad de una glucosaminoglucano sintasa en una célula vegetal.

Por consiguiente, la presente invención también se refiere a células vegetales modificadas genéticamente o a plantas modificadas genéticamente que contienen una primera molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o una GFAT bacteriana y una segunda molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa.

En el contexto de la presente invención, se entiende que la expresión “proteína que tiene la actividad de una glucosaminoglucano sintasa” significa una proteína que utiliza UDP-GlcNAc o UDP-N-acetilgalactosamina, un epímero de UDP-GlcNAc, como sustrato para la síntesis de un glucosaminoglucano. La proteína que tiene la actividad de una glucosaminoglucano sintasa es preferentemente una hialuronano sintasa, condroitina sintasa, heparosano/heparina sintasa, queratano sintasa o quitina sintasa.

Las moléculas de ácido nucleico y las secuencias de proteína correspondientes que codifican para glucosaminoglucano sintasas son conocidas por los expertos en la técnica y se describen como hialuronano sintasa por ejemplo de virus (por ejemplo *Paramecium bursaria Chlorella* Virus 1, EMBL U42580.3, PB42580, documento US 20030235893), como condroitina sintasa por ejemplo de mamíferos (por ejemplo *Homo sapiens*, documentos WO 03 012099, US 2005048604, US 2006052335), bacteria (por ejemplo *Escherichia coli*, US2003109693, EP 1283259, *Pasteurella multocida* US 2003104601), como quitina sintasa por ejemplo de bacteria (por ejemplo *Azorhizobium caulinodans* EMBLCDS:AAB51164), de hongos (por ejemplo *Chaetomium globosum* EMBLCDS:EAQ92361, *Aspergillus nidulans* EMBL AB000125, *Arthroderma benhamiae* EMBLCDS:BAB32692 *Neurospora crassa* EMBL M73437.4), de insectos (por ejemplo *Aedes aegypti* EMBLCDS:EAT46081, *Tribolium castaneum* EMBLCDS:AAQ55061), nematodos (por ejemplo *Dirofilaria immitis* EMBL AF288618, *Caenorhabditis elegans* EMBL AY874871), de virus (por ejemplo *Chorella* virus EMBLCDS:BAB83509, *Paramecium bursaria Chorella* virus CVK2 EMBLCDS:BAE48153), como heparina/heparosano sintasa por ejemplo de bacteria (por ejemplo *Pasteurella*

multocida EMBL AF425591, AF439804, US 20030099967, *Escherichia coli* X77617.1).

5 La segunda molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa es preferentemente una molécula de ácido nucleico recombinante. Las realizaciones preferidas de moléculas de ácido nucleico recombinantes ya han sido descritas y se utilizarán en el presente documento en su forma correspondiente.

10 En una realización preferida adicional, la segunda molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa se caracteriza en que los codones están modificados comparados con los codones de la molécula de ácido nucleico que codifica la proteína que tiene la actividad de una glucosaminoglucano sintasa del organismo progenitor. Particularmente preferentemente, los codones de las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una glucosaminoglucano sintasa están modificados de tal forma que se adaptan a la frecuencia de uso de los codones de la célula vegetal o planta en cuyo genoma se integran o se van a integrar.

15 Lo que se manifestó anteriormente para las moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GAFT-2 o de una GFAT bacteriana con respecto a la modificación de los codones de una molécula de ácido nucleico se aplicará en el presente documento de una forma correspondiente.

La presente invención se refiere además a plantas que contienen células vegetales de acuerdo con la invención. Dichas plantas pueden generarse mediante regeneración de las células vegetales de acuerdo con la invención.

La presente invención también se refiere a partes de plantas de acuerdo con la invención que contienen células vegetales de acuerdo con la invención.

20 En el contexto de la presente invención, se entiende que las expresiones “partes de planta” o “partes de planta” significan, por ejemplo, partes de plantas procesables utilizadas en la producción de productos alimenticios o alimentos para ganado, utilizados como fuente de materia prima para procedimientos industriales (por ejemplo, para el aislamiento de los derivados de glucosamina o glucosaminoglucanos), como fuente de materia prima para la preparación de farmacéuticos y como fuente de materia prima para la preparación de productos cosméticos.

25 En el contexto de la presente invención, se entiende que las expresiones “partes de planta” o “partes de plantas” significan, por ejemplo, partes de plantas consumibles que sirven como alimentos para el hombre o que se utilizan como alimentos para animales.

Las “partes de planta” o “partes de plantas” preferidas son frutos, raíces de almacenamiento u otras, flores, capullos, brotes, hojas o vástagos, preferentemente semillas, frutos, granos o tubérculos.

30 La presente invención también se refiere al material de propagación de plantas de acuerdo con la invención. Preferentemente, el material de propagación de acuerdo con la invención contiene células vegetales de acuerdo con la invención, de manera particularmente preferente células vegetales modificadas genéticamente de acuerdo con la invención.

35 En el presente documento, la expresión “material de propagación” comprende aquellos componentes de la planta que son adecuados para generar progenie mediante la ruta vegetativa o generativa. Los adecuados para la propagación vegetativa son, por ejemplo, cortezas, cultivos de callos, rizomas o tubérculos. Otro material de propagación incluye, por ejemplo, frutos, semillas, granos, plántulas, cultivos celulares, etc. El material de propagación toma preferentemente la forma de tubérculos, frutos, granos o semillas.

40 Una ventaja adicional de la presente invención es el hecho de que las partes de las plantas de acuerdo con la invención tienen un contenido superior de derivados de glucosamina N-acetilada que las plantas conocidas. Por consiguiente, las plantas de acuerdo con la invención son particularmente adecuadas para su uso directo como alimentos comestibles/alimentos para animales o para la preparación de productos alimenticios/alimentos para animales que tienen un efecto profiláctico o terapéutico (por ejemplo, para la profilaxis de la artrosis). Ya que las plantas de acuerdo con la invención tienen un contenido superior de derivados de glucosamina N-acetilada en comparación con las plantas conocidas, las cantidades de partes cosechables, el material de propagación, las partes procesables o las partes consumibles de las plantas de acuerdo con la invención utilizadas para la preparación de productos alimenticios/alimentos para animales tienen un contenido aumentado de derivados de glucosamina N-acetilada que se pueden reducir. Si las partes consumibles de plantas modificadas genéticamente de acuerdo con la invención se consumen, por ejemplo, directamente en forma de los denominados nutracéuticos, se puede lograr un efecto positivo incluso mediante el consumo de pequeñas cantidades de la sustancia. Esto puede ser de particular importancia entre otras cosas en la producción de alimentos para animales ya que los alimentos para animales con contenido demasiado alto de componentes vegetales no son adecuados como alimentos para animales para varias especies de animal. Además, las células vegetales de acuerdo con la invención o las plantas de acuerdo con la invención tienen la ventaja de que también se pueden usar por vegetarianos o para preparar comida kosher. Por lo tanto, es posible administrar los alimentos que tienen un contenido elevado de glucosamina N-acetilada incluso a personas que siguen estos estilos de vida mencionados.

55

Se sabe que la N-acetilglucosamina tiene un efecto estimulante en el crecimiento de bifidobacterias (Liepke y col., 2002, Eur. J. Biochem. 269, 712-718). Además, se ha demostrado que la N-acetilglucosamina sirve como sustrato para los lactobacilos (por ejemplo, *Lactobacillus casei* de las subespecies *paracasei*) del intestino de pescado (Adolfo Bucio Galindo, 2004, Proefschrift, Wageningen Universiteit, ISBN 90-5808-943-6). Por consiguiente, la N-acetilglucosamina tiene un efecto positivo sobre la bacteria probiótica. Ya que las células vegetales de acuerdo con la invención, las plantas de acuerdo con la invención o las partes de planta de acuerdo con la invención tienen un contenido elevado de N-acetilglucosamina, deberán tener, por consiguiente, un efecto positivo en el crecimiento de bacterias probióticas y de esta forma ser adecuadas para su uso como productos alimenticios/alimentos para animales prebióticos para el hombre y animales.

La presente invención se refiere además a un procedimiento para la producción de una planta genéticamente modificada que comprende las siguientes etapas:

- a) la introducción de una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de la isoforma II (GFAT-2) o la codificación para una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT bacteriana) en una célula vegetal.
- b) la regeneración de una planta a partir de las células vegetales obtenidas de acuerdo con la etapa a).
- c) si es adecuado, la generación de plantas adicionales con la ayuda de las plantas de acuerdo con la etapa b).

La presente invención se refiere además a procedimientos para la producción de una planta que sintetiza glucosaminoglucano, en los que

a) se modifica genéticamente una célula vegetal, en donde la modificación genética comprende los siguientes pasos i a ii en cualquier orden o se llevan a cabo cualquier combinación de los siguientes pasos i a ii de manera individual o simultánea.

- i) la introducción de una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT bacteriana) en una célula vegetal.
- ii) introducir una molécula de ácido nucleico exógena que codifica una glucosaminoglucano sintasa en una célula vegetal.

b) se regenera una planta a partir de células vegetales que comprenden la modificación genética de acuerdo con las etapas

- i) a) i
- ii) a) ii
- iii) a) i y a) ii,

c) la introducción en células vegetales de plantas de acuerdo con la etapa

- i) b) i una modificación genética de acuerdo con la etapa a) ii,
- ii) b) ii una modificación genética de acuerdo con la etapa a) i,

y regenerar una planta

d) si es adecuado, generar plantas adicionales con la ayuda de las plantas obtenidas de acuerdo con cualquiera de las etapas b) iii o c) i o c)ii.

Con respecto a la introducción de las moléculas de ácido nucleico exógenas de acuerdo con la etapa a) del procedimiento para la producción de una planta genéticamente modificada o de acuerdo con las etapas a) o c) de los procedimientos para la producción de una planta que sintetiza glucosaminoglucano en una célula vegetal, esta introducción puede, en principio, ser cualquier tipo de introducción de moléculas de ácido nucleico adecuadas para integrar una molécula de ácido nucleico exógena en una célula vegetal o planta. Los procedimientos ya han sido descritos anteriormente y se pueden aplicar en el presente documento de una forma correspondiente.

Con respecto a la molécula de ácido nucleico exógena para una proteína que tiene la actividad de una GFAT-2 o la codificación para una proteína que tiene la actividad de una GFAT bacteriana de acuerdo con la etapa a) del procedimiento para producir una planta genéticamente modificada o con respecto a la molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa de acuerdo con la etapa a) ii) del procedimiento para producir una planta que sintetiza glucosaminoglucano, ya se han descrito varias posibles realizaciones de las moléculas de ácido nucleico respectivas en el contexto con células vegetales de acuerdo con la invención y plantas de acuerdo con la invención. Todas estas realizaciones preferidas que ya se han descrito también pueden usarse para llevar a cabo los procedimientos de acuerdo con la invención mencionada.

La regeneración de las plantas dependiendo del procedimiento de acuerdo con la etapa b) y/o c) de los procedimientos de acuerdo con la invención se pueden llevar a cabo usando procedimientos conocidos por los

expertos en la técnica (descritos, por ejemplo en "Plant Cell Culture Protocols", 1999, edit. por R. D. Hall, Humana Press, ISBN 0-89603-549-2).

5 La generación de plantas adicionales dependiendo del procedimiento de acuerdo con la etapa c) o d) del procedimiento de acuerdo con la invención se pueden llevar a cabo, por ejemplo, mediante propagación vegetativa (por ejemplo, mediante cortezas, tubérculos o mediante cultivo de callo y regeneración de plantas intactas) o mediante propagación generativa. En este contexto, la propagación generativa preferentemente tiene lugar en condiciones controladas, es decir, se hibridan entre sí y se multiplican plantas seleccionadas con características específicas. La selección preferentemente tiene lugar de tal forma que las plantas, dependiendo del procedimiento de acuerdo con las etapas b) o d), tienen las modificaciones introducidas en la etapa a).

10 En una realización preferida adicional, los procedimientos de acuerdo con la invención para producir una planta genéticamente modificada se utilizan para producir plantas de acuerdo con la invención.

La presente invención también proporciona plantas que pueden obtenerse mediante procedimientos de acuerdo con la invención para preparar una planta genéticamente modificada.

15 La presente invención se refiere además a un procedimiento para producir glucosaminoglucanos que comprenden la etapa de extracción de glucosaminoglucanos de células vegetales modificadas genéticamente de acuerdo con la invención, de plantas modificadas genéticamente de acuerdo con la invención, de material de propagación de acuerdo con la invención, de partes de plantas de acuerdo con la invención o de plantas que se obtienen mediante un procedimiento de acuerdo con la invención para preparar una planta genéticamente modificada que sintetiza glucosaminoglucanos. El procedimiento de acuerdo con la invención se usa preferentemente para producir condroitina, hialuronano, quitina o heparina (heparosano), particularmente para producir hialuronano.

20 Preferentemente, dicho procedimiento también comprende la etapa de cosechar las células vegetales modificadas genéticamente cultivadas de acuerdo con la invención, las plantas modificadas genéticamente de acuerdo con la invención, el material de propagación de acuerdo con la invención, las partes de plantas de acuerdo con la invención antes de la extracción del glucosaminoglucano y de manera particularmente preferente además la etapa de cultivar las células vegetales modificadas genéticamente de acuerdo con la invención o plantas modificadas genéticamente de acuerdo con la invención antes de la cosecha.

25 A diferencia de los tejidos bacterianos o animales, los tejidos vegetales no contienen ninguna enzima de degradación de glucosaminoglucano. Por consiguiente, la extracción de glucosaminoglucanos de tejidos vegetales es posible usando procedimientos relativamente simples. En caso necesario, los extractos acuosos de células vegetales o tejidos que contienen glucosaminoglucanos pueden purificarse adicionalmente usando procedimientos conocidos por los expertos en la técnica, tales como, por ejemplo, precipitación repetida con etanol. Un procedimiento preferido para la purificación, por ejemplo, hialuronano se describe en el punto 5 de Procedimientos Generales.

35 La presente invención también proporciona el uso de células vegetales modificadas genéticamente de acuerdo con la invención, plantas modificadas genéticamente de acuerdo con la invención, material de propagación de acuerdo con la invención, partes de planta de acuerdo con la invención o plantas que se obtienen mediante un procedimiento de acuerdo con la invención, para producir una planta genéticamente modificada que sintetiza glucosaminoglucano para producir glucosaminoglucanos.

40 La presente invención también proporciona el uso de moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT-2 o que codifican una proteína que tiene la actividad de una GFAT bacteriana para la preparación de una planta genéticamente modificada.

La presente invención se refiere además a una composición que comprende células vegetales modificadas genéticamente de acuerdo con la invención.

45 En el presente documento, es inmaterial si las células vegetales están intactas o ya no están intactas debido a que han sido destruidas, por ejemplo, mediante procesamiento. Las composiciones son preferentemente productos alimenticios, suplementos alimenticios, o alimento para animales, productos farmacéuticos o cosméticos.

50 La presente invención preferentemente proporciona composiciones de acuerdo con la invención que comprenden moléculas de ácido nucleico recombinantes, caracterizándose las moléculas de ácido nucleico recombinantes por que comprenden moléculas de ácido nucleico que codifican una proteína que tiene la actividad enzimática de una GFAT-2 o de una proteína que tiene la actividad de GFAT.

55 Una integración estable de las moléculas de ácido nucleico exógenas en un genoma de una célula vegetal o una planta da como resultado moléculas de ácido nucleico exógenas flanqueadas después de la integración en el genoma de la célula vegetal o de la planta por las secuencias de ácido nucleico genómico de la planta. Por consiguiente, en una realización preferida, las composiciones de acuerdo con la invención se caracterizan por que las moléculas de ácido nucleico recombinantes presentes en la composición de acuerdo con la invención están flanqueadas por secuencias de ácido nucleico genómicas de plantas.

En el presente documento, las secuencias genómicas de ácido nucleico de plantas pueden ser cualquier secuencia naturalmente presente en el genoma de la célula vegetal de la planta utilizada para la preparación de la composición.

5 Que las moléculas de ácido nucleico recombinantes estén presentes en las composiciones de acuerdo con la invención puede demostrarse usando procedimientos conocidos por los expertos en la técnica, tales como, por ejemplo, procedimientos basados en hibridación o, preferentemente procedimientos basados en la PCR (reacción en cadena de la polimerasa).

10 Preferentemente, las composiciones de acuerdo con la invención comprenden al menos un 0,05%, preferentemente al menos un 0,1%, de manera particularmente preferente al menos un 0,5%, de manera muy particularmente preferente al menos un 1,0%, de derivados de glucosamina N-acetilada.

Preferentemente, las composiciones de acuerdo con la invención comprenden como máximo un 10%, preferentemente como máximo un 5%, de manera particularmente preferente como máximo un 3%, de manera muy particularmente preferente como máximo un 2%, de derivados de glucosamina N-acetilada.

15 Las composiciones de acuerdo con la invención ofrecen la ventaja de que tienen un contenido aumentado de derivados de glucosamina N-acetilada o un contenido aumentado de glucosaminoglucanos en comparación con composiciones que no comprenden células vegetales modificadas genéticamente. La N-acetil glucosamina tiene un efecto estimulante en el crecimiento de bifidobacterias (Liepke y col., 2002, Eur. J. Biochem. 269, 712-718). Además, se ha demostrado que la N-acetil glucosamina sirve como un sustrato para lactobacilos (por ejemplo, *Lactobacillus casei* de la subespecie *paracasei*) del intestino de pescado (Adolfo Bucio Galindo, 2004, Proefschrift, Wageningen Universiteit, ISBN 90-5808-943-6). Por consiguiente, la N-acetilglucosamina tiene un efecto positivo sobre bacterias prebióticas. Ya que las composiciones de acuerdo con la invención tienen un contenido de N-acetil glucosamina aumentado, deberán tener un efecto positivo en el crecimiento de bacterias prebióticas.

25 La invención además proporciona procedimientos para la preparación de una composición de acuerdo con la invención usando células vegetales de acuerdo con la invención, plantas de acuerdo con la invención, material de propagación de acuerdo con la invención, partes de planta de acuerdo con la invención o plantas que se obtienen mediante un procedimiento de acuerdo con la invención para producir una planta genéticamente modificada. Los procedimientos para preparar una composición de acuerdo con la invención son preferentemente procedimientos para producir productos alimenticios, alimentos para animales o suplementos alimenticios.

30 Los procedimientos para producir productos alimenticios, alimentos para animales, suplementos alimenticios, productos farmacéuticos o productos cosméticos son conocidos por los expertos en la técnica y comprenden entre otras cosas, pero no están exclusivamente limitados a, la trituración o molienda de plantas de acuerdo con la invención o de partes de planta de acuerdo con la invención.

La presente invención también proporciona composiciones que se pueden obtener mediante un procedimiento para preparar una composición de acuerdo con la invención.

35 La presente invención se refiere también al uso de células vegetales modificadas genéticamente de acuerdo con la invención o de plantas modificadas genéticamente de acuerdo con la invención para preparar una composición de acuerdo con la invención.

Una realización preferida de composiciones de acuerdo con la invención son harinas.

40 Las partes de plantas se procesan frecuentemente en harina. Los ejemplos de partes de plantas que se usan para preparar harina son, por ejemplo, tubérculos de plantas de patata y granos de plantas de cereal. Para producir harinas a partir de plantas de cereal, los granos que contienen el endospermo de estas plantas se muelen y se tamizan. En el caso de otras plantas que no contienen ningún endospermo sino, por ejemplo, tubérculos o raíces de almacenaje, la harina por lo general se produce mediante la trituración, secado y posterior molienda de las partes relevantes de las plantas. Las células vegetales de acuerdo con la invención y las plantas de acuerdo con la invención tienen un contenido aumentado de derivados de glucosamina N-acetilada o glucosaminoglucanos en comparación con células vegetales o plantas conocidas. Las harinas preparadas a partir de células vegetales de acuerdo con la invención, de plantas de acuerdo con la invención, de material de propagación de acuerdo con la invención o de partes de plantas de acuerdo con la invención por consiguiente contienen igualmente una proporción aumentada de derivados de glucosamina N-acetilada o glucosaminoglucanos.

50 Por consiguiente, la presente invención se refiere además a harinas que se obtienen a partir de células vegetales de acuerdo con la invención, de plantas de acuerdo con la invención o de partes de plantas de acuerdo con la invención. Las partes de planta preferidas de acuerdo con la invención para producir harinas son tubérculos y granos que contienen endospermos. En el contexto de la presente invención, se da particular preferencia a los granos de plantas de la familia *Poaceae* (sistemática), de manera especialmente preferente, los granos originados a partir de plantas de maíz, arroz, o trigo.

55

La presente invención se refiere además a harinas de acuerdo con la invención que tienen un contenido de derivados de glucosamina N-acetilada de por lo menos 10 μmol por gramo, preferentemente de al menos 20 μmol por gramo, más preferentemente de al menos 25 μmol por gramo, de manera particularmente preferente de al menos 30 μmol por gramo, de manera muy particularmente preferente de al menos 35 μmol por gramo y especialmente preferentemente de al menos 40 μmol por gramo.

Las harinas de acuerdo con la invención tienen preferentemente un contenido de derivados de glucosamina N-acetilada de como máximo 250 μmol por gramo de peso fresco, preferentemente de como máximo 200 μmol por gramo de peso fresco, de manera particularmente preferente de como máximo 150 μmol por gramo de peso fresco, de manera muy particularmente preferente de como máximo 100 μmol por gramo de peso fresco y especialmente preferentemente de como máximo 50 μmol por gramo de peso fresco.

En el contexto de la presente invención, se entiende que el término "harina" significa un polvo obtenido mediante la molienda de plantas o de partes de plantas. Si es adecuado, las plantas o partes de planta se secan antes de la molienda y después de la molienda además se trituran y/o tamizan.

En comparación con la harinas convencionales, las harinas de acuerdo con tienen la ventaja de que se pueden usar para producir productos alimenticios, tales como, por ejemplo, artículos horneados, que tienen un contenido aumentado de derivados de glucosamina N-acetilada o glucosaminoglucanos sin ser necesario añadir derivados de glucosamina N-acetilada o glucosaminoglucanos obtenidos de fuentes de materia prima animal o fúngica a las harinas. Las desventajas del uso de derivados de glucosamina N-acetilada o glucosaminoglucanos aislados de las fuentes de materias primas mencionadas, tales como, por ejemplo, el riesgo de que puedan contener patógenos o sustancias alergénicas, que ya se han mencionado adicionalmente anteriormente.

La presente invención además proporciona un procedimiento para producir harinas que comprende la etapa de moler las células vegetales de acuerdo con la invención, las plantas de acuerdo con la invención o las partes de planta de acuerdo con la invención.

Las harinas pueden producirse mediante la molienda de partes de planta. El experto en la técnica sabe cómo producir harinas. Un procedimiento para producir harinas también comprende preferentemente la etapa de cosechar las plantas cultivadas de acuerdo con la invención o las partes de plantas de acuerdo con la invención y/o el material de propagación de acuerdo con la invención y de manera particularmente preferente además la etapa de cultivar las plantas de acuerdo con la invención antes de la cosecha.

En una realización adicional de la presente invención, el procedimiento para producir harinas comprende el procesamiento de las plantas de acuerdo con la invención, de las partes de plantas de acuerdo con la invención o del material de propagación de acuerdo con la invención. En el presente documento, el procesamiento puede, por ejemplo, ser tratamiento por calor y/o secado. El tratamiento por calor seguido por secado del material tratado con calor se usa, por ejemplo, cuando se producen harinas de raíces de almacenamiento, tubérculos, tales como, por ejemplo, tubérculos de patata, antes de la molienda. La trituración de las plantas de acuerdo con la invención, de las partes de plantas de acuerdo con la invención o el material de propagación de acuerdo con la invención también pueden constituir el procesamiento en el sentido de la presente invención. La eliminación de tejido vegetal, es decir, por ejemplo, la eliminación de la vaina de los granos, antes de la molienda también constituye un procesamiento anterior a la molienda en el sentido de la presente invención.

En Una realización adicional de la presente invención, el procedimiento para producir harinas comprende el procesamiento del material molido después de la molienda.

En el presente documento, el material molido puede, por ejemplo, tamizarse después de la molienda, por ejemplo para producir diferentes tipos de harina.

La presente invención además proporciona el uso de células vegetales de acuerdo con la invención, de plantas de acuerdo con la invención, de partes de plantas de acuerdo con la invención o de material de propagación de acuerdo con la invención para producir harinas.

Descripción de las secuencias

SEC ID N° 1: Secuencia de ácido nucleico que codifica una hialuronano sintasa de *virus 1 de Paramecium bursaria Chlorella*.

SEC ID N° 2: Secuencia de aminoácidos de una hialuronano sintasa de *virus 1 de Paramecium bursaria Chlorella*. La secuencia de aminoácido que se muestra puede proceder de la SEC ID N° 1.

SEC ID N° 3: Secuencia sintética de ácido nucleico que codifica una hialuronano sintasa de *virus 1 de Paramecium bursaria Chlorella*. Los codones de la secuencia mostrada se sintetizaron de tal modo que se adaptan al uso de codones en célula vegetales. La secuencia de ácido nucleico representada codifica una proteína con la secuencia de aminoácido representada como SEC ID N° 2.

SEC ID N° 4: Secuencia de ácido nucleico que codifica una proteína con la actividad de una GFAT-1 de ratón.

