

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 575 912**

51 Int. Cl.:

C12N 9/32	(2006.01)
C12P 21/06	(2006.01)
C12N 15/00	(2006.01)
C12Q 1/68	(2006.01)
C12N 1/20	(2006.01)
C12N 1/00	(2006.01)
C12N 15/86	(2006.01)
C12P 19/04	(2006.01)
C12P 21/08	(2006.01)
C07H 21/04	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.12.2007 E 12152656 (0)**

97 Fecha y número de publicación de la concesión europea: **09.03.2016 EP 2479266**

54 Título: **Amilasas y glucoamilasas, ácidos nucleicos que las codifican y métodos para formarlas y utilizarlas**

30 Prioridad:

21.12.2006 US 877068 P
02.03.2007 US 892823 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
04.07.2016

73 Titular/es:

BASF ENZYMES LLC (50.0%)
3550 John Hopkins Court
San Diego, CA 92121, US y
SYNGENTA PARTICIPATIONS AG (50.0%)

72 Inventor/es:

SLUPSKA, MALGORZATA;
ZHANG, YAN;
HAZLEWOOD, GEOFF;
CHANG, CATHY;
LUGINBUHL, PETER;
BURKE, ELLEN;
CAYOQUETTE, MICHELLE;
GUNAVARDENA, UVINI;
GHASSEMIAN, MAJID y
SILVERSTONE, ARON

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 575 912 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Amilasas y glucoamilasas, ácidos nucleicos que las codifican y métodos para formarlas y utilizarlas

Campo técnico

5 Esta invención se relaciona con biología molecular y celular y bioquímica. En un aspecto, la invención está dirigida a polipéptidos que tienen una actividad de amilasa, polinucleótidos que codifican los polipéptidos, y métodos para elaborar y usar estos polinucleótidos y polipéptidos. En un aspecto, los polipéptidos de la invención se pueden usar como amilasas alfa, por ejemplo, para catalizar la hidrólisis de los polisacáridos que comprenden monómeros de glucosa, tales como almidón (un polímero de monómeros de glucosa unidos por enlaces 1,4-alfa o 1,6-alfa), en azúcares. En un aspecto, la invención está dirigida a polipéptidos que tienen actividad termoestable de amilasa, 10 incluyendo actividad de alfa amilasas. En un aspecto, los polipéptidos de la invención se pueden usar como alfa amilasas para catalizar la hidrólisis de los polisacáridos tales como almidón en azúcares, tales como glucosa. La invención también está dirigida a constructos de ácido nucleico, vectores y células huésped que comprenden las secuencias de ácido nucleico de la invención así como también métodos recombinantes para la producción de los polipéptidos de la invención. La invención también está dirigida al uso de amilasas de la invención en procesos de 15 conversión de polisacáridos (por ejemplo, almidón), incluyendo la producción de jarabe de maíz con alto contenido de fructosa (HFCS), etanol, dextrosa y jarabes de dextrosa.

Antecedentes

20 El almidón es un carbohidrato complejo usualmente encontrado en la dieta humana. La estructura del almidón es de polímeros de glucosa enlazados mediante enlaces glucosídicos α -1,4 y β -1,6. A nivel comercial, se usan las glucoamilasas para hidrolizar adicionalmente el almidón de maíz, que ya ha sido parcialmente hidrolizado con una alfa-amilasa. La glucoamilasa más ampliamente utilizada es producida a partir el hongo *Aspergillus niger*, uno de los problemas con el uso comercial de esta enzima es su termoestabilidad relativamente baja.

25 En general, el procesamiento del almidón hasta fructosa consiste de cuatro etapas: licuefacción del almidón granulado, sacarificación del almidón licuado en dextrosa, purificación e isomerización hasta fructosa. El objetivo de un proceso de licuefacción del almidón es convertir una suspensión concentrada de gránulos de polímero de almidón en una solución de dextrinas solubles con longitud de cadena más corta de baja viscosidad. Esta etapa es esencial para el manejo conveniente con equipamiento estándar y para la conversión eficiente hasta glucosa u otros azúcares. Para licuar el almidón granulado, es necesario gelatinizar los gránulos elevando la temperatura del almidón granulado aproximadamente por encima de 72°C. El proceso de calentamiento rompe instantáneamente los 30 gránulos de almidón insolubles para producir una solución de almidón soluble en agua. La solución de almidón solubilizada es luego licuada por la amilasa. Un gránulo de almidón está compuesto de: 69-74% de amilopectina, 26-31% de amilosa, 11-14% de agua, 0,2-0,4% de proteína, 0,5-0,9% de lípido, 0,05-0,1% de ceniza. 0.02-0,03% de fósforo, 0,1% de pentosano. Aproximadamente el 70% de un gránulo es amorfo y el 30% es cristalino.

35 El enranciamiento de los productos horneados (tales como pan) ha sido reconocido como un problema que se torna más serio a medida que transcurre más tiempo entre el momento de la preparación del producto panificado y el momento de consumo. El término enranciamiento se utiliza para describir cambios indeseables para el consumidor en las propiedades del producto panificado después de salir del horno, tales como incremento en la firmeza de la miga, una reducción de la elasticidad de la miga y cambios en la corteza, que se torna dura y curtida. La firmeza de la miga del pan se incrementa de manera adicional durante el almacenamiento hasta alcanzar un nivel que se 40 considera negativo. El incremento en la firmeza de la miga, que es considerado el aspecto más importante del enranciamiento, es reconocido por el consumidor mucho antes de que el producto panificado se haya tornado inadecuado para el consumo.

45 Existe la necesidad en la industria de nuevas amilasas, por ejemplo, amilasas ácidas, útiles para varios usos incluyendo procesos comerciales de licuefacción de almidón de maíz o fabricación mejorada que tenga características nuevas o mejoradas de desempeño con respecto a las enzimas estándar de la industria, por ejemplo, de *Bacillus licheniformis*. También existe la necesidad en la industria de identificar las amilasas y glucoamilasas capaces de hidrolizar de manera eficiente el almidón granulado (por ejemplo almidón granulado sin purificar) a bajas temperaturas sin la necesidad de un etapa de gelatinización del almidón a alta temperatura; la enzimas de la invención, por ejemplo, amilasas, se pueden utilizar para satisfacer esta necesidad.

50 También existe la necesidad de nuevas amilasas que tengan utilidad en productos para lavadoras automáticas de vajillas (ADW) y detergente para lavandería. En los productos para ADW, la amilasa actuará a pH 10-11 y a 45-60 °C en presencia de agentes quelantes de calcio y condiciones oxidativas. Para lavandería, se requerirá actividad a pH 9-10 y 40°C en la matriz apropiada de detergente. Las amilasas también son útiles en el desencolado textil, procesos de elaboración de cerveza, modificación de almidón en la industria de pulpa y papel y otros procesos 55 descritos en el arte.

Las entradas de base de datos y Q0C881 EAU29552 muestran la secuencia de aminoácidos y la secuencia de ácido nucleico, respectivamente, de una alfa-amilasa de *Aspergillus terreus*

Resumen

5 La invención provee ácidos nucleicos aislados, sintéticos o recombinantes que comprenden una secuencia de ácido nucleico que tiene al menos 98%, 99%, o más, o identidad completa de secuencia (100%) con un ácido nucleico de la invención, por ejemplo, un ejemplo de ácido nucleico de la invención. En un aspecto, el ácido nucleico codifica al menos un polipéptido que tiene actividad de amilasa y las identidades de secuencia se determinan mediante análisis con un algoritmo de comparación de secuencias o por inspección visual. En otro aspecto, la divulgación provee ácidos nucleicos para uso como sondas, moléculas inhibitoras (por ejemplo, ARNi antisentido, tales como siARN, microARN o miARN), regulación de transcripción o de traducción y similares.

Los ejemplos de ácidos nucleicos de la invención incluyen ácidos nucleicos aislados, sintéticos o recombinantes, que comprenden una secuencia de ácido nucleico como la expuesta en la SEQ ID NO: 51.

15 Los ejemplos de ácidos nucleicos de la invención también incluyen ácidos nucleicos aislados, sintéticos o recombinantes, que codifican un polipéptido de la invención, por ejemplo, un ejemplo de polipéptido que tiene una secuencia como la expuesta en la SEQ ID NO: 52.

En un aspecto, el polipéptido de la invención tiene actividad de alfa amilasa. En una realización, los polipéptidos de la invención pueden catalizar la hidrólisis de polisacáridos y/o oligosacáridos que comprenden monómeros de glucosa, tales como almidón (un polímero de monómeros de glucosa unido por enlaces 1,4-alfa o 1,6-alfa).

20 Las amilasas de la invención se pueden usar a nivel comercial en las etapas iniciales (licuefacción) del procesamiento de polisacárido, oligosacárido o almidón; en la molienda húmeda de maíz; en la producción de alcohol; como agentes de limpieza en matrices de detergentes; en la industria textil para descolado del almidón; en aplicaciones de panadería y pastelería; en la industria de las bebidas; en yacimientos petrolíferos en procesos de perforación; en el entintado de papel reciclado y en alimentación animal. Las amilasas de la invención se pueden usar en descolado textil, en procesos de elaboración de cerveza, en la modificación de polisacáridos, oligosacáridos o almidón en la industria de pulpa y papel y en otros procesos. Por ejemplo, la invención provee métodos para sacarificación por licuefacción como se ilustra en la Figura 5 usando polipéptidos de la invención.

30 Las amilasas de la invención se pueden usar para catalizar la hidrólisis de polisacáridos, por ejemplo, almidones, en azúcares; o para hidrolizar enlaces -1,4-glucosídicos internos en almidón para producir maltodextrinas de menor peso molecular. Debido al rompimiento de los polisacáridos y/o oligosacáridos, por ejemplo, almidones, es importante en el sistema digestivo y en procesos de preparación comercial, las amilasas de la invención se usan en alimentos y piensos y en procesos para su preparación, y en o como auxiliares digestivos. Las amilasas de la invención se pueden usar en las etapas iniciales (licuefacción) del procesamiento de almidón; en la molienda húmeda de maíz; en la producción de alcohol; como agentes de limpieza en matrices de detergentes; en la industria textil para descolado del almidón; en aplicaciones de panadería y pastelería; en la industria de las bebidas; en yacimientos petrolíferos en procesos de perforación; en el entintado de papel reciclado; y en la alimentación animal.

40 Las enzimas de la divulgación pueden tener actividad de glucoamilasa que actúa en forma exo y se pueden usar para hidrolizar adicionalmente el almidón de maíz que ya ha sido parcialmente hidrolizado con una alfa-amilasa (que es un polipéptido de la invención) para producir glucosa; y en un aspecto de este proceso de la divulgación, la glucosa es convertida en una mezcla de glucosa y fructosa por medio de una enzima de glucosa isomerasa. En otro aspecto de este proceso de la divulgación, se enriquece esta mezcla con fructosa para producir un jarabe de maíz con alto contenido de fructosa. En aspectos alternativos, se usan los polipéptidos de la invención en cualquier etapa del procesamiento de polisacáridos, oligosacáridos o almidón hasta fructosa, por ejemplo, incluyendo las cuatro etapas: licuefacción de almidón granulado, sacarificación del polisacárido, oligosacárido o almidón licuado en dextrosa, purificación e isomerización hasta fructosa. Un aspecto de la invención que usa al menos un polipéptido de la invención comprende un proceso de licuefacción del polisacárido, oligosacárido o almidón para convertir una suspensión concentrada de gránulos de polímero de polisacárido, oligosacárido o almidón en una solución de dextrinas solubles de longitud de cadena más corta de baja viscosidad.

50 La invención también provee un proceso de licuefacción enzimática usando al menos un polipéptido de la invención que comprende ajustar el pH de una suspensión granulada de polisacárido, oligosacárido o almidón hasta el pH óptimo de una enzima de la invención (por ejemplo, una amilasa de la invención) que va a ser utilizada, por ejemplo, entre 6,0 y 6,5, o entre aproximadamente 5,5 y 7,0; se puede agregar hidróxido de calcio, hidróxido de sodio o carbonato de sodio para este propósito (la adición de hidróxido de calcio tiene la ventaja de proveer también iones de calcio que son conocidos por estabilizar la alfa-amilasa contra la inactivación). En un aspecto, después de la adición de la amilasa (por ejemplo, una alfa-amilasa) de la invención, la suspensión se bombea a través de un chorro de vapor para elevar instantáneamente la temperatura entre 80 °C y 115 °C y el almidón es gelatinizado en

forma inmediata y, debido a la presencia de la alfa-amilasa, despolimerizado a través de hidrólisis aleatoria del enlace (1-4) glucosídico por la alfa-amilasa hasta una masa de fluido que es fácilmente bombeada.

En aspectos alternativos a este proceso de licuefacción de polisacáridos y/o oligosacáridos (por ejemplo, almidón), se añade una amilasa de la invención a la suspensión del polisacárido (por ejemplo, almidón), se mantiene la suspensión a una temperatura de 80-100°C para hidrolizar parcialmente los gránulos, por ejemplo, gránulos de almidón y se bombea la suspensión de polisacárido/almidón parcialmente hidrolizada a través de un chorro a temperaturas que exceden aproximadamente 105°C para gelatinizar completamente cualquier estructura granulada remanente. Después de enfriar el polisacárido/almidón gelatinizado, se puede realizar una segunda adición de una amilasa de la invención (por ejemplo, una alfa-amilasa) para hidrolizar de manera adicional el polisacárido/almidón.

En un aspecto, la divulgación también provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa con la novedad en común de que son derivados de cultivos mixtos, por ejemplo, de fuentes ambientales. La invención provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa aislados de cultivos mixtos que comprenden una secuencia de ácido nucleico que tiene al menos aproximadamente 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, o identidad completa de secuencia (100%) con un ejemplo de ácido nucleico de la invención sobre una región de al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 o más, residuos, en donde el ácido nucleico codifica al menos un polipéptido que tiene actividad de amilasa y/o glucoamilasa, y las identidades de secuencia se determinan mediante análisis con un algoritmo de comparación de secuencias o por inspección visual.

En un aspecto, la divulgación provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa aislados de cultivos mixtos, por ejemplo, de fuentes ambientales, que comprenden un ácido nucleico de la invención, por ejemplo, un ejemplo de ácido nucleico de la invención, por ejemplo, una secuencia como la expuesta en la SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, etc., y subsecuencias de las mismas, por ejemplo, al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 o más residuos de longitud, o sobre la longitud completa de un gen o transcritor; o, un ácido nucleico que codifica un polipéptido de la divulgación.

En un aspecto, la divulgación también provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa con la novedad en común de que se derivan de fuentes ambientales (véase la Tabla 1 a continuación, segunda columna para los ejemplos de las secuencias aisladas de fuentes "desconocidas" o ambientales), por ejemplo, fuentes ambientales mixtas. En un aspecto, la divulgación provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa aislados de fuentes ambientales, por ejemplo, fuentes ambientales mixtas, que comprenden una secuencia de ácido nucleico que tiene al menos aproximadamente 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, o identidad completa de secuencia (100%) con un ejemplo de ácido nucleico de la invención sobre una región de al menos aproximadamente 25, 30, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550 o más residuos, en donde el ácido nucleico codifica al menos un polipéptido que tiene actividad de amilasa y/o glucoamilasa y las identidades de secuencia se determinan mediante análisis con un algoritmo de comparación de secuencias o por inspección visual.

En un aspecto, la invención provee ácidos nucleicos que codifican amilasa y que codifican glucoamilasa aislados de fuentes ambientales, por ejemplo, fuentes ambientales mixtas, que comprenden un ácido nucleico de la invención, por ejemplo, un ejemplo de una secuencia de ácido nucleico de la invención como se expone en la SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 9, SEQ ID NO: 11, etc., la SEQ ID NO: 583, SEQ ID NO: 585 y subsecuencias de las mismas, por ejemplo, al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 205, 210, 215, 220, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305, 310, 315, 320, 325, 330, 335, 340, 345, 350, 355, 360, 365, 370, 375, 380, 385, 390, 395, 400, 405, 410, 415, 420, 425, 430, 435, 440, 445, 450, 455, 460, 465, 470, 475, 480, 485, 490, 495, 500, 505, 510, 515, 520, 525, 530, 535, 540, 545, 550, 555, 560, 565, 570, 575, 580, 585, 590, 595, 600, 605, 610, 615, 620, 625, 630, 635, 640, 645, 650, 655, 660, 665, 670, 675, 680, 685, 690, 695, 700, 705, 710, 715, 720, 725, 730, 735, 740, 745, 750, 755, 760, 765, 770, 775, 780, 785, 790, 795, 800, 805, 810, 815, 820, 825, 830, 835, 840, 845, 850, 855, 860, 865, 870, 875, 880, 885, 890, 895, 900, 905, 910, 915, 920, 925, 930, 935, 940, 945, 950, 955, 960, 965, 970, 975, 980, 985, 990, 995, o más, o identidad completa de secuencia (100%) con un ejemplo de ácido nucleico de la invención sobre una región de al menos aproximadamente 25, 30, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 o más residuos de longitud, o sobre la longitud completa de un gen o transcritor; o un ácido nucleico que codifica un polipéptido de la divulgación.

En un aspecto, la divulgación también provee glucoamilasas y amilasas, y ácidos nucleicos que codifican amilasa y que codifican glucoamilasa, con la novedad en común de que se derivan de fuentes de arqueas, incluyendo las amilasas y/o glucoamilasas derivadas de arqueas de la invención.

En un aspecto, el algoritmo de comparación de secuencias es un algoritmo BLAST versión 2.2.2 donde se establece una configuración de filtración para blast II -p blast p-d "nr pataa" -FF, y todas las otras opciones se establecen en forma predeterminada.

Otro aspecto de la divulgación es un ácido nucleico aislado, sintético o recombinante, que incluye al menos 10 bases consecutivas de una secuencia de ácido nucleico de la invención, secuencias sustancialmente idénticas a la misma, y las secuencias complementarias de la misma.

5 En un aspecto, la actividad de la amilasa comprende una actividad de α -amilasa o de β -amilasa, incluyendo la capacidad para hidrolizar enlaces alfa-1,4-glucosídicos internos en almidón para producir maltodextrinas de menor peso molecular. En un aspecto, la actividad de α -amilasa incluye hidrolizar enlaces alfa-1,4-glucosídicos internos en almidón al azar. La actividad de la glucoamilasa y/o amilasa puede comprender una actividad de α -amilasa, una actividad de β -amilasa, una actividad de 1,4- α -D-glucano glucohidrolasa, una actividad de exoamilasa, una actividad de glucano α -maltotetrahidrolasa, una actividad de maltasa, una actividad de isomaltasa, una actividad de glucano 1,4, α -glucosidasa, una actividad de α -glucosidasa, una actividad de sacarosa o una actividad de agarasa (por ejemplo, una actividad β -agarosa).

15 La actividad de amilasa puede comprender la hidrólisis de enlaces glucosídicos. En un aspecto, los enlaces glucosídicos comprenden un enlace α -1,4-glucosídico. En otro aspecto, los enlaces glucosídicos comprenden un enlace α -1,6-glucosídico. En un aspecto, la actividad de amilasa comprende la hidrólisis de enlaces glucosídicos en polisacáridos, por ejemplo, almidones, por ejemplo, almidón licuado. La actividad de amilasa puede comprender además la hidrólisis de enlaces glucosídicos en maltodextrinas. En un aspecto, la actividad de amilasa comprende la escisión de una unidad de maltosa o D-glucosa de un extremo no reductor del almidón.

20 En un aspecto, el ácido nucleico aislado, sintético o recombinante, codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa que es termoestable. El polipéptido puede retener una actividad de amilasa bajo condiciones que comprenden un rango de temperatura cualquiera entre aproximadamente 0°C hasta aproximadamente 37 °C, o, entre aproximadamente 37 °C hasta aproximadamente 95 °C o más, por ejemplo, 98 °C, 100 °C o más; entre aproximadamente 55 °C hasta aproximadamente 85 °C, entre aproximadamente 70 °C hasta aproximadamente 95 °C, o, entre aproximadamente 90 °C hasta aproximadamente 95 °C. Por ejemplo, el ejemplo de polipéptido que tiene una secuencia como se expone en la SEQ ID NO: 437 es termoestable, reteniendo 50% de actividad después de 25 minutos a 100 °C en ausencia de calcio agregado.

30 En otro aspecto, el ácido nucleico aislado, sintético o recombinante, codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa que es termotolerante. El polipéptido puede retener una actividad de amilasa y/o glucoamilasa después de exposición a una temperatura en el rango de más de 37 °C hasta aproximadamente 95 °C o cualquiera en el rango de más de 55 °C hasta aproximadamente 85 °C. En un aspecto, el polipéptido retiene una actividad de amilasa y/o glucoamilasa después de exposición a una temperatura en el rango de más de 90 °C hasta aproximadamente 95 °C a un pH de 4,5.

35 La divulgación provee ácidos nucleicos aislados, sintéticos o recombinantes, que comprenden una secuencia que híbrida bajo condiciones rigurosas con un ácido nucleico de la invención, por ejemplo, un ejemplo de un ácido nucleico de la invención, un ácido nucleico que comprende una secuencia como se expone en la SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5. SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO: 25, SEQ ID NO: 27, SEQ ID NO: 29, SEQ ID NO: 31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37, SEQ ID NO: 39, SEQ ID NO: 41, SEQ ID NO: 43, SEQ ID NO: 45, SEQ ID NO: 47, SEQ ID NO: 49, SEQ ID NO: 51, SEQ ID NO: 53, SEQ ID NO: 55, SEQ ID NO: 57, SEQ ID NO: 59, SEQ ID NO: 61, SEQ ID NO: 63, SEQ ID NO: 65, SEQ ID NO: 67, SEQ ID NO: 69, SEQ ID NO: 71, SEQ ID NO: 73, SEQ ID NO: 75, SEQ ID NO: 77, SEQ ID NO: 79, SEQ ID NO: 80, SEQ ID NO: 81 y/o SEQ ID NO: 82, y/o fragmentos o subsecuencias de las mismas. En un aspecto, el ácido nucleico codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa. El ácido nucleico puede ser de al menos aproximadamente 25, 30, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500 o más residuos de longitud o la longitud completa del gen o del transcripto. En un aspecto, las condiciones rigurosas incluyen una etapa de lavado que comprende un lavado en 0,2X SSC a una temperatura de aproximadamente 65 °C durante aproximadamente 15 minutos. En un aspecto, las condiciones rigurosas comprenden hibridación bajo condiciones que comprenden un regulador que tiene NaCl 0,15 M durante 15 minutos a 72 °C.

50 La invención provee una sonda de ácido nucleico para la identificación de un ácido nucleico que codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa, en donde la sonda comprende al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 o más, bases consecutivas de una secuencia que comprende una secuencia de la invención, o fragmentos o subsecuencias de la misma, en donde la sonda identifica al ácido nucleico mediante enlazamiento o hibridación. La sonda puede comprender un oligonucleótido que comprende al menos aproximadamente 10 a 50, aproximadamente 20 a 60, aproximadamente 30 a 70, aproximadamente 40 a 80, o aproximadamente 60 a 100 bases consecutivas de una secuencia que comprende una secuencia de la invención, o fragmentos o subsecuencias de la misma.

- La invención provee una sonda de ácido nucleico para la identificación de un ácido nucleico que codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa, en donde la sonda comprende un ácido nucleico que comprende una secuencia de al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 o más residuos que tiene al menos aproximadamente 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más, o identidad de secuencia completa (100%) con un ácido nucleico de la invención, en donde las identidades de secuencia se determinan mediante análisis con un algoritmo de comparación de secuencias o por inspección visual.
- 10 La sonda puede comprender un oligonucleótido que comprende al menos aproximadamente 10 a 50, aproximadamente 20 a 60, aproximadamente 30 a 70, aproximadamente 40 a 80, o aproximadamente 60 a 100 bases consecutivas de una secuencia de ácido nucleico de la invención, o una subsecuencia de la misma.
- La invención provee un par de secuencias de cebadores de amplificación para amplificación de un ácido nucleico que codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa, en donde el par cebador es capaz de amplificar un ácido nucleico que comprende una secuencia de la invención, o fragmentos o subsecuencias de las mismas. Uno de cada elemento del par de amplificación de la secuencia cebadora puede comprender un oligonucleótido que comprende al menos aproximadamente 10 a 50 bases consecutivas de la secuencia.
- 20 La invención provee métodos para amplificación de un ácido nucleico que codifica un polipéptido que tiene una actividad de amilasa y/o glucoamilasa que comprende la amplificación de un ácido nucleico plantilla con un par de amplificación de la secuencia cebadora capaz de amplificar una secuencia de ácido nucleico de la invención, o fragmentos o subsecuencias de la misma.
- La invención provee casetes de expresión (que incluyen, por ejemplo, vectores y vehículos de clonación) que comprenden un ácido nucleico de la invención o una subsecuencia del mismo. En un aspecto, el casete de expresión puede comprender al ácido nucleico que está operativamente enlazado a un promotor. El promotor puede ser un promotor viral, bacteriano, de mamífero o de una planta. En un aspecto, el promotor de la planta puede ser un promotor de patata, arroz, maíz, trigo, tabaco o cebada. El promotor puede ser un promotor constitutivo. El promotor constitutivo puede comprender CaMV35S. En otro aspecto, el promotor puede ser un promotor inducible. En un aspecto, el promotor puede ser un promotor específico de tejido o un promotor regulado por el medio ambiente o uno regulado por desarrollo. Por lo tanto, el promotor puede ser, por ejemplo, un promotor específico de semilla, específico de hoja, específico de raíz, específico del tallo o inducido por abscisión. En un aspecto, el casete de expresión puede comprender, además, un vector de expresión de una planta o de un virus de una planta. En las formas de realización alternativas, el promotor de la planta es un promotor de globulina específico del embrión de la semilla de maíz (véase, por ejemplo, Belanger (1991) Molecular basis for allelic polymorphism of the maize Globulin-1 gene. Genetics 129: 863 - 872); o un promotor de γ -zeína específico del endospermo de la semilla de maíz: (véase, por ejemplo, Lopes (1995) Identification of two opaque2 modifier loci in Quality Protein Maize. Mol. Gen. Genet. 247: 603 - 613); o un promotor GTL1 específico del endospermo de la semilla de arroz (véase, por ejemplo, Takaiwa (1991) Analysis of the 5' flanking region responsible for the endosperm-specific expression of a rice glutelin chimeric gene in transgenic tobacco. Plant Mol. Biol. 16: 49 - 58).
- 40 La invención provee vehículos de clonación que comprenden un casete de expresión (por ejemplo, un vector) de la invención o un ácido nucleico de la invención. El vehículo de clonación puede ser un vector viral, un plásmido, un fago, un fagémido, un cósmido, un fósmido, un bacteriófago o un cromosoma artificial. El vector viral puede comprender un vector de adenovirus, un vector retroviral o un vector viral adenoasociado. El vehículo de clonación puede comprender un cromosoma artificial bacteriano (BAC), un plásmido, un vector derivado del bacteriófago P1 (PAC), un cromosoma artificial de levadura (YAC), o un cromosoma artificial de mamífero (MAC).
- 45 La invención provee células transformadas (o "células huésped") que comprenden un ácido nucleico de la invención o un casete de expresión (por ejemplo, un vector) de la invención, o un vehículo de clonación de la invención. En un aspecto, la célula transformada puede ser una célula bacteriana, una célula fúngica, una célula de mamífero, una célula de levadura, una célula de insecto o una célula vegetal. En un aspecto, la célula vegetal puede ser de cualquier planta, por ejemplo plantas usadas para forraje y/o pienso para cualquier animal, incluyendo rumiantes, o como fuente de materia prima para producir energía o combustible. Las plantas de interés particular pueden incluir plantas de cultivo y plantas de materia prima, por ejemplo, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y *Miscanthus*, cultivos de leguminosas, por ejemplo, guisantes, frijol y soja, raíces/tubérculos almidonosos, por ejemplo, patata, batata, mandioca, taro, canna, remolacha azucarera, caña de azúcar y similares.
- 55 La divulgación provee animales no humanos transgénicos que comprenden un ácido nucleico de la divulgación o un casete de expresión (por ejemplo, un vector) de la invención. En un aspecto, el animal es un ratón. La invención provee células o líneas celulares aisladas de estos animales no humanos transgénicos.

La divulgación provee plantas transgénicas que comprenden un ácido nucleico de la divulgación o un casete de expresión (por ejemplo, un vector) de la divulgación. La planta transgénica puede ser cualquier planta, pero en una forma de realización la planta se sería utilizada para forraje y/o pienso para cualquier animal o como materia prima para producir energía o combustible, tal como, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, 5 sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y *Miscanthus*, cultivos de leguminosas, por ejemplo, guisante, frijol y soja, raíces/tubérculos almidonosos, por ejemplo, patata, batata, mandioca, taro, canna, remolacha azucarera, caña de azúcar y similares.

La divulgación provee semillas transgénicas que comprenden un ácido nucleico de la divulgación o un casete de expresión (por ejemplo, un vector) de la divulgación. La semilla transgénica puede ser de cualquier planta, pero en 10 una forma de realización la planta se usa para forraje y/o alimento para cualquier animal o como materia prima para producir energía o combustible, tal como, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y *Miscanthus*, cultivos de leguminosas, por ejemplo, guisante, frijol y soja, raíz/tubérculo almidonoso, por ejemplo, patata, batata, mandioca, taro, caña, remolacha azucarera, caña de azúcar y similares.

15 Cualquier planta, parte de la planta, tejido de la planta, semilla de la planta o célula vegetal puede ser usada para la introducción de un nucleótido de la divulgación, ya sea en forma estable (por ejemplo, como una planta transgénica, o célula o línea celular derivada de la misma) o transitoria; por lo tanto, la divulgación provee plantas, partes de plantas, tejidos de plantas, semillas de plantas y células vegetales que comprenden un ácido nucleico y/o polipéptido de la divulgación, en donde la planta puede ser (pero no se limita a) maíz (*Zea mays*), *Brassica* sp. (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), incluyendo aquellas especies de *Brassica* útiles como fuentes de aceite de semilla, tales como canola, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*), 20 *Sorghum vulgare*), mijo (por ejemplo, mijo perla (*Pennisetum glaucum*), mijo poroso (*Panicum miliaceum*), mijo cola de zorro (*Setaria italica*), mijo dedo (*Eleusine coracana*)), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), cacahuetes (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), batata (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea* spp.), coco (*Cocos nucifera*), piña (*Ananas comosus*), cítricos (*Citrus* spp.), cacao (*Theobroma cacao*), té (*Camellia sinensis*), banana (*Musa* spp.), aguacate (*Persea americana*), higo (*Ficus casica*), guayaba (*Psidium guajava*), mango (*Mangifera indica*), oliva (*Olea europaea*), papaya (*Carica papaya*), nuez de la India (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendra (*Prunus amygdalus*), remolacha azucarera (*Beta* sp., por ejemplo, *Beta vulgaris*), caña de azúcar (*Saccharum* spp.) 30 *andropogoneae* (pastos), *Chenopodiaceae* (plantas que dan flor), avena, cebada, vegetales, plantas ornamentales y coníferas; y vegetales, por ejemplo, tomates (*Lycopersicon esculentum*), lechuga (por ejemplo, *Lactuca sativa*), judías verdes (*Phaseolus vulgaris*), judías lima (*Phaseolus limensis*), guisantes (*Lathyrus* spp.) y miembros del género *Cucumis* tales como pepino (*C. sativus*), cantalupo (*C. cantalupensis*) y melón (*C. melo*); y plantas ornamentales, incluyendo azalea (*Rhododendron* spp.), hortensia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasansensis*), rosas (*Rosa* spp.), tulipanes (*Tulipa* spp.), narcisos (*Narcissus* spp.), petunias (*Petunia hybrida*), clavel (*Dianthus caryophyllus*), flor de Pascua (*Euphorbia pulcherrima*), canna (*Cannaceae* spp.) y crisantemo; y las coníferas que pueden ser utilizadas, incluyendo, por ejemplo, pinos tales como pino taeda (*Pinus taeda*), pino de incienso (*Pinus elliotii*), pino ponderosa (*Pinus ponderosa*), pino logdepole (*Pinus contorta*) y pino de Monterrey (*Pinus radiata*), abeto Douglas (*Pseudotsuga menziesii*); abeto occidental (*Tsuga canadensis*); picea Sitka (*Picea glauca*); secoya (*Sequoia sempervirens*); abetos reales tales como abeto plateado (*Abies amabilis*) y abeto balsámico (*Abies balsamea*); y cedros tales como cedro rojo occidental (*Thuja plicata*) y cedro amarillo de Alaska (*Chamaecyparis nootkatensis*); y plantas leguminosas, incluyendo, pero sin limitarse a, frijoles y guisantes; donde en 35 aspectos alternativos los frijoles pueden incluir guar, algarroba, fenogreco, soja, habichuelas, arvejas, frijol mungo, frijol lima, habas, lentejas, garbanzos, etc. y las legumbres pueden incluir, pero sin limitarse a, *Arachis*, por ejemplo, cacahuetes, Vicia, por ejemplo, veza corona, veza peluda, frijol adzuki, frijol mungo y garbanzos, lupines, por ejemplo, lupino, trébol rojo, *Phaseolus*, por ejemplo, frijol común y frijol lima, Pisum, por ejemplo, frijol de campo, meliloto, por ejemplo, trébol, Medicago, por ejemplo, alfalfa, loto, por ejemplo, trébol, lentejas, por ejemplo, lenteja y falso índigo; también se incluye pastos para forraje y pasto césped, tales como alfalfa, pasto varilla (*Panicum virgatum*), *Miscanthus*, pasto ovillo, festuca alta, ballico perenne, pasto encorvado rastrero y agróstide. 50

En aspectos alternativos, plantas, partes de plantas, tejidos de plantas, semillas de plantas y células vegetales que comprenden un ácido nucleico y/o un polipéptido de la divulgación también incluyen plantas de cultivo y plantas utilizadas para producir energía o combustible (por ejemplo, etanol u otros biocombustibles, por ejemplo, bioetanoles, biopropanoles, biobutanoles o biodiesel), por ejemplo, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y *Miscanthus*, cultivos de leguminosas, por ejemplo, guisante, frijol y soja, raíces/tubérculos almidonosos, por ejemplo, patata, batata, mandioca, taro, canna, caña de azúcar y/o remolacha azucarera y similares. 55

La invención provee un oligonucleótido antisentido que comprende una secuencia de ácido nucleico complementaria a o capaz de hibridar bajo condiciones rigurosas con un ácido nucleico de la invención. La invención provee métodos para inhibir la traducción de un mensaje de amilasa en una célula que comprende la administración a la célula o la 60

expresión en una célula de un oligonucleótido antisentido que comprende una secuencia de ácido nucleico complementaria a o capaz de hibridar bajo condiciones rigurosas con un ácido nucleico de la invención.

5 La invención provee un polipéptido aislado, sintético o recombinante, que comprende una secuencia de aminoácidos que tiene al menos 99%, o más, o identidad de secuencia completa (100%) con un ejemplo de polipéptido o péptido de la invención. Las secuencias de polipéptidos o péptidos de la invención incluyen la SEQ ID NO: 52 y/o subsecuencias de la misma y variantes de la misma, por ejemplo, en al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, o más residuos de longitud, o sobre la longitud completa de una enzima. Los ejemplos de secuencias de polipéptidos o péptidos de la invención incluyen una 10 secuencia codificada por el ácido nucleico de la invención. Los ejemplos de secuencias de polipéptidos o péptidos de la invención incluyen polipéptidos o péptidos específicamente unidos por un anticuerpo de la invención. En un aspecto, un polipéptido de la invención tiene al menos una actividad de amilasa, por ejemplo, una actividad de alfa amilasa.

15 En formas de realización alternativas, los polipéptidos de la divulgación carecen de una secuencia señal (secuencia líder) y/o una molécula que enlaza un carbohidrato. En formas de realización alternativas, los polipéptidos de la divulgación comprenden, además, una o más secuencias heterólogas, que pueden comprender una secuencia señal heteróloga (secuencia líder), un dominio catalítico heterólogo (CD) (es decir, sitio activo), o un módulo que enlaza a un carbohidrato heterólogo, o epítipo, etiqueta de purificación o rótulo. En un aspecto, la secuencia señal heteróloga, el módulo de enlazamiento del carbohidrato heterólogo o el dominio catalítico heterólogo (CD) se deriva de otra enzima amilasa (una amilasa diferente de una enzima de esta invención), o se deriva de una enzima que no es amilasa.

20 En formas de realización alternativas, los polipéptidos de la invención tienen una actividad de amilasa o se pueden usar para generar anticuerpos que se enlazan de manera específica con un ejemplo de polipéptido de la invención (SEQ ID NO: 52).

25 En formas de realización alternativas, los polipéptidos de la invención pueden ser sintéticos o en una forma péptido mimética.

Otro aspecto de la invención es un péptido o polipéptido aislado, sintético o recombinante, que incluye al menos 10, 15, 20, 25, 30, 35, 40, 45, 50, 75, 100, 125, 150, 175, 200, 225, 250, 300, 350, 400, 450 o 500 o más residuos de aminoácidos consecutivos de una secuencia de péptidos o polipéptidos de la invención, secuencias sustancialmente idénticas a la misma y las secuencias complementarias de la misma.

30 En un aspecto, la actividad de amilasa de un polipéptido o péptido de la invención comprende una actividad de α -amilasa, incluyendo la capacidad para hidrolizar enlaces alfa-1,4-glucosídicos internos en almidón para producir maltodextrinas de menor peso molecular. En un aspecto, la actividad de α -amilasa incluye la hidrólisis de enlaces alfa-1,4-glucosídicos internos en almidón al azar. La actividad de amilasa comprende una actividad de α -amilasa. La actividad de amilasa puede comprender la hidrólisis de enlaces glucosídicos. En un aspecto, los enlaces glucosídicos comprenden un enlace α -1,4-glucosídico. En otro aspecto, los enlaces glucosídicos comprenden un enlace α -1,6-glucosídico. En un aspecto, la actividad de amilasa comprende la hidrólisis de enlaces glucosídicos en almidón, por ejemplo, almidón licuado. La actividad de amilasa puede comprender, además, la hidrólisis de enlaces glucosídicos en maltodextrinas. En un aspecto, la actividad de amilasa comprende la escisión de una unidad de maltosa o de D-glucosa del extremo no reductor del almidón.

35 En un aspecto, la actividad de amilasa de la divulgación comprende una actividad de glucoamilasa, que puede comprender la catálisis de la hidrólisis de enlaces glucosídicos. La actividad de glucoamilasa de la divulgación puede comprender la catálisis de la liberación hidrolítica, por etapas, de la D-glucosa de los extremos no reductores del almidón u otras dextrinas relacionadas. La actividad de la glucoamilasa puede comprender una de actividad 1,4- α -D-glucono glucohidrolasa. La actividad de glucoamilasa puede comprender la catálisis de la hidrólisis de maltodextrinas, lo que resulta en la generación de glucosa libre. La actividad de glucoamilasa puede comprender una actividad de exoamilasa. La actividad de glucoamilasa puede comprender una actividad de α -amilasa o β -amilasa. Los enlaces glucosídicos hidrolizados pueden comprender enlaces α -1,4-glucosídicos o enlaces α -1,6-glucosídicos. La actividad de glucoamilasa puede comprender la hidrólisis de enlaces glucosídicos en un almidón. La actividad de glucoamilasa puede comprender además la hidrólisis de enlaces glucosídicos en el almidón para producir maltodextrinas. La actividad de glucoamilasa puede comprender la escisión de una unidad de maltosa o D-glucosa del extremo no reductor del almidón.

40 En un aspecto, la actividad de amilasa puede ser termoestable. El polipéptido puede retener una actividad de amilasa bajo condiciones que comprenden un rango de temperatura entre aproximadamente 37°C hasta aproximadamente 95 °C, entre aproximadamente 55 °C hasta aproximadamente 85 °C, entre aproximadamente 70 °C hasta aproximadamente 95 °C, o entre aproximadamente 90 °C hasta aproximadamente 95 °C. En otro aspecto, la actividad de amilasa puede ser termotolerante. El polipéptido puede retener actividad de amilasa después de la

exposición a una temperatura en el rango de más de 37 °C hasta aproximadamente 95 °C, o en el rango de más de 55 °C hasta aproximadamente 85 °C. En un aspecto, el polipéptido puede retener actividad de amilasa después de la exposición a una temperatura en el rango de más de 90 °C hasta aproximadamente 95 °C a un pH de 4,5.

5 En un aspecto, la actividad de amilasa es termoestable, por ejemplo, en donde el polipéptido retiene una actividad de amilasa bajo condiciones que comprenden un rango de temperatura de aproximadamente -100°C hasta aproximadamente -80°C, aproximadamente -80°C hasta aproximadamente -40°C, aproximadamente -40°C hasta aproximadamente -20°C, aproximadamente -20°C hasta aproximadamente 0°C, aproximadamente 0 °C hasta aproximadamente 5 °C, aproximadamente 5 °C hasta aproximadamente 15 °C, aproximadamente 15 °C hasta aproximadamente 25 °C, aproximadamente 25 °C hasta aproximadamente 37 °C, aproximadamente 37 °C hasta aproximadamente 45°C. aproximadamente 45°C hasta aproximadamente 55°C, aproximadamente 55 °C hasta aproximadamente 70°C, aproximadamente 70 °C hasta aproximadamente 75 °C, aproximadamente 75°C hasta aproximadamente 85°C, aproximadamente 85°C hasta aproximadamente 90 °C, aproximadamente 90°C hasta aproximadamente 95°C, aproximadamente 95°C hasta aproximadamente 100 °C, aproximadamente 100°C hasta aproximadamente 105°C, aproximadamente 105°C hasta aproximadamente 110°C, aproximadamente 110°C hasta aproximadamente 120°C, o 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100°C, 101°C, 102 °C, 103 °C, 104 °C, 105°C, 106°C, 107°C, 108°C, 109°C. 110 °C, 111 °C, 112 °C, 113°C, 114°C, 115°C o más. En algunas formas de realización, el polipéptido termoestable de acuerdo con la invención retiene actividad, por ejemplo, una actividad de amilasa, a una temperatura en los rangos descritos anteriormente, hasta aproximadamente pH 3,0, aproximadamente pH 3,5, aproximadamente pH 4,0, aproximadamente pH 4,5, aproximadamente pH 5,0, aproximadamente pH 5,5, aproximadamente pH 6,0, aproximadamente pH 6,5, aproximadamente pH 7,0, aproximadamente pH 7,5, aproximadamente pH 8,0, aproximadamente pH 8,5, aproximadamente pH 9,0, aproximadamente pH 9,5, aproximadamente pH 10,0, aproximadamente pH 10,5, aproximadamente pH 11,0, aproximadamente pH 11,5, aproximadamente pH 12,0 o más.

25 En un aspecto, la actividad de amilasa es termotolerante, por ejemplo, en donde el polipéptido retiene una actividad de amilasa después de la exposición a una temperatura en el rango de aproximadamente -100°C hasta aproximadamente -80°C, aproximadamente -80°C hasta aproximadamente -40°C, aproximadamente -40°C hasta aproximadamente -20°C, aproximadamente -20°C hasta aproximadamente 0°C, aproximadamente 0 °C hasta aproximadamente 5 °C, aproximadamente 5 °C hasta aproximadamente 15 °C, aproximadamente 15 °C hasta aproximadamente 25 °C, aproximadamente 25 °C hasta aproximadamente 37 °C, aproximadamente 37 °C hasta aproximadamente 45 °C, aproximadamente 45 °C hasta aproximadamente 55 °C, aproximadamente 55 °C hasta aproximadamente 70 °C, aproximadamente 70 °C hasta aproximadamente 75 °C, aproximadamente 75 °C hasta aproximadamente 85 °C. aproximadamente 85°C hasta aproximadamente 90 °C, aproximadamente 90 °C. hasta aproximadamente 95 °C. aproximadamente 95 °C hasta aproximadamente 100 °C, aproximadamente 100°C hasta aproximadamente 105 °C, aproximadamente 105 °C hasta aproximadamente 110 °C, aproximadamente 110 °C hasta aproximadamente 120 °C, o 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C o más. Los polipéptidos termotolerantes de acuerdo con la invención pueden retener actividad, por ejemplo actividad de amilasa, después de la exposición a una temperatura en el rango de aproximadamente -100 °C hasta aproximadamente -80 °C, aproximadamente -80 °C hasta aproximadamente -40 °C, aproximadamente -40 °C hasta aproximadamente -20 °C, aproximadamente -20 °C hasta aproximadamente 0 °C, aproximadamente 0 °C hasta aproximadamente 5 °C, aproximadamente 5 °C hasta aproximadamente 15 °C, aproximadamente 15 °C hasta aproximadamente 25 °C, aproximadamente 25 °C hasta aproximadamente 37 °C, aproximadamente 37 °C hasta aproximadamente 45 °C, aproximadamente 45°C hasta aproximadamente 55°C, aproximadamente 55 °C hasta aproximadamente 70 °C, aproximadamente 70 °C hasta aproximadamente 75 °C, aproximadamente 75 °C hasta aproximadamente 85 °C, aproximadamente 85 °C hasta aproximadamente 90 °C, aproximadamente 90 °C hasta aproximadamente 95 °C, aproximadamente 95 °C hasta aproximadamente 100 °C, aproximadamente 100 °C hasta aproximadamente 105 °C, aproximadamente 105 °C hasta aproximadamente 110 °C, aproximadamente 110 °C hasta aproximadamente 120 °C, o 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 104 °C, 105 °C, 106 °C, 107 °C, 108 °C, 109 °C, 110 °C, 111 °C, 112 °C, 113 °C, 114 °C, 115 °C o más. En algunas formas de realización, los polipéptidos termotolerantes de acuerdo con la invención retienen actividad, por ejemplo actividad de amilasa, después de la exposición a una temperatura en los rangos descritos anteriormente, hasta aproximadamente pH 3,0, aproximadamente pH 3,5, aproximadamente pH 4,0, aproximadamente pH 4,5, aproximadamente pH 5,0, aproximadamente pH 5,5, aproximadamente pH 6,0, aproximadamente pH 6,5, aproximadamente pH 7,0, aproximadamente pH 7,5, aproximadamente pH 8,0, aproximadamente pH 8,5, aproximadamente pH 9,0, aproximadamente pH 9,5, aproximadamente pH 10,0, aproximadamente pH 10,5, aproximadamente pH 11,0, aproximadamente pH 11,5, aproximadamente pH 12,0 o más.

60 En un aspecto, la actividad de amilasa de los polipéptidos codificados por ácidos nucleicos de la invención retienen actividad bajo condiciones ácidas que comprenden aproximadamente pH 6,5, pH 6, ,pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 o menos pH (más ácidas), o, retiene una actividad de amilasa después de la exposición a condiciones ácidas que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 o un pH menor (más ácido); o, retienen actividad bajo condiciones básicas que comprenden aproximadamente pH 7, pH 7,5, pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11, pH 11,5, pH 12, pH 12,5 o más (más básico) o, retienen

una actividad de amilasa después de exposición a condiciones básicas que comprenden aproximadamente pH 7, pH 7,5 pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11, pH 11,5, pH 12, pH 12,5 o más (más básico). En un aspecto, las actividad de amilasa de los polipéptidos codificados por ácidos nucleicos de la invención retienen actividad a una temperatura de al menos aproximadamente 80 °C, 81 °C, 82 °C, 83 °C, 84 °C, 85 °C, 86 °C, 87 °C, 88 °C, 89 °C, 90 °C, 91 °C, 92 °C, 93 °C, 94 °C, 95 °C, 96 °C, 97 °C, 98 °C, 99 °C, 100 °C, 101 °C, 102 °C, 103 °C, 103,5 °C, 104 °C, 105 °C, 107 °C, 108 °C, 109 °C o 110 °C, o más y un pH básico de al menos aproximadamente pH 7,5 pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11, pH 11,5, pH 12, pH 12,5 o más (más básico).

En un aspecto, el polipéptido aislado, sintético o recombinante, puede comprender al polipéptido de la divulgación que carece de una secuencia señal. En un aspecto, el polipéptido aislado, sintético o recombinante, puede comprender al polipéptido de la divulgación que comprende una secuencia señal heteróloga, tal como una secuencia señal de amilasa o no amilasa heteróloga.

En un aspecto, la invención provee una secuencia señal que comprende un péptido como se expone en la Tabla 1. En un aspecto, la invención provee una secuencia señal que consiste de un péptido como se expone en la Tabla 1. En un aspecto, la invención provee proteínas químicas que comprenden un primer dominio que comprende una secuencia señal de la invención y al menos un segundo dominio. La proteína puede ser una proteína de fusión. El segundo dominio puede comprender una enzima. La enzima puede ser cualquier glucoamilasa y/o amilasa (por ejemplo, una amilasa de la invención, u otra amilasa o glucoamilasa).

En un aspecto, la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención comprende una actividad específica hasta aproximadamente 37 °C en el rango desde aproximadamente 1 hasta aproximadamente 1200 unidades por miligramo de proteína, o, aproximadamente 100 hasta aproximadamente 1000 unidades por miligramo de proteína. En otro aspecto, la actividad enzimática (por ejemplo, la actividad de **amilasa**) de una enzima de esta invención comprende una actividad específica de aproximadamente 100 hasta aproximadamente 1000 unidades por miligramo de proteína, o, de aproximadamente 500 hasta aproximadamente 750 unidades por miligramo de proteína. En forma alternativa, la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención comprende una actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 750 unidades por miligramo de proteína, o, de aproximadamente 500 hasta aproximadamente 1200 unidades por miligramo de proteína. En un aspecto, la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención comprende una actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 500 unidades por miligramo de proteína, o, de aproximadamente 750 hasta aproximadamente 1000 unidades por miligramo de proteína. En otro aspecto, la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención comprende una actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 250 unidades por miligramo de proteína. En forma alternativa, la actividad enzimática (por ejemplo, la actividad de glucoamilasa) de una enzima de esta invención comprende una actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 100 unidades por miligramo de proteína. En otro aspecto, la termotolerancia comprende la retención de al menos la mitad de la actividad específica de la actividad enzimática (por ejemplo, la actividad de glucoamilasa) de una enzima de esta invención a 37 °C después de calentamiento a temperatura elevada, tal como a una temperatura de aproximadamente 0 °C hasta aproximadamente 20 °C, aproximadamente 20 °C hasta aproximadamente 37 °C, aproximadamente 37 °C hasta aproximadamente 50 °C, aproximadamente 50 °C hasta aproximadamente 70 °C, aproximadamente 70 °C hasta aproximadamente 75 °C, aproximadamente 75 °C hasta aproximadamente 80 °C, aproximadamente 80 °C hasta aproximadamente 85 °C, aproximadamente 85 °C hasta aproximadamente 90 °C, aproximadamente 90 °C hasta aproximadamente 95 °C, aproximadamente 95 °C hasta aproximadamente 100 °C, aproximadamente 100 °C hasta aproximadamente 110 °C, o más. En forma alternativa, la termotolerancia puede comprender retención de la actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 1200 unidades por miligramo de proteína, o de aproximadamente 500 hasta aproximadamente 1000 unidades por miligramo de proteína, después de calentamiento a una temperatura elevada. En otro aspecto, la termotolerancia puede comprender la retención de actividad específica a 37 °C en el rango de aproximadamente 1 hasta aproximadamente 500 unidades por miligramo de proteína después de calentamiento a una temperatura elevada, como se describió anteriormente.

La divulgación provee polipéptidos aislados, sintéticos o recombinantes, de la invención, en donde el polipéptido comprende al menos un sitio de glicosilación. En un aspecto, la glicosilación puede ser una glicosilación ligada a N. En un aspecto, el polipéptido puede ser glicosilado después de ser expresado en *P. pastoris* o *S. pombe*. La invención también provee métodos para agregar glicosilación a un polipéptido, ya sea después de la traducción o químicamente, para cambiar la propiedad de los polipéptidos, por ejemplo, su estabilidad térmica, solubilidad, tendencia a agregarse y similares.

En un aspecto, el polipéptido puede retener la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención bajo condiciones que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 o menos pH (más ácido). En otro aspecto, el polipéptido puede retener la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención bajo condiciones que comprenden aproximadamente pH 7, pH 7,5 pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11,0, pH 11,5, pH 12,

pH 12,5 o más pH (más básico). En un aspecto, el polipéptido puede retener la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención después de la exposición a condiciones que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 o menos pH (más ácido). En otro aspecto, el polipéptido puede retener la actividad enzimática (por ejemplo, la actividad de amilasa) de una enzima de esta invención después de la exposición a condiciones que comprenden aproximadamente pH 7, pH 7,5, pH 8,0, pH 8.5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11,0, pH 11,5, pH 12, pH 12,5 o más pH (más básico).

La divulgación provee preparaciones de proteínas que comprenden un polipéptido de la invención, en donde la preparación de la proteína comprende un líquido, un sólido o un gel.

La divulgación provee heterodímeros que comprenden un polipéptido de la invención y un segundo dominio. En un aspecto, el segundo dominio puede ser un polipéptido y el heterodímero puede ser una proteína de fusión. En un aspecto, el segundo dominio puede ser un epítipo o una etiqueta. En un aspecto, la invención provee homodímeros que comprenden un polipéptido de la invención.

La divulgación provee polipéptidos inmovilizados que tienen actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención, un polipéptido codificado por un ácido nucleico de la invención, o un polipéptido que comprende un polipéptido de la invención y un segundo dominio. En un aspecto, el polipéptido puede ser inmovilizado en una célula, un metal, una resina, un polímero, una cerámica, un vidrio, un microelectrodo, una partícula gráfica, una perla, un gel, una placa, una disposición o un tubo capilar.

La divulgación provee arreglos que comprenden un ácido nucleico inmovilizado de la divulgación. La divulgación provee arreglos que comprenden un anticuerpo de la divulgación.

La divulgación provee anticuerpos aislados, sintéticos o recombinantes, que se enlazan de manera específica a un polipéptido de la invención o a un polipéptido codificado por un ácido nucleico de la invención. El anticuerpo puede ser un anticuerpo monoclonal o policlonal. La divulgación provee hibridomas que comprenden un anticuerpo de la invención, por ejemplo, un anticuerpo que se enlaza de manera específica a un polipéptido de la invención o a un polipéptido codificado por un ácido nucleico de la invención.

La divulgación provee suplementos alimenticios para un animal que comprenden un polipéptido de la invención, por ejemplo, un polipéptido codificado por el ácido nucleico de la invención. En un aspecto, el polipéptido en el suplemento alimenticio puede ser glicosilado. La divulgación provee matrices de suministro de enzimas comestibles que comprenden un polipéptido de la invención, por ejemplo, un polipéptido codificado por el ácido nucleico de la invención. En un aspecto, la matriz de suministro comprende una perla. En un aspecto, el polipéptido puede ser glicosilado. En un aspecto, la actividad de amilasa es termotolerante. En otro aspecto, la actividad de amilasa es termoestable.

La divulgación provee un método para aislar o identificar un polipéptido que tiene actividad de amilasa que comprende las etapas de: (a) proveer un anticuerpo de la invención; (b) proveer una muestra que comprende polipéptidos; y (c) poner en contacto la muestra de la etapa (b) con el anticuerpo de la etapa (a) bajo condiciones en donde el anticuerpo se puede enlazar de manera específica al polipéptido, aislando o identificando de este modo un polipéptido que tiene actividad de amilasa.

La divulgación provee métodos para formar un anticuerpo antiamilasa que comprende administrar a un animal no humano un ácido nucleico de la invención o un polipéptido de la invención o subsecuencias de los mismos en una cantidad suficiente para generar una respuesta inmune humoral, elaborando de este modo un anticuerpo antiglucoamilasa y/o antiamilasa. La invención provee métodos para elaborar una antiamilasa inmune que comprende administrar a un animal no humano un ácido nucleico de la invención o un polipéptido de la invención o subsecuencias de los mismos en una cantidad suficiente para generar una respuesta inmune.

La invención provee métodos para producir un polipéptido recombinante que comprende las etapas de: (a) proveer un ácido nucleico de la invención operativamente enlazado a un promotor; y (b) expresar el ácido nucleico de la etapa (a) bajo condiciones que permitan la expresión del polipéptido, produciendo de este modo un polipéptido recombinante. El método puede comprender, además, transformar una célula huésped con el ácido nucleico de la etapa (a) seguido de expresar el ácido nucleico de la etapa (a), produciendo de este modo un polipéptido recombinante en una célula transformada.

La divulgación provee métodos para identificar un polipéptido que tiene actividad de amilasa que comprende las siguientes etapas: (a) proveer un polipéptido de la invención; o un polipéptido codificado por un ácido nucleico de la invención; (b) proveer un sustrato de amilasa; y (c) poner en contacto el polipéptido o un fragmento o variante del mismo de la etapa (a) con el sustrato de la etapa (b) y detectar una reducción en la cantidad de sustrato o un incremento en la cantidad de un producto de reacción, en donde una reducción en la cantidad del sustrato o un

incremento en la cantidad del producto de reacción detecta un polipéptido que tiene actividad de amilasa. En un aspecto, el sustrato puede ser un polisacárido, oligosacárido o almidón, por ejemplo, un almidón licuado.

5 La divulgación provee métodos para identificar un sustrato de amilasa que comprende las siguientes etapas: (a) proveer un polipéptido de la invención; o un polipéptido codificado por un ácido nucleico de la invención; (b) proveer un sustrato de prueba; y (c) poner en contacto el polipéptido de la etapa (a) con el sustrato de prueba de la etapa (b) y detectar una reducción en la cantidad de sustrato o un incremento en la cantidad del producto de reacción, en donde una reducción en la cantidad del sustrato o un incremento en la cantidad de un producto de reacción identifica al sustrato de prueba como un sustrato de amilasa.

10 La divulgación provee métodos para determinar si un compuesto de prueba se enlaza de manera específica a un polipéptido que comprende las siguientes etapas: (a) expresar un ácido nucleico o un vector que comprende el ácido nucleico bajo condiciones permisivas para traducción del ácido nucleico en un polipéptido, en donde el ácido nucleico comprende un ácido nucleico de la divulgación, o, proveer un polipéptido de la divulgación; (b) proveer un compuesto de prueba; (c) poner en contacto el polipéptido con el compuesto de prueba; y (d) determinar si el compuesto de prueba de la etapa (b) se enlaza de manera específica al polipéptido.

15 La divulgación provee métodos para identificar un modulador de una actividad de amilasa que comprende las siguientes etapas: (a) proveer un polipéptido de la divulgación o un polipéptido codificado por un ácido nucleico de la divulgación; (b) proveer un compuesto de prueba; (c) poner en contacto el polipéptido de la etapa (a) con el compuesto de prueba de la etapa (b) y medir la actividad de la amilasa, en donde un cambio en la actividad de la amilasa medida en presencia del compuesto de prueba comparado con la actividad en ausencia del compuesto de evaluación provee una determinación de que el compuesto de evaluación modula la actividad de amilasa o glucoamilasa. En un aspecto, la actividad de amilasa puede ser medida proveyendo un sustrato de amilasa y detectar una reducción en la cantidad del sustrato o un incremento en la cantidad del producto de reacción, o, un incremento en la cantidad del sustrato o una reducción en la cantidad del producto de reacción. Una reducción en la cantidad del sustrato o un incremento en la cantidad del producto de reacción con el compuesto de prueba en comparación con la cantidad de sustrato o del producto de reacción sin el compuesto de prueba identifica el compuesto de prueba como activador de la actividad de amilasa o de glucoamilasa. Un incremento en la cantidad del sustrato o una reducción en la cantidad del producto de reacción con el compuesto de prueba comparado con la cantidad de sustrato o producto de reacción sin el compuesto de prueba identifica al compuesto de prueba como un inhibidor de la actividad de amilasa.

30 La divulgación provee sistemas computarizados que comprenden un procesador y un dispositivo de almacenamiento de datos en donde dicho dispositivo de almacenamiento de datos tiene almacenada una secuencia de polipéptidos o una secuencia de ácido nucleico de la divulgación (por ejemplo, un polipéptido codificado por un ácido nucleico de la divulgación). En un aspecto, el sistema de cómputo puede comprender, además, un algoritmo de comparación de secuencias y un dispositivo de almacenamiento de datos que tienen al menos una secuencia de referencia almacenada. En otro aspecto, el algoritmo de comparación de secuencias comprende un programa de ordenador que indica polimorfismos. En un aspecto, el sistema de cómputo puede comprender, además, un identificador que identifica una o más características en dicha secuencia. La divulgación provee medios legibles para el ordenador que tienen almacenada una secuencia de polipéptidos o una secuencia de ácido nucleico de la invención. La divulgación provee métodos para identificar una característica en una secuencia que comprende las etapas de: (a) leer la secuencia usando un programa de ordenador que identifica una o más características en una secuencia, en donde la secuencia comprende una secuencia de polipéptidos o una secuencia de ácido nucleico de la divulgación; y (b) identificar una o más características en la secuencia con el programa del ordenador. La divulgación provee métodos para comparar una primera secuencia con una segunda secuencia que comprende las etapas de: (a) leer la primera secuencia y la segunda secuencia a través del uso de un programa de ordenador que compara secuencias, en donde la primera secuencia comprende una secuencia de polipéptido o una secuencia de ácido nucleico de la divulgación; y (b) determinar las diferencias entre la primera secuencia y la segunda secuencia con el programa de ordenador. La etapa de determinar diferencias entre la primera secuencia y la segunda secuencia puede comprender, además, la etapa de identificación de polimorfismos. En un aspecto, el método puede comprender, además, un identificador que identifica una o más características en una secuencia. En otro aspecto, el método puede comprender leer la primera secuencia usando un programa de ordenador e identificar una o más características en la secuencia.

55 La divulgación provee métodos para aislar o recuperar un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa de una muestra del ambiente que comprende las etapas de: (a) proveer un par de secuencias del cebador de amplificación para amplificar un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa y/o de glucoamilasa, en donde el par cebador es capaz de amplificar un ácido nucleico de la divulgación; (b) aislar un ácido nucleico de la muestra del ambiente o tratar la muestra del ambiente de manera tal que el ácido nucleico en la muestra sea accesible para hibridación con el par cebador de amplificación; y, (c) combinar el ácido nucleico de la etapa (b) con el par cebador de amplificación de la etapa (a) y amplificar el ácido nucleico de la muestra del ambiente, aislando o recuperando de este modo un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa de una muestra del ambiente. Uno o cada elemento del par de secuencias del cebador de
60

amplificación pueden comprender un oligonucleótido que comprende al menos aproximadamente 10 a 50 bases consecutivas de una secuencia de la divulgación.

5 La divulgación provee métodos para aislar o recuperar un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa de una muestra del ambiente que comprende las etapas de: (a) proveer una sonda de polinucleótido que comprende un ácido nucleico de la divulgación o una subsecuencia del mismo; (b) aislar un ácido nucleico de la muestra del ambiente o tratar la muestra del ambiente de manera tal que el ácido nucleico en la muestra sea accesible para hibridación con una sonda de polinucleótido de la etapa (a); (c) combinar el ácido nucleico aislado o la muestra del ambiente tratada de la etapa (b) con la sonda de polinucleótido de la etapa (a); y (d) aislar un ácido nucleico que hibrida de manera específica con la sonda de polinucleótido de la etapa (a), aislando o recuperando de este modo un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa de una muestra del ambiente. La muestra del ambiente puede comprender una muestra acuosa, una muestra líquida, una muestra de suelo, una muestra de aire o una muestra biológica. En un aspecto, la muestra biológica puede derivar de una célula bacteriana, una célula de protozoo, una célula de insecto, una célula de levadura, una célula vegetal, una célula de hongo o una célula de mamífero.

15 La divulgación provee métodos para generar una variante de un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa y/o de glucoamilasa que comprende las etapas de: (a) proveer un ácido nucleico plantilla que comprende un ácido nucleico de la divulgación; y (b) modificar, suprimir o agregar uno o más nucleótidos en la secuencia plantilla, o una combinación de los mismos, para generar una variante del ácido nucleico plantilla. En un aspecto, el método puede comprender, además, expresar la variante de ácido nucleico para generar una variante de polipéptido de amilasa. Las modificaciones, adiciones o supresiones se pueden introducir por medio de un método que comprende PCR propensa a error, arrastre, mutagénesis dirigida al oligonucleótido, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casetes, mutagénesis de ensamblaje recursivo, mutagénesis de ensamblaje exponencial, mutagénesis específica del sitio, reensamble de genes, mutagénesis de saturación del sitio génico (GSSM), reensamble del ligamiento sintético (SLR) o una combinación de los mismos. En otro aspecto, las modificaciones, adiciones o supresiones son introducidas por medio de un método que comprende recombinación, recombinación de secuencia recursiva, mutagénesis de ADN modificado por fosfotioato, mutagénesis de la plantilla que contiene uracilo, mutagénesis dúplex con espacios, mutagénesis de reparación de falta de correlación puntual, mutagénesis de cepa huésped de reparación deficiente, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de selección-restricción, mutagénesis de restricción-purificación, síntesis génica artificial, mutagénesis de ensamblaje, creación de multímeros de ácido nucleico quimérico y una combinación de los mismos.

En un aspecto, el método puede repetirse en forma iterativa hasta que se produce una amilasa que tiene una actividad alterada o diferente o una estabilidad alterada o diferente de aquella de un polipéptido codificado por el ácido nucleico plantilla. En un aspecto, la variante del polipéptido de amilasa es termotolerante y retiene alguna actividad después de ser expuesta a una temperatura elevada. En otro aspecto, la variante del polipéptido de amilasa tiene mayor glicosilación en comparación con la amilasa codificada por un ácido nucleico plantilla. En forma alternativa, la variante del polipéptido de amilasa tiene actividad de amilasa bajo una temperatura alta, en donde la amilasa o glucoamilasa codificada por el ácido nucleico plantilla no es activa a temperatura alta. En un aspecto, el método puede repetirse en forma iterativa hasta que se produce una secuencia codificadora de amilasa que tiene un uso de codón alterado con relación al del ácido nucleico plantilla. En otro aspecto, el método puede repetirse en forma iterativa hasta que se produce un gen de amilasa que tiene nivel inferior o superior de expresión o estabilidad del mensaje con respecto al del ácido nucleico plantilla.

La divulgación provee métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa para incrementar su expresión en una célula huésped; el método comprende las siguientes etapas: (a) proveer un ácido nucleico de la divulgación que codifica un polipéptido que tiene actividad de amilasa; y (b) identificar un codón no preferido o menos preferido en el ácido nucleico de la etapa (a) y reemplazarlo con un codón usado en forma neutral o preferida que codifica el mismo aminoácido que el codón reemplazado, en donde un codón preferido es un codón sobrerrepresentado en las secuencias codificadoras en los genes en la célula huésped y un codón no preferido o menos preferido es un codón subrepresentado en las secuencias codificadoras en los genes en la célula huésped, modificando de este modo el ácido nucleico para incrementar su expresión en una célula huésped.

La divulgación provee métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tienen actividad de amilasa; el método comprende las siguientes etapas: (a) proveer un ácido nucleico de la invención; y, (b) divulgación un codón en el ácido nucleico de la etapa (a) y reemplazarlo con un codón diferente que codifica el mismo aminoácido que el codón reemplazado, modificando de este modo los codones en un ácido nucleico que codifica una amilasa.

La divulgación provee métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa para incrementar su expresión en una célula huésped, el método comprendiendo las siguientes etapas: (a) proveer un ácido nucleico de la divulgación que codifica un polipéptido de amilasa; y, (b) identificar un

codón no preferido o menos preferido en el ácido nucleico de la etapa (a) y reemplazarlo con un codón usado en forma preferida o neutra que codifica el mismo aminoácido que el codón reemplazado, en donde un codón preferido es un codón preferido sobrerrepresentado en secuencias codificadoras en genes en la célula huésped y un codón no preferido o menos preferido es un codón subrepresentado en secuencias codificadoras en genes en la célula huésped, modificando de este modo el ácido nucleico para incrementar su expresión en una célula huésped.

La divulgación provee métodos para modificar un codón en un ácido nucleico que codifica un polipéptido que tiene actividad de amilasa para reducir su expresión en una célula huésped, el método comprendiendo las siguientes etapas: (a) proveer un ácido nucleico de la divulgación; e (b) identificar al menos un codón preferido en el ácido nucleico de la etapa (a) y reemplazarlo con un codón no preferido o menos preferido que codifica el mismo aminoácido que el codón reemplazado, en donde un codón preferido es un codón sobrerrepresentado en secuencias codificadoras en genes en una célula huésped y un codón no preferido o menos preferido es un codón subrepresentado en secuencias codificadoras en genes en la célula huésped, modificando de este modo el ácido nucleico para reducir su expresión en una célula huésped. En un aspecto, la célula huésped puede ser una célula bacteriana, una célula de hongo, una célula de insecto, una célula de levadura, una célula vegetal o una célula de mamífero.

La divulgación provee métodos para producir una biblioteca génica de ácidos nucleicos que codifican una pluralidad de sitios activos de amilasa modificados o sitios de enlazamiento de sustratos, en donde los sitios activos modificados o sitios de enlazamiento de sustratos se derivan de un primer ácido nucleico que comprende una secuencia que codifica un primer sitio activo o un primer sitio de enlazamiento del sustrato, el método comprendiendo las siguientes etapas: (a) proveer un primer ácido nucleico que codifica un primer sitio activo o primer sitio de enlazamiento de sustratos, en donde la primera secuencia de ácido nucleico comprende una secuencia que se hibrida bajo condiciones rigurosas con un ácido nucleico de la divulgación y el ácido nucleico codifica un sitio activo de amilasa o un sitio de enlazamiento del sustrato de amilasa o glucoamilasa; (b) proveer un conjunto de oligonucleótidos mutagénicos que codifican variantes de aminoácidos de ocurrencia natural en una pluralidad de codones objetivo en el primer ácido nucleico; y, (c) usar el conjunto de oligonucleótidos mutagénicos para generar un conjunto de variantes de ácidos nucleicos que codifican el sitio activo o que codifican el sitio de enlazamiento del sustrato, que codifica un rango de variaciones de aminoácidos en cada codón de aminoácidos que experimentó mutagénesis produciendo de este modo una biblioteca de ácidos nucleicos que codifican una pluralidad de sitios activos de amilasa modificados o sitios de enlazamiento de sustratos. En un aspecto, el método comprende provocar mutagénesis del primer ácido nucleico de la etapa (a) por medio de un método que comprende un sistema de evolución dirigida optimizada, mutagénesis de saturación del sitio génico (GSSM), reensamble de ligamiento sintético (SLR), PCR propensa a error, arrastre, mutagénesis dirigida al oligonucleótido, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casete, mutagénesis de ensamblaje recursivo, mutagénesis de ensamble exponencial, mutagénesis específica del sitio, reensamble génico, mutagénesis de saturación del sitio génico (GSSM), reensamble de ligamiento sintético (SLR) y una combinación de los mismos. En otro aspecto, el método comprende someter a mutagénesis al primer ácido nucleico de la etapa (a) o variantes por medio de un método que comprende recombinación, recombinación de secuencia recursiva, mutagénesis del ADN modificado por fosfotoato, mutagénesis de la plantilla que contiene uracilo, mutagénesis dúplex con espacios, mutagénesis de reparación de falta de correlación puntual, mutagénesis de cepa huésped de reparación deficiente, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de selección-restricción, mutagénesis de restricción-purificación, síntesis génica artificial, mutagénesis de ensamblaje, creación de multímeros de ácido nucleico quimérico y una combinación de los mismos.

La divulgación provee métodos para elaborar una molécula pequeña que comprende las siguientes etapas: (a) proveer una pluralidad de enzimas biosintéticas capaces de sintetizar o modificar una molécula pequeña, en donde una de las enzimas comprende una enzima amilasa o glucoamilasa codificada por un ácido nucleico de la divulgación; (b) proveer un sustrato para al menos una de las enzimas de la etapa (a); y (c) hacer reaccionar el sustrato de la etapa (b) con las enzimas bajo condiciones que facilitan una pluralidad de reacciones biocatalíticas para generar una molécula pequeña mediante una serie de reacciones biocatalíticas. La divulgación provee métodos para modificar una molécula pequeña que comprende las siguientes etapas: (a) proveer una enzima amilasa, en donde la enzima comprende un polipéptido de la divulgación, o, un polipéptido codificado por un ácido nucleico de la divulgación, o una subsecuencia del mismo; (b) proveer una molécula pequeña; y (c) hacer reaccionar la enzima de la etapa (a) con la molécula pequeña de la etapa (b) bajo condiciones que facilitan una reacción enzimática catalizada por la enzima amilasa, modificando de este modo una molécula pequeña mediante una reacción enzimática de amilasa. En un aspecto, el método puede comprender una pluralidad de sustratos de molécula pequeña para la enzima de la etapa (a), generando de este modo una biblioteca de moléculas pequeñas modificadas producidas por al menos una reacción enzimática catalizada por la enzima amilasa. En un aspecto, el método puede comprender una pluralidad de enzimas adicionales bajo condiciones que facilitan una pluralidad de reacciones biocatalíticas por la enzimas para formar una biblioteca de moléculas pequeñas modificadas producidas por la pluralidad de reacciones enzimáticas. En otro aspecto, el método puede comprender, además, la etapa de evaluar la biblioteca para determinar si una molécula pequeña modificada particular que exhibe una actividad deseada está presente en la biblioteca. La etapa de evaluar la biblioteca puede comprender, además, las etapas de eliminar en forma sistemática todas menos una de las reacciones biocatalíticas usadas para producir una porción de

la pluralidad de las moléculas pequeñas modificadas en la biblioteca mediante la evaluación de la porción de la molécula pequeña modificada en cuanto a la presencia o ausencia de la molécula pequeña modificada particular con una actividad deseada e identificar al menos una reacción biocatalítica específica que produce la molécula pequeña modificada particular de actividad deseada.

- 5 La divulgación provee métodos para determinar un fragmento funcional de una enzima amilasa que comprende las etapas de: (a) proveer una enzima amilasa, en donde la enzima comprende un polipéptido de la divulgación, o un polipéptido codificado por un ácido nucleico de la divulgación, o una subsecuencia del mismo; y (b) suprimir una pluralidad de residuos de aminoácidos de la secuencia de la etapa (a) y evaluar el resto de la subsecuencia en cuanto a la actividad de amilasa, determinando de este modo un fragmento funcional de una enzima amilasa o glucoamilasa. En un aspecto, la actividad de amilasa se mide al proveer un sustrato de amilasa y detectar una reducción en la cantidad del sustrato o un incremento en la cantidad de un producto de reacción.

- 15 La divulgación provee métodos para manipulación genética de la célula completa de fenotipos nuevos o modificados usando análisis de flujo metabólico en tiempo real; el método comprende las siguientes etapas: (a) elaborar una célula modificada mediante la modificación de la composición genética de una célula, en donde la composición genética es modificada por la adición a la célula de un ácido nucleico de la divulgación; (b) cultivar la célula modificada para generar una pluralidad de células modificadas; (c) medir al menos un parámetro metabólico de las células mediante el control del cultivo celular de la etapa (b) en tiempo real; y, (d) analizar los datos de la etapa (c) para determinar si el parámetro medido difiere de una medición comparable en una célula no modificada bajo condiciones similares, identificando de este modo un fenotipo modificado por ingeniería genética en la célula usando análisis de flujo metabólico en tiempo real. En un aspecto, la composición genética de las células puede ser modificada por medio de un método que comprende la supresión de una secuencia o modificación de una secuencia en la célula, o inactivación de la expresión de un gen. En un aspecto, el método puede comprender, además, seleccionar una célula que comprende un fenotipo recientemente modificado por ingeniería genética. En otro aspecto, el método puede comprender cultivar la célula seleccionada, generando de este modo una nueva cepa celular que comprende un fenotipo recientemente modificado por ingeniería genética.

- 20 La invención provee métodos para hidrolizar un polisacárido, oligosacárido o almidón, que comprende las siguientes etapas: (a) proveer un polipéptido que tenga actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; y (c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde el polipéptido hidroliza al polisacárido, oligosacárido o almidón. En un aspecto, la composición que comprende al polisacárido, oligosacárido o almidón, que comprende un enlace α -1,4-glucosídico o un enlace α -1,6-glucosídico. En un aspecto, la actividad de la amilasa es una actividad α -amilasa. En un aspecto, la actividad α -amilasa hidroliza enlaces internos en un almidón u otro polisacárido.

- 30 La invención provee métodos para licuar o remover un polisacárido, oligosacárido o almidón, de una composición que comprende las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; y (c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde el polipéptido remueve o licua el polisacárido, oligosacárido o almidón.

- 40 La divulgación provee métodos para incrementar la termotolerancia o termoestabilidad de un polipéptido de amilasa, el método comprendiendo la glicosilación de un polipéptido de amilasa, en donde el polipéptido comprende al menos treinta aminoácidos contiguos de un polipéptido de la divulgación; o un polipéptido codificado por una secuencia de ácido nucleico de la divulgación, incrementando de este modo la termotolerancia o termoestabilidad del polipéptido de amilasa. En un aspecto, la actividad específica de amilasa puede ser termoestable o termotolerante a una temperatura en el rango de más de aproximadamente 37 °C hasta aproximadamente 95 °C.

- 45 La divulgación provee métodos para sobreexpresar un polipéptido recombinante de amilasa en una célula que comprende expresar un vector que comprende un ácido nucleico que comprende un ácido nucleico de la divulgación o una secuencia de ácido nucleico de la divulgación, en donde las identidades de secuencia se determinan por análisis con un algoritmo de comparación de secuencias o por inspección visual, en donde la sobreexpresión se efectúa mediante el uso de un promotor de alta actividad, un vector dicistrónico o por amplificación génica del vector.

- 50 La invención provee composiciones detergentes que comprenden un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la invención, en donde el polipéptido comprende actividad de amilasa. En un aspecto, la amilasa puede ser una amilasa de superficie no activa. En otro aspecto, la amilasa puede ser una amilasa de superficie activa.

- 55 La invención provee métodos para lavar un objeto que comprende las siguientes etapas: (a) proveer una composición que comprende un polipéptido que tiene una actividad de amilasa y/o glucoamilasa, en donde el polipéptido comprende: un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la

invención; (b) proveer un objeto; y (c) poner en contacto el polipéptido de la etapa (a) y el objeto de la etapa (b) bajo condiciones en donde la composición puede lavar al objeto.

5 La invención provee métodos para hidrolizar un polisacárido, oligosacárido o almidón, por ejemplo, en un pienso o alimento antes del consumo por parte de un animal, que comprende las siguientes etapas: (a) obtener una composición, por ejemplo, un material para pienso, que comprende un polisacárido, oligosacárido o almidón, en donde el polipéptido comprende: un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la invención; y (b) agregar el polipéptido de la etapa (a) a la composición, por ejemplo, al material alimenticio o del pienso, en una cantidad suficiente durante un período de tiempo suficiente para causar la hidrólisis del polisacárido, oligosacárido o almidón, hidrolizando de este modo al polisacárido, oligosacárido o almidón. En un aspecto, el pienso o el alimento comprenden arroz, maíz, cebada, trigo, legumbres, o patata.

10 Por ejemplo, en una forma de realización, la divulgación provee composiciones que comprenden una combinación de una amilasa y una glucoamilasa (donde la amilasa es la enzima de esta invención) para la hidrólisis de un polisacárido, oligosacárido o almidón, por ejemplo, almidón de maíz, arroz, cebada, trigo, legumbres, o patata. En un aspecto, la carga de enzima para una combinación de una amilasa y una glucoamilasa comprende una proporción de 1:10 de amilasa:glucoamilasa; en donde en una forma de realización, una carga de enzima total está entre 0,015% - 0,0255% de enzima necesaria para hidrolizar completamente 33% de una harina, por ejemplo, una harina de maíz. Los ejemplos de rangos alternativos para carga son de aproximadamente 0,01% (p/p) para hidrólisis de 23% de almidón de maíz purificado y entre 0,015% - 0,2% (p/p) para hidrólisis de almidón de maíz en 33% de harina de maíz.

15 20 La invención provee métodos para desencolado textil que comprende las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la invención; (b) proveer una textil; y (c) poner en contacto el polipéptido de la etapa (a) y el textil de la etapa (b) bajo condiciones en donde la amilasa puede desencolar el textil.

25 La invención provee métodos para desentintar papel o fibras que comprenden las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer una composición que comprende papel o fibra; y (c) poner en contacto el polipéptido de la etapa (a) y la composición de la etapa (b) bajo condiciones en donde el polipéptido puede desentintar el papel o la fibra.

30 La invención provee métodos para tratamiento de fibras lignocelulósicas que comprenden las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer una fibra lignocelulósica; y (c) poner en contacto el polipéptido de la etapa (a) y la fibra de la etapa (b) bajo condiciones en donde el polipéptido puede tratar la fibra, mejorando de este modo las propiedades de la fibra.

35 La invención provee métodos para producir un jarabe de alto contenido de maltosa o alto contenido de glucosa que comprende las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende una enzima de la invención; (b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; y (c) poner en contacto el polipéptido de la etapa (a) y el textil de la etapa (b) bajo condiciones en donde el polipéptido de la etapa (a) puede licuar la composición de la etapa (b) produciendo de este modo un hidrolizado de polisacárido, oligosacárido o almidón soluble, y hidrolizar y sacarificar el polisacárido, oligosacárido o almidón soluble, y produciendo así el hidrolizado el jarabe. En un aspecto, el almidón puede ser de arroz, maíz, cebada, trigo, legumbres, patata o batata.

40 45 La invención provee métodos para mejorar el flujo de los fluidos de producción que comprenden polisacárido, por ejemplo, que contienen almidón, que comprende las siguientes etapas: (a) proveer un polipéptido que tiene actividad de glucoamilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer fluido de producción; y (c) poner en contacto el polipéptido de la etapa (a) y el fluido de producción de la etapa (b) bajo condiciones en donde la glucoamilasa puede hidrolizar el polisacárido, oligosacárido o almidón, en el fluido de producción y mejorando así su flujo al reducir su densidad. En un aspecto, el fluido de producción puede ser de una formación subterránea.

50 55 La invención provee composiciones antienranciamiento que comprenden un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la invención. La divulgación provee métodos para prevenir el enranciamiento de los productos horneados que comprenden las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la divulgación; (b) proveer una composición que contiene un polisacárido, oligosacárido o almidón, usado para hornear productos; (c) combinar el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde el polipéptido puede hidrolizar el polisacárido, oligosacárido o almidón, en la composición usada para hornear previniendo así el enranciamiento del producto horneado. En un aspecto, el producto horneado puede ser pan.

- La invención provee métodos para usar amilasa en la elaboración de cerveza o producción de alcohol que comprende las siguientes etapas: (a) proveer un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende un polipéptido de la invención; (b) proveer una composición que contiene un polisacárido, oligosacárido o almidón y usarlo para elaboración de cerveza o en la producción del alcohol; (c) combinar el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde el polipéptido puede hidrolizar el polisacárido, oligosacárido o almidón, en la composición usada para elaboración de cerveza o en la producción de alcohol. En un aspecto, la composición que contiene un polisacárido, oligosacárido o almidón, puede ser cerveza.
- La divulgación provee métodos para la elaboración de una planta transgénica que comprende las siguientes etapas: (a) introducir una secuencia de ácido nucleico heteróloga en la célula, en donde la secuencia de ácido nucleico heteróloga comprende una secuencia de ácido nucleico de la divulgación, produciendo de este modo una célula vegetal transformada; y (b) producir una planta transgénica a partir de la célula transformada. En un aspecto, la etapa (a) puede comprender, además, introducir la secuencia de ácido nucleico heteróloga por electroporación o microinyección de protoplastos de células vegetales. En otro aspecto, la etapa (a) puede comprender además la introducción de la secuencia de ácido nucleico heteróloga directamente al tejido vegetal por bombardeo de partículas de ADN. En forma alternativa, la etapa (a) puede comprender además la introducción de la secuencia de ácido nucleico heteróloga en el ADN de la célula vegetal usando un huésped de *Agrobacterium tumefaciens*. En un aspecto, la célula vegetal puede ser una célula de patata, maíz, arroz, trigo, tabaco o cebada.
- La divulgación provee métodos para expresar una secuencia de ácido nucleico heteróloga en una célula vegetal que comprende las siguientes etapas: (a) transformar la célula vegetal con una secuencia de ácido nucleico heteróloga enlazada operativamente a un promotor, en donde la secuencia heteróloga de ácido nucleico comprende un ácido nucleico de la divulgación; (b) cultivar la planta bajo condiciones en donde la secuencia heteróloga de ácido nucleico se expresa en la célula vegetal.
- La divulgación también provee un proceso para preparar una masa o un producto horneado preparado a partir de la masa que comprende agregar una amilasa de la divulgación a la masa en una cantidad que sea efectiva para retrasar el enranciamiento del pan. La divulgación también provee una masa que comprende dicha amilasa y una premezcla que comprende harina junto con dicha amilasa. Finalmente, la invención provee un aditivo enzimático para horneado, que contiene dicha amilasa o glucoamilasa. El uso de la amilasa de acuerdo con la presente invención provee un efecto anti-enranciamiento mejorado como el medido, por ejemplo, mediante menos firmeza de la miga, reteniendo la elasticidad de la miga, mejor capacidad de rebanado (por ejemplo, menos migas, miga no gomosa), mejor gusto o sabor.
- Las composiciones que comprenden enzimas de la invención (por ejemplo, que comprenden los polipéptidos, ácidos nucleicos de esta invención) se pueden formular en una variedad de formas, por ejemplo, como líquidos, geles, píldoras, tabletas, aerosoles, polvos, alimento, pellas de pienso o formas encapsuladas, que incluyen formas nanoencapsuladas. Cualquiera de estas formas de realización puede ser diseñada o formulada de manera adicional como composiciones de liberación retardada o "liberación controlada".
- La invención provee composiciones de liberación retardada ("liberación controlada") que comprenden un ingrediente deseado recubierto por un revestimiento de polímero de látex (o equivalente). En un aspecto, el ingrediente deseado comprende una enzima, por ejemplo, una enzima de la invención. En un aspecto, una composición recubierta comprende un fármaco o composición farmacéutica. En un aspecto, el ingrediente deseado comprende una molécula pequeña, un fármaco, un polisacárido, un lípido, un ácido nucleico, una vitamina, un antibiótico o un insecticida. En un aspecto, el ingrediente deseado comprende una pella o una matriz, por ejemplo, una pella o una matriz que comprende un material comestible (por ejemplo, como alimento o pienso animal, o suplemento o medicamento). La divulgación también provee métodos para la "liberación controlada" o "liberación retardada" de una composición, en donde la composición es recubierta por un revestimiento de polímero de látex (o equivalente).
- En un aspecto, el revestimiento de polímero de látex comprende una pintura de látex, o equivalente. El revestimiento de polímero de látex puede comprender un (met)acrilato, un acetato de vinilo, un estireno, un etileno, un cloruro de vinilo, un butadieno, un cloruro de vinilideno, un versatato de vinilo, un propionato de vinilo, un acrilato de t-butilo, un acrilonitrilo, un neopreno, un maleato, un fumarato, equivalentes de los mismos, combinaciones de los mismos y/o derivados de los mismos.
- La divulgación provee métodos para la liberación retardada o liberación controlada de una composición que comprende: (i) (a) proveer una composición y, proveer un revestimiento de polímero de látex; y (b) recubrir la composición con el revestimiento de polímero de látex; (ii) el método de (i), en donde la composición comprende un fármaco o composición farmacéutica; o (iii) el método de (i) o (ii), en donde la composición comprende o consiste del polipéptido de la divulgación, o un polipéptido codificado por el ácido nucleico de la divulgación.
- La invención provee fluidos para perforación de pozos petrolíferos que comprenden los polipéptidos de la invención. La invención provee métodos para cambiar la viscosidad de una composición que comprenden: (i) (a) proveer una composición y un polipéptido de la invención y, proveer una composición; y (b) tratar la composición con un

polipéptido de la invención; o (ii) el método de (i), en donde la composición comprende un suelo o lodo de perforación.

5 La divulgación provee métodos para ayudar a llevarse el lodo de perforación que comprende: (a) proveer una composición y un polipéptido de la divulgación y un lodo de perforación; y (b) tratar el lodo de perforación con una composición que comprende un polipéptido de la divulgación.

La invención provee soluciones para bi blanqueamiento que comprenden un polipéptido de la invención. La invención provee métodos para bi blanqueamiento de una composición que comprende: (i) (a) proveer una composición y un polipéptido de la invención; y (b) tratar la composición con un polipéptido de la invención; o (ii) el método de (i), en donde la composición es un producto de papel o pulpa.

10 La invención provee métodos para elaborar un combustible que comprende: (i) (a) proveer un polipéptido de la invención; (b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; y (c) poner en contacto el polipéptido de (a) con la composición de (b) bajo condiciones en donde el polipéptido hidroliza el polisacárido, oligosacárido o almidón; o (ii) el método de (i) en donde el polipéptido es una enzima termoestable; (iii) el método de (i) en donde el combustible está basado en etanol.

15 La invención provee desinfectantes que comprenden un polipéptido de la invención.

La invención provee agentes de biodefensa o biodesintoxicación que comprenden un polipéptido de la invención.

La invención provee productos lácteos que comprenden un polipéptido de la invención.

20 La invención provee métodos para procesar un material de biomasa que comprende lignocelulosa que comprenden (i) (a) proveer una composición que comprende un polipéptido de la invención y, proveer un material de biomasa; y (b) poner en contacto la composición que comprende un polipéptido de la invención, con el material de biomasa; (ii) el método de (i), en donde el material de biomasa comprende o se deriva de un cultivo agrícola, o es un subproducto de un alimento o de la producción de pienso, o es un producto de desecho lignocelulósico, o es un residuo vegetal o un papel desechable o un producto de papel desechable; (iii) el método de (i) o (ii), en donde el polipéptido tiene actividad que comprende una actividad de amilasa; (iv) el método de cualquiera de (ii), en donde el residuo vegetal
25 comprende tallos, hojas, cáscaras, cortezas, mazorcas, madera, astillas de madera, pulpa de madera y aserrín, o, el residuo de papel comprende papel de fotocopia usado o descartado, papel para impresión por ordenador, papel de cuadernos, papel de anotadores, papel de mecanografía, diarios, revistas, cartulina y materiales de embalaje con base en papel; o (v) el método de (i) a (iv), en donde el procesamiento del material de biomasa genera un bioetanol, biopropanol, biobutanol, o un biodiesel; (vi) el método de (i) a (v), en donde el material de biomasa comprende una
30 lignocelulosa.

La invención provee materiales de biomasa que comprenden un polipéptido de la invención.

35 La divulgación provee métodos para elaborar bioetanol, biopropanol, biobutanol, o un biodiesel que comprenden: (i) (a) proveer un polipéptido de la invención, y, proveer una composición que comprende un polisacárido o oligosacárido; y (b) poner en contacto la composición que comprende un polisacárido u oligosacárido con un polipéptido de la invención; (ii) el método de (i), en donde la composición que comprende un polisacárido o oligosacárido comprende una planta, producto vegetal o derivado vegetal; (iii) el método de (ii), en donde la planta o producto vegetal comprende plantas o productos vegetales de caña de azúcar, remolacha o remolacha azucarera, trigo, maíz, soja, patata, arroz o cebada; (iv) el método de cualquiera de (i) a (iii), en donde el polipéptido tiene actividad que comprende actividad de amilasa; o (v) el método de cualquiera de (i) a (iv), en donde el polisacárido u
40 oligosacárido comprende un azúcar fermentable.

45 La invención provee métodos para elaborar un combustible que comprende (i) (a) proveer un polipéptido de la invención y, proveer una composición que comprende un azúcar fermentable; y (b) poner en contacto la composición que comprende un azúcar fermentable con un polipéptido de la invención; (ii) el método de (i), en donde la composición que comprende un azúcar fermentable comprende una planta, producto vegetal o derivado vegetal; (iii) el método de (ii), en donde la planta o producto vegetal comprende plantas o productos vegetales de caña de azúcar, remolacha o remolacha azucarera, trigo, maíz, soja, patata, arroz o cebada; (iv) el método de cualquiera de (i) a (iii), en donde el polipéptido tiene actividad que comprende una actividad de amilasa; o (v) el método de cualquiera de (i) a (iv), en donde el combustible comprende un bioetanol o una mezcla de gasolina-etanol, o comprende un bioetanol, biopropanol, biobutanol, o un biodiesel.

50 La divulgación provee combustibles que comprenden (a) un polipéptido de la invención; (b) el combustible de (a), en donde el polipéptido tiene actividad que comprende actividad de amilasa, glucoamilasa, glucosidasa, por ejemplo actividad de alfa-glucosidasa o beta-glucosidasa; (c) el combustible de (a) o (b), en donde el combustible deriva de un material vegetal, o el combustible se deriva de una patata, soja (semilla oleaginosa), cebada, centeno, maíz,

avena, trigo, remolacha o caña de azúcar; o (d) el combustible de cualquiera de (a) a (c), en donde el combustible comprende un bioetanol o una mezcla de gasolina-etanol, o comprende un bioetanol, biopropanol, biobutanol, o un biodiesel.

- 5 La invención provee métodos para producir un azúcar (por ejemplo, un monosacárido); el método comprendiendo: (i) (a) proveer al menos un polipéptido que tiene actividad de amilasa o de glucoamilasa; (b) proveer una composición que comprende un polisacárido o un oligosacárido; y (c) poner en contacto la composición de la etapa (b) con el polipéptido de la etapa (a), y así generar azúcares; (ii) el método de (i), en donde la composición que comprende un polisacárido o un oligosacárido comprende un almidón; (iii) el método de cualquiera de (i) o (ii), en donde el polisacárido, oligosacárido y/o azúcar comprende o es un azúcar fermentable; (iv) el método de cualquiera de (i), (ii),
10 o (iii), que comprende además fermentar el azúcar para producir un alcohol; o (v) el método de (iv), en donde el alcohol es un etanol, propanol o butanol.

Los detalles de una o más formas de realización de la invención se exponen en los dibujos adjuntos y la descripción que viene a continuación. Otras características, objetivos y ventajas de la invención serán evidentes a partir de la descripción y los dibujos y de las reivindicaciones.

- 15 Descripción de las figuras

La Figura 1 es un diagrama de bloques de un sistema de cómputo.

La Figura 2 es un diagrama de flujo que ilustra un aspecto de un proceso para comparar una nueva secuencia de nucleótido o proteína con una base de datos de secuencias a fin de determinar los niveles de homología entre la nueva secuencia y las secuencias en la base de datos.

- 20 La Figura 3 es un diagrama de flujo que ilustra un aspecto de un proceso en un ordenador para determinar si dos secuencias son homólogas.

La Figura 4 es un diagrama de flujo que ilustra un aspecto de un proceso identificador 300 para detectar la presencia de una característica en una secuencia.

- 25 La Figura 5 ilustra un ejemplo de un método de la invención para sacarificación por licuefacción, como se discute en detalle a continuación.

La Figura 6 ilustra un ejemplo de un proceso de molienda de maíz en húmedo de la invención usando al menos una enzima de la invención, como se discute en detalle a continuación.

La Figura 7, Figura 8 y Figura 9 ilustran ejemplos de métodos de procesamiento de un polisacárido, por ejemplo, almidón (por ejemplo, procesos industriales), como se discute en detalle a continuación.

- 30 La Figura 10 y la Figura 11 ilustran la influencia del pH en el rango de entre aproximadamente pH 3,5 a 6,0 sobre la hidrólisis de almidón para siete (7) glucoamilasas y dos (2) amilasas de esta divulgación, como se discute en detalle a continuación.

- 35 La Figura 12 ilustra cómo las enzimas de la invención se pueden usar en la producción de etanol a partir de maíz mediante molienda en seco, incluyendo su uso tanto en "procesos convencionales" como en los "procesos simultáneos de licuefacción, sacarificación y fermentación", como se discute en detalle a continuación.

La Figura 13, la Figura 14, la Figura 15, la Figura 16, la Figura 17, la Figura 18, la Figura 19 y la Figura 20, ilustran comparaciones de velocidad de ejemplos de enzimas de la divulgación en diferentes reguladores de almacenamiento, como se discute en detalle en el Ejemplo 1, a continuación.

- 40 La Figura 21 ilustra un SDS PAGE que muestra los resultados de proteólisis del ejemplo de la enzima de la SEQ ID NO: 52 mediante pepsina; y la Figura 22 ilustra la caracterización de los péptidos generados en esta digestión con pepsina de la SEQ ID NO: 52; y la Figura 23A ilustra el esquema de aislamiento del péptido pequeño usado para los péptidos generados por la proteólisis de la SEQ ID NO: 52 mediante pepsina; La Figura 23B ilustra un SDS PAGE de los resultados del esquema de aislamiento del péptido pequeño; la Figura 23C ilustra el perfil de LC/MS de la fracción eluida C18 RP; la Figura 23D ilustra la secuencia de los péptidos identificados por el análisis de LC MS/MS;
45 la Figura 23E y la Figura 23F ilustran los sequins (motivos) "Asn-Xaa-Ser/Thr" en la producción de la secuencia (resaltados en azul); las asparaginas que se predice que están N-glicosiladas están resaltadas en rojo, como se discute en detalle en el Ejemplo 22, a continuación.

La Figura 24 ilustra la Tabla 1, que resume los datos que comparan la velocidades iniciales de hidrólisis del almidón de maíz granulado y del almidón soluble (dextrina) mediante ejemplos de enzimas de la divulgación y una enzima glucoamilasa de *A. niger* como punto de referencia, como se discute en detalle en el Ejemplo 23, a continuación.

5 La Figura 25 ilustra el efecto de la temperatura sobre la actividad del ejemplo de glucoamilasa de la SEQ ID NO: 20 y la enzima como "punto de referencia" de *A. niger* con almidón soluble (dextrina) como sustrato, como se discute en detalle en el Ejemplo 23, a continuación.

10 La Figura 26 ilustra las tablas que resumen la eficiencia de los protocolos de purificación de la enzima de ejemplo sobre las enzimas seleccionadas de esta divulgación y los correspondientes datos de actividad sobre, entre otros, el almidón sin purificar y almidón soluble comparando la enzima purificada y no purificada, como se discute en detalle en el Ejemplo 18, a continuación.

La Figura 27 ilustra la base teórica del procedimiento de ensayo de α -amilasa CERALFA^{MR}, como se discute en detalle en el Ejemplo 27, a continuación.

15 La Figura 28A muestra los datos que demuestran el efecto de la temperatura sobre la actividad de por ejemplo, un ejemplo de glucoamilasa de la divulgación con almidón granulado como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

La Figura 28B muestra los datos que demuestran el efecto de la temperatura sobre la actividad de por ejemplo, un ejemplo de glucoamilasa de la divulgación con almidón soluble (dextrina) como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

20 La Figura 29A muestra los datos que demuestran el efecto del pH sobre la actividad de por ejemplo, un ejemplo de glucoamilasa de la divulgación con almidón granulado como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

La Figura 29B muestra los datos que demuestran el efecto del pH sobre la actividad de por ejemplo, un ejemplo de glucoamilasa de la divulgación con almidón soluble (dextrina) como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

25 La Figura 30 muestra los datos que demuestran el efecto de la temperatura sobre la hidrólisis de almidón por las α -amilasas caracterizadas, como se discute en detalle en el Ejemplo 29, a continuación.

La Figura 31A muestra los datos que demuestran el efecto del pH sobre la actividad de los ejemplos de α -amilasas y/o glucoamilasas de la divulgación con almidón granulado como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

30 La Figura 31B muestra los datos que demuestran el efecto del pH sobre la actividad de los ejemplos de α -amilasas y/o glucoamilasas de la divulgación con almidón soluble como sustrato, como se discute en detalle en el Ejemplo 29, a continuación.

Símbolos de referencia iguales en las diferentes figuras indican elementos iguales.

Descripción detallada

35 La invención provee enzimas amilasa, por ejemplo una alfa amilasa, polinucleótidos que codifican las enzimas, métodos para elaborar y usar estos polinucleótidos y polipéptidos. La invención está dirigida a polipéptidos novedosos que tienen una actividad de amilasa, por ejemplo, una actividad de alfa amilasa y ácidos nucleicos que los codifican. Los polipéptidos de la invención pueden ser usados en una variedad de contextos de diagnóstico, terapéutico, e industriales. Los polipéptidos de la invención se pueden usar, por ejemplo, como un aditivo para un detergente, para procesar alimentos y para síntesis química utilizando una reacción inversa. Adicionalmente, los polipéptidos de la invención pueden ser usados en el tratamiento de textiles, producción de alcohol y como aditivos para alimentos o pienso animal.

45 En un aspecto, la amilasa de la invención es activa a alta y/o baja temperatura, o en un amplio rango de temperatura. Por ejemplo, puede ser activa en las temperaturas que oscilan entre 20 °C y 90 °C, entre 30 °C y 80 °C, o entre 40 °C y 70 °C. La invención también provee amilasas que tiene actividad en pH alcalinos o en pHs ácidos, por ejemplo, baja acides del agua. En aspectos alternativos, las Amilasas de la invención pueden tener actividad en pH ácidos tan bajos como pH 5,5, pH 5,0, pH 4,5, pH 4,0 y pH 3,5. En aspectos alternativos, las Amilasas de la invención pueden tener actividad en pH alcalinos tan altos como pH 8, pH 9,5, pH 10, pH 10,5 y pH 11. En un aspecto, las Amilasas de la invención son activas en el rango de temperatura de entre aproximadamente 40 °C y 50 aproximadamente 70 °C bajo condiciones de baja actividad en agua (bajo contenido de agua). Por ejemplo, la

invencción provee amilasas con la capacidad de hidrolizar un polisacárido, oligosacárido o almidón, por ejemplo, un almidón granulado (inclusive almidón granulado sin purificar), a bajas temperaturas, por ejemplo, en el rango de aproximadamente 30 °C a 40 °C; a bajo pH, por ejemplo, en el rango de aproximadamente pH 3,5 a pH 6,0; y, a bajas temperaturas y bajo pH, por ejemplo, en el rango de aproximadamente 30 °C a 40 °C y a bajo pH en el rango de aproximadamente pH 3,5 a pH 6,0.

La divulgación también provee métodos para modificar adicionalmente las Amilasas de ejemplo de la divulgación para generar proteínas con propiedades deseables. Por ejemplo, las amilasas generadas por los métodos de la divulgación pueden tener actividad enzimática, estabilidad térmica, pH/ perfil de actividad, pH/perfil de estabilidad (tal como mayor estabilidad a valores de pH bajos, por ejemplo pH<6 o pH<5, o altos, por ejemplo pH>9), estabilidad a la oxidación, dependencia de Ca²⁺, actividad específica alterados y similares. La divulgación permite alterar cualquier propiedad de interés. Por ejemplo, la alteración puede resultar en una variante que, comparada con una enzima madre, tiene perfiles de actividad enzimática alterada, o de actividad de pH o temperatura alterados.

Generación y manipulación de ácidos nucleicos

En un aspecto, la invencción provee ácidos nucleicos aislados, sintéticos o recombinantes, que comprenden una secuencia de ácido nucleico que tiene al menos 98%, 99%, o más, o identidad de secuencia completa (100%) con un ácido nucleico de la invencción. En un aspecto, el ácido nucleico codifica al menos un polipéptido que tiene actividad de alfa amilasa.

Los ácidos nucleicos (inclusive oligonucleótidos), polipéptidos o proteínas "sintéticos" de la invencción incluyen aquellos preparados mediante cualquier síntesis química, por ejemplo, como se describe a continuación.

Las expresiones "ácido nucleico" o "secuencia de ácido nucleico" incluyen oligonucleótidos, nucleótidos, polinucleótidos, o un fragmento de cualquiera de los mismos, ADN o ARN (por ejemplo, ARNm, ARNr, ARNt) de origen genómico, recombinante o sintético que pueden ser monocatenarios o bicatenarios y pueden representar una hebra sentido o antisentido, para el ácido nucleico peptídico (PNA), o para cualquier material del tipo ADN o ARN, de origen natural o sintético, incluyendo, por ejemplo, ARNi tal como ARNmi o ARNsi, ribonucleoproteínas (por ejemplo, iRNP). El término abarca ácidos nucleicos, es decir, oligonucleótidos que contienen análogos conocidos de nucleótidos naturales. El término también abarca estructuras del tipo de ácido nucleico con estructuras sintéticas, véase por ejemplo, Mata (1997) Toxicol. Appl. Pharmacol. 144: 189 - 197; Strauss- Soukup (1997) Biochemistry 36: 8692 - 8698; Samstag (1996) Antisense Nucleic Acid Drug Dev 6: 153 - 156.

La invencción provee ácidos nucleicos "recombinantes" y sintéticos que pueden incluir ácidos nucleicos adyacentes al ácido nucleico "estructural" al cual no es adyacente en su ambiente natural. En un aspecto, los ácidos nucleicos representan el 5% o más del número de insertos de ácido nucleico en una población de "moléculas estructurales" de ácido nucleico. Las "moléculas estructurales" de acuerdo con la divulgación incluyen ácidos nucleicos tales como vectores de expresión, ácidos nucleicos de autorreplicación, virus, ácidos nucleicos integradores y otros vectores o ácidos nucleicos usados para mantener o manipular un inserto de ácido nucleico de interés. En un aspecto, los ácidos nucleicos enriquecidos representan el 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, o más del número de insertos de ácido nucleico en la población de moléculas estructurales recombinantes. Los polipéptidos o proteínas "recombinantes" se refieren a los polipéptidos o proteínas producidos por técnicas de ADN recombinante; por ejemplo, producidos a partir de células transformadas por un constructo de ADN exógeno que codifica el polipéptido o proteína deseados.

"Oligonucleótido" incluye ya sea un polidesoxinucleótido monocatenario o dos cadenas de polidesoxinucleótido complementarias, que pueden ser sintetizadas químicamente (es decir, como ácidos nucleicos sintéticos). En formas de realización alternativas, los ácidos nucleicos y oligonucleótidos sintéticos de la invencción no tienen fosfato 5' y, por lo tanto, no se ligan a otro oligonucleótido sin agregar un fosfato con ATP en presencia de una quinasa. En formas de realización alternativas, un oligonucleótido sintético se puede ligar a un fragmento que no ha sido desfosforilado.

El término "gen" incluye una secuencia de ácido nucleico que comprende un segmento de ADN involucrado en la producción de un producto de transcripción (por ejemplo, un mensaje), que a su vez es traducido para producir una cadena de polipéptido, o regula la transcripción, reproducción o estabilidad génica. Los genes pueden incluir regiones precedentes o posteriores a la región codificadora, tales como guías y remolque, promotores y potenciadores, así como también, donde sea aplicable, secuencias de intervinientes (intrones) entre segmentos de codificación individuales (exones).

La invencción provee ácidos nucleicos aislados y recombinantes, incluyendo casetes de expresión tales como vectores de expresión que codifican los polipéptidos de la invencción. La divulgación provee sondas que comprenden

o consisten de los ácidos nucleicos de la invención. La divulgación también incluye métodos para descubrir nuevas secuencias de amilasa usando los ácidos nucleicos de la divulgación. La divulgación también incluye métodos para inhibir la expresión de genes de amilasa, transcritos y polipéptidos usando los ácidos nucleicos de la divulgación. También se proveen métodos para modificar los ácidos nucleicos de la divulgación, por ejemplo, mediante reensamblaje de ligación sintética, sistema de evolución dirigida optimizada y/o mutagénesis de saturación del sitio génico (GSSM).

Los ácidos nucleicos de la invención pueden ser elaborados, aislados y/o manipulados, por ejemplo, mediante clonación y expresión de bibliotecas de ADNc, amplificación de ADN mensajero o genómico por PCR y similares. Al llevar a la práctica los métodos de la divulgación, se pueden modificar los genes homólogos mediante manipulación de un ácido nucleico plantilla, como se describe aquí. La invención se puede llevar a la práctica junto con cualquier método o protocolo o dispositivo conocidos en el arte, que se describen en la literatura científica y de patentes.

Técnicas generales

Los ácidos nucleicos usados para llevar a la práctica esta invención, ya sea ARN, ARNi, ácido nucleico antisentido, ADNc, ADN genómico, vectores, virus o híbridos de los mismos, pueden ser aislados a partir de una variedad de fuentes, modificadas genéticamente, amplificados, y/o expresados / generados de forma recombinante. Los polipéptidos recombinantes generados a partir de estos ácidos nucleicos pueden ser aislados o clonados en forma individual y evaluados en cuanto a una actividad deseada. Se puede usar cualquier sistema de expresión recombinante, incluyendo sistemas de expresión de células de bacterianas, de mamífero, levadura, insecto o planta.

La divulgación provee métodos para optimizar enzimas de amilasa y secuencias de ácido nucleico que codifican enzimas, por ejemplo, elaborando "variantes de secuencias", que comprenden el uso de secuencias de la invención usando, por ejemplo, "mutagénesis de saturación" o "GSSM" (incluye un método que usa cebadores de oligonucleótidos degenerados para introducir mutaciones puntuales en un polinucleótido, como se describe en detalle a continuación); "sistema de evolución dirigida optimizada" o "evolución dirigida optimizada" (incluye un método para reensamblar fragmentos de secuencias de ácido nucleico relacionadas, por ejemplo, genes relacionados y que se explican en detalle a continuación) y/o "reensamblaje de ligación sintética" o "SLR" (incluye un método para ligar fragmentos de oligonucleótidos en forma no estocástica y que se explican en detalle, a continuación). "Variante" incluye polinucleótidos o polipéptidos de la divulgación modificados en uno o más pares de bases, codones, intrones, exones, o residuos de aminoácidos (respectivamente) que aún retienen la actividad biológica de una amilasa de la divulgación. Las variantes pueden ser producidas por cualquier cantidad de medios, incluyendo métodos tales como, por ejemplo, PCR propensa a error, arrastre, mutagénesis dirigida al oligonucleótido, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casetes, mutagénesis de ensamblaje recursivo, mutagénesis de ensamblaje exponencial, mutagénesis específica del sitio, reensamblaje génico, GSSM y cualquier combinación de los mismos. En la presente invención se proveen técnicas para producir variantes de amilasa que tienen actividad a un pH o temperatura, por ejemplo, que es diferente de un ejemplo de enzima de esta divulgación, o una amilasa de tipo silvestre.

Los ácidos nucleicos de la invención pueden ser completa o parcialmente sintéticos y, en aspectos alternativos, pueden ser sintetizados *in vitro* mediante técnicas conocidas de síntesis química, como se describe, por ejemplo, en Adams (1983) J. Am. Chem. Soc. 105: 661; Belousov (1997) Nucleic Acids Res. 25: 3440 - 3444; Frenkel (1995) Free Radic. Biol. Med. 19: 373 - 380; Blommers (1994) Biochemistry 33: 7886 - 7896; Narang (1979) Meth. Enzymol. 68: 90; Brown (1979) Meth. Enzymol. 68: 109; Beaucage (1981) Tetra. Lett. 22: 1859; patente de los Estados Unidos No. 4.458.066.

Las técnicas para manipulación de ácidos nucleicos, tales como, por ejemplo, subclonación, sondas de marcación (por ejemplo, marcación aleatoria de cebadores usando polimerasas Klenow, traducción de muesca, amplificación), secuenciación, hibridación y similares, están bien descritas en la literatura científica y de patentes, véase, por ejemplo, Sambrook, ed., Molecular Cloning: A Laboratory Manual (2da Ed.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989); Current Protocols In Molecular Biology, Ausubel, ed. John Wiley & Sons, Inc., New York (1997); Laboratory Techniques in Biochemistry and Molecular Biology: Hybridization with Nucleic Acid Probes, Parte I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.Y. (1993).

Otros medios útiles para obtener y manipular ácidos nucleicos usados para llevar a la práctica los métodos de la invención consisten en clonar muestras genómicas y, si se desea, cribar y clonar nuevamente los insertos aislados o amplificados, por ejemplo, de clones genómicos o clones de ADNc. Las fuentes de ácido nucleico usadas en los métodos de la invención incluyen bibliotecas genómicas o de ADNc contenidas, por ejemplo, en cromosomas artificiales de mamíferos (MAC), véase, por ejemplo, las patentes de los Estados Unidos Nos. 5.721.118; 6.025.155; cromosomas artificiales humanos, véase, por ejemplo, Rosenfeld (1997) Nat. Genet. 15: 333 - 335; cromosomas artificiales de levaduras (YAC); cromosomas artificiales bacterianos (BAC); cromosomas artificiales P1, véase, por ejemplo, Woon (1998) Genomics 50: 306 - 316; vectores derivados de P1 (PAC), véase, por ejemplo, Kern (1997) Biotechniques 23: 120 - 124; cósmidos, virus recombinantes, fagos o plásmidos.

En un aspecto, un ácido nucleico que codifica un polipéptido de la invención es ensamblado en fase apropiada con una secuencia guía capaz de dirigir la secreción del polipéptido traducido o un fragmento del mismo.

La divulgación provee proteínas de fusión y ácidos nucleicos que las codifican. Un polipéptido de la divulgación puede ser fusionado con un péptido o polipéptido heterólogo, tal como péptidos de identificación del terminal N que imparten las características deseadas, tales como mayor estabilidad o purificación simplificada. Los péptidos y polipéptidos de la divulgación también pueden ser sintetizados y expresados como proteínas de fusión con uno o más dominios adicionales ligados a los mismos, por ejemplo, para producir un péptido más inmunogénico, para aislar más rápidamente un péptido sintetizado de manera recombinante, para identificar y aislar anticuerpos y células B que expresan anticuerpos, y similares. Los dominios que facilitan la detección y purificación incluyen, por ejemplo, péptidos quelantes de metal tales como trectos de polihistidina y módulos de histidina-triptófano que permiten la purificación sobre metales inmovilizados, dominios de proteína A que permiten la purificación en inmunoglobulina inmovilizada y el dominio utilizado en el sistema de purificación de extensión / afinidad FLAGS (Immunex Corp, Seattle, WA). La inclusión de secuencias de enlace que pueden ser escindidas tales como el Factor Xa o enteroquinasa (Invitrogen, San Diego, CA) entre un dominio de purificación y el péptido o polipéptido que comprende el motivo para facilitar la purificación. Por ejemplo, un vector de expresión puede incluir una secuencia de ácido nucleico que codifica el epítipo enlazada a seis residuos de histidina seguido por una tiorredoxina y un sitio de escisión de enteroquinasa (véase por ejemplo, Williams (1995) *Biochemistry* 34: 1787 - 1797; Dobeli (1998) *Protein Expr. Purif.* 12: 404 - 414). Los residuos de histidina facilitan la detección y la purificación mientras que el sitio de escisión de enteroquinasa provee un medio para purificar el epítipo del resto de la proteína de fusión. La tecnología relacionada con los vectores que codifican las proteínas de fusión y la aplicación de las proteínas de fusión se describen bien en la literatura científica y de patentes, véase por ejemplo, Kroll (1993) *DNA Cell Biol.*, 12: 441 - 53.

Secuencias de control de transcripción y traducción

La invención provee secuencias de ácido nucleico (por ejemplo, ADN) de la invención enlazadas en forma operativa a la(s) secuencia(s) de control de expresión (por ejemplo, de transcripción o de traducción), por ejemplo, promotores o potenciadores, para dirigir o modular la síntesis/expresión de ARN. Los promotores que se usan para la práctica de esta invención incluyen todas las secuencias capaces de dirigir la transcripción de una secuencia codificadora en una célula, por ejemplo, una célula vegetal. Por lo tanto, los promotores usados en los constructos de la invención incluyen elementos de control de transcripción que actúan de forma cis y secuencias reguladoras que están implicadas en la regulación o modulación de la duración y/o velocidad de transcripción de un gen. Por ejemplo, un promotor usado para la práctica de esta invención puede ser un elemento de control de la transcripción que actúa de forma cis, incluyendo un potenciador, un promotor, un terminador de la transcripción, un origen de replicación, una secuencia de integración de cromosomas, regiones no traducidas 5' y 3', o una secuencia intrónica, que participan en la regulación de la transcripción. Estas secuencias que actúan de forma cis por lo general interactúan con las proteínas u otras biomoléculas para llevar a cabo (activar/desactivar, regular, modular, etc.) la transcripción. Los promotores "constitutivos" usados para la práctica de esta invención pueden ser aquellos que dirigen la expresión de manera continua bajo la mayoría de las condiciones y estados ambientales de desarrollo o diferenciación celular. Los promotores usados para la práctica de esta invención pueden ser promotores "inducibles" o "regulables" que dirigen la expresión del ácido nucleico de la invención bajo la influencia de condiciones ambientales o condiciones de desarrollo. Los ejemplos de las condiciones ambientales que pueden afectar la transcripción mediante promotores inducibles incluyen condiciones anaeróbicas, temperatura elevada, sequía, o la presencia de luz.

En una forma de realización, una secuencia promotora puede estar "operativamente enlazada a" una secuencia codificadora de la invención, por ejemplo, cuando la ARN polimerasa que inicia la transcripción en el promotor transcribirá la secuencia codificadora en ARNm.

Los promotores usados para la práctica de esta invención incluyen promotores "específicos del tejido", que son elementos de control de la transcripción que solamente son activos en células, tejidos u órganos particulares, por ejemplo, en plantas o animales. La regulación específica del tejido puede ser lograda mediante ciertos factores intrínsecos que aseguran que los genes que codifican proteínas específicas para un determinado tejido sean expresados. Se sabe que dichos factores existen en mamíferos y plantas a fin de permitir que los tejidos específicos se desarrollen.

La(s) secuencia(s) de control de expresión usados para la práctica de esta invención pueden estar en un vector de expresión. Los ejemplos de promotores bacterianos incluyen *lacZ*, T3, T7, *gpt*, *lambda PR*, μ L y *trp*. Los ejemplos de promotores eucariotas incluyen al promotor inmediato temprano del CMV, la timidina quinasa del HSV, temprano y tardío del SV40, LTR de retrovirus y metalotioneína I de ratón. Los promotores adecuados para expresar un polipéptido en las bacterias incluyen los promotores *lac* o *trp* de *E. coli*, el promotor *lacI*, el promotor *lacZ*, el promotor T3, el promotor T7, el promotor *gpt*, el promotor *lambda PR*, el promotor *lambda* μ L, promotores de operones que codifican enzimas glicolíticas tales como 3-fosfoglicerato quinasa (PGK), y el promotor de fosfatasa ácida. Los promotores eucariotas incluyen al promotor inmediato temprano del CMV, al promotor de timidina quinasa del HSV, promotores de choque térmico, al promotor temprano y tardío del SV40, LTR de retrovirus y al promotor de

metalotioneína-I de ratón. También se pueden usar otros promotores que se sabe que controlan la expresión de genes en células procariotas y eucariotas o sus virus.

Promotores vegetales específicos del tejido

5 La invención provee casetes de expresión que pueden ser expresados en una forma específica del tejido, por ejemplo, que pueden expresar una amilasa de la invención en una forma específica del tejido. La divulgación también provee plantas o semillas que expresan una amilasa de la invención en una forma específica del tejido. La especificidad del tejido puede ser específica de la semilla, específica del tallo, específica de la hoja, específica de la raíz, específica del fruto, y similares.

10 En un aspecto, un promotor constitutivo tal como el promotor 35S del CaMV puede ser usado para expresión en partes específicas de la planta o la semilla, o en toda la planta. Por ejemplo, para sobreexpresión, se puede emplear un fragmento del promotor vegetal que dirigirá la expresión de un ácido nucleico en algunos o en todos los tejidos de una planta, por ejemplo, una planta regenerada. Dichos promotores son denominados aquí promotores "constitutivos" y son activos bajo la mayoría de las condiciones y estados ambientales de desarrollo o diferenciación celular. Los ejemplos de promotores constitutivos incluyen la región de iniciación de transcripción de 35S del virus del mosaico de la coliflor (CaMV), el promotor 1' o 2' derivado de T-ADN de *Agrobacterium tumefaciens* y otras regiones de iniciación de transcripción de diferentes genes vegetales conocidos para la persona versada en la materia. Dichos genes incluyen, por ejemplo, *ACT11* de *Arabidopsis* (Huang (1996) Plant Mol. Biol. 33: 125 - 139); *Cat3* de *Arabidopsis* (GenBank No. U43147, Zhong (1996) Mol. Gen. Genet. 251: 196 - 203); el gen que codifica la proteína desaturasa portadora de esteroil-acilo de *Brassica napus* (GenBank No. X74782, Solocombe (1994) Plant Physiol. 104: 1167 - 1176); *Gpc1* de maíz (GenBank No. X15596; Martinez (1989) J. Mol. Biol. 208: 551 - 565); el *Gpc2* de maíz (GenBank No. U45855, Manjunath (1997) Plant Mol. Biol. 33: 97 - 112); promotores vegetales descritos en las patentes de los Estados Unidos Nos. 4.962.028 y 5.633.440.

25 La invención usa promotores específicos del tejido o constitutivos derivados de virus que pueden incluir, por ejemplo, al promotor subgenómico de tobamovirus (Kumagai (1995) Proc. Natl. Acad. Sci. EUA 92: 1679 - 1683; el virus baciliforme del tungro del arroz (RTBV), que se replica solamente en células del floema en plantas de arroz infectadas, con su promotor que dirige la expresión fuerte del gen reportero específico de floema; el promotor del virus del mosaico de la vena de la mandioca (CVMV), con la actividad más alta en elementos vasculares, en células mesófilas de la hoja y en las puntas de las raíces (Verdaguer (1996) Plant Mol. Biol. 31: 1129 - 1139).

30 En forma alternativa, los promotores de las plantas pueden dirigir la expresión del ácido nucleico que expresa amilasa en un tejido, órgano o tipo de célula específico (es decir, promotores específicos del tejido) o, bien puede estar bajo control ambiental o de desarrollo más preciso o bajo el control de un promotor inducible. Los ejemplos de condiciones ambientales que pueden afectar la transcripción incluyen condiciones anaeróbicas, temperatura elevada, la presencia de luz, o asperjado con compuestos químicos/hormonas. Por ejemplo, la invención incorpora el promotor inducible por sequía del maíz (Busk (1997) ver más arriba); el promotor inducible por frío, sequía y alto contenido de sal de la patata (Kirch (1997) Plant Mol. Biol. 33: 897 - 909).

40 Los promotores específicos del tejido pueden promover la transcripción únicamente dentro de un cierto marco de tiempo de la etapa de desarrollo en ese tejido. Véase, por ejemplo, Blazquez (1998) Plant Cell 10 : 791 - 800, que caracteriza al promotor del gen LEAFY de *Arabidopsis*. Véase también Cardon (1997) Plant J 12: 367 - 77, que describe el factor de transcripción SPL3, que reconoce un motivo de una secuencia conservada en la región promotora del gen AP1 A de identidad del meristemo floral de *A. thaliana*; y Mandel (1995) Plant Molecular Biology, Vol. 29. páginas 995 - 1004, que describe al promotor eIF4 del meristemo. Se pueden usar los promotores específicos del tejido que son activos durante todo el ciclo de vida de un tejido particular. En un aspecto, los ácidos nucleicos de la invención están operativamente enlazados a un promotor activo principalmente únicamente en células de fibra de algodón. En un aspecto, los ácidos nucleicos de la invención están operativamente enlazados a un promotor principalmente activo durante las etapas de elongación de las células de fibra de algodón, por ejemplo, como lo describe Rinehart (1996), ver más arriba. Los ácidos nucleicos pueden estar operativamente enlazados al promotor del gen FbI2A que es preferencialmente expresado en células de fibra de algodón (Ibidem). Véase también, John (1997) Proc. Natl. Acad. Sci. EUA 89 : 5769 - 5773; John y colaboradores, patentes de los Estados Unidos Nos. 5.608.148 y 5.602.321, que describen promotores específicos de la fibra de algodón y métodos para la construcción de plantas transgénicas de algodón. También se pueden utilizar promotores específicos de la raíz para expresar los ácidos nucleicos de la invención. Los ejemplos de promotores específicos de la raíz incluyen al promotor del gen para la alcohol deshidrogenasa (DeLisle (1990) Int. Rev. Cytol. 123: 39 - 60). Otros promotores que pueden ser usados para expresar los ácidos nucleicos de la invención incluyen, por ejemplo, promotores específicos del óvulo, específico del embrión, específicos del endospermo, específicos del integumento, específicos del recubrimiento de la semilla, o alguna combinación de los mismos; un promotor específico de la hoja (véase, por ejemplo, Busk (1997) Plant J. 11: 1285 - 1295, que describe un promotor específico de la hoja en maíz); el promotor ORF13 de *Agrobacterium rhizogenes* (que exhibe alta actividad en raíces, véase, por ejemplo, Hansen (1997) ver más arriba); un promotor específico del polen del maíz (véase, por ejemplo, Guerrero (1990) Mol. Gen. Genet. 224: 161 -168); se puede utilizar un promotor del tomate activo durante la maduración de la fruta, senescencia y abscisión

de las hojas y, en un menor grado, de las flores, (véase, por ejemplo, Blume (1997) Plant J. 12: 731 - 746); un promotor específico del pistilo del gen SK2 de la patata (véase, por ejemplo, Ficker (1997) Plant Mol. Biol. 35: 425 - 431); el gen Blec4 del guisante, que es activo en el tejido epidérmico de ápices del brote vegetativo y floral de alfalfa transgénica haciendo que sea una herramienta útil para dirigir la expresión de genes foráneos a la capa epidérmica de las fibras o brotes de crecimiento activo; el gen BEL1 específico del óvulo (véase, por ejemplo, Reiser (1995) Cell 83: 735 - 742. GenBank No. U39944); y/o, el promotor en Klee, patente de los Estados Unidos No. 5.589.583, que describe una región promotora de la planta que es capaz de conferir altos niveles de transcripción en tejido meristemático y/o células de rápida división.

En forma alternativa, se utilizan promotores vegetales que son inducibles después de exposición a hormonas de la planta, tales como auxinas, para expresar los ácidos nucleicos de la invención. Por ejemplo, la invención puede utilizar el fragmento del promotor E1 de elementos sensibles a la auxina (AuxRE) en la soja (*Glycine max* L.) (Liu (1997) Plant Physiol. 115: 397 - 407); el promotor GSTG de *Arabidopsis* sensible a la auxina (también sensible al ácido salicílico y al peróxido de hidrógeno) (Chen (1996) Plant J. 10: 955 - 966); el promotor parC inducible por auxina del tabaco (Sakai (1996) 37: 906 - 913); un elemento de respuesta a la biotina de la planta (Streit (1997) Mol. Plant Microbe Interact. 10: 933 - 937); y, al promotor sensible al ácido abscísico de la hormona del estrés (Sheen (1996) Science 274: 1900 - 1902).

Los ácidos nucleicos de la invención también pueden estar enlazados operativamente a los promotores vegetales que son inducibles después de exposición a reactivos químicos que pueden ser aplicados a la planta, tales como herbicidas o antibióticos. Por ejemplo, se puede usar el promotor In2-2 del maíz, activado por agentes de seguridad del herbicida bencenosulfonamida (De Veylder (1997) Plant Cell Physiol. 38: 568 - 577); la aplicación de diferentes agentes de seguridad del herbicida induce diferentes patrones de expresión génica, incluyendo expresión en la raíz, hidatodos y el meristema apical del brote. La secuencia codificadora puede estar bajo el control, por ejemplo, de un promotor inducible por tetraciclina, por ejemplo, como se describe con las plantas transgénicas de tabaco que contienen el gen para la arginina descarboxilasa de *Avena sativa* L. (avena) (Masgrau (1997) Plant J. 11: 465 - 473); o, un elemento sensible al ácido salicílico (Stange (1997) Plant J. 11: 1315 - 1324). Utilizando promotores inducidos químicamente (por ejemplo, inducidos por hormonas o por pesticidas) es decir, promotores sensibles a un compuesto químico que puede ser aplicado a la planta transgénica en el campo, se puede inducir la expresión de un polipéptido de la invención en una etapa particular de desarrollo de la planta. Por lo tanto, la invención también provee plantas transgénicas que contienen un gen inducible que codifica para polipéptidos de la invención cuyo rango de huésped se limita a especies vegetales objetivo, tales como maíz, arroz, cebada, trigo, patata, caña de azúcar, remolacha azucarera, u otros cultivos, inducibles en cualquiera etapa del desarrollo del cultivo.

La persona versada en el arte reconoce que un promotor vegetal específico del tejido puede dirigir la expresión de secuencias operativamente enlazadas en tejidos diferentes al tejido objetivo. Por lo tanto, un promotor específico del tejido es uno que dirige la expresión preferentemente en el tejido o tipo de célula objetivo, pero también puede conducir a alguna expresión en otros tejidos.

Los ácidos nucleicos de la invención también pueden estar operativamente enlazados a promotores vegetales que son inducibles después de exposición a reactivos químicos. Estos reactivos incluyen, por ejemplo, herbicidas, auxinas sintéticas, o antibióticos que pueden ser aplicados, por ejemplo, asperjados, sobre plantas transgénicas. La expresión inducible de los ácidos nucleicos que producen amilasa de la invención le permitirá al agricultor seleccionar plantas con la proporción óptima de almidón / azúcar. De este modo se puede controlar el desarrollo de las partes de la planta. De esta forma, la invención provee los medios para facilitar la cosecha de las plantas y de las partes de la planta. Por ejemplo, en diferentes formas de realización, se usa el promotor In2-2 del maíz, activado por agentes de seguridad del herbicida bencenosulfonamida (De Veylder (1997) Plant Cell Physiol. 38: 568 - 577); la aplicación de diferentes agentes de seguridad del herbicida induce diferentes patrones de expresión génica, incluyendo la expresión en la raíz, hidatodos y el meristema apical del brote. Las secuencias codificadoras de la invención también están bajo el control del promotor inducible por tetraciclina, por ejemplo, como se describe con las plantas transgénicas de tabaco que contienen el gen para la arginina descarboxilasa de *Avena sativa* L. (avena) (Masgrau (1997) Plant J. 11: 465 - 473); o, un elemento sensible al ácido salicílico (Stange (1997) Plant J. 11: 315 - 1324).

Si se desea una expresión apropiada del polipéptido, una región se debe incluir una región de poliadenilación en el extremo 3' de la región codificadora. La región de poliadenilación se puede derivar del gen natural, de una variedad de otros genes vegetales, o de genes en el T- ADN Agrobacteriano.

La invención provee "casetes de expresión" que comprenden cualquier secuencia de la invención "operativamente enlazada" a un regulador de transcripción; por ejemplo, en donde "operativamente enlazada" como se utiliza aquí, se refiere a una relación funcional entre dos o más segmentos de ácido nucleico (por ejemplo, ADN). Por lo general, se refiere a la relación funcional de una secuencia reguladora de la transcripción con una secuencia transcrita. Por ejemplo, un promotor está operativamente enlazado a una secuencia codificadora, tal como un ácido nucleico de la invención, si estimula o modula la transcripción de la secuencia codificadora en una célula huésped apropiada u otro sistema de expresión. En general, las secuencias reguladoras de la transcripción del promotor que están

operativamente enlazadas a una secuencia transcrita están físicamente contiguas a la secuencia transcrita, es decir, están actuando en forma cis. Sin embargo, algunas secuencias reguladoras de transcripción, tales como potenciadores, no necesitan estar físicamente contiguas o ubicadas cerca a las secuencias codificadoras cuya transcripción ellas potencian.

5 Modificación de las secuencias codificadoras y de las secuencias adyacentes

La expresión transgénica en plantas de genes derivados de fuentes heterólogas puede implicar la modificación de aquellos genes para lograr y optimizar su expresión en plantas. En particular, los ORF bacterianos que codifican enzimas separadas pero que son codificados por el mismo transcripto en el microbio nativo son mejor expresados en plantas sobre transcriptos separados. Para lograr esto, cada ORF microbiano se aísla en forma individual y se clona en un casete que provee una secuencia del promotor vegetal en el extremo 5' del ORF y un terminador de transcripción de la planta en el extremo 3' del ORF. La secuencia del ORF aislada incluye preferiblemente el codón ATG de inicio y el codón de terminación pero puede incluir una secuencia adicional más allá del codón ATG de inicio y el codón STOP. Además, el ORF puede ser truncado pero aún retener la actividad requerida; para los ORF particularmente largos, las versiones truncadas que retienen actividad pueden ser preferibles para la expresión en organismos transgénicos. Por "promotor vegetal" y "terminador de la transcripción de la planta" se pretende hacer referencia a promotores y terminadores de la transcripción que operan en células vegetales. Esto incluye promotores y terminadores de la transcripción que pueden derivarse de fuentes no vegetales tales como virus (un ejemplo es el virus del mosaico de la coliflor).

En algunos casos, no es necesaria la modificación de las secuencias codificadoras del ORF y la secuencia adyacente. Es suficiente con aislar un fragmento que contiene el ORF de interés e insertarlo secuencia abajo de un promotor de la planta. Por ejemplo, Gaffney y colaboradores (Science 261: 754 - 756 (1993)) han expresado el gen *nahG* de *Pseudomonas* en plantas transgénicas bajo el control del promotor 35S del CaMV y el terminador *tml* del CaMV exitosamente sin modificación de la secuencia codificadora y con nucleótidos del gen de *Pseudomonas* secuencia arriba del ATG aún unido y nucleótidos secuencia abajo del codón STOP aún unido al ORF de *nahG*. Preferiblemente, ya que se debe dejar unida una secuencia microbiana adyacente pequeña secuencia arriba del codón ATG y secuencia abajo del codón STOP. En la práctica, tal construcción puede depender de la disponibilidad de los sitios de restricción.

En otros casos, la expresión de los genes derivados de fuentes microbianas puede causar problemas en la expresión. Estos problemas han sido bien caracterizados en el arte y son particularmente comunes con genes derivados de ciertas fuentes, tales como *Bacillus*. Estos problemas pueden aplicarse a la secuencia de nucleótidos de esta invención y la modificación de estos genes se puede realizar usando técnicas bien conocidas en el arte. Pueden encontrarse los siguientes problemas:

Uso de codones

El uso preferido de codones en plantas difiere del uso preferido de codones en ciertos microorganismos. La comparación del uso de codones dentro de un ORF microbiano clonado para el uso en genes vegetales (y en particular genes de la planta objetivo) permitirá la identificación de los codones dentro del ORF que preferiblemente deben ser cambiados. Típicamente, la evolución de la planta ha tendido hacia una preferencia fuerte de los nucleótidos C y G en la posición de la tercera base de las monocotiledóneas, mientras que las dicotiledóneas a menudo usan los nucleótidos A o T en esta posición. Mediante la modificación de un gene para incorporar el uso de codones preferidos para una especie transgénica particular objetivo, se superan muchos de los problemas descritos a continuación para el contenido de GC/AT y el empalme ilegítimo.

Contenido de GC/AT

Los genes de las plantas por lo general tienen un contenido de GC de más del 35%. Las secuencias del ORF que son ricas en nucleótidos A y T pueden causar varios problemas en las plantas. En primer lugar, se cree que los motivos de ATTTA causan desestabilización de los mensajes y se encuentran en el extremo 3' de muchos ARNm de corta vida. En segundo lugar, se cree que la ocurrencia de señales de poliadenilación tales como AATAAA en posiciones inapropiadas dentro del mensaje causan truncamiento prematuro de la transcripción. Además, las monocotiledóneas pueden reconocer secuencias ricas en AT como sitios de empalme (véase a continuación).

Secuencias adyacentes a la metionina de inicio

Las plantas difieren de los microorganismos en que sus mensajes no poseen un sitio de enlazamiento definido del ribosoma. Por el contrario, se cree que los ribosomas se unen al extremo 5' del mensaje y hacen un barrido para el primer ATG disponible en el cual se inicia la traducción. Sin embargo, se cree que existe una preferencia por ciertos nucleótidos adyacentes al ATG y esa expresión de genes microbianos puede ser reforzada por la inclusión de un iniciador de la traducción de consenso en eucariotas en el ATG. Clontech (catálogo 1993/1994, página 210) ha

sugerido una secuencia como iniciadora de la traducción de consenso para la expresión del gen *uidA* de *E. coli* en plantas. Además, Joshi, N.A.,R. 15: 6643 - 6653 (1987), ha comparado muchas secuencias de plantas adyacentes al ATG y sugiere otra secuencia de consenso. En situaciones donde se han encontrado dificultades en la expresión de los ORF microbianos en plantas, la inclusión de una de estas secuencias en el ATG de inicio puede mejorar la traducción. En dichos casos, los tres últimos nucleótidos del consenso pueden no ser apropiados para la inclusión en la secuencia modificada debido a su modificación del segundo residuo AA. Las secuencias preferentes adyacentes a la metionina de inicio pueden diferir entre diferentes especies de plantas. Un estudio de 14 genes de maíz localizados en la base de datos de GenBank suministró los siguientes resultados:

Posición antes del ATG de inicio en 14 genes de maíz:

	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1
C3	8	4	6	2	5	6	0	10	7	
T3	0	3	4	3	2	1	1	1	0	
A2	3	1	4	3	2	3	7	2	3	
G6	3	6	0	6	5	4	6	1	5	

Este análisis se puede realizar para las especies de plantas deseadas dentro de las cuales está siendo incorporada la secuencia de nucleótidos y la secuencia adyacente al ATG modificado para incorporar los nucleótidos preferidos.

Extracción de sitios de empalme ilegítimos

Los genes clonados de fuentes no vegetales y no optimizados para expresión en plantas también pueden contener motivos que pueden ser reconocidos en plantas como sitios de empalme 5' o 3' y ser escindidos, generando así, mensajes truncados o suprimidos. Estos sitios pueden ser removidos utilizando las técnicas conocidas en el arte.

Las técnicas para la modificación de secuencias codificadoras y de secuencias adyacentes son conocidas en el arte. En casos donde la expresión inicial de un ORF microbiano es baja y se considera apropiada para hacer alteraciones a la secuencia como se describió anteriormente, entonces se puede lograr la construcción de genes sintéticos de acuerdo con los métodos conocidos en el arte. Estos son descritos, por ejemplo, en las divulgaciones de las patentes publicadas EP 0 385 962 (de Monsanto), EP 0 359 472 (de Lubrizol) y la publicación WO 93/07278 (de Ciba-Geigy). En la mayoría de los casos, es preferible evaluar la expresión de las construcciones génicas usando protocolos de ensayos transitorios (que son bien conocidos en el arte) antes de su transferencia a plantas transgénicas.