- SEC ID N° 5: Secuencia de aminoácidos de una proteína con la actividad de GFAT-1 de ratón. La secuencia de aminoácidos representada puede proceder de la SEC ID N° 4.
- SEC ID N° 6: Secuencia de ácido nucleico que codifica una proteína con la actividad de una GFAT-2 de ratón.
- 5 SEC ID N° 7: Secuencia de aminoácidos de una proteína con la actividad de GFAT-2 de ratón. La secuencia de aminoácidos representada puede proceder de la SEC ID N° 6.
- SEC ID N° 8: Secuencia de ácido nucleico que codifica una proteína con la actividad de una GFAT bacteriana de *Escherichia coli*.
- SEC ID N° 9: Secuencia de aminoácidos de una proteína con la actividad de una GFAT de *Escherichia coli*. La secuencia de aminoácidos representada puede proceder de la SEC ID N° 8.
- 10 SEC ID N° 10: Secuencia de ácido nucleico sintética que codifica una proteína que tiene la actividad de una GFAT de *Escherichia coli*. Los codones de la secuencia mostrada se sintetizaron de tal modo que se adaptan al uso de codones en células vegetales. La secuencia de ácido nucleico mostrada codifica una proteína que tiene la secuencia de aminoácidos mostrada como SEC ID N° 9.
- 15 SEC ID N° 11: Secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*.
- SEC ID N° 12: Secuencia de aminoácido de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*. La secuencia de aminoácidos mostrada puede proceder de la SEC ID N° 11.
- 20 SEC ID N° 13: Secuencia de ácido nucleico sintética que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*. Los codones de la secuencia mostrada se sintetizaron de tal forma que se adaptaron al uso de codones en células vegetales. La secuencia de ácido nucleico mostrada codifica una proteína que tiene la secuencia de aminoácidos mostrada como la SEC ID N° 12.
- 25 SEC ID N° 14: Oligonucleótido sintético que se usó en el Ejemplo 6.
- SEC ID N° 15: Oligonucleótido sintético que se usó en el Ejemplo 6.
- SEC ID N° 16: Oligonucleótido sintético que se usó en el Ejemplo 15.
- SEC ID N° 17: Oligonucleótido sintético que se usó en el Ejemplo 15.

30 El contenido de todas las publicaciones citadas incluyendo los números de referencia de las moléculas de ácido nucleico y de las secuencias de aminoácido mencionadas para las bases de datos de secuencia se incorporan por referencia en la descripción de la solicitud.

Los procedimientos que pueden usarse en conexión con la presente invención se describen a continuación. Estos procedimientos son realizaciones específicas; sin embargo, la presente invención no se limita a estos procedimientos. Los expertos en la técnica saben que la invención puede llevarse a cabo del mismo modo mediante la modificación de los procedimientos descritos y/o mediante el reemplazo de procedimientos o partes individuales de los procedimientos por procedimientos alternativos o partes alternativas de los procedimientos.

Procedimientos Generales

1. Transformación de plantas de patata

40 Las plantas de patata se transformaron con ayuda de *Agrobacterium*, tal como se describe en Rocha-Sosa y col., (EMBO J. 8, (1989), 23-29).

2. Transformación de plantas de tomate

Las plantas de tomate se transformaron con la ayuda de *Agrobacterium* de acuerdo con el procedimiento descrito en el documento US 5.565.347.

45 3. Transformación de plantas de arroz

Las plantas de arroz se transformaron de acuerdo con el procedimiento descrito por Hiei y col., (1994, Plant Journal 6(2), 271-282).

4. Determinación del contenido de glucosaminas N-acetiladas

50 Los derivados de glucosamina N-acetilada que tienen un extremo reductor se determinaron de manera similar al procedimiento de Elson y Morgan (1933, J Biochem. 27, 1824) y al procedimiento de determinación colorimétrico mejorado de Reissig y col., (1955, Biol. Chem. 217, 959-966). El procedimiento de determinación colorimétrico se basó en una reacción del cromogen III (Muckenschnabel y col., 1998, Cancer Letters 131, 13-20) con p-dimetilaminobenzaldehído (DMAB, reactivo de Ehrlich), produciendo un producto rojo cuya concentración puede determinarse fotométricamente.

55

a) Desarrollo del material vegetal

En primer lugar, se trituró el material vegetal cosechado. Dependiendo de la cantidad de material vegetal usado, la trituración se llevó a cabo en un molino de bola oscilante de laboratorio (MM200, de Retsch, Alemania) durante 30 segundos a 30 Hz o usando una mezcladora Warring a velocidad máxima durante aproximadamente 30 segundos.

- 5 En general, se mezclaron 0,5 g de material vegetal triturado (por ejemplo, hoja, tubérculo, o grano de arroz) con 1 ml de una solución que consistía en 7% de ácido perclórico, 5 mM de EGTA y se incubó sobre hielo durante 20 minutos. La mezcla después se centrifugó (5 minutos a 16.000 x g, 4°C). El sobrenadante obtenido después de la centrifugación se extrajo y se neutralizó usando una solución que consistía en 5 M de KOH, 1 M de TEA (pH ajustado a 7,0) y después se volvió a centrifugar (5 minutos a 16.000 x g, 4°C). Después del fin de la centrifugación, el sobrenadante se absorbió, su volumen se determinó y la cantidad de derivados de glucosamina N-acetilada que tenían un extremo reductor se determinó usando el procedimiento descrito en b).

b) Determinación del contenido de derivados de glucosamina N-acetilada que tienen agentes reductores.

Se agregaron 20 µl de una solución que consistía en 0,8 M K₂B₄O₇, pH 9,6 a 100 µl del extracto vegetal obtenido mediante el procedimiento descrito en a), y, después de un mezclado vigoroso, se calentó a 95°C durante 5 minutos. Después de enfriar la mezcla a temperatura ambiente, se agregaron 0,7 ml de reactivo de Ehrlich (solución que consiste en 10 g de DMAB en 12,5 ml de HCl concentrado, 87,5 ml de ácido acético glacial; 1:10 diluido con ácido acético glacial) a la mezcla, que se mezcló otra vez y se incubó a 37°C durante 30 minutos más. La mezcla después se centrifugó a 16.000 x g durante 1 minuto, y posteriormente se determinó la densidad óptica (DO) del sobrenadante obtenido después de la centrifugación en un fotómetro a 585 nm.

c) Cálculo de la concentración de derivados de glucosamina N-acetilada

En primer lugar, se estableció una curva de calibración usando cantidades definidas de N-acetil glucosamina 6-fosfato. En este punto, se determinó la DO de las soluciones que comprendían 0 mM, 0,1 mM, 0,5 mM, 1 mM, 5 mM y 10 mM de N-acetilglucosamina 6-fosfato de acuerdo con el procedimiento descrito en b)

- 25 La curva de calibración se estableció en Microsoft Excel ajustando una línea de tendencia/regresión polinómica de segundo orden de fórmula $y = ax^2 + bx + c$ o $y = x^2 + px + q$ para los puntos medidos para concentraciones individuales. Para calcular los valores, la ecuación obtenida se resolvió para x, dando como resultado: $x = -p/2 - \text{raíz cuadrada}(p^2/4 - q)$, en donde $p = b/a$, $q = (c-y)/a$, e y es la DO medida de la muestra desconocida. Teniendo en cuenta el peso fresco utilizado, el volumen utilizado y teniendo en cuenta cualquier factor de dilución utilizado, el contenido se calculó en µmol (de la solución medida) o en µmol por g de peso fresco.

5. Aislamiento de glucosaminoglucanos de tejido vegetal usando el ejemplo de hialuronano

Para detectar la presencia de hialuronano y determinar el contenido de hialuronano en el tejido vegetal, se trabajó material vegetal del modo siguiente: se agregaron 200 µl de agua (desmineralizada, conductividad = 18 MΩ) a aproximadamente 0,3 g de material de hoja o tubérculo, y la mezcla se trituró en un molino de bola oscilante de laboratorio (MM200, de Retsch, Alemania) (30 segundos a 30 Hz). Después se agregaron 800 µl adicionales de agua (desmineralizada, conductividad = 18 MΩ), y la mezcla se mezcló bien (usando, por ejemplo, un mezclador Vortex). Los restos celulares y los componentes insolubles se separaron del sobrenadante mediante centrifugación a 16.000 x g durante 5 minutos. Una alícuota del sobrenadante obtenida se utilizó para determinar la cantidad de hialuronano.

- 40 En el caso de frutos de tomate, en cada caso se desarrolló una fruta de tomate maduro completo. Para este fin, se determinó el peso de la fruta de tomate, el tomate se trituró en una mezcladora Warring con un poco de agua, la muestra triturada se liberó restos celulares mediante centrifugación a 3.600 x g durante 30 minutos y se determinó el volumen del extracto. Una alícuota del sobrenadante obtenida se utilizó para determinar la cantidad de hialuronano.

6. Purificación de glucosaminoglucanos usando el ejemplo de hialuronano.

Después de la adición de 100 ml de agua (desmineralizada, conductividad = 18 MΩ), se trituraron aproximadamente 100 gramos de material de plantas en una mezcladora Warring a una máxima velocidad durante aproximadamente 30 segundos. Si se utilizan partes de plantas relativamente grandes, tales como, por ejemplo, tubérculos o frutos de tomate, para el aislamiento, se cortaron de antemano en trozos de un tamaño de aproximadamente 1 cm³. Posteriormente se retiró el residuo celular usando un filtro de té. El residuo celular que se había separado se volvió a suspender en 300 ml de agua (desmineralizada, conductividad = 18 MΩ) y después se volvió a retirar usando un filtro de té. Las dos suspensiones obtenidas (100 ml + 300 ml) se combinaron y centrifugaron a 13.000 x g durante 15 minutos. Se añadió NaCl al sobrenadante de centrifugación obtenido hasta que se alcanzó una concentración final del 1%. Después de que todo el NaCl se incorporase en la solución, se llevó a cabo otra precipitación mediante la adición de dos veces el volumen de etanol, mezclado vigoroso e incubación a -20°C durante toda la noche. La mezcla se centrifugó posteriormente a 13.000 g durante 15 minutos. El precipitado sedimentado obtenido después de esta centrifugación se disolvió en 100 ml de tampón (50 mM de Tris HCl, pH 8, 1 mM CaCl₂) y después se agregó proteinasa K a una concentración final de 100 µg/ml y la solución se incubó a 42°C durante 2 horas. Esto fue seguido por 10 minutos de incubación a 95°C. Una vez más, se agregó NaCl a esta solución hasta que se alcanzó

55

una concentración final del 1%. Después de que todo el NaCl se incorporase en la solución, se llevó a cabo otra precipitación mediante la adición de 2 veces el volumen de etanol, mezclado vigoroso e incubación a -20°C durante 96 horas. Esto fue seguido por 15 minutos de centrifugación a 13.000 x g. El precipitado sedimentado obtenido después de esta centrifugación se disolvió en 30 ml de agua (desmineralizada, conductividad = 18 MΩ), y una vez más, se agregó NaCl a una concentración final del 1%. Mediante la adición de dos veces el volumen de etanol, mezcla vigorosa e incubación a -20°C durante la noche, se llevó a cabo otra precipitación. El precipitado obtenido después de la centrifugación posterior a 13.000 x g durante 15 minutos se disolvió en 20 ml de agua (desmineralizada, conductividad = 18 MΩ).

Se llevó a cabo una purificación adicional mediante filtración centrífuga. Para este fin, se aplicaron en cada caso 5 ml del precipitado disuelto a un filtro de membrana (CentriconAmicon, ancho de poro 10 000 NMWL, Prod. n.º UCF8 010 96), y la muestra se centrifugó a 2.200 x g hasta que solamente quedaban aproximadamente 3 ml de la solución sobre el filtro. Dos veces más, se agregaron en cada caso 3 ml de agua (desmineralizada, conductividad = 18 MΩ) después a la solución por encima de la membrana y en cada caso se volvieron a centrifugar en idénticas condiciones, hasta que, finalmente, solamente quedaron aproximadamente 3 ml de la solución sobre el filtro. Las soluciones aún presentes sobre la membrana después de la filtración centrífuga se absorbieron, y la membrana se aclaró repetidamente (de tres a cinco veces) con aproximadamente de 1,5 ml de agua (desmineralizada, conductividad = 18 MΩ). Todas las soluciones que aún permanecían sobre la membrana y las soluciones obtenidas del aclarado se combinaron, se agregó NaCl a una concentración final del 1%, después de que se retirase toda la solución de NaCl, se agregó dos veces el volumen de etanol, se mezcló la muestra y el precipitado se obtuvo mediante almacenamiento a -20°C durante toda la noche. El precipitado obtenido después de la centrifugación posterior a 13.000 x g durante 15 minutos se disolvió en 4 ml de agua (desmineralizada, conductividad = 18 MΩ) y después se liofilizó (24 horas a una presión de 0,037 kPa, el aparato de liofilización fue Christ Alpha 1-4 de Christ, Osterode, Alemania).

7. Detección de hialuronano y determinación del contenido de hialuronano

Se detectó hialuronano usando un ensayo comercial (prueba de ácido hialurónico (HA) de Corgenix, Inc., Colorado, EE.UU., Prod. n.º 029-001) según las instrucciones del fabricante que se incorporan a la descripción del presente documento por referencia. El principio de la prueba se basó en la disponibilidad de una proteína que se une específicamente a hialuronano (HABP) y se llevó a cabo de manera similar a un ELISA, en donde una reacción de color indica el contenido de hialuronano en la muestra examinada. Por consiguiente, para la determinación cuantitativa de hialuronano, las muestras que se van a medir deberán utilizarse a una concentración tal que estén dentro de los límites establecidos (por ejemplo: dilución de la muestra en cuestión o uso de menos agua para extraer hialuronano del tejido vegetal, dependiendo de si se excedió un límite o no se alcanzó).

8. Determinación de la actividad de una GFAT

La actividad de una proteína que tiene la actividad de una GFAT se determinó tal como se describe en Rachel y col., (1996, J. Bacterial. 178 (8), 2320-2327).

Para distinguir si una proteína tiene la actividad de una GFAT-1 o GFAT-2, se utilizó el procedimiento descrito en Hu y col., (2004, J. Biol. Chem. 279 (29), 29988-29993).

9. Detección de derivados de glucosamina N-acetilada por medio de espectroscopia de masa

Para detectar los derivados de glucosamina N-acetilada mediante espectroscopia de masas, el tejido vegetal se desarrolló como en el punto 4a de Procedimientos Generales). Para obtener un extracto tan libre de sal como fuese posible, las muestras respectivas, antes del examen a través de espectroscopia de masas, se congelaron inicialmente a -20°C y se descongelaron durante la centrifugación (16.000 x g a temperatura ambiente). Para la medición, el sobrenadante se diluyó a 1:20 con una mezcla de metanol:agua a una relación de 1:1 (volumen/volumen).

Para aumentar la sensibilidad de detección para señales débiles (picos), se registraron los espectros de EM con tres diferentes sensibilidades detectoras. Sin embargo, en este caso, la respuesta del detector ya no es lineal, lo cual se observa cuando las intensidades de señal (áreas de pico) de diferentes metabolitos se comparan y deberán tenerse en cuenta. Para asegurar que las mediciones pueden compararse entre sí, se aseguró que las muestras individuales diesen intensidades de señal idénticas (en cps, cuentas por segundo) con la misma configuración del detector.

Las áreas de las señales resultantes (áreas pico) asignadas a los diferentes metabolitos se manifiestan en relación al área pico de las hexosas (m/z=179) en %. La relación de las intensidades de señal (áreas pico) en diferentes muestras se puede usar para inferir las relaciones de concentración de los derivados de glucosamina N-acetilada correspondientes en relación a la concentración de hexosas en la muestra en cuestión.

Las mediciones de EM-EM de las muestras individuales y de las sustancias de referencia correspondientes individuales (glucosamina, N-acetil glucosamina, glucosamina 6-fosfato, glucosamina 1-fosfato, N-acetil glucosamina 6-fosfato, N-acetil glucosamina 1-fosfato, UDP-N-acetilglucosamina) se llevaron a cabo en paralelo. De este modo, es posible evaluar si la señal (pico) utilizada para determinar el área es una señal generada exclusivamente por un

metabolito específico o por metabolitos isoméricos específicos que tienen la misma masa, o si la señal en cuestión se puede asignar solamente parcialmente al metabolito correspondiente o a los metabolitos isoméricos específicos correspondientes que tienen la misma masa.

5 Los espectros de EM y EM-EM se registraron de modo negativo usando un espectrómetro de masas híbrido Q-STAR Pulsar i de Applied Biosystems equipado con una fuente de nanoelectropulverización. Los iones detectados fueron principalmente iones desprotonados con una sola carga.

Las mediciones se llevaron a cabo en las siguientes condiciones:

Intervalo de masa	50-700 Da
Sensibilidad del detector:	2.000, 2.050 y 2.100

Para cada una de las configuraciones del detector, se aseguró que las muestras tuviesen intensidades de señal similares (en cps, cuentas por segundo).

10 Ejemplos

1. Adquisición de secuencias de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT de ratón

15 La secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-1 (glutamina:fuctosa 6-fosfato amidotransferasa o glucosamina 6-fosfato sintasa, EC 2.6.1.16) se compró en BioCat GMBH, Heidelberg, Alemania (n.º de art. MMM1013-65346, clon de ADNc MGC-58262, IMAGE:6742987). Este es un clon que se produce por el consorcio I.M.A.G.E. (<http://image.llnl.gov>) y se distribuye por BioCat GMBH. El ADNc codificante para una proteína que tiene la actividad de una GFAT-1 se clonó en el vector pCMV Sport 6 de Invitrogen. El plásmido se denominó IC 365-256. La secuencia de ácido nucleico que codifica la proteína que tiene la actividad de una GFAT-1 de *Mus musculus* se muestra en la SEC ID N° 4.

20 Para facilitar las etapas posteriores de clonación, la secuencia codificante de la GFAT-1 se extrajo usando *Xho* I y *Eco* RV del plásmido IC 365-256 y se clonó en el plásmido pME9, que se había cortado con la misma endonucleasa de restricción. El plásmido obtenido se denominó IC 367-256.

25 El plásmido pME9 es un vector pBlueSkript de Stratagene (n.º de prod. 212207) en donde, a diferencia del vector pBlueSkript mencionado, pME9 contiene un Sitio de Clonación Múltiple (MCS, por sus siglas en inglés) modificado que, además del MCS presente en el vector pBlueSkript, tiene un sitio de restricción de *Pac* I adicional en ambos extremos del MCS.

2. Adquisición de una secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2 de un ratón.

30 La secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2 (glutamina:fructosa 6-fosfato amidotransferasa o glucosamina 6-fosfato sintasa, EC 2.6.1.16) se compró a Invitrogen (ID de clon 4167189, clon de ADNc MGC:18324, IMAGE:4167189). Este es un clon producido por el consorcio I.M.A.G.E (<http://image.llnl.gov>) y distribuido por Invitrogen. El ADNc codificante de una proteína que tiene la actividad de una GFAT-2 se clonó en el vector pCMV Sport 6. El plásmido se denominó IC 369-256. La secuencia de ácido nucleico que codifica la proteína tiene la actividad de una GFAT-2 de *Mus musculus* que se muestra como SEC ID N° 6.

35 3. Síntesis de las secuencias de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT bacteriana de *Escherichia coli*

40 La secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT bacteriana (glutamina:fructosa 6-fosfato amidotransferasa o glucosamina 6-fosfato sintasa, gms, EC 2.6.1.16) de *Escherichia coli* se sintetizó por Emtelechon GMBH y se clonó en el vector pCR4Topo de Invitrogen (n.º de prod. K4510-20). El plásmido obtenido se denominó IC 373-256. La secuencia de ácido nucleico sintética que codifica la proteína que tiene la actividad de una GFAT bacteriana de *Escherichia coli* se muestra como SEC ID N° 10. La secuencia de ácido nucleico correspondiente originalmente aislada de *Escherichia coli* se muestra como SEC ID N° 8.

4. Síntesis de moléculas de ácido nucleico que codifican una hialuronano sintasa del Virus 1 de *Paramecium bursaria Chlorella*

45 La secuencia de ácido nucleico que codifica una hialuronano sintasa del virus 1 de *Paramecium bursaria Chlorella* se sintetizó por Medigenomix GMBH (Múnich, Alemania) y se clonó en el vector pCR2.1 de Invitrogen (n.º de prod. K2000-01). El plásmido obtenido se denominó IC 323-215. La secuencia de ácido nucleico que codifica la proteína HAS del virus 1 de *Paramecium bursaria Chlorella* se muestra como SEC ID N° 3. La secuencia de ácido nucleico correspondiente originalmente aislada del virus 1 de *Paramecium bursaria Chlorella* se muestra como SEC ID N° 1.

50 5. Síntesis de moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del Virus 1 de *Paramecium bursaria Chlorella*

La secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*, se sintetizó por Entelechon GMBH y se clonó en el vector pCR4Topo de Invitrogen (n.º de prod. K4510-20). El plásmido obtenido se denominó IC 339-222. La secuencia de ácido nucleico sintética que codifica la proteína tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella* se muestra como SEC ID N° 13. La secuencia de ácido nucleico correspondiente originalmente aislada del virus de *Paramecium bursaria Chlorella* se muestra como SEC ID N° 11.

6. Preparación del vector de expresión de planta IR 47-71

El plásmido pBinAR es un derivado del plásmido de vector binario pBin19 (Bevan, 1984, Nucl Acids Res 12:8711-8721) que se construyó del siguiente modo:

Se aisló un fragmento de 529 pb que comprende los nucleótidos 6909-7437 del promotor 35S del virus de mosaico de la coliflor como un fragmento *EcoR I/Kpn I* del plásmido pDH51 (Pietrzak y col., 1986 Nucleic Acids Res. 14, 5858) y se ligó entre los sitios de restricción *EcoR I* y *Kpn I* del polienlazador de pUC18. Esto dio el plásmido pUC18-35S. Con la ayuda de las endonucleasas de restricción *Hind III* y *Pvu II*, se aisló un fragmento de 192 pb que comprende la señal de poliadenilación (extremo 3') del gen octopin sintasa (Gen 3) del ADN-T del plásmido Ti pTiACH5 (Gielen y col., 1984, EMBO Journal 3, 835-846) (nucleótidos 11749-11939) del plásmido pAGV40 (Herrera-Estrella y col., 1983 Nature, 303, 209-213). Después de la adición de los enlazadores *Sph I* al sitio de restricción *Pvu II*, el fragmento se ligó entre los sitios de restricción de *Sph I* y *Hind III* de pUC18-35S. Esto dio el plásmido pA7. A partir de este plásmido, se extrajeron el polienlazador completo que comprende el promotor 35S y el terminador OCS con *EcoR I* y *Hind III* y se ligaron en el vector pBin19 cortado de manera adecuada. Esto dio el vector de expresión en plantas pBinAR (Höfgen y Willmitzer, 1990, Plant Science 66, 221-230).

El promotor del gen Patatin B33 de *Solanum tuberosum* (Rocha-Sosa y col., 1989, EMBO J. 8, 23-29) se ligó como el fragmento *Dra I* (nucleótidos -1512+14) en el vector pUC19, el cual se había cortado con *Sst I* y cuyos extremos se habían despuntado con la ayuda de ADN polimerasa de T4. Esto dio el plásmido pUC19-B33. Usando *EcoR I* y *Sma I*, el promotor B33 se extirpo de este plásmido y se ligó en un vector pBinAR cortado de manera adecuada. Esto dio el vector de expresión de planta pBinB33.

Para facilitar las etapas de clonación adicionales, se amplió el MCS (Sitio de Clonación Múltiple). Para este fin, se sintetizaron dos oligonucleótidos complementarios, se calentaron a 95°C durante 5 minutos y se enfriaron lentamente hasta temperatura ambiente, y se clonó el fragmento bicatenario en los sitios de restricción de *Sal I* y *Kpn I* de pBinB33. Los oligonucleótidos utilizados para este fin tenían la secuencia siguiente:

5'-TCG ACA GGC CTG GAT CCT TAA TTA AAC TAG TCT CGA GGA GCT CGG TAC-3'

5'-CGA GCT CCT CGA GAC TAG TTT AAT TAA GGA TCC AGG CCT G-3'

El plásmido obtenido se denominó IR 47-71.

7. Preparación del vector de expresión de planta pBinARHyg

Usando las endonucleasas de restricción *EcoR I* y *Hind III*, se extrajeron el fragmento que comprende el promotor 35S, el terminador OCS y el sitio de clonación múltiple completo del plásmido pA7 y se clonaron en el vector pBIBHyg (Becker, 1990, Nucleic Acids Res. 18, 203) que se había cortado con la mismas endonucleasas de restricción. El plásmido obtenido se denominó pBinARHyg.

8. Preparación del vector de clonación IC 317-204

Se aislaron fragmentos de ácido nucleico que comprenden el terminar OCS del plásmido IR 47-71 usando endonucleasas de restricción *Xho I* y *Hind III* y se clonaron en el vector pBlueSkript KS (de Stratagene, n.º de prod. 212207), que se había cortado con las mismas endonucleasas de restricción. El plásmido obtenido se denominó IC 306-204.

Los fragmentos de ácido nucleico que comprenden el promotor B33 se aislaron del plásmido IR 47-71 usando las endonucleasas de restricción *Bam HI* y *Eco RI* y se clonaron en el vector pBlueSkript KS (de Stratagene, n.º de prod. 212207), se había cortado con las mismas endonucleasas de restricción. El plásmido obtenido se denominó IC 314-204. A partir de IC 306-204, se aisló el terminador OCS aisló usando la endonucleasa de restricción *Bam HI* y se clonó en el plásmido IC 314-204, que se había cortado con las mismas endonucleasas de restricción. El plásmido obtenido se denominó IC 317-204.