Promotores vegetales

Las composiciones de la invención pueden contener secuencias de ácido nucleico para transformación y expresión en una planta de interés. Las secuencias de ácido nucleico pueden estar presentes en constructos de ADN o casetes de expresión. En formas de realización alternativas, "casete de expresión" como se utiliza aquí significa una molécula de ácido nucleico capaz de dirigir la expresión de una secuencia particular de nucleótidos en una célula huésped apropiada, que comprende un promotor operativamente enlazado a la secuencia de nucleótidos de interés, que está operativamente enlazada a señales de terminación. También comprende típicamente secuencias requeridas para una traducción apropiada de la secuencia de nucleótidos. La región codificadora usualmente codifica para una proteína de interés pero también puede codificar para un ARN funcional de interés, por ejemplo ARN antisentido o un ARN no traducido, en dirección sentido o antisentido. El casete de expresión que comprende la secuencia de nucleótidos de interés puede ser quimérico, lo que significa que al menos uno de sus componentes es heterólogo con respecto a al menos uno de sus otros componentes. El casete de expresión también puede ser uno que sea de origen natural pero que ha sido obtenido en una forma recombinante útil para expresión heteróloga. En las formas de realización alternativas, el casete de expresión es heterólogo con respecto al huésped, es decir, la secuencia de ADN particular del casete de expresión no se presenta naturalmente en la célula huésped y debe haber sido introducida en la célula huésped o en un ancestro de la célula huésped mediante un evento de transformación. La expresión de la secuencia de nucleótidos en el casete de expresión puede estar bajo el control de un promotor constitutivo o de un promotor inducible que inicia la transcripción únicamente cuando la célula huésped se expone a algunos estímulos externos particulares. En formas de realización alternativas, el promotor también puede ser específico para un tejido u órgano o estado de desarrollo particular.

La presente divulgación abarca la transformación de plantas con casetes de expresión capaces de expresar polinucleótidos. En formas de realización alternativas, el casete de expresión incluirá en la dirección 5'-3' de transcripción, una región de inicio de la transcripción y la traducción (es decir, un promotor) y un polinucleótido de interés. El casete de expresión puede comprender opcionalmente una región de terminación de la transcripción y la traducción (es decir, región de terminación) que actúa en las plantas. En algunas formas de realización, el casete de expresión comprende un gen marcador seleccionable para permitir la selección de transformantes estables. Los

constructos de expresión de la invención también pueden comprender una secuencia guía y/o una secuencia que permite la expresión inducible del polinucleótido de interés. Véase, Guo y colaboradores (2003) *Plant J.* 34: 383 - 92 y Chen y colaboradores (2003) *Plant J.* 36: 731 - 40 para ejemplos de secuencias que permiten la expresión inducible.

- 5 En formas de realización alternativas, las secuencias reguladoras del constructo de expresión están operativamente enlazadas al polinucleótido de interés. Por "operativamente enlazadas" se entiende un enlace funcional entre un promotor y una segunda secuencia en donde la secuencia del promotor inicia y media la transcripción de la secuencia de ADN correspondiente a la segunda secuencia. En formas de realización alternativas, "operativamente enlazada" significa que las secuencias de nucleótidos que son enlazadas son contiguas.
- 10 Cualquier promotor capaz de dirigir la expresión en una planta de interés puede ser usado en la práctica de la invención. El promotor puede ser nativo o análogo o foráneo o heterólogo con respecto al huésped de la planta. En formas de realización alternativas, los términos "heterólogo" y "exógeno" cuando se utilizan aquí para referirse a una secuencia de ácido nucleico (por ejemplo una secuencia de ADN o ARN) o un gen, se refiere a una secuencia que se origina a partir de una fuente externa a la célula huésped particular o, si es a partir de la misma fuente, se modifica a partir de su forma original. Por lo tanto, un gen heterólogo en una célula huésped incluye un gen que es endógeno a la célula huésped particular pero que ha sido modificado. Los términos también incluyen copias múltiples de origen no natural de una secuencia de ADN de origen natural. Por lo tanto, los términos se refieren a un segmento de ADN que es externo o heterólogo a la célula, u homólogo a la célula pero en una posición dentro del ácido nucleico de la célula huésped dentro del no se encuentra ordinariamente el elemento. Los segmentos de ADN
- 20 exógeno se expresan para producir polipéptidos exógenos.

En formas de realización alternativa, una secuencia de ácido nucleico "homóloga" (por ejemplo ADN) es una secuencia de ácido nucleico (por ejemplo ADN o ARN) naturalmente asociada con una célula huésped dentro de la cual se introduce.

- 25 La elección de promotores que van a ser incluidos puede depender de varios factores, incluyendo, pero sin limitarse a, eficiencia, capacidad de selección, capacidad de inducción, nivel deseado de expresión y expresión preferencial en el tejido o la célula. Es una cuestión de rutina para la persona capacitada en el arte modular la expresión de una secuencia mediante selección y posicionamiento apropiado de los promotores y otras regiones reguladoras relacionadas con esa secuencia.

- 30 En formas de realización alternativa, los promotores adecuados inician la transcripción únicamente, o predominantemente, en ciertos tipos celulares. Por lo tanto, como se utiliza aquí, un promotor preferencial del tipo de célula o tejido es uno que dirige la expresión preferencialmente en el tejido objetivo, pero que puede conducir también a alguna expresión en otros tipos de células o tejidos. Los métodos para identificar y caracterizar regiones promotoras en ADN genómico de plantas incluyen, por ejemplo, aquellos descritos en las siguientes referencias: Jordano, y colaboradores, *Plant Cell*, 1: 855 - 866 (1989); Bustos, y colaboradores, *Plant Cell*, 1: 839 - 854 (1989); Green, y colaboradores, *EMBO J.*, 7, 4035 - 4044 (1988); Meier, y colaboradores, *Plant Cell*, 3, 309 - 316 (1991); y Zhang, y colaboradores, *Plant Physiology* 110: 1069 - 1079 (1996).
- 35

- En formas de realización alternativa, se pueden utilizar promotores y/o genes regulados preferidos del tejido que han sido reportados en plantas. Los genes reportados preferidos del tejido que pueden ser usados en formas de realización alternativas incluyen los genes que codifican las proteínas de almacenamiento de las semillas (tales como napina, cruciferina, beta - conglucina y faseolina, prolaminas, glutelinas, globulinas y zeinas), zeinas o proteínas de cuerpos oleosos (tales como oleosina), o genes involucrados en la biosíntesis de ácidos grasos (incluyendo la proteína portadora de acilo, esteroil-ACP desaturasa y desaturasas de ácido graso (fad 2-1)) y otros genes expresados durante el desarrollo del embrión (tales como Bce4, véase, por ejemplo, el documento EP 255378 y Kridl y colaboradores, (1991) *Seed Science Research*, 1: 209). Los ejemplos de promotores específicos del tejido
- 40 que pueden ser utilizados para la práctica de esta invención, que han sido descritos, incluyen la lectina (Vodkin, *Prog. Clin. Biol. Res.*, 138; 87 (1983); Lindstrom y colaboradores, (1990) *Dev. Genet.*, 11: 160), alcohol deshidrogenasa 1 de maíz (Dennis y colaboradores, *Nucleic Acid Res.*, 12: 3983 (1984)), complejo de cosecha liviano de maíz (véase, por ejemplo, Simpson, (1986) *Science*, 233: 34; Bansal (1992) *Proc. Natl. Acad. Sci. EUA* 89:23654), proteína de choque térmico de maíz (véase, por ejemplo, Odell y colaboradores, (1985) *Nature*, 313: 810; subunidad pequeña de la RuBP carboxilasa de guisante (véase, por ejemplo, Poulsen y colaboradores, (1986) *Mol. Gen. Genet.*, 205: 193 -200; Cashmore y colaboradores, (1983) *Gen. Eng. of Plants*, Plenum Press, New York, 29 - 38); manopina sintasa de plásmido Ti (véase, por ejemplo, Langridge y colaboradores, (1989) *Proc. Natl. Acad. Sci. EUA*, 86: 3219 - 3223), nopalina sintasa de plásmido Ti (Langridge y colaboradores, (1989) *Proc. Natl. Acad. Sci. EUA*, 86: 3219 - 3223), chalcona isomerasa de petunia (véase, por ejemplo, vanTunen (1988) *EMBO J.* 7: 1257);
- 50 proteína 1 rica en glicina de frijol (véase, por ejemplo, Keller (1989) *Genes Dev.* 3: 1639); 35S truncado del CaMV (véase, por ejemplo, Odell (1.985) *Nature* 313: 810); patatina de patata (véase, por ejemplo, Wenzler (1989) *Plant Mol. Biol.* 13: 347; célula de raíz (véase, por ejemplo, Yamamoto (1990) *Nucleic Acid Res.*, 18: 7449); zeína de maíz (véase, por ejemplo, Reina (1990) *Nucleic Acid Res.*, 18: 6425; Lopes y colaboradores (1995) *Mol. Gen. Genet.* 247: 603 - 613; Kriz (1987) *Mol. Gen. Genet.* 207: 90; Wandelt (1989) *Nucleic Acid Res.*, 17: 2354; Langridge (1983) *Cell*,
- 55

34: 1015; Reina (1990) Nucleic Acid Res., 18: 7449), promotor gpp de ADP (véase, por ejemplo, patente de los Estados Unidos No. 7.102.057); globulina-1 (véase, por ejemplo, Belanger (1991) Genetics 129: 863); α -globulina (Sunilkumar, y colaboradores (2002). Transgenic Res., 11: 347 - 359); α -tubulina; cab (véase, por ejemplo, Sullivan (1989) Mol. Gen. Genet., 215: 431); PEPCasa (véase por ejemplo, Hudspeth & Grula, (1989) Plant Molec. Biol., 12: 579 - 589); promotores asociados al complejo del gen R (Chandler y colaboradores, (1989) Plant Cell, 1: 1175); promotor de vicilina de guisante (Czako y colaboradores, (1992) Mol. Gen. Genet., 235: 33; patente de los Estados Unidos No. 5.625.136); promotor GTL1 (Takaiwa y colaboradores (1991) Plant Mol. Biol. 16 (1), 49 - 58); promotores de la chalcona sintasa (Franken y colaboradores, (1991) EMBO J., 10: 2605); y/o un promotor GY1 (Sims Goldberg (1989) Nuc. Acid Res. 17(11) 4368) y similares.

10 En formas de realización alternativa, se utiliza una clase de promotores preferidos del fruto expresados en o durante la antétesis a través del desarrollo del fruto, al menos hasta el comienzo de la maduración, por ejemplo, como se discute en la patente de los Estados Unidos 4.943.674. El promotor para el gen de la poligalacturonasa es activo en la maduración del fruto. También se puede utilizar un gen para la poligalacturonasa, por ejemplo, como se describe en la patente de los Estados Unidos No. 4.535.060, la patente de los Estados Unidos No. 4.769.061, la patente de los Estados Unidos No. 4.801.590 y la patente de los Estados Unidos No. 5.107.065.

20 Otros ejemplos de promotores preferidos del tejido que pueden ser usados incluyen aquello que dirigen la expresión en células de la hoja después del daño a la hoja (por ejemplo, por insectos masticadores), en tubérculos (por ejemplo, promotor del gen de la patatina), y en células de fibra (un ejemplo de una proteína de célula de fibra regulada por desarrollo es E6 (John & Crow (1992) PNAS 89: 5769 - 5773). El gen E6 es más activo en fibra, aunque se encuentran bajos niveles de transcritos en la hoja, el óvulo y la flor.

25 En formas de realización alternativa, se pueden utilizar promotores activos en tejido fotosintético con el fin de dirigir la transcripción en tejidos verdes tales como hojas y vástagos; estos promotores son adecuados cuando dirigen la expresión únicamente o predominantemente en dichos tejidos. En formas de realización alternativa, el promotor puede conferir expresión constitutivamente a través de toda la planta, o diferencialmente con respecto a los tejidos verdes, o diferencialmente con respecto a la etapa de desarrollo del tejido verde en la cual ocurre la expresión, o en respuesta a estímulos externos.

30 Los ejemplos de promotores que pueden ser usados para llevar a cabo esta invención incluyen los promotores de ribulosa-1,5-bisfosfato carboxilasa (RbcS) tales como el promotor RbcS de alerce oriental (*Larix laricina*), el promotor cab6 de pino (Yamamoto y colaboradores (1994) Plant Cell Physiol. 35: 773 - 778), el promotor del gen Cab-1 de trigo (Fejes y colaboradores (1990) Plant Mol. Biol. 15: 921 - 932). el promotor CAB-1 de espinaca (Lubberstedt y colaboradores (1994) Plant Physiol. 104: 997 - 1006), el promotor cabIR del arroz (Luan y colaboradores (1992) Plant Cell 4: 971 - 981), el promotor de la piruvato ortofosfato diquinasa (PPDK) del maíz (Matsuoka y colaboradores (1993) Proc. Natl. Acad. Sci EUA 90: 9586 - 9590), el promotor Lhcb1*2 del tabaco (Cerdan y colaboradores (1997) Plant Mol. Biol. 33: 245 - 255), el promotor de SUC2 sacarosa-H⁺ symporter de *Arabidopsis thaliana* (Truernit y colaboradores (1995) Plant 196: 564 - 570) y promotores de proteína de membrana tilacoide de la espinaca (psaD, psaF, psaE, PC, FNR, atpC, atpD, cab, rbcS). Otros promotores que dirigen la transcripción en tallos, hojas y tejido verde se describen en la publicación de la patente de los Estados Unidos No. 2007/0006346.

40 La especificidad de tejido de algunos promotores "preferidos del tejido" puede no ser absoluta; en formas de realización alternativas se utilizan genes reporteros tales como Gus o proteína fluorescente verde, proteína ciano fluorescente, proteína fluorescente amarilla o proteína fluorescente roja. En formas de realización alternativas, se puede lograr la expresión preferida del tejido con expresión "defectuosa" por medio de una combinación de diferentes promotores preferidos del tejido. Se pueden utilizar otros promotores preferidos del tejido, y pueden ser aislados, por una persona versada en el arte (véase la patente de los Estados Unidos No. 5.589.379).

45 En un aspecto, se utilizan los promotores vegetales que son inducibles después de la exposición a hormonas de la planta, tales como auxinas, para expresar los ácidos nucleicos de la invención. Por ejemplo, la invención puede utilizar un fragmento del promotor E1 de elementos de respuesta a la auxina (AuxRE) en la soja (*Glycine max* L.) (Liu (1997) Plant Physiol. 115: 397 - 407); el promotor GST6 de *Arabidopsis* sensible a la auxina (también sensible al ácido salicílico y al peróxido de hidrógeno) (Chen (1996) Plant J. 10: 955 - 966); el promotor parC inducible por auxina de tabaco (Sakai (1996) 37: 906 - 913); un elemento de respuesta a la biotina de la planta (Streit (1997) Mol. Plant Microbe Interact. 10: 933 - 937); y el promotor sensible al ácido abscísico de la hormona del estrés (Sheen (1996) Science 274: 1900 - 1902).

55 Los ácidos nucleicos de la invención también pueden estar operativamente enlazados a promotores de la planta que son inducibles ante la exposición a reactivos químicos que pueden ser aplicados a la planta, tales como herbicidas o antibióticos. Por ejemplo, sistemas de expresión génica que son activados en presencia de un ligando químico, incluyendo etanol, tal como se puede encontrar en los documentos WO 96/27673; WO 93/01294; WO 94 /03619; WO 02/061102. Se puede usar el promotor In2-2 de maíz, activado por agentes de seguridad del herbicida bencenosulfonamida (De Veylder (1997) Plant Cell Physiol. 38: 568 - 577); la aplicación de diferentes agentes de seguridad del herbicida induce distintos patrones de expresión génica, incluyendo la expresión en la raíz, hidatodos

y el meristema apical del vástago. La secuencia codificadora puede estar bajo el control, por ejemplo, de un promotor inducible por tetraciclina, por ejemplo, como se describe con plantas transgénicas de tabaco que contienen el gen para la arginina decarboxilasa de *Avena sativa* L. (avena) (Masgrau (1997) Plant J. 11: 465 - 473); estrógeno, tal como, el receptor de ecdisona (WO 01/52620) o, un elemento sensible al ácido salicílico (Stange (1997) Plant J. 11: 1315 - 1324). Utilizando promotores inducidos químicamente (por ejemplo, inducidos por hormona o pesticida), es decir, promotores sensibles a un compuesto químico que puede ser aplicado a la planta transgénica en el campo, se puede inducir la expresión de un polipéptido de la invención en una etapa particular del desarrollo de la planta.

Los ejemplos de algunos promotores constitutivos que pueden ser usados para la práctica de esta invención y que han sido descritos, incluyen actina 1 de arroz (Wang y colaboradores (1992) Mol. Cell. Biol. 12: 3399; la patente de los estados Unidos No. 5.641.876); otras isoformas de actina (McEiroy y colaboradores (1990) Plant Cell 2: 163 - 171 y McEiroy y colaboradores (1991) Mol. Gen. Genet. 231: 150 - 160); 35S del CaMV (Odell y colaboradores (1985) Nature, 313: 810); 19S del CaMV (Lawton y colaboradores (1987) Plant Mol. Biol. 9: 315 - 324; la patente de los Estados Unidos No. 5.639.949); nos (Ebert y colaboradores (1987) PNAS EUA 84: 5745 - 5749); Adh (Walker y colaboradores (1987) PNAS EUA 84: 6624 - 6628), sacarosa sintasa (Yang & Russell (1990) PNAS EUA 87: 4144 - 4148); y los promotores de ubiquitina (por ejemplo, girasol Binet y colaboradores (1991) Plant Science 79: 87 - 94; maíz Christensen y colaboradores (1989) Plant Molec. Biol. 12: 619 - 632; y *Arabidopsis* - Callis y colaboradores, J. Biol. Chem. (1990) 265: 12486 - 12493; y Norris y colaboradores, Plant Mol. Biol. (1993) 21: 895 - 906.

Se pueden una variedad de terminadores de la transcripción en casetes de expresión para la realización de esta invención. Estos terminadores de la transcripción son responsables por la terminación de la transcripción más allá del transgén y la correcta poliadenilación del ARNm. La región de terminación puede ser nativa con la región de iniciación de la transcripción, puede ser nativa con la secuencia de ADN operativamente enlazada de interés, puede ser nativa con el huésped de la planta o se puede derivar de otra fuente (es decir, foránea o heteróloga al promotor, la secuencia de ADN de interés, el huésped de la planta o cualquier combinación de los mismos). Los terminadores apropiados de la transcripción son aquellos que se saben que actúan en plantas e incluyen al terminador 35S del CaMV, el terminador tml, el terminador de nopalina sintasa y el terminador E9 rbcS de guisante. Estos pueden ser utilizados tanto en monocotiledóneas como en dicotiledóneas. Además, se puede utilizar un terminador de la transcripción nativo del gen.

Se ha encontrado que numerosas secuencias refuerzan la expresión génica a partir de la unidad de transcripción y estas secuencias se pueden usar junto con los genes de esta invención para incrementar su expresión en plantas transgénicas. Por ejemplo, diferentes secuencias de intrones han mostrado que refuerzan la expresión, particularmente en células monocotiledóneas. Por ejemplo, se ha encontrado que los intrones del gen Adhl de maíz refuerzan de manera significativa la expresión del gen de tipo silvestre bajo su promotor cognado cuando se introducen en células de maíz.

En formas de realización alternativa, se pueden utilizar secuencias guía no traducidas derivadas de virus para reforzar la expresión, y estas son particularmente efectivas en células dicotiledóneas. En formas de realización alternativas, se pueden utilizar r secuencias guía del virus del mosaico del tabaco (TMV, la "secuencia W"), el virus del moteado clorótico del maíz (MCMV,) y virus del mosaico de la alfalfa (AMV) y estas han mostrado ser efectivas en el reforzamiento de la expresión (por ejemplo Gallie y colaboradores, Nucl. Acids Res. 15: 8693 - 8711 (1987); Skuzeski y colaboradores, Plant Molec. Biol. 15: 65 - 79 (1990)).

40 Direccionamiento del producto génico dentro de la célula

En formas de realización alternativa se utilizan diversos mecanismos para dirigir los productos génicos y se sabe que estos existen en las plantas; y se han caracterizado en forma detallada las secuencias que controlan el funcionamiento de estos mecanismos. Se han caracterizado secuencias que provocan el direccionamiento de productos génicos a otros compartimientos celulares. Las secuencias amino terminales pueden ser responsables por dirigir una proteína de interés a cualquier compartimiento celular, tal como, una vacuola, mitocondria, peroxisoma, cuerpos proteicos, retículo endoplasmático, cloroplasto, gránulo de almidón, amiloplasto, apoplasto o la pared celular de una planta (por ejemplo, Unger y colaboradores Plant Molec. Biol. 13: 411 - 418 (1989); Rogers y colaboradores (1985) Proc. Natl. Acad. Sci. EUA 82: 6512 - 651; patente de Estados Unidos No. 7.102.057; el documento WO 2005/096704). En formas de realización alternativa, se utilizan diferentes secuencias señal, por ejemplo, la secuencia señal puede ser una secuencia señal del terminal N de cera, una secuencia señal del terminal N de γ -zeína, un dominio de enlazamiento de almidón, un dominio de enlazamiento de almidón del terminal C, una secuencia de direccionamiento al cloroplasto, que importa la proteína madura al cloroplasto (Comai y colaboradores (1988) J. Biol. Chem. 263: 15104 - 15109; van den Broeck, y colaboradores (1985) Nature 313: 358 - 363; patente de los Estados Unidos No. 5.639.949) o una secuencia señal de secreción de las células de aleurona (Koehler & Ho, Plant Cell 2: 769 - 783 (1990)). En formas de realización alternativa, se utilizan secuencias del terminal amino junto con secuencias del terminal carboxilo; éstas son responsables del direccionamiento vacuolar de productos génicos (Shinshi y colaboradores (1990) Plant Molec. Biol. 14: 357 - 368).

La secuencia señal seleccionada puede incluir el sitio de escisión conocido, y la fusión construida debe tener en cuenta cualquiera de los aminoácido después del (de los) sitio(s) de escisión, que se requieren para la escisión. En algunos casos, este requerimiento puede ser cumplido por la adición de un pequeño número de aminoácidos entre el sitio de escisión y el ATG del transgén o, en forma alternativa, el reemplazo de algunos aminoácidos dentro de la secuencia del transgén. Estas técnicas de construcción son bien conocidas en el arte y son aplicables de igual forma a cualquier compartimiento celular.

Los mecanismos anteriormente descritos para direccionamiento celular pueden ser utilizados no solo junto con sus promotores cognados, sino también junto con promotores heterólogos para efectuar un objetivo de direccionamiento celular específico bajo la regulación transcripcional de un promotor que tiene un patrón de expresión diferente a aquel del promotor del cual se deriva la señal de direccionamiento.

Vectores de expresión y vehículos de clonación

La invención provee vectores de expresión y vehículos de clonación que comprenden ácidos nucleicos de la invención, por ejemplo, secuencias que codifican la amilasa de la invención. Los vectores de expresión y vehículos de clonación de la invención pueden comprender partículas virales, baculovirus, fagos, plásmidos, fagémidos, cósmidos, fósmodos, cromosomas artificiales bacterianos, ADN viral (por ejemplo, vacuna, adenovirus, virus de la viruela aviar, pseudorrabia y derivados del SV40), cromosomas artificiales basados en P1, plásmidos de levadura, cromosoma artificial de levadura y cualquier otro vector específico para los huéspedes específicos de interés (tales como *Bacillus*, *Aspergillus* y levadura). Los vectores de la invención pueden incluir secuencias de ADN sintéticas, cromosómicas y no cromosómicas. Aquellos capacitados en el arte conocen un gran número de vectores adecuados, y que se encuentran comercialmente disponibles. Los ejemplos de vectores incluyen: bacterianos: vectores pQE (Qiagen), plásmidos pBLUESCRIPT^{MR} (pBluescript), vectores pNH, (vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); eucariotas: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Sin embargo, cualquier otro plásmido o vector puede ser usado siempre que sea replicable y viable en el huésped. Se pueden emplear los vectores de número de copia bajo o número de copia alto con la presente invención. Los "plásmidos" se pueden encontrar comercialmente disponibles, públicamente disponibles en forma restringida, o pueden ser construidos a partir de plásmidos disponibles de acuerdo con los procedimientos publicados. Se conocen en el arte plásmidos equivalente a los descritos aquí y serán evidentes para aquellas personas versada en el arte.

La invención provee "casetes de expresión" que comprenden cualquier secuencia de la invención "operativamente enlazada" a un regulador de transcripción; el término "casete de expresión" como se usa aquí puede referirse a una secuencia de nucleótidos que es capaz de afectar la expresión de un gen estructural (por ejemplo, una secuencia que codifica a una proteína, tal como una amilasa de la invención) en un huésped compatible con tales secuencias. Los casetes de expresión incluyen al menos un promotor operativamente enlazado con la secuencia que codifica al polipéptido; y, opcionalmente, con otras secuencias, por ejemplo, señales de terminación de transcripción. También se pueden utilizar factores adicionales necesarios o útiles para efectuar la expresión, por ejemplo, reforzadores. Por lo tanto, los casetes de expresión también incluyen plásmidos, vectores de expresión, virus recombinantes, cualquier forma de un vector de "ADN desnudo" recombinante, y similares.

Un "vector" comprende un ácido nucleico que puede infectar, transfectar, transducir en forma transitoria o permanente una célula. Se reconocerá que un vector puede ser un ácido nucleico desnudo, o un ácido nucleico que forma un complejo con una proteína o un lípido. El vector comprende, opcionalmente, ácidos nucleicos virales o bacterianos, y/o proteínas y/o membranas (por ejemplo, una membrana celular, una envoltura de un lípido viral, etc.). Los vectores usados para la realización de esta invención esta invención incluyen, pero no se limitan a replicones (por ejemplo, replicones de ARN, bacteriófagos) a los cuales se pueden unir los fragmentos de ADN y replicarse. Los vectores usados para la realización de esta invención incluyen, pero no se limitan a ARN, ADN o ARN lineal o circular que se replica por sí mismo en forma autónoma (por ejemplo, plásmidos, virus y similares, véase, por ejemplo, la patente de los Estados Unidos No. 5.217.879) e incluyen tanto los plásmidos de expresión como los de no expresión. Cuando se describe un microorganismo recombinante o cultivo celular como hospedador de un "vector de expresión", este incluye tanto ADN circular y lineal extracromosómico como ADN que ha sido incorporado en el(los) cromosoma(s) huésped. Cuando un vector es mantenido por una célula huésped, el vector puede o bien ser replicado en forma estable por las células durante la mitosis como una estructura autónoma, o es incorporado en el genoma del huésped.

El vector de expresión puede comprender un promotor, un sitio de enlazamiento del ribosoma para inicio de traducción y un terminador de la transcripción. El vector también puede incluir secuencias apropiadas para amplificación de la expresión. Los vectores de expresión de mamíferos pueden comprender un origen de replicación, cualquiera de los sitios necesarios de enlazamiento del ribosoma, un sitio de poliadenilación, sitios aceptores y donantes de empalme, secuencias de terminación de la transcripción y secuencias no transcriptas que flanquean a 5'. En algunos aspectos, las secuencias de ADN derivadas de los sitios de empalme y poliadenilación de SV40 pueden ser usadas para proporcionar los elementos genéticos no transcritos requeridos.

En un aspecto, los vectores de expresión contienen uno o más genes marcadores seleccionables para permitir la selección de células huésped que contienen al vector. Tales marcadores seleccionables incluyen genes que codifican dihidrofolato reductasa o genes que confieren resistencia a la neomicina para el cultivo de células eucariotas, genes que confieren resistencia a tetraciclina o ampicilina en *E. coli* y el gen TRP1 de *S. cerevisiae*. Las regiones promotoras pueden ser seleccionadas a partir de cualquier gen deseado usando vectores de cloranfenicol transferasa (CAT) u otros vectores con marcadores seleccionables.

Los vectores para expresar el polipéptido o un fragmento del mismo en células eucariotas también pueden contener reforzadores para incrementar los niveles de expresión. Los reforzadores son elementos que actúan de forma cis de ADN, usualmente de aproximadamente 10 hasta aproximadamente 300 pb de longitud que actúan sobre un promotor para incrementar su transcripción. Los ejemplos incluyen al reforzador de SV40 en el lado tardío del origen de replicación de 100 a 270 pb, el reforzador del promotor temprano de citomegalovirus, el reforzador de poliovirus sobre el lado tardío del origen de replicación y los reforzadores del adenovirus.

Se puede insertar una secuencia de ácido nucleico en un vector mediante una variedad de procedimientos. En general, se liga la secuencia en la posición deseada en el vector después de la digestión del inserto y del vector con las endonucleasas de restricción apropiadas. En forma alternativa, se pueden ligar los extremos romos tanto del inserto como del vector. Se conocen en el arte una variedad de técnicas de clonación, por ejemplo, como se describe en Ausubel y Sambrook. Tales procedimientos y otros se consideran que están dentro del alcance de las personas versadas en el arte.

El vector puede estar en la forma de un plásmido, una partícula viral o un fago. Otros vectores incluyen secuencias de ADN, cromosómico y no cromosómico y sintéticas, derivadas de SV40; plásmidos bacterianos, ADN de fago, baculovirus, plásmidos de levadura, vectores derivados de combinaciones de plásmidos y ADN de fago, ADN viral tal como vacuna, adenovirus, virus de la viruela aviar y pseudorrabia. Se describen una variedad de vectores de expresión y clonación para uso con huéspedes procariotas y eucariotas, por ejemplo, por Sambrook.

Los vectores bacterianos particulares que pueden ser usados incluyen los plásmidos comercialmente disponibles que comprenden los elementos genéticos del vector de clonación bien conocido pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suecia), GEM1 (Promega Biotec, Madison, WI, EUA), pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, DR540, pRIT5 (Pharmacia), pKK232-8 y pCM7. Los vectores eucariotas particulares incluyen pSV2CAT, pOG44, pXT1, pSG (Stratagene), pSVK3, pBPV, pMSG y pSVL (Pharmacia). Sin embargo, se puede utilizar cualquier otro vector, mientras sea replicable y viable en la célula huésped.

Los ácidos nucleicos de la invención pueden ser expresados en casetes de expresión, vectores o virus, y expresados en forma transitoria o estable en células vegetales y semillas. Un ejemplo de un sistema de expresión transitorio utiliza sistemas de expresión episomales, por ejemplo, ARN viral del virus del mosaico de la coliflor (CaMV) generado en el núcleo mediante transcripción de un minicromosoma episomal que contiene ADN superenrollado, véase, por ejemplo, Covey (1990) Proc. Natl. Acad. Sci. EUA 87:1633-1637. En forma alternativa, se pueden insertar secuencias codificadoras, es decir, todas las secuencias o sub fragmentos de las secuencias de la invención en un genoma de la célula huésped de la planta que se convierte en parte integral del ADN cromosómico del huésped. Se pueden expresar transcritos sentido o antisentido en esta forma. Un vector que comprende las secuencias (por ejemplo, promotores o regiones codificadoras) de ácidos nucleicos de la invención puede comprender un gen marcador que confiere un fenotipo seleccionable sobre una célula o una semilla de una planta. Por ejemplo, el marcador puede codificar resistencia a los biocidas, particularmente resistencia a los antibióticos, tal como resistencia a kanamicina, G418, bleomicina, higromicina, o resistencia a herbicidas, tal como resistencia a la clorosulfurona o Basta.

Los vectores de expresión capaces de expresar ácidos nucleicos y proteínas en plantas son bien conocidos en el arte y pueden incluir, por ejemplo, vectores de *Agrobacterium* spp., virus X de patata (véase, por ejemplo, Angell (1997) EMBO J. 16: 3675 - 3684), virus del mosaico del tabaco (véase, por ejemplo, Casper (1996) Gene 173: 69 - 73), virus de atrofia arbustiva del tomate (véase, por ejemplo, Hillman (1989) Virology 169: 42 - 50), virus del grabado del tabaco (véase, por ejemplo, Dolja (1997) Virology 234: 243 - 252), virus del mosaico dorado del frijol (véase, por ejemplo, Morinaga (1993) Microbiol Immunol. 37: 471 - 476), virus del mosaico de la coliflor (véase, por ejemplo, Cecchini (1997) Mol. Plant Microbe Interact. 10: 1094 - 1101), elemento transponible Ac/Ds del maíz (véase, por ejemplo, Rubin (1997) Mol. Cell. Biol. 17: 6294 - 6302; Kunze (1996) Curr. Top. Microbiol. Immunol. 204: 161 - 194) y el elemento transponible mutador del supresor (Spm) del maíz (véase, por ejemplo, Schlappi (1996) Plant Mol. Biol. 32: 717 - 725); y derivados de los mismos.

En un aspecto, el vector de expresión puede tener dos sistemas de replicación para permitirle que se mantenga en dos organismos, por ejemplo en células de mamífero o de insecto para expresión en un huésped procariota para clonación y amplificación. Además, para integración de los vectores de expresión, el vector de expresión puede contener al menos una secuencia homóloga al genoma de la célula huésped. Puede contener dos secuencias homólogas que flanquean al constructo de expresión. El vector de integración puede ser dirigido a un locus

específico en la célula huésped mediante la selección de la secuencia homóloga apropiada para inclusión en el vector. Los constructos para vectores de integración son conocidos en el arte.

Los vectores de expresión de la invención también pueden incluir un gen marcador seleccionable para permitir la selección de cepas bacterianas que han sido transformadas, por ejemplo, genes que vuelven a las bacterias resistentes a fármacos tales como ampicilina, cloranfenicol, eritromicina, kanamicina, neomicina y tetraciclina. Los marcadores seleccionables pueden incluir, también, genes biosintéticos, tales como aquellos en las rutas biosintéticas de la histidina, triptófano y leucina.

Células huésped y células transformadas

La invención también provee una célula transformada, que comprende una secuencia de ácido nucleico de la invención, por ejemplo, una secuencia que codifica una amilasa de la invención, o un vector de la invención. La célula huésped puede ser cualquiera de las células huésped familiares para aquellos capacitados en el arte, incluyendo células procariotas, células eucariotas, tales como células bacterianas, células de hongos, células de levadura, células de mamífero, células de insecto, o células de plantas. Los ejemplos de células bacterianas incluyen cualquier especie dentro del género de *Escherichia*, *Bacillus*, *Streptomyces*, *Salmonella*, *Pseudomonas* y *Staphylococcus*, incluyendo, por ejemplo, *Escherichia coli*, *Lactococcus lactis*, *Bacillus subtilis*, *Bacillus cereus*, *Salmonella typhimurium*, *Pseudomonas fluorescens*. Los ejemplos de células de levadura incluyen a cualquiera de especie *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Schwanniomyces*, incluyendo *Pichia pastoris*, *Saccharomyces cerevisiae*, o *Schizosaccharomyces pombe*. Los ejemplos de células de insectos incluyen cualquier especie de *Spodoptera* o *Drosophila*, inclusive *Drosophila S2* y *Spodoptera Sf9*. Los ejemplos de células de animales incluyen CHO, COS o melanoma de Bowes o cualquier línea celular humana o de ratón. La selección de un huésped apropiado se encuentra dentro de las habilidades de aquellos capacitados en el arte. Las técnicas para transformación de una variedad de especies de plantas superiores son bien conocidas y están descritas en la literatura técnica y científica. Véase, por ejemplo, Weising (1988) Ann. Rev. Genet. 22: 421 - 477; patente de los Estados Unidos No. 5.750.870.

El vector puede ser introducido en las células huésped usando cualquiera de una variedad de técnicas, incluyendo transformación, transfección, transducción, infección viral, pistolas génicas, o transferencia génica mediada por Ti. Los métodos particulares incluyen transfección de fosfato de calcio, transfección mediada por DEAE-Dextrano, lipofección o electroporación (Davis, L., Dibner, M., Battey, L., Basic Methods in Molecular Biology, (1986)).

En un aspecto, los ácidos nucleicos o vectores de la invención son introducidos en las células para cribado, por lo tanto, los ácidos nucleicos ingresan a las células en una forma adecuada para expresión posterior del ácido nucleico. El método de introducción está grandemente dictado por el tipo de célula objetivo. Los ejemplos de métodos incluyen precipitación de CaPO₄, fusión de liposomas, lipofección (por ejemplo, LIPOFECTIN^{MR}), electroporación, infección viral, etc. Los ácidos nucleicos candidatos pueden integrarse en forma estable en el genoma de la célula huésped (por ejemplo, con introducción retroviral) o puede existir ya sea en forma transitoria o estable en el citoplasma (es decir a través del uso de plásmidos tradicionales, utilizando secuencias reguladoras estándar, marcadores de selección, etc.). Al igual que muchos cribados farmacéuticamente importantes se requieren objetivos celulares de mamíferos humanos o modelo, se prefieren vectores retrovirales capaces de transfectar tales objetivos.

Cuando sea apropiado, las células huésped modificadas por ingeniería genética pueden ser cultivadas en medios nutrientes convencionales modificados según sea conveniente para activar los promotores, seleccionar los transformantes o amplificar los genes de la invención. Después de la transformación de una cepa huésped adecuada y el crecimiento de la cepa huésped hasta una densidad celular apropiada, el promotor seleccionado puede ser inducido por medios apropiados (por ejemplo, cambio de temperatura o inducción química) y las células pueden ser cultivadas durante un período adicional para permitirles producir el polipéptido deseado o un fragmento del mismo.

Las células pueden ser cosechadas por centrifugación, rotas por medios físicos o químicos, y el extracto sin purificar resultante es retenido para purificación adicional. Las células microbianas empleadas para expresión de proteínas pueden ser rotas mediante cualquier método conveniente, incluyendo ciclo de congelación-descongelación, sonicación, ruptura mecánica, o el uso de agentes de lisado celular. Dichos métodos son conocidos para aquellos entrenados en el arte. El polipéptido expresado o un fragmento del mismo puede ser recuperado y purificado a partir de los cultivos celulares recombinantes mediante métodos que incluyen precipitación con sulfato de amonio o etanol, extracción con ácidos, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de hidroxilapatita y cromatografía de lectina. Las etapas de repliegamiento de la proteína pueden ser usadas, según sea necesario, para completar la configuración del polipéptido. Si se desea, se puede emplear cromatografía líquida de alto rendimiento (HPLC) para las etapas finales de purificación.

Se pueden emplear también diferentes sistemas de cultivo de células de mamíferos para expresar la proteína recombinante. Los ejemplos de sistemas de expresión de mamífero incluyen líneas celulares COS-7 de fibroblastos de riñón de mono y otras líneas celulares capaces de expresar proteínas de un vector compatible, tales como las líneas celulares C127, 3T3, CHO, HeLa y BHK.

- 5 Los constructos en las células huésped se pueden usar de manera convencional para producir el producto génico codificado por la secuencia recombinante. Dependiendo del huésped empleado en un procedimiento de producción recombinante, los polipéptidos producidos por las células huésped que contienen al vector pueden ser glicosilados o pueden ser no glicosilados. Los polipéptidos de la invención pueden o no incluir también un residuo de aminoácido inicial de metionina.
- 10 Los sistemas de traducción libres de células también se pueden emplear para producir un polipéptido de la invención. Los sistemas de traducción libres de células pueden usar los ARNm transcritos a partir de un constructo de ADN que comprende un promotor operativamente enlazado a un ácido nucleico que codifica al polipéptido o un fragmento del mismo. En algunos aspectos, el constructo de ADN puede ser alineado antes de llevar a cabo una reacción de transcripción *in vitro*. El ARNm transcrito es luego incubado con un extracto de traducción apropiado
- 15 libre de células, tal como un extracto de reticulocito de conejo, para producir el polipéptido deseado o un fragmento del mismo.

Los vectores de expresión pueden contener uno o más genes marcadores seleccionables para proporcionar un rasgo fenotípico para selección de las células huésped transformadas tal como resistencia a la dihidrofolato reductasa o a la neomicina para un cultivo células eucariotas, o tal como una resistencia a la tetraciclina o la

20 ampilicina en *E. coli*.

Amplificación de ácidos nucleicos

En la práctica de la invención, se pueden reproducir mediante amplificación los ácidos nucleicos de la invención y los ácidos nucleicos que codifican los polipéptidos de la invención. También se puede utilizar la amplificación para clonar o modificar los ácidos nucleicos de la invención. Por lo tanto, la divulgación provee pares de secuencia de del

25 cebador de amplificación para la amplificación de ácidos nucleicos de la invención. Alguien capacitado en el arte puede diseñar pares de secuencias del cebador de amplificación para cualquier parte de o la longitud completa de estas secuencias.

También se pueden utilizar reacciones de amplificación para cuantificar la cantidad de ácido nucleico en una muestra (tal como la cantidad de mensaje en una muestra celular), etiquetar el ácido nucleico (por ejemplo, para aplicarlo a un arreglo o a una transferencia), detectar el ácido nucleico, o cuantificar la cantidad de un ácido nucleico

30 específico en una muestra. En un aspecto de la divulgación, se amplifican el mensaje aislado de una célula o una biblioteca de ADNc.

La persona versada en el arte puede seleccionar y diseñar cebadores de amplificación de oligonucleótidos adecuados. Los métodos de amplificación también son conocidos en el arte e incluyen, por ejemplo, reacción en

35 cadena de la polimerasa, PCR (véase, por ejemplo, PCR Protocols, A Guide To Methods And Applications, ed. Innis, Academic Press, N.Y. (1990) y PCR Strategies (1995), ed. Innis, Academic Press, Inc., N.Y., reacción en cadena de la ligasa (LCR) (véase, por ejemplo, Wu (1989) Genomics 4: 560; Landegren (1988) Science 241: 1077; Barringer (1990) Gene 89: 117); amplificación de la transcripción (véase, por ejemplo, Kwoh (1989) Proc. Natl. Acad. Sci. EUA 86: 1173); y, replicación autosostenida de la secuencia (véase, por ejemplo, Guatelli (1990) Proc. Natl. Acad. Sci.

40 EUA 87: 1874); amplificación de Q Beta replicasa (véase, por ejemplo, Smith (1997) J. Clin. Microbiol. 35: 1477 - 1491), ensayo de amplificación automatizado de Q-beta replicasa (véase, por ej., Burg (1996) Mol. Cell. Probes 10: 257 - 271) y otras técnicas mediadas por ARN polimerasa (por ejemplo, NASBA, Cangene, Mississauga, Ontario); véase además Berger (1987) Methods Enzymol. 152: 307 - 316', Sambrook; Ausubel; patentes de los Estados Unidos Nos. 4.683.195 y 4.683.202; Sooknanan (1995) Biotechnology 13: 563 - 564.

45 Determinación del grado de identidad de secuencia

La invención provee ácidos nucleicos que comprenden las secuencias que tienen al menos 98%, 99%, o más, o identidad de secuencia completa (100%) con un ácido nucleico de ejemplo de la invención. La invención provee polipéptidos que comprenden secuencias que tienen al menos aproximadamente 99% o más, o identidad de

50 secuencia completa (100%) con un polipéptido de ejemplo de la invención. El grado de identidad de secuencia (homología) se puede determinar usando cualquier programa de ordenador y parámetros asociados, incluyendo aquellos descritos aquí, tales como BLAST 2.2.2 o FASTA versión 3.0t78, con los parámetros predeterminados.

La expresión "sustancialmente idénticos" en el contexto de dos ácidos nucleicos puede referirse a dos o más secuencias que tienen, por ejemplo, al menos 98%, 99%, o más identidad de nucleótidos (secuencia), cuando se comparan y se alinean para máxima correspondencia, cuando se mide usando cualquier algoritmo de comparación

de secuencia conocido, como se discute en detalle a continuación, o por medio de inspección visual. En aspectos alternativos, la divulgación provee secuencias de ácido nucleico y de polipéptidos que tienen identidad sustancial con un ejemplo de secuencia de la invención sobre una región de al menos aproximadamente 10, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 o más residuos, o una región que oscila entre aproximadamente 50 residuos hasta la longitud completa del ácido nucleico o polipéptido. Las secuencias de ácido nucleico de la invención pueden ser sustancialmente idénticas sobre la longitud completa de una región que codifica al polipéptido.

Las secuencias homólogas también incluyen secuencias de ARN en las cuales las uridinas reemplazan a las timinas en las secuencias de ácido nucleico. Las secuencias homólogas pueden ser obtenidas usando cualquiera de los procedimientos descritos aquí o pueden resultar de la corrección de un error de secuenciación. Se apreciará que las secuencias de ácido nucleico como las expuestas aquí puedan ser representadas en un formato tradicional de carácter individual (véase, por ejemplo, Stryer, Lubert. *Biochemistry*, 3ra Ed., W. H Freeman & Co., New York) o en cualquier otro formato que registre la identidad de los nucleótidos en una secuencia.

Se utilizan en este aspecto de la invención diferentes programas de comparación de secuencias identificadas aquí. Se pueden evaluar las identidades de las secuencias de proteína y/o de ácido nucleico identificadas (homologías) usando cualquiera entre una variedad de algoritmos y programas de comparación de secuencia conocidos en el arte. Tales algoritmos y programas incluyen, pero no se limitan a, TBLASTN, BLASTP, FASTA, TFASTA y CLUSTALW (Pearson y Lipman, *Proc. Natl. Acad. Sci. EUA* 85(8): 2444 - 2448, 1988; Altschul y colaboradores, *J. Mol. Biol.* 215(3): 403 - 410, 1990; Thompson y colaboradores, *Nucleic Acid Res.* 22(2): 4673 - 4680, 1994; Higgins y colaboradores, *Methods Enzymol.* 266: 383 - 402, 1996; Altschul y colaboradores, *J. Mol. Biol.* 215(3): 403 - 410, 1990; Altschul y colaboradores, *Nature Genetics* 3: 266 - 272, 1993).

La homología o identidad de secuencia se puede medir usando software de análisis de secuencia (por ejemplo, el paquete de software de análisis de secuencia del Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, WI 53705). Tal software empareja secuencias similares al asignarles grados de homología a diferentes supresiones, sustituciones y otras modificaciones. Los términos "homología" e "identidad" en el contexto de dos o más secuencias de ácidos nucleicos o de polipéptidos, se refiere a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales cuando se comparan y se alinean para una correspondencia máxima en una ventana de comparación o región designada como se mide usando cualquier número de algoritmos de comparación de secuencias o mediante alineación manual e inspección visual. Para comparación de secuencias, una secuencia puede actuar como una secuencia de referencia, por ejemplo, una secuencia de la invención, para lo cual se comparan las secuencias de prueba. Cuando se utiliza un algoritmo de comparación de secuencias, se ingresan las secuencias de prueba y de referencia en un ordenador, se designan las subsecuencias coordinadas si es necesario, y se designan los parámetros del programa de algoritmo de secuencia. Se pueden utilizar los parámetros de programa predeterminados, o se pueden designar parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidad de secuencia para las secuencias de prueba con relación a la secuencia de referencia, con base en los parámetros del programa.

Una "ventana de comparación", como se usa aquí, incluye la referencia a un segmento de cualquiera de los números de residuos contiguos. Por ejemplo, en aspectos alternativos de la divulgación, los residuos contiguos que oscilan desde 20 hasta la longitud completa de un ejemplo de una secuencia de polipéptido o de ácido nucleico de la invención se comparan con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alinean de forma óptima. Si la secuencia de referencia tiene la identidad de secuencia requerida con respecto a un polipéptido de ejemplo o ácido nucleico de la invención, por ejemplo 98%, 99% o más de identidad de secuencia con una secuencia de la invención, esa secuencia está dentro del alcance de la invención. En realizaciones alternativas, subsecuencias que varían de aproximadamente 20 a 600, aproximadamente 50 a 200, y aproximadamente 100 a 150 se comparan con una secuencia de referencia del mismo número de posiciones contiguas después de que las dos secuencias se alinean de manera óptima.

Los métodos de alineación de secuencia para comparación son bien conocidos en el arte. En aspectos alternativos, la alineación óptima de secuencias para comparación se puede llevar a cabo, por ejemplo, mediante el algoritmo de homología local de Smith & Waterman, *Adv. Appl. Math.* 2: 482, 1981, mediante el algoritmo de alineación de homología de Needleman & Wunsch, *J. Mol. Biol.* 48: 443, 1970, mediante el método de búsqueda de similitudes de Pearson & Lipman, *Proc. Natl. Acad. Sci. EUA* 85: 2444, 1988, por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el paquete de software de Wisconsin Genetics, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual. Otros algoritmos para determinación homología o identidad incluyen, por ejemplo, además de un programa BLAST (Basic Local Alignment Search Tool en el National Center for Biological Information), ALIGN, AMAS (Analysis of Multiply Aligned Sequences), AMPS (Protein Multiple Sequence Alignment), ASSET (Aligned Segment Statistical Evaluation Tool), BANDS, BESTSCOR, BIOSCAN (Biological Sequence Comparative Analysis Node), BLIMPS (Blocks IMProved Searcher), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, el algoritmo de Smith-Waterman, DARWIN, algoritmo Las Vegas, FNAT (Forced Nucleotide

Alignment Tool), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Fristensky Sequence Analysis Package), GAP (Global Alignment Program), GENAL, GIBBS, GenQuest, ISSC (Sensitive Sequence Comparación), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construcción & Analysis workbench), MAP (Multiple Alignment Program), MBLKP, MBLKN, PIMA (Pattern-Induced Multi-sequence Alignment), SAGA (Sequence Alignment by Genetic Algorithm) y WHAT-IF. Tales programas de alineación también se pueden utilizar para la criba de bases de datos genómicas para identificar secuencias de polinucleótidos que tienen secuencias sustancialmente idénticas. Una cantidad de bases de datos de genomas están disponibles, por ejemplo, una parción sustancial del genoma humano está disponible como parte del Human Genome Sequencing Project (Gibbs, 1995). Varios genomas han sido secuenciados, por ejemplo, de *M. genitalium* (Fraser y colaboradores, 1995), *M. jannaschii* (Bult y colaboradores, 1996), *H. influenzae* (Fleischmann y colaboradores, 1995), *E. coli* (Blattner y colaboradores, 1997) y levadura (*S. cerevisiae*) (Mewes y colaboradores, 1997) y *D. melanogaster* (Adams y colaboradores, 2000). También se han hecho progresos significativos en la secuenciación de los genomas del organismo modelo, tal como ratón, *C. elegans* y *Arabidopsis* sp. Las bases de datos que contienen información genómica anotada con alguna información funcional son mantenidas por diferentes organizaciones y son accesibles por Internet.

Los algoritmos BLAST, BLAST 2.0 y BLAST 2.2.2 también pueden ser utilizados para la práctica de la divulgación. Se describen, por ejemplo, en Altschul (1977) Nuc. Acids Res. 25: 3389 - 3402; Altschul (1990) J. Mol. Biol. 215: 403 - 410. El software para realizar el análisis por BLAST se encuentra públicamente disponible a través del National Center for Biotechnology Information. Este algoritmo implica identificar primero pares de secuencia de alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia pregunta, que o bien coincide o satisface algún puntaje T del umbral valorado como positivo cuando se alinea con una palabra de la misma longitud en una secuencia de la base de datos. T se refiere al umbral de puntuación de la palabra vecina (Altschul (1990) ver más arriba). Estas coincidencias de la palabra vecina inicial actúan como semillas para iniciar búsquedas para encontrar los HSP más largos que las contienen. Las coincidencias de palabras se extienden en ambas direcciones a lo largo de cada secuencia tan lejos como se pueda aumentar el puntaje de alineación acumulado. Los puntajes acumulados se calculan utilizando, para secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre >0). Para secuencias de aminoácidos, se utiliza una matriz de puntuación para calcular el puntaje acumulado. La extensión de las coincidencias de palabras en cada dirección se detienen cuando: el puntaje de alineación acumulado cae en una cantidad X de su valor máximo alcanzado; el puntaje acumulado va hasta cero o menos, debido a la acumulación de una o más alineaciones de residuos de puntuación negativa; o se llega al final de cada secuencia. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) utiliza como parámetros predeterminados una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas hebras. Para las secuencias de aminoácidos, el programa BLASTP utiliza como parámetros predeterminados una longitud de palabra de 3 y expectativas (E) de 10 y la matriz de puntuación BLOSUM62 (véase Henikoff Henikoff (1989) Proc. Natl. Acad. Sci. EUA 89: 10915) alineaciones (B) de 50, expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas hebras. El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin & Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que provee una indicación de la probabilidad por la cual una coincidencia entre dos secuencias de nucleótidos o de aminoácidos ocurriría por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de la suma más pequeña en una comparación del ácido nucleico de prueba con el ácido nucleico de referencia es aproximadamente menor a 0,2, más preferentemente aproximadamente menor a 0,01 y lo más preferible aproximadamente menor a 0,001. En un aspecto, las homologías de secuencia de proteínas y de ácidos nucleicos se evalúan utilizando Basic Local Alignment Search Tool ("BLAST"). Por ejemplo, se pueden utilizar cinco programas BLAST específicos para llevar a cabo la siguiente tarea: (1) BLASTP y BLAST3 comparan una secuencia pregunta de aminoácidos contra una base de datos de secuencia de proteína; (2) BLASTN compara una secuencia pregunta de nucleótidos contra una base de datos de secuencia de nucleótidos; (3) BLASTX compara los productos de traducción conceptual de seis marcos de una secuencia pregunta de nucleótidos (ambas hebras) contra una base de datos de secuencia de proteína; (4) TBLASTN compara una secuencia pregunta de proteína contra una base de datos de secuencia de nucleótidos traducida en todos los seis marcos de lectura (ambas hebras); y, (5) TBLASTX compara las traducciones de seis marcos de una secuencia pregunta de nucleótidos contra las traducciones de seis marcos de una base de datos de secuencias de nucleótidos. Los programas BLAST identifican secuencias homólogas mediante la identificación de segmentos similares, que se denominan aquí como "pares de segmentos de alta puntuación", entre una secuencia pregunta de ácido nucleico o aminoácidos y una secuencia de prueba que se obtiene preferiblemente de una base de datos de secuencias de proteínas o de ácidos nucleicos. Los pares de segmentos de alta puntuación se identifican preferiblemente (es decir, se alinean) por medio de una matriz de puntuación, muchos de los cuales son conocidos en el arte. Preferiblemente, la matriz de puntuación usada es la matriz BLOSUM62 (Gonnet y colaboradores, Science 256: 1443 - 1445, 1992; Henikoff y Henikoff, Proteins 17: 49-61, 1993), Menos preferiblemente, se pueden utilizar también las matrices PAM o PAM250 (véase, por ejemplo, Schwartz y Diahoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation).

- En un aspecto de la divulgación, para determinar si un ácido nucleico tiene la identidad de secuencia requerida para estar dentro del alcance de la invención, se utilizan los programas NCBI BLAST 2.2.2, con opciones de parámetros predeterminados para blastp. Existen aproximadamente 38 opciones de configuración en el programa BLAST 2.2.2.
- 5 En este aspecto de ejemplo de la divulgación, se usan todos los valores predeterminados excepto para la configuración de filtración predeterminada (es decir, se fijan todos los parámetros predeterminados excepto la filtración, que se fija en OFF); en su lugar, se usa una configuración "-F F", que desactiva la filtración. El uso de la filtración en forma predeterminada a menudo conduce a violaciones de Karlin-Altschul debido a la corta longitud de la secuencia.
- Los valores predeterminados usados en este aspecto de la divulgación que sirve como ejemplo y para determinar los valores en la Figura 3, como se discutió anteriormente, incluyen:
- 10 "Filtro para baja complejidad: ON
- Tamaño de palabra: 3
- Matriz: Blosum62
- Costos de los espacios: Existencia:11
- 15 Extensión: 1"
- Otros parámetros predeterminados pueden ser: filtro para baja complejidad OFF, tamaño de palabra de 3 para proteína, matriz BLOSUM62, penalidad por existencia de espacios de -11 y una penalidad por extensión de espacio de -1. Un ejemplo de configuración del programa NCBI BLAST 2.2.2 tiene el parámetro de opción "-W" en 0. Esto significa que, si no se ajusta, el tamaño de palabra es por defecto 3 para las proteínas y 11 para los nucleótidos.
- 20 Sistemas computarizados y productos de programas de cómputo
- Para determinar e identificar identidades de secuencia, homologías estructurales, motivos y similares *in silico*, se puede almacenar, registrar y manipular la secuencia de la invención en cualquier medio que pueda ser leído y accedido por un ordenador. Por lo tanto, la divulgación provee ordenadores, sistemas de cómputo, medios legibles por el ordenador, productos de programas de cómputo y similares, registradas o almacenadas con respecto a las
- 25 secuencias de ácido nucleico y polipéptidos de la invención. Como se utilizan aquí, las palabras "registrado" y "almacenado" se refieren a un proceso para almacenamiento de información en un medio computarizado. Una persona capacitada en la materia puede adoptar cualquiera de los métodos conocidos para registro de la información en un medio legible por el ordenador para generar producciones que comprenden una o más de las secuencias de ácido nucleico y/o de polipéptidos de la invención.
- 30 La divulgación también provee ordenadores y procesadores que comprenden productos de programas de computación que comprenden las secuencias de la invención; y como se utiliza aquí, los términos "ordenador", "programa de cómputo" y "procesador" se utilizan en su contexto general más amplio e incorporan todos esos dispositivos, como se describe en detalle a continuación. Una "secuencia codificadora de" o una "secuencia que codifica" un polipéptido o proteína particular, es una secuencia de ácido nucleico que se transcribe y traduce en un
- 35 polipéptido o proteína cuando se coloca bajo el control de secuencias reguladoras apropiadas.
- Otro aspecto de la divulgación es un medio legible por el ordenador que tiene registrado al respecto al menos una secuencia de ácido nucleico y/o de polipéptido de la invención. Los medios legibles por el ordenador incluyen medios legibles en forma magnética, medios legibles en forma óptica, medios legibles en forma electrónica y medios magnéticos/ópticos. Por ejemplo, los medios legibles por ordenador pueden ser un disco duro, un disco blando, una
- 40 cinta magnética, CD-ROM, disco digital versátil (DVD), memoria de acceso aleatoria (RAM), o memoria únicamente de lectura (ROM) así como también otros tipos de medios conocidos por personas capacitadas en el arte.
- Aspectos de la divulgación incluyen sistemas (por ejemplo, sistemas basados en Internet), particularmente sistemas computarizados, que almacenan y manipulan las secuencias e información de las secuencias descritas aquí. Un ejemplo de un sistema de cómputo 100 se ilustra en forma de un diagrama de bloque de la Figura 1. Como se utiliza
- 45 aquí, "un sistema de cómputo" se refiere a los componentes del hardware, componentes del software y componentes de almacenamiento de datos usados para analizar una secuencia de nucleótidos o de polipéptidos de la invención. El sistema de cómputo 100 puede incluir un procesador para procesamiento, acceso y manipulación de los datos de la secuencia. El procesador 105 puede ser cualquier tipo conocido de unidad central de procesamiento, tal como, por ejemplo, el Pentium III de Intel Corporation, o un procesador similar de Sun, Motorola, Compaq, AMD o
- 50 International Business Machines. El sistema de cómputo 100 es un sistema de propósito general que comprende el procesador 105 y uno o más componentes internos de almacenamiento de datos 110 para almacenamiento de datos y uno o más dispositivos de recuperación de datos para recuperar los datos almacenados en los componentes

de almacenamiento de datos. Una persona capacitada en la materia puede apreciar fácilmente que cualquiera de los sistemas de computarizados disponibles actualmente es adecuado.

En un aspecto, el sistema de cómputo 100 incluye un procesador 105 conectado a un bus que está conectado a una memoria principal 115 (preferiblemente implementada como RAM) y uno o más dispositivos internos de almacenamiento de datos 110, tal como un disco duro y/o otros medios legibles por el ordenador que tiene datos registrados en él. El sistema de cómputo 100 también puede incluir uno o más dispositivos de recuperación de datos 118 para leer los datos almacenados en los dispositivos internos de almacenamiento de datos 110. El dispositivo de recuperación de datos 118 puede representar, por ejemplo, un dispositivo para disco blando, un dispositivo para disco compacto, un dispositivo para disco magnético, o un modem capaz de conectarse a un sistema remoto de almacenamiento de datos (por ejemplo, a través de la Internet) etc. En algunas formas de realización, el dispositivo de almacenamiento interno de datos 110 es un medio removible legible por el ordenador tal como un disco blando, un disco compacto, una cinta magnética. etc. que contiene lógica de control y/o datos registrados en el mismo. El sistema de cómputo 100 puede, incluir convenientemente, o ser programado mediante un software apropiado para leer los datos de lógica de control y/o los datos del componente de almacenamiento de datos una vez insertado en el dispositivo de recuperación de datos. El sistema de cómputo 100 incluye una pantalla 120 que se usa para mostrar el resultado a un usuario del ordenador. También se debe observar que el sistema de cómputo 100 puede estar enlazado a otros sistemas de cómputo 125a-c en una red o red de área ancha para proporcionar acceso centralizado al sistema de cómputo 100. El software para acceder y procesar las secuencias de nucleótidos o de aminoácidos de la invención puede residir en la memoria principal 115 durante la ejecución. En algunos aspectos, el sistema de cómputo 100 puede comprender, además, un algoritmo de comparación de secuencias para comparar una secuencia de ácido nucleico de la invención. El algoritmo y la(s) secuencia(s) se pueden almacenar en un medio legible por el ordenador. Un "algoritmo de comparación de secuencias" se refiere a uno o más programas que se implementan (en forma local o remota) sobre el sistema de cómputo 100 para comparar una secuencia de nucleótidos con otras secuencias de nucleótidos y/o los compuestos almacenados en el medio de almacenamiento de datos. Por ejemplo, el algoritmo de comparación de secuencias puede comparar las secuencias de nucleótidos de la invención almacenadas en un medio legible por el ordenador para las secuencias de referencia almacenadas en un medio legible por el ordenador para identificar homologías o motivos estructurales.

Los parámetros usados con los algoritmos anteriores pueden ser adaptados dependiendo de la longitud de la secuencia y del grado de homología estudiado. En algunos aspectos, los parámetros pueden ser los parámetros predeterminados usados por los algoritmos en ausencia de instrucciones del usuario. La Figura 2 es un diagrama de flujo que ilustra un aspecto de un proceso 200 para comparar una nueva secuencia de nucleótidos o de proteínas con una base de datos de secuencias a fin de determinar los niveles de homología entre la nueva secuencia y las secuencias en la base de datos. La base de datos de las secuencias puede ser una base de datos privada almacenada en el sistema de cómputo 100, o una base de datos pública tal como GENBANK que se encuentra disponible a través de Internet. El proceso 200 comienza en un estado de inicio 201 y luego pasa a un estado 202 en donde la nueva secuencia a ser comparada se almacena en una memoria en un sistema de cómputo 100. Como se discutió anteriormente, la memoria podría ser cualquier tipo de memoria, inclusive RAM o un dispositivo de almacenamiento interno. El proceso 200 luego pasa a un estado 204 en donde una base de datos de secuencias se abre para análisis y comparación. El proceso 200 luego pasa a un estado 206 en donde la primera secuencia almacenada en la base de datos se lee en una memoria en el ordenador. Se realiza luego una comparación en un estado 210 para determinar si la primera secuencia es igual a la segunda secuencia. Es importante notar que este paso no se limita a realizar una comparación exacta entre la nueva secuencia y la primera secuencia en la base de datos. Los métodos son conocidos por aquellos versados en el arte para comparar dos secuencias de nucleótidos o de proteínas, incluso si no son idénticas. Por ejemplo, se pueden introducir espacios en una secuencia a fin de elevar el nivel de homología entre las dos secuencias evaluadas. Los parámetros que controlan si se introducen espacios u otras características en una secuencia durante la comparación, por lo general son ingresados por el usuario del sistema de cómputo. Una vez que se ha realizado una comparación de las dos secuencias en el estado 210, se hace una determinación en un estado de decisión 210 de si las dos secuencias son iguales. Desde luego, el término "iguales" no se limita a las secuencias que son absolutamente idénticas. Las secuencias que están dentro de los parámetros de homología ingresados por el usuario se marcarán como "iguales" en el proceso 200. Si se determina que las dos secuencias son iguales, el proceso 200 pasa a un estado 214 en donde se muestra al usuario el nombre de la secuencia de la base de datos. Este estado notifica al usuario que la secuencia con el nombre visualizado cumple con las restricciones de homología que fueron ingresadas. Una vez que el usuario visualiza el nombre de la secuencia almacenada, el proceso 200 pasa a un estado de decisión 218 en donde se determina si existen más secuencias en la base de datos. Si no existen más secuencias en la base de datos, entonces el proceso 200 termina en un estado final 220. Sin embargo, si existen más secuencias en la base de datos, entonces el proceso 200 pasa a un estado 224 en donde un indicador se mueve a la próxima secuencia en la base de datos de manera tal que pueda ser comparada con la nueva secuencia. De esta forma, se alinea y compara la secuencia nueva con cada secuencia en la base de datos. Debe observarse que si se hizo una determinación en el estado de decisión 212 de que las secuencias no eran homólogas, entonces el proceso 200 pasará inmediatamente al estado de decisión 218 a fin de determinar si cualquier otra secuencia estaba disponible en la base de datos para comparación. Por lo tanto, un aspecto de la divulgación es un sistema de cómputo que comprende un procesador, un dispositivo de almacenamiento de datos que tiene almacenado en el mismo una secuencia de ácido nucleico de

la invención y un comparador de secuencia para realizar la comparación. El comparador de secuencia puede indicar un nivel de homología entre las secuencias comparadas o identificar los motivos estructurales, o puede identificar los motivos estructurales en las secuencias que se comparan con estos códigos de ácido nucleico y con códigos de polipéptido. La Figura 3 es un diagrama de flujo que ilustra una realización de un proceso 250 en un ordenador para determinar si dos secuencias son homólogas. El proceso 250 comienza en un estado de inicio 252 y luego pasa a un estado 254 en donde se almacena en una memoria una primera secuencia a ser comparada. Luego se almacena en una memoria una segunda secuencia a ser comparada en un estado 256. El proceso 250 luego pasa a un estado 260 en donde el primer carácter en la primera secuencia es leído y luego pasa a un estado 262 en donde el primer carácter de la segunda secuencia es leído. Debe entenderse que si la secuencia es una secuencia de nucleótidos, entonces el carácter es normalmente A, T, C, G o U. Si la secuencia es una de secuencia de proteína, entonces puede ser un código de aminoácidos de una sola letra de manera tal que la primera y la segunda secuencia puedan ser fácilmente comparadas. Se hace luego una determinación en un estado de decisión 264 de si los dos caracteres son iguales. Si lo son, entonces el proceso 250 pasa a un estado 268 en donde son leídos los próximos caracteres en la primera y segunda secuencias. Se determina entonces si los próximos caracteres son iguales. Si lo son, entonces el proceso 250 continúa este circuito hasta que dos caracteres no son iguales. Si se determina que los próximos dos caracteres no son iguales, el proceso 250 pasa a un estado de decisión 274 para determinar si existen más caracteres en cualquier secuencia para ser leídos. Si no existen más caracteres por ser leídos, entonces el proceso 250 pasa a un estado 276 en donde se muestra al usuario el nivel de homología entre la primera y la segunda secuencia. El nivel de homología se determina al calcular la proporción de caracteres entre las secuencias que no eran iguales del número total de secuencias en la primera secuencia. Por lo tanto, si cada carácter en una primera secuencia de nucleótidos 100 se alinea con cada carácter en una segunda secuencia, el nivel de homología sería del 100%.

En forma alternativa, el programa de computación puede comparar una secuencia de referencia con una secuencia de la invención para determinar si las secuencias difieren en una o más posiciones. El programa puede registrar la longitud e identidad de los residuos de aminoácidos o nucleótidos insertados, suprimidos o sustituidos con respecto a la secuencia ya sea de la referencia o de la invención. El programa de computación puede ser un programa que determina si una secuencia de referencia contiene un polimorfismo de un solo nucleótido (SNP) con respecto a una secuencia de la invención, o, si una secuencia de la invención comprende un SNP de una secuencia conocida. Por lo tanto, en algunos aspectos, el programa de computación es un programa que identifica SNPs. El método puede ser implementado por los sistemas computarizados descritos anteriormente y el método ilustrado en la Figura 3. El método puede ser realizado leyendo una secuencia de la invención y las secuencias de referencia a través del uso del programa de cómputo e identificar las diferencias con el programa de cómputo.

En otros aspectos, el sistema basado en un ordenador comprende un identificador para identificar características en un ácido nucleico o polipéptido de la invención. Un "identificador" se refiere a uno o más programas que identifican ciertas características en una secuencia de ácido nucleico. Por ejemplo, un identificador puede comprender un programa que identifica un marco de lectura abierto (ORF) en una secuencia de ácido nucleico. La Figura 4 es un diagrama de flujo que ilustra un aspecto de un proceso identificador 300 para detectar la presencia de una característica en una secuencia. El proceso 300 comienza en un estado de inicio 302 y luego pasa a un estado 304 en donde una primera secuencia que va a ser verificada en cuanto a sus características se almacena en una memoria 115 en el sistema de cómputo 100. El proceso 300 pasa luego a un estado 306 en donde se abre una base de datos de características de secuencia. Dicha base de datos incluiría una lista de cada atributo de la característica junto con el nombre de la característica. Por ejemplo, un nombre de característica podría ser "Codón de inicio" y el atributo sería "ATG". Otro ejemplo sería el nombre de la característica "Caja TAATAA" y el atributo de la característica sería "TAATAA". Un ejemplo de dicha base de datos es producido por la University of Wisconsin Genetics Computer Group. En forma alternativa, las características pueden ser motivos estructurales de polipéptidos tales como hélices alfa, láminas beta, o motivos de polipéptidos funcionales tales como sitios activos enzimáticos, motivos hélice-vuelta- hélice u otros motivos conocidos por aquellos capacitados en el arte. Una vez que se abre la base de datos de las características en el estado 306, el proceso 300 pasa a un estado 308 en donde la primera característica es leída de la base de datos. Luego se realiza una comparación del atributo de la primera característica con la primera secuencia en un estado 310. Luego se determina en un estado de decisión 316 si se encontró el atributo de la característica en la primera secuencia. Si se encontró el atributo, entonces el proceso 300 pasa a un estado 318 en donde se muestra al usuario el nombre de la característica encontrada. El proceso 300 luego pasa a un estado de decisión 320 en donde se determina si existen más características en la base de datos. Si no existen más características, entonces el proceso 300 termina en un estado final 324. Sin embargo, si existen más características en la base de datos, entonces el proceso 300 lee la próxima característica de la secuencia en un estado 326 y regresa al estado 310 en donde el atributo de la próxima característica es comparado contra la primera secuencia. Si el atributo de la característica no se encuentra en la primera secuencia en el estado de decisión 316, el proceso 300 pasa directamente al estado de decisión 320 a fin de determinar si existen más características en la base de datos. Por lo tanto, en un aspecto, la divulgación provee un programa de computador que identifica los marcos de lectura abierta (ORFs).