9. Preparación del vector de expresión de planta IC 341-222 que comprende una secuencia de ácido nucleico codificante de una hialuronano sintasa del Virus 1 de *Paramecium bursaria Chlorella*

Mediante la digestión de restricción con *BamHI* y *Xho I*, las moléculas de ácido nucleico que comprenden la secuencia codificante de hialuronano sintasa se aislaron a partir del plásmido IC 323-215 y se clonaron en los sitios de restricción *BamHI* y *Xho I* del plásmido IR 47-71. El vector de expresión de planta obtenido se denominó IC 341-

222.

10. Preparación de los vectores de expresión 349-222 que comprenden la secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del Virus 1 de *Paramecium bursaria Chlorella*

5 Usando la digestión de restricción con *BamH* I y *Kpn* I, las moléculas de ácido nucleico que comprenden la secuencia codificante para una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 *Paramecium bursaria Chlorella*, se aislaron a partir del plásmido IC 339-222 y se clonaron en el plásmido pA7, que había sido cortado con las mismas endonucleasas de restricción. El plásmido obtenido se denominó IC 342-222.

10 Mediante la digestión de restricción con *Xba* I y *Kpn* I, las moléculas de ácido nucleico que comprenden la secuencia codificante de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*, se aislaron a partir del plásmido IC 342-222 y se clonaron en el vector de expresión pBinAR Hyg, que se había cortado con *Xba* I y *Kpn* I. El plásmido obtenido se denominó IC 349-222.

15 11. Preparación de los vectores de expresión de planta IC 376-271 que comprenden secuencias de ácido nucleico codificantes para una proteína que tiene la actividad de una GFAT-1 de ratón y para una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*.

Un fragmento de ácido nucleico que comprende el promotor B33 y terminador OCS, habiéndose aislado dicho fragmento a partir de IC 317-204 mediante digestión de restricción usando *Eco* RI, se clonó en el sitio de restricción de *Eco* RI del plásmido IC 349-222. En el presente documento, se aseguró la orientación de cabeza a cabeza de los promotores (25S y B33). El vector obtenido se denominó IC 354-22.

20 Para obtener un vector de expresión de planta que comprende una secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-1 de ratón, se aisló la secuencia codificante de la proteína que tiene la actividad de una GFAT-1 de ratón mediante digestión de restricción con *Xho* I y *Eco* RV de IC 365-256 y se clonó en el plásmido IC 354-222, que se había cortado con *Xho* I y *Ecl136* II. El vector de expresión de planta obtenido se denominó IC 376-256.

25 12. Preparación del vector de expresión de planta IC 372-256 que comprende las secuencias de ácido nucleico codificantes de una proteína que tiene la actividad de una GFAT-2 de ratón y de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del virus 1 de *Paramecium bursaria Chlorella*

30 Se aisló un fragmento de ácido nucleico que comprende la secuencia codificante de la proteína que tiene la actividad de una GFAT-2 de ratón a partir de IC 369-256 mediante digestión de restricción con *Xho* I y *Eco* RV y se clonó en el plásmido IC- 354-222, que se había cortado con *Xho* I y *Ecl136* II. El vector de expresión de planta obtenido se denominó IC 372-256.

13. Preparación del vector de expresión de planta 375-271 que comprende las secuencias de ácido nucleico codificantes de una proteína que tiene la actividad de una GFAT de *Escherichia coli* y de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa del Virus 1 *Paramecium bursaria Chlorella*

35 Se aisló un fragmento de ácido nucleico que comprende la secuencia codificante de la proteína que tiene la actividad de una GFAT de *Escherichia coli* a partir de IC 373-256 mediante digestión de restricción con *Xho* I y *Eco* RV y se clonó en el plásmido IC 354-222, que se había cortado con *Xho* I y *Ecl136* II. El vector de expresión de planta obtenido se denominó IC 375-271.

40 14. Preparación del vector de expresión de planta IC 398-311 que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una GFAT de *Escherichia coli*

Mediante la digestión de restricción con *Ecl136* I y *Xho* I, se aisló la secuencia codificante de la proteína que tiene la actividad de una GFAT bacteriana de *E. coli* a partir del plásmido IC 373-256 y se ligó en los sitios de restricción *Sma* I y *Sal* I del vector pBinAR Hyg. El vector de expresión de planta obtenido se denominó IC 398-311.

15. Preparación del vector de expresión de planta IC 386-299

45 Mediante PCR usando ADN genómico aislado de hojas de arroz (*Oryza sativa*, planta de cultivo M2002) usando ADN polimerasa (Expand High Fidelity PCR Systems, Roche n.º de prod.: 1732641), se aisló el ADN del promotor de prolamina de arroz (n.º de ref. EMBL D63901, Sha y col., 1996, Biosci. Biotech. Biochem. 60, 335-337, Wu y col., 1998. Plant Cell Physiol. 39(8), 885-889). El amplicón obtenido a partir de esta reacción PCR se clonó en el vector pCR 2.1 usando el kit de clonación TA (Invitrogen n.º de prod.: KNM2040-01). El plásmido obtenido se denominó MI 4-154. Las condiciones usadas para la amplificación del ADN que codifica para el promotor de prolamina: Se utilizaron las condiciones y tampones indicados por el fabricante y 50 ng del ADN total.

Mezcla 0.83 µM dNTP

0.25 µM del cebador prol-F1

5'-AAAACTAGTTCTACATCGGCTTAGGTGTAGCAACACG

0.25 µm del cebador prol-R1
5'-AAAAGATATCTGTTGTTGGATTCTACTACTATGCTTCAA

Condiciones de Reacción:

5	Etapa 1	15 s
	Etapa 2	60°C 15 s
	Etapa 3	72°C 45 s

10 En primer lugar, se llevó a cabo la reacción de acuerdo con las etapas 1 a 3 usando 35 repeticiones (ciclos). Después de terminar la reacción, la mezcla de reacción se enfrió a 4°C. La clonación posterior en el vector pCR 2.1 usando el kit de clonación TA (Invitrogen n.º de prod.: KNM2040-01) se llevó a cabo siguiendo las condiciones indicadas por el fabricante. El plásmido que comprende el promotor de prolamina de arroz se denominó MI 4-154.

15 Un fragmento de ácido nucleico que comprende la secuencia codificante de la proteína que tiene la actividad de GFAT-2 de ratón se aisló mediante digestión de restricción usando las endonucleasas de restricción *Not* I y *Kpn* I del plásmido IC 369-256 y se clonó en los vectores pMCS5 (comprado a MoBiTec), que había sido digerido con *Not* I y *Kpn* I. El plásmido obtenido se denominó IC 385-299. En la etapa siguiente, el fragmento de ácido nucleico que comprende la secuencia codificante de la proteína que tiene la actividad de una GFAT-2 de ratón se aisló mediante digestión de restricción con las endonucleasas de restricción *Xho* I y *Hpa* I de IC 385-299 y se clonó en el plásmido MI 9-154, que se había cortado con *Xho* I y *Ecl*136 II. El vector de expresión de planta obtenido se denominó IC 386-299. El vector de partida de la preparación del vector MI 9-154 es el plásmido ML 18-56 (documento WO 05/030941). Un MCS sintetizado mediante dos oligonucleótidos, y que tiene los extremos adherentes adecuados y que comprende los sitios de restricción *Pst* I, *Sac* I, *Bln* I, *Xho* I, *Hpa* I, *Spe* I y *Hind* III se introdujo en el plásmido ML 18-56, que había digerido con *Hind* III y *Pst* I. El vector obtenido se denominó MI 8-154.

20 Mediante la digestión con *Eco* RV y *Spe* I, el promotor de prolamina se aisló a partir de MI 4-154 y se ligó en el vector MI 8-154, que se había digerido con *Hpa* I y *Spe* I. El vector obtenido se denominó MI 9-154.

25 16. Plantas de patata que comprenden una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT bacteriana

a) Transformación de las plantas de patata

30 Las plantas de patata (variedad de cultivo Désirée) se transformaron mediante el procedimiento proporcionado en el punto 1 de Procedimientos Generales usando el vector de expresión de planta IC 398-311, que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una GFAT bacteriana de *Escherichia coli* bajo el control del promotor del gen patatin B33 de *Solanum Tuberosum* (Rocha-Sosa y col., 1989, EMBO J. 8, 23-29). Las líneas transgénicas obtenidas, que se transformaron con el plásmido IC 398-311, se denominaron 432 ES.

b) Análisis de las líneas 432 ES

35 Las plantas de la línea 432 ES se cultivaron en un invernadero con tierra en macetas de 6 cm. En cada caso se desarrollaron aproximadamente 0,3 g a 0,8 g de material de hoja, cosechado de plantas individuales de acuerdo con el procedimiento descrito en el punto 4 de Procedimientos Generales, y se determinó el contenido de los derivados de glucosamina N-acetilada. Para las plantas individuales que tienen un contenido aumentado de derivados de glucosamina N-acetilada, se obtuvieron los siguientes resultados:

40 Tabla 1: Cantidad de derivados de glucosamina N-acetilada (en µmol por gramo de peso fresco) medido en hojas de plantas transgénicas independientes de la línea 432 ES. La columna 1 se refiere en cada caso a la planta, obtenida independientemente de la transformación, a partir de la cual se cosechó el material ("ts" se refiere a plantas que no han sido transformadas).

Planta	µmol/g FW	Planta	µmol/g FW
432ES 1	2,97	432ES 25	4,45
432ES 2	0,51	432ES 26	1,81
432ES 4	2,19	432ES 27	1,75
432ES 5	3,99	432ES 28	0,45
432ES 6	6,20	432ES 32	4,56
432ES 7	2,98	432ES 33	3,64
432ES 8	0,48	432ES 35	3,64
432ES 9	11,48	432ES 37	6,67
432ES 10	0,30	432ES 38	0,95

(continuación)

Planta	$\mu\text{mol/g FW}$	Planta	$\mu\text{mol/g FW}$
432ES 11	6,89	432ES 40	8,69
432ES 12	5,45	432ES 42	1,47
432ES 13	0,23	432ES 43	5,41
432ES 14	0,80	432ES 44	6,33
432ES 15	1,75	432ES 45	3,39
432ES 16	4,87	ts 1	0,05
432ES 18	3,38	ts 2	0,26
432ES 19	6,38	ts 3	0,17
432ES 21	1,42		
432ES 22	9,73		
432ES 23	5,88		

Estos resultados demuestran que las plantas que tienen una molécula de ácido nucleico exógena codifican una proteína que tiene la actividad de una GFAT bacteriana tienen un contenido considerablemente mayor de derivados de glucosamina N-acetilada que las plantas de tipo silvestre no transformadas correspondientes.

- 5 17. Plantas de arroz que comprenden una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2

a) Transformación de las plantas de arroz

10 Las plantas de arroz (variedad M202) se transformaron de acuerdo con el procedimiento proporcionado en el punto 3 de procedimientos generales con el vector de expresión de planta IC 386-299, que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una GFAT-2 de ratón bajo el control del promotor del polipéptido de prolamina de 13 kDa. Las líneas transgénicas obtenidas, que se transforman con el plásmido IC 386-299, se denominaron GAOS0788.

b) Análisis de las líneas GAOS0788

15 Las plantas independientes, originadas de la transformación con el plásmido IC 386-299, de la línea GAOS788 se cultivaron en tierra en un invernadero. A partir de cada planta, se cosecharon aproximadamente 20-25 semillas maduras (granos), se retiraron las vainas con un extractor de vainas (Laboratory Paddy sheller, Grainman, Miami, Florida, USA) y aproximadamente 7 semillas de arroz pardo (grupos) de cada línea se trituraron en un molino de bola oscilante de laboratorio (MM200, de Retsch, Alemania, 30 segundos a 30 Hz), dando como resultado una
20 harina. Usando el procedimiento descrito en el punto 4 de procedimientos generales, se determinó el contenido de derivados de glucosamina N-acetilada. Para las plantas individuales que tienen un contenido aumentado de derivados de glucosamina N-acetilada, se obtuvieron los siguientes resultados:

25 Tabla 2: Cantidad de derivados de glucosamina N-acetilada (en μmol por gramo de peso fresco) medido en grupos de semillas maduras de plantas transgénicas independientes de la línea GAOS0788. La columna 1 se refiere a la planta, independientemente obtenida de la transformación, a partir de la cual se cosechó el material (aquí, "control" se refiere a plantas transformadas con un plásmido que no tiene una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT. Las cantidades no detectables se marcaron como "n.d.".

Planta	$\mu\text{mol/g FW}$	Planta	$\mu\text{mol/g FW}$
GAOS0788-00101	14,50	GAOS0788-02301	7,82
GAOS0788-00202	17,36	GAOS0788-02401	20,89
GAOS0788-00301	14,46	GAOS0788-02501	6,67
GAOS0788-00501	23,07	GAOS0788-02601	7,34
GAOS0788-00602	7,75	GAOS0788-02701	4,31
GAOS0788-00701	4,44	GAOS0788-02802	8,02
GAOS0788-00802	17,43	GAOS0788-02901	4,74
GAOS0788-00901	10,13	GAOS0788-03001	4,36
GAOS0788-01001	6,38	GAOS0788-03101	11,83
GAOS0788-01202	8,32	GAOS0788-03202	2,76

(continuación)

Planta	$\mu\text{mol/g FW}$	Planta	$\mu\text{mol/g FW}$
GAOS0788-01401	8,64	GAOS0788-03302	12,82
GAOS0788-01502	2,97	Control	n.d.
GAOS0788-01602	8,15	Control	n.d.
GAOS0788-01701	16,50		
GAOS0788-02002	5,65		
GAOS0788-02202	5,15		

c) Análisis de las semillas individuales de las plantas GAOS0788-02401 y GAOS0788-00501

5 Las semillas cosechadas en el Ejemplo b) procedentes de plantas obtenidas directamente después de la transformación, siendo estas plantas heterocigóticas con respecto a los sitios de integración respectivos de los ADN-T en cuestión. Por consiguiente, como resultado de las leyes de Mendel de herencia, los grupos de semillas analizados contenían semillas que comprendían varias cantidades de ADN-T en cuestión, también siendo posible que las semillas individuales que no tienen ningún ADN-T integrado mediante transformación estén presentes en los respectivos grupos. Por lo tanto, las semillas pardas individuales de las plantas de la línea GAOS0788-02401 y de las plantas de la línea GAOS0788-00501 se examinaron cada una mediante el procedimiento descrito en el punto 4 de Procedimientos Generales respecto su contenido de derivados de glucosamina N-acetilada. Se obtuvieron los siguientes resultados:

10 Tabla 3: Cantidad de derivados de glucosamina N-acetilada (en μmol por gramo de peso fresco) de semillas individuales de las plantas de las líneas GAOS9788-02401 y GAOS0788-00501. En cada caso, la columna 1 se refiere a la planta, independientemente obtenida mediante la transformación, a partir de la cual las semillas individuales se cosecharon y se analizaron (aquí, "control" se refiere a semillas de plantas transformadas con una construcción que no comprende una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT). Las cantidades no detectables se marcaron como "n.d."

Muestra	$\mu\text{mol/g FW}$
Semilla 1 GAOS0788-02401	n.d.
Semilla 2 GAOS0788-02401	22,41
Semilla 3 GAOS0788-02401	38,47
Semilla 4 GAOS0788-02401	16,57
Semilla 5 GAOS0788-02401	17,67
Semilla 6 GAOS0788-02401	3,79
Semilla 7 GAOS0788-02401	10,14
Semilla 8 GAOS0788-02401	18,70
Semilla 1 GAOS0788-00501	n.d.
Semilla 2 GAOS0788-00501	17,20
Semilla 3 GAOS0788-00501	19,89
Semilla 4 GAOS0788-00501	15,47
Semilla 5 GAOS0788-00501	9,31
Semilla 6 GAOS0788-00501	20,88
Semilla 7 GAOS0788-00501	25,31
Semilla 8 GAOS0788-00501	31,92
Semilla 9 GAOS0788-00501	28,82
Semilla 10 GAOS0788-00501	43,35
Semilla de control 1	n.d.
Semilla de control 2	n.d.
Semilla de control 3	n.d.
Semilla de control 3	n.d.

Los resultados obtenidos muestran que las harinas de semillas (granos) de plantas de raíz tienen una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2 que tiene un contenido

considerablemente mayor de derivados de glucosamina N-acetilada en comparación con las harinas producidas de plantas que no tienen una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2.

5 18. Síntesis de derivados de glucosamina N-acetilada en plantas de tomate transformadas con moléculas de ácido nucleico que codifican varias isoformas de una proteína que tiene la actividad de una GFAT

a) Producción de plantas de tomate que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1

10 Las plantas de tomate (variedad de cultivo MoneyMaker) se transformaron mediante el procedimiento dado en el punto 2 de Procedimientos Generales con el vector de expresión de plantas IC 376-271, que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una glucosa deshidrogenasa y una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1. Las líneas transgénicas obtenidas, que se transformaron con el plásmido 376-271, se denominaron 420 ES. Las proteínas que tienen la actividad de una UDP-glucosa deshidrogenasa, catalizan la síntesis de UDP-GlcA de UDP-glucosa. Además de GlcNAc, algunas glucosaminoglucano sintasas, tales como, por ejemplo, hialuronano sintasa, necesitan UDP-GlcA como sustrato.

15 b) Análisis de las líneas 420 ES

20 Las plantas de la línea 420 ES se cultivaron en hidrocultivos en macetas en un invernadero. En cada caso aproximadamente 5 g del material de planta, cosechado de plantas individuales, se desarrolló usando el procedimiento descrito en el punto 4 de Procedimientos Generales, y el contenido de derivados de glucosamina N-acetilada se determinó. En el presente documento, por planta, se llevaron a cabo una diversidad de mediciones independientes para cada muestra trabajada. Se obtuvieron los siguientes resultados para las plantas individuales:

Planta	µmol/g FW	Media [µmol/FW]
420ES 1 a	0,15	0,15
420ES 1 b	0,10	
420ES 1 c	0,15	
420ES 2 a	0,11	0,10
420ES 2 b	0,09	
420ES 2 c	0,10	
420ES 3 a	0,29	0,28
420ES 3 b	0,30	
420ES 3 c	0,25	
420ES 4 a	0,20	0,18
420ES 4 b	0,10	
420ES 4 c	0,16	
420ES 5 a	0,10	0,09
420ES 5 b	0,08	
420ES 5 c	0,09	
420ES 6 a	0,24	0,27
420ES 6 c	0,29	
420ES 7 b	1,12	1,31
420ES 7 c	1,50	
420ES 8 a	0,05	0,06
420ES 8 b	0,05	
420ES 8 c	0,06	

Planta	µmol/g FW	Media [µmol/FW]
420ES 16 a	0,08	0,07
420ES 16 b	0,06	
420ES 16 c	0,07	
420ES 17 a	0,08	0,07
420ES 17 b	0,07	
420ES 17 c	0,08	
420ES 18 a	0,07	0,08
420ES 18 b	0,09	
420ES 18 c	0,09	
420ES 19 a	0,03	0,03
420ES 19 b	0,00	
420ES 19 c	0,05	
420ES 20 a	0,04	0,06
420ES 20 b	0,07	
420ES 20 c	0,05	
420ES 22 a	0,08	0,08
420ES 22 b	0,07	
420ES 22 c	0,08	
420ES 23 a	0,14	0,13
420ES 23 b	0,11	
420ES 23 c	0,13	
420ES 24 a	0,05	0,05

(continuación)

Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]	Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]
420ES 9 a	0,02	0,02	420ES 24 c	0,04	0,06
420ES 9 b	0,01		420ES 24 c	0,05	
420ES 9 c	0,02		420ES 25 a	0,05	
420ES 10 a	0,05	0,04	420ES 25 b	0,07	0,09
420ES 10 b	0,03		420ES 25 c	0,06	
420ES 10 c	0,05		420ES 26 a	0,13	
420ES 11 a	0,06	0,06	420ES 26 b	0,06	0,08
420ES 11 b	0,10		420ES 26 c	0,08	
420ES 11 c	0,03		420ES 27 a	0,09	
420ES 12 a	0,09	0,08	420ES 27 b	0,10	0,01
420ES 12 b	0,06		420ES 27 c	0,05	
420ES 13 a	0,02		420ES 28 a	0,01	
420ES 13 b	0,01	0,02	420ES 28 b	0,02	0,08
420ES 13 c	0,03		420ES 28 c	0,01	
420ES 14 a	0,02		420ES 29 a	0,09	
420ES 14 b	0,04	0,03	420ES 29 b	0,07	0,03
420ES 14 c	0,04		420ES 29 c	0,07	
420ES 15 a	0,05		420ES 30 a	0,04	
420ES 15 b	0,06	0,06	420ES 30 b	0,03	0,10
420ES 15 c	0,06		420ES 30 c	0,01	
			ts 7a	0,09	
				0,01	
		ts 7b	0,11		
		ts 7c	0,09		
				0,01	
		ts 12a	0,02		
		ts 12b	n.d.		
			ts 12c	0,03	

Tabla 4: Cantidad de derivados de glucosamina N-acetilada (en μmol por gramo de peso fresco) medida en hojas de plantas transgénicas independientes de la línea 420 ES.

5 La columna 1 se refiere a la planta, independientemente de su origen de la transformación, a partir de la cual el material se cosecho (en el presente documento, "ts" se refiere a plantas no transformadas). La extensión de los nombres de las plantas por a, b o c denota mediciones independientes llevadas a cabo para las muestras desarrolladas en cuestión. Las cantidades no detectables se marcan como "n.d."

10 Estos resultados muestran que las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1 y codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa, tienen un contenido de derivados de glucosamina N-acetilada que es ligeramente mayor que la de las plantas de tipo silvestre no transformadas correspondientes.

c) Producción de plantas de tomate que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2

5 Las plantas de tomate (variedad de cultivo MoneyMaker) se transformaron mediante el procedimiento dado en el punto 2 de Procedimientos Generales con el vector de expresión de planta IC 372-256 que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2. Las líneas transgénicas obtenidas, que se transformaron con el plásmido IC 372-256, se denominaron 421 ES.

d) Análisis de las líneas 421 ES

10 Las plantas de la línea 421 ES se cultivaron en hidrocultivos en macetas en un invernadero. En cada caso aproximadamente 5 g del material de planta, se cosechó de plantas individuales, se desarrolló usando el procedimiento descrito en el punto 4 de Procedimientos Generales, y el contenido de derivados de glucosamina N-acetilada se determinó. En el presente documento, por planta, se llevó a cabo una diversidad de mediciones independientes para cada muestra trabajada. Se obtuvieron los siguientes resultados para las plantas individuales:

Planta	µmol/g FW	Media [µmol/FW]
421ES 1 a	0,60	0,67
421ES 1 b	0,61	
421ES 1 c	0,80	
421ES 3 a	1,07	1,10
421ES 3 b	1,07	
421ES 3 c	1,17	
421ES 4 a	2,22	2,00
421ES 4 b	1,88	
421ES 4 c	1,89	
421ES 5 a	0,79	0,87
421ES 5 b	1,07	
421ES 5 c	0,74	
421ES 6 a	0,62	0,74
421ES 6 b	0,76	
421ES 6 c	0,85	
421ES 7 a	1,20	1,01
421ES 7 b	1,01	
421ES 7 c	0,84	
421ES 9 a	0,35	0,40
421ES 9 b	0,36	
421ES 9 c	0,48	
421ES 10 a	0,80	0,16
421ES 10 b	0,18	
421ES 10 c	0,22	
421ES 11 a	2,96	2,78
421ES 11 b	2,61	
421ES 11 c	2,78	

Planta	µmol/g FW	Media [µmol/FW]
421ES 27 a	0,11	0,13
421ES 27 b	0,16	
421ES 27 c	0,12	
421ES 28 a	0,01	0,01
421ES 28 b	n.d.	
421ES 28 c	0,01	
421ES 29 a	0,35	0,40
421ES 29 b	0,46	
421ES 29 c	0,39	
421ES 31 a	0,14	0,14
421ES 31 b	0,16	
421ES 31 c	0,11	
421ES 32 a	0,04	0,04
421ES 32 b	0,01	
421ES 32 c	0,05	
421ES 33 a	0,12	0,11
421ES 33 b	0,08	
421ES 33 c	0,13	
421ES 34 a	0,32	0,34
421ES 34 b	0,37	
421ES 34 c	0,34	
421ES 35 a	0,20	0,21
421ES 35 b	0,24	
421ES 35 c	0,17	
421ES 36 a	0,07	0,06
421ES 36 b	0,07	
421ES 36 c	0,03	

(continuación)

Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]	Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]
421ES 12 a	1,13	0,96	421ES 37 a	0,12	0,12
421ES 12 b	0,82		421ES 37 b	0,11	
421ES 12 c	0,93		421ES 37 c	0,14	
421ES 19 a	0,04	0,04	421ES 38 a	0,32	0,34
421ES 19 b	0,03		421ES 38 b	0,34	
421ES 21 a	0,21	0,25	421ES 38 c	0,37	
421ES 21 b	0,36		ts 8 a	n.d.	n.d.
421ES 21 c	0,19		ts 8 c	n.d.	
421ES 23 a	0,01	0,01	ts 13 a	n.d.	n.d.
421ES 23 b	0,01		ts 13 b	n.d.	
421ES 23 c	0,02		ts 13c	n.d.	
421ES 26 a	0,18	0,16			
421ES 26 b	0,19				
421ES 26 c	0,10				

Tabla 5: Cantidad de derivados de glucosamina N-acetilada (en μmol por gramo de peso fresco) medida en hojas de plantas transgénicas independientes de la línea 421 ES. La columna 1 se refiere a la planta, independientemente originada de la transformación, a partir de la cual el material se cosecho (aquí, "ts" se refiere a las plantas no transformadas). La extensión de los nombres de las plantas por a, b o c denota mediciones independientes llevadas a cabo para las muestras desarrolladas en cuestión. Las cantidades no detectables se marcan como "n.d."

Estos resultados muestran que las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 y la codificación para una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa, tienen un contenido de derivados de glucosamina N-acetilada que es considerablemente mayor que la de las plantas de tipo silvestre no transformadas correspondientes.

e) Producción de plantas de tomate que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT bacteriana

Las plantas de tomate (variedad de cultivo Moneymaker) se transformaron mediante el procedimiento dado en el punto 2 de Procedimientos Generales con el vector de expresión de planta IC 375-271 que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT bacteriana. Las líneas transgénicas obtenidas, las cuales se transformaron con el plásmido IC 375-271, se denominaron 422 ES.

f) Análisis de las líneas 422 ES

Las plantas de la línea 422 ES se cultivaron en hidrocultivos en macetas en un invernadero. En cada caso aproximadamente 5 g del material de planta, se cosechó de plantas individuales, se desarrolló usando el procedimiento descrito en el punto 4 de Procedimientos Generales, y el contenido de derivados de glucosamina N-acetilada se determinó. Aquí, por planta, una pluralidad de mediciones independientes se llevó a cabo para cada muestra desarrollada. Se obtuvieron los siguientes resultados para las plantas individuales:

Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]	Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]
422ES 2 a	13,96	14,50	422ES 13 a	13,11	10,89
422ES 2 b	13,39		422ES 13 b	10,54	
422ES 2 c	16,13		422ES 13 c	9,02	

(continuación)

Planta	µmol/g FW	Media [µmol/FW]	Planta	µmol/g FW	Media [µmol/FW]
422ES 3 a	0,28	0,29	422ES 14 a	7,68	7,75
422ES 3 b	0,29		422ES 14 b	8,05	
422ES 3 c	0,30		422ES 14 c	7,52	
422ES 4 a	0,20	9,97	422ES 16 a	14,02	14,45
422ES 4 b	0,13		422ES 16 b	13,35	
422ES 4 c	0,21		422ES 16 c	15,98	
422ES 5 a	10,57	9,97	422ES 17 a	10,79	9,72
422ES 5 b	9,74		422ES 17 b	9,99	
422ES 5 c	9,60		422ES 17 c	8,37	
422ES 6 a	16,58	16,20	422ES 18 a	3,09	4,20
422ES 6 b	16,11		422ES 18 b	4,55	
422ES 6 c	15,91		422ES 18 c	4,96	
422ES 7 a	3,13	2,99	422ES 19 a	6,43	5,99
422ES 7 b	2,64		422ES 19 b	4,94	
422ES 7 c	3,19		422ES 19 c	6,59	
422ES 8 a	16,50	14,70	422ES 20 a	15,85	15,50
422ES 8 b	14,32		422ES 20 b	15,87	
422ES 8 c	13,27		422ES 20 c	14,79	
422ES 9 a	9,76	9,72	422ES 21 a	0,20	0,35
422ES 9 b	9,33		422ES 21 c	0,24	
422ES 9 c	10,07		ts 9 a	0,36	0,23
422ES 11 a	5,80	ts 9 b	0,19		
422ES 11 b	5,34	ts 9 c	0,13		
422ES 11 c	5,05	5,40	ts 14 a	n.d.	n.d.
422ES 12 a	11,57		ts 14 b	n.d.	
422ES 12 b	11,65		ts 14 c	n.d.	
422ES 12 c	13,46	12,23			

Tabla 6: Cantidad de derivados de glucosamina N-acetilada (en µmol por gramo de peso fresco) medida en hojas de plantas transgénicas independientes de la línea 422 ES. La columna 1 se refiere a la planta, independientemente de su origen de la transformación, a partir de la cual el material se cosecho (en el presente documento, “ts” se refiere a las plantas no transformadas). La extensión de los nombres de las plantas por a, b o c denota mediciones independientes llevadas a cabo para las muestras desarrolladas en cuestión. Las cantidades no detectables se marcan como “n.d.”.

g) Análisis de frutos de las líneas 420 ES, 421 ES y 422 ES

Los frutos maduros se cosecharon de plantas seleccionadas de las líneas 420 ES, 421 ES y 422 RS. Se cosecharon varios frutos de tomate completas de plantas individuales y se desarrollaron el procedimiento descrito en el punto 4 de Procedimientos Generales, y se determinó el contenido de derivados de glucosamina N-acetilada. Aquí, las mediciones independientes se llevaron a cabo para diferentes frutos de una planta. Se obtuvieron los siguientes resultados para las plantas individuales:

ES 2 562 211 T3

Planta	µmol/g FW	Media [µmol/FW]	Planta	µmol/g FW	Media [µmol/FW]
420ES 2 I	0,01	0,01	422ES 2 I	3,17	4,26
420ES 2 II	0,01		422ES 2 II	3,74	
420ES 3 I	0,07	0,06	422ES 2 III	5,79	
420ES 3 II	0,06		422ES 2 IV	4,90	
420ES 4 I	0,04	0,04	422ES 2 V	3,73	
420ES 4 II	n.d.		422ES 5 I	2,12	3,41
420ES 4 III	0,04		422ES 5 II	1,76	
420ES 6 I	0,09	0,05	422ES 5 III	1,99	
420ES 6 II	0,03		422ES 5 IV	3,26	
420ES 6 III	0,04		422ES 5 V	5,27	
420ES 7 I	0,01	0,03	422ES 5 VI	4,49	
420ES 7 II	0,04		422ES 5 VII	4,95	
420ES 7 III	0,05		422ES 6 I	7,41	3,34
420ES 7 IV	0,03		422ES 9 I	3,67	
420ES 8 I	0,03	0,04	422ES 9 II	3,02	1,98
420ES 8 II	0,04		422ES 11 I	2,55	
420ES 8 III	0,04		422ES 11 II	1,92	
420ES 12 I	0,00	0,04	422ES 11 III	1,47	3,50
420ES 12 II	0,07		422ES 12 I	3,76	
420ES 17 I	0,05	0,05	422ES 12 II	9,80	
420ES 17 II	0,06		422ES 12 III	9,39	
420ES ts 7 I	n.d.	0,03	422ES 13 I	5,79	7,57
420ES ts 7 II	0,04		422ES 13 II	5,04	
420ES ts 7 III	0,03		422ES 13 III	5,11	
421ES 4 II	0,85	0,86	422ES 14 I	4,08	2,30
421ES 4 III	0,67		422ES 16 I	2,62	
421ES 4 IV	0,79		422ES 16 II	2,72	
421ES 4 V	1,02		422ES 16 III	5,45	
421ES 5 I	0,35	0,53	422ES 17 I	7,25	0,02
421ES 5 II	0,72		422ES 17 II	7,89	
421ES 5 III	0,67		422ES 18 I	n.d.	
421ES 5 IV	0,45		422ES 18 II	2,56	
421ES 5 V	0,48		422ES 18 II	2,04	
421ES 21 I	2,02	1,17	422ES ts 9 I	0,02	0,02
421ES 21 II	0,92		422ES ts 9 II	0,01	
421ES 21 III	0,96		422ES ts 9 III	0,00	
			422ES ts 9 IV	0,05	

(continuación)

Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]	Planta	$\mu\text{mol/g}$ FW	Media [$\mu\text{mol}/\text{FW}$]
421ES 21 IV	0,79	0,76	422ES ts 9 V	n.d.	0,05
421ES 25 I	0,61		422ES ts 14 I	0,05	
421ES 25 II	0,75		422ES ts 14 II	n.d.	
421ES 25 III	0,91		422ES ts 14 III	n.d.	
421ES 27 I	0,86	0,89			
421ES 27 II	0,91				
421ES 27 III	0,90				
421ES 29 I	0,48	0,76			
421ES 29 II	0,52				
421ES 29 III	0,52				
421ES 29 IV	1,53				
421ES 33 I	0,74	0,67			
421ES 33 II	0,83				
421ES 33 III	0,45				
421ES 35 I	0,48	0,77			
421ES 35 II	0,79				
421ES 35 III	0,87				
421ES 35 IV	0,95				
421ES 38 I	0,97	1,21			
421ES 38 II	1,35				
421ES 38 III	1,29				
421ES ts 13 I	0,03	0,04			
421ES ts 13 II	0,05				
421ES ts 13 III	n.d.				

Tabla 7: Cantidad de derivados de glucosamina N-acetilada (en μmol por gramo de peso fresco) medida en frutos de plantas transgénicas independientes de las líneas 420 ES, 421 ES y 422 ES. La columna 1 se refiere a la planta, independientemente del origen de la transformación, a partir de la cual el material se cosecho (en el presente documento, “ts” se refiere a las plantas no transformadas). La extensión de los nombres de las plantas por medio de números en latín denota diferentes frutos de la planta en cuestión. Las cantidades no detectables se marcan como “n.d.”.

5 Estos resultados muestran que las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT bacteriana y que codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa, tienen un contenido de derivados de glucosamina N-acetilada que es considerablemente mayor que las plantas de tipo silvestre no transformadas correspondientes. En comparación con las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1 y que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa, las plantas que comprenden una molécula de ácido nucleico exógena que codifican una proteína que tienen la actividad de una GFAT-2 y que codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa tiene un contenido aún más alto de derivados de glucosamina N-acetilada. Esto es cierto tanto para el material de hoja como para frutos de las plantas en cuestión.

h) Análisis de derivados de glucosamina N-acetilada de la línea 422 ES a través de espectroscopia de masa

20 Los extractos de diferentes frutos individuales de la planta con el nombre 422 ES 13 se examinaron a través de espectroscopia de masas de acuerdo con el procedimiento descrito en el punto 9 de Procedimientos Generales para

la presencia de derivados de glucosamina N-acetilada. Se obtuvieron los siguientes resultados:

Detector	Muestra	Masa (m/z) con metabolitos asociados				
		178 GlcN	220 GlcNAc	258 GlcN-P	300 GlcN-P	302.5 UDP-GlcNAc
cps:9.5-10 e4 (d.v.:2000)	422 ES 13 I	0,08	5,08	0,56	1,44	0,14
	422 ES 13 II	0,09	6,41	0,61	1,48	0,14
	422 ES 13 III	0,09	5,95	1,05	1,51	0,18
	ts	0,06	0,05	0,10	0,06	0,00
cps:1.6-1.8 e5 (d.v.:2050)	422 ES 13 I	0,37	10,69	1,91	4,31	0,42
	422 ES 13 II	0,37	13,92	2,10	4,49	0,45
	422 ES 13 III	0,30	12,96	3,43	3,98	0,55
	ts	0,30	0,25	0,49	0,38	0,03
cps:2.0-2.2 e5 (d.v.:2100)	422 ES 13 I	0,71	18,77	3,95	8,70	0,73
	422 ES 13 II	0,68	21,81	3,88	8,32	0,67
	422 ES 13 III	0,48	19,82	6,25	7,14	0,84
	ts	0,55	0,53	1,05	0,94	0,05

Tabla 8: Detección de metabolitos de glucosamina (GlcN), N-acetilglucosamina (GlcNAc), fosfato de glucosamina (GlcN-P), fosfato de N-acetilglucosamina (GlcNAc-P) y UDP-N-acetilglucosamina (UDP-GlcNAc) en frutos de la planta 422 ES 13 a través de espectroscopia de masa. Lo que se muestra es la proporción de la intensidad de señal (área pico) obtenido del metabolito manifestado en el espectro de masas, con base en la intensidad de señal de las hexosas (m/z=179) obtenidas en la misma medición, en porcentaje. Las diferentes mediciones se llevaron a cabo con la configuración del detector manifestadas con respecto a la sensibilidad (“d.v.”) y la intensidad de señal (“cps”) (columna 1). La columna 2 denota la planta, independientemente de su origen de la transformación, a partir de la cual se cosecho el material (en el presente documento, “ts” se refiere a la plantas no transformadas). La extensión de los nombres de las plantas por medio de números en latín denota diferentes frutos de la planta en cuestión.

En paralelo, a través de mediciones EM-EM de las muestras 422 ES 13 I y las frutos de una planta de tipo silvestre (ts) usando las sustancias de referencia (glucosamina, N-acetilglucosamina, glucosamina 6-fosfato, glucosamina 1-fosfato, N-acetilglucosamina 6-fosfato, N-acetilglucosamina 1-fosfato, UDP-N-acetilglucosamina) se analizaron si las intensidades de señal detectadas (áreas pico en cuestión del espectro EM fue realmente debido a la presencia del metabolito correspondiente o de los metabolitos isoméricos correspondientes de la misma masa, o si las intensidades de señal en cuestión en el espectro EM fueron posiblemente causadas por la interferencia a través de señales de otra sustancia. Se hacen las siguientes observaciones:

Glucosamina (GlcN, m/z=178): Las cantidades más altas de GlcN detectada en el espectro EM de las muestras 422 ES 13 I, y ts estuvieron en la escala del límite de detección inferior. En el espectro EM, no se notaron diferencias significativas entre las muestras 422 ES 13 I y ts. Por consiguiente, no fue posible determinar en ningún grado de certeza si las muestras contuvieron GlcN.

N-acetilglucosamina (GlcNAc, m/z=220): Las diferencias más significativas en el espectro EM de las muestras 422 ES 13 y la muestra ts se encontró para este metabolito. En el espectro EM de las muestras 422 ES 13 I, 422 ES 13 II y 422 ES 13 III, se detectaron considerables cantidades de GlcNAc. El espectro EM-EM correspondiente para la muestra 422 ES 13 I corresponde al espectro de la sustancia de referencia (N-acetilglucosamina) y tiene, si existe alguno, solamente muy pequeñas cantidades de sustancias que pueden interferir con la señal relevante en el espectro EM: En contraste, en el espectro EM-EM de la muestra ts mostró que GlcNAc está solamente presente en rastros, si existe alguno. El espectro EM-EM claramente demostró que la intensidad de señal determinada por m/z=220 de la muestra ts, en el espectro EM fue el resultado principal de la interferencia de otras sustancias con la señal.

Fosfatos de glucosamina (GlcN-P, m/z=258): La intensidad de señal del espectro EM para la muestra ts es considerablemente inferior que para las muestras 422 ES 13 I, 422 ES 13 II y 422 ES 13 III. Todas las muestras medidas por EM-EM mostraron que la señal para m/z=258 no solamente es debido a la presencia de GlcN-P sino también a la interferencia de la señal a través de otra sustancia. El espectro EM-EM de la muestra ts mostró que solamente están presentes rastros de GlcN-P, si existe alguno. En contraste, la señal correspondiente para la muestra 422 ES 13 I en el espectro EM-EM mostró la presencia de cantidades significativas de GlcN-P en la señal relevante del espectro EM.

Fosfato de N-acetilglucosamina (GlcNAc-P, m/z=300): para la muestra ts, las intensidades de señal para m/z=300 en el espectro EM son sustancialmente inferiores que para las muestras 422 ES 13 I, 422 ES 13 II y

422 ES 13 III. Los valores determinados por EM-EM para la muestra ts muestran que, si existe alguno, solamente están presentes rastros de GlcNAc-P. En contraste, para la muestra 422 ES 13 I, fue posible demostrar a través de la medición EM-EM que la parte predominante de la intensidad de señal determinada por $m/z=300$ en el espectro EM de esta muestra es debido a GlcNAc-P.

5 UPD-N-Acetilglucosamina (UDP-GlcNAc, $m/z=302.5$): En las intensidades de señal, de tipo silvestre del espectro EM son considerablemente inferiores que en las muestras 422 ES 13 I, 422 ES II y 422 ES III. El espectro EM correspondiente muestra que en todas las muestras una cierta parte de la intensidad de señal del espectro EM no solamente es debido a la presencia de UDP-GlcNAc, sino también debido a la interferencia de señal a través de otra sustancia. Sin embargo, las mediciones EM-EM mostraron que en comparación con las sustancias de interferencia de señal, la proporción de UDP-GlcNAc en el espectro EM de la muestra 422 ES 13 I es sustancialmente más alta que para la muestra ts.

19. Producción de plantas que sintetizan glucosaminoglucanos

Para determinar si las plantas que tienen un contenido aumentado de derivados de glucosamina N-acetilada son adecuadas para la producción de plantas que tienen un contenido aumentado de glucosaminoglucano, primero se generaron las plantas que expresan una glucosaminoglucano sintasa (hialuronano sintasa).

a) Plantas que comprenden una molécula de ácido nucleico que codifican una proteína que tiene la actividad de una hialuronano sintasa

Las plantas de patata (variedad de cultivo Désirée) y las plantas de tomate (variedad de cultivo Moneymaker) se transformaron mediante el procedimiento dado en el punto 1 de Procedimientos Generales (plantas de patata) y en el punto 2 de Procedimientos Generales (plantas de tomate) respectivamente, con el vector de expresión de planta IC 341-222, que comprende una secuencia de ácido nucleico codificante de una proteína que tiene la actividad de una hialuronano sintasa del virus 1 de *Paramecium bursaria Chlorella* bajo el control del promotor del gen patatin B33 de *Solanum Tuberosum* (Rocha-Sosa y col., 1989, EMBO J. 8, 23-29). Las líneas transgénicas obtenidas, las cuales se transformaron con el plásmido IC 341-22, se denominaron 365 ES (plantas de patata) y 367 ES (plantas de tomate), respectivamente.

b) Análisis de las líneas 365 ES

Las plantas individuales de la línea 365 ES se cultivaron en tierra en macetas de 6 cm en un invernadero. En cada caso aproximadamente 0,3 g del material de los tubérculos de patata de las plantas individuales se desarrollaron usando el procedimiento descrito en el punto 5 de procedimientos generales. La cantidad de hialuronano presente en los extractos de plantas respectivos se determinó usando el procedimiento descrito en el punto 7 de procedimientos generales. Aquí, el sobrenadante obtenido después de la centrifugación se diluyó 1:10 para determinar el contenido de hialuronano. Los resultados siguientes se obtuvieron para las plantas seleccionadas:

Tabla 9: Cantidad de hialuronano (en μg por gramo de peso fresco) producida por medio de plantas transgénicas independientemente seleccionadas de la línea 365 ES. La columna 1 se refiere a la planta a partir de la cual se cosecho el material de tubérculo (aquí, "ts" se refiere a plantas no transformadas). La columna 2 manifiesta el valor de la cantidad de hialuronano determinado en hojas de plantas en cuestión. Las cantidades no detectables se marcaron como "n.d."

Planta	Hialuronano [$\mu\text{g/g}$ FW]
365 ES 13	47
365 ES 74	68
ts	n.d.

c) Análisis de las plantas de la línea 367 ES

40 De diferentes de plantas de tomate seleccionadas de la línea 367 ES que había sido cultivadas en tierra en un invernadero, en cada caso se cosechó una hoja y se congeló en nitrógeno líquido. El desarrollo adicional y la determinación del contenido de hialuronano se llevó a cabo como se describe en el Ejemplo 19b) para tubérculos de plantas de patata. Se obtuvieron los siguientes resultados:

Tabla 10: Cantidad de hialuronano (en µg por gramo de peso fresco) producida por medio de plantas transgénicas independientemente seleccionadas de la línea 367 ES. La columna 1 se refiere a la planta a partir de la cual se cosecho el material de tubérculo (aquí, "ts" se refiere a plantas no transformadas). La columna 2 manifiesta el valor de la cantidad de hialuronano determinado en hojas de plantas en cuestión.

Planta	Hialuronano [µg/g FW]
367 ES 25	57,19
365 ES 42	88,99
ts	0,06

5

20. Plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa

10 Algunas glucosaminoglucano sintasas (tales como, por ejemplo, hialuronano sintasa) requieren, como sustrato, derivados de glucosamina N-acetilada y UDP-GlcA. Por consiguiente, primero se generaron plantas que tienen una actividad aumentada de una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y una actividad aumentada de una proteína que tiene la actividad de una hialuronano sintasa.

a) Producción de plantas de patata

15 Las plantas de patata de la línea 365 ES 74 (véase Ejemplo 19b)) se transformaron usando el procedimiento dado en el punto 1 de Procedimientos Generales con respecto al vector de expresión de planta IC 349-222 que comprende la secuencia de ácido nucleico que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa bajo con el control del promotor 35S. Las líneas transgénicas obtenidas, las cuales se transformaron con el plásmido IC 349-222, se denominaron 423 ES.

b) Análisis de las plantas de la línea 423 ES

20 Las plantas de la línea 423 ES se cultivaron en tierra en macetas de 6 cm en un invernadero. En cada caso aproximadamente de 0,3 g a 0,8 g del material de hoja, se cosechó se plantas individuales, se desarrolló usando el procedimiento descrito en el punto 5 de Procedimientos Generales, y el contenido de hialuronano se determinó usando el procedimiento descrito en el punto 7 de Procedimientos Generales. Para las plantas individuales que tienen un contenido aumentado de derivados de N-acetilglucosamina, se obtuvieron los siguientes resultados:

Planta	Hialuronano [µg/g FW]	Planta	Hialuronano [µg/g FW]
423ES 1	328,73	423ES 49	285,19
423ES 3	210,38	423ES 51	213,97
423ES 5	340,99	423ES 53	328,76
423ES 6	250,88	423ES 54	358,23
423ES 7	214,53	423ES 55	154,06
423ES 8	309,22	423ES 59	276,32
423ES 9	253,31	423ES 60	498,70
423ES 10	229,61	423ES 61	300,97
423ES 11	234,40	423ES 62	292,08
423ES 12	480,22	423ES 65	230,38
423ES 13	253,63	423ES 67	267,54
423ES 14	221,77	423ES 68	370,08
423ES 15	202,46	ts 1	0,38
423ES 17	281,46	ts 2	0,12
423ES 18	310,41	ts 3	0,07

25

(continuación)

Planta	Hialuronano [µg/g FW]
423ES 19	268,91
423ES 20	394,04
423ES 21	462,64
423ES 24	438,33
423ES 25	419,50
423ES 26	342,89
423ES 27	383,32
423ES 28	236,83

Planta	Hialuronano [µg/g FW]
ts 4	n.d.
ts 5	0,47
ts 6	n.d.
ts 7	0,05
ts 8	0,05
ts 9	0,10
ts 10	n.d.
365ES 74-1	348,43

Planta	Hialuronano [µg/g FW]
423ES 29	332,63
423ES 32	254,88
423ES 33	283,31
423ES 35	276,60
423ES 36	308,85
423ES 38	307,72
423ES 41	259,89
423ES 43	244,62
423ES 47	229,25
423ES 48	238,22

Planta	Hialuronano [µg/g FW]
365ES 74-2	214,59
365ES 74-3	391,88
365ES 74-4	442,60
365ES 74-5	293,01
365ES 74-6	323,47
365ES 74-7	464,21
365ES 74-8	341,32
365ES 74-9	338,93
365ES 74-10	438,55

Tabla 11: Cantidad de hialuronano (en µg por gramo de peso fresco) medido en hojas de plantas transgénicas independientes de la línea 423 ES. La columna 1 se refiere en cada caso a las plantas, independientemente originadas de la transformación, a partir de las cuales se cosechó el material (en el presente documento, “ts 1” a “ts 10” se refiere a plantas no transformadas independientes). Para la comparación, los valores de 10 diferentes provienen de plantas de la línea 365 ES utilizadas como línea de partida para la transformación (365 ES-1 a 365 ES-10) se muestran. Las cantidades no detectables se marcan como “n.d.”.

5 Se pueden ver a partir de los resultados que las plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una UDP-glucosamina deshidrogenasa y una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una hialuronano sintasa no sintetizan ninguna cantidad significativa estadísticamente aumentada de hialuronano en comparación con plantas que solamente tienen una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una hialuronano sintasa.

10 21. Plantas que sintetizan cantidades aumentadas de glucosaminoglucano

a) Producción de plantas de tomate que sintetizan cantidades aumentadas de glucosaminoglucano

20 Las plantas de tomate de las líneas 367 ES 25 (véase ejemplo 19c)), tienen una molécula de ácido nucleico que codifica para una hialuronano sintasa se transformaron de nuevo usando el procedimiento dado en el punto 2 de Procedimientos Generales con los vectores de expresión de planta IC 372-256 o IC 375-271 que comprenden moléculas de ácido nucleico que codifican diferentes isoformas de las proteínas que tienen la actividad de una GFAT.

Las plantas de tomate transgénicas obtenidas después de la transformación de la línea 367 ES 25, con el plásmido IC 372-256 (GFAT-2), se denominaron 399 ES.