Una secuencia de polipéptidos o de ácido nucleico de la invención puede ser almacenada y manipulada en una variedad de programas de procesamiento de datos en una variedad de formatos. Por ejemplo, se puede almacenar

una secuencia como texto en un archivo de procesamiento de texto, tal como Microsoft WORD o WORDPERFECT o como un archivo ASCII en una variedad de programas de bases de datos familiares para aquellos capacitados en el arte, tales como DB2^{MR}, SYBASE^{MR}, u ORACLE^{MR}. Asimismo, muchos programas de computación y bases de datos pueden ser usados como algoritmos de comparación de secuencias, identificadores o fuentes de secuencias de nucleótidos de referencia o secuencias de polipéptidos a ser comparadas con una secuencia de ácido nucleico de la invención. Los programas y las bases de datos usados para la práctica de la divulgación incluyen, pero no se limitan a: MACPATTERN^{MR} (EMBL), DiscoveryBase (Molecular Applications Group), GENEMINE^{MR} (Molecular Applications Group), LOOK^{MR} (Molecular Applications Group), MACLOOK^{MR} (Molecular Applications Group), BLAST y BLAST2 (NCBI), BLASTN y BLASTX (Altschul y colaboradores, J. Mol. Biol. 215: 403, 1990), FASTA (Pearson y Lipman, Proc. Natl. Acad. Sci. EUA, 85: 2444, 1988), FASTDBTM (Brutlag y colaboradores. Comp. App. Biosci. 6: 237 - 245, 1990), Catalyst (Molecular Simulations Inc.), CATALYST^{MR} / SHAPE^{MR} (Molecular Simulations Inc.), CERIU2.DB ACCESS^{MR} (Molecular Simulations Inc.), HYPOGEN^{MR} (Molecular Simulations Inc.), INSIGHT II^{MR}, (Molecular Simulations Inc.), DISCOVER^{MR} (Molecular Simulations Inc.), CHARMM^{MR} (CHARMm^{MR}) (Molecular Simulations Inc.), Felix (Molecular Simulations Inc.), DELPHI^{MR}, (Molecular Simulations Inc.), QUANTEM^{MR}, (Molecular Simulations Inc.), HOMOLOG^{MR} (Molecular Simulations Inc.), MODELER^{MR} (Molecular Simulations Inc.), ISIS^{MR} (Molecular Simulations Inc.), QUANTA^{MR} / Protein Design (Molecular Simulations Inc.), WEBLAB^{MR} (Molecular Simulations Inc.), WEBLAB DIVERSITY EXPLORER^{MR} (Molecular Simulations Inc), GENE EXPLORER^{MR} (Molecular Simulations Inc.), SEQFOLD^{MR} (Molecular Simulations Inc.), la base de datos de MDL Available Chemicals Directory, la base de datos de MDL Drug Data Report, la base de datos de Comprehensive Medicinal Chemistry, la base de datos de Derwent's World Drug Index, la base de datos de BioByteMasterFile, la base de datos del GenBank y la base de datos de Genseqn. Muchos otros programas y bases de datos serán evidentes para aquella persona versada en el arte a partir de la presente divulgación.

Los motivos que pueden ser detectados usando los programas anteriormente mencionados incluyen secuencias que codifican cremalleras de leucina, motivos de hélice-vuelta-hélice, sitios de glicosilación, sitios de ubiquitinación, hélices alfa, y láminas beta, secuencias señal que codifican péptidos señal que dirigen secreción de las proteínas codificadas, secuencias implicadas en la regulación de la transcripción tal como cajas homeocajas, tramos ácidos, sitios enzimáticos activos, sitios de enlazamiento de sustratos y sitios de escisión enzimáticos.

Hibridación de ácidos nucleicos

La divulgación provee ácidos nucleicos aislados, sintéticos o recombinantes, que se hibridan bajo condiciones rigurosas con una secuencia de ejemplo de la invención, o un ácido nucleico que codifica un polipéptido de la invención. Las condiciones rigurosas pueden ser condiciones rigurosas altas, condiciones rigurosas medias, condiciones rigurosas bajas, incluyendo las condiciones de rigurosidad alta y reducida descritas aquí. En un aspecto, es la rigurosidad de las condiciones de lavado que establecen las condiciones que determinan si un ácido nucleico está dentro del alcance de la invención, como se discute a continuación.

Los protocolos de "hibridación" usados para la práctica de esta invención incluyen procesos mediante los cuales una hebra de ácido nucleico se une a una hebra complementaria a través de apareamiento de bases. Las reacciones de hibridación pueden ser sensibles y selectivas de manera tal que se puede identificar una secuencia de interés particular incluso en muestras en las que está presente en bajas concentraciones. Las condiciones rigurosas pueden ser definidas, por ejemplo, por las concentraciones de sal o formamida en las soluciones de prehibridación e hibridación, o por la temperatura de hibridación y son conocidas en el arte. Por ejemplo, se puede incrementar la rigurosidad al reducir la concentración de la sal, incrementar la concentración de formamida, o elevar la temperatura de hibridación, alterando el tiempo de hibridación, como se describe en detalle, a continuación. En aspectos alternativos, los ácidos nucleicos de la divulgación se definen por su capacidad de hibridar bajo diversas condiciones de astringencia (por ejemplo, alto, medio y bajo), tal como se expone aquí.

En formas de reaxilización alternativas, los ácidos nucleicos de la divulgación se definen por su capacidad para hibridarse bajo condiciones rigurosas que pueden ser de aproximadamente cinco residuos y la longitud completa de ácido nucleico de la invención; por ejemplo, pueden ser de al menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, o más residuos de longitud. Los ácidos nucleicos más cortos que la longitud completa también están incluidos. Estos ácidos nucleicos pueden ser útiles, por ejemplo, como sondas de hibridación, sondas de marcación, sondas de oligonucleótido para la PCR, ARNi, antisentido o secuencias que codifican péptidos de enlazamiento a anticuerpos (epítomos), motivos, sitio activos y similares.

En un aspecto, los ácidos nucleicos de la divulgación se definen por su capacidad para hibridarse bajo condiciones de alta rigurosidad que comprenden condiciones de aproximadamente 50% de formamida aproximadamente entre 37 °C y 42 °C. En un aspecto, los ácidos nucleicos de la divulgación se definen por su capacidad para hibridarse bajo condiciones de rigurosidad reducida que comprenden condiciones de aproximadamente 35% a 25% de formamida aproximadamente entre 30 °C y 35 °C.

En forma alternativa, los ácidos nucleicos de la divulgación se definen por su capacidad para hibridarse bajo condiciones de alta rigurosidad que comprenden condiciones a 42 °C en 50% de formamida, 5X SSPE, 0,3% de SDS y una secuencia repetitiva que bloquea al ácido nucleico, tal como cot-1 o ADN de esperma de salmón (por ejemplo, 200 µg/ml de ADN de esperma de salmón cortado y desnaturalizado). En un aspecto, se definen ácidos nucleicos de la divulgación por su capacidad para hibridarse bajo condiciones de rigurosidad reducida que comprenden 35% de formamida a una temperatura reducida de 35 °C.

Después de la hibridación, se puede lavar el filtro con 6X SSC, 0,5% de SDS a 50 °C. Estas condiciones son consideradas condiciones "moderadas" por encima de 25% de formamida y condiciones "bajas" por debajo de 25% de formamida. Un ejemplo específico de condiciones "moderadas" de hibridación es cuando se realiza la hibridación con 30% de formamida. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación se realiza con 10% de formamida.

El rango de temperatura correspondiente a un nivel particular de rigurosidad puede hacerse adicionalmente más estrecho calculando la proporción de purina con respecto a la pirimidina del ácido nucleico de interés y ajustando en forma apropiada la temperatura. Los ácidos nucleicos de la divulgación también se definen por su capacidad para hibridarse bajo condiciones de rigurosidad alta, media y baja, como se expone en Ausubel y Sambrook. Las variaciones en los rangos y condiciones anteriormente mencionados son conocidos en el arte. Las condiciones de hibridación son discutidas de manera adicional a continuación.

El procedimiento anterior puede ser modificado para identificar los ácidos nucleicos que tienen niveles reducidos de homología con respecto a la secuencia de la sonda. Por ejemplo, para obtener ácidos nucleicos de homología reducida con respecto a la sonda detectable, se pueden utilizar condiciones menos rigurosas. Por ejemplo, se puede reducir la temperatura de hibridación en incrementos de 5 °C, desde 68 °C hasta 42 °C en un regulador de hibridación que tiene una concentración de Na⁺ de aproximadamente 1 M. Después de la hibridación, se puede lavar el filtro con 2X SSC, 0,5% de SDS a la temperatura de hibridación. Estas condiciones se considera que son condiciones "moderadas" por encima de 50 °C y condiciones "bajas" por debajo de 50 °C. Un ejemplo específico de condiciones de hibridación "moderadas" es cuando la hibridación anterior se realiza a 55 °C. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación se realiza a 45 °C.

En forma alternativa, la hibridación se puede llevar a cabo en reguladores, tales como 6X SSC, que contienen formamida a una temperatura de 42 °C. En este caso, la concentración de formamida en el regulador de hibridación se puede reducir en incrementos del 5% desde 50% hasta 0% para identificar clones que tienen niveles reducidos de homología con respecto a la sonda. Después de la hibridación, se puede lavar el filtro con 6X SSC, 0,5% de SDS a 50°C. Estas condiciones se considera que son condiciones "moderadas" por encima del 25% de formamida y condiciones "bajas" por debajo del 25% de formamida. Un ejemplo específico de condiciones "moderadas" de hibridación es cuando la hibridación anterior se realiza con 30% de formamida. Un ejemplo específico de condiciones de hibridación "de rigurosidad baja" es cuando la hibridación se realiza con 10% de formamida.

Sin embargo, la selección de un formato de hibridación no es crítica - es la rigurosidad de las condiciones de lavado que establecen las condiciones que determinan si un ácido nucleico está dentro del alcance de la divulgación. Las condiciones de lavado usadas para identificar los ácidos nucleicos dentro del alcance de la divulgación incluyen, por ejemplo: una concentración de sal de aproximadamente 0,02 molar a un pH de 7 y una temperatura de al menos aproximadamente 50 °C o aproximadamente 55 °C hasta aproximadamente 60 °C; o, una concentración de sal de aproximadamente 0,15 M de NaCl a 72 °C durante aproximadamente 15 minutos; o una concentración de sal de aproximadamente 0,2X SSC a una temperatura de al menos aproximadamente 50 °C o aproximadamente 55 °C hasta aproximadamente 60 °C durante aproximadamente 15 hasta aproximadamente 20 minutos; o, el complejo de hibridación se lava dos veces con una solución con una concentración de sal de aproximadamente 2X SSC que contiene 0,1% de SDS a temperatura ambiente durante 15 minutos y luego se lava dos veces por 0,1X SSC que contiene 0,1% de SDS a 68 °C durante 15 minutos; o condiciones equivalentes. Véase Sambrook, Tijssen y Ausubel para una descripción del regulador de SSC y condiciones equivalentes.

Estos métodos se pueden usar para aislar los ácidos nucleicos de la divulgación.

Sondas de oligonucleótidos y métodos para usarlas

La divulgación también provee sondas de ácido nucleico que se pueden usar, por ejemplo, para identificar ácidos nucleicos que codifican un polipéptido con actividad de amilasa o fragmentos del mismo, o para identificar genes de amilasa. En un aspecto, la sonda comprende al menos 10 bases consecutivas de un ácido nucleico de la divulgación. De manera alternativa, una sonda de la divulgación puede ser al menos de aproximadamente 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 150 o aproximadamente de 10 a 50, aproximadamente de 20 a 60, aproximadamente de 30 a 70 bases consecutivas de una secuencia como se expone en un ácido nucleico de la divulgación. Las sondas identifican un ácido nucleico por enlazamiento y/o hibridación. Las sondas se pueden usar en los arreglos de la divulgación, véase la discusión a continuación, incluyendo, por ejemplo, arreglos capilares. Las sondas de la divulgación también se pueden usar para aislar otros ácidos nucleicos o polipéptidos.

Las sondas de la divulgación se pueden usar para determinar si una muestra biológica, tal como una muestra de suelo, contiene un organismo que tiene una secuencia de ácido nucleico de la invención o un organismo a partir del cual se obtuvo el ácido nucleico. En tales procedimientos, se obtiene una muestra biológica que potencialmente aloja al organismo a partir del cual se aisló el ácido nucleico, y se obtienen los ácidos nucleicos de la muestra. Los ácidos nucleicos se ponen en contacto con la sonda bajo condiciones que permiten que la sonda se hibride específicamente con cualquier secuencia complementaria presente en la muestra. Cuando sea necesario, se pueden determinar las condiciones que permiten que la sonda se hibride específicamente con secuencias complementarias al poner la sonda en contacto con las secuencias complementarias de las muestras que se sabe que contienen la secuencia complementaria, así como también secuencias de control que no contienen la secuencia complementaria. Las condiciones de hibridación, tales como la concentración de sal del regulador de hibridación, la concentración de formamida del regulador de hibridación, o la temperatura de hibridación, se pueden variar para identificar las condiciones que permiten que la sonda se hibride específicamente con ácidos nucleicos complementarios (véase discusión sobre las condiciones de hibridación específicas).

Si la muestra contiene el organismo a partir del cual se aisló el ácido nucleico, se detecta luego la hibridación específica de la sonda. La hibridación puede ser detectada mediante la marcación de la sonda con un agente detectable tal como un isótopo radioactivo, un colorante fluorescente o una enzima capaz de catalizar la formación de un producto detectable. Muchos métodos para usar las sondas marcadas para detectar la presencia de ácidos nucleicos complementarios en una muestra son familiares para aquellos capacitados en el arte. Estos incluyen transferencias del tipo Southern, transferencias del tipo Northern, procedimientos de hibridación de colonias y transferencias puntuales. Los protocolos para cada uno de estos procedimientos son proporcionados en Ausubel y Sambrook.

En forma alternativa, se puede usar más de una sonda (al menos una que sea capaz de hibridarse en forma específica con cualquier secuencia complementaria presente en la muestra de ácido nucleico), en una reacción de amplificación para determinar si la muestra contiene un organismo que contiene una secuencia de ácido nucleico de la divulgación (por ejemplo, un organismo a partir del cual se aisló el ácido nucleico). En un aspecto, las sondas comprenden oligonucleótidos. En un aspecto, la reacción de amplificación puede comprender una reacción PCR. Los protocolos de PCR se describen en Ausubel y Sambrook (véase la discusión sobre las reacciones de amplificación). En tales procedimientos, los ácidos nucleicos en la muestra se ponen en contacto con las sondas, se realiza la reacción de amplificación y se detecta cualquier producto de amplificación. El producto de amplificación se puede detectar mediante la realización de electroforesis en gel sobre los productos de reacción y tificando el gel con un intercalador tal como bromuro de etidio. En forma alternativa, se pueden marcar una o más de las sondas con un isótopo radioactivo y se puede detectar la presencia de un producto de amplificación radioactivo mediante autorradiografía después de la electroforesis de gel.

Las sondas derivadas de las secuencias cerca de los extremos 3' o 5' de una secuencia de ácido nucleico de la invención también se pueden usar en procedimientos de paseo cromosómico para identificar los clones que contienen secuencias adicionales, por ejemplo, secuencias genómicas. Tales métodos permiten el aislamiento de genes que codifican proteínas adicionales de interés del organismo huésped.

En un aspecto, las secuencias de ácido nucleico de la divulgación se usan como sondas para identificar y aislar ácidos nucleicos relacionados. En algunos aspectos, los ácidos nucleicos relacionados así identificados pueden ser ADNc o ADN genómicos de organismos diferentes de aquellos a partir de los cuales se aisló primero el ácido nucleico de la invención. En dichos procedimientos, se pone en contacto una muestra de ácido nucleico con la sonda bajo condiciones que permiten que la sonda se hibride en forma específica con las secuencias relacionadas. La hibridación de la sonda con ácidos nucleicos del organismo relacionado se detecta luego usando cualquiera de los métodos descritos anteriormente.

En las reacciones de hibridación de ácido nucleico, las condiciones usadas para alcanzar un nivel particular de rigurosidad pueden variar, dependiendo de la naturaleza de los ácidos nucleicos que están siendo hibridados. Por ejemplo, se pueden considerar la longitud, el grado de complementariedad, la composición de la secuencia de nucleótidos (por ejemplo, el contenido de GC versus el de AT) y el tipo de ácido nucleico (por ejemplo, ARN versus ADN) de la regiones de hibridación de los ácidos nucleicos al seleccionar las condiciones de hibridación. Una consideración adicional es si se inmoviliza uno de los ácidos nucleicos, por ejemplo, sobre un filtro. La hibridación se puede llevar a cabo bajo condiciones de rigurosidad baja, rigurosidad moderada o rigurosidad alta. Como ejemplo de hibridación de ácido nucleico, se hibrida previamente primero una membrana polimérica que contiene ácidos nucleicos desnaturalizados inmovilizados durante 30 minutos a 45 °C en una solución que consiste de 0,9 M de NaCl, 50 mM de NaH₂PO₄, pH 7,0, 5,0 mM de Na₂EDTA, 0,5% de SDS, 10X reactivo de Denhardt y 0,5 mg/ml de ácido polirriboadenílico. Se pueden añadir luego a la solución aproximadamente 2 X 10⁷ cpm (actividad específica 4-9 X 10⁸ cpm/μg) de una sonda de oligonucleótido marcada en el extremo con ³²P. Después de 12-16 horas de incubación, se lava la membrana durante 30 minutos a temperatura ambiente (RT) en 1X SET (150 mM de NaCl, 20 mM de clorhidrato de Tris, pH 7,8, 1 mM de Na₂EDTA) que contiene 0,5% de SDS, seguido de un lavado de 30 minutos en 1X SET fresco a Tm-10 °C para la sonda de oligonucleótido. Se expone luego la membrana a una película autorradiográfica para detección de las señales de hibridación.

Al variar la rigurosidad de las condiciones de hibridación usadas para identificar ácidos nucleicos, tales como ADNc o ADN genómicos, que hibridan con la sonda detectable, los ácidos nucleicos que tienen diferentes niveles de homología con respecto a la sonda pueden ser identificados y aislados. La rigurosidad puede variar al realizarse la hibridación a diferentes temperaturas por debajo de las temperaturas de fusión de las sondas. La temperatura de fusión, T_m , es la temperatura (bajo una fuerza iónica y pH definidos) a la cual 50% de la secuencia objetivo se hibrida con una sonda perfectamente complementaria. Se seleccionan condiciones muy rigurosas para que sean iguales o aproximadamente 5 °C inferiores a la T_m para una sonda particular. La temperatura de fusión de la sonda se puede calcular usando las siguientes fórmulas de ejemplo. Para las sondas entre 14 y 70 nucleótidos de longitud la temperatura de fusión (T_m) se calcula usando la fórmula: $T_m = 81,5 + 16,6(\log [Na^+]) + 0,41(\text{fracción G+C}) - (600/N)$ donde N es la longitud de la sonda. Si la hibridación se lleva a cabo en una solución que contiene formamida, se puede calcular la temperatura de fusión usando la ecuación: $T_m = 81,5 + 16,6(\log [Na^+]) + 0,41(\text{fracción G+C}) - (0,63\% \text{ de formamida}) - (600/N)$ donde N es la longitud de la sonda. Se puede llevar a cabo la prehibridación en 6X SSC, 5X reactivo de Denhardt, 0,5% de SDS, 100 µg/ml de ADN de esperma de salmón fragmentado desnaturalizado o 6X SSC, 5X reactivo de Denhardt, 0,5% de SDS, 100 µg/ml de ADN de esperma de salmón fragmentado desnaturalizado, 50% de formamida. Las fórmulas para SSC y el reactivo de Denhardt y otras soluciones se encuentran enlistadas, por ejemplo, en Sambrook.

La hibridación se realiza al agregar la sonda detectable a las soluciones de prehibridación mencionadas anteriormente. Cuando la sonda comprende AND bicatenario, se desnaturaliza antes de la adición a la solución de hibridación. El filtro entra en contacto con la solución de hibridación durante un período de tiempo suficiente para permitir que la sonda se hibride con ADNc o ADN genómicos que contienen secuencias complementarias con los mismos u homólogas a los mismos. Para las sondas de más de 200 nucleótidos de longitud, se puede llevar a cabo la hibridación a 15 - 25 °C por debajo de la T_m . Para las sondas más cortas, tales como las sondas de oligonucleótidos, se puede realizar la hibridación a 5 - 10 °C por debajo de la T_m . En un aspecto, las hibridaciones en 6X SSC se realizan aproximadamente a 68 °C. En un aspecto, las hibridaciones en soluciones que contienen 50% de formamida se realizan aproximadamente a 42 °C. Todas las hibridaciones anteriores serían consideradas bajo condiciones de alta rigurosidad.

Después de la hibridación, se lava el filtro para remover cualquier sonda detectable no enlazada en forma específica. La rigurosidad usada para lavar los filtros también puede variar dependiendo de la naturaleza de los ácidos nucleicos que están siendo hibridados, la longitud de los ácidos nucleicos que están siendo hibridados, el grado de complementariedad, la composición de la secuencia de nucleótidos (por ejemplo, contenido de GC versus AT) y el tipo de ácido nucleico (por ejemplo, ARN versus ADN). Los ejemplos de lavados con condiciones de rigurosidad progresivamente superiores son los siguientes: 2X SSC, 0,1% de SDS a temperatura ambiente durante 15 minutos (rigurosidad baja); 0,1X SSC, 0,5% de SDS a temperatura ambiente durante 30 minutos a 1 hora (rigurosidad moderada); 0,1X SSC, 0,5% de SDS durante 15 a 30 minutos entre la temperatura de hibridación y 68 °C (rigurosidad alta); y 0,15 M de NaCl durante 15 minutos a 72°C (rigurosidad muy alta). Se puede llevar a cabo un lavado final de baja rigurosidad en 0,1X SSC a temperatura ambiente. Los ejemplos anteriores son meramente ilustrativos de un conjunto de condiciones que se pueden usar para lavar los filtros. Alguien capacitado en el arte sabría que existen numerosas recetas para lavados de diferentes rigurosidad.

Los ácidos nucleicos que se hibridaron con la sonda se pueden identificar por autorradiografía u otras técnicas convencionales. El procedimiento anteriormente mencionado puede ser modificado para identificar los ácidos nucleicos que tienen niveles reducidos de homología con respecto a la secuencia de la sonda. Por ejemplo, para obtener ácidos nucleicos de homología reducida con respecto a la sonda detectable, se pueden usar condiciones menos rigurosas. Por ejemplo, la temperatura de hibridación puede ser reducida en incrementos de 5 °C desde 68 °C hasta 42 °C en un regulador de hibridación que tiene una concentración de Na^+ de aproximadamente 1 M. Después de la hibridación, se puede lavar el filtro con 2X SSC, 0,5% de SDS a la temperatura de hibridación. Estas condiciones se consideran condiciones "moderadas" por sobre los 50 °C y condiciones "bajas" por debajo de los 50 °C. Un ejemplo de condiciones de hibridación "moderadas" es cuando la hibridación anterior se realiza a 55 °C. Un ejemplo de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación anterior se realiza a 45 °C.

En forma alternativa, los protocolos de hibridación se pueden llevar a cabo en reguladores, tal como 6X SSC, que contienen formamida a una temperatura de 42 °C. En este caso, la concentración de formamida en el regulador de hibridación puede reducirse en incrementos del 5% desde 50% hasta 0% para identificar clones que tienen niveles decrecientes de homología con respecto a la sonda. Después de la hibridación, se puede lavar el filtro con 6X SSC, 0,5% de SDS a 50 °C. Estas condiciones son consideradas como condiciones "moderadas" por encima del 25% de formamida y condiciones "bajas" por debajo del 25% de formamida. Un ejemplo específico de condiciones de hibridación "moderadas" es cuando la hibridación anterior se realiza con 30% de formamida. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación anterior se realiza con 10% de formamida.

Estas sondas y métodos de la divulgación se pueden usar para aislar ácidos nucleicos que tienen una secuencia con al menos 98% o 99% de identidad de secuencia ("homología") con una secuencia de ácido nucleico de la invención que sirve de ejemplo. Se puede medir la homología usando un algoritmo de alineación, como se discute aquí. Por

ejemplo, los polinucleótidos homólogos pueden tener una secuencia codificadora que es una variante alélica de ocurrencia natural de una de las secuencias codificadoras descritas aquí. Dichas variantes alélicas pueden tener una sustitución, supresión o adición de uno o más nucleótidos cuando se comparan con un ácido nucleico de la invención.

5 Adicionalmente, se pueden usar las sondas y métodos de la divulgación para aislar ácidos nucleicos que codifican polipéptidos que tienen al menos 98% o 99%, de identidad de secuencia (homología) con un polipéptido de la invención, como se determina usando un algoritmo de alineación de secuencias (por ejemplo, tal como el algoritmo FASTA versión 3.0t78 con los parámetros por defecto, o un programa BLAST 2.2.2 con los ajustes de ejemplo como se expone aquí).

10 Inhibición de la expresión de amilasa

La invención provee ácidos nucleicos complementarios (por ejemplo, secuencias antisentido) con las secuencias de ácido nucleico de la invención, por ejemplo, ácidos nucleicos que comprenden ribozimas, ARNsi, ARNmi antisentido. Las secuencias antisentido son capaces de inhibir el transporte, empalme o transcripción de genes que codifican para amilasa y genes que codifican para glucoamilasa. La inhibición se puede efectuar a través del direccionamiento del ADN genómico o ARN mensajero. La transcripción o función del ácido nucleico objetivo se puede inhibir, por ejemplo, por hibridación y/o escisión. Un conjunto particularmente útil de inhibidores provisto por la presente invención incluye oligonucleótidos que son capaces de enlazar ya sea un gen de amilasa o un mensaje; en cualquier caso, evitando o inhibiendo la producción o la función de la amilasa. La asociación puede ser a través de la hibridación de una secuencia específica. Otra clase útil de inhibidores incluye oligonucleótidos que causen inactivación o escisión de mensaje de amilasa. El oligonucleótido puede tener actividad enzimática, que causa dicha escisión, tal como ribozimas. Los oligonucleótidos pueden ser modificado o conjugado químicamente con una enzima o composición capaz de escindir el ácido nucleico complementario. Un conjunto de muchos de tales oligonucleótidos diferentes puede ser cribado para aquellos con la actividad deseada.

Los métodos y composiciones de la invención para inhibir la expresión de amilasa, glucoamilasa, glucosidasa y otras enzimas que hidrolizan polisacáridos pueden tener una variedad de aplicaciones industriales. Por ejemplo, la inhibición de la expresión de glucosidasa puede hacer más lenta o prevenir el deterioro. El deterioro puede ocurrir cuando los polisacáridos, lípidos o polipéptidos, por ejemplo, polisacáridos estructurales, son degradados enzimáticamente. Esto puede conducir a un deterioro, o descomposición de frutas y vegetales. En un aspecto, el uso de composiciones de la divulgación que inhiben la expresión y/o actividad de las glucosidasas, por ejemplo, anticuerpos, oligonucleótidos antisentido, ribozimas y ARNi, se utilizan para hacer más lento o prevenir el deterioro. Por lo tanto, en un aspecto, la divulgación provee métodos y composiciones que comprenden la aplicación sobre una planta o producto vegetal (por ejemplo, una fruta, semilla, raíz, hoja, etc.) anticuerpos, oligonucleótidos antisentido, ribozimas y ARNi de la invención, por ejemplo, para retrasar o prevenir el deterioro, o para otro propósito. Estas composiciones también pueden ser expresadas por la planta (por ejemplo, una planta transgénica) o de otro organismo (por ejemplo, una bacteria u otro microorganismo transformado con un gen glucosidasa de la invención).

Oligonucleótidos antisentido

La divulgación provee oligonucleótidos antisentido capaces de enlazar el mensaje de amilasa que puede inhibir la actividad proteolítica dirigiendo el ARNm. La estrategias para diseñar oligonucleótidos antisentido están bien descritas en la literatura científica y de patente, y el experto en la técnica puede diseñar tales oligonucleótidos de amilasa usando los novedosos reactivos de la divulgación. Por ejemplo, los protocolos de paseo de genes/ mapeo de ARN para cribar los oligonucleótidos antisentido efectivos son conocidos en el arte, véase, por ejemplo, Ho (2000) *Methods Enzymol.* 314: 168 - 183, que describe un ensayo de mapeo de ARN que se basa en técnicas moleculares estándar para proveer un método fácil y confiable para selección de una secuencia antisentido potente. Véase también Smith (2000) *Eur. J. Pharm. Sci.* 11: 191 - 198.

Los ácidos nucleicos de ocurrencia natural se usan como oligonucleótidos antisentido. Los oligonucleótidos antisentido pueden ser de cualquier longitud; por ejemplo, en aspectos alternativos, los oligonucleótidos antisentido están entre aproximadamente 5 a 100, aproximadamente 10 a 80, aproximadamente 15 a 60, aproximadamente 18 a 40. La longitud óptima puede ser determinada por cribado de rutina. Los oligonucleótidos antisentido pueden estar presentes en cualquier concentración. La concentración óptima se puede determinar por cribado de rutina. Se conocen una amplia variedad de análogos de ácidos nucleicos y nucleótidos de origen no natural, sintéticos, que pueden manejar este problema potencial. Por ejemplo, se pueden usar ácidos nucleicos peptídicos (PNA) que contienen estructuras base no iónicas, tales como unidades de N-(2-aminoetil) glicina. También se pueden utilizar oligonucleótidos antisentido que tienen enlaces fosforotioato, como se describe en los documentos WO 97/03211; WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144: 189 - 197; *Antisense Therapeutics*, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Los oligonucleótidos antisentido que tienen análogos de la estructura base de ADN sintético provistos por la divulgación, también pueden incluir ácidos nucleicos de fosforditioato, metilfosfonato,

fosforamidato, fosfotriester de alquilo, sulfamato, 3'-tioacetato, metileno(metilimino), 3'-N-carbamato y morfolino carbamato, como se describió anteriormente.

Se puede usar la metodología de química combinatoria para crear un gran número de oligonucleótidos que pueden ser fácilmente cribados para los oligonucleótidos específicos que tienen especificidades y afinidades de
 5 enlazamiento apropiadas hacia cualquier objetivo, tal como las secuencias de amilasa sentido y antisentido de la invención (véase, por ejemplo, Gold (1995) J. of Biol. Chem. 270: 13581 - 13584).

Ribozimas inhibidoras

La divulgación provee ribozimas capaces de enlazar un mensaje de amilasa. Estas ribozimas pueden inhibir la actividad de amilasa, por ejemplo, dirigiendo el ARNm. Las estrategias para diseñar ribozimas y seleccionar la
 10 secuencia antisentido específica de amilasa para direccionamiento son descritas en la literatura científica y de patentes, y el experto en la técnica puede diseñar tales ribozimas usando los novedosos reactivos de la divulgación. Las ribozimas actúan enlazándose a un ARN objetivo a través de la porción de enlazamiento al ARN objetivo de una ribozima que se mantiene muy cerca a una porción enzimática del ARN que escinde el ARN objetivo. Por lo tanto, la ribozima reconoce y enlaza un ARN objetivo a través de apareamiento de base complementarias y, una vez
 15 enlazadas al sitio correcto, actúa en forma enzimática para escindir e inactivar el ARN objetivo. La escisión de un ARN objetivo de dicha forma destruirá su habilidad de dirigir la síntesis de una proteína codificada si la escisión ocurre en la secuencia codificadora. Una vez que una ribozima ha enlazado y escindido su ARN objetivo, puede ser liberada de ese ARN para enlazar y escindir nuevos objetivos en forma repetida.

En algunas circunstancias, la naturaleza enzimática de una ribozima puede ser ventajosa sobre otras tecnologías, tales como tecnología antisentido (donde una molécula de ácido nucleico simplemente se enlaza con un ácido nucleico objetivo para bloquear su transcripción, traducción o asociación con otra molécula) ya que la concentración efectiva de ribozima necesaria para efectuar un tratamiento terapéutico puede ser menor e aquella de un oligonucleótido antisentido. Esta ventaja potencial refleja la habilidad de la ribozima de actuar enzimáticamente. Por lo tanto, una molécula única de ribozima es capaz de escindir muchas moléculas de ARN objetivo. Además, una
 25 ribozima es típicamente un inhibidor altamente específico, con la especificidad de inhibición que depende no solamente del mecanismo de enlazamiento de apareamiento de bases, sino también del mecanismo mediante el cual la molécula inhibe la expresión del ARN con el cual se enlaza. Es decir, la inhibición es causada por la escisión del ARN objetivo y, por lo tanto se define la especificidad como la relación de la velocidad de escisión del ARN objetivo sobre la velocidad de escisión del ARN no objetivo. Este mecanismo de escisión depende de factores adicionales a aquellos implicados en el apareamiento de bases. Por lo tanto, la especificidad de acción de una
 30 ribozima puede ser mayor que aquella del oligonucleótido antisentido que enlaza al mismo sitio del ARN.

La ribozima, de la divulgación por ejemplo, una molécula de ARN de ribozima enzimática, se puede formar en un motivo cabeza de martillo, un motivo de horquilla, tal un motivo del virus de la hepatitis delta, un motivo del intrón del grupo I y/o un ARN del tipo de la ARNasaP junto con una secuencia guía de ARN. Los ejemplos de motivos cabeza de martillo son descritas, por ejemplo, por Rossi (1992) Aids Research and Human Retroviruses 8:183; motivos de horquilla por Hampel (1989) Biochemistry 28:4929 y Hampel (1990) Nuc. Acids Res. 18:299; el motivo del virus de la hepatitis delta por Perrotta (1992) Biochemistry 31:16; el motivo ARNasaP por Guerrier-Takada (1983) Cell 35:849; y el intrón del grupo I por Cech, Patente de los Estados Unidos No. 4.987.071. Aquellos capacitados en el arte se darán cuenta que una ribozima de la divulgación, por ejemplo, una de molécula enzimática de ARN de esta
 35 invención, puede tener un sitio de enlazamiento específico del sustrato complementario a una o más de las regiones de ARN del gen objetivo. Una ribozima de la divulgación puede tener una secuencia de nucleótidos dentro o alrededor de ese sitio de enlazamiento del sustrato que imparte una actividad de escisión al ARN con la molécula.

ARN de interferencia (ARNi)

En un aspecto, la divulgación provee una molécula de ARN de inhibición, así llamada molécula de "ARNi", comprende una secuencia de una enzima amilasa de la invención (que incluye tanto hebras sentido como antisentido). La molécula de ARNi puede comprender una molécula de ARN bicatenaria (ARNbi), por ejemplo, ARNsi y/o ARNmi. La molécula de ARNi comprende una molécula de ARN bicatenario (ARNbi). El ARNi puede inhibir la expresión de un gen de amilasa. En un aspecto, el ARNi, por ejemplo, el ARNsi y/o el ARNmi, es aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más de nucleótidos dúplex de longitud. El
 50 ARNi puede entrar a una célula y causar la degradación de un ARN monocatenario (ARNmc) de secuencias similares o idénticas, incluyendo ARNm endógeno. Cuando se expone una célula a ARN bicatenario (ARNbi), se degrada selectivamente el ARNm del gen homólogo mediante un proceso llamado ARN de interferencia (ARNi). Un posible mecanismo básico detrás del ARNi es el rompimiento de un ARN bicatenario (ARNbi) que coincide con una secuencia de un gen específico en pequeñas piezas llamadas ARN de interferencia corta, lo que desencadena la degradación del ARNm que coincide con su secuencia. En un aspecto, los ARNi de la invención se utilizan en terapias de silenciamiento génico, véase, por ejemplo, Shuey (2002) Drug Discov. Today 7:1040-1046. En un aspecto, la divulgación provee métodos para degradar selectivamente el ARN utilizando los ARNi de la invención. Los procesos se pueden llevar a cabo *in vitro*, *ex vivo* o *in vivo*. En un aspecto, las moléculas de ARNi de la

divulgación se pueden utilizar para generar una mutación de pérdida de función en una célula, un órgano o un animal. Los métodos para elaborar y utilizar moléculas de ARNi, por ejemplo, ARNsi y/o ARNm, para degradar selectivamente ARN son bien conocidos en el arte, véase, por ejemplo, las Patentes de los Estados Unidos No. 6.506.559; 6.511.824; 6.515.109; 6.489.127.

5 Modificación de ácidos nucleicos

La divulgación provee métodos de generación de variantes de los ácidos nucleicos de la divulgación, por ejemplo, aquellos que codifican una amilasa. Estos métodos se pueden repetir o usar en diferentes combinaciones para generar amilasas que tienen una actividad alterada o diferente o una estabilidad alterada o diferente de aquella de una amilasa codificada por el ácido nucleico plantilla. Estos métodos también se pueden repetir o usar en diferentes combinaciones, por ejemplo, para generar variaciones en la expresión génica/del mensaje, traducción del mensaje o estabilidad del mensaje. En otro aspecto, la composición genética de una célula es alterada, por ejemplo, por modificación de un gen homólogo *ex vivo*, seguido de su re inserción en la célula.

Un ácido nucleico de la divulgación, puede ser alterado por cualquier medio. Por ejemplo, métodos aleatorios o estocásticos, o no estocásticos, o métodos de "evolución dirigida", véase, por ejemplo, la Patente de los Estados Unidos No. 6,361,974. Los métodos para mutación aleatoria de genes son bien conocidos en el arte, véase, por ejemplo, la Patente de los Estados Unidos No. 5.830.696. Por ejemplo, se puede utilizar mutagénesis para mutar aleatoriamente un gen. La mutagénesis incluyen, por ejemplo, luz ultravioleta o irradiación gamma, o un mutágeno químico, por ejemplo, mitomicina, ácido nitroso, psoralenos fotoactivados, solos o en combinación, para inducir rompimientos de ADN susceptibles de reparación por recombinación. Otros mutágenos químicos incluyen, por ejemplo, bisulfito de sodio, ácido nitroso, hidroxilamina, hidrazina o ácido fórmico. Otros mutágenos son análogos de precursores nucleótidos, por ejemplo, nitrosoguanidina, 5-bromouracilo, 2-aminopurina o acridina. Estos agentes se pueden añadir a una reacción PCR en lugar de los precursores nucleótidos mutando así la secuencia. También se pueden utilizar agentes intercalación tales como proflavina, acriflavina, quinacrina y similares.

Se puede utilizar cualquier técnica en biología molecular, por ejemplo, mutagénesis por PCR aleatoria, véase, por ejemplo, Rice (1992) Proc. Natl. Acad. Sci. EUA 89:5467 - 5471; o mutagénesis combinatoria de múltiples casetes, véase, por ejemplo, Cramer (1995) Biotechniques 18:194 - 196. Alternativamente, se pueden reensamblar ácidos nucleicos, por ejemplo, genes, después de una fragmentación aleatoria, o "estocástica", véase, por ejemplo, las Patentes de los Estados Unidos Nos. 6.291.242; 6.287.862; 6.287.861; 5,955.358; 5.830.721; 5.824.514; 5.811.238; 5.605.793. En aspectos alternativos, se introducen modificaciones, adiciones o supresiones mediante PCR propensa a error, arrastre, mutagénesis dirigida a oligonucleótido, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casetes, mutagénesis de ensamblaje recursivo, mutagénesis de ensamblaje exponencial, mutagénesis específica de sitio, reensamblaje génico, mutagénesis saturada del sitio génico (GSSM), reensamblaje de ligación sintético (SLR), recombinación, recombinación de secuencia recursiva, mutagénesis de ADN modificado por fosfoato, mutagénesis de plantilla que contiene uracilo, mutagénesis dúplex con espacios, mutagénesis de reparación por falta de coincidencia puntual, mutagénesis de hebra huésped de reparación deficiente, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de selección de restricción, mutagénesis de purificación de restricción, síntesis génica artificial, mutagénesis de ensamblaje, creación del multímeros de ácido nucleico quimérico, y/o una combinación de estos y otros métodos.

Las siguientes publicaciones describen una variedad de procedimientos y/o métodos de recombinación recursivos, los cuales pueden ser incorporados en los métodos de la divulgación: Stemmer (1999) "Molecular breeding of viruses for targeting and other clinical properties" Tumor Targeting 4:1 - 4; Ness (1999) Nature Biotechnology 17:893 - 896; Chang (1999) "Evolution of a cytokine using DNA family shuffling" Nature Biotechnology 17:793 - 797; Minshull (1999) "Protein evolution by molecular breeding" Current Opinion in Chemical Biology 3:284 - 290; Christians (1999) "Directed evolution of timidine kinase for AZT fosforylation using ADN family shuffling" Nature Biotechnology 17:259 - 264; Cramer (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391:288-291; Cramer (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling", Nature Biotechnology 15:436 - 438; Zhang (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" Proc. Natl. Acad. Sci. USA 94:4504 - 4509; Patten y colaboradores (1997) "Applications of DNA shuffling to Pharmaceuticals and Vaccines" Current Opinion in Biotechnology 8:724 - 7313; Cramer y colaboradores (1996) "Construction and evolution of antibody- phage libraries by DNA shuffling" Nature Medicine 2:100 - 103; Gates y colaboradores (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer'" Journal of Molecular Biology 255:373 - 386; Stemmer (1996) "Sexual PCR and Assembly PCR" En: The Encyclopedia of Molecular Biology. VCH Publishers, New York, páginas 447 - 457; Cramer y Stemmer (1995) "Combinatorial multiple cassette mutagénesis creates all the permutations of mutant and wildtype cassettes" Biotechniques 18:194 - 195; Stemmer y colaboradores (1995) "Single-step assembly of a gene and entire plasmid form large numbers of oligodeoxyribonucleotides" Gene. 164:249 - 53; Stemmer (1995) "The Evolution of Molecular Computation" Science 270: 1510; Stemmer (1995) "Searching Sequence Space" Bio/Technology 13:549 - 553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" Nature 370:389 - 391; y Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution." Proc. Natl. Acad. Sci. EUA 91:10747 - 10751.

Los métodos mutacionales de generación de diversidad incluyen, por ejemplo, mutagénesis dirigida a sitio (Ling y colaboradores (1997) "Approaches to DNA mutagénesis: an overview" *Anal Biochem.* 254(2): 157 - 178; Dale y colaboradores (1996) "Oligonucleotide-directed random mutagénesis usando the phosphorothioate method" *Methods Mol. Biol.* 57:369 - 374; Smith (1985) "In vitro mutagénesis" *Ann. Rev. Genet.* 19:423 - 462; Botstein & Shortle (1985) Strategies and applications of in vitro mutagénesis" *Science* 229:1193 - 1201; Carter (1986) "Site-directed mutagénesis" *Biochem. J.* 237z1-7; y Kunkel (1987) "The efficiency of oligonucleotide directed mutagénesis" en *Nucleic Acids & Molecular Biology* (Eckstein, F. y Lilley, D. M. J. eds., Springer Verlag, Berlín)); mutagénesis utilizando plantillas que contienen uracilo (Kunkel (1985) Rapid and efficient site-specific mutagénesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* 82:488 - 492; Kunkel y colaboradores (1987) "Rapid and efficient site-specific mutagénesis without phenotypic selection" *Methods in Enzymol.* 154, 367 - 382; y Bass y colaboradores (1988) "Mutant Trp repressors with new ADN-binding specificities" *Science* 242:240 - 245); mutagénesis dirigida al oligonucleótido (Methods' in *Enzymol.* 100: 468 - 500 (1983); *Methods in Enzymol.* 154: 329 - 350 (1987); Zoller Smith (1982) "Oligonucleotide-directed mutagénesis usando M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" *Nucleic Acid Res.* 10:6487 - 6500; Zoller & Smith (1983) "Oligonucleotide-directed mutagénesis of DNA fragments cloned into M13 vectors" *Methods in Enzymol.* 100:468-500; y Zoller & Smith (1987) Oligonucleotide-directed mutagénesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" *Methods in Enzymol.* 154:329 - 350); mutagénesis de ADN modificado por fosforotioato (Taylor y colaboradores (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked ADN" *Nucl. Acids Res.* 13: 8749 - 8764; Taylor y colaboradores (1985) "The rapid generation of oligonucleotide- directed mutations at high frequency using phosphorothioate -modified ADN" *Nucl. Acids Res.* 13: 8765 - 8787 (1985); Nakamaye (1986) "inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagénesis " *Nucl. Acids Res.* 14: 9679-9698; Sayers y colaboradores (1988) "Y-T Exonuc/eases in phosphorothioate-based oiigonucleotide- directed mutagénesis" *Nucl. Acids Res.* 16:791 - 802; y Sayers y colaboradores (1988) "Strand specific cleavage of phosphorothioate-containing ADN by reaction with restriction endonucieases in the presence of etidium bromide" *Nucl. Acids Res.* 16; 803 - 814); mutagénesis utilizando ADN dúplex con espacios (Kramer y colaboradores (1984) The gapped duplex DNA approach to oiigonucleotide- directed mutagénesis constructions" *Nucl. Acids Res.* 12: 9441 - 9456; Kramer & Fritz (1987) *Methods in Enzymol.* Oligonucleotide-directed mutations constructions via gapped duplex DNA" 154:2350-367; Kramer y colaboradores (1988) "Improved enzymatic in vitro reactions in the gapped duplex ADN approach to oligonuc/eotide-directed mutations constructions " *Nucl. Acids Res.* 16: 7207; y Fritz y colaboradores (1988) "Oligonucleotide-directed mutatio'ns constructions: a gapped duplex DNA procedure without enzymatic reaction *in vitro*" *Nucl. Acids Res.* 16: 6987 - 6999).

Los protocolos adicionales que pueden ser utilizados incluyen reparación por falta de coincidencia puntual (Kramer (1984) "Point Mismatch Repair" *Cell* 38:879 - 887), mutagénesis que utiliza hebras huésped deficientes en reparación (Carter y colaboradores (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" *Nucl. Acids Res.* 13: 4431 - 4443; y Carter (1987) "Improved oligonucleotide- directed mutagenesis using M13 vectors" *Methods in Enzymol.* 154: 382 - 403), mutagenesis por supresión (Eghtedarzadeh (1986) "Use of oligonucleotides for generating large deletions" *Nucl. Acids Res.* 14: 5115), selección por restricción y selección por restricción y purificación por restricción (Wells y colaboradores (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" *Phil. Trans. R. Soc. Lond. A* 317: 415 - 423), mutagénesis mediante síntesis génica total (Nambiar y colaboradores (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" *Science* 223: 1299 - 1301; Sakamar y Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" *Nucl. Acids Res.* 14: 6361 - 6372; Wells y colaboradores (1985) Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" *Gene* 34:315 - 323; y Grundstrom y colaboradores (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" *Nucl. Acids Res.* 13: 3305 - 3316), reparación por ruptura bicatenaria (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" *Current Opinion in Biotechnology* 4:450 - 455. "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagénesis" *Proc. Natl. Acad. Sci. USA.* 83:27177 - 7181). Los detalles adicionales sobre muchos de los métodos anteriormente pueden encontrarse en *Methods in Enzymology* Volume 154, que también describe controles útiles para resolución de problemas con diferentes métodos de mutagenesis.

Los protocolos que pueden ser utilizados se describen, por ejemplo, en la Patente de los Estados Unidos No. 5.605.793 de Stemmer (Feb. 25, 1997), *Methods for In Vitro Recombination*"; la Patente de los Estados Unidos No. 5.811.238 de Stemmer y colaboradores (Sep. 22, 1998) *Methods for Generating Polinucleotides having Desired Characteristics by Iterative Selection and Recombination*"; la Patente de los Estados Unidos No. 5.830.721 de Stemmer y colaboradores (Nov. 3, 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly"; la Patente de los Estados Unidos No. 5.834.252 de Stemmer, y colaboradores (Nov. 10, 1998) "End- Complementary Polimerase Reaction"; la Patente de los Estados Unidos No. 5.837.458 de Minshull, y colaboradores (Nov. 17. 1998). *Methods and Compositions for Cellular and Metabolic Engineering*"; el documento WO 95/22625, Stemmer y Cramer, *Mutagenesis by Random Fragmentation and Reassembly*"; el documento WO 96/33207 de Stemmer y Lipschutz *End Complementary Polimerase Chain Reaction*"; el documento WO 97/20678 de Stemmer y Cramer *Methods for Generating Polinuc/eotides having Desired Characteristics by Iterative Selection and Recombination*"; el documento WO 97/35966 de Minshull y Stemmer, "Methods and Compositions for Cellular and Metabolic

Engineering"; el documento WO 99/41402 de Punnonen y colaboradores "Targeting of Genetic Vaccine Vectors"; el documento WO 99/41383 de Punnonen y colaboradores "Antigen Library Immunization"; el documento WO 99/41369 de Punnonen y colaboradores "Genetic Vaccine Vector Engineering"; el documento WO 99/41368 de Punnonen y colaboradores "Optimization of Immunomodulatory Properties of Genetic Vaccines"; el documento EP 752008 de Stemmer y Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly"; el documento EP 0932670 de Stemmer "Evolving Cellular ADN Uptake by Recombinación de secuencia recursiva"; el documento WO 99/23107 de Stemmer y colaboradores, "Modification of Virus Tropism and Host Range by Viral Genome Shuffling"; el documento WO 99/21979 de Apt y colaboradores, "Human Papi/iomavirus Vectors"; el documento WO 98/31837 por del Cardíare y colaboradores Evolution of Whole Cel/s and Organisms by Recursive Sequence Recombination"; el documento WO 98/27230 de Patten y Stemmer, "Methods and Compositions for Poiipeptide Engineering"; el documento WO 98/27230 de Stemmer y colaboradores. Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection", el documento WO 00/00632, Methods for Generating Highly Diverse Libraries", el documento WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucieotide Sequence Banks and Resulting Sequences", el documento WO 98/42832 de Arnold y colaboradores, Recombination of Polynucleotide Sequences Using Random or Defined Primers", el documento WO 99/29902 de Arnold y colaboradores, "Method for Creating Polynucleotide and Polypeptide Sequences", el documento WO 98/41653 de Vind, "An in Vitro Method for Construction of a DNA Library", el documento WO 98/41622 de Borchert y colaboradores, "Method for Constructing a Library Using DNA Shuffling", y el documento WO 98/42727 por Pati y Zarling, "Sequence Alterations using Homologous Recombination".

Los protocolos que pueden ser utilizados (que proporcionan detalles relacionados con difernetes métodos que generan diversidad) se describen, por ejemplo, en la Solicitud de Patente de los Estados Unidos serial No. (USSN) 09/407.800, "Shuffling Of Codon Aitered Genes" por Patten y colaboradores presentada el 28 de septiembre de 1999; "Evolution Of Whole Cel/s And Organisms By Recursive Sequence Recombination" por del Cardíare y colaboradores, Patente de los Estados Unidos No. 6.379.964; "O/ígonucleotide Mediated Nucleic Acid Recombination" por Crameri y colaboradores. Patentes de los Estados Unidos Nos. 6.319.714; 6.368.861; 6.376.246; 6.423.542; 6.426.224 y el documento PCT/USOO/01203; "Use Of Codon-Varied Oligonucleotide Synthesis For Synthetic Shuffling" por Welch y colaboradores, Patente de los Estados Unidos No. 6.436.675; "Methods For Making Character Strings, Polynuc/eolides & Polypeptides Having Desired Characteristics" por Selifonov y colaboradores, presentada el 18 de enero de 2000, (PCT/USOO/O1202) y, por ejemplo "Methods For Making Character Strings, Polynucleotides Polypeptides Having Desired Characteristics" por Selifonov y colaboradores, presentada el 18 de julio de 2000 (Solicitud de Patente de los Estados Unidos serial No. 09/618.579); "Methods Of Populating Data Structures For Use In Evolutionary Simulations" por Selifonov y Stemmer, presentada el 18 de enero de 2000 (PCT/USOO/01138); y "Single-Stranded Nuc/eic Acid Template-Mediated Recombination And Nucleic Acid Fragment Isolation" por Affhoiter, presentada el 6 de septiembre de 2000 (Solicitud Patente de los Estados Unidos serial No. 09/656.549); y las Patentes de los Estados Unidos Nos. 6.177.263; 6.153.410.

Se utilizan métodos no estocásticos o de "evolución dirigida" incluyen, por ejemplo, mutagénesis de saturación del sitio génico (GSSM), reensamblaje de ligación sintética (SLR), o una combinación de los mismos para modificar los ácidos nucleicos de la divulgación para generar amilasas con propiedades nuevas o alteradas (por ejemplo, actividad bajo condiciones muy ácidas o alcalinas, temperaturas altas, y similares). Los polipéptidos codificados por los ácidos nucleicos modificados e pueden cribar por una actividad antes de analizar por actividad proteolítica u otra actividad. Se puede utilizar cualquier protocolo o modalidad de ensayo, por ejemplo, utilizando una plataforma de arreglo capilar. Véase, por ejemplo, Patentes de los Estados Unidos Nos. 6.361.974; 6.280.926; 5.939.250.

Mutagenesis por saturación, o GSSM

La divulgación también provee métodos para elaborar enzimas usando mutagénesis de saturación de sitio génico, o GSSM, como se describe aquí y también en las Patentes de los Estados Unidos Nos. 6.171.820 y 6.579.258.

En un aspecto, se utilizan cebadores de codones que contienen una secuencia de N,N,G/T degenerada para introducir mutaciones puntuales en un polinucleótido, por ejemplo una amilasa de la divulgación, para generar un conjunto de polipéptidos de progenie en los cuales está representada un rango completo de sustituciones de aminoácidos individuales en cada posición del aminoácidos, por ejemplo, un residuo de aminoácidos en un sitio activo de una enzima o sitio de enlazamiento de un ligando destinado a ser modificado. Estos oligonucleótidos pueden comprender una primera secuencia homóloga contigua, una secuencia de N,N,G/T degenerada y, opcionalmente, una segunda secuencia homóloga. Los productos de la traducción de la progenie secuencia abajo incluyen todos los cambios de aminoácidos posibles en cada sitio de aminoácidos a lo largo del polipéptido, que la degeneración de la secuencia de N,N,G/T incluye codones para todos los 20 aminoácidos. En un aspecto, uno de tales oligonucleótido degenerados (compuesto por ejemplo de, un casete de N,N,G/T degenerado) se utiliza para someter cada codón original en una plantilla de un polinucleótido madre a un rango completo de sustituciones de codones. En otro aspecto, se utilizan al menos dos casetes degenerados, ya sea en el mismo oligonucleótido o no, para someter al menos dos codones originales en una plantilla de polinucleótido madre a un rango completo de sustituciones de codones. Por ejemplo, más de una secuencia de N,N,G/T puede estar contenida en un oligonucleótido para introducir mutaciones de aminoácidos en más de un sitio. Esta pluralidad de secuencias de

N,N,G/T puede estar directamente contigua, o separada por una o más secuencias adicionales de nucleótidos. En otro aspecto, los oligonucleótidos útiles para introducción de adiciones y supresiones se pueden utilizar ya sea solos o en combinación con los codones que contienen una secuencia de N,N,G/T, para introducir cualquier combinación o permutación de adiciones, supresiones y/o sustituciones de aminoácidos.

- 5 En un aspecto, se realiza una mutagénesis simultánea de dos o más posiciones contiguas de aminoácidos utilizando un oligonucleótido que contiene tripletes contiguos de N,N,G/T, es decir una secuencia degenerada de (N,N,G/T)_n. En otro aspecto, se utilizan casetes degenerados que tienen menos degeneración que la secuencia de N,N,G/T. Por ejemplo, puede ser deseable en algunos casos utilizar (por ejemplo en un oligonucleótido) una secuencia degenerada de un triplete que comprende únicamente un N, en donde dicho N puede estar en la primera, segunda o
10 tercera posición del triplete. Se puede utilizar cualquier otra base incluyendo cualquier combinación y permutación de la misma en las dos posiciones restantes del triplete. Alternativamente, puede ser deseable en algunos casos utilizar (por ejemplo en un oligo) una secuencia degenerada del triplete N,N,N.

- En un aspecto, el uso de tripletes degenerados (por ejemplo, tripletes N,N,G/T) permite una generación sistemática y fácil de un rango completo de posibles aminoácidos naturales (para un total de 20 aminoácidos) en cada uno y cada
15 posición de aminoácidos en un polipéptido (en aspectos alternativos, los métodos también incluyen la generación de menos de todas las sustituciones posibles por residuo de aminoácidos, o codón, posición). Por ejemplo, para un polipéptido de 100 aminoácidos, se pueden generar 2000 especies distintas (es decir 20 posibles aminoácidos por posición X 100 posiciones de aminoácidos). A través del uso de un oligonucleótido o conjunto de oligonucleótidos que contienen un triplete degenerado de N,N,G/T, 32 secuencias individuales pueden codificar para todos los 20
20 posibles aminoácidos naturales. Por lo tanto, en un segundo recipiente de reacción en el cual se somete una secuencia de polinucleótidos madre a mutagénesis por saturación utilizando al menos uno de tales oligonucleótidos, se generan 32 distintos polinucleótidos progenie distintos que codifican 20 polipéptidos distintos. En contraste, el uso de un oligonucleótido no degenerado en mutagénesis dirigida al sitio conduce a únicamente un producto del polipéptido descendiente por recipiente de reacción. Se pueden utilizar opcionalmente oligonucleótidos no
25 degenerados en combinación con los cebadores degenerados divulgados; por ejemplo, se pueden utilizar oligonucleótidos no degenerados para generar mutaciones puntuales específicas en un polinucleótido de trabajo. Esto proporciona un medio para generar mutaciones puntuales silenciosas específicas, mutaciones puntuales que conducen a los correspondientes cambios de aminoácidos y mutaciones puntuales que causan la generación de codones de terminación y la correspondiente expresión de fragmentos de polipéptidos.

- 30 En un aspecto, cada recipiente de reacción de mutagénesis por saturación contiene polinucleótidos que codifican al menos 20 moléculas de polipéptido descendiente (por ejemplo, amilasas) de tal manera que todos los 20 aminoácidos naturales estén representados en una posición específica del aminoácido correspondiente a la posición del codón sometido a mutagénesis en el polinucleótido madre (otros aspectos utilizan menos de todas las 20 combinaciones naturales). Los polipéptidos progenie degenerados 32 veces generados a partir de cada recipiente de reacción de mutagénesis por saturación pueden ser sometidos a amplificación clonal (por ejemplo clonado en un
35 huésped adecuado, por ejemplo, huésped de E. coli, utilizando por ejemplo, un vector de expresión) y sometidos a cribado de expresión. Cuando un polipéptido descendiente individual se identifica mediante cribado para mostrar un cambio favorable en propiedad (cuando se compara con el polipéptido madre, tal como una actividad proteolítica incrementada bajo condiciones ácidas o alcalinas), puede ser secuenciado para identificar la sustitución de
40 aminoácidos correspondientemente favorable contenida en el mismo.

- En un aspecto, después de someter a mutagénesis todas y cada una de las posiciones de aminoácidos en un polipéptido madre utilizando mutagénesis por saturación como se divulgó aquí, se pueden identificar cambios favorables de aminoácidos en más de una posición de aminoácidos. Una o más moléculas progenie nuevas pueden ser generadas, que contienen una combinación de todo o parte de estas sustituciones favorables de aminoácidos.
45 Por ejemplo, si se identifican 2 cambios específicos favorables de aminoácidos en cada una de las 3 posiciones de aminoácidos en un polipéptido, las permutaciones incluyen 3 posibilidades en cada posición (sin cambio en el aminoácido original, y cada uno de los dos cambios favorables) y 3 posiciones. Por lo tanto, existen 3 x 3 x 3 o 27 posibilidades totales, incluyendo 7 que fueron previamente examinadas - 6 mutaciones puntuales individuales (es decir, 2 en cada una de las tres posiciones) y sin cambio en ninguna posición.

- 50 En otro aspecto, se puede utilizar mutagénesis de saturación del sitio junto con otros medios estocásticos o no estocásticos para variar la secuencia, por ejemplo, reensamblaje de ligación sintética (véase más adelante), arrastre, quimerización, recombinación y otros procesos de mutagénesis y agentes de mutagénesis. Esta divulgación provee el uso de cualquier proceso de mutagénesis, incluyendo mutagénesis por saturación, de manera iterativa.

Reensamblaje de ligación sintética (SLR)

- 55 La divulgación provee un sistema de modificación génica no estocástica denominado "reensamblaje de ligación sintética", o simplemente "SLR", un "proceso de evolución dirigida", para generar polipéptidos, por ejemplo, amilasas o anticuerpos de la invención, con propiedades nuevas o alteradas. El SLR es un método para ligar en forma conjunta fragmentos de oligonucleótidos en forma no estocástico. Este método difiere del arrastre estocástico de

oligonucleótido en que los bloques de construcción de ácido nucleico no son arrastrados, concatenados o quimerizados al azar, sino que son ensamblados en forma no estocástica. Véase, por ejemplo, las Patentes de los Estados Unidos Nos. 6.773.900, 6.740.506, 6.713.282, 6.635.449, 6.605.449, 6.537.776.

5 En un aspecto, el SLR comprende las siguientes etapas: (a) proveer un polinucleótido plantilla, en donde el polinucleótido plantilla comprende una secuencia que codifica un gen homólogo; (b) proveer una pluralidad de polinucleótidos de bloques de construcción, en donde los polinucleótidos de bloques de construcción están diseñados para reensamblaje de entrecruzamiento con el polinucleótido plantilla en una secuencia predeterminada y un polinucleótido de bloque de construcción comprende una secuencia que es una variante del gen homólogo y una secuencia homóloga con el polinucleótido plantilla que flanquea la variante de la secuencia; (c) combinar un polinucleótido de bloque de construcción con un polinucleótido plantilla de manera tal que el polinucleótido de bloque de construcción de reensamblajes de entrecruzamiento con el polinucleótido plantilla para generar polinucleótidos que comprenden variaciones de secuencia génica homóloga.

15 El SLR no depende de la presencia de altos niveles de homología entre los polinucleótidos que van a ser reordenados. Por lo tanto, este método se puede usar para generar bibliotecas en forma no estocástica (o conjuntos) de moléculas progenie que comprenden más de 10^{100} quimeras diferentes. El SLR se puede usar para generar bibliotecas que comprenden más de 10^{1000} quimeras de progenie diferentes. Por lo tanto, los aspectos de la presente divulgación incluyen métodos no estocásticos de la producción de un conjunto de moléculas de ácido nucleico quimérico finalizado que rebaja una orden de montaje global que es escogido por el diseño. Este método incluye las etapas de generar mediante el diseño de una pluralidad de bloques de construcción de ácido nucleico específica que tienen extremos ligables mutuamente compatibles reparados, y el ensamblaje de estos bloques de construcción de ácidos nucleicos, de tal manera que se consigue una orden general de ensamblaje diseñado

25 Los extremos que pueden ligarse en forma mutuamente compatible de los bloques de construcción de ácido nucleico que van a ser ensamblados se considera que son "útil" para este tipo de ensamblaje ordenado si permiten que los bloques de construcción se acoplen en órdenes predeterminados. Por lo tanto, el orden de ensamblaje completo en el cual los bloques de construcción de ácido nucleico pueden ser acoplados se especifica mediante el diseño de los extremos que pueden ser ligados. Si se va a utilizar más de una etapa de ensamblaje, entonces el orden de ensamblaje completo en el cual los bloques de construcción de ácido nucleico pueden ser acoplados también es especificado por orden secuencial de la(s) etapa(s) de ensamblaje. En un aspecto, las piezas de construcción hibridadas son tratadas con una enzima, tal como una ligasa (por ejemplo ADN ligasa T4), para lograr el enlazamiento covalente de las piezas de construcción.

35 En un aspecto, el diseño de los bloques de construcción de oligonucleótidos se obtiene mediante el análisis de un conjunto de plantillas de secuencias de ácido nucleico progenitores que sirven como base para producir un conjunto de progenie de polinucleótidos quiméricos finalizados. Estas plantillas de oligonucleótidos madres sirven por lo tanto como fuente de información de secuencia que ayuda en el diseño de bloques de construcción de ácido nucleico que se van a someter a mutagénesis, por ejemplo, quimerizada o de arrastre. En un aspecto de este método, las secuencias de una pluralidad de plantillas progenitoras de ácido nucleico se alinean para seleccionar uno o más puntos de demarcación. Los puntos de demarcación se pueden localizar en un área de homología, y constan de uno o más nucleótidos. Estos puntos de demarcación son preferentemente compartidos por al menos dos de las plantilla madre. Los puntos de demarcación se pueden utilizar por lo tanto, para delinear los límites de los bloques de construcción de oligonucleótidos para ser generados en orden para reordenar los polinucleótidos madre. Los puntos de demarcación identificados y seleccionados en las moléculas progenitoras sirven como puntos de quimerización potenciales en el ensamblaje de las moléculas finales de la progenie quiméricas. Un punto de demarcación pueden ser un área de homología (que consta de al menos una base nucleótida homóloga) compartida por al menos dos secuencias madre de polinucleótidos. En forma alternativa, un punto de demarcación puede ser un área de homología que es compartida por al menos la mitad de las secuencias madre de polinucleótidos, o, puede ser un área de homología que es compartida por al menos dos tercios de las secuencias madre de polinucleótidos. Aún más preferentemente, un punto de demarcación útil es un área de homología que es compartida por al menos tres cuartas partes de las secuencias madre de polinucleótidos, o, puede ser compartida por al menos todas las secuencias madre de polinucleótidos. En un aspecto, un punto de demarcación es un área de homología que es compartida por todas las secuencias madre de polinucleótidos.

55 En un aspecto, se realiza un proceso de reensamblaje de ligación exhaustivamente con el fin de generar una biblioteca exhaustiva de la progenie de polinucleótidos quiméricos. En otras palabras, todas las combinaciones posibles combinaciones ordenadas de los bloques de construcción de ácido nucleico están representadas en el conjunto de moléculas de ácido nucleico quimérico finalizado. Al mismo tiempo, en otro aspecto, el orden de ensamblaje (es decir el orden de ensamblaje de cada bloque de construcción en la secuencia 5' a 3' de cada ácido nucleico quimérico finalizado) en cada combinación se hace por diseño (o en forma no estocástica) como se describió anteriormente. Debido a la naturaleza no estocástica de esta invención, se reduce grandemente la posibilidad de productos secundarios no deseados.

En otro aspecto, el método de reensamblaje de ligación se realiza en forma sistemática. Por ejemplo, el método se realiza a fin de generar una biblioteca sistemáticamente compartimentalizada de moléculas de la progenie, con compartimientos que pueden ser sistemáticamente cribados, por ejemplo uno por uno. En otras palabras, a través del uso selectivo y juicioso de bloques de construcción específicos de ácido nucleico, acoplados con el uso selectivo y juicioso de reacciones de ensamblaje secuencialmente intensificadas, se puede lograr un diseño donde se elaboran conjuntos específicos de productos de la progenie en cada uno de los diferentes recipientes de reacción. Esto permite realizar un procedimiento de examen y cribado sistemático. Por lo tanto, estos métodos permiten examinar sistemáticamente un número potencialmente muy grande de moléculas de la progenie en grupos más pequeños. Debido a esta posibilidad de llevar a cabo quimerizaciones en una forma que sea muy flexible, pero también exhaustiva y sistemática, particularmente cuando existe un bajo nivel de homología entre las moléculas progenitoras, estos métodos permiten la generación de una biblioteca (o conjunto) compuesta de un gran número de moléculas de la progenie. Debido a la naturaleza no estocástica de la presente invención de reensamblaje de ligación, las moléculas generadas de la progenie comprenden preferentemente una biblioteca de moléculas de ácido nucleico quimérico finalizadas que tienen un orden de ensamblaje total que se escoge por diseño. Los métodos de mutagénesis por saturación y de evolución dirigida optimizada también se pueden utilizar para generar diferentes especies moleculares de la progenie. Se aprecia que la divulgación provee libertad de selección y control en cuanto a la selección de los puntos de demarcación, el tamaño y el número de los bloques de construcción de ácido nucleico, y el tamaño y el diseño de los acoplamientos. Se aprecia, adicionalmente, que el requisito de homología intermolecular es muy relajado para la operabilidad de esta divulgación. De hecho, incluso los puntos de demarcación pueden ser elegidos en áreas de poca o ninguna homología intermolecular. Por ejemplo, a causa de bamboleo de codones, es decir la degeneración de codones, sustituciones de nucleótidos se pueden introducir en bloques de construcción de ácido nucleico sin alterar el aminoácido codificado originalmente en la plantilla progenitora correspondiente.