Las plantas de tomate transgénicas obtenidas después de la transformación de la línea 367 ES 25 con el plásmido IC 375-271 (GFAT bacteriano), se denominaron 405 ES.

b) Análisis de las líneas 399 ES y 405 ES

5 Los frutos maduros se cosecharon de diferentes plantas de tomate de las líneas 399 ES y 405 ES cultivadas en tierra en un invernadero, y se determinó el contenido de hialuronano como se describe en el punto 7 de Procedimientos Generales. Se obtuvieron los siguientes resultados:

10 Tabla 12: Cantidad de hialuronano (“HA” en µg por gramo de peso fresco) medido en frutos de plantas transgénicas independientes de las líneas 399 ES y 405 ES. La columna 1 se refiere a las plantas, independientemente de su origen de la transformación, a partir de las cuales se cosechó el material (en el presente documento, “ts” se refiere a plantas no transformadas). Para la comparación, los valores de diferentes progenies de las plantas de la línea 367 ES utilizadas como la línea de partida para la transformación se muestran. Las extensiones de los nombres de las plantas por números en latín denotan diferentes frutos de la planta en cuestión. Las cantidades no detectables se marcan como “n.d.”.

Muestra	HA [µg/g FW]	Media [µg/g FW]	Muestra	HA [µg/g FW]	Media [µg/g FW]
399ES 1 I	63,38	87,02	405ES 5 I	207,20	254,94
399ES 1 II	96,45		405ES 5 II	302,67	
399ES 1 III	101,23		405ES 10 I	1232,38	1074,94
399ES 11 I	388,83	292,79	405ES 10 II	917,50	
399ES 11 II	244,01		ts I	0,86	0,46
399ES 11 III	254,91		ts II	0,06	
399ES 11 IV	285,72		367ES 25-8 I	136,67	155,70
399ES 11 V	297,99		367ES 25-8 II	174,72	
399ES 11 VI	285,29		367ES 25-9 I	37,76	
ts I	0,02	0,01			
ts II	0,02				
ts III	0,01				
ts IV	n.d.				
367ES 25-1 I	9,77	12,04			
367ES 25-1 II	8,21				
367ES 25-1 III	18,04				
367ES 25-1 IV	13,66				
367ES 25-1 V	10,33				
367ES 25-2 I	9,31	11,96			
367ES 25-2 II	10,55				
367ES 25-2 III	11,53				
367ES 25-2 IV	16,54				
367ES 25-2 V	11,86				
367ES 25-3 I	6,99	8,51			
367ES 25-3 II	7,94				
367ES 25-3 III	9,23				
367ES 25-3 IV	7,09				
367ES 25-3 V	11,28				

Estos resultados muestran las plantas que comprenden moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa y codifican una proteína que tienen la actividad de una UDP-glucosa deshidrogenasa y codifican una proteína que tienen la actividad de una GFAT-2 o una GFAT bacteriana que sintetiza considerablemente cantidades mayores de glucosaminoglucanos que las plantas que tienen solamente una molécula de ácido nucleico exógena que codifica para una glucosaminoglucano sintasa.

5

c) Producción de plantas de patata que sintetizan cantidades aumentadas de glucosaminoglucano

Las plantas de patata de las líneas 365 ES 74 (véase Ejemplo 19 b)) comprenden una molécula de ácido nucleico que codifica para una hialuronano sintasa se transformaron de nuevo usando el procedimiento manifestado en el punto de Procedimientos Generales con los vectores de expresión de planta IC 376-271, IC 372-256 o IC 375-271 que comprenden moléculas de ácido nucleico que codifican para diferentes isoformas de proteínas que tienen la actividad de una GFAT.

10

Las plantas de patata transgénicas obtenidas después de la transformación de la línea 365 ES 74 con el plásmido IC 376-271 (GFAT-1), se denominaron 409 ES.

15

Las plantas de patata transgénicas obtenidas después de la transformación de la línea 365 ES 74 con el plásmido IC 372-256 (GFAT-2), se denominaron 396 ES.

Las plantas de patata transgénicas obtenidas después de la transformación de la línea 365 ES 74 con el plásmido IC 375-271 (GFAT bacteriano), se denominaron 404 ES.

d) Análisis de las líneas 396 ES, 404 ES y 409 ES

20

Se cosechó material de hoja y/o tubérculo de diferentes plantas de patata de las líneas 396 ES (GFAT-2), 404 ES (GFAT bacteriano) y 409 ES (GFAT-1) cultivadas en tierra en un invernadero, y se determinó el contenido de hialuronano como se describe en el punto 7 de Procedimientos Generales. Se obtuvieron los siguientes resultados para las plantas de la línea 409 ES:

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
409 ES 2	54,01	
409 ES 3	68,75	212,24
409 ES 4	59,80	111,54
409 ES 5	26,90	
409 ES 6	38,01	182,39
409 ES 7	25,80	95,68
409 ES 8	51,92	99,35
409 ES 9	48,43	168,61
409 ES 10	52,52	

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
409 ES 26	38,81	
409 ES 27	24,71	126,74
409 ES 28	66,95	
409 ES 29	79,58	164,66
ts-1	n.d.	
ts-2	n.d.	
ts-3	n.d.	
ts-4	n.d.	
365 ES 74-1	25,19	

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
409 ES 13	55,87	
409 ES 14	45,91	143,96
409 ES 15	52,76	
409 ES 16	60,28	
409 ES 22	69,47	114,97
409 ES 23	108,67	

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
365 ES 74-2	31,15	
365 ES 74-3	72,96	
365 ES 74-4	35,98	
365 ES 74-5	40,18	123,66
365 ES 74-6	37,70	

Tabla 13: Cantidad de hialuronano (“HA” en µg por gramo de peso fresco) medido en hojas y tubérculos de plantas transgénicas independientes de la línea 409 ES. La columna 1 se refiere a las plantas, independientemente de su origen de la transformación, a partir de las cuales se cosechó el material (en el

25

presente documento, “ts” se refiere a plantas no transformadas). Los valores para las diferentes progenies de las plantas de la línea 365 ES 74, utilizadas como la línea de partida para la transformación, se muestran para comparación Las cantidades no detectables se marcan como “n.d.”.

Se obtuvieron los siguientes resultados para las plantas de la línea 396 ES:

- 5 Tabla 14: Cantidad de hialuronano (“HA” en µg por gramo de peso fresco) medido en hojas y tubérculos de plantas transgénicas independientes de la línea 396 ES. La columna 1 se refiere a las plantas, independientemente de su origen de la transformación, a partir de las cuales se cosechó el material (en el presente documento, “ts” se refiere a plantas no transformadas). Los valores para las diferentes progenies de las plantas de la línea 365 ES 74, utilizadas como la línea de partida para la transformación, se muestran para comparación.

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]	Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
396 ES 2		470,93	396 ES 51	1160,57	
396 ES 9		735,40	396 ES 57	428,33	
396 ES 11	938,33		396 ES 57	807,97	
396 ES 15		393,64	365 ES 74-1	265,10	
396 ES 16	416,43		366 ES 74-2		91,84
396 ES 17	426,79		365 ES 74-3	193,40	
396 ES 23	271,85		367 ES 74-4		175,48
396 ES 24	443,57		365 ES 74-5	73,90	
396 ES 25	801,58		368 ES 74-6		168,68
396 ES 28		484,76	365 ES 74-7	67,58	
396 ES 30		224,06	369 ES 74-8		121,89
396 ES 32	941,89		365 ES 74-9	62,23	
396 ES 33		1295,98	370 ES 74-10		275,24
396 ES 34	796,79		365 ES 74-11	134,56	
396 ES 36	204,49		ts-1	0,07	2,27
396 ES 36	860,54		ts-2	0,11	
396 ES 42		1445,51	ts-3	0,12	1,07
396 ES 44		1312,56	ts-4	0,04	0,78
396 ES 48	461,05		ts-5	0,10	
396 ES 49	538,75		ts-5	0,24	
396 ES 50	619,23				

- 10 Se obtuvieron los siguientes resultados para las plantas de 404 ES:

Tabla 15: Cantidad de hialuronano (en µg por gramo de peso fresco) medido en hojas de plantas transgénicas independientes de la línea 404 ES. La columna 1 se refiere a las plantas, independientemente de su origen de la transformación, a partir de las cuales se cosechó el material (en el presente documento, “ts” se refiere a las plantas no transformadas). Los valores para las diferentes progenies de las plantas de la línea 365 ES 74, utilizadas como la línea de partida para la transformación, se muestran para comparación.

15

Planta	Hialuronano en hojas [µg/g FW]	Planta	Hialuronano en hojas [µg/g FW]
404 ES 1	801,14	404 ES 37	533,89
404 ES 6	365,15	404 ES 38	651,12
404 ES 7	218,42	404 ES 39	353,74

(continuación)

Planta	Hialuronano en hojas [µg/g FW]	Planta	Hialuronano en hojas [µg/g FW]
404 ES 8	521,92	404 ES 40	371,88
404 ES 9	366,46	404 ES 42	849,43
404 ES 10	226,83	404 ES 43	479,34
404 ES 11	231,39	404 ES 44	921,11
404 ES 13	1547,12	404 ES 46	846,81
404 ES 14	616,79	404 ES 48	302,54
404 ES 15	832,32	ts-1	0,20
404 ES 20	581,11	ts-2	0,30
404 ES 21	489,73	ts-3	0,19
404 ES 23	817,91	ts-4	0,39
404 ES 24	434,06	ts-5	0,20
404 ES 26	205,00	365 ES 74-1	72,44
404 ES 28	369,96	365 ES 74-2	135,60
404 ES 29	1146,68	365 ES 74-3	19,56
404 ES 34	310,76	365 ES 74-4	114,83
404 ES 35	1388,51	365 ES 74-5	73,77
404 ES	1095,11		

Estos resultados muestran las plantas que comprenden moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa y codifican una proteína que tienen la actividad de una UDP-glucosa deshidrogenasa y codifican una proteína que tienen la actividad de una GFAT-2 o una GFAT bacteriana que sintetiza considerablemente cantidades mayores de glucosaminoglucanos que las plantas que tienen solamente una molécula de ácido nucleico exógena que codifica para una glucosaminoglucano sintasa y codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y codifican una proteína que tiene la actividad de una GFAT-1.

5 e) Producción de plantas que comprenden moléculas de ácido nucleico exógenas que codifican para una hialuronano sintasa y una proteína que tiene la actividad de una GFAT bacteriana

Las plantas de patata de la línea 365 ES 74 (véase Ejemplo 19 b)), comprenden una molécula de ácido nucleico que codifica para una hialuronano sintasa se transformaron de nuevo usando el procedimiento dado en el punto 1 de Procedimientos Generales con el vector de expresión de planta IC 398-311 que comprende moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una GFAT bacteriana. Las líneas que se originan de esta transformación se denominaron 433 ES.

f) Análisis de la línea 433 ES

Se cosechó el material de hoja y/o tubérculo de diferentes plantas de patata de la línea 433 ES cultivadas en tierra en un invernadero, y el contenido de hialuronano se determinó como se describe en el punto 7 de Procedimientos Generales. Se obtuvieron los siguientes resultados para las plantas de la línea 433 ES:

Tabla 16: Cantidad de hialuronano ("HA" en µg por gramo de peso fresco) medido en hojas y tubérculos de plantas transgénicas independientes de la línea 433 ES. La columna 1 se refiere a las plantas, independientemente de su origen de la transformación, a partir de las cuales se cosechó el material (en el presente documento, "ts" se refiere a las plantas no transformadas). Los valores para las diferentes progenies de las plantas de la línea 365 ES 74, utilizadas como la línea de partida para la transformación, se muestran para comparación. Los valores para la línea 365 ES 74 corresponden a aquellas de la Tabla 14, ya que todas las plantas se cultivaron simultáneamente en un invernadero.

Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]	Planta	HA en hojas [µg/g FW]	HA en tubérculos [µg/g FW]
433ES 1	111,84	126,70	433ES 28	1850,99	294,98
433ES 3	303,34	203,16	433ES 30	2512,40	
433ES 4	3142,41		433ES 31	3337,54	
433ES 5	312,98	825,96	433ES 32	1583,60	
433ES 7	1492,94		433ES 34	3552,44	
433ES 8	914,03		433ES 35	5419,43	
433ES 9	1858,68		433ES 36	902,01	
433ES 10	357,90		433ES 37	829,35	
433ES 11	5962,82		433ES 38	1536,55	
433ES 12		662,99	ts-1	0,40	n.d.
433ES 13	626,52	624,33	ts-2	0,34	n.d.
433ES 14	665,23		ts-3	n.d.	
433ES 15	601,36		365 ES 74-1	265,1	
433ES 16	3416,94		366 ES 74-2		91,84
433ES 18	781,02		365 ES 74-3	193,5	
433ES 19	3294,09		367 ES 74-4		175,48
433ES 20	1348,85	975,18	365 ES 74-5	73,9	
433ES 21	937,92		368 ES 74-6		166,68
433ES 22	1086,45		365 ES 74-7	67,58	
433ES 23	1327,28		369 ES 74-8		121,89
433ES 24	340,80	76,00	365 ES 74-9	62,23	
433ES 25	1529,95		370 ES 74-10		275,24
433ES 26	375,53		365 ES 74-11	134,56	
433ES 27		425,65			

5 Estos resultados muestran que las plantas comprenden moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa y codifican una proteína que tiene la actividad de una GFAT bacteriana que sintetiza cantidad considerablemente mayores de glucosaminoglucano que las plantas que tienen solamente moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa.

22. Resumen de los resultados

Los resultados en el Ejemplo 16 muestran que las plantas que comprenden una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT bacteriana tienen contenidos considerablemente aumentados de derivados de glucosamina N-acetilada comparados con plantas de tipo silvestre no transformadas.

10 Los resultados en el Ejemplo 17 muestran que las plantas que comprenden una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una GFAT-2 tienen contenidos considerablemente superiores de derivados de glucosamina N-acetilada que las plantas de tipo silvestre no transformadas.

15 Todas las plantas transformadas descritas en el Ejemplo 18 tienen, además de las moléculas de ácido nucleico que codifican para diferentes isoformas de una proteína que tiene la actividad de una GFAT, en cada caso la misma molécula de ácido nucleico que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa. Por consiguiente, la diferencia esencial de las plantas transformadas descritas en el Ejemplo 18 son diferentes moléculas de ácido nucleico exógenas que codifican para diferentes isoformas de una proteína que tienen la actividad de una GFAT. El Ejemplo 18 b) muestra que el contenido de los derivados de glucosamina N-acetilada en las plantas que tienen una molécula de ácido nucleico exógena para una proteína que tiene la actividad de una GFAT-2 se incrementa solamente de forma ligera comparada con las plantas no transformadas.

20 Además, se puede ver en el Ejemplo 18 d) que las plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2, tiene un contenido considerablemente superior de derivados de glucosamina N-acetilada que las plantas de tipo silvestre no transformadas. El contenido de derivados de glucosamina N-acetilada en las plantas que tienen una molécula de ácido nucleico exógena

codifican una proteína que tiene la actividad de una GFAT-2 es también considerablemente mayor que las plantas que tienen una molécula de ácido nucleico exógena que codifica con una proteína que tiene la actividad de una GFAT-1.

5 Además, como se puede ver en los Ejemplos 18 f) y 18 g) las plantas que comprenden una molécula de ácido nucleico exógena que codifican una proteína que tiene la actividad de una GFAT bacteriana tienen contenidos aún mayores de derivados de glucosamina N-acetilada que las plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2.

10 Los resultados en el Ejemplo 21 f) muestran que las plantas que comprenden moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa que codifican una proteína que tiene la actividad de una GFAT bacteriana sintetizan cantidades considerablemente mayores de glucosaminoglucano que las plantas que tienen solamente moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa.

15 De esta forma, se puede concluir que la cantidad de glucosaminoglucanos sintetizados en las plantas se puede incrementar considerablemente a través de la generación de plantas que, además de las moléculas de ácido nucleico exógenas codifican para una glucosaminoglucano sintasa, adicionalmente comprenden moléculas de ácido nucleico exógenas que codifican una proteína que tiene la actividad de una GFAT bacteriana.

20 Todas las plantas transformadas cuyos resultados se muestran en el Ejemplo 21 b) y 21 d) tienen, además de moléculas de ácido nucleico que codifican para diferentes isoformas de una proteína que tiene la actividad de una GFAT, también moléculas de ácido nucleico exógenas que codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y moléculas de ácido nucleico exógenas que codifican para una glucosaminoglucano sintasa. La diferencia esencial entre las plantas transformadas cuyos resultados se muestran en los Ejemplos 21 b) y 21 d) por consiguiente consiste en las diferentes moléculas de ácido nucleico que codifican para las diferentes isoformas de una proteína que tiene la actividad de una GFAT.

25 Los resultados mostrados en el Ejemplo 21 b) muestran que el contenido de glucosaminoglucano en las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o codifican una proteína que tiene la actividad de una GFAT bacteriana se incrementa considerablemente en comparación con las plantas que solamente tienen la actividad de una glucosaminoglucano sintasa.

30 Los resultados mostrados en el Ejemplo 21 d) muestran que las plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1 contiene una cantidad ligeramente mayor de glucosaminoglucanos que las plantas que tienen solamente actividad de una glucosaminoglucano sintasa. En contraste, el contenido glucosaminoglucano en plantas que comprenden una molécula de ácido nucleico exógena que codifican una proteína que tiene la actividad de una GFAT-2 es considerablemente mayor que las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-1. Además, se puede ver que el Ejemplo 21 d) que las plantas individuales que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT bacteriana contienen cantidades aún mayores de glucosaminoglucanos que las plantas que tienen una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2.

40 Los resultados en el Ejemplo 20 b) muestran que las plantas que comprenden una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y una molécula de ácido nucleico que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa no tienen ninguna cantidad significativamente incrementa en forma estadística de glucosaminoglucano en comparación con las plantas que comprenden solamente una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glucosaminoglucano sintasa.

45 Para concluir, los resultados mostrados indican que los incrementos considerables en las cantidades de glucosaminoglucanos en las plantas que comprenden moléculas de ácido nucleico exógenas codifican una proteína que tiene la actividad de una UDP-glucosa deshidrogenasa y codifican una proteína que tiene la actividad de una glucosaminoglucano sintasa y tiene la actividad de una GFAT-2 o tiene la actividad de una GFAT bacteriana no es debido a la presencia de las moléculas de ácido nucleico exógenas que tienen la actividad de una UDP-glucosa deshidrogenasa, sino a la presencia de moléculas de ácido nucleico que tienen la actividad de una GFAT-2 o que tienen la actividad de una GFAT bacteriana.

50 Ya que las hialuronano sintasas utilizadas en forma ilustrativa como proteínas que tienen la actividad de una glucosaminoglucano sintasa requieren, como sustratos, tanto UDP-GlcNAc como UDP-GlcA, también se puede concluir de los resultados mostrados en las cantidades aumentadas de hialuronano (glucosaminoglucanos) son debido a cantidades aumentadas de derivados de glucosamina N-acetilada y no a cantidades aumentadas de UDP-GlcA en esas plantas.

55

Listado de secuencias

<110> Bayer CropScience GmbH

5 <120> Plantas con un contenido de amino azucares aumentado

<130> BCS 06-5010 PCT

10 <150> EP05090279.0
<151> 05-10-2005

<150> US60/725.388
<151> 11-10-2005

15 <150> EP06090177.4
<151> 22-09-2006

<160> 17

20 <170> PatentIn versión 3.3

<210> 1
<211> 1707
<212> ADN

25 <213> Virus 1 *Paramecium bursaria Chlorella*

<220>
<221> CDS
<222> (1)..(1707)

30 <300>
<308> PB42580
<309> 24-121995
<313> (50903)..(52609)

35 <400> 1

atg ggt aaa aat ata atc ata atg gtt tcg tgg tac acc atc ata act 48
Met Gly Lys Asn Ile Ile Ile Met Val Ser Trp Tyr Thr Ile Ile Thr
1 5 10 15

tca aat cta atc gcg gtt gga gga gcc tct cta atc ttg gct ccg gca 96
Ser Asn Leu Ile Ala Val Gly Gly Ala Ser Leu Ile Leu Ala Pro Ala
20 25 30

att act ggg tat gtt cta cat tgg aat att gct ctc tcg aca atc tgg 144
Ile Thr Gly Tyr Val Leu His Trp Asn Ile Ala Leu Ser Thr Ile Trp
35 40 45

ES 2 562 211 T3

gga gta tca gct tat ggt att ttc gtt ttt ggg ttt ttc ctt gca caa	192
Gly Val Ser Ala Tyr Gly Ile Phe Val Phe Gly Phe Phe Leu Ala Gln	
50 55 60	
ggt tta ttt tca gaa ctg aac agg aaa cgt ctt cgc aag tgg att tct	240
Val Leu Phe Ser Glu Leu Asn Arg Lys Arg Leu Arg Lys Trp Ile Ser	
65 70 75 80	
ctc aga cct aag ggt tgg aat gat gtt cgt ttg gct gtg atc att gct	288
Leu Arg Pro Lys Gly Trp Asn Asp Val Arg Leu Ala Val Ile Ile Ala	
85 90 95	
gga tat cgc gag gat cct tat atg ttc cag aag tgc ctc gag tct gta	336
Gly Tyr Arg Glu Asp Pro Tyr Met Phe Gln Lys Cys Leu Glu Ser Val	
100 105 110	
cgt gac tct gat tat ggc aac gtt gcc cgt ctg att tgt gtg att gac	384
Arg Asp Ser Asp Tyr Gly Asn Val Ala Arg Leu Ile Cys Val Ile Asp	
115 120 125	
ggt gat gag gac gat gat atg agg atg gct gcc gtt tac aag gcg atc	432
Gly Asp Glu Asp Asp Asp Met Arg Met Ala Ala Val Tyr Lys Ala Ile	
130 135 140	
tac aat gat aat atc aag aag ccc gag ttt gtt ctg tgt gag tca gac	480
Tyr Asn Asp Asn Ile Lys Lys Pro Glu Phe Val Leu Cys Glu Ser Asp	
145 150 155 160	
gac aag gaa ggt gaa cgc atc gac tct gat ttc tct cgc gac att tgt	528
Asp Lys Glu Gly Glu Arg Ile Asp Ser Asp Phe Ser Arg Asp Ile Cys	
165 170 175	
gtc ctc cag cct cat cgt gga aaa cgg gag tgt ctt tat act ggg ttt	576
Val Leu Gln Pro His Arg Gly Lys Arg Glu Cys Leu Tyr Thr Gly Phe	
180 185 190	
caa ctt gca aag atg gac ccc agt gtc aat gct gtc gtt ctg att gac	624
Gln Leu Ala Lys Met Asp Pro Ser Val Asn Ala Val Val Leu Ile Asp	
195 200 205	
agc gat acc gtt ctc gag aag gat gct att ctg gaa gtt gta tac cca	672
Ser Asp Thr Val Leu Glu Lys Asp Ala Ile Leu Glu Val Val Tyr Pro	
210 215 220	
ctt gca tgc gat ccc gag atc caa gcc gtt gca ggt gag tgt aag att	720
Leu Ala Cys Asp Pro Glu Ile Gln Ala Val Ala Gly Glu Cys Lys Ile	
225 230 235 240	
tgg aac aca gac act ctt ttg agt ctt ctc gtc gct tgg cgg tac tat	768

ES 2 562 211 T3

Trp Asn Thr Asp Thr Leu Leu Ser Leu Leu Val Ala Trp Arg Tyr Tyr	245	250	255	
tct gcg ttt tgt gtg gag agg agt gcc cag tct ttt ttc agg act gtt				816
Ser Ala Phe Cys Val Glu Arg Ser Ala Gln Ser Phe Phe Arg Thr Val	260	265	270	
cag tgc gtt ggg ggg cca ctg ggt gcc tac aag att gat atc att aag				864
Gln Cys Val Gly Gly Pro Leu Gly Ala Tyr Lys Ile Asp Ile Ile Lys	275	280	285	
gag att aag gac ccc tgg att tcc cag cgc ttt ctt ggt cag aag tgt				912
Glu Ile Lys Asp Pro Trp Ile Ser Gln Arg Phe Leu Gly Gln Lys Cys	290	295	300	
act tac ggt gac gac cgc cgg cta acc aac gag atc ttg atg cgt ggt				960
Thr Tyr Gly Asp Asp Arg Arg Leu Thr Asn Glu Ile Leu Met Arg Gly	305	310	315	320
aaa aag gtt gtg ttc act cca ttt gct gtt ggt tgg tct gac agt ccg				1008
Lys Lys Val Val Phe Thr Pro Phe Ala Val Gly Trp Ser Asp Ser Pro	325	330	335	
acc aat gtg ttt cgg tac atc gtt cag cag acc cgc tgg agt aag tcg				1056
Thr Asn Val Phe Arg Tyr Ile Val Gln Gln Thr Arg Trp Ser Lys Ser	340	345	350	
tgg tgc cgc gaa att tgg tac acc ctc ttc gcc gcg tgg aag cac ggt				1104
Trp Cys Arg Glu Ile Trp Tyr Thr Leu Phe Ala Ala Trp Lys His Gly	355	360	365	
ttg tct gga att tgg ctg gcc ttt gaa tgt ttg tat caa att aca tac				1152
Leu Ser Gly Ile Trp Leu Ala Phe Glu Cys Leu Tyr Gln Ile Thr Tyr	370	375	380	
ttc ttc ctc gtg att tac ctc ttt tct cgc cta gcc gtt gag gcc gac				1200
Phe Phe Leu Val Ile Tyr Leu Phe Ser Arg Leu Ala Val Glu Ala Asp	385	390	395	400
cct cgc gcc cag aca gcc acg gtg att gtg agc acc acg gtt gca ttg				1248
Pro Arg Ala Gln Thr Ala Thr Val Ile Val Ser Thr Thr Val Ala Leu	405	410	415	
att aag tgt ggg tat ttt tca ttc cga gcc aag gat att cgg gcg ttt				1296
Ile Lys Cys Gly Tyr Phe Ser Phe Arg Ala Lys Asp Ile Arg Ala Phe	420	425	430	
tac ttt gtg ctt tat aca ttt gtt tac ttt ttc tgt atg att ccg gcc				1344
Tyr Phe Val Leu Tyr Thr Phe Val Tyr Phe Phe Cys Met Ile Pro Ala				

ES 2 562 211 T3

Ile Thr Gly Tyr Val Leu His Trp Asn Ile Ala Leu Ser Thr Ile Trp
 35 40 45

Gly Val Ser Ala Tyr Gly Ile Phe Val Phe Gly Phe Phe Leu Ala Gln
 50 55 60

Val Leu Phe Ser Glu Leu Asn Arg Lys Arg Leu Arg Lys Trp Ile Ser
 65 70 75 80

Leu Arg Pro Lys Gly Trp Asn Asp Val Arg Leu Ala Val Ile Ile Ala
 85 90 95

Gly Tyr Arg Glu Asp Pro Tyr Met Phe Gln Lys Cys Leu Glu Ser Val
 100 105 110

Arg Asp Ser Asp Tyr Gly Asn Val Ala Arg Leu Ile Cys Val Ile Asp
 115 120 125

Gly Asp Glu Asp Asp Asp Met Arg Met Ala Ala Val Tyr Lys Ala Ile
 130 135 140

Tyr Asn Asp Asn Ile Lys Lys Pro Glu Phe Val Leu Cys Glu Ser Asp
 145 150 155 160

Asp Lys Glu Gly Glu Arg Ile Asp Ser Asp Phe Ser Arg Asp Ile Cys
 165 170 175

Val Leu Gln Pro His Arg Gly Lys Arg Glu Cys Leu Tyr Thr Gly Phe
 180 185 190

Gln Leu Ala Lys Met Asp Pro Ser Val Asn Ala Val Val Leu Ile Asp
 195 200 205

Ser Asp Thr Val Leu Glu Lys Asp Ala Ile Leu Glu Val Val Tyr Pro
 210 215 220

ES 2 562 211 T3

Leu Ala Cys Asp Pro Glu Ile Gln Ala Val Ala Gly Glu Cys Lys Ile
 225 230 235 240

Trp Asn Thr Asp Thr Leu Leu Ser Leu Leu Val Ala Trp Arg Tyr Tyr
 245 250 255

Ser Ala Phe Cys Val Glu Arg Ser Ala Gln Ser Phe Phe Arg Thr Val
 260 265 270

Gln Cys Val Gly Gly Pro Leu Gly Ala Tyr Lys Ile Asp Ile Ile Lys
 275 280 285

Glu Ile Lys Asp Pro Trp Ile Ser Gln Arg Phe Leu Gly Gln Lys Cys
 290 295 300

Thr Tyr Gly Asp Asp Arg Arg Leu Thr Asn Glu Ile Leu Met Arg Gly
 305 310 315 320

Lys Lys Val Val Phe Thr Pro Phe Ala Val Gly Trp Ser Asp Ser Pro
 325 330 335

Thr Asn Val Phe Arg Tyr Ile Val Gln Gln Thr Arg Trp Ser Lys Ser
 340 345 350

Trp Cys Arg Glu Ile Trp Tyr Thr Leu Phe Ala Ala Trp Lys His Gly
 355 360 365

Leu Ser Gly Ile Trp Leu Ala Phe Glu Cys Leu Tyr Gln Ile Thr Tyr
 370 375 380

Phe Phe Leu Val Ile Tyr Leu Phe Ser Arg Leu Ala Val Glu Ala Asp
 385 390 395 400

Pro Arg Ala Gln Thr Ala Thr Val Ile Val Ser Thr Thr Val Ala Leu
 405 410 415

Ile Lys Cys Gly Tyr Phe Ser Phe Arg Ala Lys Asp Ile Arg Ala Phe

ES 2 562 211 T3

	420		425		430														
Tyr	Phe	Val	Leu	Tyr	Thr	Phe	Val	Tyr	Phe	Phe	Cys	Met	Ile	Pro	Ala				
		435					440					445							
Arg	Ile	Thr	Ala	Met	Met	Thr	Leu	Trp	Asp	Ile	Gly	Trp	Gly	Thr	Arg				
	450						455				460								
Gly	Gly	Asn	Glu	Lys	Pro	Ser	Val	Gly	Thr	Arg	Val	Ala	Leu	Trp	Ala				
465					470					475									480
Lys	Gln	Tyr	Leu	Ile	Ala	Tyr	Met	Trp	Trp	Ala	Ala	Val	Val	Gly	Ala				
				485					490					495					
Gly	Val	Tyr	Ser	Ile	Val	His	Asn	Trp	Met	Phe	Asp	Trp	Asn	Ser	Leu				
			500					505					510						
Ser	Tyr	Arg	Phe	Ala	Leu	Val	Gly	Ile	Cys	Ser	Tyr	Ile	Val	Phe	Ile				
		515					520					525							
Val	Ile	Val	Leu	Val	Val	Tyr	Phe	Thr	Gly	Lys	Ile	Thr	Thr	Trp	Asn				
	530					535					540								
Phe	Thr	Lys	Leu	Gln	Lys	Glu	Leu	Ile	Glu	Asp	Arg	Val	Leu	Tyr	Asp				
545					550					555					560				
Ala	Thr	Thr	Asn	Ala	Gln	Ser	Val												
							565												

<210> 3
 <211> 1707
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Secuencia sintética que codifica la proteína hialuronansintasa del virus *Paramecium bursaria Chlorella*

10

<400> 3

ES 2 562 211 T3

atgggtaaga acattatcat tatgggtgcc tgggtacacaa ttattacaag taatctcatc 60
 gcagttggty gtgcatctct tattctcgct ccagctatca ctggatatgt tcttcaactgg 120
 aacatcgccc tctcaactat ttggggagtt tccgcatatg gtatTTTTgt tttcgggttc 180
 tttttggctc aggttctggt ctcaagagctc aatcgtaaga gactcaggaa gtggattagc 240
 cttagaccaa aggggtggaa tgacgttcgt ctgctgtca ttatcgctgg ctaccgtgaa 300
 gatccttaca tgtttcaaaa gtgcttggaa tcagttaggg atagtgatta tggcaacgtc 360
 gctagactga tctgtgtgat tgatggagat gaggacgacg atatgaggat ggcagctggt 420
 tataaggcta tctataatga taacattaag aagcctgaat ttgttctttg cgagtctgat 480
 gacaaggaag gagaacggat tgattcagat ttctcactg atactctgct tctccaacct 540
 catcgtggga agcgtgaatg tctttataca ggtttccaac tcgccaaaat ggacccatca 600
 gtgaacgctg tggttcttat cgatagtgat actgtgctgg agaaagatgc tatcttggag 660
 gttgtttacc ctcttgctg tgatctgaa attcaagctg tggctggaga gtgcaagatc 720
 tggaaacacag atactcttct ttctctgctt gtcgcatgga gatattactc cgcattctgt 780
 gtggagagga gcgctcaatc ctttttccgt accgttcaat gcgttggtyg tcttttggga 840
 gcttacaaaa ttgatatcat caaggagatt aaggacccat ggattagtca aaggtttctt 900
 ggtcagaagt gcacttatgg cgatgatcgt agattgacta acgaaatcct tatgaggggc 960
 aagaaagctg tttttactcc atttgcctg gcgatgctctg attcacctac aaatgttttc 1020
 cgttatattg tgcaacaaac acgttggagt aagagctggt gtagggagat ctggtacact 1080
 ttgttcgctg cttggaagca cgggcttagc ggaatttggc ttgcttttga atgcctttac 1140
 cagattacat actttttctt ggtgatctat ttgttttcac gtcttgccgt cgaggctgac 1200
 cctagagcac agactgcaac tgtgattggt tctactacag tcgcacttat taagtgtggc 1260
 tatttcagtt ttagagcaaa agatattaga gccttctatt ttgttttgta cacatttggt 1320
 tatttctttt gcatgatcc agctcgtatt accgctatga tgaccttgty ggacatcgga 1380
 tggggaacta gaggtggtaa cgaaaagcct tctgtgggaa caagggtggc cctttgggca 1440
 aaacaatatc tcatcgccta catgtggtgg gccgctgctg ttggtgccgg agtgtactca 1500

ES 2 562 211 T3

atcgttcata actggatggt tgactggaac tctttgagct atcgtttcgc tcttgtgggt 1560
 atttgttctt acattgtttt catcgtgatt gtgctcgttg tgtatttcac tggtaaaatc 1620
 acaacctgga atttcactaa acttcaaaag gaattgattg aagacagggt tctgtatgat 1680
 gctactacca acgcccagtc agtttaa 1707

5 <210> 4
 <211> 2298
 <212> ADN
 <213> *Mus musculus*

10 <220>
 <221> CDS
 <222> (150)..(2192)

15 <300>
 <308> BC050762.1
 <309> 08-03-2005
 <313> (150)..(2195)

<400> 4
 gagagcgaag cgagcgcctga gtcggactgt cgggtctgag ctgtcgcac ccagagtcct 60

ctcattgcca ccaccccggc ccgagctcac cctcgttct gaagctctcc gcgcgcccga 120

cagctcagcc ctcgcccgtg accaacatc atg tgc ggt ata ttt gct tat tta 173
 Met Cys Gly Ile Phe Ala Tyr Leu
 1 5

aat tac cat gtt cct cga aca aga cga gaa atc ttg gag aca cta atc 221
 Asn Tyr His Val Pro Arg Thr Arg Arg Glu Ile Leu Glu Thr Leu Ile
 10 15 20

aaa ggc ctt cag aga ctg gaa tac aga gga tat gat tct gct ggt gtg 269
 Lys Gly Leu Gln Arg Leu Glu Tyr Arg Gly Tyr Asp Ser Ala Gly Val
 25 30 35 40

gga ctt gac gga ggc aat gac aaa gac tgg gaa gcc aac gcc tgc aaa 317
 Gly Leu Asp Gly Gly Asn Asp Lys Asp Trp Glu Ala Asn Ala Cys Lys
 45 50 55

atc cag ctg att aag aag aaa gga aaa gtt aag gca ctg gat gaa gaa 365
 Ile Gln Leu Ile Lys Lys Lys Gly Lys Val Lys Ala Leu Asp Glu Glu
 60 65 70

ES 2 562 211 T3

gtt cac aaa caa caa gat atg gac ttg gat ata gaa ttt gat gtg cat	413
Val His Lys Gln Gln Asp Met Asp Leu Asp Ile Glu Phe Asp Val His	
75 80 85	
ctt gga ata gct cat acc cgt tgg gcg aca cat gga gaa ccc aat cct	461
Leu Gly Ile Ala His Thr Arg Trp Ala Thr His Gly Glu Pro Asn Pro	
90 95 100	
gtc aat agt cac ccc cag cgc tct gat aaa aat aat gaa ttc att gtt	509
Val Asn Ser His Pro Gln Arg Ser Asp Lys Asn Asn Glu Phe Ile Val	
105 110 115 120	
att cat aat gga atc atc acc aac tac aaa gac ttg aaa aag ttt ctg	557
Ile His Asn Gly Ile Ile Thr Asn Tyr Lys Asp Leu Lys Lys Phe Leu	
125 130 135	
gaa agc aaa ggc tat gac ttt gaa tct gaa aca gac aca gaa acc att	605
Glu Ser Lys Gly Tyr Asp Phe Glu Ser Glu Thr Asp Thr Glu Thr Ile	
140 145 150	
gcc aag ctc gtc aag tac atg tat gac aac tgg gag agc cag gac gtc	653
Ala Lys Leu Val Lys Tyr Met Tyr Asp Asn Trp Glu Ser Gln Asp Val	
155 160 165	
agt ttt acc acc ttg gtg gag aga gtt atc caa caa ttg gaa ggc gcc	701
Ser Phe Thr Thr Leu Val Glu Arg Val Ile Gln Gln Leu Glu Gly Ala	
170 175 180	
ttt gct ctt gtg ttt aaa agt gtc cat ttt ccc ggg caa gca gtt ggc	749
Phe Ala Leu Val Phe Lys Ser Val His Phe Pro Gly Gln Ala Val Gly	
185 190 195 200	
aca agg cga ggt agc cct ctc ttg att ggt gtg cgg agt gaa cat aag	797
Thr Arg Arg Gly Ser Pro Leu Leu Ile Gly Val Arg Ser Glu His Lys	
205 210 215	
ctt tct aca gat cac att ccg att ctg tac aga aca ggc aaa gac aag	845
Leu Ser Thr Asp His Ile Pro Ile Leu Tyr Arg Thr Gly Lys Asp Lys	
220 225 230	
aaa gga agc tgc ggt ctt tcc cgt gtg gac agc acg aca tgc ctg ttc	893
Lys Gly Ser Cys Gly Leu Ser Arg Val Asp Ser Thr Thr Cys Leu Phe	
235 240 245	
cct gtt gag gaa aag gca gtt gaa tat tac ttt gct tct gat gca agt	941
Pro Val Glu Glu Lys Ala Val Glu Tyr Tyr Phe Ala Ser Asp Ala Ser	
250 255 260	

ES 2 562 211 T3

gcc gtg ata gag cac acc aat cgt gtc atc ttt ctg gaa gat gat gat 989
 Ala Val Ile Glu His Thr Asn Arg Val Ile Phe Leu Glu Asp Asp Asp
 265 270 275 280

gtt gca gca gtg gtg gat ggc cgt ctc tct atc cac cga att aaa cga 1037
 Val Ala Ala Val Val Asp Gly Arg Leu Ser Ile His Arg Ile Lys Arg
 285 290 295

act gca gga gac cat cct ggc cga gct gtg caa act ctc cag atg gag 1085
 Thr Ala Gly Asp His Pro Gly Arg Ala Val Gln Thr Leu Gln Met Glu
 300 305 310

ctc cag cag atc atg aag ggc aac ttt agt tca ttt atg cag aag gaa 1133
 Leu Gln Gln Ile Met Lys Gly Asn Phe Ser Ser Phe Met Gln Lys Glu
 315 320 325

att ttt gag cag cca gaa tct gtt gtg aac aca atg aga gga aga gtc 1181
 Ile Phe Glu Gln Pro Glu Ser Val Val Asn Thr Met Arg Gly Arg Val
 330 335 340

aat ttt gat gac tac act gtg aat ttg gga ggt ttg aaa gat cac att 1229
 Asn Phe Asp Asp Tyr Thr Val Asn Leu Gly Gly Leu Lys Asp His Ile
 345 350 355 360

aag gag atc cag cgg tgt cgg cgg ttg att ctt att gct tgt ggc aca 1277
 Lys Glu Ile Gln Arg Cys Arg Arg Leu Ile Leu Ile Ala Cys Gly Thr
 365 370 375

agt tac cac gct ggt gtg gca acc cgt cag gtc ctg gag gag ctg acc 1325
 Ser Tyr His Ala Gly Val Ala Thr Arg Gln Val Leu Glu Glu Leu Thr
 380 385 390

gag ctg ccc gtg atg gtg gag ctt gcc agt gac ttc ttg gat aga aac 1373
 Glu Leu Pro Val Met Val Glu Leu Ala Ser Asp Phe Leu Asp Arg Asn
 395 400 405

act cca gtc ttt cga gat gat gtt tgc ttt ttc att agt caa tca ggc 1421
 Thr Pro Val Phe Arg Asp Asp Val Cys Phe Phe Ile Ser Gln Ser Gly
 410 415 420

gag aca gct gac acc ctg atg gga ctt cgt tac tgt aag gag aga gga 1469
 Glu Thr Ala Asp Thr Leu Met Gly Leu Arg Tyr Cys Lys Glu Arg Gly
 425 430 435 440

gcc tta act gtg ggg atc aca aat aca gtc ggc agt tct ata tca agg 1517
 Ala Leu Thr Val Gly Ile Thr Asn Thr Val Gly Ser Ser Ile Ser Arg
 445 450 455

gag aca gat tgc ggg gtt cat att aat gct ggt cct gag att ggc gtg 1565

ES 2 562 211 T3

Glu Thr Asp Cys Gly Val His Ile Asn Ala Gly Pro Glu Ile Gly Val	
460	465
470	
gcc agt aca aag gca tac acc agc cag ttt gtg tcc ctc gtg atg ttt	1613
Ala Ser Thr Lys Ala Tyr Thr Ser Gln Phe Val Ser Leu Val Met Phe	
475	480
485	
gct ctc atg atg tgt gat gac agg atc tcc atg caa gag aga cgc aaa	1661
Ala Leu Met Met Cys Asp Asp Arg Ile Ser Met Gln Glu Arg Arg Lys	
490	495
500	
gag atc atg ctc gga ctg aag cga ctg ccg gac ttg att aag gaa gtg	1709
Glu Ile Met Leu Gly Leu Lys Arg Leu Pro Asp Leu Ile Lys Glu Val	
505	510
515	520
ctg agc atg gat gat gaa atc cag aag ctg gcg acg gag ctt tac cac	1757
Leu Ser Met Asp Asp Glu Ile Gln Lys Leu Ala Thr Glu Leu Tyr His	
525	530
535	
cag aag tcg gtc ctg ata atg ggg cgg ggc tac cat tat gct aca tgc	1805
Gln Lys Ser Val Leu Ile Met Gly Arg Gly Tyr His Tyr Ala Thr Cys	
540	545
550	
ctt gaa ggg gct ctg aaa atc aag gag att act tat atg cat tcg gaa	1853
Leu Glu Gly Ala Leu Lys Ile Lys Glu Ile Thr Tyr Met His Ser Glu	
555	560
565	
ggc atc ctt gct ggt gag ctc aag cac ggc cct ctg gcc ttg gtg gac	1901
Gly Ile Leu Ala Gly Glu Leu Lys His Gly Pro Leu Ala Leu Val Asp	
570	575
580	
aag ttg atg cct gtc atc atg atc atc atg cga gac cac act tat gcc	1949
Lys Leu Met Pro Val Ile Met Ile Ile Met Arg Asp His Thr Tyr Ala	
585	590
595	600
aag tgc cag aac gct ctt cag cag gtg gtt gca cgg cag ggg cgt cca	1997
Lys Cys Gln Asn Ala Leu Gln Gln Val Val Ala Arg Gln Gly Arg Pro	
605	610
615	
gtc gtg atc tgt gat aag gag gat act gag acc att aag aat aca aaa	2045
Val Val Ile Cys Asp Lys Glu Asp Thr Glu Thr Ile Lys Asn Thr Lys	
620	625
630	
agg aca atc aag gtg ccc cac tca gtg gac tgc ttg cag ggc att ctc	2093
Arg Thr Ile Lys Val Pro His Ser Val Asp Cys Leu Gln Gly Ile Leu	
635	640
645	
agt gtg att ccc ctg cag ctg ctg gct ttc cac ctg gct gtg ctg aga	2141
Ser Val Ile Pro Leu Gln Leu Leu Ala Phe His Leu Ala Val Leu Arg	

ES 2 562 211 T3

```

        650                655                660
ggc tac gat gtt gat ttt cca cgg aat ctt gcc aaa tct gta aca gta      2189
Gly Tyr Asp Val Asp Phe Pro Arg Asn Leu Ala Lys Ser Val Thr Val
665                670                675                680

gag taacagacac ctgaaactta agacagttaa gcaacacgag ataccttttg      2242
Glu

tatttaaatt tttgatttaa actatcaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa      2298

```

5
 <210> 5
 <211> 681
 <212> PRT
 <213> *Mus musculus*

<400> 5

```

Met Cys Gly Ile Phe Ala Tyr Leu Asn Tyr His Val Pro Arg Thr Arg
1                5                10                15

```

```

Arg Glu Ile Leu Glu Thr Leu Ile Lys Gly Leu Gln Arg Leu Glu Tyr
                20                25                30

```

```

Arg Gly Tyr Asp Ser Ala Gly Val Gly Leu Asp Gly Gly Asn Asp Lys
                35                40                45

```

```

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

```

```

Lys Val Lys Ala Leu Asp Glu Glu Val His Lys Gln Gln Asp Met Asp
65                70                75                80

```

```

Leu Asp Ile Glu Phe Asp Val His Leu Gly Ile Ala His Thr Arg Trp
                85                90                95

```

```

Ala Thr His Gly Glu Pro Asn Pro Val Asn Ser His Pro Gln Arg Ser
                100                105                110

```

```

Asp Lys Asn Asn Glu Phe Ile Val Ile His Asn Gly Ile Ile Thr Asn

```

10

ES 2 562 211 T3

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

ES 2 562 211 T3

Phe Ser Ser Phe Met Gln Lys Glu Ile Phe Glu Gln Pro Glu Ser Val
 325 330 335

Val Asn Thr Met Arg Gly Arg Val Asn Phe Asp Asp Tyr Thr Val Asn
 340 345 350

Leu Gly Gly Leu Lys Asp His Ile Lys Glu Ile Gln Arg Cys Arg Arg
 355 360 365

Leu Ile Leu Ile Ala Cys Gly Thr Ser Tyr His Ala Gly Val Ala Thr
 370 375 380

Arg Gln Val Leu Glu Glu Leu Thr Glu Leu Pro Val Met Val Glu Leu
 385 390 395 400

Ala Ser Asp Phe Leu Asp Arg Asn Thr Pro Val Phe Arg Asp Asp Val
 405 410 415

Cys Phe Phe Ile Ser Gln Ser Gly Glu Thr Ala Asp Thr Leu Met Gly
 420 425 430

Leu Arg Tyr Cys Lys Glu Arg Gly Ala Leu Thr Val Gly Ile Thr Asn
 435 440 445

Thr Val Gly Ser Ser Ile Ser Arg Glu Thr Asp Cys Gly Val His Ile
 450 455 460

Asn Ala Gly Pro Glu Ile Gly Val Ala Ser Thr Lys Ala Tyr Thr Ser
 465 470 475 480

Gln Phe Val Ser Leu Val Met Phe Ala Leu Met Met Cys Asp Asp Arg
 485 490 495

Ile Ser Met Gln Glu Arg Arg Lys Glu Ile Met Leu Gly Leu Lys Arg
 500 505 510

ES 2 562 211 T3

Leu Pro Asp Leu Ile Lys Glu Val Leu Ser Met Asp Asp Glu Ile Gln
 515 520 525

Lys Leu Ala Thr Glu Leu Tyr His Gln Lys Ser Val Leu Ile Met Gly
 530 535 540

Arg Gly Tyr His Tyr Ala Thr Cys Leu Glu Gly Ala Leu Lys Ile Lys
 545 550 555 560

Glu Ile Thr Tyr Met His Ser Glu Gly Ile Leu Ala Gly Glu Leu Lys
 565 570 575

His Gly Pro Leu Ala Leu Val Asp Lys Leu Met Pro Val Ile Met Ile
 580 585 590

Ile Met Arg Asp His Thr Tyr Ala Lys Cys Gln Asn Ala Leu Gln Gln
 595 600 605

Val Val Ala Arg Gln Gly Arg Pro Val Val Ile Cys Asp Lys Glu Asp
 610 615 620

Thr Glu Thr Ile Lys Asn Thr Lys Arg Thr Ile Lys Val Pro His Ser
 625 630 635 640

Val Asp Cys Leu Gln Gly Ile Leu Ser Val Ile Pro Leu Gln Leu Leu
 645 650 655

Ala Phe His Leu Ala Val Leu Arg Gly Tyr Asp Val Asp Phe Pro Arg
 660 665 670

Asn Leu Ala Lys Ser Val Thr Val Glu
 675 680

<210> 6
 <211> 2049
 <212> ADN
 <213> *Mus musculus*

5

<220>
 <221> CDS
 <222> (1)..(2046)