De forma alternativa, se puede alterar un codón de tal manera que se altera la codificación para un aminoácido original. Esta divulgación provee que tales sustituciones se pueden introducir en el bloque de construcción de ácido nucleico con el fin de incrementar la incidencia de puntos de demarcación homólogos intermoleculares y por lo tanto permitir que se logren un mayor número de acoplamientos entre los bloques de construcción, lo que a su vez permite que se generen un mayor número de moléculas quiméricas de la progenie.

En otro aspecto, la naturaleza sintética de la etapa en la cual se generan los bloques de construcción permite el diseño y la introducción de nucleótidos (por ejemplo, uno o más nucleótidos, que pueden ser, por ejemplo, codones o intrones o secuencias reguladoras) que pueden ser posteriormente removidos opcional, en un proceso *in vitro* (por ejemplo por medio de mutagénesis) o en un proceso *in vivo* (por ejemplo mediante la utilización de la habilidad de empalme génico de un organismo huésped). Se aprecia que en muchos casos la introducción de estos nucleótidos también puede ser deseables por muchas otras razones además del beneficio potencial de crear un punto de demarcación útil.

En un aspecto, se utiliza un bloque de construcción de ácido nucleico para introducir un intrón. Por lo tanto, se introducen intrones funcionales en un gen fabricado por el hombre de acuerdo con los métodos descritos aquí. El(los) intrón(es) introducido(s) artificialmente puede(n) ser funcional(es) en células huésped para empalme génico casi en la misma forma que los intrones de origen natural sirven funcionalmente en empalme génico.

40 Sistema de evolución dirigido optimizado

La divulgación provee un sistema de modificación génico no estocástica llamado "sistema de evolución dirigido optimizado" para generar polipéptidos, por ejemplo, amilasas de la invención, con propiedades nuevas o alteradas. La evolución dirigida optimizada está dirigida al uso de ciclos repetitivos de reordenamiento reductivo, recombinación y selección que permite la evolución molecular dirigida de ácidos nucleicos a través de recombinación. La evolución dirigida optimizada permite la generación de una gran población de secuencias quiméricas evolucionadas, en donde la población generada está significativamente enriquecida por secuencias que tienen un número predeterminado de eventos de entrecruzamiento.

Un evento de entrecruzamiento es un punto en una secuencia quimérica donde ocurre un cambio en la secuencia de una variante madre por otra variante madre. Tal punto está normalmente en la unión donde los oligonucleótidos de dos progenitores se ligan entre sí para formar una secuencia única. Este método permite el cálculo de las concentraciones correctas de secuencias de oligonucleótidos de manera que la población quimérica final de secuencias está enriquecida por el número escogido de eventos de entrecruzamiento. Esto permite un mayor control sobre las variantes quiméricas escogidas que tienen un número predeterminado de eventos de entrecruzamiento.

Además, este método proporciona un medio conveniente para explorar una tremenda cantidad de posibles variante de empalme de proteína en comparación con otros sistemas. Previamente, si se generan, por ejemplo, 10^{13} moléculas quiméricas durante una reacción, sería extremadamente difícil evaluar una cantidad tan grande de variantes quiméricas para una actividad particular. Además, una porción significativa de la población de la progenie tendría un número muy alto de eventos de entrecruzamiento que resultarían en proteínas que probablemente

tendrían mayores niveles de una actividad particular. Mediante el uso de estos métodos, se puede enriquecer la población de moléculas quiméricas con aquellas variantes que tienen un número particular de eventos de entrecruzamiento. Por lo tanto, aunque se pueden generar aún 10^{13} moléculas quiméricas durante una reacción, cada una de las moléculas escogidas para análisis adicional probablemente tenga, por ejemplo, únicamente tres eventos de entrecruzamiento. Como la población de la progenie resultante puede ser sesgada para tener un número predeterminado de eventos de entrecruzamiento, se reducen los límites sobre la variedad funcional entre las moléculas quiméricas. Esto permite un número más manejable de variables cuando se calcula qué oligonucleótido de los polinucleótidos madre originales pueden ser responsables por afectar un rasgo particular.

Un método para crear una secuencia de polinucleótidos de la progenie quimérica es crear oligonucleótidos correspondientes a fragmentos o porciones de cada secuencia madre. Cada oligonucleótido incluye preferentemente una región única de superposición para que la mezcla de los oligonucleótidos entre sí resulte en una nueva variante que tenga cada fragmento de oligonucleótido ensamblado en el orden correcto. En forma alternativa, se pueden encontrar los protocolos para llevar a cabo estos métodos de la invención en las patentes de los Estados Unidos Nos. 6.773.900, 6.740.506, 6.713.282, 6.635.449, 6.605.449, 6.537.776, 6.361.974.

El número de oligonucleótidos generados para cada variante madre guarda relación con el número total de entrecruzamientos resultantes en la molécula quimérica que es creada en última instancia. Por ejemplo, se podrían proveer tres variantes de secuencias de nucleótidos madre para someterse a una reacción de ligación a fin de encontrar una variante quimérica que tenga, por ejemplo, mayor actividad a temperatura alta. A manera de ejemplo, se puede generar un conjunto de 50 secuencias de oligonucleótidos que corresponden a cada una de las porciones de cada variante madre. Por lo tanto, durante el proceso de reensamblaje de ligación, podría haber hasta 50 eventos de entrecruzamiento dentro de cada una de las secuencias quiméricas. La probabilidad de que cada uno de los polinucleótidos quiméricos generados contenga oligonucleótidos de cada variante madre en orden alternante es muy baja. Si cada fragmento de oligonucleótido está presente en la reacción de ligación en la misma cantidad molar, es posible que en algunas posiciones, los oligonucleótidos del mismo polinucleótido madre se ligan en forma próxima entre sí y por lo tanto no generen un evento de entrecruzamiento. Si la concentración de cada oligonucleótido de cada progenitor se mantiene constante durante cualquier etapa de ligación en este ejemplo, existe $1/3$ de posibilidad (asumiendo 3 progenitores) de que un oligonucleótido de la misma variante madre se ligue dentro de la secuencia quimérica y no produzca ningún entrecruzamiento.

Por lo tanto, se puede determinar una función de probabilidad de densidad (PDF) para predecir la población de eventos de entrecruzamiento que probablemente ocurrirán durante cada etapa en una reacción de ligación, dado un número determinado de variantes parentales, un número de oligonucleótidos correspondientes a cada variante y las concentraciones de cada variante durante cada etapa en la reacción de ligación. Las estadísticas y matemáticas detrás de la determinación de la PDF se describe más adelante. Mediante la utilización de estos métodos, se puede calcular dicha función de probabilidad de densidad, y por lo tanto, enriquecer la población de la progenie quimérica para un número predeterminado de eventos de entrecruzamiento que resultan de una reacción particular de ligación. Además, se puede determinar un número objetivo de eventos de entrecruzamiento, y se puede programar luego el sistema y se pueden calcular las cantidades de partida de cada oligonucleótido madre durante cada etapa en la reacción de ligación para obtener una función de probabilidad de densidad que se centre en el número predeterminado de eventos de entrecruzamiento. Estos métodos están dirigidos al uso de ciclos repetidos de reordenamiento reductivo, recombinación y selección que permiten la evolución molecular dirigida de un ácido nucleico que codifica un polipéptido a través de recombinación. Este sistema permite la generación de una gran población de secuencias quiméricas evolucionadas, en donde la población generada es significativamente enriquecida por las secuencias que tienen un número predeterminado de eventos de entrecruzamiento. Un evento de entrecruzamiento es un punto en una secuencia quimérica donde ocurre un cambio en la secuencia de una variante madre a otra variante madre. Dicho punto está normalmente en la unión donde los oligonucleótidos de dos progenitores se ligan entre sí para formar una secuencia única. El método permite el cálculo de las concentraciones correctas de secuencias de oligonucleótidos para que la población quimérica final de secuencias sea enriquecida para el número elegido de eventos de entrecruzamiento. Esto provee más control al elegir variantes quiméricas que tienen un número predeterminado de eventos de entrecruzamiento.

Además, estos métodos proporcionan un medio conveniente para explorar una tremenda cantidad del espacio posible de la variante de proteína en comparación con otros sistemas. Usando los métodos descritos aquí se puede enriquecer la población de moléculas quiméricas por aquellas variantes que tienen un número particular de eventos de entrecruzamiento. Por lo tanto, aunque se puede generar aún 10^{13} moléculas quiméricas durante una reacción, cada una de las moléculas elegidas para análisis adicionales tiene más probablemente, por ejemplo, solamente tres eventos de entrecruzamiento. Como la población de la progenie resultante se puede sesgar para tener un número predeterminado de eventos de entrecruzamiento, se reducen los límites sobre la variedad funcional entre las moléculas quiméricas. Esto provee un número más manejable de variables cuando se calcula qué oligonucleótido de los polinucleótidos originales madre puede ser responsables de afectar un rasgo particular.

En un aspecto, el método crea una secuencia de polinucleótidos de progenie quimérica mediante la creación de oligonucleótidos que corresponden a fragmentos o porciones de cada secuencia madre. Cada oligonucleótido

incluye preferentemente una región única de superposición para que la mezcla de los oligonucleótidos entre sí resulte en una nueva variante que tenga cada fragmento de oligonucleótido ensamblado en el orden correcto. Véanse también las patentes de los Estados Unidos Nos. 6.773.900, 6.740.506, 6.713.282, 6.635.449, 6.605.449, 6.537.776, 6.361.974.

5 Determinación de eventos de entrecruzamiento

Los aspectos de la divulgación incluyen un sistema y software que recibe una función deseada de probabilidad de densidad de entrecruzamiento (PDF), el número de genes madre que van a ser reensamblados, y el número de fragmentos en el reensamblaje como entradas. La salida de este programa es un "fragmento de PDF " que se puede utilizar para determinar una receta para producir los genes reensamblados y la PDF de entrecruzamiento estimada de esos genes. El procesamiento descrito aquí se realiza preferentemente en MATLAB^{MR} (The Mathworks, Natick, Massachusetts) un lenguaje de programación y un ambiente de desarrollo para computación técnica.

Procesos iterativos

En la práctica de la divulgación, estos procesos pueden ser repetidos en forma iterativa. Por ejemplo, se identifica, se aísla nuevamente, se modifica de nuevo, se analiza nuevamente con respecto a la actividad un ácido nucleico (o el ácido nucleico) responsable por un fenotipo alterado o nuevo de amilasa. Este proceso puede ser repetido en forma iterativa hasta que se modifica por ingeniería genética un fenotipo deseado. Por ejemplo, se puede modificar por ingeniería genética una ruta catabólica o anabólica bioquímica completa en una célula, incluyendo, por ejemplo, la actividad de hidrólisis del almidón.

En forma similar, si se determina que un oligonucleótido particular no tiene ningún efecto sobre el rasgo deseado (por ejemplo, un nuevo fenotipo de amilasa), puede ser removido como una variable al sintetizar oligonucleótidos madre mayores que incluyen la secuencia e va a ser removida. Como la incorporación de la secuencia en una secuencia más larga previene cualquier evento de entrecruzamiento, no habrá ninguna variación de esta secuencia en los polinucleótidos de la progenie. Esta práctica iterativa de determinar cuáles oligonucleótidos están más relacionados con el rasgo deseado y cuáles no están relacionados, permite la exploración más eficiente de todas las posibles variantes de proteína que pueden proveer un rasgo o actividad particulares.

Arrastre *in vivo*

Se puede utilizar el arrastre *in vivo* de moléculas en métodos de la divulgación que proveen variantes de polipéptidos de la divulgación, por ejemplo, anticuerpos, amilasas y similares. Se puede realizar el arrastre *in vivo* utilizando la propiedad natural de las células para recombinar multimeros. Aunque la recombinación *in vivo* ha proporcionado la principal ruta natural para la diversidad molecular, la recombinación genética sigue siendo un proceso relativamente complejo que involucra 1) el reconocimiento de homologías; 2) etapas de incisión de una hebra, invasión de una hebra y etapas metabólicos que conducen a la producción de quiasma recombinante; y finalmente 3) la resolución del quiasma en moléculas recombinadas discretas. La formación del quiasma requiere el reconocimiento de secuencias homólogas.

En un aspecto, la divulgación provee un método para producir un polinucleótido híbrido a partir al menos un primer polinucleótido (por ejemplo, una amilasa de la invención) y un segundo polinucleótido (por ejemplo, una enzima, tal como una amilasa de la invención o cualquier otra amilasa, o una etiqueta o un epítipo). La divulgación puede ser usada para producir un polinucleótido híbrido mediante la introducción de al menos un primer polinucleótido y un segundo polinucleótido que comparten al menos una región de homología parcial de secuencia en una célula huésped adecuada. Las regiones de homología de secuencia parcial promueven procesos que generan una reorganización de las secuencias produciendo un polinucleótido híbrido. El término "polinucleótido híbrido", como se utiliza aquí, es cualquier secuencia de nucleótidos que se genera a partir del método de la presente divulgación y contiene una secuencia de al menos dos secuencias originales de polinucleótidos. Tales polinucleótidos híbridos pueden resultar de los eventos de recombinación intermolecular que promueven la integración de secuencia entre moléculas de ADN. Además, tales polinucleótidos híbridos pueden resultar de procesos de reordenamiento reductivo intramolecular que utilizan secuencias repetidas para alterar una secuencia de nucleótidos dentro de una molécula de ADN.

Producción de variantes de secuencia

La divulgación también proporciona métodos adicionales para elaborar variantes de secuencias de las secuencias de ácido nucleico (por ejemplo, amilasas) de la divulgación. La divulgación también proporciona métodos adicionales para aislar amilasas utilizando los ácidos nucleicos y polipéptidos de la divulgación. En un aspecto, la divulgación proporciona variantes de una secuencia de codificación de amilasa (por ejemplo, un gen, ADNc o mensaje) de la divulgación, que pueden ser alterados por cualquier medio, incluyendo, por ejemplo, métodos aleatorios o estocásticos, o, métodos no estocástico, o de "evolución dirigida," como se describe más arriba.

Las variantes aisladas pueden ser de origen natural. La variante también se puede crear *in vitro*. Las variantes pueden ser creadas usando técnicas de manipulación genética tales como mutagénesis dirigida al sitio, mutagénesis química aleatoria, procedimientos de supresión de Exonucleasa III y técnicas de clonación estándar. En forma alternativa, dichas variantes, fragmentos, análogos o derivados pueden ser creados utilizando procedimientos de modificación o síntesis química. Otros métodos para elaborar variantes también son familiares para las personas versadas en el arte. Estos incluyen procedimientos en los cuales las secuencias de ácido nucleico obtenidas a partir de aislados naturales son modificadas para generar ácidos nucleicos que codifican polipéptidos que tienen características que refuerzan su valor en las aplicaciones de laboratorio o industriales. En tales procedimientos se generan y caracterizan, un gran número de variantes de secuencias que tienen una o más diferencias de nucleótidos con respecto a la secuencia obtenida del aislado natural. Estas diferencias de nucleótidos pueden resultar en cambios de aminoácidos con respecto a los polipéptidos codificados por los ácidos nucleicos de los aislados naturales.

Por ejemplo, las variantes pueden ser creadas usando PCR propenso a error. En la PCR propensa a error, se realiza la PCR bajo condiciones donde la fidelidad de copiado de la ADN polimerasa es baja, de manera tal que se obtiene un alto índice de mutaciones puntuales lo largo de la longitud total del producto de la PCR. La PCR propensa a error se describe, por ejemplo, en Leung, D.W. y colaboradores, Technical, 1: 11 - 15, 1989) y Caldwell, R. C. Joyce G.F., PCR Methods Applic., 2: 28 - 33, 1992. En resumen, en tales procedimientos, se mezclan los ácidos nucleicos que van a experimentar mutagénesis con cebadores de la PCR, regulador de reacción, MgCl₂, MnCl₂, Taq polimerasa y una concentración apropiada de los dNTP para lograr un alto índice de mutación puntual a lo largo de toda la longitud del producto de la PCR. Por ejemplo, se puede llevar a cabo la reacción utilizando 20 fmoles de ácido nucleico que va a experimentar mutagénesis. 30 pmol de cada cebador de la PCR, un regulador de reacción que comprende 50 mM de KCl, 10 mM de Tris HCl (pH 8,3) y 0,01% de gelatina, 7 mM de MgCl₂, 0.5 mM de MnCl₂, 5 unidades de Taq polimerasa, 0,2 mM de dGTP, 0.2 mM de dATP, 1 mM de dCTP y 1 mM de dTTP. Se puede realizar la PCR por 30 ciclos de 94 °C durante 1 min, 45 °C durante 1 min y 72 °C durante 1 min. Sin embargo, se apreciará que estos parámetros pueden ser variados según sea conveniente. Se clonan los ácidos nucleicos e experimentaron la mutación en un vector apropiado y se evalúan las actividad de los polipéptidos codificados por los ácidos nucleicos que experimentaron la mutación.

También se pueden crear variantes utilizando mutagénesis dirigida al oligonucleótido para generar mutaciones específicas del sitio en cualquier ADN clonado de interés. La mutagénesis de oligonucleótidos s descrita, por ejemplo, en Reidhaar-Olson (1988) Science 241: 53 - 57. En resumen, en tales procedimientos, se sintetizan una pluralidad de oligonucleótidos bicatenarios que portan una o más mutaciones que son introducidas en el ADN clonado e insertadas en el ADN clonado que va a experimentar la mutagénesis. Se recuperan los clones que contienen el ADN que experimentó la mutación y se evalúan las actividades de los polipéptidos que codifican.

Otro método para generar variantes es la PCR de ensamblaje. La PCR de ensamblaje implica el ensamblaje de un producto de PCR a partir de una mezcla de fragmentos pequeños de ADN. Un gran número de diferentes reacciones de PCR se realizan en paralelo en el mismo vial, con los productos de una reacción cebando a los productos de otra reacción. La PCR de ensamblaje se describe en, por ejemplo, la Patente de los Estados Unidos No. 5.965.408.

Incluso otro método para generar variantes es la mutagénesis de la PCR sexual. En la mutagénesis de la PCR sexual, se produce una recombinación homóloga forzada entre moléculas de ADN de secuencia diferentes pero altamente relacionadas de ADN *in vitro*, como resultado de una fragmentación aleatoria en la molécula de ADN con base en homología de secuencia, seguido por la fijación del entrecruzamiento mediante la extensión del cebador en una reacción PCR. Se describe la mutagénesis de PCR sexual, por ejemplo, en Stemmer (1994) Proc. Natl. Acad. Sci. EUA 91: 10747 - 10751. En resumen, en tales procedimientos, se digieren una pluralidad de ácidos nucleicos que van a ser recombinados con ADNasa para generar fragmentos que tienen un tamaño promedio de 50 - 200 nucleótidos. Los fragmentos del tamaño promedio deseado se purifican y resuspenden en una mezcla de la PCR. Se realiza la PCR bajo condiciones que facilitan la recombinación entre los fragmentos de ácido nucleico. Por ejemplo, se puede realizar la PCR mediante la resuspensión de los fragmentos purificados a una concentración de 10 - 30 ng/μl en una solución de 0,2 mM de cada dNTP, 2,2 mM de MgCl₂, 50 mM KCl, 10 mM de Tris HCl, pH 9,0 y 0,1% de Triton X-100. Se añaden 2,5 unidades de Taq polimerasa por 100:1 de mezcla de reacción y se lleva a cabo la PCR utilizando el siguiente régimen: 94 °C durante 60 segundos, 94 °C durante 30 segundos, 50 - 55 °C durante 30 segundos, 72 °C durante 30 segundos (30 - 45 veces) y 72 °C durante 5 minutos. Sin embargo, se apreciará que estos parámetros pueden ser variados según sea conveniente. En algunos aspectos, se pueden incluir oligonucleótidos en las reacciones PCR. En otros aspectos, el fragmento Klenow de la ADN polimerasa I se puede utilizar en un primer conjuntos de reacciones PCR y se puede utilizar Taq polimerasa en un grupo posterior de reacciones PCR. Las secuencias recombinantes se aíslan y se evalúan las actividades de los polipéptidos que las codifican.

Se pueden crear también variantes mediante mutagénesis *in vivo*. En algunos aspectos, se generan mutaciones aleatorias en una secuencia de interés mediante la propagación de la secuencia de interés en una cepa bacteriana, tal como cepa de *E. coli*, que porta mutaciones en una o más de las rutas de reparación del ADN. Tales cepas "mutadoras" tienen una mayor tasa de mutación aleatoria que aquellas del progenitor de tipo silvestre. La

propagación del ADN en una de estas cepas generará eventualmente mutaciones aleatorias dentro del ADN. Las cepas mutadoras adecuadas para uso en mutagénesis *in vivo* se describen, por ejemplo, en la publicación PCT No. WO 91/16427.

5 También se pueden generar variantes utilizando mutagénesis de casetes. En la mutagénesis de casetes, se reemplaza una pequeña región de una molécula de ADN bicatenaria con un "casete" de oligonucleótido sintético que difiere de la secuencia nativa. El oligonucleótido contiene a menudo una secuencia nativa completamente y/o parcialmente aleatoria.

10 También se puede utilizar mutagénesis de ensamble recursivo para generar variantes. La mutagénesis de ensamble recursivo es un algoritmo para modificación genética de la proteína (mutagénesis de proteína) desarrollado para producir diversas poblaciones de mutantes fenotípicamente relacionados cuyos miembros difieren en la secuencia de aminoácidos. Este método utiliza un mecanismo de retroalimentación para controlar rondas sucesivas de mutagénesis de casete combinatorial. La mutagénesis de ensamble recursivo se describe, por ejemplo, en Arkin (1992) Proc. Natl. Acad. Sci. EUA 89:7811 - 7815.

15 En algunos aspectos, se crean variantes utilizando mutagénesis de ensamblajes exponenciales. La mutagénesis de ensamblajes exponenciales es un proceso para generar bibliotecas combinatorias con un alto porcentaje de mutantes únicos y funcionales, en donde pequeños grupos de residuos son aleatorizados en paralelo para identificar, en cada posición alterada, los aminoácidos que producen proteínas funcionales. La mutagénesis de ensamble exponencial es descrita, por ejemplo, en Delegrave (1993) Biotechnology Res. 11: 1548 - 1552. Las mutagénesis aleatoria y dirigida a sitio son descritas, por ejemplo, en Arnold (1993) Current Opinion in Biotechnology 20 4: 450 - 455.

En algunos aspectos se crean variantes utilizando variantes de arrastre en donde porciones de una pluralidad de ácidos nucleicos que codifican polipéptidos distintos se fusionan entre sí para crear secuencias de ácido nucleico quimérico que codifican polipéptidos quiméricos como se describe, por ejemplo, en las Patentes de los Estados Unidos Nos. 5.965.408, 5.939.250 (véase también la discusión, más arriba).

25 La divulgación también provee variantes de polipéptidos de la invención (por ejemplo, amilasas) que comprenden secuencias en las cuales se sustituyen uno o más residuos de aminoácidos (por ejemplo, de un ejemplo de un polipéptido de la invención) con un residuo de aminoácido conservado o no conservado (por ejemplo, un residuo de aminoácido conservado) y tal residuo de aminoácido sustituido puede o no ser uno codificado por el código genético. Sustituciones conservadoras son aquellas que sustituyen un aminoácido dado en un polipéptido por otro aminoácido 30 de características similares. Por lo tanto, los polipéptidos de la divulgación incluyen aquellos con sustituciones conservadoras de secuencias de la invención, por ejemplo, los ejemplos de polipéptidos de la invención, incluyendo, pero no limitándose a los siguientes reemplazos: reemplazos de aminoácidos alifáticos tales como Alanina, Valina, Leucina e Isoleucina con otro aminoácido alifático; reemplazo de una Serina con una Treonina o vice versa; el reemplazo de un residuo ácido tal como Acido aspártico y Ácido glutámico con otro residuo ácido; el reemplazo de 35 un residuo que contiene un grupo amida, tal como Asparagina y Glutamina, con otro residuo que contiene un grupo amida; el intercambio de un residuo básico tal como Lisina y Arginina con otro residuo básico; y el reemplazo de un residuo aromático tal como Fenilalanina, Tirosina con otro residuo aromático. Otras variantes son aquellas en las cuales uno o más de los residuos de aminoácidos de los polipéptidos de la invención incluyen un grupo sustituyente. Una sustitución de conservadora de aminoácido también puede comprender la sustitución de un aminoácido por otro de la misma clase (por ejemplo, la sustitución de un aminoácido hidrófobo tal como isoleucina, valina, leucina, o metionina, por otro, o sustitución de un aminoácido polar por otro, tal como la sustitución de arginina por lisina, ácido glutámico por ácido aspártico o de glutamina por asparagina). Se pueden suprimir uno o más aminoácidos, por ejemplo, de una amilasa, lo que resulta en la modificación de la estructura del polipéptido, sin alterar significativa su actividad biológica. Por ejemplo, se pueden remover aminoácidos del terminal amino o el terminal carboxilo que no 45 se requieren para la actividad de la amilasa.

Otras variantes son aquellas en las cuales el polipéptido se asocia con otro compuesto, tal como un compuesto para incrementar la vida media del polipéptido, por ejemplo, polietilenglicol.

50 Variantes adicionales son aquellas en las cuales se fusionan aminoácidos adicionales se fusionan con el polipéptido, tal como una secuencia guía, una secuencia secretoria, una secuencia de proproteína o una secuencia que facilita la purificación, enriquecimiento o estabilización del polipéptido.

En algunos aspectos, las variantes, fragmentos, derivados y análogos de los polipéptidos de la divulgación retienen la misma función o actividad biológica que los ejemplos de polipéptidos, por ejemplo, actividad de amilasas, como se describen aquí. En otros aspectos, la variante, fragmento, derivado o análogo incluye una proproteína, de tal manera que la variante, fragmento, derivado o análogo pueden ser activados por escisión de la porción proproteína para 55 producir un polipéptido activo.

Optimización de codones para lograr altos niveles de expresión proteína en células huésped

La divulgación provee métodos para modificar los ácidos nucleicos que codifican amilasa y que codifican glucoamilasa para modificar el uso del codón y los ácidos nucleicos que codifican una amilasa optimizada por el codón y que codifican glucoamilasa, incluyendo la SEQ ID NO: 79, SEQ ID NO:80, SEQ ID NO: 81 y SEQ ID NO:82; estos ejemplos de ácidos nucleicos que codifican amilasa y que codifican glucoamilasa fueron generados como se discute en detalle en el Ejemplo 28, más adelante, a partir de las SEQ ID NO: 51, SEQ ID NO:3, SEQ ID NO:47 y SEQ ID NO:25, respectivamente

Enzima	SEQ ID NO: de tipo silvestre	SEQ ID NO: de codón optimizado
Amilasa	51	79
Amilasa	3	80
Glucoamilasa	47	81
Glucoamilasa	25	82

En un aspecto, la divulgación provee métodos para modificar codones en un ácido nucleico que codifica una amilasa para incrementar o reducir su expresión en una célula huésped. La divulgación también provee ácidos nucleicos que codifican una amilasa modificada para incrementar su expresión en una célula huésped, amilasa modificada de este modo y métodos para elaborar las amilasas modificadas. El método comprende identificar un codón "no preferido" o uno "menos preferido" en un ácido nucleico que codifica amilasa y que codifica glucoamilasa y reemplazar uno o más de estos codones no preferidos o menos preferidos con un "codón preferido" que codifica el mismo aminoácido que el codón reemplazado y al menos un codón no preferido o menos preferido en el ácido nucleico ha sido reemplazado por un codón preferido que codifica el mismo aminoácido. Un codón preferido es un codón sobrerrepresentado en las secuencias codificadoras en los genes en la célula huésped y un codón no preferido o menos preferido es un codón subrepresentado en secuencias codificadoras en genes en la célula huésped.

Las células huésped para expresar los ácidos nucleicos, casetes de expresión y vectores de la invención incluyen bacterias, levadura, hongos, células vegetales, células de insectos y células de mamíferos. Por lo tanto, la divulgación provee métodos para optimizar el uso de codones en todas estas células, ácidos nucleicos de codón alterado y polipéptidos elaborados por los ácidos nucleicos de codón alterado. Los ejemplos de células huésped incluyen bacterias Gram negativas, tal como cualquier especie del género *Escherichia* o *Pseudomonas*, por ejemplo, *Escherichia coli* y *Pseudomonas fluorescens*; o bacterias Gram positivas, tal como cualquier especie del género *Bacillus*, *Streptomyces*, *Lactococcus*, *Lactobacillus*, por ejemplo, *Bacillus cereus*, *Lactobacillus gasseri*, *Lactococcus lactis*, *Lactococcus cremoris*, *Bacillus subtilis*. Los ejemplos de células huésped también incluyen organismos eucariotas, por ejemplo, diferentes levaduras, tales como *Schizosaccharomyces* sp., *Aspergillus* sp., *Hansenula* sp., *Kluyveromyces* sp., *Pichia* sp. y *Saccharomyces* sp., incluyendo, por ejemplo, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris* y *Kluyveromyces lactis*, *Hansenula polymorpha*, *Aspergillus niger*, células de mamífero y líneas celulares y células de insecto y líneas celulares. Por lo tanto, la divulgación también incluye ácidos nucleicos y polipéptidos optimizados para expresión en estos organismos y especies.

Por ejemplo, los codones de un ácido nucleico que codifica una amilasa aislada de una célula bacteriana se modifican de manera tal que el ácido nucleico sea expresado en forma óptima en una célula bacteriana diferente de la bacteria de la cual se derivó la amilasa, una levadura, un hongo, una célula vegetal, una célula de insecto o una célula de mamífero. Los métodos para optimizar codones son conocidos en el arte, véase, por ejemplo, la patente de los Estados Unidos No. 5.795.737; Baca (2000) Int. J. Parasitol. 30: 113 - 118; Hale (1998) Protein Expr. Purif. 12: 185 - 188; Narum (2001) Infect. Immun. 69: 7250 - 7253. Véase también Narum (2001) Infect. Immun. 69: 7250 - 7253, que describe la optimización de codones en sistemas de ratones; Outchkourov (2002) Protein Expr. Purif. 24: 18 - 24, que describe la optimización de codones en levadura; Feng (2000) Biochemistry 39: 15399 - 15409, que describe la optimización de codones en *E. coli*; Humphreys (2000) Protein Expr. Purif. 20: 252 - 264, que describe la optimización del uso de codones que afecta la secreción en *E. coli*.

Animales no humanos transgénicos

La divulgación provee animales no humanos transgénicos que comprenden un ácido nucleico, un polipéptido (por ejemplo, una amilasa), un casete de expresión o un vector o una célula transfectada o transformada de la divulgación. La divulgación también provee métodos para elaborar y usar estos animales no humanos transgénicos.

Los animales no humanos transgénicos pueden ser, por ejemplo, cabras, conejos, ovejas, cerdos, vacas, ratas y ratones, que comprenden los ácidos nucleicos de la divulgación. Estos animales se pueden usar, por ejemplo, como modelos *in vivo* para estudiar la actividad de la amilasa, o como modelos para cribar agentes que cambian la

actividad de la amilasa *in vivo*. Las secuencias codificadoras para los polipéptidos que van a ser expresados en los animales no humanos transgénicos pueden ser diseñadas para ser constitutivas o para estar bajo el control de factores reguladores de transcripción inducibles o específicos del desarrollo o específicos del tejido. Los animales no humanos transgénicos pueden ser diseñados y generados usando cualquier método conocido en el arte; véase, por ejemplo, las patentes de los Estados Unidos Nos. 6.211.428, 6.187.992, 6.156.952, 6.118.044, 6.111.166, 6.107.541, 5.959.171, 5.922.854, 5.892.070, 5.880.327, 5.891.698, 5.639.940, 5.573.933, 5.387.742, 5.087.571, que describen la elaboración y el uso de células y óvulos transformados, y de ratones, ratas, conejos, ovejas, cerdos y vacas transgénicos. Véase también, por ejemplo, Pollock (1999) *J. Immunol. Methods* 231: 147 - 157, que describe la producción de proteínas recombinantes en la leche de animales lecheros transgénicos; Baguisi (1999) *Nat. Biotechnol.* 17: 456 - 461, que demuestra la producción de cabras transgénicas. La patente de los Estados Unidos No. 6.211.428, describe la elaboración y el uso de mamíferos no humanos transgénicos que expresan en sus cerebros un constructo de ácido nucleico que comprende una secuencia de ADN. La patente de los Estados Unidos No. 5.387.742, describe la inyección de secuencias clonadas de AND recombinante o sintético en óvulos de ratón fertilizados, implantando los óvulos inyectados en hembras pseudopreñadas y permitiendo que se desarrollen hasta término los ratones transgénicos cuyas células expresan proteínas relacionadas con la patología de la enfermedad de Alzheimer. La patente de los Estados Unidos No. 6.187.992, describe la formación y el uso de un ratón transgénico cuyo genoma comprende una ruptura del gen que codifica la proteína precursora amiloidea (APP).

Se pueden utilizar también "animales nuligénicos". Por ejemplo, en un aspecto, los animales transgénicos o modificados de la invención comprenden un "animal nuligénico", por ejemplo, un "ratón nuligénico", manipulado genéticamente para no expresar un gen endógeno, que es reemplazado con un gen que expresa una amilasa de la divulgación, o una proteína de fusión que comprende una amilasa de la divulgación.

Plantas y semillas transgénicas

La divulgación provee plantas y semillas transgénicas que comprenden un ácido nucleico, un polipéptido (por ejemplo, una gluconasa, manasa o xilanasa), un casete vector de expresión o una célula transformada o transfectada de la invención. La divulgación también provee productos vegetales, por ejemplo, aceites, semillas, hojas, extractos y similares, que comprenden un ácido nucleico y/o un polipéptido (por ejemplo, una glucanasa, manasa o xilanasa) de la invención. La planta transgénica puede ser una dicotiledónea (una dicot) o monocotiledónea (una monocot). La divulgación provee plantas transgénicas con un sabor, contenido de sólidos y/o textura modificados, en donde esa modificación es generada al expresar al menos una enzima de la invención ya sea en forma constitutiva o selectiva en la planta transgénica (o semilla, o fruto, etc.), como se describe, por ejemplo, en la solicitud de patente de los Estados Unidos No. 20060195940.

La divulgación también provee métodos para la elaboración y el uso de estas plantas y semillas transgénicas. La planta transgénica o célula vegetal que expresa un polipéptido de la presente divulgación puede ser construida de acuerdo con cualquier método conocido en el arte. Véase, por ejemplo, la patente de los Estados Unidos No. 6.309.872.

Los ácidos nucleicos y constructos de expresión de la divulgación se pueden introducir en una célula vegetal por cualquier medio. El término "introducir" en el contexto de un polinucleótido, por ejemplo, un constructo de nucleótidos de interés, pretende dar a entender que presenta a la planta el polinucleótido de tal manera que el polinucleótido gana acceso al interior de una célula de la planta. Cuando se introduce más de un polinucleótido, estos polinucleótidos pueden ser ensamblados como parte de un único constructo de nucleótidos, o como constructos de nucleótidos separados, y pueden ser ubicados en el mismo o en diferentes vectores de transformación. Por consiguiente, estos polinucleótidos se pueden introducir en la célula huésped de interés en un único evento de transformación, en eventos de transformación separados, o, por ejemplo, en plantas, como parte de un protocolo de fitomejoramiento. Los métodos de la divulgación no dependen de un método particular para introducir uno o más polinucleótidos en una planta, solamente que el(los) polinucleótido(s) gana(n) acceso al interior de al menos una célula de la planta. Los métodos para introducir polinucleótidos en las plantas son conocidos en el arte incluyendo, pero sin limitarse a, métodos de transformación transitoria, métodos de transformación estable y métodos mediados por virus.

La "transformación transitoria" en el contexto de un polinucleótido quiere decir que un polinucleótido es introducido en la planta y no se integra en el genoma de la planta.

Por "introducir de manera estable" o "introducido en forma estable" en el contexto de un polinucleótido introducido en una planta quiere decir que el polinucleótido introducido se incorpora de manera estable en el genoma de la planta y, por lo tanto, la planta es transformada en forma estable con el polinucleótido.

En formas de realización alternativas, "transformación estable" o "transformado en forma estable" quiere decir que un polinucleótido, por ejemplo, un constructo de nucleótido descrito aquí, introducido en una planta se integra en el genoma de la planta y es capaz de ser heredado por la progenie de la misma, más particularmente, por la progenie de múltiples generaciones sucesivas. La introducción en el genoma de una planta deseada puede ser tal que la

enzima es regulada por los elementos de control de transcripción o traducción endógenos. Las técnicas de transformación tanto para las monocotiledóneas como para las dicotiledóneas son conocidas en el arte.

Los ácidos nucleicos de la invención se pueden usar para conferir rasgos deseados esencialmente en cualquier planta. Los ácidos nucleicos de la divulgación pueden ser utilizados para manipular las rutas metabólicas de una planta con el fin de optimizar o alterar la expresión del huésped de glucanasa, mananasa, o xilanas. Pueden cambiar la actividad de la amilasa, glucoamilasa, glucanasa, mananasa, xilanas en una planta. En forma alternativa, una amilasa, glucoamilasa, glucanasa, mananasa, xilanas o de la divulgación puede ser usada en la producción de una planta transgénica para producir un compuesto producido no naturalmente por esa planta. Esto puede reducir los costos de producción o crear un producto novedoso. En una forma de realización, la enzima de la divulgación puede ser expresada de manera tal que la enzima no entrará en contacto con su sustrato hasta que sea deseado. Por ejemplo, una enzima de la divulgación puede ser dirigida y retenida en el retículo endoplasmático de una célula vegetal. La retención de la enzima, en el retículo endoplasmático de las célula, evitará que la enzima entre en contacto con su sustrato. La enzima y el sustrato pueden ser luego puestas en contacto a través de cualquier medio capaz de romper la arquitectura subcelular, tal como, trituración, molienda, calentamiento y similares. Véanse, los documentos WO 98/11235, WO 2003/18766 y WO 2005/096704.

Los genes marcadores seleccionables pueden ser agregados al constructo génico a fin de identificar las células vegetales o tejidos que han integrado exitosamente el transgén. Esto puede ser necesario ya que lograr la incorporación y expresión de genes en las células vegetales es un evento raro, que ocurre solamente en un pequeño porcentaje de los tejidos o células objetivo. Los genes marcadores seleccionables codifican proteínas que proveen resistencia a agentes que son normalmente tóxicos para las plantas, tales como antibióticos o herbicidas. Solamente las células vegetales que han integrado el gen marcador seleccionable sobrevivirán cuando se cultiven en un medio que contiene el antibiótico o herbicida apropiados. Los marcadores de selección usados en forma rutinaria en la transformación incluyen el gen nptII, que confiere resistencia a la kanamicina y antibióticos relacionados (Messing & Vierra. *Gene* 19: 259 - 268 (1982); Bevan y colaboradores, *Nature* 304: 184 - 187 (1983)), el gen bar, que confiere resistencia al herbicida fosfotricina (White y colaboradores. *Nucl. Acids Res* 18: 1062 (1990), Spencer y colaboradores *Theor. Appl. Genet* 79: 625 - 631 (1990)), el gen hph, que confiere resistencia al antibiótico higromicina (Blochinger Diggelmann, *Mol Cell Biol* 4: 2929 - 2931), el gen dhfr, que confiere resistencia al metatretaxato (Bourouis y colaboradores, *EMBO J.* 2(7): 1099 - 1104 (1983)), el gen EPSPS, que confiere resistencia a glifosato (patentes de los Estados Unidos Nos. 4.940.935 y 5.188.642).

En forma alternativa, el material de la planta transgénica puede ser identificado a través de un sistema de selección positiva, tal como el sistema que utiliza el gen para la manosa-6-fosfato isomerasa, que provee la habilidad para metabolizar manosa (patentes de los Estados Unidos Nos. 5.767.778 y 5.994.629).

En un aspecto, la elaboración de plantas o semillas transgénicas comprende incorporar las secuencias de la divulgación y, en forma opcional, genes marcadores en un constructo de expresión objetivo (por ejemplo, un plásmido), junto con el posicionamiento del promotor y las secuencias terminadoras. Esto puede implicar transferir el gen modificado en la planta a través de un método adecuado. Una o más de las secuencias de la divulgación pueden ser combinadas con las secuencias que confieren resistencia a insectos, enfermedades, sequía, mayor rendimiento, mayor calidad nutricional del grano, mayor rendimiento de etanol y similares.

Por ejemplo, se puede introducir un constructo directamente en el ADN genómico de la célula vegetal usando técnicas tales como electroporación y microinyección de protoplastos de las células vegetales, o los constructos se pueden introducir directamente en el tejido vegetal usando métodos balísticos, tales como bombardeo de partículas de ADN. Por ejemplo, véase, por ejemplo, Christou (1997) *Plant Mol. Biol.* 35: 197 - 203; Pawlowski (1996) *Mol. Biotechnol.* 6: 17 - 30; Klein (1987) *Nature* 327: 70 - 73; Takumi (1997) *Genes Genet. Syst.* 72: 63 - 69, que discuten el uso del bombardeo de partículas para introducir transgenes en el trigo; y Adam (1997) ver más arriba, para uso del bombardeo de partículas para introducir las YAC en células vegetales. Por ejemplo, Rinehart (1997) ver más arriba, utilizó bombardeo de partículas para generar plantas de algodón transgénicas. Los aparatos para acelerar partículas se describen en la patente de los Estados Unidos No. 5.015.580; y, el instrumento de aceleración de partículas POS-2000a comercialmente disponible de BioRad (Biolistics); véase también, John, patente de los Estados Unidos No. 5.608.148; y Ellis, patente de los Estados Unidos No. 5. 681.730, que describen la transformación mediada por partículas de gimnospermas.

En un aspecto, los protoplastos pueden ser inmovilizados e inyectados con ácidos nucleicos, por ejemplo, un constructo de expresión. Si bien la regeneración de la planta a partir de los protoplastos no es fácil con los cereales, la regeneración de la planta es posible en las legumbres usando embriogénesis somática de los callos derivados del protoplasto. Los tejidos organizados pueden ser transformados con ADN desnudo usando la técnica de pistola de genes, donde se coloca el ADN sobre microproyectiles de tungsteno, que se disparan a razón de 1/100 con respecto al tamaño de las células, llevando el ADN al interior de las células y organelos. El tejido transformado se induce luego para regeneración, usualmente por embriogénesis somática. Esta técnica ha sido exitosa en varias especies de cereales incluyendo maíz y arroz.

Los ácidos nucleicos, por ejemplo, constructos de expresión, también pueden ser introducidos también en las células vegetales usando virus recombinantes. Las células vegetales pueden ser transformadas usando vectores virales, tales como, por ejemplo, vectores derivados del virus del mosaico del tabaco (Rouwendal (1997) Plant Mol. Biol. 33:989,-999), véase Porta (1996) "Use of viral replicons for the expression of genes in plants" Mol. Biotechnol. 5: 209 - 221.

En forma alternativa, los ácidos nucleicos, por ejemplo, un constructo de expresión, pueden ser combinados con regiones flanqueadoras de T-ADN adecuadas e introducidos en un vector huésped de *Agrobacterium tumefaciens* convencional. Las funciones de virulencia del huésped de *Agrobacterium tumefaciens* dirigirán la inserción del constructo y el marcador adyacente en el ADN de la célula vegetal cuando la célula es infectada por la bacteria. Las técnicas de transformación mediadas por *Agrobacterium tumefaciens*, que incluyen el desarmado y uso de vectores binarios, son descritas en la literatura científica. Véase, por ejemplo, Horsch (1984) Science 233: 496 -98; Fraley (1983) Proc. Natl. Acad. Sci. EUA 80: 4803 (1983); Gene Transfer to Plants, Potrykus, ed. (Springerlag, Berlin 1995). El ADN en una célula de *A. tumefaciens* está contenido en el cromosoma bacteriano así como también en otra estructura conocida como un plásmido Ti (de inducción tumoral). El plásmido Ti contiene una región de ADN denominada T-ADN (~20 kb de longitud) que es transferido a la célula vegetal en el proceso de infección y una serie de genes vir (virulencia) que dirigen el proceso de infección. *A. tumefaciens* puede infectar únicamente una planta a través de heridas: cuando un tallo o raíz de una planta es herido, emite ciertas señales químicas, en respuesta a lo cual, los genes vir de *A. tumefaciens* se activan y dirigen una serie de eventos necesarios para la transferencia del T-ADN del plásmido Ti al cromosoma de la planta. El T-ADN luego ingresa en la célula vegetal a través de la herida. Una especulación es que el T-ADN espera hasta que el ADN de la planta esté siendo replicado o transcrito, luego se inserta en el ADN expuesto de la planta. A fin de usar *A. tumefaciens* como vector transgénico, la sección de inducción tumoral del T-ADN debe ser removida, al tiempo que se retienen las regiones limítrofes del T-ADN y los genes vir. El transgén luego se inserta entre las regiones limítrofes del T-ADN donde es transferido a la célula vegetal y se integra en los cromosomas de la planta.

La divulgación provee la transformación de las plantas monocotiledóneas usando los ácidos nucleicos de la divulgación, incluyendo cereales importantes, véase Hiei (1997) Plant Mol. Biol. 35: 205 - 218. Véase además, por ejemplo, Horsch, Science (1984) 233: 496; Fraley (1983) Proc. Natl. Acad. Sci. EUA 80: 4803; Thykjaer (1997) ver más arriba; Park (1996) Plant Mol. Biol. 32: 1135 - 1148, que discuten la integración del T-ADN en ADN genómico. Véase también D'Halluin, patente de los Estados Unidos No. 5.712.135, que describe un proceso para la integración estable de un ADN que comprende un gen que es funcional en una célula de un cereal u otra planta monocotiledónea.

En un aspecto, el tercer paso puede implicar la selección y regeneración de plantas completas capaces de transmitir el gen objetivo incorporado a la próxima generación. Dichas técnicas de regeneración se basan en la manipulación de ciertas fitohormonas en un medio de cultivo tisular, por lo general basándose en un marcador biocida y/o herbicida que ha sido introducido junto con las secuencias de nucleótidos deseadas. La regeneración de la planta a partir de protoplastos cultivados se describe en Evans y colaboradores, Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, páginas. 124 - 176, MacMillan Publishing , Company, New York, 1983; y Binding, Regeneration of Plantas, Plant Protoplasts, páginas. 21 - 73, CRC Press, Boca Raton, 1985. La regeneración también puede ser obtenida a partir de callos de la planta, explantes, órganos o partes de la misma. Dichas técnicas de regeneración se describen en general en Klee (1987) Ann. Rev. of Plant Phys. 38: 467- 486. Para obtener las plantas completas a partir de tejidos transgénicos tales como embriones inmaduros, se pueden cultivar bajo condiciones ambientales controladas en una serie de medios que contienen nutrientes y hormonas, un proceso conocido como cultivo de tejido. Una vez que las plantas completas son generadas y producen semillas, comienza la evaluación de la progenie.

Después de que el casete de expresión es incorporado en forma estable en plantas transgénicas, se puede introducir en otras plantas por cruzamiento sexual. Se puede utilizar cualquiera entre una cantidad de técnicas de estándar de fitomejoramiento, dependiendo de la especie que va a ser cruzada. Véase, por ejemplo, Welsh J. R., Fundamentals of Plant Genetics and Breeding, John Wiley & Sons, NY (1981); Crop Breeding, Madera D. R. (Ed.) American Society of Agronomy Madison, Wis. (1983); Mayo O., The Theory of Plant Breeding, Segunda Edición, Clarendon Press, Oxford (1987); Singh, D. P., Breeding for Resistance to Diseases and Insect Pests, Springer-Verlag, NY (1986); y Wricke y Weber, Quantitative Genetics and Selection Plant Breeding, Walter de Gruyter and Co. Berlín (1986).

Como la expresión transgénica de los ácidos nucleicos de la divulgación genera cambios fenotípicos, las plantas que comprenden los ácidos nucleicos recombinantes de la divulgación pueden ser cruzadas de manera sexual con una segunda planta para obtener un producto final. Por lo tanto, la semilla de la divulgación puede ser derivada a partir de un cruce entre dos plantas transgénicas de la divulgación, o un cruce entre una planta de la divulgación y otra planta. Los efectos deseados (por ejemplo, expresión de los polipéptidos de la divulgación para producir una planta en la cual se altera el comportamiento de floración) pueden ser mejorados cuando ambas plantas progenitoras expresan los polipéptidos (por ejemplo, glucanasa, manasa o xilanasa) de la divulgación. Los efectos deseados pueden pasar a futuras generaciones de plantas a través de medios de propagación estándares.

Cualquier planta se puede usar para la introducción del nucleótido de interés, incluyendo, pero sin limitarse a, maíz (*Zea mays*), *Brassica* sp. (por ejemplo, *B. napus*, *B. rapa*, *B. juncea*), particularmente las especies *Brassica* útiles como fuente de aceite de semilla, tal como canola, alfalfa (*Medicago sativa*), arroz (*Oryza sativa*), centeno (*Secale cereale*), sorgo (*Sorghum bicolor*, *Sorghum vulgare*), mijo (por ejemplo, mijo perlado (*Pennisetum glaucum*), mijo proso (*Panicum miliaceum*), mijo cola de zorro (*Setaria italica*), mijo dedo (*Eleusine coracana*), girasol (*Helianthus annuus*), cártamo (*Carthamus tinctorius*), trigo (*Triticum aestivum*), soja (*Glycine max*), tabaco (*Nicotiana tabacum*), patata (*Solanum tuberosum*), cacahuets (*Arachis hypogaea*), algodón (*Gossypium barbadense*, *Gossypium hirsutum*), batata (*Ipomoea batatas*), mandioca (*Manihot esculenta*), café (*Coffea* spp.), coco (*Cocos nucifera*), piña (*Ananas comosus*), árboles cítricos (*Citrus* spp.), cacao (*Theobroma cacao*), té (*Camellia sinensis*), banana (*Musa* spp.), aguacate (*Persea americana*), higo (*Ficus casica*), guayaba (*Psidium guajava*), mango (*Mangifera indica*), oliva (*Olea europaea*), papaya (*Carica papaya*), nuez de la India (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), almendra (*Prunus amygdalus*), remolacha azucarera (*Beta vulgaris*), caña de azúcar (*Saccharum* spp.), avena, cebada, vegetales, plantas ornamentales y coníferas.

Los vegetales pueden incluir tomates (*Lycopersicon esculentum*), lechuga (por ejemplo, *Lactuca sativa*), judías verdes (*Phaseolus vulgaris*), judías lima (*Phaseolus limensis*), guisantes (*Lathyrus* spp.) y miembros del género *Cucumis* tales como pepino (*C. sativus*), cantalupo (*C. cantalupensis*) y melón almizcleño (*C. melo*). Las plantas ornamentales pueden incluir azalea (*Rhododendron* spp.), hortensia (*Macrophylla hydrangea*), hibisco (*Hibiscus rosasanensis*), rosas (*Rosa* spp.), tulipanes (*Tulipa* spp.), narcisos (*Narcissus* spp.), petunias (*Petunia hybrida*), claveles (*Dianthus caryophyllus*), flor de Pascua (*Euphorbia pulcherrima*), canna (*Cannaceae* spp.) y crisantemo. Las coníferas que pueden ser utilizadas, incluyen, por ejemplo, pinos tales como pino taeda (*Pinus taeda*), pino de incienso (*Pinus elliotii*), pino ponderosa (*Pinus ponderosa*), pino logdepole (*Pinus contorta*) y pino de Monterrey (*Pinus radiata*), abeto Douglas (*Pseudotsuga menziesii*); abeto occidental (*Tsuga canadensis*); picea Sitka (*Picea glauca*); secoya (*Sequoia sempervirens*); abetos reales tales como abeto plateado (*Abies amabilis*) y abeto balsámico (*Abies balsantea*); y cedros tales como cedro rojo occidental (*Thuja plicata*) y cedro amarillo de Alaska (*Chamaecyparis nootkatensis*); y plantas leguminosas, incluyendo, pero sin limitarse a, frijoles y guisantes; donde en aspectos alternativos los frijoles pueden incluir guar, algarroba, fenogreco, soja, habichuelas, arvejas, frijol mungo, frijol lima, habas, lentejas, garbanzos, etc. y las legumbres pueden incluir, pero sin limitarse a, *Arachis*, por ejemplo, cacahuets, Vicia, por ejemplo, veza corona, veza peluda, frijol adzuki, frijol mungo y garbanzos, lupines, por ejemplo, lupino, trébol rojo, *Phaseolus*, por ejemplo, frijol común y frijol lima, Pisum, por ejemplo, frijol de campo, meliloto, por ejemplo, trébol, *Medicago*, por ejemplo, alfalfa, loto, por ejemplo, trébol, lentejas, por ejemplo, lenteja y falso índigo; también se incluye pastos para forraje y pasto césped, tales como alfalfa, pasto varilla (*Panicum virgatum*), *Miscanthus*, pasto ovido, festuca alta, ballico perenne, pasto encorvado rastrero y agróstide.

Las plantas de interés particular pueden incluir plantas de cultivo y plantas usadas para producir energía o combustible, por ejemplo, maíz, alfalfa, girasol, *Brassica*, soja, algodón, cártamo, cacahuete, sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y *Miscanthus*, cultivos de leguminosas, por ejemplo, guisante, frijol y soja, raíz/tubérculo almidonoso, por ejemplo, patata, batata, mandioca, taro, canna, remolacha azucarera, caña de azúcar y similares.

En formas de realización alternativas, los ácidos nucleicos de la divulgación son expresados en plantas que contienen células de fibras, incluyendo, por ejemplo, algodón, árbol de algodón de seda (*Kapok*, *Ceiba pentandra*), sauce del desierto, Chaparral, ontina, balsa, ramio, cáñamo de la india, cáñamo común, rosella, yute, sisal, abacá y lino. En formas de realización alternativas, las plantas transgénicas de la divulgación pueden ser miembros del género *Gossypium*, incluyendo miembros de cualquier especie *Gossypium*, tal como *G. arboreum*; *G. herbaceum*, *G. barbadense* y *G. hirsutum*.

La divulgación también provee plantas transgénicas que se utilizan para producir grandes cantidades de los polipéptidos (por ejemplo, glucanasa, manasa o xilanasa) de la divulgación. Por ejemplo, véase Palmgren (1997) Trends Genet. 13: 348; Chong (1997) Transgenic Res. 6: 289 - 296 (que producen proteína beta-caseína de la leche humana en plantas de patata transgénicas usando un promotor, manopina sintasa bidireccional inducible por auxina (más 1',2') con métodos de transformación de disco de hoja mediados por *Agrobacterium tumefaciens*).

Usando procedimientos conocidos, una persona versada en el arte puede cribar las plantas de la divulgación detectando el incremento o reducción de ARNm transgénico o proteína en las plantas transgénicas. Los medios para detectar y cuantificar los ARNm o proteínas son bien conocidos en el arte.

Polipéptidos y péptidos

En un aspecto, la invención provee polipéptidos aislados, sintéticos o recombinantes, que tienen una identidad de secuencia (por ejemplo, al menos 99%, o más, o identidad de secuencia completa (100%)) con la secuencia de ejemplo de la invención, por ejemplo, SEQ ID NO: 52.

En una forma de realización, los polipéptidos de la invención pueden catalizar la hidrólisis de polisacáridos que comprenden monómeros de glucosa, tales como almidón (un polímero de monómero de glucosa unido por enlaces

1,4-alfa o 1,6-alfa). En un aspecto, el polipéptido tiene una actividad de alfa amilasa, y el término "amilasa" como se usa aquí, también incluye actividad de enzima que cataliza la hidrólisis de un polisacárido, oligosacárido o almidón. Las amilasas de la invención incluyen polipéptidos que tienen una actividad de α -amilasa. Por ejemplo, una actividad de amilasa de la invención incluye α -amilasa, que incluye la habilidad para hidrolizar enlaces alfa-1,4-glucosídicos internos en almidón para producir maltodextrinas de peso molecular más pequeño. En un aspecto, la actividad de la α -amilasa incluye hidrolizar enlaces de alfa-1-4-glucosídicos internos en almidón al azar

En una forma de realización, los polipéptidos de la divulgación se pueden usar para generar un anticuerpo que se enlaza de manera específica a (es específico de) un polipéptido de la invención, por ejemplo, un ejemplo de enzima de la divulgación (por ejemplo, la SEQ ID NO: 2, SEQ ID NO:4, etc.

10 En un aspecto, la actividad de glucoamilasa de un polipéptido de la divulgación comprende la catálisis de la hidrólisis de enlaces glucosídicos. La actividad de glucoamilasa puede comprender catalizar la liberación hidrolítica por etapas de la D-glucosa a partir de los extremos no reductores de almidón u otras dextrinas relacionadas. La actividad de glucoamilasa puede comprender una actividad de 1,4- α -D-glucano glucohidrolasa. La actividad de glucoamilasa puede comprender la catálisis de la hidrólisis de maltodextrinas que resultan en la generación de glucosa libre. La actividad de glucoamilasa puede comprender una actividad de exoamilasa. La actividad de glucoamilasa puede comprender actividad una de α -amilasa o de β -amilasa. Los enlaces glucosídicos hidrolizados pueden comprender enlaces α -1,4-glucosídicos o enlaces α -1,6-glucosídicos. La actividad de glucoamilasa puede comprender hidrolizar enlaces glucosídicos en un polisacárido, oligosacárido o almidón. La actividad de glucoamilasa puede comprender además, hidrolizar enlaces glucosídicos en el almidón para producir maltodextrinas. La actividad de glucoamilasa puede comprender escindir una unidad de maltosa o de D-glucosa del extremo no reductor del polisacárido, oligosacárido o almidón.

En un aspecto, la invención provee alfa-amilasas (α -amilasas) que son enzimas que actúan en forma endo que pueden hidrolizar almidón, un polímero de monómero de glucosa unido por enlaces 1,4-alfa o 1,6-alfa, a maltodextrinas cortas. En un aspecto, la divulgación provee glucoamilasas que son hidrolasas que actúan en forma exo que pueden liberar beta-D-glucosa de los extremos no reductores del almidón y sacáridos relacionados. Las amilasas de esta invención se pueden usar a nivel comercial para licuar y sacarificar el almidón durante la producción de etanol usando procesos tales como el proceso de molienda en seco. En un aspecto de un proceso de la divulgación, el proceso de molienda en seco muele maíz entero (la suspensión) que se somete a una temperatura elevada (para promover la gelatinización de almidón) y es hidrolizada por una o más amilasas termoestables, que incluyen al menos una enzima de la divulgación, lo que resulta en un polisacárido, por ejemplo, almidón, licuefacción. En un aspecto, el polisacárido hidrolizado, por ejemplo, almidón, es digerido adicionalmente por una glucoamilasa de la divulgación, que puede ser agregada una vez que la temperatura de la masa desciende; y en un aspecto, la glucosa liberada del polisacárido, por ejemplo, almidón, es fermentada hasta etanol por levadura agregada a la terminación del polisacárido, por ejemplo, almidón, proceso de hidrólisis, o durante la sacarificación del polisacárido, por ejemplo, almidón.

En un aspecto, las glucosidasas (por ejemplo, glucoamilasas, alfa glucosidasas) de la divulgación hidrolizan los enlaces internos de polisacáridos, por ejemplo, enlaces α -1,4- y α -1,6-glucosídicos en un polisacárido, oligosacárido o almidón, para producir maltodextrinas de menor peso molecular. En un aspecto, esta hidrólisis es muy aleatoria. Por lo tanto, la divulgación provee métodos para producir maltodextrinas de menor peso molecular. Las glucosidasas de la divulgación se pueden usar en establecimientos industriales y en laboratorios para hidrolizar un polisacárido, oligosacárido o almidón, o cualquier compuesto que comprenda maltodextrina para una variedad de propósitos. Estas glucosidasas se pueden usar en forma solas para proveer hidrólisis específica o pueden ser combinadas con otras glucosidasas para proveer un "cóctel" con un amplio espectro de actividad. Los ejemplos de usos incluyen la remoción o hidrólisis parcial o completa de un polisacárido, oligosacárido o almidón, o cualquier compuesto que comprenda maltodextrina a partir de muestras biológicas, alimenticias, de pienso animal, o farmacéuticas o industriales.

Por ejemplo, las glucosidasas (por ejemplo, las glucoamilasas) de la divulgación pueden ser formuladas en detergentes para lavandería para ayudar a la remoción de tinturas que comprenden polisacáridos, por ejemplo, que contienen almidón. Las glucosidasas de la divulgación se pueden usar como agentes de limpieza en matrices de detergentes (véase aplicaciones industriales más adelante). Las glucosidasas de la divulgación se pueden usar en las etapas iniciales (licuefacción) del polisacárido, por ejemplo, almidón, procesamiento en la molienda de maíz en húmedo, en la producción del alcohol, en la industria textil para descolado del almidón, en aplicaciones de panadería y pastelería, en la industria de las bebidas, en procesos de perforación en campos petroleros; en el entintado de papel reciclado; y en pienso animal.

55 Las glucosidasas de la divulgación (por ejemplo, glucoamilasas) pueden tener actividad de glucosidasa bajo diferentes condiciones, por ejemplo, de pH y/o temperatura extremos, agentes de oxidación y similares. La divulgación provee métodos que conducen a preparaciones alternativas de glucosidasas con diferentes eficiencias y estabildades catalíticas, por ejemplo, en cuanto a temperatura, agentes oxidantes y cambio en las condiciones de lavado. En un aspecto, las variantes de glucosidasa pueden ser producidas usando técnicas de mutagénesis dirigida

al sitio y/o mutagénesis aleatoria. En un aspecto, la evolución dirigida se puede usar para producir una gran variedad de variantes de glucosidasa con especificidad y estabilidad alternativas.

Las proteínas de la divulgación también son útiles como reactivos de investigación para identificar moduladores de amilasa y/o glucoamilasa, por ejemplo, activadores o inhibidores de la actividad de amilasa y/o glucoamilasa. En resumen, se añaden las muestras de prueba (compuestos, caldos, extractos y similares) a los ensayos de amilasa y/o de glucoamilasa para determinar su habilidad para inhibir la escisión del sustrato. Los inhibidores identificados de este modo se pueden usar en la industria y en investigación para reducir o prevenir la proteólisis indeseada. Como sucede con la amilasa y/o la glucoamilasa, los inhibidores pueden ser combinados para incrementar el espectro de actividad.

- 5 Una actividad de amilasa de la invención también incluye hidrolizar un polisacárido, oligosacárido o almidón, a altas temperaturas, bajas temperaturas, pH alcalino y pH ácido. Por ejemplo, en un aspecto, la invención provee polipéptidos y los ácidos nucleicos que los codifican, que tienen una actividad de amilasa que es termoestable. El polipéptido puede retener una actividad de amilasa bajo condiciones que comprenden un rango de temperatura de entre aproximadamente 37 °C hasta aproximadamente 95 °C; entre aproximadamente 55 °C hasta aproximadamente 85 °C, entre aproximadamente 70 °C hasta aproximadamente 95 °C, o, entre aproximadamente 90 °C hasta aproximadamente 95 °C. En otro aspecto, un polipéptido de la invención puede tener una actividad de amilasa que es termotolerante. El polipéptido puede retener una actividad de amilasa después de la exposición a una temperatura en el rango de más de 37 °C hasta aproximadamente 95 °C o en el rango de más de 55 °C hasta aproximadamente 85 °C. En un aspecto, el polipéptido conserva una actividad de amilasa después de la exposición a una temperatura en el rango de mayor que 90 °C hasta aproximadamente 95 °C a pH 4,5.

- 15 La invención provee "aminoácidos" o "secuencias de aminoácidos de la invención", que incluyen una secuencia de oligopéptido, péptido, polipéptido, o proteína, o para un fragmento, porción o subunidad de cualquiera de estas, y para moléculas sintéticas, recombinantes o de origen natural. Los términos "polipéptido" y "proteína" incluyen aminoácidos unidos entre sí por enlaces peptídicos o enlaces peptídicos modificados, es decir, isómeros peptídicos o enlaces no peptídicos (enlaces sintéticos, polipéptidos sintéticos) y pueden contener aminoácidos modificados diferente a los 20 aminoácidos codificados por genes. El término "polipéptido" también incluye fragmentos de péptidos y polipéptidos, motivos y similares. El término también incluye polipéptidos glicosilados. Los péptidos y polipéptidos de la invención también incluyen todas las formas "miméticas" y "peptidomiméticas" como se describe con mayor detalle, más adelante.

- 25 El término "aislado" incluye un material removido de su ambiente original, por ejemplo, el ambiente natural si es de origen natural. Por ejemplo, un polinucleótido o polipéptido de origen natural presente en un animal vivo no es aislado, pero el mismo polinucleótido o polipéptido, separado de parte o todo el material coexistente en el sistema natural, es aislado. Dichos polinucleótidos podrían ser parte de un vector y/o dichos polinucleótidos o polipéptidos podrían ser parte de una composición y aún así estar aislados de manera tal que dicho vector o composición no sea parte de su ambiente natural. Como se usa aquí, un material o composición aislados pueden ser también una composición "purificada", es decir, que no requiere pureza absoluta; en cambio, se usa como definición relativa. Los ácidos nucleicos individuales obtenidos de una biblioteca pueden ser purificados en forma convencional hasta homogeneidad electroforética. En aspectos alternativos, la invención provee ácidos nucleicos que han sido purificados de ADN genómico o de otras secuencias en una biblioteca u otro ambiente en al menos uno, dos, tres, cuatro, cinco o más órdenes de magnitud.

- 30 La divulgación también provee "variantes de amilasa" y "variantes de glucoamilasa" que pueden comprender una secuencia de aminoácidos que se deriva de la secuencia de aminoácidos de una "amilasa precursora", por ejemplo, en un aspecto, una secuencia de ejemplo de la divulgación (por ejemplo, la SEQ ID NO:4, etc., o cualquier polipéptido de esta divulgación. La glucoamilasa o amilasa precursora también puede incluir glucoamilasas o amilasas de origen natural y amilasas recombinantes. La secuencia de aminoácidos de la variante de glucoamilasa o amilasa se puede "derivar" de la secuencia de aminoácidos de la glucoamilasa o amilasa precursora por sustitución, supresión o inserción de uno o más aminoácidos de la secuencia de aminoácidos precursora. Dicha modificación puede ser de la "secuencia de ADN precursora" que codifica la secuencia de aminoácidos de la amilasa precursora en vez de la manipulación de la enzima amilasa precursora por sí misma. Los métodos adecuados para dicha manipulación de la secuencia de ADN precursora incluyen los métodos divulgados aquí, así como también los métodos conocidos para aquellos capacitados en el arte.

- 35 Las actividades de las secuencias que sirven de ejemplo de la divulgación se enlistan en la tabla ("Tabla 1") inmediatamente a continuación. Para ayudar a la lectura de la tabla, por ejemplo, en la primera fila las SEQ ID NO: 1, 2, representan el ejemplo de polipéptido de la divulgación que tiene una secuencia como se expone en la SEQ ID NO: 2, codificada por, por ejemplo, la SEQ ID NO: 1; y esta secuencia de ejemplo fue aislada inicialmente a partir de *Cochliobolus heterostrophus*, ATCC 48331; no se predice ninguna secuencia señal (pero bajo ciertas condiciones celulares *in vivo*, la secuencia puede tener una secuencia señal); el polipéptido tiene actividad de "glucosidasa", que puede ser designada en forma más específica como actividad de la enzima "amilasa"; y el correspondiente número "EC" para las enzimas amilasa (un número EC es el número asignado a un tipo de enzima de acuerdo con un

esquema de nomenclatura de enzimas estandarizada desarrollada por la Enzyme Commission of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology, o IUBMB); "CMB20" que designa el "dominio de enlazamiento de carbohidratos"; y la última columna indica una "fuente genética", o la fuente de la secuencia de ejemplo tal como se determina por análisis de homología del ARN 18/16S de las célula a partir de la cual se aisló inicialmente. En la segunda fila, la SEQ ID NO: 11, 12, representa el ejemplo de polipéptido de la divulgación, que tiene una secuencia como se expone en la SEQ ID NO: 12, codificada por, por ejemplo, la SEQ ID NO: 11; y esta secuencia que sirve de ejemplo se aisló inicialmente de una fuente desconocida; el "sitio SS", designa los residuos amino terminales que son la secuencia señal y para la SEQ ID NO: 12 son los 21 residuos de aminoácidos del terminal amino que forman la secuencia señal (o, MFNQVLYGLAATALWQQQVVA, es decir, los residuos 1 a 21 de la SEQ ID NO: 12); y el polipéptido tiene actividad de "glicosidasa", que puede ser designada en forma más específica como una actividad de enzima "glicoamilasa"; teniendo esta enzima un número EC de 3.2.1.3; con "dominio de enlazamiento del carbohidrato"; y una coincidencia genética del ARN 18/16S de la célula a partir de la cual se aisló inicialmente es *Fusarium equiseti*.