10

ES 2 562 211 T3

Asp	Thr	Glu	Thr	Ile	Ala	Lys	Leu	Ile	Lys	Tyr	Val	Phe	Asp	Asn	Arg		
145					150					155					160		
gag	act	gag	gac	ata	acg	ttt	tcc	aca	ttg	gtc	gaa	aga	gtc	att	cag		528
Glu	Thr	Glu	Asp	Ile	Thr	Phe	Ser	Thr	Leu	Val	Glu	Arg	Val	Ile	Gln		
				165					170					175			
cag	ttg	gaa	ggc	gcc	ttt	gca	ctg	gtt	ttc	aag	agt	att	cac	tac	ccg		576
Gln	Leu	Glu	Gly	Ala	Phe	Ala	Leu	Val	Phe	Lys	Ser	Ile	His	Tyr	Pro		
			180					185					190				
gga	gaa	gct	gtc	gcc	acg	agg	aga	ggc	agc	ccc	ttg	ctc	atc	ggg	gta		624
Gly	Glu	Ala	Val	Ala	Thr	Arg	Arg	Gly	Ser	Pro	Leu	Leu	Ile	Gly	Val		
		195					200					205					
cga	agc	aaa	tac	aaa	ctc	tcc	aca	gag	cag	atc	ccc	gtc	tta	tat	ccg		672
Arg	Ser	Lys	Tyr	Lys	Leu	Ser	Thr	Glu	Gln	Ile	Pro	Val	Leu	Tyr	Pro		
	210					215				220							
aca	tgc	aat	atc	gag	aat	gtg	aag	aat	atc	tgc	aag	act	agg	atg	aag		720
Thr	Cys	Asn	Ile	Glu	Asn	Val	Lys	Asn	Ile	Cys	Lys	Thr	Arg	Met	Lys		
225					230					235					240		
aga	ctg	gac	agc	tcc	acc	tgc	ctg	cac	gct	gtg	ggc	gat	aaa	gct	gtg		768
Arg	Leu	Asp	Ser	Ser	Thr	Cys	Leu	His	Ala	Val	Gly	Asp	Lys	Ala	Val		
				245					250					255			
gaa	ttc	ttc	ttt	gct	tct	gat	gca	agt	gcc	atc	ata	gaa	cac	acc	aac		816
Glu	Phe	Phe	Phe	Ala	Ser	Asp	Ala	Ser	Ala	Ile	Ile	Glu	His	Thr	Asn		
			260					265					270				
cgg	gtc	atc	ttc	tta	gaa	gat	gat	gat	atc	gct	gca	gtg	gct	gat	ggg		864
Arg	Val	Ile	Phe	Leu	Glu	Asp	Asp	Asp	Ile	Ala	Ala	Val	Ala	Asp	Gly		
		275					280					285					
aaa	ctc	tcc	att	cac	cga	gtc	aag	cgc	tca	gct	act	gat	gac	ccc	tcc		912
Lys	Leu	Ser	Ile	His	Arg	Val	Lys	Arg	Ser	Ala	Thr	Asp	Asp	Pro	Ser		
	290					295					300						
cga	gcc	atc	cag	acc	ttg	cag	atg	gaa	ctg	cag	caa	ata	atg	aaa	ggt		960
Arg	Ala	Ile	Gln	Thr	Leu	Gln	Met	Glu	Leu	Gln	Gln	Ile	Met	Lys	Gly		
305					310					315					320		
aac	ttc	agc	gca	ttt	atg	cag	aag	gag	atc	ttc	gag	cag	cca	gaa	tca		1008
Asn	Phe	Ser	Ala	Phe	Met	Gln	Lys	Glu	Ile	Phe	Glu	Gln	Pro	Glu	Ser		
			325						330					335			
gtt	ttt	aat	acc	atg	aga	ggt	cgg	gtg	aat	ttt	gag	acc	aac	aca	gtg		1056
Val	Phe	Asn	Thr	Met	Arg	Gly	Arg	Val	Asn	Phe	Glu	Thr	Asn	Thr	Val		

ES 2 562 211 T3

ctc ctg ggt ggc ttg aag gac cat ttg aaa gag atc cga cga tgc cga	1104
Leu Leu Gly Gly Leu Lys Asp His Leu Lys Glu Ile Arg Arg Cys Arg	
355 360 365	
agg ctc att gtg att ggc tgt gga acc agc tac cat gcc gct gtg gct	1152
Arg Leu Ile Val Ile Gly Cys Gly Thr Ser Tyr His Ala Ala Val Ala	
370 375 380	
aca cgg caa gtc tta gag gaa ctg acc gag ctg cct gtg atg gtt gaa	1200
Thr Arg Gln Val Leu Glu Glu Leu Thr Glu Leu Pro Val Met Val Glu	
385 390 395 400	
ctt gcc agt gac ttt ctg gac agg aac aca cct gtg ttc agg gat gac	1248
Leu Ala Ser Asp Phe Leu Asp Arg Asn Thr Pro Val Phe Arg Asp Asp	
405 410 415	
gtt tgc ttt ttc ata agc caa tca ggt gag act gca gac acg ctc ctg	1296
Val Cys Phe Phe Ile Ser Gln Ser Gly Glu Thr Ala Asp Thr Leu Leu	
420 425 430	
gcg ctg cga tac tgt aag gat cga ggt gcg ctg acc gtg ggc atc acc	1344
Ala Leu Arg Tyr Cys Lys Asp Arg Gly Ala Leu Thr Val Gly Ile Thr	
435 440 445	
aac acc gtg ggt agc tcc atc tcc cgg gag act gac tgt ggc gtc cac	1392
Asn Thr Val Gly Ser Ser Ile Ser Arg Glu Thr Asp Cys Gly Val His	
450 455 460	
atc aac gca ggg ccc gag att ggg gtg gcc agc acc aag gcg tac acc	1440
Ile Asn Ala Gly Pro Glu Ile Gly Val Ala Ser Thr Lys Ala Tyr Thr	
465 470 475 480	
agc cag ttc atc tct ctg gtg atg ttt ggt ttg atg atg tct gaa gat	1488
Ser Gln Phe Ile Ser Leu Val Met Phe Gly Leu Met Met Ser Glu Asp	
485 490 495	
cga att tct cta cag aac agg aga caa gag atc atc cgt ggc ctc aga	1536
Arg Ile Ser Leu Gln Asn Arg Arg Gln Glu Ile Ile Arg Gly Leu Arg	
500 505 510	
tct tta ccg gag ctg atc aaa gaa gtg ctg tcc ctg gat gag aag atc	1584
Ser Leu Pro Glu Leu Ile Lys Glu Val Leu Ser Leu Asp Glu Lys Ile	
515 520 525	
cat gac ttg gcc ctg gag ctc tac aca caa agg tct ctc ctc gtg atg	1632
His Asp Leu Ala Leu Glu Leu Tyr Thr Gln Arg Ser Leu Leu Val Met	
530 535 540	

ES 2 562 211 T3

gga cgg gga tat aac tat gcc aca tgt ctg gaa ggt gcc ttg aaa att 1680
 Gly Arg Gly Tyr Asn Tyr Ala Thr Cys Leu Glu Gly Ala Leu Lys Ile
 545 550 555 560

aag gag ata acc tac atg cat tca gaa ggt atc cta gcc gga gag ctg 1728
 Lys Glu Ile Thr Tyr Met His Ser Glu Gly Ile Leu Ala Gly Glu Leu
 565 570 575

aag cac ggg ccc ctt gct ctc gtc gac aag cag atg cca gtc atc atg 1776
 Lys His Gly Pro Leu Ala Leu Val Asp Lys Gln Met Pro Val Ile Met
 580 585 590

gtc atc atg aag gat cct tgc ttt gcc aag tgc cag aat gcc ctg cag 1824
 Val Ile Met Lys Asp Pro Cys Phe Ala Lys Cys Gln Asn Ala Leu Gln
 595 600 605

cag gtc act gcc cgc cag ggt cgc cca atc ata ctg tgt tcc aag gat 1872
 Gln Val Thr Ala Arg Gln Gly Arg Pro Ile Ile Leu Cys Ser Lys Asp
 610 615 620

gac acc gag agc tcc aag ttt gca tat aaa acc att gaa ctt ccc cac 1920
 Asp Thr Glu Ser Ser Lys Phe Ala Tyr Lys Thr Ile Glu Leu Pro His
 625 630 635 640

aca gtg gac tgt ctc cag ggt atc ctg agc gtg att cca ctc cag ctt 1968
 Thr Val Asp Cys Leu Gln Gly Ile Leu Ser Val Ile Pro Leu Gln Leu
 645 650 655

ctg tcc ttc cac ctg gct gtc ctc cga ggt tat gat gtt gac ttc ccc 2016
 Leu Ser Phe His Leu Ala Val Leu Arg Gly Tyr Asp Val Asp Phe Pro
 660 665 670

aga aac cta gcc aag tct gtc act gtg gaa tga 2049
 Arg Asn Leu Ala Lys Ser Val Thr Val Glu
 675 680

<210> 7
 <211> 682
 <212> PRT
 <213> *Mus musculus*

<400> 7

Met Cys Gly Ile Phe Ala Tyr Met Asn Tyr Arg Val Pro Lys Thr Arg
 1 5 10 15

5

10

ES 2 562 211 T3

Lys Glu Ile Phe Glu Thr Leu Ile Arg Gly Leu Gln Arg Leu Glu Tyr
 20 25 30

Arg Gly Tyr Asp Ser Ala Gly Val Ala Ile Asp Gly Asn Asn His Glu
 35 40 45

Val Lys Glu Arg His Ile His Leu Val Lys Lys Arg Gly Lys Val Lys
 50 55 60

Ala Leu Asp Glu Glu Leu Tyr Lys Gln Asp Ser Met Asp Leu Lys Val
 65 70 75 80

Glu Phe Glu Thr His Phe Gly Ile Ala His Thr Arg Trp Ala Thr His
 85 90 95

Gly Val Pro Asn Ala Val Asn Ser His Pro Gln Arg Ser Asp Lys Asp
 100 105 110

Asn Glu Phe Val Val Ile His Asn Gly Ile Ile Thr Asn Tyr Lys Asp
 115 120 125

Leu Arg Lys Phe Leu Glu Ser Lys Gly Tyr Glu Phe Glu Ser Glu Thr
 130 135 140

Asp Thr Glu Thr Ile Ala Lys Leu Ile Lys Tyr Val Phe Asp Asn Arg
 145 150 155 160

Glu Thr Glu Asp Ile Thr Phe Ser Thr Leu Val Glu Arg Val Ile Gln
 165 170 175

Gln Leu Glu Gly Ala Phe Ala Leu Val Phe Lys Ser Ile His Tyr Pro
 180 185 190

Gly Glu Ala Val Ala Thr Arg Arg Gly Ser Pro Leu Leu Ile Gly Val
 195 200 205

Arg Ser Lys Tyr Lys Leu Ser Thr Glu Gln Ile Pro Val Leu Tyr Pro

ES 2 562 211 T3

210	215	220																	
Thr	Cys	Asn	Ile	Glu	Asn	Val	Lys	Asn	Ile	Cys	Lys	Thr	Arg	Met	Lys				
225					230					235					240				
Arg	Leu	Asp	Ser	Ser	Thr	Cys	Leu	His	Ala	Val	Gly	Asp	Lys	Ala	Val				
				245					250						255				
Glu	Phe	Phe	Phe	Ala	Ser	Asp	Ala	Ser	Ala	Ile	Ile	Glu	His	Thr	Asn				
			260					265					270						
Arg	Val	Ile	Phe	Leu	Glu	Asp	Asp	Asp	Ile	Ala	Ala	Val	Ala	Asp	Gly				
		275					280						285						
Lys	Leu	Ser	Ile	His	Arg	Val	Lys	Arg	Ser	Ala	Thr	Asp	Asp	Pro	Ser				
290						295					300								
Arg	Ala	Ile	Gln	Thr	Leu	Gln	Met	Glu	Leu	Gln	Gln	Ile	Met	Lys	Gly				
305					310					315					320				
Asn	Phe	Ser	Ala	Phe	Met	Gln	Lys	Glu	Ile	Phe	Glu	Gln	Pro	Glu	Ser				
				325					330					335					
Val	Phe	Asn	Thr	Met	Arg	Gly	Arg	Val	Asn	Phe	Glu	Thr	Asn	Thr	Val				
			340					345					350						
Leu	Leu	Gly	Gly	Leu	Lys	Asp	His	Leu	Lys	Glu	Ile	Arg	Arg	Cys	Arg				
		355					360					365							
Arg	Leu	Ile	Val	Ile	Gly	Cys	Gly	Thr	Ser	Tyr	His	Ala	Ala	Val	Ala				
370						375					380								
Thr	Arg	Gln	Val	Leu	Glu	Glu	Leu	Thr	Glu	Leu	Pro	Val	Met	Val	Glu				
385				390					395						400				
Leu	Ala	Ser	Asp	Phe	Leu	Asp	Arg	Asn	Thr	Pro	Val	Phe	Arg	Asp	Asp				
				405					410					415					

ES 2 562 211 T3

Val Cys Phe Phe Ile Ser Gln Ser Gly Glu Thr Ala Asp Thr Leu Leu
 420 425 430

Ala Leu Arg Tyr Cys Lys Asp Arg Gly Ala Leu Thr Val Gly Ile Thr
 435 440 445

Asn Thr Val Gly Ser Ser Ile Ser Arg Glu Thr Asp Cys Gly Val His
 450 455 460

Ile Asn Ala Gly Pro Glu Ile Gly Val Ala Ser Thr Lys Ala Tyr Thr
 465 470 475 480

Ser Gln Phe Ile Ser Leu Val Met Phe Gly Leu Met Met Ser Glu Asp
 485 490 495

Arg Ile Ser Leu Gln Asn Arg Arg Gln Glu Ile Ile Arg Gly Leu Arg
 500 505 510

Ser Leu Pro Glu Leu Ile Lys Glu Val Leu Ser Leu Asp Glu Lys Ile
 515 520 525

His Asp Leu Ala Leu Glu Leu Tyr Thr Gln Arg Ser Leu Leu Val Met
 530 535 540

Gly Arg Gly Tyr Asn Tyr Ala Thr Cys Leu Glu Gly Ala Leu Lys Ile
 545 550 555 560

Lys Glu Ile Thr Tyr Met His Ser Glu Gly Ile Leu Ala Gly Glu Leu
 565 570 575

Lys His Gly Pro Leu Ala Leu Val Asp Lys Gln Met Pro Val Ile Met
 580 585 590

Val Ile Met Lys Asp Pro Cys Phe Ala Lys Cys Gln Asn Ala Leu Gln
 595 600 605

ES 2 562 211 T3

Gln Val Thr Ala Arg Gln Gly Arg Pro Ile Ile Leu Cys Ser Lys Asp
 610 615 620

Asp Thr Glu Ser Ser Lys Phe Ala Tyr Lys Thr Ile Glu Leu Pro His
 625 630 635 640

Thr Val Asp Cys Leu Gln Gly Ile Leu Ser Val Ile Pro Leu Gln Leu
 645 650 655

Leu Ser Phe His Leu Ala Val Leu Arg Gly Tyr Asp Val Asp Phe Pro
 660 665 670

Arg Asn Leu Ala Lys Ser Val Thr Val Glu
 675 680

<210> 8
 <211> 1830
 <212> ADN
 <213> *Escherichia coli*

<220>
 <221> CDS
 <222> (1)..(1827)

<300>
 <308> U00096.2
 <309> 08-09-2005
 <313> (3909862)..(3911691)

<400> 8

atg tgt gga att gtt ggc gcg atc gcg caa cgt gat gta gca gaa atc 48
 Met Cys Gly Ile Val Gly Ala Ile Ala Gln Arg Asp Val Ala Glu Ile
 1 5 10 15

ctt ctt gaa ggt tta cgt cgt ctg gaa tac cgc gga tat gac tct gcc 96
 Leu Leu Glu Gly Leu Arg Arg Leu Glu Tyr Arg Gly Tyr Asp Ser Ala
 20 25 30

ggc ctg gcc gtt gtt gat gca gaa ggt cat atg acc cgc ctg cgt cgc 144
 Gly Leu Ala Val Val Asp Ala Glu Gly His Met Thr Arg Leu Arg Arg
 35 40 45

5

10

15

ES 2 562 211 T3

ctc ggt aaa gtc cag atg ctg gca cag gca gcg gaa gaa cat cct ctg	192
Leu Gly Lys Val Gln Met Leu Ala Gln Ala Ala Glu His Pro Leu	
50 55 60	
cat ggc ggc act ggt att gct cac act cgc tgg gcg acc cac ggt gaa	240
His Gly Gly Thr Gly Ile Ala His Thr Arg Trp Ala Thr His Gly Glu	
65 70 75 80	
cct tca gaa gtg aat gcg cat ccg cat gtt tct gaa cac att gtg gtg	288
Pro Ser Glu Val Asn Ala His Pro His Val Ser Glu His Ile Val Val	
85 90 95	
gtg cat aac ggc atc atc gaa aac cat gaa ccg ctg cgt gaa gag cta	336
Val His Asn Gly Ile Ile Glu Asn His Glu Pro Leu Arg Glu Glu Leu	
100 105 110	
aaa gcg cgt ggc tat acc ttc gtt tct gaa acc gac acc gaa gtg att	384
Lys Ala Arg Gly Tyr Thr Phe Val Ser Glu Thr Asp Thr Glu Val Ile	
115 120 125	
gcc cat ctg gtg aac tgg gag ctg aaa caa ggc ggg act ctg cgt gag	432
Ala His Leu Val Asn Trp Glu Leu Lys Gln Gly Gly Thr Leu Arg Glu	
130 135 140	
gcc gtt ctg cgt gct atc ccg cag ctg cgt ggt gcg tac ggt aca gtg	480
Ala Val Leu Arg Ala Ile Pro Gln Leu Arg Gly Ala Tyr Gly Thr Val	
145 150 155 160	
atc atg gac tcc cgt cac ccg gat acc ctg ctg gcg gca cgt tct ggt	528
Ile Met Asp Ser Arg His Pro Asp Thr Leu Leu Ala Ala Arg Ser Gly	
165 170 175	
agt ccg ctg gtg att ggc ctg ggg atg ggc gaa aac ttt atc gct tct	576
Ser Pro Leu Val Ile Gly Leu Gly Met Gly Glu Asn Phe Ile Ala Ser	
180 185 190	
gac cag ctg gcg ctg ttg ccg gtg acc cgt cgc ttt atc ttc ctt gaa	624
Asp Gln Leu Ala Leu Leu Pro Val Thr Arg Arg Phe Ile Phe Leu Glu	
195 200 205	
gag ggc gat att gcg gaa atc act cgc cgt tcc gta aac atc ttc gat	672
Glu Gly Asp Ile Ala Glu Ile Thr Arg Arg Ser Val Asn Ile Phe Asp	
210 215 220	
aaa act ggc gcg gaa gta aaa cgt cag gat atc gaa tcc aat ctg caa	720
Lys Thr Gly Ala Glu Val Lys Arg Gln Asp Ile Glu Ser Asn Leu Gln	
225 230 235 240	
tat gac gcg ggc gat aaa ggc att tac cgt cac tac atg cag aaa gag	768

ES 2 562 211 T3

Tyr	Asp	Ala	Gly	Asp	Lys	Gly	Ile	Tyr	Arg	His	Tyr	Met	Gln	Lys	Glu		
				245					250					255			
atc	tac	gaa	cag	ccg	aac	gcg	atc	aaa	aac	acc	ctt	acc	gga	cgc	atc		816
Ile	Tyr	Glu	Gln	Pro	Asn	Ala	Ile	Lys	Asn	Thr	Leu	Thr	Gly	Arg	Ile		
			260					265					270				
agc	cac	ggt	cag	gtt	gat	tta	agc	gag	ctg	gga	ccg	aac	gcc	gac	gaa		864
Ser	His	Gly	Gln	Val	Asp	Leu	Ser	Glu	Leu	Gly	Pro	Asn	Ala	Asp	Glu		
		275					280					285					
ctg	ctg	tcg	aag	gtt	gag	cat	att	cag	atc	ctc	gcc	tgt	ggg	act	tct		912
Leu	Leu	Ser	Lys	Val	Glu	His	Ile	Gln	Ile	Leu	Ala	Cys	Gly	Thr	Ser		
	290						295				300						
tat	aac	tcc	ggt	atg	gtt	tcc	cgc	tac	tgg	ttt	gaa	tcg	cta	gca	ggt		960
Tyr	Asn	Ser	Gly	Met	Val	Ser	Arg	Tyr	Trp	Phe	Glu	Ser	Leu	Ala	Gly		
305					310					315					320		
att	ccg	tgc	gac	gtc	gaa	atc	gcc	tct	gaa	ttc	cgc	tat	cgc	aaa	tct		1008
Ile	Pro	Cys	Asp	Val	Glu	Ile	Ala	Ser	Glu	Phe	Arg	Tyr	Arg	Lys	Ser		
				325					330						335		
gcc	gtg	cgt	cgt	aac	agc	ctg	atg	atc	acc	ttg	tca	cag	tct	ggc	gaa		1056
Ala	Val	Arg	Arg	Asn	Ser	Leu	Met	Ile	Thr	Leu	Ser	Gln	Ser	Gly	Glu		
				340				345						350			
acc	gcg	gat	acc	ctg	gct	ggc	ctg	cgt	ctg	tcg	aaa	gag	ctg	ggg	tac		1104
Thr	Ala	Asp	Thr	Leu	Ala	Gly	Leu	Arg	Leu	Ser	Lys	Glu	Leu	Gly	Tyr		
			355				360					365					
ctt	ggt	tca	ctg	gca	atc	tgt	aac	gtt	ccg	ggg	tct	tct	ctg	gtg	cgc		1152
Leu	Gly	Ser	Leu	Ala	Ile	Cys	Asn	Val	Pro	Gly	Ser	Ser	Leu	Val	Arg		
	370					375					380						
gaa	tcc	gat	ctg	gcg	cta	atg	acc	aac	gcg	ggg	aca	gaa	atc	ggc	gtg		1200
Glu	Ser	Asp	Leu	Ala	Leu	Met	Thr	Asn	Ala	Gly	Thr	Glu	Ile	Gly	Val		
385					390					395					400		
gca	tcc	act	aaa	gca	ttc	acc	act	cag	tta	act	gtg	ctg	ttg	atg	ctg		1248
Ala	Ser	Thr	Lys	Ala	Phe	Thr	Thr	Gln	Leu	Thr	Val	Leu	Leu	Met	Leu		
				405					410					415			
gtg	gcg	aag	ctg	tct	cgc	ctg	aaa	ggg	ctg	gat	gcc	tcc	att	gaa	cat		1296
Val	Ala	Lys	Leu	Ser	Arg	Leu	Lys	Gly	Leu	Asp	Ala	Ser	Ile	Glu	His		
			420					425					430				
gac	atc	gtg	cat	ggg	ctg	cag	gcg	ctg	ccg	agc	cgt	att	gag	cag	atg		1344
Asp	Ile	Val	His	Gly	Leu	Gln	Ala	Leu	Pro	Ser	Arg	Ile	Glu	Gln	Met		

ES 2 562 211 T3

435	440	445	
ctg tct cag gac aaa cgc att gaa gcg ctg gca gaa gat ttc tct gac			1392
Leu Ser Gln Asp Lys Arg Ile Glu Ala Leu Ala Glu Asp Phe Ser Asp			
450	455	460	
aaa cat cac gcg ctg ttc ctg ggc cgt ggc gat cag tac cca atc gcg			1440
Lys His His Ala Leu Phe Leu Gly Arg Gly Asp Gln Tyr Pro Ile Ala			
465	470	475	480
ctg gaa ggc gca ttg aag ttg aaa gag atc tct tac att cac gct gaa			1488
Leu Glu Gly Ala Leu Lys Leu Lys Glu Ile Ser Tyr Ile His Ala Glu			
	485	490	495
gcc tac gct gct ggc gaa ctg aaa cac ggt ccg ctg gcg cta att gat			1536
Ala Tyr Ala Ala Gly Glu Leu Lys His Gly Pro Leu Ala Leu Ile Asp			
	500	505	510
gcc gat atg ccg gtt att gtt gtt gca ccg aac aac gaa ttg ctg gaa			1584
Ala Asp Met Pro Val Ile Val Val Ala Pro Asn Asn Glu Leu Leu Glu			
	515	520	525
aaa ctg aaa tcc aac att gaa gaa gtt cgc gcg cgt ggc ggt cag ttg			1632
Lys Leu Lys Ser Asn Ile Glu Glu Val Arg Ala Arg Gly Gly Gln Leu			
	530	535	540
tat gtc ttc gcc gat cag gat gcg ggt ttt gta agt agc gat aac atg			1680
Tyr Val Phe Ala Asp Gln Asp Ala Gly Phe Val Ser Ser Asp Asn Met			
545	550	555	560
cac atc atc gag atg ccg cat gtg gaa gag gtg att gca ccg atc ttc			1728
His Ile Ile Glu Met Pro His Val Glu Glu Val Ile Ala Pro Ile Phe			
	565	570	575
tac acc gtt ccg ctg cag ctg ctg gct tac cat gtc gcg ctg atc aaa			1776
Tyr Thr Val Pro Leu Gln Leu Leu Ala Tyr His Val Ala Leu Ile Lys			
	580	585	590
ggc acc gac gtt gac cag ccg cgt aac ctg gca aaa tcg gtt acg gtt			1824
Gly Thr Asp Val Asp Gln Pro Arg Asn Leu Ala Lys Ser Val Thr Val			
	595	600	605
gag taa			1830
Glu			