Tabla 1

SEQ ID NO:	Fuente	Sitio SS	Secuencia señal	Actividad general	Un ejemplo de actividad específica	Número EC	CBM20	Fuente (como se determinó mediante ARN 18/16S)
1, 2	Cochliobolus heterostrophus ATCC 48331			Glicosidasa	amilasa	3.2.1.1	CBM20	Cochliobolus heterostrophus
11, 12	Desconocida	AA 1-21	MFNQVLYGLAATALWQGQVVA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	Fusarium equiseti 100%
13, 14	Fusarium verticillioides GZ3639	AA 1-21	MFTQILYGLTALSALQGQVTA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	Fusarium verticillioides GZ3639
15, 16	Cochliobolus heterostrophus ATCC 48332	AA 1-20	MLSKILLPVVALAASANAHG	Glicosidasa	glucoamilasa		CBM20	Cochliobolus heterostrophus
17, 18	Fusarium verticillioides GZ3639			Glicosidasa	glucoamilasa	3.2.1.3	CBM20	Fusarium verticillioides GZ3639
19,20	Desconocida	AA 1-22	MLTLNVL TALLAPIVLSALPA	Glicosidasa	glucoamilasa	3.2.1.3	no	Penicillium chrysogenum 100%
21, 22	Desconocida	AA 1-18	MVLARLAWLAGLVSTAVA	Glicosidasa	amilasa	3.2.1.1	no	Penicillium expansum 99%

SEQ ID NO:	Fuente	Sitio SS	Secuencia señal	Actividad general	Un ejemplo de actividad específica	Número EC	CBM20	Fuente (como se determinó mediante ARN 18/16S)
23, 24	Desconocida	AA 1-20	MKLSHTLTALLPLICTVSA	Glicosidasa	amilasa	3.2.1.1	no	<i>Penicillium</i>
25, 26	Desconocida	AA 1-21	MTISRLSSVLFALALQGSALA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>chrysogenum</i> 100%
27, 28	Desconocida	AA 1-20	MYLSSAFLLGSLALQSVLG	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>Fusarium</i> <i>verruculosum</i> 100%
29, 30	Desconocida	AA 1-21	MLFSSLLRALSASLLAGAVQQG	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>merismoides</i> 99%
3, 4	Cochliobolus heterostrophus ATCC 48331	AA 1-20	MLLLNIFTTLFFYITCIVSA	Glicosidasa	amilasa	3.2.1.1	CBM20	<i>Phoma herbarum</i> 99%
31, 32	Desconocida	AA 1-18	MVLARLAWLAGLVSTAIA	Glicosidasa	amilasa	3.2.1.1	no	<i>Cochliobolus heterostrophus</i>
33, 34	Desconocida	AA 1-18	MVGFNLTLLAPPAALS	Glicosidasa	glucoamilasa	3.2.1.3	no	<i>chrysogenum</i> 100%
35, 36	Desconocida	AA 1-20	MAPRFWTTLCALTLGSAALA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>Penicillium herquei</i> 99%
								<i>Fusarium oxysporum</i> 100%

SEQ ID NO:	Fuente	Sitio SS	Secuencia señal	Actividad general	Un ejemplo de actividad específica	Número EC	CBM20	Fuente (como se determinó mediante ARN 18/16S)
37, 38	Desconocida	AA 1-19	MAPRFWIALWALTFGQAIA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>Cordyceps</i>
39, 40	Desconocida	AA 1-20	MAPRFWTALWALTIGHAWA	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>ophioglossoides</i> 99%
41, 42	Desconocida			Glicosidasa	glucoamilasa	3.2.1.3	no	<i>Penicillium chrysogenum</i> 100%
43, 44	Cochliobolus heterostrophus ATCC 48332	AA 1-23	MHTHSFVQASTVLSLLTAGQ	Glicosidasa	α -glucosidasa			<i>Cucurbitaria berberidis</i> 98%
45, 46	Desconocida	AA 1-19	MKLLQLAALVASLSPFTNA	Glicosidasa	amilasa	3.2.1.1	no	<i>Cochliobolus heterostrophus</i>
47, 48	Desconocida	AA 1-20	MTRILTLALHGLALVQSVVG	Glicosidasa	glucoamilasa	3.2.1.3	CBM20	<i>Fusarium equiseti</i> 100%
49, 50	Aspergillus terreus	AA 1-18	MSFFLSCLYLSLCGSALA	Glicosidasa	amilasa	3.2.1.1	no	<i>Aspergillus versicolor</i> 99%

SEQ ID NO:	Fuente	Sitio SS	Secuencia señal	Actividad general	Un ejemplo de actividad específica	Número EC	CBM20	Fuente (como se determinó mediante ARN 18/16S)
5, 6	<i>Cochliobolus heterostrophus</i> ATCC 48331			Glicosidasa	α -glucosidasa			<i>Cochliobolus heterostrophus</i>
51, 52	<i>Aspergillus terreus</i>	AA 1-20	MKWTFSLLLLLSVFGQATHA	Glicosidasa	amilasa	3.2.1.1	CBM20	<i>Aspergillus terreus</i>
53, 54	<i>Aspergillus terreus</i>	AA 1-20	MKLSRALTVLLHLTSTALA	Glicosidasa	amilasa	3.2.1.1	no	<i>Aspergillus terreus</i>
55, 56	Desconocida	AA 1-27	MLKQFTKRLITLTSLLALVLVAPLAS A	Glicosidasa		3.2.1.1		
57, 58	Desconocida			Glicosidasa		3.2.1.10		
59, 60	Desconocida			Glicosidasa		3.2.1.10		
61, 62	Desconocida	AA 1-20	MVAGFGLYGAALLTPMAAQA	Glicosidasa		3.2.1.1		
63, 64	Desconocida			Glicosidasa		3.2.1.20		
65, 66	Desconocida	AA 1-25	MKLYLALVLLAVASIGLLSTPVG	Glicosidasa		3.2.1.1		
67, 68	Desconocida	AA 1-23	MKNTISALVAGMVLGFASNAMA	Glicosidasa		3.2.1.1		
69, 70	Desconocida	AA 1-44	MINRPGTGASGRPQRSATSWQS RNGGWLLASLLAVCFATAPVRA	Glicosidasa		3.2.1.1		

SEQ ID NO:	Fuente	Sitio SS	Secuencia señal	Actividad general	Un ejemplo de actividad específica	Número EC	CBM20	Fuente (como se determinó mediante ARN 18/16S)
7, 8	Cochliobolus heterostrophus ATCC			Glicosidasa	glucoamilasa	3.2.1.3	no	Cochliobolus heterostrophus
71, 72	Desconocida			Glicosidasa		3.2.1.20		
73, 74				Glicosidasa		3.2.1.3		
75, 76	Desconocida	AA 1-24		Glicosidasa	amilasa	3.2.1.1		
9, 10	Desconocida	AA 1-22	MLTLNVL TALLAPGVLSSALPA	Glicosidasa	glucoamilasa	3.2.1.3	no	Penicillium expansum 99%

Los polipéptidos de la invención también pueden ser más cortos que la longitud completa de los polipéptidos que sirven de ejemplo. En aspectos alternativos, la invención provee polipéptidos (péptidos, fragmentos) que oscilan en tamaño entre aproximadamente 5 y la longitud completa de un polipéptido, por ejemplo, una enzima, tal como una amilasa; los ejemplos de tamaños son aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, o más residuos, por ejemplo, residuos contiguos de un ejemplo de amilasa de la invención. Los péptidos de la invención pueden ser útiles, por ejemplo, como sondas de marcación, antígenos, tolerágenos, motivos, sitios activos de amilasa.

Los polipéptidos y péptidos de la invención pueden ser aislados a partir de fuentes naturales, ser polipéptidos sintéticos o generados en forma recombinante. Los péptidos y proteínas pueden ser expresados en forma recombinante *in vitro* o *in vivo*. Los péptidos y polipéptidos de la invención pueden ser elaborados y aislados usando cualquier método conocido en el arte. Los polipéptidos y péptidos de la invención también pueden ser sintetizados, en forma completa o en parte, usando métodos químicos conocidos en el arte. Véase por ejemplo, Caruthers (1980) Nucleic Acid Res. Symp. Ser. 215 - 223; Horn (1980) Nucleic Acid Res. Symp. Ser. 225 - 232; Banga, A. K., Therapeutic Peptides and Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co., Lancaster, PA. Por ejemplo, la síntesis peptídica puede ser realizada usando varias técnicas en fase sólida (véase, por ejemplo, Roberge (1995) Science 269: 202; Merrifield (1997) Methods Enzymol. 289: 3 - 13) y se puede lograr una síntesis automatizada, por ejemplo, usando el Peptide Synthesizer ABI 431A (Perkin Elmer) de acuerdo con las instrucciones provistas por el fabricante.

Los péptidos y polipéptidos de la invención también pueden ser glicosilados. La glicosilación puede ser agregada de manera posterior a la traducción ya sea a nivel químico o por mecanismos biosintéticos celulares, en donde este último incorpora el uso de motivos de glicosilación conocidos, que pueden ser nativos con respecto a la secuencia o pueden ser agregados como un péptido o agregados en la secuencia codificadora del ácido nucleico. La glicosilación puede estar enlazada a O o enlazada a N. La glicosilación puede ser agregada a cualquier polipéptido de la invención para generar una enzima que es más termotolerante o termoestable que la enzima "progenitora" (a la cual se le agregó la glicosilación). La glicosilación puede ser agregada ya sea por mecanismos químicos o por mecanismos biosintéticos celulares.

La invención provee amilasas que tienen un amplio rango de actividad específica sobre un amplio rango de temperaturas, por ejemplo, en aproximadamente 37 °C en el rango de aproximadamente 10 a 10.000, o, 100 a aproximadamente 1000 unidades por miligramo de proteína. Las amilasas de la invención también pueden tener actividad a temperaturas tan altas como 120 °C. En aspectos alternativos, la amilasa utilizada en estos métodos está activa a estas temperaturas, por ejemplo, activa a temperaturas en un rango de entre aproximadamente 80 °C a aproximadamente 115 °C, entre aproximadamente 100 °C a aproximadamente 110 °C, y de aproximadamente 105 °C a aproximadamente 108 °C. Sin embargo, las amilasas de la invención también pueden tener actividad a bajas temperaturas, por ejemplo, tan bajas como 4 °C a 5 °C.

La T_m de una enzima de la invención puede ser cambiada (por ejemplo, puede ser cambiada entre aproximadamente 10 °C y 90 °C) mediante activación térmica. Por ejemplo, la T_m de la SEQ ID NO: 336/337 puede ser cambiada aproximadamente a 17°C hasta 87°C mediante activación térmica: por ejemplo, preincubación a 80°C durante 5 minutos.

Los péptidos y polipéptidos de la divulgación, como se definió anteriormente, incluyen todas las formas "miméticas" y "peptidomiméticas". Los términos "mimético" y "peptidomimético" se refieren a un compuesto químico sintético que tiene sustancialmente las mismas características estructurales y/o funcionales de los polipéptidos de la divulgación. El mimético puede estar ya sea completamente compuesto de análogos no naturales, sintéticos, de aminoácidos, o, es una molécula quimérica de aminoácidos peptídicos parcialmente naturales y análogos parcialmente no naturales de aminoácidos. El mimético también puede incorporar cualquier cantidad de sustituciones conservadoras de aminoácidos naturales siempre que dichas sustituciones tampoco alteren sustancialmente la estructura y/o actividad del mimético. Como sucede con los polipéptidos de la divulgación que son variantes conservadoras, la experimentación de rutina determinará si la estructura y/o función de un mimético no sea sustancialmente alterada. Por lo tanto, en un aspecto, una composición mimética está dentro del alcance de la divulgación si tiene una actividad de amilasa.

Las composiciones miméticas de polipéptidos de la divulgación pueden contener cualquier combinación de componentes estructurales no naturales (por ejemplo, pueden ser completamente o parcialmente sintéticos o "miméticos"). En un aspecto alternativo, las composiciones miméticas de la divulgación incluyen uno o todos los siguientes tres grupos estructurales: a) grupos de enlazamiento a un residuo diferentes de los enlaces de amida naturales ("enlace peptídico"); b) residuos no naturales en lugar de residuos de aminoácidos de origen natural; o c) residuos que inducen mimetismo estructural secundario, es decir, para inducir o estabilizar una estructura secundaria, por ejemplo, una vuelta beta, una vuelta gamma, una lámina beta, una conformación de hélice alfa, y similares. Por ejemplo, un polipéptido de la divulgación puede ser caracterizado como un mimético cuando todos o algunos de sus residuos se unen por medios químicos que no sean los enlaces peptídicos naturales. Los residuos

péptido- miméticos individuales pueden estar unidos por enlaces peptídicos, otros enlaces químicos o medios de acoplamiento, tales como por ejemplo, glutaraldehído, ésteres de N-hidroxisuccinimida, maleimidias bifuncionales, N,N'-diciclohexilcarbodiimida (DCC) o N,N'-diisopropilcarbodiimida (DIC). Los grupos de enlazamiento que pueden ser una alternativa para el enlazamiento de amida tradicional ("enlace peptídico") incluyen, por ejemplo, cetometileno (por ejemplo, -C(=O)-CH₂- por -C(=O)-NH-), aminometileno (CH₂-NH), etileno, olefina (CH=CH), éter (CH₂-O), tioéter (CH₂-S), tetrazol (CN₄-), tiazol, retroamida, tioamida, o éster (véase, por ejemplo, Spatola (1983) en Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, páginas 267 - 357, "Peptide Backbone Modifications", Marcell Dekker, NY).

Un polipéptido de la divulgación también puede ser caracterizado como un mimético por contener todos o algunos de los residuos no naturales en lugar de los residuos de aminoácidos de origen natural. Los residuos no naturales se describen en la literatura científica y de patentes; unos pocos ejemplos no naturales de composiciones útiles como miméticos de residuos de aminoácidos naturales y guanidinas se describen a continuación. Los agentes miméticos de los aminoácidos aromáticos pueden ser generados mediante reemplazo por, por ejemplo, D- o L- naftilalanina; D- o L- fenilglicina; D- o L-2 tieneilalanina; D- o L-1, -2, 3-, o 4- pireneilalanina; D- o L-3 tieneilalanina; D- o L-(2-piridinil)-alanina; D- o L-(3-piridinil)-alanina; D- o L-(2-pirazinil)-alanina; D- o L-(4-isopropil)-fenilglicina; D-(trifluorometil)-fenilglicina; D-(trifluorometil)-fenilalanina; D-p-fluoro-fenilalanina; D- o L-p-bifenilfenilalanina; D- o L-p-metoxi-bifenilfenilalanina; D- o L-2- indol(alquil)alaninas; y, D- o L-alquilalaninas, en donde el alquilo puede ser metilo, etilo, propilo, hexilo, butilo, pentilo, isopropilo, isobutilo, sec-isotilo, iso-pentilo o aminoácidos no ácidos sustituidos o no sustituidos. Los anillos aromáticos de un aminoácido no natural incluyen, por ejemplo, anillos aromáticos de tiazolilo, tiofenilo, pirazolilo, bencimidazolilo, naftilo, furanilo, pirrolilo y piridilo.

Los agentes miméticos de aminoácidos ácidos pueden ser generados mediante sustitución, por ejemplo, aminoácidos no carboxilados, mientras que mantienen una carga negativa; (fosfono)alanina; treonina sulfatada. Los grupos laterales carboxilo (por ejemplo, aspartilo o glutamilo) también pueden ser modificados en forma selectiva por reacción con carbodiimidias (R'-N-C-N-R') tales como, por ejemplo, 1-ciclohexil-3(2-morfolinil-(4-etil) carbodiimida o 1-etil-3(4-azonia-4,4-dimetolpentil) carbodiimida. El aspartilo o glutamilo también pueden ser convertidos en residuos de asparaginilo y glutaminilo por reacción con iones amonio. Los agentes miméticos de los aminoácidos básicos pueden ser generados por sustitución con, por ejemplo, (además de lisina y arginina) los aminoácidos ornitina, citrulina, o ácido (guanidino)-acético, o ácido (guanidino)alquil-acético, donde alquilo se definió anteriormente. El derivado de nitrilo (por ejemplo, que contiene la fracción CN en lugar de COOH) puede ser sustituido por asparagina o glutamina. Los residuos de asparaginilo y glutaminilo pueden ser desaminados con los correspondientes residuos de aspartilo o glutamilo. Los agentes miméticos del residuo arginina pueden ser generados por reacción del arginilo por ejemplo con, uno o mas reactivos convencionales, incluyendo, por ejemplo, fenilgloxal, 2,3-butanodiona, 1,2-ciclohexanediona, o ninhidrina, preferiblemente bajo condiciones alcalinas. Los agentes miméticos de los residuos de tirosina pueden ser generados por reacción del tirosilo, por ejemplo, con compuestos de diazonio aromático o tetranitrometano. El N-acetilimidazol y tetranitrometano se pueden usar para formar la especie O-acetil tirosilo y derivados 3-nitro, respectivamente. Los agentes miméticos de los residuos de cisteína pueden ser generados por reacción de residuos de cisteinilo, por ejemplo, con alfa-haloacetatos tales como ácido 2-cloroacético o cloroacetamida y las aminas correspondientes, para producir derivados carboximetilo o carboxiamidometilo. Los agentes miméticos del residuo de cisteína también pueden ser generados por reacción de residuos de cisteinilo, por ejemplo, con bromo-trifluoroacetona, ácido de alfa-bromo-beta-(5-imidazolil)propiónico; cloroacetil fosfato, N-alquilmaleimidias, disulfuro de 3-nitro-2-piridilo; disulfuro de metil 2-piridilo; p-cloromercuribenzoato; 2-cloromercuri-4 nitrofenol; o, cloro-7-nitrobenzo-oxa-1,3-diazol. Los agentes miméticos de lisina pueden ser generados (y residuos amino terminales pueden ser alterados) mediante reacción del lisinilo, por ejemplo, con anhídridos de ácido succínico u otros ácidos carboxílicos. Los agentes miméticos de los residuos de lisina y otros que contienen alfa-amino también pueden ser generados por reacción con imidoésteres, tales como metil picolinimidato, piridoxal fosfato, piridoxal, cloroborohidruro, ácido trinitro-bencenosulfónico, O-metilisourea, 2,4,-pentanediona y reacciones catalizadas por transamidasa con glioxilato. Los agentes miméticos de metionina pueden ser generados por reacción por ejemplo, con sulfóxido de metionina. Los agentes miméticos de prolina incluyen, por ejemplo, ácido pipercolico, ácido tiazolidin carboxílico, 3- o 4- hidroxil prolina, deshidroprolina, 3- o 4-metilprolina, o 3,3-dimetilprolina. Los agentes miméticos de los residuos de histidina pueden ser generada por reacción del histidilo, por ejemplo, con dietilprocarbonato o bromuro de para-bromofenacilo. Otros miméticos incluyen, por ejemplo, aquellos generados por hidroxilación de prolina y lisina; fosforilación de los grupos hidroxilo de residuos de serilo o treonilo; metilación de los grupos alfa-amino de lisina, arginina e histidina; acetilación de amina N-terminal, metilación de los residuos amida de la cadena principal o sustitución con N-metil aminoácidos; o amidación de grupos carboxilo del terminal C.

Un residuo, por ejemplo, un aminoácido, de un polipéptido de la divulgación también puede ser reemplazado por un aminoácido (o residuo peptidomimético) de la quiralidad opuesta. Por lo tanto, cualquier aminoácido de origen natural en la configuración L (que también puede denominarse como R o S, dependiendo de la estructura de la entidad química) puede ser reemplazado con un aminoácido del mismo tipo estructural químico o un peptidomimético, pero de quiralidad opuesta, denominada como el aminoácido D pero que también puede ser denominada como la forma R o S.

La divulgación también provee métodos para modificar los polipéptidos de la invención ya sea por medio de procesos naturales, tales como procesamiento post traducción (por ejemplo, fosforilación, acilación, etc.), o por técnicas de modificación químicas y los polipéptidos modificados resultantes. Las modificaciones pueden aparecer en cualquier lado del polipéptido, incluyendo la estructura principal del péptido, las cadenas laterales del aminoácido y los terminales amino o carboxilo. Se apreciará que el mismo tipo de modificación puede estar presente en el mismo grado o en diferentes grados en varios sitios en un polipéptido dado. También, un polipéptido dado puede tener muchos tipos de modificaciones. Las modificaciones incluyen acetilación, acilación, ADP-ribosilación, amidación, unión covalente de flavina, unión covalente de una fracción hemo, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un lípido o derivado de lípido, unión covalente de un fosfatidilinositol, ciclización de entrecruzamiento, formación de enlace de disulfuro, desmetilación, formación de entrecruzamientos covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicosilación, formación de anclaje a GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, pegilación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfación y adición mediada por ARN de transferencia de aminoácidos a proteína tal como arginilación. Véase, por ejemplo, Creighton, T.E., *Proteins - Structure and Molecular Properties* 2da Ed., W.H. Freeman and Company, New York (1993); *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York, páginas 1 - 12 (1983).

Los métodos de síntesis química peptídica en fase sólida también se pueden usar para sintetizar los polipéptidos o fragmentos de la invención. Tales métodos se conocen en el arte desde los comienzos de la década de 1960 (Merrifield, R. B., *J. Am. Chem. Soc.*, 85: 2149 - 2154, 1963) (Véase también Stewart, J. M. and Young, J. D., *Solid Phase Peptide Synthesis*, 2da Ed., Pierce Chemical Co., Rockford, Ill., página 11 - 12) y han sido recientemente empleados en kits de diseño y síntesis de péptidos de laboratorios, comercialmente disponibles (Cambridge Research Biocemicals). Tales kits de laboratorio comercialmente disponibles han utilizado generalmente las enseñanzas de H. M. Geysen y colaboradores, *Proc. Natl. Acad. SCL, USA*, 81: 3998 (1984) y permiten la síntesis los péptidos en las puntas de una multitud de "barras" o "alfileres" todos los cuales están conectados a una sola placa. Cuando se utiliza dicho sistema, se invierte una placa de barras o alfileres y se inserta en una segunda placa de los correspondientes pozos o reservorios, que contienen soluciones para unir o anclar un aminoácido apropiado a las puntas de los alfileres o barras. Mediante la repetición de dicha etapa del proceso, es decir, invertir e insertar las puntas de las alfileres o barras en soluciones apropiadas, se construyen los aminoácidos en los péptidos deseados. Además, se encuentran disponibles una cantidad de sistemas de síntesis peptídica FMOC. Por ejemplo, el ensamblaje de un polipéptido o fragmento se puede llevar a cabo en un soporte sólido usando un sintetizador de péptidos automatizado de Applied Biosystems, Inc. Modelo 431A^{MR}. Dicho equipo provee fácil acceso a los péptidos de la invención, ya sea mediante síntesis directa o por síntesis de una serie de fragmentos que se pueden acoplar usando otras técnicas conocidas.

La invención provee nuevas alfa amilasas, los ácidos nucleicos que las codifican, y métodos para usarlas. En un aspecto, los polipéptidos de la invención tienen actividad de glucoamilasa, como se describe aquí, incluyendo, por ejemplo, la habilidad de hidrolizar polisacáridos, oligosacáridos y/o almidones, en azúcares. En un aspecto, los polipéptidos de la invención tienen una actividad de amilasa. En aspectos alternativos, las amilasas de la invención tienen actividades que han sido modificadas a partir de aquellas del ejemplo de amilasas descritas aquí.

La invención incluye amilasas de la invención con y sin secuencias señal (incluyendo secuencias señal de la invención, véase por ejemplo, la Tabla 1, u otras secuencias señal y las secuencias señal en sí mismas (por ejemplo, Tabla 1). La divulgación también incluye polipéptidos (por ejemplo, proteínas de fusión) que comprenden una secuencia señal de la invención, véase, por ejemplo, la Tabla 1. El polipéptido que comprende una secuencia señal de la divulgación puede ser una amilasa y/o una glucoamilasa de la invención u otra amilasa u otra enzima u otro polipéptido.

La divulgación incluye amilasas inmovilizadas, glucoamilasas, anticuerpos de antiglucoamilasa, de antiamilasa y fragmentos de los mismos. La divulgación provee métodos para inhibir la actividad de amilasa y/o de glucoamilasa, por ejemplo, usando mutantes negativos dominantes o anticuerpos antiamilasa o antiglucoamilasa de la divulgación. La divulgación incluye hetero complejos, por ejemplo, proteínas de fusión, heterodímeros, etc., que comprenden las amilasas y/o glucoamilasas de la divulgación.

En un aspecto, alfa amilasas de la invención hidrolizan los enlaces polisacáridos o oligosacáridos internos, por ejemplo, enlaces α -1,4- y 1,6-glucosídicos en almidón para producir maltodextrinas de menor peso molecular. En un aspecto, esta hidrólisis es altamente aleatoria. Por lo tanto, la invención provee métodos para producir maltodextrinas de menor peso molecular.

Las amilasas de la invención pueden ser usadas en establecimientos industriales y de laboratorio para hidrolizar polisacáridos, oligosacáridos o almidón o cualquier compuesto que comprende maltodextrina para una variedad de propósitos. Estas amilasas se pueden usar solas para proveer una hidrólisis específica o se pueden combinar con otras amilasas para proveer un "cóctel" con un amplio espectro de actividad. Los ejemplos de usos incluyen la remoción o hidrólisis parcial o completa de polisacáridos, oligosacáridos o almidón o cualquier compuesto que comprenda maltodextrina a partir de muestras biológicas, alimenticias, pienso animal, farmacéuticas o industriales.

Por ejemplo, las amilasas de la presente invención se pueden formular en detergentes de lavandería para ayudar en la eliminación de polisacárido que comprende, por ejemplo, manchas que contienen almidón. En un aspecto, la invención provee detergentes que comprenden amilasas de la invención, incluyendo amilasas activas bajo condiciones alcalinas y métodos de fabricación y sus usos. Estas composiciones detergentes incluyen las soluciones y aplicación en lavandería y máquinas para lavado de vajilla (por ejemplo, lavado automático de vajillas). Las amilasas de la invención se pueden usar como agentes de limpieza en cualquier matriz de detergente (véanse las aplicaciones industriales más adelante). Las amilasas de la presente invención se pueden usar en las etapas iniciales de procesamiento (licuefacción) del polisacárido, por ejemplo, almidón, en la molienda húmeda de maíz, en la producción del alcohol, en la industria textil para desencolado de polisacáridos, por ejemplo, almidón, en aplicaciones de panadería y pastelería, en la industria de las bebidas, en campos petroleros en procesos de perforación; en el entintado de papel reciclado; y en piensos animales.

Las amilasas de la invención pueden tener actividad de amilasa bajo diferentes condiciones, por ejemplo, bajo pH y/o temperatura extremos, agentes oxidantes y similares. La divulgación provee métodos que conducen a preparaciones de amilasas alternativas con diferentes eficiencias y estabildades catalíticas, por ejemplo, con respecto a temperatura, agentes oxidantes y cambios en las condiciones de lavado. En un aspecto, se pueden producir variantes de amilasas usando técnicas de mutagénesis dirigida al sitio y/o mutagénesis aleatoria. En un aspecto, se puede usar evolución dirigida para producir una gran variedad de variantes de amilasa con especificidades y estabilidad alternativas.

La divulgación también provee métodos para descubrir nuevas amilasas usando los ácidos nucleicos y polipéptidos de la invención. En un aspecto, se criban bibliotecas de fagos lambda para el descubrimiento de amilasas con base en la expresión. En un aspecto, la divulgación usa bibliotecas de fagos lambda en el cribado para permitir la detección de clones tóxicos; el acceso mejorado al sustrato; menor necesidad de manipular un huésped genéticamente, derivando el potencial para cualquier tendencia generada a partir de la escisión en masa de la biblioteca; y, crecimiento más rápido a densidades clónicas bajas. El cribado de bibliotecas de fagos lambda puede ser en fase líquida o en fase sólida. En un aspecto, la divulgación provee cribado en fase líquida. Esto permite una mayor flexibilidad en las condiciones del ensayo; flexibilidad adicional de sustrato; mayor sensibilidad para clones débiles; y facilidad de automatización sobre el cribado en fase sólida.

La divulgación provee métodos de cribado que utilizan las proteínas y ácidos nucleicos de la invención y automatización robótica para permitir la ejecución de muchos miles de reacciones biocatalíticas y ensayos de cribado en un corto período de tiempo, por ejemplo, por día, así como también garantizar un alto nivel de precisión y reproducibilidad (véase la discusión de los arreglos, más adelante). Como resultado de esto, se puede producir una biblioteca de compuestos derivados en cuestión de semanas. Para enseñanzas adicionales sobre la modificación de moléculas, incluyendo moléculas pequeñas, véase el documento PCT/U894/09174.

La presente invención incluye enzimas amilasa que son variantes de carbonil hidrolasa que no son de origen natural (por ejemplo, variantes de amilasa y/o de glucoamilasa) que tienen actividad, estabilidad, especificidad de sustrato, perfil de pH y/o desempeño proteolítico diferente característico comparado con la carbonil hidrolasa precursora a partir de la cual se deriva la secuencia de aminoácidos de la variante. Específicamente, tales variantes de amilasa tienen una secuencia de aminoácidos que no se encuentra en la naturaleza, que se deriva por sustitución de una pluralidad de residuos de aminoácidos de una amilasa y/o glucoamilasa precursora con diferentes aminoácidos. La amilasa precursora puede ser una amilasa de origen natural o una amilasa recombinante. Las variantes útiles de amilasa abarcan la sustitución de cualquiera de los L-aminoácidos de origen natural en las posiciones designadas de los residuos de aminoácidos.

Secuencias señal de la amilasa y la glucoamilasa

La divulgación provee secuencias señal que consisten de o que comprenden un péptido que tiene una secuencia que comprende los residuos 1 a 12, 1 a 13, 1 a 14, 1 a 15, 1 a 16, 1 a 17, 1 a 18, 1 a 19, 1 a 20, 1 a 21, 1 a 22, 1 a 23, 1 a 24, 1 a 25, 1 a 26, 1 a 27, 1 a 28, 1 a 29, 1 a 30 o 1 a 31, 1 a 32, 1 a 33, 1 a 34, 1 a 35, 1 a 36, 1 a 37, 1 a 38, o 1 a 39, o más largos, de un polipéptido de la invención. Por ejemplo, la divulgación provee una secuencia señal de amilasas (por ejemplo, alfa amilasas) o glucoamilasa y ácidos nucleicos que codifican esta secuencia señal, por ejemplo, ejemplos de péptidos de la divulgación que tienen secuencias como se expone en la Tabla anterior.

Las secuencias señal de amilasa y/o glucoamilasa de la divulgación pueden ser péptidos aislados, o, secuencias unidas a otro polipéptido de amilasa y/o glucoamilasa o que no es de amilasa o no es de glucoamilasa, por ejemplo, como una proteína de fusión. En un aspecto, la divulgación provee polipéptidos que comprenden secuencias señal de amilasa y/o glucoamilasa de la invención. En un aspecto, los polipéptidos que comprenden secuencias señal de amilasa y/o glucoamilasa de la divulgación comprenden secuencias heterólogas a una amilasa y/o glucoamilasa de la divulgación (por ejemplo, una proteína de fusión que comprende una secuencia señal de amilasa de la divulgación y secuencias de otra proteína amilasa o que no es amilasa). En un aspecto, la invención provee amilasas y/o glucoamilasas de la invención con secuencias señal heterólogas, por ejemplo, secuencias con una secuencia señal

de levadura. Por ejemplo, una amilasa de la invención comprende una secuencia señal heteróloga en vectores, por ejemplo, un vector de la serie pPIC (Invitrogen, Carlsbad, CA).

En un aspecto, las secuencias señal de la divulgación se identifican después de la identificación de nuevos polipéptidos de amilasa y/o glucoamilasa. Las rutas por las cuales se clasifican y transportan las proteínas a su ubicación celular apropiada a menudo se denominan como rutas de direccionamiento de proteína. Uno de los elementos más importantes en todos estos sistemas de direccionamiento es una secuencia de aminoácidos corta en el terminal amino de un polipéptido recientemente sintetizado, denominada secuencia señal. Esta secuencia señal dirige una proteína a su ubicación apropiada en la célula y se remueve durante el transporte o cuando la proteína alcanza su destino final. La mayoría de las proteínas liposomales, de membrana, o secretadas tienen una secuencia señal en el terminal amino que las marca para translocación en el lumen del retículo endoplasmático. Se han determinado más de 100 secuencias señal para las proteínas en este grupo. Las secuencias señal pueden variar en longitud desde aproximadamente 13 a 36, o entre aproximadamente 10 y 40 residuos de aminoácidos. Varios métodos de reconocimiento de las secuencias señal son conocidos por aquellos entrenados en el arte. Por ejemplo, en un aspecto, se identifican los nuevos péptidos señal de amilasa por medio de un método denominado como SIGNALP^{MR}. SignalP^{RM} utiliza una red neural combinada que reconoce tanto los péptidos señal como sus sitios de escisión; véase, por ejemplo, Nielsen (1997) "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites", Protein Engineering, vol. 10, no. 1, página 1 - 6.

Se entiende que en algunos aspectos, las amilasas de la invención pueden no tener secuencias señal. En un aspecto, la invención provee las amilasas de la invención que carecen totalmente o parte de una secuencia señal, por ejemplo, las secuencias señal de la divulgación (véase la Tabla 1). En un aspecto, la divulgación provee una secuencia de ácido nucleico que codifica una secuencia señal de una amilasa operativamente enlazada a una secuencia de ácido nucleico de una amilasa diferente o, opcionalmente, se puede desear una secuencia señal de una proteína no amilasa. La Tabla 1 muestra ejemplos de secuencias señal de la divulgación.

Amilasa y glicoamilasa prepro y secuencias señal y dominios catalíticos

Además de las secuencias señal (por ejemplo, péptidos señal (SPS)), como se discutió anteriormente, la divulgación provee dominios prepro y dominios catalíticos (CD). Los SP, dominios prepro y/o los CD de la divulgación pueden ser péptidos aislados, sintéticos o recombinantes o pueden ser parte de una proteína de fusión, por ejemplo, como dominio heterólogo en una proteína quimérica. La divulgación provee ácidos nucleicos que codifican estos dominios catalíticos (CD) (por ejemplo, "sitio activos"), dominios prepro y secuencias señal (SP, por ejemplo, un péptido que tiene una secuencia que comprende/consiste de residuos amino terminales de un polipéptido de la invención).

Las secuencias señal (SP) de amilasa y/o glucoamilasa, dominios catalíticos (CD) y/o secuencias prepro de la divulgación pueden ser péptidos aislados, o, secuencias unidas a otro polipéptido de amilasa o que no es amilasa o que no es glucoamilasa, por ejemplo, como una proteína de fusión (quimérica). En un aspecto, los polipéptidos que comprenden secuencias señal SP de amilasa y/o prepro de la divulgación comprenden secuencias heterólogas para amilasas de la invención (por ejemplo, una proteína de fusión que comprende una SP y/o una prepro de la divulgación y secuencias de otra proteína amilasa y/o glucoamilasa, o una proteína que no es amilasa o una que no es glucoamilasa). En un aspecto, la divulgación provee amilasas de la invención con secuencias CD, SP y/o prepro heterólogas, por ejemplo, secuencias con una secuencia señal de levadura. Una amilasa de la invención puede comprender un CD, SP y/o prepro heteróloga en un vector, por ejemplo, un vector de la serie pPIC (Invitrogen, Carlsbad, CA).

En un aspecto, las secuencias SP, CD, y/o prepro de la divulgación se identifican después de la identificación de nuevos polipéptidos de amilasa. Las rutas por medio de las cuales se clasifican y transportan las proteínas a sus ubicaciones celulares apropiadas a menudo son denominadas como rutas de direccionamiento de proteínas. Uno de los elementos más importantes en todos estos sistemas de direccionamiento es una secuencia corta de aminoácidos en el terminal amino de un polipéptido recientemente sintetizado llamado secuencia señal. Esta secuencia señal dirige una proteína a su ubicación apropiada en la célula y es removido durante el transporte o cuando la proteína alcanza su destino final. La mayoría de las proteínas liposomales, de membrana, o secretadas tienen una secuencia señal amino-terminal que las marca para translocación en el lumen del retículo endoplasmático. Las secuencias señal pueden variar en longitud desde 13 a 45 o más residuos de aminoácidos. Diferentes métodos de reconocimiento de las secuencias señal son conocidas por aquellos versados en el arte. Por ejemplo, en un aspecto, se identifican nuevos péptidos señal de hidrolasa por medio de un método denominado como SignalP. SignalP utiliza una red neural combinada que reconoce tanto los péptidos señal como sus sitios de escisión. (Nielsen, y colaboradores, "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites." Protein Engineering, vol. 10, no. 1, página 1 - 6 (1997).

En algunos aspectos, una amilasa de la invención puede no tener secuencias de SP y/o prepro, y/o dominios catalíticos (CD). En un aspecto, la divulgación provee amilasas que carecen de todo o parte de un SP, un CD y/o un dominio prepro. En un aspecto, la divulgación provee una secuencia de ácido nucleico que codifica una secuencia señal (SP), un CD y/o prepro de una amilasa y/o glucoamilasa operativamente enlazada a una secuencia de ácido

nucleico de una amilasa y/o glucoamilasa diferente, o, opcionalmente, pueden desearse una secuencia señal (SP), un CD y/o un dominio prepro de una proteína no amilasa o no glucoamilasa.

La divulgación también provee polipéptidos aislados, sintéticos o recombinantes que comprenden secuencias señal (SP), un dominio prepro y/o dominios catalíticos (CD) de la divulgación y secuencias heterólogas. Las secuencias heterólogas son secuencias que no asociadas en forma natural (por ejemplo, a una amilasa) con un SP, un dominio prepro y/o un CD. Las secuencias a las cuales no están naturalmente asociados el SP, el dominio prepro y/o el CD pueden estar sobre el extremo del terminal amino, el extremo del terminal carboxilo del SP, el dominio prepro y/o del CD, y/o sobre ambos extremos del SP y/o del CD. En un aspecto, la divulgación provee un polipéptido aislado, sintético o recombinante que comprende (o consiste de) un polipéptido que comprende una secuencia señal (SP), dominio prepro y/o un dominio catalítico (CD) de la divulgación con la condición de que no este asociado con ninguna secuencia a la cual este asociado naturalmente (por ejemplo, una secuencia de amilasa y/o glucoamilasa). En forma similar, en un aspecto, la divulgación provee ácidos nucleicos aislados, sintéticos o recombinantes que codifican estos polipéptidos. Por lo tanto, en un aspecto, el ácido nucleico aislado, sintético o recombinante de la divulgación comprende una secuencia codificadora para una secuencia señal (SP), un dominio prepro y/o un dominio catalítico (CD) de la invención y una secuencia heteróloga (es decir, una secuencia no asociada naturalmente con la secuencia señal (SP), el dominio prepro y/o el dominio catalítico (CD) de la invención). La secuencia heteróloga puede estar sobre el extremo terminal 3', el extremo terminal 5' y/o sobre ambos extremos de la secuencia del SP, el dominio prepro y/o del CD.

Los polipéptidos de la invención incluyen amilasas en forma activa o inactiva. Por ejemplo, los polipéptidos de la invención incluyen proproteínas antes de la "maduración" o el procesamiento de secuencias prepro, por ejemplo, por una enzima de procesamiento de proproteínas, tal como una proproteína convertasa para generar una proteína madura "activa". Los polipéptidos de la invención incluyen amilasas inactivas por otras razones, por ejemplo, antes de la "activación" por un evento de procesamiento de post-traducción, por ejemplo, una acción de endo- o exopeptidasa o proteinasa, un evento de fosforilación, una amidación, una glicosilación o una sulfación, un evento de dimerización y similares. Los métodos para identificar secuencias de dominio "prepro", los CD y secuencias señal son conocidos en el arte, véase, por ejemplo, Van de Ven (1993) Crit. Rev. Oncog. 4(2): 115 - 136. Por ejemplo, para identificar una secuencia prepro, la proteína es purificada a partir de un espacio extracelular y se determina y se compara la secuencia del terminal N de la proteína con la forma no procesada.

Los polipéptidos de la invención incluyen todas las formas activas, incluyendo las subsecuencias activas, por ejemplo, los dominios catalíticos (CD) o los sitios activos, de una enzima de la invención. En un aspecto, la invención provee dominios catalíticos o sitios activos como se expone más adelante. En un aspecto, la invención provee un péptido o polipéptido que comprende o que consiste de un dominio de un sitio activo como se predice a través del uso de una base de datos tal como Pfam (que es una gran colección de múltiples alineaciones de secuencias y modelos Markov ocultos que cubren muchas familias de proteínas comunes, la base de datos Pfam de las familias de proteína, A. Bateman, E. Birney, L. Cerruti, R. Durbin, L. Ewinger, S.R. Eddy, S. Griffiths-Jones, K.L. Howe, M. Marshall y E.L.L. Sonnhammer, Nucleic Acid Research. 30(1):276-280. 2002) o equivalente.

Amilasas y glucoamilasas híbridas y bibliotecas de péptidos

En un aspecto, la divulgación provee amilasas y/o glucoamilasas híbridas y proteínas de fusión, incluyendo bibliotecas de péptidos, que comprenden secuencias de la invención. Las bibliotecas de péptidos de la divulgación se pueden usar para aislar moduladores de péptidos (por ejemplo, activadores o inhibidores) de objetivos, tales como receptores, enzimas, sustratos de glucoamilasa. Las bibliotecas de péptidos de la divulgación se pueden usar para identificar los patrones de enlace formales de blancos, tales como ligandos, por ejemplo, citoquinas, hormonas y similares.

En un aspecto, las proteínas de fusión de la divulgación (por ejemplo, la fracción peptídica) están conformacionalmente estabilizadas (con relación a los péptidos lineales) para permitir una mayor afinidad de enlazamiento superior para los objetivos. La divulgación provee fusiones de amilasas de la invención y otros péptidos, incluyendo péptidos conocidos y aleatorios. Se pueden fusionar de tal manera que no se perturbe significativamente la estructura de las amilasas y/o glucoamilasas y el péptido está metabólicamente o estructuralmente conformacionalmente estabilizado. Esto permite la creación de una biblioteca de péptidos que sea monitoreada fácilmente por su presencia dentro de las células y su cantidad.

Las variantes de secuencia de aminoácidos de la invención pueden ser caracterizadas por una naturaleza predeterminada de la variación, una característica que las fija separadamente de una forma de origen natural, por ejemplo, una variación alélica o entre especies de una secuencia de amilasa. En un aspecto, las variantes de la invención exhiben la misma actividad cualitativa biológica que el análogo de origen natural. En forma alternativa, las variantes pueden ser seleccionadas por tener características modificadas. En un aspecto, si bien el sitio o región para introducir una variación de secuencia de aminoácidos es predeterminado, la mutación por sí misma no requiere ser predeterminada. Por ejemplo, a fin de optimizar el desempeño de una mutación en un sitio dado, se puede realizar una mutagénesis aleatoria en el codón o región objetivo y se pueden cribar las variantes de amilasa y/o

glucoamilasa expresadas para una combinación óptima de la actividad deseada. Las técnicas para hacer mutaciones de sustitución en sitios predeterminados de ADN que tiene una secuencia conocida son bien conocidas, como se discute aquí, por ejemplo, mutagénesis por cebador M13 y mutagénesis por PCR. El cribado de los mutantes puede hacerse usando ensayos de actividades proteolíticas. En aspectos alternativos, las sustituciones de aminoácidos pueden ser residuos individuales; las inserciones pueden estar en el orden de aproximadamente 1 a 20 aminoácidos, si bien se pueden realizar inserciones considerablemente mayores. Las supresiones pueden oscilar desde aproximadamente 1 hasta aproximadamente 20, 30, 40, 50, 60, 70 residuos o más. Para obtener un derivado final con propiedades óptimas, se pueden utilizar sustituciones, supresiones, inserciones, o cualquier combinación de las mismas. En general, estos cambios se realizan en pocos aminoácidos para minimizar la alteración de la molécula. Sin embargo, se pueden tolerar cambios mayores en ciertas circunstancias.

La invención provee amilasas donde la estructura de la cadena principal del polipéptido, la estructura secundaria o la terciaria, por ejemplo, una estructura de hélice alfa o de lámina beta, ha sido modificada. En un aspecto, la carga o hidrofobicidad han sido modificados. En un aspecto, el volumen de una cadena lateral ha sido modificado. Se hacen cambios sustanciales en la identidad inmunológica o de la función, seleccionando sustituciones que sean menos conservadoras. Por ejemplo, se pueden hacer sustituciones que afectan en forma más significativa: la estructura de la cadena principal del polipéptido en el área de la alteración, por ejemplo una estructura de hélice alfa o una lámina beta; una carga o un sitio hidrófobo de la molécula, que puede estar en un sitio activo; o una cadena lateral. La invención provee sustituciones en polipéptidos de la invención en donde (a) un residuo hidrofílico, por ejemplo, serilo o treonilo, es sustituido por (o mediante) un residuo hidrófobo, por ejemplo, leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (b) una cisteína o prolina es sustituida por (o mediante) cualquier otro residuo; (c) un residuo que tiene una cadena lateral electropositiva, por ejemplo, lisilo, arginilo, o histidilo, se sustituye por (o mediante) un residuo electronegativo, por ejemplo glutamilo o aspartilo; o (d) un residuo que tiene una cadena lateral voluminosa, por ejemplo, fenilalanina, se sustituye por (o mediante) uno que no tiene cadena lateral, por ejemplo glicina. Las variantes pueden exhibir la misma actividad biológica cualitativa (es decir actividad de amilasa y/o de glucoamilasa) aunque se pueden seleccionar variantes para modificar las características de las amilasas según sea necesario.

En un aspecto, amilasas de la invención comprenden epítopos o etiquetas de purificación, secuencias señal u otras secuencias de fusión, etc. En un aspecto, las Amilasas de la invención se pueden fusionar con un péptido aleatorio para formar un polipéptido de fusión. Por "fusionado" u "operativamente enlazado" se entiende aquí, que el péptido aleatorio y la amilasa se enlazan entre sí, de tal forma que se minimiza la ruptura de la estabilidad de la estructura de la amilasa, por ejemplo, ,retiene la actividad de amilasa. El polipéptido de fusión (o el polinucleótido de fusión que codifica el polipéptido de fusión), puede comprender también componentes adicionales, incluyendo múltiples péptidos en múltiples bucles.

En un aspecto, los péptidos y ácidos nucleicos que los codifican son aleatorizados, ya sea completamente aleatorizados o se parcializan en su aleatorización, por ejemplo, generalmente en la frecuencia de nucleótidos/residuos o por posición. "Aleatorizado" significa que cada ácido nucleico y péptido consiste de nucleótidos y aminoácidos esencialmente aleatorios, respectivamente. En un aspecto, los ácidos nucleicos que dan lugar a los péptidos pueden ser químicamente sintetizados y, por lo tanto, pueden incorporar cualquier nucleótido en cualquier posición. Por lo tanto, cuando los ácidos nucleicos se expresan para formar péptidos, cualquier residuo de aminoácido puede ser incorporado en cualquier posición. El proceso de síntesis se puede diseñar para generar ácidos nucleicos aleatorizados, para permitir la formación de todas o de la mayoría de las combinaciones posibles sobre la longitud del ácido nucleico, formando así una biblioteca de ácidos nucleicos aleatorizados. La biblioteca puede proveer una población suficiente y estructuralmente diversa de productos de expresión aleatorizados para afectar un rango probabilísticamente suficiente de respuestas celulares para proveer una o más células que exhiben una respuesta deseada. Por lo tanto, la divulgación provee una biblioteca de interacción lo suficientemente grande para que al menos uno de sus miembros tenga una estructura que le brinde afinidad por alguna molécula, proteína u otro factor.

Metodologías de cribado y dispositivos de monitoreo "en línea"

Al llevar a la práctica los métodos de la invención, se pueden utilizar una variedad de aparatos y metodologías para, en conjunto con los polipéptidos y ácidos nucleicos de la invención, por ejemplo, cribar polipéptidos para actividad de amilasa, cribar compuestos como moduladores potenciales, por ejemplo, activadores o inhibidores, de actividad de amilasa, para anticuerpos que se enlazan a un polipéptido de la invención, para ácidos nucleicos que hibridan con un ácido nucleico de la invención, para cribar células que expresan un polipéptido de la invención y similares.

Arreglos capilares

Los arreglos capilares, tales como la GIGAMATRIX^{MR}, Diversa Corporation, San Diego, CA, pueden ser usados en los métodos de la invención. Los ácidos nucleicos o polipéptidos de la invención pueden ser inmovilizados en o aplicados a un arreglo, incluyendo arreglos capilares. Los arreglos se pueden usar para cribar o monitorear bibliotecas de composiciones (por ejemplo, moléculas pequeñas, anticuerpos, ácidos nucleicos, etc.) por su habilidad para enlazarse con o modular la actividad de un ácido nucleico o un polipéptido de la invención. Los

arreglos capilares proporcionan otro sistema para mantener o cribar muestras. Por ejemplo, un aparato para cribar muestras puede incluir una pluralidad de capilares formados en un arreglo de capilares adyacentes, en donde cada capilar comprende al menos una pared que define un lumen para retener una muestra. El aparato puede incluir, además, material intersticial dispuesto entre capilares adyacentes en un arreglo y uno o más indicios de referencia formados dentro del material intersticial. Un capilar para cribar una muestra, en donde el capilar se adapta para ser enlazado en un arreglo de capilares, puede incluir una primera pared que define un lumen para retener la muestra, y una segunda pared formada de un material de filtración, para filtrar la energía de excitación suministrada al lumen para excitar la muestra.

Un polipéptido o ácido nucleico, por ejemplo, un ligando, puede ser introducido en un primer componente en al menos una porción de un capilar de un arreglo capilar. Cada capilar del arreglo capilar puede comprender al menos una pared que define un lumen para retener al primer componente. Se puede introducir una burbuja de aire en el capilar detrás del primer componente. Se puede introducir un segundo componente en el capilar, en donde el segundo componente se separa del primer componente mediante la burbuja de aire. Se puede introducir una muestra de interés como un primer líquido marcado con una partícula detectable en un capilar de un arreglo capilar, en donde cada capilar del arreglo capilar comprende al menos una pared que define un lumen para retener el primer líquido y la partícula detectable y en donde al menos una pared es recubierta con un material de enlazamiento para enlazar la partícula detectable con al menos una pared. El método puede incluir, además, remover el primer líquido del tubo capilar, en donde la partícula detectable enlazada se mantiene dentro del capilar y se introduce un segundo líquido en el tubo capilar.

El arreglo capilar puede incluir una pluralidad de capilares individuales que comprenden al menos una pared exterior que define un lumen. La pared exterior del capilar puede ser una o más paredes fusionadas entre sí. En forma similar, la pared puede definir un lumen que es cilíndrico, cuadrado, hexagonal o de cualquier otra forma geométrica siempre que las paredes formen un lumen para la retención de un líquido o muestra. Los capilares del arreglo capilar pueden ser mantenidos juntos en forma próxima para formar una estructura plana. Los capilares se pueden enlazar entre sí, al ser fusionados (por ejemplo en donde los capilares se hacen de vidrio), pegados, adheridos o unidos lado a lado. El arreglo capilar puede estar formado por cualquier cantidad de capilares individuales. por ejemplo, en el intervalo de 100 a 4000000 capilares. Un arreglo capilar puede formar una placa de microtitulación que tiene aproximadamente 100.000 o más capilares individuales enlazados entre sí.

Arreglos o "biochips"

Los ácidos nucleicos o polipéptidos de la invención se pueden inmovilizar a o, aplicar a un arreglo, por ejemplo, un "arreglo" o "microarreglo" o "biochip" o "chip", que en una realización comprenden una pluralidad de elementos objetivo, cada elemento objetivo comprendiendo una cantidad definida de uno o más polipéptidos (incluyendo anticuerpos) o ácidos nucleicos inmovilizados sobre un área definida de una superficie de sustrato, en donde al menos uno de los "elementos objetivo" es un polipéptido (por ejemplo, una enzima o un anticuerpo) de la invención, o un ácido nucleico de la invención.

Los arreglos de la divulgación se pueden utilizar para cribar o monitorear bibliotecas de composiciones (por ejemplo, moléculas pequeñas, anticuerpos, ácidos nucleicos, etc.) por su habilidad para unirse a o modular la actividad de un ácido nucleico o un polipéptido de la invención. Por ejemplo, un parámetro monitorizado es la expresión de un transcripto de un gen de amilasa y/o de glucoamilasa. Uno o más, o todos los transcriptos de una célula pueden ser medidos mediante hibridación de una muestra que comprende transcriptos de la célula, o ácidos nucleicos representativos de, o complementarios de transcriptos de una célula, mediante hibridación con ácidos nucleicos inmovilizados sobre un arreglo, o "biochip". Por medio del uso de un "arreglo" de ácidos nucleicos sobre un microchip, algunos o todos los transcriptos de una célula pueden ser simultáneamente cuantificados. En forma alternativa, los arreglos que comprenden ácido nucleico genómico también pueden ser utilizados para determinar el genotipo de una cepa recientemente modificada por ingeniería genética por los métodos de la invención. Los "arreglos de polipéptidos" también se pueden usar cuantificar simultáneamente una pluralidad de proteínas. La presente invención también se pueden llevar a la práctica con cualquier "arreglo" conocido, también denominado como un "microarreglo" o "arreglo de ácido nucleico" o "arreglo de polipéptido" o "arreglo de anticuerpo" o "biochip", o variación de los mismos. Los arreglos son genéricamente una pluralidad de "puntos" o "elementos objetivo", cada elemento objetivo comprendiendo una cantidad definida de una o más moléculas biológicas, por ejemplo, oligonucleótidos, inmovilizados sobre una área definida de una superficie de sustrato para enlamiento específico con una molécula de la muestra, por ejemplo, transcriptos de ARNm.

En la realización de los métodos de la divulgación, cualquier arreglo y/o método conocidos para la elaboración y uso de arreglos pueden ser incorporados en su totalidad o en parte, o variaciones de los mismos, como se describe, por ejemplo, en las patentes de los Estados Unidos Nos. 6.277.628, 6.277.489, 6.261.775, 6.258.506, 6.054.270, 6.048.695, 6.045.996, 6.022.963, 6.013.440, 5.965.452, 5.959.098, 5.856.174, 5.830.645, 5.770.456, 5.632.957, 5.556.752, 5.143.854, 5.807.522, 5.800.992, 5.744.305, 5.700.637, 5.556.752, 5.434.049; véanse también, por ejemplo, los documentos WO 99151773; WO 99/09217; WO 97/46313; WO 96/17958; véanse también, por ejemplo, Johnston (1998) Curr. Biol. 82R171 - R174; Schummer (1997) Biotechniques 23: 1087 - 1092; Kern (1997)

Biotechniques 23: 120 - 124; Solinas-Toldo (1997) Genes, Chromosomes & Cancer 20: 399 - 407; Bowtell (1999) Nature Genetics Supp. 21: 25 - 32. Véanse también las solicitudes de patente de los Estados Unidos publicadas Nos. 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765.

Métodos de cribado basados en anticuerpo y anticuerpos

5 La divulgación provee anticuerpos aislados, sintéticos o recombinantes que se unen específicamente con una amilasa de la invención. Estos anticuerpos se pueden utilizar para aislar, identificar o cuantificar la amilasa de la invención o polipéptidos relacionados. Estos anticuerpos se pueden usar para aislar otros polipéptidos dentro del alcance de la invención u otras amilasas relacionadas. Los anticuerpos pueden ser diseñados para enlazarse con un sitio activo de una amilasa. Por lo tanto, la divulgación provee métodos para inhibir amilasas usando los anticuerpos de la divulgación.

15 La divulgación provee anticuerpos que comprenden un péptido o polipéptido derivados de, modelados después o sustancialmente codificados por un gen de inmunoglobulina o genes de inmunoglobulina, o fragmentos de los mismos, capaces de enlazarse específicamente con un antígeno o epítipo, véase, por ejemplo Fundamental Immunology, Third Edition, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) J. Immunol. Methods 175: 267 - 273; Yarmush (1992) J. Biochem. Biophys. Methods 25: 85 - 97. El término anticuerpo incluye porciones que se enlazan al antígeno, es decir, "sitios de enlazamiento con el antígeno", (por ejemplo, fragmentos, subsecuencias, regiones determinantes de complementariedad (CDR)), que retienen capacidad para enlazarse con el antígeno, incluyendo (i) un fragmento Fab, un fragmento monovalente que consiste de los dominios VL, VH, CL y CH1; (ii) un fragmento F(ab')₂, un fragmento bivalente que comprende dos fragmentos Fab enlazados por un puente de disulfuro en la región bisagra; (iii) un fragmento Fd que consiste de los dominios VH y CH1; (iv) un fragmento Fv que consiste de los dominios VL y VH de un brazo individual de un anticuerpo, (v) un fragmento dAb (Ward y colaboradores, (1989) Nature 341: 544 - 546), que consiste de un dominio VH; y (vi) una región aislada determinante de complementariedad (CDR). Los anticuerpos de una sola cadena también se incluyen por referencia en el término anticuerpo".

25 Los anticuerpos se pueden utilizar en inmunoprecipitación, tinción, columnas de inmovilización y similares. Si se desea, las secuencias de ácido nucleico que codifican para antígenos específicos se pueden generar por inmunización seguida de aislamiento del polipéptidos o ácido nucleico, amplificación o clonación e inmovilización del polipéptidos sobre un arreglo de la divulgación. En forma alternativa, los métodos de la divulgación se pueden utilizar para modificar la estructura de un anticuerpo producido por una célula que va a ser modificada, por ejemplo, una afinidad de anticuerpo se puede aumentar o disminuir. Además, la habilidad para elaborar o modificar anticuerpos puede ser un fenotipo modificado por ingeniería genética en una célula por medio de los métodos de la divulgación.

35 Los métodos de inmunización, producción y aislamiento de anticuerpos (policlonales y monoclonales) son conocidos por aquellos capacitados en el arte y descritos en la literatura científica y de patentes, véase, por ejemplo, Coligan, Current Protocols In Immunology, Wiley/Greene, NY (1991); Stites (eds.) Basic And Clinical Immunology (7th ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, Monoclonal Antibodies: Principles And Practice (2d ed.) Academic Press, New York, NY (1986); Kohler (1975) Nature 256: 495; Harlow (1988) Antibodies, A Laboratory Manual. Cold Spring Harbor Publicaciones, New York. Los anticuerpos también se pueden generar *in vitro*, por ejemplo, usando sitios de enlazamiento al anticuerpo recombinante que expresan bibliotecas de despliegue en fagos, además de los métodos tradicionales *in vivo* usando animales. Véase, por ejemplo, Hoogenboom (1997) Trends Biotechnol. 15: 62 - 70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26: 27 - 45.

45 Los polipéptidos o péptidos se pueden usar para generar anticuerpos que se enlazan específicamente con los polipéptidos, por ejemplo, las ilasas de la invención. Los anticuerpos resultantes se pueden usar en procedimientos de cromatografía de inmovilización para aislar o purificar el polipéptido o para determinar si el polipéptido está presente en una muestra biológica. En tales procedimientos, se pone en contacto una preparación de proteína, tal como un extracto, o una muestra biológica con un anticuerpo capaz de enlazarse específicamente con uno de los polipéptidos de la invención.

50 En procedimientos de inmovilización, se une el anticuerpo a un soporte sólido, tal como una perla u otra matriz de columna. La preparación de proteína se pone en contacto con el anticuerpo bajo condiciones en las cuales el anticuerpo se enlaza específicamente con uno de los polipéptidos de la invención. Después de un lavado para remover las proteínas no enlazadas específicamente, se eluyen los polipéptidos enlazados específicamente.

55 La habilidad de las proteínas en una muestra biológica para enlazarse con el anticuerpo puede ser determinada usando cualquiera entre una variedad de procedimientos familiares para aquellos capacitados en el arte. Por ejemplo, el enlazamiento se puede sr determinar mediante la marcación del anticuerpo con una etiqueta detectable tal como un agente fluorescente, una etiqueta enzimática o un radioisótopo. En forma alternativa, se puede detectar el enlazamiento del anticuerpo con la muestra usando un anticuerpo secundario que tiene dicha etiqueta detectable sobre el mismo. Los ensayos particulares incluyen ensayos de ELISA, ensayos tipo sándwich, radioinmunoensayos y transferencias tipo Western.

Los anticuerpos policlonales generados contra los polipéptidos de la invención se pueden obtener por inyección directa de los polipéptidos en un animal o administrando los polipéptidos a un animal no humano. El anticuerpo así obtenido se enlazaría entonces con el polipéptido en sí mismo. En esta manera, incluso una secuencia que codifica únicamente un fragmento del polipéptido puede ser utilizada para generar anticuerpos que pueden enlazarse con el polipéptido nativo completo. Tales anticuerpos pueden ser usados luego para aislar el polipéptido de las células que expresan ese polipéptido. Para la preparación de anticuerpos monoclonales, se puede usar cualquier técnica que provea anticuerpos producidos por cultivos de líneas celulares continuas. Los ejemplos incluyen la técnica de hibridoma, la técnica de trioma, la técnica de hibridoma de células B humanas y la técnica de EBV-hibridoma (véase, por ejemplo, Cole (1985) en *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., páginas 77 - 96).

5

10 Las técnicas descritas para la producción de anticuerpos monocatenarios (véase, por ejemplo, la patente de los Estados Unidos No. 4.946.778) pueden ser adaptadas para producir anticuerpos monocatenarios para los polipéptidos de la invención. En forma alternativa, se pueden utilizar ratones transgénicos para expresar anticuerpos humanizados para estos polipéptidos o fragmentos de los mismos.

15 Los anticuerpos generados contra los polipéptidos de la invención se pueden utilizar en el cribado para polipéptidos similares (por ejemplo, una amilasa) de otros organismos y muestras. En tales técnicas, los polipéptidos del organismo se ponen en contacto con el anticuerpo y se detectan aquellos polipéptidos que se enlazan específicamente con el anticuerpo son detectados. Cualquiera de los procedimientos descritos anteriormente se pueden usar para detectar el enlazamiento del anticuerpos.

Kits

20 La divulgación provee kits que comprende las composiciones, por ejemplo, ácidos nucleicos, casetes de expresión, vectores, células, semillas transgénicas o plantas o partes de plantas, o polipéptidos (por ejemplo, una amilasa y/o una glucoamilasa) y/o anticuerpos de la divulgación. Los kits también pueden contener material instructivo que enseña las metodologías y usos industriales de la invención, como se describe aquí.

Medición de parámetro metabólicos

25 Los métodos de la divulgación proveen evolución de la célula completa o modificación genética de una célula completa, de una célula para desarrollar una nueva cepa celular que tiene un nuevo fenotipo, por ejemplo, una actividad nueva o modificada de amilasa, mediante la modificación de la composición genética de la célula. La composición genética puede ser modificada mediante la adición a la célula de un ácido nucleico de la invención. Para detectar el nuevo fenotipo, se monitorea al menos un parámetro metabólico de una célula modificada en la célula en "tiempo real" o en un marco de tiempo "en línea". En un aspecto, se monitorean una pluralidad de células, tales como un cultivo de células, en "tiempo real" o "en línea". En un aspecto, se monitorean una pluralidad de parámetros metabólicos en "tiempo real" o "en línea". Los parámetros metabólicos pueden ser monitorizados usando la amilasa de la invención.

30

35 Los análisis de flujo metabólico (MFA) se basan en un marco bioquímico conocido. Se construye una matriz metabólica linealmente independiente basada en la ley de conservación de masa y la hipótesis de estado pseudo-estable (PSSH) sobre los metabolitos intracelulares. En la realización de los métodos de la divulgación, se establecen redes metabólicas incluyendo la:

- identidad de todos los sustratos de las rutas, productos y metabolitos intermediarios
 - identidad de todas las reacciones químicas que interconvierten los metabolitos de las rutas, la estequiometría de las reacciones de las rutas.
 - identidad de todas las enzimas que catalizan las reacciones, la cinética de reacción enzimática,
 - las interacciones regulatorias entre los componentes de las rutas, por ejemplo interacciones alostéricas, interacciones enzima-enzima, etc.,
 - compartimentalización intracelular de enzimas o cualquier otra organización supramolecular de las enzimas y,
 - la presencia de cualquier gradiente de concentración de metabolitos, enzimas o moléculas efectoras o barreras de difusión para su movimiento.
- 40
- 45

Una vez que se construye la red metabólica para una determinada cepa, la presentación matemática por noción de matriz se puede introducir para estimar los flujos metabólicos intracelulares si los datos del metaboloma en línea están disponibles. El fenotipo metabólico se basa en los cambios de la red metabólica completa en una célula. El fenotipo metabólico se basa en el cambio de la utilización de las rutas con respecto a condiciones ambientales, regulación genética, estado de desarrollo y el genotipo, etc. En un aspecto de los métodos de la divulgación, después del cálculo de MFA en línea, el comportamiento dinámico de las células, su fenotipo y otras propiedades son analizados mediante la investigación de la utilización de las rutas. Por ejemplo, si se aumenta el suministro de glucosa y disminuye el de oxígeno durante la fermentación con levadura, la utilización de las rutas respiratorias se reducirá y/o detendrá, y dominará la utilización de las rutas fermentativas. El control del estado fisiológico de los

50

55

5 cultivos celulares será posible después del análisis de las rutas. Los métodos de la divulgación pueden ayudar a determinar cómo manipular la fermentación al determinar cómo cambiar el suministro de sustrato, la temperatura, el uso de inductores, etc., para controlar el estado fisiológico de las células para moverse a lo largo de una dirección deseable. Al llevar a la práctica los métodos de la divulgación, los resultados de MFA también se pueden comparar con datos de transcriptoma y proteoma para diseñar experimentos y protocolos para manipulación genética metabólica o arrastre de genes, etc.

Al llevar a la práctica los métodos de la divulgación, cualquier fenotipo modificado o nuevo puede ser conferido y detectado, incluyendo características nuevas o mejoradas en la célula. Cualquier aspecto del metabolismo o del desarrollo puede ser monitorizado.

10 Monitoreo de expresión de un transcrito de ARNm

En un aspecto de la invención, el fenotipo manipulado genéticamente comprende el aumento o la disminución de la expresión de un transcrito de ARNm (por ejemplo, un mensaje de amilasa) o generar nuevos transcritos en una célula (por ejemplo, amilasa). Esta expresión incrementada o reducida puede ser rastreada al evaluar la presencia de una amilasa y/o glucoamilasa de la invención o por ensayos de actividad de amilasa. Los transcritos de ARNm, o mensajes, también pueden ser detectados y cuantificados por cualquier método conocido en el arte, incluyendo, por ejemplo, transferencias tipo Northern, reacciones de amplificación cuantitativa, hibridación con arreglos y similares. Las reacciones de amplificación cuantitativa incluyen, por ejemplo, PCR cuantitativa, incluyendo, por ejemplo, reacción en cada de la polimerasa de transcripción reversa cuantitativa, o RT-PCR; RT-PCR cuantitativa en tiempo real, o "RT-PCR cinética en tiempo real" (véase, por ejemplo, Kreuzer (2001) Br. J. Haematol. 114: 313 - 318; Xia (2001) Transplantation 72: 907 - 914).

En un aspecto de la divulgación, el fenotipo manipulado genéticamente es generado al desactivar la expresión de un gen homólogo. La secuencia codificadora del gen o uno o más elementos de control de transcripción pueden ser desactivados, por ejemplo, promotores o potenciadores. Por lo tanto, la expresión de un transcrito puede ser completamente eliminada o solamente reducida.

25 En un aspecto de la invención, el fenotipo manipulado genéticamente comprende incrementar la expresión de un gen homólogo. Esto se puede efectuar desactivando un elemento de control negativo, incluyendo un elemento regulador de transcripción que actúa en cis o trans, o sometiendo a mutagénesis un elemento de control positivo. Uno o más, o todos los transcritos de una célula pueden ser medidos por hibridación de una muestra que comprende transcritos de la célula, o ácidos nucleicos representativos de o complementarios a transcritos de una célula, por hibridación con ácidos nucleicos inmovilizados sobre un arreglo.

30 Monitoreo de expresión de polipéptidos, péptidos y aminoácidos

En un aspecto de la invención, el fenotipo manipulado genéticamente comprende incrementar o reducir la expresión de un polipéptido (por ejemplo, una amilasa) o generar nuevos polipéptidos en una célula. Esta expresión incrementada o reducida puede ser rastreada al determinar la cantidad de amilasa presente o por ensayos de actividad de amilasa. Los polipéptidos, péptidos y aminoácidos también pueden ser detectados y cuantificados por cualquier método conocido en el arte, incluyendo, por ejemplo, resonancia magnética nuclear (RMN), espectrofotometría, radiografía (marcación radioactiva de proteína), electroforesis, electroforesis capilar, cromatografía líquida de alto rendimiento (HPLC), cromatografía de capa fina (TLC), cromatografía de hiperdifusión, diferentes métodos inmunológicos, por ejemplo, inmunoprecipitación, inmunodifusión, inmunoelectroforesis, radioinmunoensayos (RIA), ensayos por inmunoabsorción ligado a enzimas (ELISA), ensayos inmunofluorescencia, electroforesis en gel (por ejemplo, SDS-PAGE), tinción con anticuerpos, clasificador de células activado por fluorescencia (FACS), espectrometría de masas por pirólisis, espectrometría infrarroja por transformadas de Fourier, espectrometría Raman, GC-MS y espectrometrías de masas por LC-electroaspersión y por cap-LC-electroaspersión en tándem y similares. Se pueden cribar nuevas bioactividades utilizando métodos, o variaciones de los mismos, descritos en la patente de los Estados Unidos No. 6.057.103. Además, como se discute más adelante en detalle, uno o más, o todos los polipéptidos de una célula pueden ser medidos utilizando un arreglo de proteína.