<210> 9
 <211> 609
 <212> PRT
 <213> *Escherichia coli*

<400> 9

ES 2 562 211 T3

Met Cys Gly Ile Val Gly Ala Ile Ala Gln Arg Asp Val Ala Glu Ile
 1 5 10 15

Leu Leu Glu Gly Leu Arg Arg Leu Glu Tyr Arg Gly Tyr Asp Ser Ala
 20 25 30

Gly Leu Ala Val Val Asp Ala Glu Gly His Met Thr Arg Leu Arg Arg
 35 40 45

Leu Gly Lys Val Gln Met Leu Ala Gln Ala Ala Glu Glu His Pro Leu
 50 55 60

His Gly Gly Thr Gly Ile Ala His Thr Arg Trp Ala Thr His Gly Glu
 65 70 75 80

Pro Ser Glu Val Asn Ala His Pro His Val Ser Glu His Ile Val Val
 85 90 95

Val His Asn Gly Ile Ile Glu Asn His Glu Pro Leu Arg Glu Glu Leu
 100 105 110

Lys Ala Arg Gly Tyr Thr Phe Val Ser Glu Thr Asp Thr Glu Val Ile
 115 120 125

Ala His Leu Val Asn Trp Glu Leu Lys Gln Gly Gly Thr Leu Arg Glu
 130 135 140

Ala Val Leu Arg Ala Ile Pro Gln Leu Arg Gly Ala Tyr Gly Thr Val
 145 150 155 160

Ile Met Asp Ser Arg His Pro Asp Thr Leu Leu Ala Ala Arg Ser Gly
 165 170 175

ES 2 562 211 T3

Ser Pro Leu Val Ile Gly Leu Gly Met Gly Glu Asn Phe Ile Ala Ser
 180 185 190

Asp Gln Leu Ala Leu Leu Pro Val Thr Arg Arg Phe Ile Phe Leu Glu
 195 200 205

Glu Gly Asp Ile Ala Glu Ile Thr Arg Arg Ser Val Asn Ile Phe Asp
 210 215 220

Lys Thr Gly Ala Glu Val Lys Arg Gln Asp Ile Glu Ser Asn Leu Gln
 225 230 235 240

Tyr Asp Ala Gly Asp Lys Gly Ile Tyr Arg His Tyr Met Gln Lys Glu
 245 250 255

Ile Tyr Glu Gln Pro Asn Ala Ile Lys Asn Thr Leu Thr Gly Arg Ile
 260 265 270

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

Leu Leu Ser Lys Val Glu His Ile Gln Ile Leu Ala Cys Gly Thr Ser
 290 295 300

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

Ile Pro Cys Asp Val Glu Ile Ala Ser Glu Phe Arg Tyr Arg Lys Ser
 325 330 335

Ala Val Arg Arg Asn Ser Leu Met Ile Thr Leu Ser Gln Ser Gly Glu
 340 345 350

Thr Ala Asp Thr Leu Ala Gly Leu Arg Leu Ser Lys Glu Leu Gly Tyr
 355 360 365

Leu Gly Ser Leu Ala Ile Cys Asn Val Pro Gly Ser Ser Leu Val Arg

ES 2 562 211 T3

Tyr Thr Val Pro Leu Gln Leu Leu Ala Tyr His Val Ala Leu Ile Lys
 580 585 590

Gly Thr Asp Val Asp Gln Pro Arg Asn Leu Ala Lys Ser Val Thr Val
 595 600 605

Glu

<210> 10
 <211> 1830
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia sintética que codifica una proteína de *Escherichia coli* que tiene la actividad de una GFAT

<400> 10

atgtgcggaa ttgttggtgc tatcgcccaa agagacgttg ctgagatttt gtagagggt 60
 ctgccaaggc tagagtatag aggatatgac tccgctggtc tggctgtcgt tgatgctgag 120
 ggtcatatga caaggctaag aaggtagga aaggttcaga tgcttgctca ggcagctgag 180
 gaacatccat tgcattggagg tactggtatt gcacatacca ggtgggctac tcatggggag 240
 ccatcagaag ttaatgctca tccacatgtg agtgagcata tcgttgtagt tcacaatggg 300
 ataattgaaa accacgaacc attgagggaa gagttaaagg caagaggata tacttttgtg 360
 agtgagactg aactgaggt tattgcacat ttagtgaact gggactcaa acaggggggc 420
 acattgcgtg aggctgtgtt aagagctatt cctcaactta gaggtgcata cggtactggt 480
 attatggatt caagacaccc agatactctc cttgcagcta gatcaggtag tcccttggtc 540
 ataggacttg gaatgggtga aaattttatc gctagcgacc aattggcctt attgccagtt 600
 acaagacgat ttattttcct tgaagagggc gatattgctg agattactag aaggctctgtg 660
 aacatctttg ataagactgg cgctgaggtt aaacgtcagg atatcgagtc taacctcaa 720
 tacgatgctg gtgataaagg aatttacagg cattatatgc aaaaggaaat ttatgaacaa 780

ES 2 562 211 T3

ccaaatgcta tcaaaaacac acttactggc cgtattttctc atggacaggt cgattttaagc 840
 gagcttggtc ctaatgcaga cgaactgcta tcaaaagttg agcacataca gatactggca 900
 tgcggaacta gttataattc aggaatggtc tctagatact ggttcgaaag cttggcaggt 960
 ataccttgtg atgtagagat cgcttctgag tttaggtata gaaagtctgc tgtgcgtaga 1020
 aatcattaa tgattacatt atctcaatcc ggagaaacag cagatacact ggctggattg 1080
 aggctttcta aggaactcgg atatctgggt tcacttgcta tttgtaatgt accaggttcc 1140
 tcattggttc gtgaatcaga tctagcactt atgacaaatg caggaactga aataggtgtg 1200
 gcaagtacca aggctttcac aaccaactg accgtacttt taatgttggt agcaaaactc 1260
 agtcgattaa aggggctaga tgcattctatc gaacatgata ttgttcacgg gcttcaagct 1320
 ctccctcaa gaattgaaca aatgctttca caagataaga gaatagaggc attggctgaa 1380
 gatTTTTcCG acaaacatca cgcattgttt cttggacgtg gcgatcaata tccaattgca 1440
 ttggaaggag ctttgaagtt gaaagaaata agttacattc acgcagaagc atatgcagct 1500
 ggagaactca agcatggctc tttggcactc atcgacgtg acatgccctg gatcgtagtg 1560
 gctcctaata acgaactgct cgaaaagctt aatcaaata tcgaagaggt tcgagctaga 1620
 ggaggtcagc tttacgtttt cgctgaacaa gatgctggat tcgtgtcaag cgataaatg 1680
 catataattg aatgcctca cgttgaagaa gtgattgcac ctatatttta tacagtccca 1740
 ttgcaacttc tagcttacca tgttgcactt attaaaggaa ctgatgttga tcagcctaga 1800
 aacctagcaa aatctgtaac agtcgaataa 1830

5 <210> 11
 <211> 1260
 <212> ADN
 <213> Virus 1 *Paramecium bursaria Chlorella*

10 <220>
 <221> CDS
 <222> (62)..(1228)

15 <300>
 <308> U42580.4
 <309> 20-09-2004
 <313> (291749)..(292918)

<400> 11

ES 2 562 211 T3

atcaacgtga tttatatttt aaacaaagac cattcacatc tttagtactt aattaattat 60

a atg tca cga atc gca gtc gtt ggt tgt ggt tac gtc gga acc gct tgt 109
 Met Ser Arg Ile Ala Val Val Gly Cys Gly Tyr Val Gly Thr Ala Cys
 1 5 10 15

gca gta ctt ctt gct caa aaa aac gaa gtc atc gtg ctt gat att agc 157
 Ala Val Leu Leu Ala Gln Lys Asn Glu Val Ile Val Leu Asp Ile Ser
 20 25 30

gaa gac cgt gtt caa cta atc aag aac aag aag agt cca atc gag gac 205
 Glu Asp Arg Val Gln Leu Ile Lys Asn Lys Lys Ser Pro Ile Glu Asp
 35 40 45

aag gaa atc gaa gag ttt ctc gaa acg aaa gac ctg aac ctg acc gcg 253
 Lys Glu Ile Glu Glu Phe Leu Glu Thr Lys Asp Leu Asn Leu Thr Ala
 50 55 60

acg act gac aag gtt ctt gca tac gaa aac gcc gaa ttt gtc atc atc 301
 Thr Thr Asp Lys Val Leu Ala Tyr Glu Asn Ala Glu Phe Val Ile Ile
 65 70 75 80

gca acc ccg act gac tat gac gtg gtt act agg tat ttt aac acg aaa 349
 Ala Thr Pro Thr Asp Tyr Asp Val Val Thr Arg Tyr Phe Asn Thr Lys
 85 90 95

tct gtg gaa aac gtc att ggg gac gtg atc aaa aat aca cag acc cat 397
 Ser Val Glu Asn Val Ile Gly Asp Val Ile Lys Asn Thr Gln Thr His
 100 105 110

cca act atc gtg att aaa tct acc atc ccc att gga ttt gtt gat aag 445
 Pro Thr Ile Val Ile Lys Ser Thr Ile Pro Ile Gly Phe Val Asp Lys
 115 120 125

gtt cgt gag caa ttc gac tac caa aat atc att ttc tcc cca gaa ttt 493
 Val Arg Glu Gln Phe Asp Tyr Gln Asn Ile Ile Phe Ser Pro Glu Phe
 130 135 140

ctg cgt gaa ggt aga gcc ttg tat gat aat ctc tac cca tcc cgt atc 541
 Leu Arg Glu Gly Arg Ala Leu Tyr Asp Asn Leu Tyr Pro Ser Arg Ile
 145 150 155 160

atc gta gga gat gat tcc ccc att gcg ctt aag ttc gca aac ctt ctc 589
 Ile Val Gly Asp Asp Ser Pro Ile Ala Leu Lys Phe Ala Asn Leu Leu

ES 2 562 211 T3

	165		170		175		
ggt gaa ggt tct aaa act ccg ctt gcc cct gtc ctg acg atg gga act							637
Val Glu Gly Ser Lys Thr Pro Leu Ala Pro Val Leu Thr Met Gly Thr							
	180		185		190		
cgc gaa gcc gag gcc gtc aaa cta ttc tct aac acg tat ctt gca atg							685
Arg Glu Ala Glu Ala Val Lys Leu Phe Ser Asn Thr Tyr Leu Ala Met							
	195		200		205		
cga gtt gca tac ttc aac gaa cta gat aca ttc gca atg tct cac ggt							733
Arg Val Ala Tyr Phe Asn Glu Leu Asp Thr Phe Ala Met Ser His Gly							
	210		215		220		
atg aat gcg aaa gaa atc att gat ggt gtg act ctg gag cct cgc att							781
Met Asn Ala Lys Glu Ile Ile Asp Gly Val Thr Leu Glu Pro Arg Ile							
	225		230		235		240
ggt cag ggg tac tca aac cct tcg ttc ggt tat gga gct tat tgc ttt							829
Gly Gln Gly Tyr Ser Asn Pro Ser Phe Gly Tyr Gly Ala Tyr Cys Phe							
	245		250		255		
cca aag gat acg aag caa ctg ctg gct aat ttc gag gga gtg cct caa							877
Pro Lys Asp Thr Lys Gln Leu Leu Ala Asn Phe Glu Gly Val Pro Gln							
	260		265		270		
gat atc atc gga gca att gta gaa tca aat gag act cgc aag gaa gtg							925
Asp Ile Ile Gly Ala Ile Val Glu Ser Asn Glu Thr Arg Lys Glu Val							
	275		280		285		
att gtg agt gaa gta gaa aat cgt ttc ccc acg act gtt ggt gtg tat							973
Ile Val Ser Glu Val Glu Asn Arg Phe Pro Thr Thr Val Gly Val Tyr							
	290		295		300		
aag ctc gcc gct aaa gcg ggt tct gat aat ttt cgg agt tct gca att							1021
Lys Leu Ala Ala Lys Ala Gly Ser Asp Asn Phe Arg Ser Ser Ala Ile							
	305		310		315		320
gta gac ata atg gag cga ctt gca aac aag ggt tat cac att aag att							1069
Val Asp Ile Met Glu Arg Leu Ala Asn Lys Gly Tyr His Ile Lys Ile							
	325		330		335		
ttc gaa cca act gtg gaa caa ttc gaa aac ttt gaa gtt gat aac aac							1117
Phe Glu Pro Thr Val Glu Gln Phe Glu Asn Phe Glu Val Asp Asn Asn							
	340		345		350		
ctg aca aca ttt gcg act gag agc gat gta att atc gca aac aga gtt							1165
Leu Thr Thr Phe Ala Thr Glu Ser Asp Val Ile Ile Ala Asn Arg Val							
	355		360		365		

ES 2 562 211 T3

ccc gtt gaa cat cgc att ctc ttt ggt aaa aaa tta atc aca cgt gat 1213
 Pro Val Glu His Arg Ile Leu Phe Gly Lys Lys Leu Ile Thr Arg Asp
 370 375 380

gta tat ggc gat aac taaaatgttt tcaatatgat gttgtaatg at 1260
 Val Tyr Gly Asp Asn
 385

<210> 12
 <211> 389
 <212> PRT
 <213> Virus 1 *Paramecium bursaria* *Chlorella*

5

<400> 12

Met Ser Arg Ile Ala Val Val Gly Cys Gly Tyr Val Gly Thr Ala Cys
 1 5 10 15

Ala Val Leu Leu Ala Gln Lys Asn Glu Val Ile Val Leu Asp Ile Ser
 20 25 30

Glu Asp Arg Val Gln Leu Ile Lys Asn Lys Lys Ser Pro Ile Glu Asp
 35 40 45

Lys Glu Ile Glu Glu Phe Leu Glu Thr Lys Asp Leu Asn Leu Thr Ala
 50 55 60

Thr Thr Asp Lys Val Leu Ala Tyr Glu Asn Ala Glu Phe Val Ile Ile
 65 70 75 80

Ala Thr Pro Thr Asp Tyr Asp Val Val Thr Arg Tyr Phe Asn Thr Lys
 85 90 95

Ser Val Glu Asn Val Ile Gly Asp Val Ile Lys Asn Thr Gln Thr His
 100 105 110

Pro Thr Ile Val Ile Lys Ser Thr Ile Pro Ile Gly Phe Val Asp Lys
 115 120 125

10

ES 2 562 211 T3

Val Arg Glu Gln Phe Asp Tyr Gln Asn Ile Ile Phe Ser Pro Glu Phe
 130 135 140

Leu Arg Glu Gly Arg Ala Leu Tyr Asp Asn Leu Tyr Pro Ser Arg Ile
 145 150 155 160

Ile Val Gly Asp Asp Ser Pro Ile Ala Leu Lys Phe Ala Asn Leu Leu
 165 170 175

Val Glu Gly Ser Lys Thr Pro Leu Ala Pro Val Leu Thr Met Gly Thr
 180 185 190

Arg Glu Ala Glu Ala Val Lys Leu Phe Ser Asn Thr Tyr Leu Ala Met
 195 200 205

Arg Val Ala Tyr Phe Asn Glu Leu Asp Thr Phe Ala Met Ser His Gly
 210 215 220

Met Asn Ala Lys Glu Ile Ile Asp Gly Val Thr Leu Glu Pro Arg Ile
 225 230 235 240

Gly Gln Gly Tyr Ser Asn Pro Ser Phe Gly Tyr Gly Ala Tyr Cys Phe
 245 250 255

Pro Lys Asp Thr Lys Gln Leu Leu Ala Asn Phe Glu Gly Val Pro Gln
 260 265 270

Asp Ile Ile Gly Ala Ile Val Glu Ser Asn Glu Thr Arg Lys Glu Val
 275 280 285

Ile Val Ser Glu Val Glu Asn Arg Phe Pro Thr Thr Val Gly Val Tyr
 290 295 300

Lys Leu Ala Ala Lys Ala Gly Ser Asp Asn Phe Arg Ser Ser Ala Ile
 305 310 315 320

Val Asp Ile Met Glu Arg Leu Ala Asn Lys Gly Tyr His Ile Lys Ile

ES 2 562 211 T3

325 330 335
Phe Glu Pro Thr Val Glu Gln Phe Glu Asn Phe Glu Val Asp Asn Asn
340 345 350

Leu Thr Thr Phe Ala Thr Glu Ser Asp Val Ile Ile Ala Asn Arg Val
355 360 365

Pro Val Glu His Arg Ile Leu Phe Gly Lys Lys Leu Ile Thr Arg Asp
370 375 380

Val Tyr Gly Asp Asn
385

5 <210> 13
<211> 1170
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Secuencia sintética que codifica una proteína de virus de *Paramecium bursaria Chlorella* que tiene la actividad de una UDP-Glc-DH

<400> 13

atgtctcgca tagctgttgt aggatgtggc tatgtgggaa ctgcatgtgc ggttctactt 60
gctcaaaaga acgaagttat tgtgcttgat attagtgaag accgtgttca acttattaag 120
aacaagaagt ctctattga ggataaggaa atcgaagagt tcttggaac aaaggatctt 180
aatcttactg cgactacaga taaggttctt gcctacgaga acgctgagtt tgtgataatc 240
gctacaccaa ccgattacga cgttgtgact cgatatttca ataccaaatc cgtggaaaac 300
gttataggag atgttatcaa gaacactcaa acccacccta ctatcgtcat caagtccaca 360
attcccatcg gtttcggtga taaggtcaga gagcagtttg attatcaaaa cattatcttc 420
tcacctgagt tcttaaggga gggtcgtgct ctctacgata atttgtatcc gtcccgtatt 480
atcgttggcg acgattctcc tategctetc aagttcgcaa atctcttagt tgagggtagt 540
aagaccctt tggctcctgt tttgacaatg ggaaccagag aagcagaagc tgtcaagcta 600

ES 2 562 211 T3

ttctctaata cctaccttgc catgagggta gcatacttta acgaacttga tacatttgct 660
atgtcgcgatg gtatgaatgc caaggagatt atagatgggtg tcaactttaga gcccaggatc 720
ggtcaaggat attctaacce atcattcggc tatggagctt actgctttcc taaggacact 780
aagcagttgc tggcaaactt cgagggagtt cctcaagaca tcataggcgc tattgtggag 840
tcaaacgaaa caaggaaaga ggtgatagtt agtgaggtag agaatcgttt cccaacgaca 900
gtcggtgttt acaaactggc agctaaagct ggtagcgata acttcaggtc aagtgctatt 960
gtcgacatca tggaaacgct ggctaacaaa ggttaccaca ttaagatctt tgagccaact 1020
gtagagcagt tcgaaaattt cgaagttgac aataacttga caacgtttgc tactgagtca 1080
gacgttatta tcgcaaatcg tgtccctgtg gaacatagaa tcctatttgg aaagaagctc 1140
attaccagag atgtttacgg tgataattaa 1170

5 <210> 14
 <211> 48
 <212> ADN
 <213> Artificial

 10 <220>
 <223> oligonucleótido sintético

 <400> 14
 tcgacaggcc tggatccta attaaactag tctcaggag ctcggtac 48

 15 <210> 15
 <211> 40
 <212> ADN
 <213> Artificial

 20 <220>
 <223> oligonucleótido sintético

 <400> 15
 cgagctctc gagactagt taattaagga tccaggcctg 40

 25 <210> 16
 <211> 38
 <212> ADN
 <213> Artificial

 30 <220>
 <223> oligonucleótido sintético

 <400> 16
 aaaaactagt tctacatcgg cftaggtgta gcaacacg 38

 35 <210> 17
 <211> 39
 <212> ADN
 <213> Artificial

 40 <220>

ES 2 562 211 T3

<223> oligonucleótido sintético

<400> 17

aaaagatgc tgggtgga ttctactact atgctcaa 39

5

REIVINDICACIONES

1. Una célula vegetal o una planta que contiene una molécula de ácido nucleico exógena integrada de manera estable en su genoma, en la que la molécula de ácido nucleico codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT), en la que la molécula de ácido nucleico exógena codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa (GFAT bacteriana) y en la que la célula vegetal o la planta tiene un contenido de derivados de glucosamina N-acetilada de al menos 2 μmol por gramo de peso fresco, en donde la molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o que codifica una proteína que tiene la actividad de una GFAT bacteriana se selecciona del grupo que consiste en
- a) moléculas de ácido nucleico que codifican una proteína que tiene la secuencia de aminoácidos proporcionada como SEC ID N° 7 (GFAT-2) o una proteína que tiene la secuencia de aminoácidos proporcionada como SEC ID N° 9 (GFAT bacteriana);
- b) moléculas de ácido nucleico que codifican una proteína cuya secuencia es al menos un 60% idéntica a la secuencia de aminoácidos mostrada como SEC ID N° 7 (GFAT-2) o como SEC ID N° 9 (GFAT bacteriana);
- c) moléculas de ácido nucleico que comprenden la secuencia de nucleótidos de SEC ID N° 6 (GFAT-2) o de SEC ID N° 8 (GFAT bacteriana) o como SEC ID N° 10 (GFAT bacteriana) o una secuencia complementaria a la misma;
- d) moléculas de ácido nucleico que son al menos un 60% idénticas a las secuencias de ácido nucleico mostradas como a) o c);
- e) moléculas de ácido nucleico que hibridan en condiciones rigurosas con al menos una hebra de las secuencias de ácido nucleico descritas como a) o c);
- f) moléculas de ácido nucleico cuya secuencia de nucleótidos difiere de la secuencia de las moléculas de ácido nucleico mencionadas como a) o c) debido a la degeneración del código genético y
- g) moléculas de ácido nucleico que son fragmentos, variantes alélicas y/o derivados de las moléculas de ácido nucleico mencionadas en a), b), c), d), e) o f).
2. Una parte de una planta que contiene células vegetales de acuerdo con la reivindicación 1 que tiene un contenido de derivados de glucosamina N-acetilada de al menos 2 μmol por gramo de peso fresco.
3. Material de propagación de plantas según la reivindicación 1 que contiene células vegetales de acuerdo con la reivindicación 1, que tiene un contenido de derivados de glucosamina N-acetilados de al menos 2 μmol por gramo de peso fresco.
4. Un procedimiento de producción de una planta modificada genéticamente que comprende las siguientes etapas:
- a) introducción de una molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o que codifica una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa bacteriana en una célula vegetal donde la molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o que codifica una proteína que tiene la actividad de una GFAT bacteriana se selecciona del grupo que consiste en:
- i) moléculas de ácido nucleico que codifican una proteína que tiene la secuencia de aminoácidos proporcionadas como SEC ID N° 7 (GFAT-2) o una proteína que tiene la secuencia de aminoácidos proporcionada como SEC ID N° 9 (GFAT bacteriana);
- ii) moléculas de ácido nucleico que codifican una proteína cuya secuencia es al menos un 60 % idéntica a la secuencia de aminoácidos mostrada como SEC ID N° 7 (GFAT-2) o como SEC ID N° 9 (GFAT bacteriana);
- iii) moléculas de ácido nucleico que comprenden la secuencia nucleotídica mostrada como SEC ID N° 6 (GFAT-2) o como SEC ID N° 8 (GFAT bacteriana) o como SEC ID N° 10 (GFAT bacteriana) o una secuencia complementaria a esta;
- iv) moléculas de ácido nucleico que son al menos un 60% idénticas a las secuencias de ácido nucleico mostradas como a) o c);
- v) moléculas de ácido nucleico que hibridan en condiciones rigurosas con al menos una hebra de las secuencias de ácido nucleico descritas en a) o c);
- vi) moléculas de ácido nucleico cuya secuencia nucleotídica difiere de la secuencia de las moléculas de ácido nucleico mencionadas en a) o c) debido a la degeneración del código genético y
- vii) moléculas de ácido nucleico que son fragmentos, variantes alélicas y/o derivados de las moléculas de ácido nucleico mencionadas en i), ii), iii), iv), v) o vi);
- b) regeneración de una planta a partir de células vegetales obtenidas de acuerdo con la etapa a) en la que la planta tiene un contenido de derivados de glucosamina N-acetilada de al menos 2 μmol por gramo de peso fresco.
5. El uso de moléculas de ácido nucleico que codifican una proteína que tiene la actividad de una glutamina:fructosa 6-fosfato amidotransferasa de isoforma II (GFAT-2) o que codifica una proteína que tiene la actividad de una

glutamina:fructosa 6-fosfato amidotransferasa bacteriana (GFAT bacteriana), en la que la molécula de ácido nucleico exógena que codifica una proteína que tiene la actividad de una GFAT-2 o que codifica una proteína que tiene la actividad de una GFAT bacteriana se selecciona del grupo que consiste en:

- 5 a) moléculas de ácido nucleico que codifican una proteína que tiene la secuencia de aminoácidos proporcionada como SEC ID N° 7 (GFAT-2) o una proteína que tiene la secuencia de aminoácidos proporcionada como SEC ID N° 9 (GFAT bacteriana);
- b) moléculas de ácido nucleico que codifican una proteína cuya secuencia es al menos un 60% idéntica a la secuencia de aminoácidos mostrada como SEC ID N° 7 (GFAT-2) o como SEC ID N° 9 (GFAT bacteriana);
- 10 c) moléculas de ácido nucleico que comprenden la secuencia nucleotídica mostrada como SEC ID N° 6 (GFAT-2) o como SEC ID N° 8 (GFAT bacteriana) o como SEC ID N° 10 (GFAT bacteriana) o una secuencia complementaria a las mismas;
- d) moléculas de ácido nucleico que son al menos un 60% idénticas a las secuencias de ácido nucleico mostradas como a) o c);
- 15 e) moléculas de ácido nucleico que hibridan en condiciones rigurosas con al menos una hebra de las secuencias de ácido nucleico descritas en a) o c);
- f) moléculas de ácido nucleico cuya secuencia nucleotídica difiere de la secuencia de las moléculas de ácido nucleico mencionadas en a) o c) debido a la degeneración del código genético y
- g) moléculas de ácido nucleico que son fragmentos, variantes alélicas y/o derivados de las moléculas de ácido nucleico mencionadas en a), b), c), d), e) o f);

20 para la preparación de una planta modificada genéticamente.

6. Una composición que comprende células vegetales modificadas genéticamente según la reivindicación 1.

7. Un procedimiento de producción de harina que comprende la etapa de moler partes de planta según la reivindicación 2 o material de propagación según la reivindicación 3.

25 8. El uso de plantas según la reivindicación 1, de partes de plantas según la reivindicación 2 o de material de propagación según la reivindicación 3 para producir harina.

9. EL procedimiento de la reivindicación 4 que comprende una etapa adicional que consiste en la generación de plantas adicionales llevado a cabo mediante propagación vegetativa.