Aplicaciones industriales

La invención provee muchos usos industriales y aplicaciones médicas para las amilasas de la invención. Por ejemplo, la invención provee enzimas y métodos para licuar el polisacárido, por ejemplo, almidón. Muchas amilasas y/o glucoamilasas y glucosidasas usadas en los procesos para convertir polisacárido licuado, por ejemplo, almidón, en glucosa son incapaces de hidrolizar enlaces alfa (1,6), y esta deficiencia deja aproximadamente 5% del azúcar como panosa e isomaltosa. Sin embargo, en un aspecto, las enzimas de la invención pueden convertir un polisacárido, por ejemplo, almidón, en glucosa para maximizar la producción de glucosa, incluyendo la conversión de polisacárido licuado, por ejemplo, almidón, en glucosa. En un aspecto, la invención provee enzimas y métodos para hidrolizar enlaces 1,4-alfa y/o 1,6- alfa (por ejemplo, hidrólisis de almidones), e hidrólisis de panosa e isomaltosa.

Las glucosidasas de la invención se pueden utilizar en una variedad de procesos industriales, incluyendo la conversión de biomasa en combustibles (por ejemplo, los biocombustibles, tales como el bioetanol, biopropanol, biobutanol, o un biodiesel) y similares), por ejemplo, incluyendo su uso en la inicial etapas (licuefacción) del polisacárido, por ejemplo, almidón, de procesamiento, en la molienda húmeda de maíz, en la producción de alcohol, en la industria textil para polisacárido, por ejemplo, almidón, el desencolado, en aplicaciones de panadería, en la industria de bebidas, en los campos petrolíferos en los procesos de perforación; en el entintado de papel reciclado; y en pienso para animales. Por lo tanto, la invención también provee un combustible, por ejemplo, un biocombustible (tal como un bioetanol, biopropanol, biobutanol, o un biodiesel), que comprende un polipéptido de la invención

Composiciones detergentes

La invención provee composiciones detergentes que comprenden uno o más polipéptidos de la invención, por ejemplo, amilasas de la invención, tal como alfa amilasas, etc., y métodos de elaboración y utilización de estas composiciones. La invención incorpora todos los métodos de elaboración y utilización de composiciones detergentes, véase, por ejemplo, las patentes de los Estados Unidos Nos. 6.413.928, 6.399.561, 6.365.561, 6.380.147. Las composiciones detergentes pueden ser una composición acuosa de una y dos partes, una composición líquida no acuosa, un sólido moldeado, una forma granulada, una forma en partículas, una tableta comprimida, un gel y/o una suspensión y una forma en suspensión. La invención también provee métodos capaces de remoción rápida de manchas gruesas de alimentos, películas de residuos de alimentos y otras composiciones alimenticias menores que usan estas composiciones detergentes. Las Amilasas de la invención pueden facilitar la remoción de manchas que comprenden polisacáridos, por ejemplo, manchas de almidón por medio de hidrólisis catalítica de un polisacárido y/o oligosacárido, por ejemplo, almidón. Las glucoamilasas de la invención se pueden usar en detergentes para el lavado de vajilla o en detergentes para el lavado de prendas textiles.

El contenido enzimático activo real depende del método de fabricación de una composición detergente y no es crítico, asumiendo que la solución de detergente tiene la actividad enzimática deseada. En un aspecto, la cantidad de amilasa presente en la solución final está en el intervalo de aproximadamente 0,001 mg hasta 0,5 mg por gramo de la composición detergente. La enzima particular elegida para uso en el proceso y los productos de esta invención depende de las condiciones de utilidad final, que incluyen la forma física del producto, del uso del pH, del uso de la temperatura y de los tipos de mugre que van a ser degradados o alterados. La enzima se puede elegir para proveer actividad y estabilidad óptimas para un determinado conjunto de condiciones de utilidad. En un aspecto, los polipéptidos de la presente invención son activos en los rangos de pH de aproximadamente 4 hasta aproximadamente 12 y en el rango de temperatura de aproximadamente 20 °C hasta aproximadamente 95 °C. Los detergentes de la invención pueden comprender agentes tensoactivos catiónicos, no iónicos semipolares o zwitteriónicos; o mezclas de los mismos.

Las amilasas de la presente invención pueden ser formuladas en detergentes en polvo o líquidos que tienen pH entre 4,0 y 12,0 a niveles de aproximadamente 0,01 hasta aproximadamente 5% (preferiblemente 0,1% hasta 0,5%) en peso. Estas composiciones detergentes también pueden incluir otras enzimas tales como las proteasas, celulasas, lipasas o endoglicosidasas, así como también conformadores y estabilizadores conocidos. La adición de amilasas de la invención a las composiciones de limpieza convencionales no crea ninguna limitación de uso especial. En otras palabras, cualquier temperatura y pH adecuado para el detergente también es adecuado para las presentes composiciones siempre que el pH se encuentre dentro del rango anteriormente mencionado. y la temperatura esté por debajo de la temperatura de desnaturalización de la enzima descrita. Asimismo, los polipéptidos de la invención se pueden usar en una composición de limpieza sin detergentes, nuevamente, ya sea solos o en combinación con conformadores y estabilizadores.

La presente invención provee composiciones de limpieza que incluyen composiciones detergentes para limpiar superficies duras, composiciones detergentes para limpiar textiles, composiciones para el lavado de vajillas, composiciones de limpieza bucal, composiciones de limpieza de prótesis dentales y soluciones de limpieza de lentes de contacto.

En un aspecto, la invención provee un método para lavar un objeto que comprende poner en contacto el objeto con un polipéptido de la invención bajo condiciones suficientes para el lavado. En un aspecto, se usa un polipéptido de la invención (por ejemplo, una amilasa activa en medio alcalino) en un detergente, es decir, como un aditivo detergente. La composición detergente de la invención puede, por ejemplo, ser formulada como una composición detergente para lavado mecánico o manual que comprende un polipéptido de la invención. Las composiciones detergentes de la invención incluyen soluciones de lavado de prendas textiles y de vajillas (por ejemplo, lavado automático de vajillas) y aplicación. Un aditivo para lavado de prendas adecuado para pretratamiento de textiles manchados puede comprender un polipéptido de la invención. Una composición suavizante de textiles puede comprender un polipéptido de la invención. En forma alternativa, un polipéptido de la invención puede ser formulado como una composición detergente para uso en operaciones generales de limpieza de superficies duras domésticas. En aspectos alternativos, los aditivos detergentes y composiciones detergentes de la invención pueden comprender una o más enzimas tales como una proteasa, una lipasa, una cutinasa, otra amilasa, y/o glucoamilasa, una carbohidrasa, una celulasa, una pectinasa, una mananasa, una arabinasa, una galactanasa, una xilanasa, una

oxidasa, por ejemplo, una lactasa, y/o una peroxidasa. Las propiedades de la(s) enzima(s) de la invención se eligen para ser compatibles con el detergente seleccionado (es decir de pH óptimo, compatible con otros ingredientes enzimáticos y no enzimáticos, etc.) y la(s) enzima(s) está(n) presente(s) en cantidades efectivas. En un aspecto, las enzimas amilasas de la invención se utilizan para remover materiales malolientes de los textiles. Varias composiciones detergentes y métodos para su elaboración que pueden ser usados en la realización de la invención se describen en, por ejemplo, las patentes de los Estados Unidos Nos. 6.333.301; 6.329.333; 6.326.341; 6.297.038; 6.309.871; 6.204.232; 6.197.070; 5.856.164.

Tratamiento de textiles

La invención provee métodos para tratar textiles utilizando uno o más polipéptidos de la invención. Los polipéptidos de la invención se pueden usar en cualquier método para tratamiento de textiles, que son conocidos en el arte, véase, por ejemplo, la patente de los Estados Unidos No. 6.077.316. Por ejemplo, en un aspecto, la sensación y apariencia de un textil se mejora por un método que comprende poner en contacto el textil con una amilasa de la invención en una solución. En un aspecto, el textil es tratado con la solución bajo presión.

En un aspecto, las enzimas de la invención se aplican durante o después del tejido de los textiles, o durante la etapa de desencolado, o una o más etapas adicionales de procesamiento del textil. Durante el tejido de los textiles, los hilos se exponen a tensión mecánica considerable. Antes de realizar el tejido sobre telares mecánicos, los hilos trenzados a menudo están recubiertos con almidón para encolado o derivados de almidón a fin de incrementar su resistencia a la tracción y evitar la ruptura. Las enzimas de la invención pueden ser aplicadas para remover este almidón para encolado o los derivados de almidón. Una vez que los textiles han sido tejidos, el textil puede continuar a la etapa de desencolado. Esto puede estar seguido de uno o más etapas adicionales de procesamiento del textil. El desencolado es el acto de remover el encolado del textil. Después del tejido se debe remover el revestimiento de encolado antes del procesamiento del textil a fin de asegurar un resultado homogéneo y a prueba de agua. La invención provee un método de desencolado que comprende la hidrólisis enzimática del encolado mediante la acción de una enzima de la invención.

Las enzimas de la invención se pueden usar para desencolar los textil, incluyendo los textil que contienen algodón, como aditivos detergentes, por ejemplo, en composiciones acuosas. La invención provee métodos para producir una apariencia de lavado a la piedra sobre textiles y prendas de vestir de mezclilla teñidas con índigo. Para la fabricación de vestidos, el textil puede ser cortado y cosido en vestidos o prendas de vestir, que posteriormente se terminan. En particular, para la fabricación de pantalones vaqueros, se han desarrollado diferentes métodos de acabado enzimático. El acabado de la prenda de mezclilla normalmente se inicia con una etapa de desencolado enzimático, durante el cual las prendas se someten a la acción de enzimas amilolíticas a fin de proveer suavidad al textil y volver al algodón más accesible a las etapas posteriores de terminado enzimático. La invención provee métodos para el terminado de prendas de mezclilla (por ejemplo, un "proceso de biolavado a la piedra"), desencolado enzimático y para proveer suavidad a los textiles usando las amilasas de la invención. La invención provee métodos para suavizar rápidamente prendas de mezclilla en un proceso de desencolado y/o terminado.

Alimentos y procesamiento de alimentos: licuefacción de polisacáridos, por ejemplo, almidón

Las enzimas de la invención tienen numerosas aplicaciones en la industria del procesamiento de alimentos. Las amilasas de la invención se usan para el procesamiento de almidón hasta fructosa. En un aspecto alternativo, los procesos de la invención comprenden un procesamiento de un polisacárido, por ejemplo, almidón, hasta fructosa que comprende cuatro etapas: licuefacción de almidón granulado, sacarificación del polisacárido licuado, por ejemplo, almidón, en dextrosa, purificación e isomerización hasta fructosa; y una, varias o todas estas etapas pueden comprender el uso de una o más enzimas de la invención. Las enzimas de la invención se pueden usar en la producción de etanol a partir de una biomasa, por ejemplo, un maíz o un pasto, por molienda en seco o en húmedo; por ejemplo, la Figura 12 ilustra cómo las enzimas de la invención se pueden usar en la producción de etanol a partir del maíz por molienda en seco, incluyendo su uso tanto en "procesos convencionales" como en "procesos simultáneos de sacarificación y fermentación por licuefacción".

La invención provee métodos de licuefacción de un polisacárido, por ejemplo, almidón, utilizando las enzimas de la invención. Las suspensiones concentradas de gránulos poliméricos de almidón se convierten en una solución de dextrinas solubles de longitud de cadena más corta de baja viscosidad. Esta etapa es útil para el manejo conveniente con equipo estándar y para una conversión eficiente hasta glucosa u otros azúcares 10^3 . En un aspecto, el almidón granulado es licuado mediante gelatinización de los gránulos elevando la temperatura del almidón granulado aproximadamente por encima de 72°C . El proceso de calentamiento rompe instantánea los gránulos de almidón insolubles para producir una solución de almidón soluble en agua. La solución de almidón solubilizada puede ser luego licuada mediante una amilasa y/o una glucoamilasa de la invención. Por lo tanto, la invención provee procesos de licuefacción enzimática del almidón utilizando una amilasa de la invención.

La Figura 7, la Figura 8 y la Figura 9 ilustran ejemplos alternativos de procesos del almidón, incluyendo procesos de licuefacción del almidón, de la invención utilizando al menos una enzima de la invención. Por ejemplo, la Figura 7

ilustra un ejemplo de un proceso de la invención de licuefacción del almidón que comprende el tratamiento de una suspensión de almidón (por ejemplo, que tiene aproximadamente 30% a 35% de sólidos) con vapor para licuefacción primaria (por ejemplo, aproximadamente a 105 °C durante aproximadamente 5 minutos), entrada en un tanque de expansión, seguido por una licuefacción secundaria (por ejemplo, hasta aproximadamente entre 90°C y 95°C durante aproximadamente 90 minutos), involucrando cada una de estas etapas el uso de una enzima de la invención. La Figura 8 ilustra otro ejemplo de un proceso de licuefacción de almidón de la invención que comprende tratar una suspensión de almidón aproximadamente entre pH 4 a pH 5, por ejemplo, pH 4,5, ajustando el pH, la adición de calcio, licuefacción aproximadamente a un pH 5 a pH 6, por ejemplo, pH 5,4, a aproximadamente 95°C usando una alfa amilasa de la invención, seguido por otro ajuste de pH y temperatura para sacarificación aproximadamente entre pH 4 y pH 5, por ejemplo, pH 4,5, a una temperatura entre aproximadamente 60°C y 65°C usando una alfa amilasa de la invención. La Figura 9 ilustra otro ejemplo de un proceso del almidón de la invención que comprende tratar una suspensión de almidón aproximadamente entre pH 4 y pH 5, por ejemplo, pH 4,5, (ajustando opcionalmente el pH, y la adición de calcio), combinando la licuefacción y sacarificación usando una amilasa de la invención aproximadamente entre pH 4 y pH 5, por ejemplo, pH 4,5, a una temperatura aproximadamente superior a 90°C, o, aproximadamente superior a 95°C, seguido por otro ajuste de pH y temperatura para sacarificación hasta aproximadamente entre pH 4 y pH 5, por ejemplo, pH 4,5, a una temperatura aproximadamente entre 60°C y 65°C usando una glucoamilasa de la invención. En un aspecto, la licuefacción-sacarificación combinadas de la invención son un proceso en un único recipiente. En un aspecto, el proceso completo es un proceso en un único recipiente. Cualquiera de estos procesos y cualquiera de estas etapas, puede comprender también, o puede comprender en forma adicional, otra enzima de la invención (por ejemplo, una glucosidasa tal como una α -1,6-glucosidasa, una maltasa, etc.), u otra enzima tal como una pululanasa o una isomerasa.

Un ejemplo de un proceso de licuefacción enzimático implica el ajuste el pH del de una suspensión de almidón granulado entre 6,0 y 6,5 y la adición de hidróxido de calcio, hidróxido de sodio o carbonato de sodio. En un aspecto, se agrega hidróxido de calcio. Esto provee iones de calcio para estabilizar la glucoamilasa de la invención contra la inactivación. En un aspecto, después de la adición de amilasa, se bombea la suspensión a través de un chorro de vapor para elevar instantáneamente la temperatura entre 80 - 115°C. En un aspecto, se gelatiniza inmediatamente el almidón y, debido a la presencia de amilasa, se despolimeriza a través de hidrólisis aleatoria de enlaces α -1,4-glucosídicos por medio de amilasa hasta una masa fluida. La masa fluida puede ser bombeada fácilmente. La invención provee diferentes procesos de licuefacción enzimática de polisacáridos, oligosacáridos y/o almidón usando una amilasa de la invención. En un aspecto del proceso de licuefacción de la invención, se añade una amilasa al polisacárido, oligosacárido y/o suspensión de almidón y la suspensión se mantiene a una temperatura de entre aproximadamente 80° -100 °C para hidrolizar parcialmente el polisacárido, oligosacárido y/o gránulos de almidón. En un aspecto, se bombea la suspensión parcialmente hidrolizada de polisacárido, oligosacárido y/o almidón a través de un chorro a temperatura aproximadamente por encima de 105°C hasta gelatinizar completamente cualquier estructura granulada remanente. En un aspecto, después de enfriar el almidón gelatinizado, se realiza una segunda adición de amilasa para hidrolizar adicionalmente el polisacárido, oligosacárido y/o almidón.

La invención provee enzimas y procesos para hidrolizar polisacáridos, oligosacáridos y/o almidón líquido (licuado) y granulado. Dicho almidón se puede derivar de cualquier fuente, por ejemplo, maíz, trigo, milo, sorgo, centeno o bulgher. La invención se aplica a cualquier fuente de almidón en granos que sea útil para la licuefacción, por ejemplo, cualquier otra fuente de granos o vegetal conocida para producir almidón adecuado para licuefacción. Los métodos de la invención comprenden la licuefacción del almidón a partir de cualquier material natural, tal como arroz, arroz germinado, maíz, cebada, milo, trigo, legumbres y batata. El proceso de licuefacción puede hidrolizar sustancialmente el almidón para producir un jarabe. El rango de temperatura de la licuefacción puede ser cualquier temperatura de licuefacción que se sepa que es efectivo en la licuefacción del almidón. Por ejemplo, la temperatura del polisacárido, oligosacárido y/o almidón puede estar aproximadamente entre 80°C y aproximadamente 115°C, entre aproximadamente 100°C y aproximadamente 110 °C entre aproximadamente 105°C y aproximadamente 108°C. En aspectos alternativos, la amilasa usada en estos métodos es activa a estas temperaturas, por ejemplo, activa a temperaturas en un rango de entre aproximadamente 80°C hasta aproximadamente 115°C, entre aproximadamente 100°C hasta aproximadamente 110°C y desde aproximadamente 105°C hasta aproximadamente 108°C.

La invención provee métodos para licuefacción y sacarificación como se ilustra en la Figura 5. En un aspecto, las alfa amilasas de la invención se usan en la etapa de licuefacción ilustrada (algunos de los métodos industriales actuales usan α -amilasa de *B. licheniformis*). En un aspecto, el proceso tiene lugar a un pH aproximadamente de 6,0 a una temperatura en cualquier parte en el rango de entre aproximadamente 95°C hasta 105°C, por un período de tiempo entre aproximadamente 0,5 y 5 horas, por ejemplo, 60, 90 o 120 minutos. En un aspecto, en un proceso de maíz fermentado, antes de la licuefacción, se añaden celulasas, proteasas y/o proteína tioreductasas.

En un aspecto de un proceso de licuefacción de la invención, se agrega una amilasa de la invención que tiene actividad a pH de aproximadamente 4,5 (o entre aproximadamente pH 4,5 y pH 5), que puede o no ser dependiente de Ca^{2+} . La eliminación de la adición de sales al comienzo del proceso elimina la necesidad de removerlas al final

del proceso. En un aspecto de un proceso de licuación de la invención, se utiliza una amilasa que es más activa. Esto puede permitir que uno disminuya la cantidad de enzima necesaria. En un aspecto, se realizan la licuefacción y sacarificación en el mismo recipiente, como un "proceso en un solo recipiente", por ejemplo, bajo condiciones que comprenden aproximadamente 90°C hasta 95°C (o en cualquier punto entre aproximadamente 80°C y 105°C), como
 5 aproximadamente un proceso de 3 horas (o como un proceso que dura entre aproximadamente 1 y 5 horas). En este aspecto, la carga de enzima se puede cortar nuevamente a la mitad.

En un aspecto de un proceso de sacarificación de la invención, se usa una amilasa de la invención. En un aspecto, las amilasas de la invención se usan en un etapa de sacarificación (además o en lugar de una glucoamilasa de *A. niger*). En un aspecto, el proceso tiene lugar aproximadamente un pH 4,5, en un rango de temperatura de entre
 10 aproximadamente 60°C y 62°C (o en cualquier punto en el rango entre aproximadamente 50 °C y 72°C, o, entre aproximadamente 40°C y 80°C) como un proceso que dura aproximadamente entre 12 y 96 o más horas. En un aspecto, se agregan otras enzimas, por ejemplo, pululanastas para incrementar la cantidad de glucosa.

En un aspecto, una, algunas o todas las enzimas usadas en los procesos de la invención (incluyendo las enzimas de la invención) se inmovilizan, por ejemplo, se inmovilizan sobre cualquier superficie, por ejemplo, una superficie plana
 15 o una columna de enzimas, por ejemplo, inmovilizadas sobre un arreglo, una perla, fibra, poro, capilar y similares. En un aspecto, al ser inmovilizadas, pueden ser reutilizadas.

En un aspecto, la invención provee "cócteles de enzimas" usando al menos una enzima de la invención. En un aspecto, los "cócteles de enzimas" se usan en los procesos de la invención, por ejemplo, incluyendo los métodos de licuefacción y sacarificación como se ilustran en la Figura 5. Por ejemplo, en un aspecto, se usan enzimas que
 20 degradan la pared celular (CWDE), por ejemplo, para textiles, pulpa y papel y procesos de lavado de prendas de la invención, incluyendo, por ejemplo, combinaciones de celulasas, hemicelulasas, xilanasas, galactomananasas, glucomanasas, arabinofuranosidasas y otras. En un aspecto, los "cócteles de enzimas" usados en los procesos de la invención para bioblanqueamiento (por ejemplo, pulpa y papel, proceso de lavado de prendas), incluye combinaciones de lacasas, peroxidases, oxidasas y similares. En un aspecto, las enzimas que degradan la pared
 25 celular se combinan con enzimas de bio blanqueamiento y enzimas de la invención para degradar las paredes de las células vegetales para liberar agentes de color.

Otro ejemplo de un cóctel de enzimas de la invención que comprende al menos una amilasa de la invención puede hidrolizar aproximadamente más del 95% del almidón de maíz molido en azúcares fermentables en no más de 60
 30 horas a aproximadamente 30 a 40°C y a aproximadamente pH 3,5 hasta pH 5,5 en presencia de levaduras. En un aspecto, la cantidad total de proteína enzimática requerida no supera los 50 gramos/tonelada de maíz (0,005% p/p).

Las enzimas de la invención se pueden usar para tratar productos lácteos; y la invención provee productos lácteos que comprenden un polipéptido de la invención, en donde opcionalmente el producto lácteo comprende, una leche, un helado, un queso o un yogur.

La enzimas de la invención se pueden usar en procesos simultáneos de licuefacción, sacarificación y fermentación
 35 (SLSF); y las ventajas de usar las enzimas de esta invención pueden incluir:

- Menores costes de energía debido a la eliminación de la etapa a alta temperatura;
 - Rendimiento mejorado debido a una fermentación más rápida;
 - Riesgo reducido de contaminación bacteriana;
 - La levadura produce etanol más temprano que en un proceso convencional;
 - Almacenamiento y manejo reducido de la levadura a granel.
- 40

Las propiedades de las enzimas de la invención usadas en estos procesos de la invención pueden incluir: actividad para hidrolizar polisacárido, oligosacárido y/o almidón sin purificar; perfiles de actividad de temperatura y pH compatibles con las condiciones de proceso; alta actividad sobre polisacárido, oligosacárido y/o almidón no gelatinizado; y/o activación sobre la fracción "resistente" del almidón sin purificar.

45 En aspectos alternativos, las enzimas de la invención tienen actividad de amilasa y pueden hidrolizar (parcialmente o completamente) gránulos de polisacárido, oligosacárido y/o almidón que tienen estructuras complejas multinivel, incluyendo regiones amorfas y cristalinas y/o cadenas lineales y ramificadas, por ejemplo, pueden hidrolizar (parcial o completamente) el almidón sin purificar, incluyendo: RS I - almidón físicamente inaccesible; RS II - gránulos de almidón resistentes como en patata y banana sin purificar; RS III - retrogradado como en la patata cocida.

50 Procesos para producir jarabes de dextrosa de alto peso molecular

La invención provee procesos para producir jarabes de dextrosa de alto peso molecular, usando enzimas de la invención, que incluyen métodos para producir oligosacáridos que tiene grupos estrictamente de alto peso molecular de aproximadamente 20000. En un aspecto, las amilasas de la invención se pueden usar para licuar una

composición que comprende polisacárido, que comprende oligosacárido y/o que comprende almidón, por ejemplo, un almidón de maíz, para producir un patrón de oligosacárido que está estrictamente agrupado en aproximadamente 20000 de peso molecular (Bacillus amylases producirá jarabes que contienen fragmentos de peso molecular mucho mayores y los oligosacáridos de alto peso molecular no se convierten completamente en glucosa por las
5 glucoamilasas, por ejemplo, glucoamilasas de Aspergillus, durante la sacarificación).

En un aspecto, usando las amilasas de la invención para catalizar la hidrólisis de una composición que comprende almidón, por ejemplo, un almidón de maíz, se producen fragmentos de aproximadamente 20000 unidades de peso molecular. Estos fragmentos de aproximadamente 20000 unidades de peso molecular se pueden convertir rápida y completamente en glucosa. Por lo tanto, en un aspecto, los jarabes sacarificados resultantes de la licuefacción con
10 amilasas de *Bacillus* contienen menos dextrosa que los jarabes sacarificados provenientes de la licuefacción usando amilasas de la invención.

Procesos para producir maltodextrinas homogéneas

La divulgación provee procesos para producir maltodextrinas homogéneas usando las enzimas de la invención. Las maltodextrinas homogéneas producidas por los métodos de la divulgación se pueden usar en una gran variedad de
15 aplicaciones como alimentos, fármacos y revestimientos. En un aspecto, las amilasas de la invención se pueden usar para generar una composición de maltodextrina extremadamente uniforme (los procesos de fabricación convencionales que usan hidrólisis ácida o enzimática de almidón resultan en una distribución amplia de peso molecular típicamente bimodal de los oligosacáridos). Las maltodextrinas homogéneas producidas por los métodos de la divulgación tienen una distribución homogénea de peso molecular y se pueden usar en una variedad de
20 productos que comprenden maltodextrina, resultando en soluciones claras (sin turbidez), de menor viscosidad mejores propiedades de revestimiento, mejores propiedades de formación de película y similares.

En un aspecto, las amilasas de la invención se usan para licuar almidón de maíz para producir una composición uniforme que comprende maltodextrina. En un aspecto, la licuefacción se realiza a un pH de aproximadamente 4,5 hasta aproximadamente pH 6,5, por ejemplo, pH 5,0 o 5,5, a temperaturas hasta aproximadamente 105°C. La
25 composición uniforme de maltodextrina puede ser producida en el intervalo de DE de aproximadamente 5 hasta tan alto como aproximadamente 20. Los jarabes producidos por estas amilasas de la invención pueden ser filtrados, tratados con carbón vegetal y/o secados por aspersión para obtener el producto que comprende maltodextrina.

En un aspecto, una o más de otras enzimas son utilizadas junto con una composición que comprende las enzimas de la invención para uso en procesamiento de almidón hasta fructosa, licuefacción de almidón granulado y procesos
30 para producir maltodextrinas homogéneas o jarabes de dextrosa de alto peso molecular, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglicosidasas, endo-beta-1,4-lacasas, amiloglicosidasas, otrs glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipoxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidadasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidadasas, descarboxilasas, fenoloxidasas, ligninasas, pululanadas, arabinanasas, hemicelulasas, mananasas,
35 xilolacasas, xilanasas. Pectín acetil esteradasas, ramnogalacturonano acetil esteradasas, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesteradasas. celobiohidrolasas y/o transglutaminasas.

Procesos enzimáticos de molienda en seco

La invención provee procesos enzimáticos de molienda en seco usando una amilasa de la invención; los ejemplos de procesos se ilustran en la Figura 12. En la molienda en seco, se muele el grano entero y se combina con agua. Se remueve opcionalmente el germen mediante separación por flotación o técnicas equivalentes. La mezcla resultante, que contiene polisacárido, por ejemplo, almidón, fibra, proteína y otros componentes del grano, se licúa usando amilasa. En un aspecto, se realiza la licuefacción enzimática a temperaturas menores que el polisacárido,
45 por ejemplo, almidón de los procesos de licuefacción discutidos anteriormente. En un aspecto, después de la gelatinización del polisacárido, por ejemplo, almidón, se mantiene la solución a una temperatura elevada en presencia de amilasa hasta que se alcanza un DE de 10 - 20. En un aspecto, este es un período de aproximadamente 1- 3 horas. El equivalente de dextrosa (DE) es el estándar industrial para medir la concentración de los azúcares totales reductores, calculada como D-glucosa sobre peso en base seca. El almidón granulado no hidrolizado tiene un DE de virtualmente cero, mientras que el DE de D-glucosa se define como 100.

Las enzimas de la invención se pueden usar en procesos de molienda húmeda o seca de biomasa. Por ejemplo, en un aspecto de un ejemplo de un proceso de molienda en seco, se muele el grano entero y se combina con agua, y se lo trata con una enzima de la invención. El germen puede ser removido mediante separación por flotación o técnicas equivalentes. En un aspecto, la mezcla resultante, que contiene polisacárido, por ejemplo, almidón, fibra, proteína y otros componentes del grano, se licúa usando una amilasa de la invención (por ejemplo, una alfa-
55 amilasa). En un aspecto, esta licuefacción enzimática se realiza a una temperatura relativamente inferior cuando se usa el proceso de molienda en seco; sin embargo, se cree que la licuefacción a baja temperatura es menos eficiente que la licuefacción a alta temperatura en la conversión de polisacárido, por ejemplo, almidón, hasta dextrinas

solubles. Por lo tanto, en un aspecto, la invención provee una etapa adicional en donde, después de la gelatinización, la solución de polisacárido, por ejemplo, almidón, se mantiene a una temperatura elevada en presencia de una amilasa de la invención (por ejemplo, una alfa-amilasa), que en un aspecto se logra un DE de aproximadamente 10 a 20, usualmente un período de aproximadamente 1 a 3 horas (el equivalente de dextrosa (DE) es el estándar industrial para medir la concentración de los azúcares reductores totales, calculada como D-glucosa con base en el peso seco; el almidón granulado no hidrolizado tiene un DE de virtualmente cero, mientras que el DE de D-glucosa se define como 100).

En aspectos alternativos, el uso de amilasas de la invención en procesos para producción de etanol mediante molienda en seco pueden proveer ventajas operacionales, por ejemplo: reducción rápida de viscosidad de la harina de maíz en suspensión, logrando un incremento en sólidos disueltos y de rendimiento posible sin inversión adicional de capital; estabilidad térmica superior para competir mejor, lo que elimina la dosificación dividida; algunas amilasas de la invención son enzimas termoestables - y esto elimina la necesidad de dosificar antes y después de la cocción a chorro; se obtienen viscosidades menores con temperaturas de proceso superiores y permite un control microbiano mejorado en el tanque de la suspensión (el proceso se realiza a una temperatura superior, de manera tal que se matan los microbios indeseados); menor pH de licuefacción, que elimina la necesidad de ajuste del pH, disminuye la formación de escamas (el precipitado de oxalato de calcio se forma sobre los equipos, etc.; si la licuefacción se realiza a un pH bajo, existe una mayor probabilidad de formación de escamas) y se reduce la formación de subproductos.

En resumen, en aspectos alternativos, las amilasas de la invención pueden ser enzimas termoestables que pueden satisfacer las necesidades claves de la industria, por ejemplo, bajo ciertas condiciones, se reduce rápidamente la viscosidad de la suspensión de harina de maíz de sólidos muy secos, pueden ser termoestables (temperatura óptima de 95°C), pueden ser independientes del calcio, pueden ser activas a bajo pH óptimo, y pueden tolerar hasta un 30% de reciclados en contracorriente. En un aspecto, la dosis recomendada está en el rango entre aproximadamente 0,4 y 0,6 kg/ MT de almidón.

25 Procesos enzimáticos de molienda en húmedo

La divulgación provee procesos de molienda en húmedo, por ejemplo, molienda en húmedo de maíz, usando una enzima, por ejemplo, una amilasa, de la invención. La molienda en húmedo de maíz es un proceso que produce aceite de maíz, harina de gluten, pienso de gluten y polisacárido, por ejemplo, almidón. Por lo tanto, la divulgación provee métodos para elaborar aceite de maíz, harina de gluten, pienso de gluten y polisacárido, por ejemplo, almidón, usando una enzima de la invención. En un aspecto, se usa una amilasa alcalina de la invención en la licuefacción del polisacárido, por ejemplo, almidón. En un aspecto, se usa una amilasa de la invención en la sacarificación para producir glucosa. Un ejemplo de un proceso de molienda en húmedo de maíz de la invención (usando al menos una enzima de la invención), se ilustra en la Figura 6. La Figura 6 ilustra un ejemplo de un proceso de aceite de maíz de la invención que comprende remojar, desgerminar, desfibrar y separar el gluten, seguido por licuefacción usando una enzima de la invención (por ejemplo, una alfa amilasa) y sacarificación usando una enzima de la invención.

En un aspecto, se somete maíz (un grano que consiste de un revestimiento de semilla exterior (fibra), polisacárido, por ejemplo, almidón, una combinación de almidón y glucosa y el germen interior), a un proceso de cuatro etapas, que resulta en la producción de almidón. En un aspecto, se remoja, se desgermina, se desfibra y se separa el gluten el maíz. En un proceso de remojo, se retiran los solubles. El producto restante después de la extracción de los solubles se desgermina, lo que resulta en la producción de aceite de maíz y en producción de una torta oleosa, que se añade a los solubles de la etapa de remojo. El producto remanente se desfibra, y se añaden los sólidos de la fibra a la mezcla de torta oleosa/solubles. Esta mezcla de sólidos de fibra, torta oleosa y solubles forma un pienso de gluten. Después del desfibrado, se somete el resto del producto a la separación del gluten. Esta separación genera harina de gluten y almidón. Se somete luego el almidón a licuefacción y sacarificación usando polipéptidos de la invención para producir glucosa.

La Figura 6 ilustra un ejemplo de un proceso de molienda en húmedo de maíz de la invención (usando al menos una enzima de la invención). La Figura 7, la Figura 8 y la Figura 9 ilustran ejemplos alternativos de métodos de procesamiento del almidón (por ejemplo, procesos industriales), incluyendo procesos de licuefacción de almidón, de la invención (usando al menos una enzima de la invención).

Las enzimas de la invención se pueden usar en procesos de molienda en seco/húmedo de la biomasa; por ejemplo, una molienda en húmedo de maíz, produce un aceite vegetal (por ejemplo, de maíz), una harina de gluten, un pienso de gluten y/o un almidón. En un aspecto, se usa una amilasa alcalina de la invención en la licuefacción del almidón y se usa una glucoamilasa en la sacarificación, produciendo glucosa. En un aspecto, se somete la biomasa (por ejemplo, un grano de maíz, que consiste de un revestimiento exterior de la semilla exterior (fibra), almidón, una combinación de almidón y glucosa y el germen interno) a un proceso de cuatro etapas para producir almidón. La biomasa (por ejemplo, maíz) se remoja, se desgermina, se desfibra y, finalmente, se separa el gluten. En el proceso de remojo, se retiran los solubles. El producto restante después de la remoción de los solubles se desgermina, lo

que resulta en la producción de un aceite vegetal (por ejemplo, un aceite de maíz) y la producción de una torta oleosa, que se añade a los solubles de la etapa de remojo. El producto restante se desfibra y se añaden los sólidos de la fibra a la mezcla de torta oleosa/solubles. Esta mezcla de sólidos de fibra, torta oleosa y solubles forma un pienso de gluten. Después del desfibrado, se somete el resto del producto a separación del gluten. Esta separación da como resultado una harina de gluten y almidón. En un aspecto, el almidón se somete luego a licuefacción y sacarificación (por ejemplo, usando enzimas de la invención) para producir glucosa.

La divulgación también provee un proceso de alto rendimiento para producir goma de fibra de maíz de alta calidad mediante el tratamiento de la fibra de maíz con una enzima de la invención seguido por tratamiento con peróxido de hidrógeno para obtener un extracto de fibra de maíz molida. Véase, por ejemplo, la patente de los Estados Unidos No. 6.147.206.

En un aspecto, se utilizan una o más de otras enzimas junto con una composición que comprende enzimas de la invención para uso con estos procesos de molienda en seco o en húmedo, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglicosidasas, endo-beta-1,4-lacasas, amiloglicosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululanasas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín acetil esterases, ramnogalacturonano acetil esterases, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesterases, celobiohidrolasas y/o transglutaminasas.

20 Procesos anti-enranciamiento

La divulgación provee procesos anti-enranciamiento (por ejemplo, de productos horneados tales como pan) usando una glucoamilasa de la invención. La divulgación provee métodos para disminuir el incremento de la firmeza de la miga (del producto horneado) y una disminución de la elasticidad de la miga usando una amilasa de la invención. El enranciamiento de los productos horneados (tales como pan) es más serio a medida que pasa el tiempo entre el momento de la preparación del producto horneado y el momento de consumo. El término "enranciamiento" se usa para describir cambios indeseables para al consumidor en las propiedades del producto horneado después de salir del horno, tal como un incremento de la firmeza de la miga, una reducción de la elasticidad de la miga y cambios en la corteza, que se torna dura y correosa. La firmeza de la miga del pan se incrementa de manera adicional durante el almacenamiento hasta un nivel que se considera negativo. Las amilasas de la invención se usan para retardar el enranciamiento del pan como se describe, por ejemplo, en las patentes de los Estados Unidos Nos. 6.197.352, 2.615.810, 3.026.205; Silberstein (1964) Baker's Digest 38: 66 - 72.

En un aspecto, una enzima de la invención se usa para retardar el enranciamiento de los productos horneados mientras que no se hidrolice el almidón en dextrinas ramificadas. Las dextrinas ramificadas se forman por escisión de las cadenas ramificadas de las dextrinas generadas por hidrólisis de α -amilasa que no puede ser degradada adicionalmente por la α -amilasa. Esto puede producir una miga gomosa en el pan resultante. Por consiguiente, la divulgación provee un proceso para retardar el enranciamiento de los productos horneados (por ejemplo, productos horneados con levadura) que comprende agregar una enzima de la invención que comprende actividad de exoamilasa a una harina o una masa usada para producir un producto horneado.

La divulgación también provee un proceso para preparar una masa o un producto horneado preparado a partir de una masa que comprende agregar una amilasa de la invención a la masa en una cantidad que sea efectiva para retardar el enranciamiento del pan. La divulgación también provee una masa que comprende una amilasa y una premezcla que comprende harina junto con dicha amilasa. Finalmente, la divulgación provee un aditivo de horneado enzimático, que contiene dicha amilasa.

Pienso y aditivos para animales

La invención provee piensos, alimentos, aditivos para alimentos, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios que comprenden un polipéptido de esta invención para humanos y animales; y la invención provee métodos para tratar humanos y animales que utilizan piensos, alimentos, aditivos para alimentos, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios que comprenden un polipéptido de esta invención; y/o el uso de una enzima amilasa de la invención. La invención provee alimentos para humanos y/o piensos para animales, alimentos, aditivos para alimentos, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios que comprenden amilasas de la invención. En un aspecto, el tratamiento de alimentos para humanos y/o piensos para animales, alimentos, aditivos para alimentos, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios usando las enzimas amilasas de la invención puede ayudar en la disponibilidad del polisacárido, por ejemplo, almidón, alimentos para humanos y/o piensos para animales, alimentos, aditivos para alimentos, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios. Esto puede resultar en la liberación de azúcares rápidamente digeribles y fácilmente absorbidos.

El uso de una amilasa de la invención puede incrementar la capacidad digestiva de aves y animales. El uso de una amilasa de la invención puede garantizar la disponibilidad de un suministro de nutrientes adecuado para mejor crecimiento y funcionamiento. En un aspecto, las enzimas de la invención pueden ser añadidas como aditivos para piensos, o en piensos, alimentos, aditivos para alimentos, aditivos para alimentos, suplementos nutricionales y/o suplementos dietarios, para animales. En otro aspecto, los piensos, alimentos, aditivos, aditivos para alimenticios, aditivos para piensos, suplementos nutricionales y/o suplementos dietarios, por ejemplo, para pienso animal, pueden ser tratados con una amilasa de la invención antes del consumo por el animal. En otro aspecto, se puede suministrar una amilasa de la invención mediante la expresión de las enzimas directamente en los cultivos de piensos transgénicos (tal como, por ejemplo, plantas, semillas transgénicas, y similares), tal como maíz. Como se discutió anteriormente, la divulgación provee plantas transgénicas, partes de plantas y células vegetales que comprenden una secuencia de ácido nucleico que codifica un polipéptido de la invención. En un aspecto, el ácido nucleico se expresa de manera tal que una amilasa de la invención sea producida en cantidades recuperables. La amilasa de la invención puede ser recuperada de cualquier planta o parte de la planta. En forma alternativa, la planta o parte de la planta que contiene el polipéptido recombinante se pueden usar como tal para mejorar la calidad de un alimento o pienso, aditivo para pienso, suplemento nutricional y/o suplemento dietario, y similares, por ejemplo, mejorando el valor nutricional, el sabor y las propiedades reológicas, o para destruir un factor antinutritivo.

Tratamiento del papel o de la pulpa

Las enzimas de la invención pueden ser usadas en el tratamiento del papel o de la pulpa o para desentintar papel. Por ejemplo, en un aspecto, la invención provee un proceso de tratamiento de papel que utiliza amilasas de la invención. En un aspecto, la enzima de la invención se puede usar para modificar el polisacárido, por ejemplo, almidón, en el papel, convirtiéndolo así en una forma licuada. En otro aspecto, se pueden usar los componentes del papel del papel fotocopiado reciclado durante un proceso químico y de desentintado enzimático. En un aspecto, las amilasas de la invención se pueden usar en combinación con celulasas. El papel puede ser tratado por medio de los siguientes tres procesos: 1) desintegración en presencia de una enzima de la invención, 2) desintegración con un compuesto químico para desentintado y una enzima de la invención, y/o 3) desintegración después del remojo con una enzima de la invención. El papel reciclado tratado con amilasa puede tener un mayor brillo debido a la remoción de las partículas de tóner comparado con el papel tratado solo con celulasa. Si bien la invención no está limitada por ningún mecanismo particular, el efecto de una amilasa de la invención puede ser debido a su comportamiento como agente activo de superficie en la suspensión de la pulpa.

La invención provee métodos para tratamiento del papel y de la pulpa de papel usando uno o más polipeptidos de la invención. Los polipéptidos de la invención se pueden usar en cualquier método de tratamiento del papel o de la pulpa, que son bien conocidos en el arte, véanse, por ejemplo, las patentes de los Estados Unidos No. 6.241.849, 6.066.233, 5.582.681. Por ejemplo, en un aspecto, la invención provee un método para desentintar y decolorar un papel impreso que contiene una tinta, que comprende formar una pulpa con un papel impreso para obtener una suspensión de pulpa y quitar una tinta de la suspensión de la pulpa en presencia de una enzima de la invención (también se pueden agregar otras enzimas). En otro aspecto, la invención provee un método para mejorar la permeabilidad de la pulpa, por ejemplo, pulpa hecha de fibra secundaria, al agregar una mezcla enzimática que comprende una enzima de la invención (también puede incluir otras enzimas, por ejemplo, enzimas pectinasas) a la pulpa y tratar bajo condiciones que causen una reacción para producir una pulpa tratada enzimáticamente. La permeabilidad de la pulpa tratada enzimáticamente se incrementa a partir de la permeabilidad inicial de la pulpa de fibra secundaria sin pérdida de brillo.

En un aspecto, una o más de otras enzimas se usan junto con una composición que comprende las enzimas de la invención para uso con estos métodos de tratamiento de pulpa y/o papel, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululaninas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín, acetil esterinas, ramnogalacturonano acetil esterinas, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesterinas, celobiohidrolasas y/o transglutaminasas.

Redespulpado: tratamiento de los materiales lignocelulósicos

La invención también provee un método para el tratamiento de fibras lignocelulósicas, en donde las fibras son tratadas con un polipéptido de la invención, en una cantidad que es eficiente para mejorar las propiedades de las fibras; Las amilasas de la invención también se pueden usar en la producción de materiales lignocelulósicos tales como pulpa, papel y cartón, a partir del polisacárido, por ejemplo, almidón, papel y cartón residuales reforzados, especialmente cuando ocurre redespulpado a un pH por encima de 7, y y donde las amilasas pueden facilitar la desintegración del material de desecho a través de la degradación del polisacárido de refuerzo, por ejemplo, almidón. Las amilasas de la invención pueden ser útiles en un proceso para producir una pulpa para fabricación de

papel a partir del papel impreso revestido de almidón. El proceso puede ser realizado como se describe, por ejemplo, en el documento WO 95/14807.

- 5 Un ejemplo de un proceso comprende desintegrar el papel para producir una pulpa, mediante tratamiento con una enzima que degrada polisacárido, por ejemplo, que degrada almidón, antes, durante o después de desintegrar y separar las partículas de tinta de la pulpa después de la desintegración, y tratamiento con enzima. Véase también la patente de los Estados Unidos No. 6.309.871 y otras patentes de los Estados Unidos citadas aquí. Por lo tanto, la invención incluye un método para el desentintado enzimático de pulpa de papel reciclado, en donde el polipéptido se aplica en una cantidad que es eficiente para el desentintado efectivo de la superficie de la fibra.

Tratamiento de desechos

- 10 Las enzimas de la invención se pueden usar en una variedad de otras aplicaciones industriales, por ejemplo, en tratamiento de desechos. Por ejemplo, en un aspecto, la invención provee un proceso de digestión de desecho sólido que usa las enzimas de la invención. Los métodos pueden comprender reducir la masa y el volumen del desecho sólido sustancialmente no tratado. El desecho sólido se puede tratar con un proceso digestivo enzimático en presencia de una solución enzimática (incluyendo una enzima de la invención) a una temperatura controlada.
- 15 Esto traduce en una reacción sin fermentación bacteriana apreciable a partir de los microorganismos añadidos. El desecho sólido es convertido en un desecho licuado y cualquier desecho sólido residual. El desecho licuado resultante puede ser separado de dicho desecho solidificado residual. Véase por ejemplo, la patente de los Estados Unidos No. 5.709.796.

- 20 En un aspecto, una o más de otras enzimas se usan junto con una composición que comprende las enzimas de la invención para uso con estos métodos de tratamiento de desechos, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa, isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululanasas, arabinanasas, hemicelulasas, mananasas, xilolacasas,
- 25 xilanasas, pectín acetil estereras, ramnogalacturonano acetil estereras, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesterasas, celobiohidrolasas y/o transglutaminasas.

Productos para el cuidado oral

- 30 La invención provee productos para el cuidado oral que comprenden una amilasa de la invención. Los ejemplos de productos para el cuidado oral incluyen pastas dentales, cremas dentales, geles o polvos dentales, lavados bucales y dentales, formulaciones para enjuague pre y post cepillado, gomas de mascar, pastillas o caramelos. Véase, por ejemplo, la patente de los Estados Unidos No. 6.264.925.

- 35 En un aspecto, una o más de otras enzimas se usan junto con una composición para el cuidado oral que comprende enzimas de la invención y para uso con estos métodos, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, other otras glucosidasas, glucosa somerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululanasas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín acetil estereras, ramnogalacturonano acetil estereras, proteasas, peptidasas, proteinasas,
- 40 poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesterasas, celobiohidrolasas y/o transglutaminasas.

Elaboración de cerveza y fermentaciones

- 45 La invención provee métodos de elaboración de cerveza (por ejemplo, fermentación) que comprenden una amilasa de la invención. En un ejemplo de un proceso, los materiales sin purificar que contienen almidón se desintegran y se procesan para formar una malta. Una amilasa de la invención se puede usar en cualquier punto del proceso de fermentación. Por ejemplo, las amilasas de la invención se pueden usar en el procesamiento de malta de cebada. El principal material sin purificar de la elaboración de cerveza es la malta de cebada. Este puede ser un proceso de tres etapas. Primero, los granos de cebada pueden ser remojados para incrementar el contenido de agua, por ejemplo, hasta aproximadamente el 40%. Segundo, el grano puede ser germinado por incubación a 15 - 25 °C durante 3 a 6 días cuando la síntesis enzimática es estimulada bajo el control de giberelinas. Durante este tiempo, los niveles de amilasa se elevan en forma significativa. En un aspecto, las amilasas de la invención se añaden en esta etapa del proceso (o en cualquier otro). La acción de la amilasa da como resultado un incremento en los azúcares reductores fermentables. Esto se puede expresar como la potencia diastática, DP, que se puede elevar desde alrededor de 80 hasta 190 en 5 días a 12°C.
- 50

Las amilasas de la invención se pueden usar en cualquier proceso de producción de cerveza, como se describe, por ejemplo, en las patentes de los Estados Unidos Nos. 5.762.991. 5.536.650. 5.405.624. 5.021.246. 4.788.066.

En un aspecto, una o más de otras enzimas se usan junto con un alimento, pienso o bebida que comprende enzimas de la invención, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidadasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululanasas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín acetil esteradasas, ramnogalacturonano acetil esteradasas, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesteradasas, celobiohidrolasas y/o transglutaminasas.

Uso en perforación de pozos y operaciones mineras

La divulgación también incluye métodos que utilizan enzimas de la invención en pozos y en operaciones de perforación, por ejemplo, gas, petróleo u otras operaciones de minería o perforación. Por ejemplo, en un aspecto, las enzimas de la invención se utilizan para incrementar el flujo de los fluidos de producción a partir de una formación subterránea, por ejemplo, un pozo o una mina. En un aspecto, las enzimas de la invención se utilizan para remover fluidos viscosos que contienen polisacáridos y/o que contienen almidón que pueden ser nocivos, por ejemplo, fluidos formados durante las operaciones de producción. Estos fluidos que contienen polisacáridos y/o que contienen almidón pueden ser encontrados dentro de una formación subterránea que rodea un orificio de pozo completo. En un aspecto, una amilasa de la invención se usa en un fluido de perforación de pozo de petróleo para ayudar a extraer el lodo de perforación.

La invención provee métodos para cambiar la viscosidad de una composición que comprende: proveer una composición y el polipéptido de la invención y una composición; y tratar la composición con el polipéptido de la invención; y en un aspecto del método, la composición comprende un suelo o un lodo de perforación.

En un aspecto, el uso de estos métodos de la divulgación permite que los fluidos de producción (que comprenden las enzimas de la invención), fluyan desde el orificio del pozo o una mina. Los métodos pueden comprender la reducción del flujo de los fluidos de producción de la formación por debajo de los índices de flujo esperados y la formulación de un tratamiento enzimático al mezclar un fluido acuoso y un polipéptido de la invención. Los métodos pueden comprender bombear el tratamiento enzimático a un sitio deseado dentro del orificio del pozo u otro conducto perforado y permitir que el tratamiento enzimático degrade el fluido nocivo viscoso que contiene almidón. Los métodos pueden comprender la remoción del fluido desde la formación subterránea hasta el pozo o superficie del conducto. En un aspecto, el tratamiento enzimático es efectivo para atacar los enlaces alfa glucosídicos en el fluido que contiene almidón. En un aspecto, las amilasas de la invención se utilizan en la perforación minera, perforación de pozos (por ejemplo, perforación de pozos de petróleo o gas) y similares, para trasladar el lodo de perforación, por ejemplo, mientras se perfora el agujero (orificio o conducto del pozo).

Las enzimas de la invención se pueden usar en cualquier operación de perforación de pozo, conducto o mina, muchos de los cuales son conocidos en el arte. Por ejemplo, la divulgación provee métodos para introducir una enzima de la invención, que en un aspecto también puede comprender la producción de un compuesto químico de un campo petrolero o de gas, en una formación rocosa que comprende petróleo y/o gas, que comprende pasar una microemulsión que comprende la enzima (y, en un aspecto, el compuesto químico) dentro de un pozo de producción y luego en la formación. En un aspecto, se somete un pozo de producción a un tratamiento de "cierre" por el cual una composición acuosa que comprende una enzima de la invención es inyectada en el pozo de producción bajo presión y "presionada" en la formación y mantenerla allí. Véase, por ejemplo, la patente de los Estados Unidos No. 6.581,687.

En un aspecto, las amilasas de la invención usadas en otras operaciones de perforación de gas, petróleo u otras operaciones de perforación o minería son activas a alto o bajo pH y/o altas o bajas temperaturas, por ejemplo, las amilasas de la invención usadas en estos procesos son activas bajo condiciones que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5 o pH 4, o menores, o, bajo condiciones que comprenden aproximadamente pH 7, pH 7,5, pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5 o pH 11 o superior. En un aspecto, las amilasas de la invención usadas en estos procesos son activas bajo condiciones que comprenden un rango de temperatura en cualquier punto entre aproximadamente 0°C hasta aproximadamente 37°C, o entre aproximadamente 37°C hasta aproximadamente 95°C o más, o entre aproximadamente 80°C hasta aproximadamente 120°C, por ejemplo, 85°C, 90°C, 95°C, 98°C, 100°C, 105°C, 110°C, 115°C, 120°C o más.

En un aspecto, una o más de otras enzimas se usan junto con una composición que comprende enzimas de la invención para uso con estos métodos, por ejemplo, incluyendo otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidadasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas,

ligninasas, pululanastas, arabinastas, hemicelulasas, manastas, xilolastas, xilastas, pectín acetil esterastas, ramnogalacturonano acetil esterastas, proteastas, peptidastas, proteinastas, poligalacturonastas, ramnogalacturonastas. galactanastas, pectín liastas, transglutaminastas, pectín metilesterastas, celobiohidrolastas y/o transglutaminastas.

- 5 En un aspecto, las enzimas o cócteles de enzimas de la invención que se usan en estas operaciones de excavación o perforación de gas, petróleo u otros, o que incluyen cualquier otro proceso de fractura y/o lavado de pozos de gas y petróleo, son activas a alto o bajo pH y/o altas o bajas temperaturas, por ejemplo, enzimas que degradan al polímero o que degradan al polisacárido ("rompedor del polímero") de esta invención, que incluyen el uso de "cócteles" de estas y otras enzimas tales como las enzimas amilasa, glucoamilasa, xantanasa, glucosidasa y/o celulasa, o una enzima degradante de lignina, alfa amilasa, beta amilasa, glucoamilasa, dextrinasa, celulasa, celobiohidrolasa, avicelasa, carboximetilcelulasa, beta-glucanasa, glucosidasa, xilanasa, mananasa, arabinofuranosidasa, lacasa, lignina peroxidasa, pectinasa, pectato liasa, xantanasa, xantana liasa, xantana de polimerasa, pululanasa, liquenasa, paquimanasa, lipasa, proteasa, proteinasa, fitasa, peptidasa y catalasa, que incluyen el uso de "cócteles" de estas y otras enzimas, que se utilizan estos procesos son activas bajo condiciones que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5, pH 4,0, pH 3,5, pH 3,0 o menos (más ácido), o, bajo condiciones que comprenden aproximadamente pH 7, pH 7,5, pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5, pH 11,0, pH 11,5, pH 12, pH 12,5 o más (más básico). En un aspecto, las enzimas o cócteles de enzimas de la invención usados en estos procesos son activos bajo condiciones que comprenden una rango de temperatura de cualquier punto entre aproximadamente -100°C hasta aproximadamente -80°C, aproximadamente -80°C hasta aproximadamente -40°C, aproximadamente -40°C hasta aproximadamente 20°C, aproximadamente -20°C hasta aproximadamente 0°C, aproximadamente 0°C hasta aproximadamente 37°C, aproximadamente 0°C hasta aproximadamente 5°C, aproximadamente 5°C hasta aproximadamente 15°C, aproximadamente 15°C hasta aproximadamente 25°C, aproximadamente 25°C hasta aproximadamente 37°C, aproximadamente 37°C hasta aproximadamente 45°C, aproximadamente 45°C hasta aproximadamente 55°C, aproximadamente 55°C hasta aproximadamente 70°C, aproximadamente 70°C hasta aproximadamente 75°C, aproximadamente 75°C hasta aproximadamente 85°C, aproximadamente 85°C hasta aproximadamente 90°C, aproximadamente 90°C hasta aproximadamente 95°C, aproximadamente 95°C hasta aproximadamente 100°C, aproximadamente 100°C hasta aproximadamente 105°C, aproximadamente 105°C hasta aproximadamente 110°C, aproximadamente 110°C hasta aproximadamente 120°C, o 95°C, 96°C, 97°C, 98°C, 99°C, 100°C, 101°C, 102°C, 103°C, 104°C, 105°C, 106°C, 107°C, 108°C, 109°C, 110°C, 111°C, 112°C, 113°C, 114°C, 115°C, 120°C o más.
- 20
- 25
- 30 En una forma de realización, el "mecanismo de desencadenamiento del pH" comprende el uso de enzimas termofílicas, por ejemplo, una "pirolasa" tal como el polipéptido de la SEQ ID NO: 4 y/o de la SEQ ID NO: 6 (codificada por ejemplo, por la SEQ ID NO: 3 y la SEQ ID NO: 5, respectivamente). En un aspecto, la invención provee un sistema que comprende una o más enzimas atrapadas en un lodo o sus ingredientes en una forma seca, por ejemplo, un polvo de goma guar, arena o sal reguladora. En un aspecto, la enzima permanece inactiva o "menos activa" a causa de la baja temperatura del lodo, o permanece inactiva o "menos activa" debido a que la cantidad de carga enzimática se ajusta para que no ocurra degradación, u ocurra menos degradación (no deseada) (de sustrato) en forma "prematura" hasta que se caliente el lodo a una temperatura más alta que conduce a la activación de la(s) enzima(s). Esta forma de realización se puede referenciar como "un sistema enzimático atrapado con un mecanismo de desencadenamiento de temperatura".
- 35
- 40 En un aspecto, la composición y métodos de la invención se usan para degradar los polímeros en un "lodo" que, en formas de realización alternativas, comprende un medio que contiene agua en el cual la enzima, polisacárido y otros componentes se mezclan. En un aspecto, la composición y métodos de la invención se usan como una solución o polvo seco, que se puede mezclar con un ingrediente o componente del "lodo". En un aspecto, la enzima se incorpora/arrastra en los ingredientes del lodo antes de la preparación del lodo real. Por ejemplo, un ejemplo de formulación o mezcla usada para realizar esta invención comprenden un polvo de almidón, xantana o celulosa mezclada con una o más enzimas, o una mezcla de sales reguladoras y enzima(s), en donde cada uno de dichos ingredientes que contienen enzimas se pueden usar luego para preparar el lodo.
- 45

Uso de enzimas libres e inmovilizadas en operaciones de perforación y fracturamiento hidráulico:

- 50 La divulgación provee composiciones y métodos que comprenden la inclusión de enzimas de rompimiento de polímeros (degradación de polímeros), por ejemplo, degradación de polisacáridos, en una forma libre o en una forma inmovilizada, por ejemplo, en una forma inmovilizada sobre un revestimiento, por ejemplo, de una partícula, por ejemplo, de un grano de arena o un material cerámico tal como una bauxita sinterizada.

- 55 En un aspecto, las composiciones y métodos que comprenden la inclusión de enzimas de rompimiento de polímeros (degradación de polímeros), por ejemplo, degradación de polisacáridos en o sobre una resina o material similar que reviste partículas, por ejemplo, granos de arena o un material cerámico tal como una bauxita sinterizada; estas partículas (por ejemplo, granos de arena) se pueden usar como el agente de consolidación o con un agente de consolidación (por ejemplo, una arena revestida con resina o materiales cerámicos de alta resistencia) en un fluido de fracturamiento hidráulico. En un aspecto, un agente de consolidación utilizado para la práctica de esta divulgación son partículas de un cierto tamaño mezcladas con un fluido de fracturamiento para mantener las

fracturas abiertas después de un tratamiento de fracturamiento hidráulico. Además de los granos de arena de origen natural, los agentes de consolidación sintéticos o diseñados especialmente, tales como arena revestida con resina o materiales cerámicos de alta resistencia como bauxita sinterizada, también se pueden usar. Los materiales de agente de consolidación pueden ser clasificados por tamaño y esfericidad para proveer un conducto eficiente para la producción de fluidos desde el depósito hasta el orificio del pozo. Después del asentamiento de la arena en las fracturas y fisuras del pozo, las enzimas enlazadas a resina pueden difundirse y trabajar sobre el polímero concentrado y no quebrado que a menudo se deposita sobre la superficie de la formación en la terminación de las operaciones de fracturamiento. Por lo tanto, este aspecto de la divulgación puede remover, eficientemente un polisacárido, una goma xantana o guar, por ejemplo, una torta de filtración guar, de los pozos de petróleo y gas fracturados, y/o pueden mejorar la permeabilidad de la zona fracturada.

En una forma de realización, durante las operaciones de fracturamiento hidráulico, los grandes volúmenes de agua, arena, compuestos químicos auxiliares (incluyendo enzimas y las mezclas de enzimas de esta invención) y un polímero basado en polisacáridos (por ejemplo, guar y/o sus derivados) son mezclados e inyectados bajo presión en los pozos de petróleo y/o gas para "fracturar" la formación circundante y mejorar el flujo de gas o de petróleo en el orificio del pozo. Las enzimas y mezclas de enzimas como se describe aquí pueden ser usadas para hidrolizar estos polímeros de polisacáridos y reducir la viscosidad del fluido (usados en operaciones de fracturamiento hidráulico) para la mejor penetración y retorno más efectivo del fluido al final de la operación.

En una forma de realización, las composiciones y métodos de esta invención son usados en la hidrólisis enzimática de los polímeros base (por ejemplo, polímeros basados en polisacáridos, tales como goma guar, xantana y/o sus derivados); esto puede solucionar el problema donde la hidrólisis enzimática de estos polímeros base puede ser incompleta para dejar algún polímero "no quebrado" en el fluido usado en las operaciones de fracturamiento hidráulico. Como el contenido de agua en el fluido se pierde en la formación, el fluido se torna más concentrado y los polímeros no quebrados forman una torta de filtro gruesa; esta torta de filtro impide la formación de poros y reduce el flujo de petróleo o gas en el orificio del pozo. En una realización, las composiciones y métodos de esta invención se usan para romper estos tapones de la torta de filtro.

Los fluidos de fracturamiento contienen grandes cantidades de arena, por lo general denominados como el agente de consolidación. A medida que se bombea el fluido en el pozo, el agente de consolidación se asienta en las fisuras y fracturas y evita que las mismas se cierren. Esto ayuda a mejorar la porosidad y permeabilidad de la formación para un mejor flujo del gas/petróleo. Los granos de arena a menudo se revisten con diferentes resinas industriales para incrementar su resistencia mecánica y evitar que choquen bajo la presión de la formación. Por lo tanto, en una forma de realización, la divulgación provee composiciones y métodos que usan enzimas que degradan polímeros libres o inmovilizados ("ruptura de polímeros") aproximadamente, en o sobre el material de revestimiento de la arena. En un aspecto, esto se realiza atrapando la enzima en la resina o por inmovilización sobre la superficie de revestimiento. Por lo tanto, en este aspecto, la(s) enzima(s) usadas para la práctica de esta divulgación pueden permanecer en contacto con la torta de filtro proveyendo así una hidrólisis continua del polímero concentrado, removiendo la torta de las fracturas y mejorando la permeabilidad de la formación fracturada.

En un aspecto, la divulgación provee métodos que usan estas enzimas descritas en operaciones de perforación, por ejemplo, una operación de perforación típica, donde un pozo es creado perforando un orificio de 5 a 30 pulgadas (13 -76 cm) de diámetro en la tierra con una torre de perforación petrolera, que gira una broca. Una vez que el orificio es perforado, una tubería de acero (cubierta) ligeramente menor que el orificio se coloca en el orificio y se asegura con cemento. La cubierta proporciona integridad estructural al orificio del pozo recientemente perforado además de aislar zonas potencialmente peligrosas de alta presión una de otra y de la superficie.

Con estas zonas aisladas en forma segura y la formación protegida por la cubierta, el pozo puede ser perforado más profundamente (en formaciones potencialmente más inestables y violentas) con una broca más pequeña y también revestido con una tubería de tamaño menor. Un pozo puede tener 2 a 5 conjuntos de orificios de menor tamaño perforados uno dentro de otro, cada uno cementado con una cubierta.

A fin de perforar el pozo, la broca, ayudada por un esfuerzo de torsión y el peso de compresión de los anillos de perforación por encima de ella, rompe la tierra. El fluido de perforación, o "lodo", que comprende la inclusión de enzimas y mezclas de enzimas que rompen polímeros (que degradan polímeros), por ejemplo, que degradan polisacáridos, de esta invención, en una forma libre o en una forma inmovilizada, es bombeado por el interior del tubo de perforación. El fluido sale de la broca y ayuda a romper la roca, manteniendo presión en la parte superior de la broca, así como también la limpieza, enfriamiento y lubricación de la broca.

Los "cortes" generados en la roca son barridos por el fluido de perforación a medida que circulan nuevamente hacia la superficie exterior de la tubería de perforación. El fluido que comprende las enzimas y mezclas de enzimas que quiebran polímeros (degradan polímeros), por ejemplo, degradan polisacáridos, de esta invención, en una forma libre o en una forma inmovilizada, también pueden ser añadidos en esta etapa.

Los fluidos luego atraviesan "agitadores" que agitan los recortes sobre cribas permitiendo que el fluido bueno retorne a los pozos. El fluido que comprende enzimas y mezclas de enzimas que quiebran polímeros (degradan polímeros), por ejemplo, degradan polisacáridos, de esta invención, en una forma libre o en una forma inmovilizada, también pueden ser agregados en esta etapa.

5 Estos procesos de la invención pueden ser facilitados por adición de enzimas y mezclas de enzimas que quiebran polímeros (degradan polímeros), por ejemplo, degradan polisacáridos, de esta invención, en una forma libre o en una forma inmovilizada. La plataforma de perforación puede contener todo el equipamiento necesario para circular el fluido de perforación, elevar y voltear el tubo, controlar las presiones descendentes del orificio, extraer los recortes del fluido de perforación y generar energía en el sitio para estas operaciones.

10 Las enzimas, mezclas de enzimas y métodos de la invención pueden ser realizados con cualquier lodo de perforación o fluido de perforación (algunos prefieren reservar el término "fluido de perforación" para "lodos" más sofisticados y bien definidos), o cualquier fluido usado en operaciones para perforar orificios de perforación en la tierra. Las enzimas, mezclas de enzimas y métodos de la invención se pueden realizar al tiempo que se perforan los pozos de petróleo y/o gas natural y en plataformas de perforación de exploración, incluyendo el uso con orificios más simples.

15 Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclarse con y/o llevar a cabo junto con cualquier pozo u operación de perforación, por ejemplo, cuando se usa cualquier lodo, incluyendo el uso de cualquiera de los tres esquemas principales de clasificación de lodo, en donde "lodo" se usa en forma amplia y se separa en 3 categorías con base en el componente principal que forma el lodo: (1) "lodo con base en agua" (WBM), que puede subdividirse en lodos dispersos y no dispersos; (2) lodo "no acuoso" o más comúnmente "lodo con base en aceite" (OBM), incluyendo aceites sintéticos (SBM); y/o (3) lodo gaseoso o neumático.

20 La enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclarse con y/o llevar a cabo junto con cualquier pozo u operación de perforación, por ejemplo, también se pueden usar en o con:

- Pozos de producción cuando se perforan principalmente para producir petróleo o gas, una vez que se establecen la estructura y características de producción,
- pozos de aforo cuando se usan para evaluar las características (tal como tasa de flujo) de una acumulación probada de hidrocarburos,
- pozos de exploración cuando se perforan simplemente para propósitos exploratorios (reunión de información) en una área nueva,
- pozos de prueba cuando se perfora un pozo, con base en un gran elemento de expectativa, en un área de frontera donde se conoce muy poco sobre la subsuperficie.

25 La enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a cabo junto con cualquier pozo u operación de perforación, por ejemplo, también se pueden usar junto con los métodos, equipos y/u operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070089910, Hewson, y colaboradores, que describe, por ejemplo, métodos para formar un orificio perforado subterráneo soportado y usa, por ejemplo, un motor de lodo de desplazamiento positivo.

30 La enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclarse con y/o llevar a cabo junto con métodos, equipos y/u operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070084638, Bohnsack; C., y colaboradores, que describe, por ejemplo, un sistema para facilitar el flujo de los sólidos asentados con fluido de perforación desde un contenedor; el sistema incluye un aparato de boquilla a presión con al menos una boquilla desde la cual el fluido que puede fluir bajo presión, el aparato energizado de rotación para hacer girar selectivamente el aparato de boquilla a presión para que al menos una boquilla pueda moverse dentro del contenedor a medida que se bombea el fluido a través de al menos una boquilla dentro del contenedor; y, en un aspecto, el aparato de traslado para mover el aparato de boquilla a presión con respecto al contenedor a medida que se bombea fluido bajo presión hasta al menos una boquilla de rotación. También se describen los tanques de lodos y los pozos de lodo, y se pueden usar las enzimas, mezclas de enzimas y los métodos de la invención en o con cualquiera de estos fluidos, y/o en cualquiera de los tanques de lodo y pozos de lodo utilizados en estos tipos de operación.

35 La enzimas, mezclas de enzimas y métodos de la invención se pueden utilizar en, mezclar con y/o llevar a cabo junto con los métodos, equipo y/u operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070081157, Csutak; S., y colaboradores, que describe, por ejemplo, un aparato para estimar una propiedad de un fluido descendiente del orificio que comprende una fuente de luz ultravioleta (UV) para inducir luz en el fluido a una longitud de onda que produce luz de dispersión Raman a longitudes de onda que son más cortas que las longitudes de onda de fluorescencia sustancial reflejada desde el fluido en respuesta a la luz inducida; un detector que detecta un espectro de la luz de dispersión Raman y provee señales en respuesta al espectro detectado; y un procesador que procesa las señales para proveer un estimado de

una propiedad del fluido. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o con cualquiera de estos fluidos, y/o en operaciones para estimar la contaminación filtrada en un fluido de formación. Por ejemplo, estos métodos incluyen detectar dispersiones Raman a una pluralidad de longitudes de onda de al menos un componente presente en un lodo con base en aceite que no está presente en forma natural en la formación y las enzimas, mezclas de enzimas y los métodos de la invención pueden ser utilizados para ayudar en la precisión de esta detección.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a cabo junto con los métodos, equipo y/o operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070075706, Chen, S., y colaboradores, que describe, por ejemplo, métodos para evaluar una formación terrestre que comprende hacer mediciones con una herramienta de orificio descendente en un orificio de perforación en la formación terrestre; medir un factor de calidad de una antena de la herramienta de orificio descendente a profundidades donde se realizan las mediciones; y usar la Q medida y una resistividad de un lodo en el orificio de perforación y una resistividad de formación, y/o un indicador de tamaño de orificio de perforación (BSI), para estimar la resistividad de formación y BSI, que incluye medir la resistividad del lodo en el orificio de perforación. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o junto con cualquiera de estos fluidos y/o en operaciones para evaluar una formación terrestre.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a la práctica junto con los métodos, equipo y/o operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070068675, Barry, M., y colaboradores, que describe, por ejemplo, métodos para perforar y completar un orificio relleno con grava, que comprende perforar un orificio del pozo con un fluido de perforación, acondicionar el fluido de perforación, llevar las herramientas de ensamblado de relleno con grava a la profundidad en el orificio del pozo con el fluido de perforación acondicionado y rellenar con grava un intervalo del orificio del pozo con un fluido de terminación. El fluido de terminación puede ser el mismo que el fluido de perforación. Este método puede ser combinado con tecnología de cribado de arena de ruta alterna para asegurar la distribución apropiada del relleno con grava. Los fluidos apropiados para perforar, rellenar con grava e instalación de cribas de arena son esenciales para el éxito de la terminación del pozo. La planificación cuidadosa, buena preparación y ejecución de terminación son necesarias para incrementar la longevidad y productividad de terminación. Usualmente, un mínimo de tres fluidos han sido usados para perforar y completar los pozos rellenos con grava. El primer fluido es un fluido de perforación cargado de sólidos usado para perforar el intervalo de terminación. El segundo fluido es un fluido de terminación libre de sólidos usados para desplazar el fluido de perforación cargado de sólidos y activar el equipo de exclusión de arena y herramientas de relleno con grava en un ambiente generalmente libre de sólidos. El tercer fluido es un fluido portador para la grava durante el relleno con grava del intervalo de terminación. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o con cualquiera de estos fluidos (incluyendo fluidos de perforación cargados de sólidos, fluidos de terminación libres de sólidos y/o fluidos portadores) y/o en operaciones para perforar y finalizar un pozo relleno con grava.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a la práctica junto con los métodos, equipo y/o operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070066491, Bicerano, J., y colaboradores, uso de partículas en la construcción, perforación, terminación y/o estimulación de fractura de pozos de petróleo y gas natural; por ejemplo, como una monocapa parcial del agente de consolidación, un relleno de agente de consolidación, un componente integral de terminación de relleno con grava, un cojinete de bolas, un lubricante sólido, un constituyente de lodo de perforación, y/o un aditivo de cemento, incluyendo el uso de partículas de polímero termoendurecibles para uso en las aplicaciones que requieren partículas de peso liviano que poseen alta rigidez, fuerza, resistencia a temperatura, y/o resistencia a medios agresivos. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o con cualquiera de estos rellenos con grava, cojinetes de bolas, lubricantes sólidos, constituyentes de lodo de perforación, aditivos de cemento y/o las partículas de polímero termoendurecibles descritas. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o con nanorellenos y/o nanocompuestos, que incluyen morfologías de nanocompuestos heterogéneas.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a la práctica junto con los métodos, equipo y/o operaciones de perforación como se describe, por ejemplo, en la publicación de la solicitud de patente de los Estados Unidos No. 20070039735, Robertson, B., y colaboradores, que describe, por ejemplo, métodos para sellar una zona permeable en una formación subterránea, que comprende: preparar una composición de taponamiento que comprende petróleo, arcilla, cloruro de magnesio y polvo de óxido de magnesio; y poner en contacto la composición de taponamiento con agua en la formación subterránea de manera tal que la composición de taponamiento forme una masa selladora, sellando sustancialmente, de este modo, una zona permeable en la formación subterránea.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en, mezclar con y/o llevar a la práctica junto con lodos de perforación de densidad variable que comprenden materiales de partículas comprimibles, por ejemplo, como se describe en la publicación de la solicitud de patente de los Estados Unidos No. 20070027036, Polizzotti, R., y colaboradores. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o

con, por ejemplo, lodos de perforación que comprenden un material particulado comprimible en el lodo de perforación, en donde la densidad del lodo de perforación cambia debido al cambio de volumen del material particulado comprimible en respuesta a los cambios de presión o temperatura y en donde el material particulado comprimible se configura para mantener la densidad del lodo de perforación entre un gradiente de presión de poro y un gradiente de fractura basados en el cambio de volumen del material particulado comprimible en respuesta a cambios de presión a ciertas profundidades.

Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar para modificar la viscosidad del lodo de perforación solo o junto con los materiales comprimibles descritos (véase Polizzotti; R., y colaboradores), por ejemplo, para ubicar la viscosidad del fluido dentro de los requisitos de capacidad de bombeo, y/o ajustar el gradiente de presión de poro y gradiente de fractura. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar para efectuar un cambio de volumen en el lodo de perforación, por ejemplo, donde la reología del lodo de perforación se configura para lograr una reología deseada de lodo de perforación compuesto.

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan para alterar las propiedades del lodo de perforación para proveer un punto deseado de gel de lodo compuesto, por ejemplo, un punto de gel de lodo que puede suspender los recortes de rocas en una corona circular de un orificio del pozo durante las operaciones de perforación; y/o alterar la viscosidad del lodo de perforación junto con, o solamente (sin), objetos huecos comprimibles (véase Polizzotti; R., y colaboradores) para alterar los requisitos de capacidad de bombeo.

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan para alterar las propiedades de los fluidos del pozo que comprenden lodos de perforación, fluidos de limpieza del pozo, fluidos de reacondicionamiento, fluidos espaciadores, fluidos de relleno con grava, fluidos de acidificación y/o fluidos de fracturamiento. En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan para facilitar la perforación, terminación y/o estimulación de una formación subterránea usando un fluido de densidad variable y para modificar el fluido de densidad variable.

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan en los métodos de perforación, terminación y/o estimulación de las formaciones subterráneas usando un fluido de densidad variable, por ejemplo, al modificar y/o "ajustar" la densidad del fluido; por ejemplo, un método (véase Polizzotti; R., y colaboradores) que comprende las etapas de: introducir un fluido que tiene una densidad que varía en función de la presión en la formación subterránea, donde el fluido comprende un fluido base y una porción de partículas elásticas; y perforar, finalizar y/o estimular una formación subterránea usando el fluido de densidad variable (que puede comprender las enzimas, mezclas de enzimas de la invención, o que han sido modificadas por los métodos de la invención).

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención son usadas con los métodos y composiciones como se describe en la patente de los Estados Unidos No. 4.099.583, que describe, por ejemplo, un sistema de perforación de gradiente dual, donde un fluido más liviano se inyecta en el anillo de retorno del lodo (por lo general en el tubo de subida) u otra vía para reducir la densidad del lodo del punto de inyección en forma ascendente y las enzimas, mezclas de enzimas y métodos de la invención pueden modificar y/o "ajustar" la densidad de este fluido.

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan con los métodos y composiciones como se describe en la patente de los Estados Unidos No. 6.530.437 y la patente de los Estados Unidos No. 6.588.501, que describen un método de perforación multigradiente y un aparato para la reducción de presión hidrostática en los tubos de subida submarinos; y la patente de los Estados Unidos No. 6.422.326, la patente de los Estados Unidos No. 6.156.708, las patentes de los Estados Unidos Nos. 5.910.467 y 5.881.826, que describen la adición de varios afromes fluidos para las formulaciones del lodo de perforación.

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se usan con los métodos y composiciones como se describe en la patente de los Estados Unidos No. 6.497.289, que describe el uso de caños expandibles sólidos, por ejemplo, como sistemas tubulares que funcionan en un pozo y se expanden.

En formas de realización alternativas, las enzimas, mezclas de enzimas y métodos de la invención se usan para adecuar la densidad del lodo de perforación con la profundidad de manera tal que el peso efectivo del lodo permanezca entre el gradiente de presión de poro y de fractura en todas las profundidades. La variación requerida en la densidad del lodo se puede lograr al cambiar las propiedades de los fluidos con las enzimas, mezclas de enzimas y métodos de la invención para modificar/cambiar el volumen y densidad, para efectuar un cambio en respuesta a la presión. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar con cualquier componente particulado, por ejemplo, diferentes formas, tales como esferas, cubos, pirámides, esferoides oblatos o prolados, cilindros, cojinetes y/u otras formas o estructuras. Las enzimas, mezclas de enzimas y métodos de la invención se pueden usar con cualquier componente particulado, por ejemplo, objetos huecos comprimibles que se llenan con gas presurizado, o materiales sólidos comprimibles u objetos como se describe en Polizzotti; R., y colaboradores, ver más arriba.

En aspectos alternativos, las enzimas, mezclas de enzimas y métodos de la invención se pueden usar en o con cualquier pozo u operación de perforación, por ejemplo, que incluye perforación direccional, a menudo conocida como perforación inclinada, para perforar pozos no verticales, incluyendo los tres grupos principales usados en cualquiera de las perforaciones direccionales; Perforación Direccional de Campo Petrolífero, Perforación Direccional de Instalación de Utilidades (comúnmente conocida como H.D.D./ perforación direccional horizontal/taladrado direccional); y/o perforación direccional dentro de la beta (metano en lecho de carbón).

En un aspecto, las enzimas, mezclas de enzimas y métodos de la invención se pueden usar junto con el sondeo de pozo, una técnica usada en la industria del petróleo y gas para registrar las propiedades de la roca y del fluido para encontrar zonas de hidrocarburos en las formaciones geológicas dentro de la corteza terrestre. El sondeo se puede realizar para medir el efecto de llevar a la práctica los métodos de esta invención, por ejemplo, bombear fluidos que comprenden las enzimas o mezclas de enzimas de esta invención en un pozo. Un procedimiento de sondeo puede consistir en descender una "herramienta de sondeo" sobre el extremo de un cable en un pozo petrolero (u orificio) para medir las propiedades de la roca y de los fluidos de la formación. Luego se realiza una interpretación de estas mediciones para localizar y cuantificar las zonas de profundidad potenciales que contienen petróleo y gas (hidrocarburos). Las herramientas de sondeo desarrolladas con el paso de los años miden propiedades eléctricas, acústicas, radioactivas, electromagnéticas y otras propiedades de las rocas y sus fluidos contenidos. El sondeo es usualmente realizado a medida que las herramientas de sondeo se retiran del orificio. Este dato se registra en un registro impreso llamado "sondeo de pozo" y se transmite normalmente en forma digital a las oficinas. El sondeo de pozos se realiza en diferentes intervalos durante la perforación del pozo y cuando se hace la perforación total, que podría estar en un intervalo de profundidades de 300 m hasta 8000 m (1000 pies a 25,000 pies) o más.

Además de los métodos, enzimas o mezclas de enzimas descritas aquí, los métodos, los lodos de enzimas u otros fluidos de perforación utilizados para la práctica de esta invención pueden comprender (el uso de) un lodo de perforación con base en agua que puede comprender una arcilla bentonita (gel) y en algunos aspectos, y también comprende aditivos tales como sulfato de bario (barita), carbonato de calcio (tiza) o hematita. Varios espesantes también pueden ser usados para influenciar la viscosidad del fluido, por ejemplo, lignosulfonatos, goma xantana, goma guar, glicol, carboximetilcelulosa, celulosa polianiónica (PAC), o almidón. Las enzimas o mezclas de enzimas descritas aquí, usadas para llevar a la práctica esta invención se pueden usar para modificar las propiedades de (por ejemplo, la viscosidad de) los fluidos, por ejemplo, para modificar las propiedades de los lignosulfonatos, goma xantana, goma guar, glicol, carboximetilcelulosa, celulosa polianiónica (PAC), o almidón.

Los métodos, enzimas o mezclas de enzimas descritos aquí, usados para llevar a la práctica esta invención se pueden usar para modificar las propiedades de los defloculantes, que se usan para reducir la viscosidad de lodos con base en arcilla; los polielectrolitos aniónicos, por ejemplo, acrilatos, polifosfatos, lignosulfonatos (Lig) o derivados de ácido tánico tales como Quebracho ("lodo rojo" fue el nombre para una mezcla a base de Quebracho, denominada por el color de las sales rojas de ácido tánico; se usaba comúnmente en 1940 a 1950, luego se tornó obsoleto cuando entraron en uso los lignosulfatos).

Los métodos, enzimas o mezclas de enzimas descritas aquí, usados para llevar a la práctica esta invención se pueden usar en (por ejemplo, añadir a) inyectores de agua para inyectar agua en una formación, ya sea para mantener la presión del reservorio o simplemente descartar el agua producida con un hidrocarburo (por ejemplo, porque incluso después del tratamiento, sería demasiado oleoso y demasiado salino como para ser considerado limpio para vertimiento, por ejemplo, vertimiento al agua o en una fuente de agua fresca en el caso de pozos en tierra). Por lo tanto, los métodos y composiciones (por ejemplo, mezclas de enzimas, enzimas inmovilizadas) de esta invención se usan con inyección de agua como elemento de manejo de reservorio y descarte de agua producida.

Los métodos, enzimas o mezclas de enzimas descritos aquí, usados para llevar a la práctica esta invención se pueden usar en (por ejemplo, agregarse a) productores de acuíferos, por ejemplo, como en reservorios de agua producido en forma intencional para reinyección (por ejemplo, en una perforación de pozo) para manejar presión, esto es, en efecto, mover agua del reservorio desde donde no es tan útil hasta donde es más útil. Estos pozos por lo general se usan solamente si el agua producida desde el productor de petróleo o gas es insuficiente para los fines de manejo del reservorio. Por lo tanto, en un aspecto, los métodos y composiciones (por ejemplo, mezclas de enzimas, enzimas inmovilizadas) de esta invención se usan con agua producida de acuífaro y/o agua de mar.

50 Composiciones de liberación retardada

La invención provee composiciones de liberación retardada o "liberación controlada" que comprende una composición deseada recubierta por un polímero de látex, por ejemplo, una pintura de látex, o equivalente. Las composiciones de liberación retardada/ liberación controlada de la invención pueden comprender cualquier composición deseada, que incluye enzimas y cualquier ingrediente activo, incluyendo moléculas pequeñas, fármacos, polisacáridos, lípidos, ácidos nucleicos, vitaminas, antibióticos, insecticidas y similares. En un aspecto, el revestimiento no se disolverá con rapidez a una temperatura relativamente baja pero se descompondrá para liberar la composición deseada (por ejemplo, enzima) a una temperatura relativamente superior.

La divulgación provee métodos para la liberación retardada/ liberación controlada de las composiciones en donde la composición es recubierta por un polímero de látex, por ejemplo, una pintura de látex, o equivalente.

5 Las composiciones de liberación retardada / liberación controlada y métodos de la invención pueden ser usadas para una variedad de aplicaciones médicas e industriales, por ejemplo, en un aspecto, las composiciones enzimáticas de liberación retardada / liberación controlada de la invención comprenden enzimas implicadas en fluidos de fracturamiento de goma guar en operaciones mejoradas de recuperación de petróleo. La aplicación degradante de goma guar en un campo petrolífero de la invención se facilita por un recubrimiento que no se disolverá fácilmente a baja temperatura pero se descompondrá para liberar la enzima a temperaturas superiores.

10 En otro aspecto, las composiciones enzimáticas de liberación retardada/ liberación controlada de la invención comprenden piensos para animales o suplementos nutricionales que comprenden, por ejemplo, enzimas, vitaminas, antibióticos y/u otros alimentos, fármaco o suplementos nutricionales. Estos compuestos activos en los piensos animales o suplementos nutricionales están protegidos de las condiciones de granulación o digestión gástrica por medio del recubrimiento sobre una composición de liberación retardada/ liberación controlada de la invención.

15 En un aspecto, la liberación es una liberación activada por temperatura, por ejemplo, la composición deseada (por ejemplo, enzimática) se libera a una temperatura elevada, por ejemplo, entre aproximadamente 37°C hasta aproximadamente 95°C o más, por ejemplo, 85°C, 90°C, 95°C, 98°C, 100°C o más. La velocidad de liberación puede ser controlada por el espesor o la cantidad de polímero de látex o "barrera", aplicada a la composición deseada, por ejemplo, una pella o matriz que comprende la composición deseada. Por lo tanto, la invención provee pellas o matrices que tienen un rango de espesor de polímero de látex o equivalente y métodos para usarlas.

20 La invención provee composiciones enzimáticas de liberación retardada/ liberación controlada, por ejemplo, en un aspecto, que comprende una enzima de la invención. En un aspecto, la invención provee una enzima (por ejemplo, una enzima de la invención), o una composición en pellas que comprende una enzima (por ejemplo, una enzima de la invención), recubierta con un polímero de látex, por ejemplo, una pintura de látex, o equivalente. En un aspecto, la divulgación provee métodos para elaborar composiciones enzimáticas de liberación retardada que comprende el recubrimiento de una enzima (por ejemplo, una enzima de la invención), o una composición en pellas que comprende una enzima (por ejemplo, una enzima de la invención), con un polímero de látex, por ejemplo, una pintura de látex, o equivalente. En un aspecto, la divulgación provee métodos para elaborar composiciones de liberación retardada/ liberación controlada que comprenden el revestimiento de un compuesto deseado con un polímero de látex, por ejemplo, una pintura de látex, o equivalente.

30 Los polímeros de látex que se usan en las composiciones de liberación retardada/ liberación controlada (por ejemplo, composiciones enzimáticas de liberación retardada/ liberación controlada) de la invención incluyen, pero no se limitan a, diferentes tipos tales como los siguientes: acrílicos; alquídicos; celulosas; indenos de cumarona; epóxidos; ésteres; hidrocarburos; maleicos; melaminas; resinas naturales; resinas oleosas; fenólicos; poliamidas; poliésteres; rosinas; siliconas; estirenos; terpenos; ureas; uretanos; vinilos; y similares. Los polímeros de látex que se usan en las composiciones de liberación retardada de la invención también incluyen, pero no se limitan a, uno o más homo o copolímeros que contienen uno o más de los siguientes monómeros: (met)acrilatos; acetato de vinilo; estireno; etileno; cloruro de vinilo; butadieno; cloruro de vinilideno; versatato de vinilo; propionato de vinilo; acrilato de t-butilo; acrilonitrilo; neopreno; maleatos; fumaratos; y similares, incluyendo derivados plastificados u otros derivados de los mismos.

40 La cantidad de polímeros de látex usados en la composición de látex de la invención no es crítica, pero puede ser cualquier cantidad después de los procedimientos bien establecidos usando polímeros de látex. En aspectos alternativos, la cantidad de polímero de látex seco es al menos de aproximadamente 1, o, de aproximadamente 2 hasta aproximadamente 50, o, de aproximadamente 3 hasta aproximadamente 40 por ciento en peso de la composición de látex total. La composición de látex de la invención puede contener en forma opcional otros componentes tal como aquellos generalmente usados en las composiciones de látex. Estos componentes adicionales incluyen, pero no se limitan a, uno o más de los siguientes: solventes tales como hidrocarburos alifáticos o aromáticos, alcoholes, ésteres, cetonas, glicoles, éteres de glicol, nitroparafinas o similares; pigmentos; agentes de relleno, secadores; agentes de aplanamiento; plastificantes; estabilizantes; dispersantes; agentes tensoactivos; agentes de viscosidad incluyendo espesantes asociativos poliméricos, espesantes con base en polisacáridos y así sucesivamente; agentes de suspensión; agentes de control de flujo; desespumantes; agentes anti-descascaramiento; y conservantes; extensores; agentes de formación de película; entrecruzadores; mejoradores de superficie; inhibidores de corrosión; y otros ingredientes útiles en las composiciones de látex. En un aspecto, se provee composiciones de látex de la invención que tienen reología y estabilidad mejoradas mediante la combinación del polímero de látex y un polisacárido con agua después de los procedimientos establecidos. Véase, por ejemplo
55 las patentes de los Estados Unidos Nos. 6.372.901 y 5.610.225.

En un aspecto, en la elaboración de una composición que comprende pellas o matrices de la invención que comprende una composición activa, por ejemplo, una enzima (por ejemplo, una enzima de la invención), recubierta con un polímero de látex, por ejemplo, una pintura de látex, o equivalente, la composición activa (por ejemplo,

enzima) se embebe en el cuerpo de la pella (en un aspecto, una mayoría o toda la composición activa (por ejemplo, enzimática)) se embebe en la pella. Por lo tanto, se pueden usar compuestos químicos duros, por ejemplo, el recubrimiento de látex, que puede ser un inactivador del ingrediente activo deseado, para recubrir la superficie de la pella o matriz. La composición del recubrimiento puede ser descompuesta por agentes tales como calor, ácido,
 5 base, presión, enzimas, otros compuestos químicos y similares, para tener una liberación controlada de la actividad enzimática deseada desencadenada por la exposición al agente de degradación del recubrimiento.

En un aspecto, una composición activa, por ejemplo, una enzima (por ejemplo, una enzima de la invención, u otra enzima, por ejemplo una manasa), se dispersa en una matriz de harina de maíz y/o almidón de maíz (por ejemplo, como una pella). Esta mezcla (por ejemplo, una pella) se desintegra en un lapso diez minutos a temperatura
 10 ambiente en agua (por ejemplo, aproximadamente a 22 °C) para liberar toda la composición activa, (100%) por ejemplo, libera toda la actividad enzimática. A temperaturas superiores, se incrementa la velocidad de liberación. Esta no es una velocidad aceptable de desintegración para muchos usos.

Sin embargo, como una composición de liberación retardada/ liberación controlada de la invención, es decir, cuando esta mezcla se recubre con un polímero de látex, por ejemplo, una pintura de látex, o equivalente, la desintegración
 15 de la mezcla (por ejemplo, una pella, una matriz) se retardada. La velocidad y el grado de liberación se pueden controlar mediante el espesor del recubrimiento (barrera) aplicado a la pella o matriz. Por ejemplo, una partícula recubierta liberará solamente 30% de la actividad después de seis horas en agua a 22°C. A 60°C, 50% de la enzima se libera en 90 minutos. A 80°C, 80% de la enzima se libera durante una hora.

En un aspecto, una o más de otras enzimas se añaden a una composición de liberación retardada/ liberación controlada de la invención, por ejemplo, que incluye otras amilasas, beta-galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidadas, amilasas glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululanadas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín acetil esteradas, ramnogalacturonano acetil esteradas, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesteradas, celobiohidrolasas y/o transglutaminasas.
 20
 25

Conversión de biomasa y producción de biocombustibles limpios

La invención provee enzimas y métodos para la conversión de biomasa en combustibles (por ejemplo, bioetanol, biopropanol, biobutanol, o biodiésel) y compuestos químicos. Por lo tanto, las composiciones y métodos de la
 30 invención proveen alternativas efectivas y sostenibles para el uso de productos a base de petróleo. La invención provee organismos que expresan enzimas de la invención para la participación en ciclos químicos que implican la conversión de biomasa natural. En un aspecto, las enzimas y métodos para la conversión se usan en los ensamblajes enzimáticos para la despolimerización eficiente de polímeros de biomasa en fracciones de carbono metabolizables. Como se discutió anteriormente, la invención provee métodos para descubrir e implementar lo más
 35 efectivo de las enzimas para permitir estos nuevos procesos de "conversión de biomasa" y procesos industriales de energía alternativos.

En un aspecto, los polipéptidos de la invención se usan en procesos para convertir biomasa lignocelulósica y/o de almidón en etanol. La invención también provee procesos para elaborar etanol ("bioetanol"), propanol ("biopropanol"), butanol ("biobutanol"), o combustible diesel ("biodiésel"), a partir de las composiciones que
 40 comprenden biomasa lignocelulósica y/o de almidón. El material de biomasa de lignocelulosa y/o almidón puede ser obtenido a partir de cultivos agrícolas, como subproducto de la producción de alimentos o piensos, o como producto de desechos lignocelulósicos, tales como residuos vegetales y residuos de papel. Los ejemplos de residuos vegetales adecuados para tratamiento con los polipéptidos de la invención incluyen tallos, hojas, cáscaras, cortezas, mazorcas y similares, así como también madera, astillas de madera, pulpa de madera y aserrín. Los ejemplos de
 45 residuos de papel adecuados para tratamiento con los polipéptidos de la invención incluyen residuos papel de fotocopia, papel para impresión por ordenador, papel de cuadernos, papel de anotadores, papel de mecanografía y similares, así como también diarios, revistas, cartón y materiales de embalaje basados en papel.

En un aspecto, las enzimas y métodos de la invención se pueden usar junto con medios más "tradicionales" para elaborar etanol a partir de biomasa, por ejemplo, como métodos que comprenden hidrolizar materiales lignocelulósicos al someter el material lignocelulósico seco en un reactor a un catalizador incluido en una solución
 50 diluida de un ácido fuerte y una sal metálica; esto puede disminuir la energía de activación o la temperatura, de hidrólisis de celulosa para obtener mayores rendimientos de azúcar; véanse, por ejemplo, las patentes de los Estados Unidos Nos. 6.660.506; 6.423.145.

Otro ejemplo de un método que incorpora el uso de las enzimas de la invención comprende catalizar la hidrólisis de polisacáridos que comprende monómeros de glucosa, tales como almidón (un polímero de monómeros de glucosa unido por enlaces 1,4-alfa o 1,6- alfa), en azúcares; esto se puede usar junto con la enzima para hidrolizar material lignocelulósico que contiene hemicelulosa, celulosa y lignina. En un aspecto, la biomasa se somete a una primera
 55

etapa de hidrólisis en un medio acuoso a una temperatura y una presión escogidos para efectuar la despolimerización primaria de hemicelulosa sin una despolimerización importante de celulosa a glucosa. Este etapa genera una suspensión en la cual la fase acuosa líquida contiene monosacáridos disueltos generados a partir de la despolimerización de hemicelulosa y una fase sólida que contiene celulosa y lignina. Una segundo paso de la etapa de hidrólisis puede comprender condiciones de manera tal que al menos una parte principal de la celulosa se despolimerice, resultando dicha etapa en una fase acuosa líquida que contiene productos de despolimerización disueltos/ solubles de celulosa, que pueden ser hidrolizados con las enzimas de esta invención. Véase, por ejemplo, la patente de los Estados Unidos No. 5.536.325. Las enzimas de la invención pueden ser agregadas en cualquier etapa de este proceso que sirve como ejemplo.

Otro ejemplo de un método que incorpora las enzimas de la invención comprende procesar un material de biomasa que contiene lignocelulosa por una o más etapas de hidrólisis con ácido diluido con aproximadamente 0,4% a 2% de ácido fuerte; y tratar un componente lignocelulósico sólido no reaccionado del material de biomasa hidrolizado de ácido por delignificación alcalina para producir precursores para termoplásticos biodegradables y derivados. Véase, por ejemplo, la patente de los Estados Unidos No. 6.409.841. Las enzimas de la invención pueden ser agregadas en cualquier etapa de este proceso que sirve como ejemplo.

Otro ejemplo de un método que incorpora el uso de las enzimas de la invención comprende prehidrolizar material lignocelulósico en un reactor de prehidrólisis; agregar un líquido ácido al material lignocelulósico sólido para formar una mezcla; calentar la mezcla a una temperatura de reacción; mantener la temperatura de reacción durante un tiempo suficiente para fraccionar el material lignocelulósico en una porción soluble que contiene al menos aproximadamente el 20% de la lignina del material lignocelulósico y una fracción sólida que contiene celulosa; remover una porción solubilizada de la fracción sólida cuando esté en o cerca de la temperatura de reacción, en donde la celulosa en la fracción sólida se hace más propensa a la digestión enzimática; y recuperar una porción solubilizada. Véase, por ejemplo, la patente de los Estados Unidos No. 5.705.369. Las enzimas de la invención pueden ser agregadas a cualquier etapa de este proceso que sirve como ejemplo.

La invención provee métodos para elaborar composiciones de combustible para motores (por ejemplo, para motores de ignición por chispa) basadas en hidrocarburos líquidos mezclados con un alcohol de grado combustible elaborado usando una enzima o un método de la invención. En un aspecto, los combustibles elaborados mediante el uso de una enzima de la invención comprenden, por ejemplo, mezclas de etanol con gas de carbón líquido o gas natural líquido. En un aspecto, un cosolvente es un 2-metiltetrahidrofurano (MTHF) derivado de biomasa. Véase, por ejemplo, la patente de los Estados Unidos No. 6.712.866.

Los métodos de la invención para la degradación enzimática de lignocelulosa, por ejemplo, para la producción de etanol a partir del material lignocelulósico, también pueden comprender el uso de tratamiento ultrasónico del material de biomasa; véase, por ejemplo, la patente de los Estados Unidos No. 6.333.181.

Otro ejemplo de un proceso para la elaboración de un biocombustible que comprende, por ejemplo, un bioetanol, biopropanol, biobutanol, o un biodiésel, usando las enzimas de la invención comprende pretratar un material de partida que comprende una materia prima lignocelulósica que comprende al menos hemicelulosa y celulosa. En un aspecto, el material de partida comprende papatas, soja (semilla oleaginosa), cebada, centeno, maíz, avena, trigo, remolacha o caña de azúcar o un componente o residuo o subproductos de la producción de alimentos o piensos. El material de partida ("materia prima") reacciona bajo condiciones que rompen la estructura de las fibras de la planta para efectuar al menos una hidrólisis parcial de la hemicelulosa y celulosa. Las condiciones de ruptura pueden comprender, por ejemplo, someter el material de partida a una temperatura promedio de 180°C a 270°C a pH de 0,5 a 2,5 durante un período de aproximadamente 5 segundos a 60 minutos; o, temperatura de 220°C a 270°C, a un pH de 0,5 a 2,5 durante un período de 5 segundos a 120 segundos, o equivalente. Esto genera una materia prima con mayor accesibilidad para ser digerida por una enzima, por ejemplo, una amilasa, de esta invención. Patente de los Estados Unidos No. 6.090.595.

Los Ejemplos de condiciones para hidrólisis enzimática de material lignocelulósico incluyen reacciones a temperaturas entre aproximadamente 30°C y 48°C, y/o un pH entre aproximadamente 4,0 y 6,0. Otros ejemplos de condiciones incluyen una temperatura entre aproximadamente 30°C y 60°C y un pH entre aproximadamente 4,0 y 8,0.

En un aspecto de estos procesos para generación de biocombustibles (tal como un bioetanol, biopropanol, biobutanol, o un biodiésel) de la invención usando al menos una enzima de la invención, se añaden una o más de otras enzimas, por ejemplo, otras amilasas, beta- galactosidasas, catalasas, lacasas, celulasas, endoglucosidasas, endo-beta-1,4-lacasas, amiloglucosidasas, otras glucosidasas, glucosa isomerasas, glucosiltransferasas, lipasas, fosfolipasas, lipooxigenasas, beta-lacasas, endo-beta-1,3(4)-lacasas, cutinasas, peroxidasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, decarboxilasas, fenoloxidasas, ligninasas, pululananas, arabinanasas, hemicelulasas, mananasas, xilolacasas, xilanasas, pectín acetil esterases, ramnogalacturonano acetil esterases, proteasas, peptidasas, proteinasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectín liasas, transglutaminasas, pectín metilesterasas, celobiohidrolasas y/o transglutaminasas.

Por lo tanto, la invención provee métodos para procesar un material de biomasa que comprende lignocelulosa que comprende poner en contacto una composición que comprende un polipéptido de la invención, en donde el material de biomasa comprende en forma opcional o se deriva de un cultivo agrícola, o es un subproducto de la producción de un alimento o un pienso, o es un producto de desecho lignocelulósico, o es un residuo vegetal o un residuo de papel o producto residual de papel, y el polipéptido tiene actividad que comprenden actividad de amilasa, y opcionalmente, el residuo de la planta comprende tallos, hojas, cáscaras, certezas, mazorcas, madera, astillas de madera, pulpa de madera y aserrín y, en opcionalmente, el residuo de papel comprende papel para fotocopias usado o descartado, papel para impresión por ordenador, papel de cuadernos, papel de anotadores, papel de mecanografía, diarios, revistas, cartón y materiales de embalaje basados en papel y, en opcionalmente, el procesamiento del material de biomasa genera un bioetanol. La invención provee material de biomasa que comprende un polipéptido de la invención.

La invención provee métodos para elaborar biocombustible (tal como un bioetanol, biopropanol, biobutanol, o un biodiésel) que comprende poner en contacto una composición que comprende un azúcar fermentable con un polipéptido de la invención, en donde, opcionalmente, la composición que comprende un azúcar fermentable comprende una planta, producto vegetal o derivado vegetal y, opcionalmente, la planta o producto vegetal comprende plantas o productos vegetales de caña de azúcar, remolacha o remolacha azucarera, trigo, maíz, soja, patata, arroz o cebada, y opcionalmente, el polipéptido tiene actividad que comprende amilasa, glucoamilasa, glucosidasa, por ejemplo, actividad de alfa-glucosidasa o de beta-glucosidasa.

La invención provee métodos para elaborar un combustible (tal como un bioetanol, biopropanol, biobutanol, o un biodiésel) que comprende poner en contacto una composición que comprende un azúcar fermentable con un polipéptido de la invención, en donde, opcionalmente, la composición que comprende un azúcar fermentable comprende una planta, producto vegetal o derivado vegetal y, en forma opcional, la planta o producto vegetal comprende plantas o productos vegetales de caña de azúcar, remolacha o remolacha azucarera, trigo, maíz, soja, patata, arroz o cebada, y el polipéptido tiene actividad que comprende actividad amilasa, y opcionalmente, el combustible comprende un bioetanol o una mezcla de gasolina-etanol. La divulgación provee combustibles que comprenden un polipéptido de la invención, en donde el polipéptido tiene actividad que comprende actividad de amilasa, en donde, opcionalmente, el combustible deriva de un material vegetal que, opcionalmente, comprende patata, soja (semilla oleaginosa), cebada, centeno, maíz, avena, trigo, remolacha o caña de azúcar y, opcionalmente, el combustible comprende un bioetanol, biopropanol, biobutanol, biodiésel y/o una mezcla de gasolina-etanol.

En otro aspecto, el material vegetal que comprende la enzima como se describe aquí se puede usar en un proceso industrial para producir combustible o energía. Las enzimas expresadas en las plantas se pueden agregar a, mezclar en o asperjar sobre la materia prima. En forma alternativa, las enzimas podrían ser expresadas directamente en la materia prima. En una forma de realización, el material vegetal que expresa las enzimas podría ser triturado, molido, calentado, etc., a fin de romper la integridad física de las células u órganos vegetales que contienen la enzima, liberando de este modo la enzima que entra en contacto con el sustrato. Los ejemplos de fuentes de material vegetal incluyen, pero no se limitan a, maíz, alfalfa, girasol, Brassica, soja, algodón, cártamo, cacahuate, sorgo, trigo, avena, centeno, mijo, cebada, arroz, coníferas, pastos, por ejemplo, pasto varilla y Miscanthus, cultivos de leguminosas, por ejemplo, guisante, frijol y soja, raíz/tubérculo almidonoso, por ejemplo, patata, batata, mandioca, taro, canna, y remolacha azucarera y similares.

La invención provee polipéptidos, que incluyen amilasas de la invención y métodos para la conversión de una biomasa o cualquier material lignocelulósico (por ejemplo, cualquier composición que comprende celulosa, hemicelulosa y lignina), en un combustible (por ejemplo, bioetanol, biopropanol, biobutanol, biopropanol, biometanol, biodiésel), además de piensos, alimentos y compuestos químicos. Por ejemplo, en un aspecto, una enzima de la divulgación tiene actividad de β -glucosidasa para liberar D-glucosa de dímeros celobiosa. En un aspecto, las enzimas tienen actividad exo- o endo-beta-glucanasa.

Por lo tanto, las composiciones y métodos de la invención proveen alternativas o adjuntos efectivos y sostenibles para el uso de productos basados en petróleo, por ejemplo, como una mezcla de un biocombustible tal como biometanol, bioetanol, biopropanol, biobutanol y similares, como combustible diésel, gasolina, querosén y similares. La invención provee organismos que expresan las enzimas de la invención para participación en ciclos químicos que implican conversión de biomasa natural. En un aspecto, las enzimas y métodos para la conversión se usan en ensamblajes enzimáticos para la despolimerización eficiente de polisacáridos, polímeros celulósicos y/o hemicelulósicos hasta fracciones de carbono metabolizables (por ejemplo, fermentables). La invención provee métodos para descubrir e implementar lo más efectivo de las enzimas para permitir estos nuevos procesos de "conversión de biomasa" y procesos industriales de energía alternativos.

Las composiciones y métodos de la invención se pueden usar para proveer alternativas o adjuntos efectivos y sostenibles para uso de productos basados en petróleo, por ejemplo, como una mezcla de bioetanol, biopropanol, biobutanol, biopropanol, biometanol y/o biodiésel y gasolina. La invención provee organismos que expresan enzimas de la invención para participación en ciclos químicos que implican conversión de biomasa natural. La invención

proporciona métodos para descubrir e implementar lo más efectivo de las enzimas para permitir estos nuevos e importantes procesos de "conversión de biomasa" y procesos industriales de energía alternativos.

La invención proporciona métodos, enzimas y mezclas de enzimas o "cócteles" de la invención, para procesar un material, por ejemplo un material de biomasa, que comprende un celooligosacárido, un oligómero de arabinoxilano, una lignina, una lignocelulosa, un xilano, un glucano, una celulosa y/o un azúcar fermentable que comprende poner en contacto la composición con un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención, en donde, en forma opcional, el material se deriva de un cultivo agrícola (por ejemplo, trigo, cebada, patatas, pasto varilla, madera de álamo), es un subproducto de la producción de un alimento o un pienso, es un producto de desecho lignocelulósico, o es un residuo vegetal o un residuo de papel o producto residual de papel y, en forma opcional, el residuo vegetal comprende tallos, hojas, cáscaras, cortezas, maíz o mazorcas de maíz, maíz de forraje, fibra de maíz, heno, rastrojo (por ejemplo rastrojo de arroz o rastrojo de trigo), bagazo de la caña de azúcar, pulpa de remolacha azucarera, pulpa de cítrico y cáscaras de cítricos, madera, aclarados de madera, astillas de madera, pulpa de madera, desechos de pulpa, desechos de madera, virutas de madera y aserrín, desechos de construcción y/o demolición y escombros (por ejemplo madera, virutas de madera y aserrín) y, en forma opcional, el residuo de papel comprende papel para fotocopias usado o descartado, papel para impresión por ordenador, papel de cuadernos, papel de anotadores, papel de mecanografía, diarios, revistas, cartón y materiales de embalaje basados en papel y materiales de papel reciclado. Asimismo, se pueden usar los desechos urbanos, por ejemplo la fracción de papel de desecho sólido municipal, desecho de madera municipal y desecho orgánico municipal, junto con otros materiales que contienen azúcar, almidón, y/o celulosa. En forma opcional, el procesamiento del material, por ejemplo el material de biomasa, genera un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol.

En forma alternativa, el polipéptido de la invención puede ser expresado en el material vegetal de biomasa o materia prima en sí misma.

Los métodos de la invención también incluyen tomar el material lignocelulósico convertido (procesado por las enzimas de la invención) y convertirlo en un combustible (por ejemplo un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol, o biodiésel) por fermentación y/o por síntesis química. En un aspecto, los azúcares producidos se fermentan y/o los productos no fermentables se gasifican.

Los métodos de la invención también incluyen convertir algas, aceites vegetales vírgenes, aceites vegetales de desecho, grasas animales y grasas (por ejemplo cebo, manteca de cerdo y grasa amarilla), o aguas residuales, usando enzimas de la invención y convirtiéndolos en combustible (por ejemplo un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol, o biodiésel) por fermentación y/o por síntesis química o conversión.

Las enzimas de la invención (incluyendo, por ejemplo, organismos, tales como microorganismos, por ejemplo, levadura o bacterias, que elaboran y, en algunos aspectos, secretan enzimas recombinantes de la invención) se pueden usar en o incluir/ integrar en cualquier etapa de cualquier proceso de conversión de biomasa, por ejemplo, en cualquier etapa, o diferentes etapas o incluir en todas las etapas, o todos los siguientes métodos de procesos de conversión de biomasa, o todas estas alternativas de biocombustible:

Combustión directa: la incineración de material por calor directo y es la tecnología de biomasa más simple; puede ser muy económica si la fuente de biomasa está cerca.

Pirólisis: es la degradación térmica de biomasa por calor en ausencia de oxígeno. En un aspecto, la biomasa se calienta a una temperatura entre aproximadamente 800 y 1400 grados Fahrenheit, pero no se introduce oxígeno para sustentar la combustión lo que resulta en la creación de gas, aceite combustible y carbón vegetal.

Gasificación: la biomasa se puede usar para producir metano a través de calentamiento o digestión anaeróbica. El Sintegas, una mezcla de monóxido de carbono e hidrógeno, se puede derivar de la biomasa.

Gas de basurero: se genera por la descomposición (digestión anaeróbica) de basura incinerada en los basureros. Cuando los residuos orgánicos se descomponen, generan gas que consiste de aproximadamente 50% de metano, el principal componente del gas natural.

Digestión anaeróbica: convierte la materia orgánica en una mezcla de metano, el componente principal del gas natural y dióxido de carbono. En un aspecto, la biomasa tal como agua residual (aguas de baño), estiércol, o desechos del procesamiento de alimentos, se mezcla con agua y se alimenta en un tanque de digestión sin aire.

50 Fermentación

Fermentación de alcohol: el alcohol combustible es producido al convertir masa celulósica y/o almidón en azúcar, fermentación del azúcar en alcohol, luego separar la mezcla de agua y alcohol por destilación. Las materias primas

tales como cultivos dedicados (por ejemplo, trigo, cebada, patata, pasto varilla, madera de álamo), residuos y desechos agrícolas (por ejemplo rastrojo de arroz, forraje de maíz, rastrojo de trigo, bagazo de la caña de azúcar, cáscaras de arroz, fibra de maíz, pulpa de remolacha azucarera, pulpa de cítrico y cáscaras de cítrico), desechos forestales (por ejemplo aclarados de madera dura y madera blanda, residuos de madera dura y madera blanda de las operaciones de madereras, virutas de madera y aserrín), desechos urbanos (por ejemplo fracción de papeles de desecho sólido municipal, residuos de madera municipales, residuos orgánicos municipales), desechos de madera (por ejemplo, desechos de aserrín, desechos del molido de pulpa, desechos de construcción, desechos de demolición, virutas de madera y aserrín) y residuos de papel u otros materiales que contienen azúcar, almidón, y/o celulosa se pueden convertir en azúcares y luego en alcohol por fermentación con levadura. En forma alternativa, los materiales que contienen azúcares se pueden convertir directamente en alcohol por fermentación.

Transesterificación: Una ejemplo de reacción para convertir aceite en biodiésel se denomina transesterificación. El proceso de transesterificación hace reaccionar un alcohol (como metanol) con los aceites triglicéridos contenidos en aceites vegetales, grasas animales o grasas recicladas, formando alquil ésteres de ácido graso (biodiésel) y glicerina. La reacción requiere calor y un catalizador básico fuerte, tal como hidróxido de sodio o hidróxido de potasio.

Biodiesel: el biodiésel es una mezcla de alquil ésteres de ácido graso elaborados a partir de aceites vegetales, grasas animales o grasas recicladas. El biodiésel se puede usar como un combustible para vehículos en su forma pura, pero por lo general se usa como un aditivo del diésel petróleo para reducir los niveles de partículas, monóxido de carbono, hidrocarburos y los tóxicos del aire de los vehículos que funcionan con diésel.

Hidrólisis: incluye hidrólisis de un compuesto, por ejemplo, una biomasa, tal como un material lignocelulósico, catalizado usando una enzima de la presente invención.

Cogeneración: es la producción simultánea de más de una forma de energía utilizando un solo combustible y una sola instalación. En un aspecto, la cogeneración de biomasa tiene crecimiento más potencial que la generación de biomasa sola porque la cogeneración produce tanto calor como electricidad.

En un aspecto, los polipéptidos de la invención tienen actividad enzimática (que incluyen por ejemplo, una actividad de amilasa) para generar un combustible (por ejemplo un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol, o biodiésel) a partir de un material orgánico, por ejemplo, una biomasa, tal como composiciones derivadas de plantas y animales, incluyendo cualquier cultivo agrícola u otra materia prima renovable, un residuo agrícola o un desecho animal, los componentes orgánicos de los residuos municipales e industriales, o residuos o escombros de construcción o demolición, o microorganismos tales como algas o levaduras.

En un aspecto, los polipéptidos de la invención se usan en procesos para convertir biomasa lignocelulósica en un combustible (por ejemplo un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol, o biodiésel) o, bien se usan en procesos para hidrolizar o digerir biomateriales de manera tal que se pueden usar como un combustible (por ejemplo un bioalcohol, por ejemplo, un bioetanol, biometanol, biobutanol o biopropanol, o biodiesel), o para hacer más fácil que la biomasa sea procesada en un combustible.

En un aspecto alternativo, los polipéptidos de la invención, incluyendo la mezcla de enzimas o "cócteles" de la invención, se usan en procesos para un proceso de transesterificación que hace reaccionar un alcohol (como etanol, propanol, butanol, metanol) con un aceite triglicérido contenido en un aceite vegetal, grasa animal o grasas recicladas, formando alquil ésteres de ácido graso (biodiésel) y glicerina. En un aspecto, el biodiésel se elabora a partir de aceite de soja o aceites de cocina reciclados. Las grasas de animales, otros aceites vegetales y otros aceites reciclados también se pueden usar para producir biodiesel, dependiendo de sus costos y disponibilidad. En otro aspecto, las mezclas de todas las clases de grasas y aceites se usan para producir un combustible biodiésel de la invención.

Las enzimas de la invención, incluyendo la mezcla de enzimas o "cócteles" de la invención, también se pueden usar en el refinamiento de glicerina. El subproducto de glicerina contiene un catalizador no reaccionado y jabones que se neutralizan con un ácido. El agua y alcohol se remueven para producir 50% a 80% de glicerina cruda. Los contaminante restantes incluyen grasas y aceites no reaccionados, que pueden ser procesos usando los polipéptidos de la invención. En grandes plantas de biodiésel de la invención, la glicerina puede ser purificada en forma adicional, por ejemplo, hasta un 99% o más de pureza, para las industrias cosméticas y farmacéuticas.

Los combustibles (que incluyen bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o biodiéseles) elaborados usando los polipéptidos de la invención, que incluyen la mezcla de enzimas o "cócteles" de la invención, se pueden usar con oxigenatos combustibles para mejorar las características de combustión. Agregar oxígeno genera combustión más completa, que reduce emisiones de monóxido de carbono. Este es otro beneficio ambiental de reemplazar combustibles de petróleo con biocombustibles (por ejemplo, un combustible de la invención). Un biocombustible elaborado usando las composiciones y/o métodos de esta invención se pueden

combinar con gasolina para formar una mezcla E10 (aproximadamente 5% a 10% de etanol y aproximadamente 90% a 95% de gasolina), pero se pueden usar en concentraciones superiores tal como E85 o en su forma pura. Un biocombustible elaborado usando las composiciones y/o métodos de esta invención se puede mezclar con diésel de petróleo para formar una mezcla B20 (20% de biodiésel y 80% de diésel de petróleo), si bien los niveles de la mezcla se pueden usar hasta B100 (biodiésel puro).

La invención también provee procesos para elaborar biocombustibles (incluyendo bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o biodiéseles) a partir de las composiciones que comprenden biomasa lignocelulósica. El material de biomasa de lignocelulosa puede ser obtenido a partir de cultivos agrícolas, como un subproducto de la producción de alimentos o piensos, o como productos de desechos lignocelulósicos, tales como residuos vegetales, residuos de papel o residuos o escombros de construcción y/o demolición. Los ejemplos de fuentes vegetales o residuos vegetales adecuados para tratamiento con los polipéptidos de la invención incluyen algas marinas, algas, granos, semillas, tallos, hojas, cáscaras, cortezas, mazorcas de maíz, forraje de maíz, rastrojo, pastos (por ejemplo, pasto de la India, tal como *Sorghastrum nutans*; o, pasto varilla, por ejemplo, *Panicum species*, tales como *Panicum virgatum*) y similares, así como también madera, astillas de madera, pulpa de madera y aserrín. Los ejemplos de residuos de papel adecuados para tratamiento con los polipéptidos de la invención incluyen papel de fotocopia desechado, papel para impresión por ordenador, papel de cuadernos, papel de anotadores, papel de mecanografía y similares, así como también diarios, revistas, cartón y materiales de embalaje basados en papel. Los ejemplos de residuos y escombros de construcción y demolición incluyen madera, trozos de madera, virutas de madera y aserrín.

En una forma de realización, las enzimas, que incluyen la mezcla de enzimas o "cócteles" de la invención y métodos de la invención se pueden usar junto con medios más "tradicionales" para elaborar etanol, metanol, propanol, butanol, propanol y/o diésel a partir de la biomasa, por ejemplo, como métodos que comprenden hidrolizar materiales lignocelulósicos al someter material lignocelulósico seco en un reactor a un catalizador que comprende por una solución diluida de un ácido fuerte y una sal metálica; esto puede disminuir la energía de activación, o la temperatura, de hidrólisis de celulosa para obtener mayores rendimientos de azúcar; véanse, por ejemplo, las patentes de los Estados Unidos Nos. 6.660.506 y 6.423.145.

Otro ejemplo de un método que incorpora el uso de enzimas de la invención, incluyendo las mezclas de enzimas o "cócteles" de la invención, comprende hidrolizar material lignocelulósico que contiene hemicelulosa, celulosa y lignina, o cualquier otro polisacárido que puede ser hidrolizado por una enzima de esta invención, al someter el material a un primer paso de la etapa de hidrólisis en un medio acuoso a una temperatura y una presión elegida para efectuar la despolimerización primaria de la hemicelulosa sin una despolimerización importante de celulosa en glucosa. Este etapa resulta en una suspensión en la cual la fase acuosa líquida contiene monosacáridos disueltos resultantes de la despolimerización de la hemicelulosa y una fase sólida que contiene celulosa y lignina. Un segundo paso de la etapa de hidrólisis puede comprender condiciones de tal manera que al menos una porción importante de la celulosa se despolimeriza, dicho etapa resultando en una fase acuosa líquida que contiene productos de despolimerización disueltos/solubles de celulosa. Véase, por ejemplo, la patente de los Estados Unidos No. 5.536.325. Las enzimas de la invención (que incluyen las mezclas o "cócteles" de enzimas de la invención) se pueden agregar en cualquier etapa de este proceso que sirve como ejemplo.

Otro ejemplo de un método que incorpora el uso de las enzimas de la invención, incluyendo la mezcla de enzimas o "cócteles" de la invención, comprende procesar un material de biomasa que contiene lignocelulosa por una o más etapas de hidrólisis con ácido diluido con aproximadamente 0,4% a 2% de ácido fuerte; y tratar un componente lignocelulósico sólido no reaccionado del material de biomasa hidrolizado con ácido por delignificación alcalina para producir precursores para termoplásticos biodegradables y derivados. Véase, por ejemplo, la patente de los Estados Unidos No. 6.409.841. Las enzimas de la invención se pueden agregar en cualquier etapa de este proceso que sirve como ejemplo.

Otro ejemplo de un método que incorpora el uso de enzimas de la invención, incluyendo la mezcla de enzimas o "cócteles" de la invención, comprende prehidrolizar material lignocelulósico en un reactor de prehidrólisis; añadir un líquido ácido al material lignocelulósico sólido para formar una mezcla; calentar la mezcla a temperatura de reacción; mantener la temperatura de reacción durante tiempo suficiente para fraccionar el material lignocelulósico en una porción solubilizada que contiene al menos aproximadamente 20% de la lignina del material lignocelulósico y una fracción sólida que contiene celulosa; remover una porción solubilizada de la fracción sólida mientras esté en o cerca de la temperatura de reacción en donde la celulosa en la fracción sólida se torna más propensa a digestión enzimática; y recuperar una porción solubilizada. Véase, por ejemplo, la patente de los Estados Unidos No. 5.705.369. Las enzimas de la invención pueden ser agregadas en cualquier etapa de este proceso que sirve como ejemplo.

La invención provee métodos para elaborar composiciones de combustible para motores (por ejemplo, para motores de ignición por chispa) basadas en hidrocarburos líquidos mezclados con un alcohol de grado combustible elaborado usando una enzima o un método de la invención. En un aspecto, los combustibles elaborados por el uso de una enzima de la invención comprenden, por ejemplo, mezclas de etanol con gas de carbón líquido o gas natural líquido.

En un aspecto, un cosolvente es un 2- metiltetrahidrofurano (MTHF) derivado de biomasa. Véase, por ejemplo, la patente de los Estados Unidos No. 6.712.866

5 En un aspecto, los métodos de la invención para la degradación enzimática de lignocelulosa, por ejemplo, para la producción de biocombustibles (incluyendo bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o biodiéseles) a partir de material lignocelulósico, también pueden comprender el uso de tratamiento ultrasónico del material de biomasa; véase, por ejemplo, la patente de los Estados Unidos No. 6.333.181.

10 En otro aspecto, los métodos de la invención para producir biocombustibles (incluyendo bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o biodiéseles) a partir de un sustrato celulósico comprende proveer una mezcla de reacción en la forma de una suspensión que comprende sustrato celulósico, una enzima de esta invención y un agente de fermentación (por ejemplo, en un recipiente de reacción, tal como un bioreactor de alimentación de sólidos semicontinua), y la mezcla de reacción se hace reaccionar bajo condiciones suficientes para iniciar y mantener una reacción de fermentación (como se describe, por ejemplo, en la solicitud de patente de los Estados Unidos No. 20060014260). En un aspecto, los cálculos experimentales o teóricos pueden determinar una frecuencia de alimentación óptima. En un aspecto, las cantidades adicionales del sustrato celulósico y la enzima se provee en el recipiente de reacción a un intervalo(s) de acuerdo con la frecuencia de alimentación optimizada.

20 Un un ejemplo de un proceso para elaboración de biocombustibles (incluyendo bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o biodiéseles) de la invención se describe en las publicaciones de solicitudes de patente de los Estados Unidos Nos. 20050069998; 20020164730; y en un aspecto comprende las etapas de triturar la biomasa lignocelulósica (por ejemplo, a un tamaño de 15 - 30 mm), someter el producto obtenido a pretratamiento por explosión de vapor (por ejemplo, a una temperatura de 190 - 230°C) entre 1 y 10 minutos en un reactor; recolectar el material pretratado en un ciclón o producto de fabricación relacionado; y separar las fracciones líquidas y sólidas por filtración en un filtro prensa, introduciendo la fracción sólida en un depósito de fermentación y agregando una o más enzimas de la invención, por ejemplo, una amilasa, enzima (por ejemplo, disueltas en regulador de citrato, pH 4,8).

25 Otro ejemplo de un proceso para elaborar biocombustibles (incluyendo bioalcoholes tales como bioetanoles, biometanoles, biobutanoles o biopropanoles, o, biodiéseles) de la invención que comprende bioetanoles, biometanoles, biobutanoles o biopropanoles usando enzimas de la invención comprende pretratar un material de partida que comprende una materia prima lignocelulósica que comprende al menos hemicelulosa y celulosa. En un aspecto, el material de partida comprende patata, soja (semilla oleaginosa), cebada, centeno, maíz, avena, trigo, remolacha o caña de azúcar o un componente o desecho o subproducto de la producción de alimentos o piensos. El material de partida ("materia prima") se hace reaccionar a condiciones que rompen la estructura de las fibras de la planta para efectuar al menos una hidrólisis parcial de la hemicelulosa y celulosa. Las condiciones de ruptura pueden comprender, por ejemplo, someter el material de partida a una temperatura promedio de 180°C a 270°C a pH de 0,5 a 2,5 durante un período de aproximadamente 5 segundos a 60 minutos; o, temperatura de 220°C a 270 °C, a pH de 0,5 a 2,5 durante un periodo de 5 segundos a 120 segundos, o equivalente. Esto genera una materia prima con mayor accesibilidad para ser digerida por una enzima, por ejemplo, una enzima celulasa de la invención. Patente de los Estados Unidos No. 6.090.595.

40 Los ejemplos de condiciones para usar las enzimas de la invención en la hidrólisis de material lignocelulósico incluyen reacciones a temperaturas entre aproximadamente 30°C y 48°C, y/o un pH entre aproximadamente 4,0 y 6,0. Otros ejemplos de condiciones incluyen una temperatura entre aproximadamente 30°C y 60°C y un pH entre aproximadamente 4,0 y 8,0.

45 Las amilasas de la invención se pueden usar en la conversión de biomasa en combustibles y en la producción de etanol, por ejemplo, como se describe en las solicitud de patente PCT Nos. WO0043496 y WO8100857. Las amilasas de la invención se pueden usar para producir azúcares fermentables y biomasa que contiene glucano que puede ser convertidos en etanol de combustible.

Composiciones farmacéuticas, desinfectantes y suplementos dietarios.

50 La invención también provee composiciones farmacéuticas, desinfectantes y suplementos dietarios (por ejemplo, auxiliares dietarios) que comprenden enzimas de la invención (por ejemplo, enzimas que tienen actividad de amilasa). En un aspecto, las composiciones farmacéuticas y suplementos dietarios (por ejemplo, auxiliares dietarios) son formuladas para ingestión oral, por ejemplo, para mejorar la digestibilidad de alimentos pienso que tengan un alto contenido de almidón, celulosa o un componente lignocelulósico.

55 Los compuestos de tratamientos periodontal pueden comprender una enzima de la invención, por ejemplo, como se describe en la patente de los Estados Unidos No. 6.776.979. Las composiciones y métodos para el tratamiento o la profilaxis del síndrome de intestino ácido pueden comprender una enzima de la invención, por ejemplo, como se describe en la patente de los Estados Unidos No. 6.468.964.

En otro aspecto, los vendajes para heridas, implantes y similares comprenden enzimas antimicrobianas (por ejemplo, que actúan como antibióticos), que incluyen una enzima de la invención. Las enzimas de la invención también se pueden usar en vendajes de alginato, vendajes de barrera antimicrobiana, vendajes para quemaduras, vendas de compresión, herramientas de diagnóstico, vendajes en gel, vendajes hidroselectivos, vendajes 5 hidrocelulares (espuma), vendajes hidrocoloides, vendajes I.V, cubiertas para incisiones, vendajes de baja adherencia, vendajes de absorción de olores. vendas en suspensión, vendajes postoperatorios, control de cicatrices, cuidados de la piel, vendajes de película transparente y/o torniquete de heridas. Las enzimas de la invención se pueden usar en limpieza de heridas, preparación de la superficie de la herida, para tratar úlceras de presión, úlceras en piernas, quemaduras, úlceras de pies diabéticos, cicatrices, fijación IV, heridas quirúrgicas y heridas menores. 10 Las enzimas de la invención se pueden usar en composiciones estériles de desbridamiento enzimático, por ejemplo, ungüentos. En varios aspectos, la celulasa es formulada como una tableta, gel, píldora, implante, líquido, atomizador, polvo, alimento, pellas de piensos o como una formulación encapsulada.

Aplicaciones de biodefensa

En otros aspectos, las enzimas de la invención, por ejemplo, glucoamilasas, se pueden usar en biodefensa, por 15 ejemplo, destrucción de esporas o bacterias. El uso de las enzimas de la invención en aplicaciones de biodefensa ofrecen un beneficio significativo, ya que pueden ser desarrolladas muy rápidamente contra cualquier agente biológico potencialmente nocivo actualmente desconocido o del futuro. Asimismo, las enzimas de la invención se pueden usar para la descontaminación de los ambientes afectados. En un aspecto, la invención provee un agente de biodefensa o biodesintoxicante que comprende un polipéptido de la invención que tiene actividad de amilasa.

20 La invención se describirá además con referencia a los siguientes ejemplos; sin embargo, debe entenderse que la invención no se limita a dichos ejemplos.

Ejemplos

Ejemplo 1: Identificación y caracterización de α -amilasas termoestables

Los siguientes ejemplos describen ejemplos de métodos para determinar si un polipéptido está dentro del alcance de 25 la invención. Se pueden llevar a cabo programas de cribado bajo condiciones de pH bajo y neutro. Los análisis bioinformáticos y de secuencia de ADN pueden clasificar las amilasas.

Estudios bioquímicos

El análisis bioquímico de clones genómicos de amilasa se puede usar para determinar si cualquiera tiene un pH 30 óptimo menor a pH 6. Los lisados de estos clones genómicos pueden ser evaluados en cuanto a la tolerancia térmica por incubación a 70°C, 80°C, 90°C o 100°C durante 10 minutos y la medición de actividad residual a un pH de 4,5. Aquellos clones que retienen > 50% de actividad después del tratamiento térmico a 80°C son elegidos para análisis adicional. Estos clones se pueden incubar a 90°C durante 10 minutos a un pH de 6,0 y 4,5 y se analizan su actividad residual a un pH de 4,5. La actividad térmica de los clones con actividad residual después del tratamiento térmico a 90°C a pH 4,5 se puede medir a temperatura ambiente, 70°C y 90°C a pH 4,5.

35 En un aspecto, las enzimas degradantes de almidón se criban en cuanto a la actividad sobre almidón sin purificar y almidón "resistente" para cualquier forma de actividad de hidrolasa, incluyendo actividad de amilasa, pululanasa, ciclodextrina glucosiltransferasa, glucoamilasa o cualquier otra actividad de glucosidasa. En un aspecto, las enzimas activas identificadas se caracterizan, por ejemplo, por actividad específica y especificidad por azúcares ramificados y/o oligosacáridos más largos.

40 En un aspecto, los aislados de hongos se investigan en cuanto a las enzimas novedosas usando, por ejemplo, sondas y/o procesos de descubrimiento de la divulgación, por ejemplo, cribando bibliotecas de ADNg y/o ADNc obtenidas de animales, microorganismos o insectos, por ejemplo, los contenidos de las tripas de insectos que atacan y consumen granos almacenados, incluyendo el cribado de bibliotecas ambientales para las enzimas sub-representadas. Las combinaciones de enzimas se pueden evaluar.

45 En un aspecto, se realizan ensayos de estabilidad: por ejemplo, las enzimas purificadas en su propio regulador de almacenamiento, comparado con la actividad de enzima fresca, si se analiza; por ejemplo, alícuotas de 20 μ l almacenadas a temperaturas inferiores, por ejemplo, entre aproximadamente 4°C a -20°C a -80°C; y en un aspecto, la actividad es analizada nuevamente en almidón granulado y soluble en forma mensual.

50 Las amilasas se pueden evaluar bajo una variedad de condiciones. En los siguientes protocolos se puede usar maíz dentado amarillo N°2 como una fuente de almidón.

Ejemplo de un ensayo de licuefacción

Una suspensión de almidón que comprende 35% de sólidos secos ("DS") se somete a licuefacción primaria durante cinco minutos bajo diferentes temperaturas en el rango de 95°C a 119°C (por ejemplo, hasta aproximadamente 110°C), con una concentración enzimática de entre 0,2 a 0,8 gramo/kilogramo (g/kg) de DS de almidón, con adición de calcio en el rango de entre cero y 30 partes por millón (ppm), a un pH 4,0 a pH 5,6. La licuefacción secundaria comprende condiciones de 120 minutos a 95°C.

Ejemplo de ensayo de sacarificación

La sacarificación se analiza inicialmente usando 35% de sólidos secos ("DS") (suspensión de almidón) y glucoamilasa AMG 300 L (Novozymes A/S, Dinamarca) a razón de 0,225 AGU/gramo de DS (AGU= unidades de amiloglicosidasa o glucoamilasa), pH 4,3, a 60°C durante 44 horas.

10 En un aspecto, los ejemplos de amilasas de la invención se usan en un rango de dosificación entre 0,5 a 0,7 kg / MT de DS de almidón.

La invención provee métodos para elaborar endulzantes nutritivos usando las enzimas de la invención, por ejemplo, procesos que comprenden los protocolos de licuefacción y sacarificación anteriormente descritos usando cualquier amilasa de la invención. En un aspecto, el rango de dosificación para una enzima de la invención en estos procesos es entre aproximadamente 0,5 y 0,7 gramos por kg de DS de almidón, una temperatura de chorro (por ejemplo, usando un quemador a chorro) de aproximadamente 110°C, pH 4,5, sin adición de calcio.

Producción de etanol por molienda seca

La invención provee métodos para producción de etanol por molienda seca usando las enzimas de la invención. Al evaluar las enzimas de la invención para uso en la producción de etanol por molienda seca, particularmente, la licuefacción de harina de maíz molido en seco, se puede usar un reactor a escala de laboratorio con harina de maíz obtenido de una molienda seca comercial. La amilasa TERMAMYL^{MR} SC (Novozymes A/S, Dinamarca) se puede usar como punto de referencia competitivo. En aspectos alternativos, las condiciones óptimas son 85°C, pH 5,7. Se pueden estudiar cinco variables independientes: temperatura (en un rango entre 80°C y 100°C), dosis enzimática entre 0,2 y 1,0 g/kg de almidón, pH 4,4 a 6,0, calcio en un rango entre 0 ppm a 200 ppm y una contracorriente reciclada entre aproximadamente 0% y 40%.

30 A 95°C, en algunas formas de realización, las amilasas de la invención pueden reducir la viscosidad de una harina de maíz molida en seco más rápidamente que la amilasa TERMAMYL^{MR} SC (Novozymes A/S, Dinamarca) en sus condiciones óptimas, inclusive a 85°C. La tasa de reducción de la viscosidad por amilasas puede ser influenciado más por dosis de la enzima y la temperatura. Los intervalos óptimos alternativos pueden estar en el rango de 0.4 a 0,6 g/kg de almidón, con una temperatura óptima a 95°C.

En algunas formas de realización, las amilasas de la invención pueden ser efectivas a un pH menor y una temperatura mayor que la de la amilasa TERMAMYL^{MR} SC (Novozymes A/S, Dinamarca) a un pH en el rango entre pH 4,4 y pH 5,6. La adición de calcio puede tener un efecto mínimo sobre la reducción de la tasa de viscosidad a 95°C.

35 Las actividades de las enzimas purificadas se compararon en diferentes reguladores de almacenamiento, como se enlista a continuación, después de 1 semana de incubación a 37°C. El regulador con la menor pérdida de actividad comparada con la actividad de la misma enzima mantenida a +4°C se seleccionó como el regulador de almacenamiento de elección. Los ejemplos de las condiciones de ensayo para evaluar la actividad de amilasa, por ejemplo, para determinar si un polipéptido de la invención retiene actividad bajo condiciones particulares, incluyen (MPB: Metilparabeno):

- 20% de Glucosa, 0,1% de MPB en PBS. pH 7,2 a 37°C
- 20% de Glucosa, 0,1% de MPB en regulador de acetato, pH 5 a 37°C
- 20% de Sacarosa, 0,1% de MPB en PBS, pH 7,2 a 37°C
- 20% de Sacarosa, 0,1% de MPB en regulador de acetato, pH 5 a 37°C
- 45 ▪ 0,1% de MPB en PBS, pH 7,2 a 37°C
- 0,1% de MPB en regulador de acetato, pH 5 a 37°C
- PBS, pH 7,2 a 37°C
- Regulador de acetato, pH 5 a 37°C
- 50 ▪ 20% de Glucosa, 0,1% de MPB en PBS, pH 7,2 a 4°C
- 20% de Glucosa, 0,1% de MPB en regulador de acetato, pH 5 a 4°C
- 20% de Sacarosa, 0,1% de MPB en PBS, pH 7,2 a 4°C
- 20% de Sacarosa, 0,1% de MPB en Regulador de acetato, pH 5 a 4°C
- 20% de Maltosa en PBS (fosfato de sodio 50 mM pH 7,5; NaCl 100 mM)
- 0,1% de MPB en PBS, pH 7,2 a 4°C

- 0,1% de MPB en regulador de acetato, pH 5 a 4°C
- PBS, pH 7,2 a 4°C
- Regulador de acetato, pH 5 a 4°C

5 Los ejemplos de enzimas de la invención que han sido analizadas bajo estos ejemplos de condiciones de ensayo incluyen: la SEQ ID NO: 52. se seleccionó 20% de sacarosa en 1X PBS, pH 7 con 0,1% de metil parabeno (SMP*) y para la SEQ ID NO: 18, se seleccionó 20% de sacarosa en acetato de sodio 50 mM con cloruro de sodio 150 mM y 0,1% de metil parabeno (SMA**). La Figura 13 ilustra una comparación de la velocidad inicial (usando Vmax) del ejemplo de la enzima de la SEQ ID NO: 70 en diferentes reguladores de almacenamiento (como se indica en la figura); 0,15 µg de proteína total por reacción; la Figura 14 ilustra una comparación de la velocidad inicial (usando Vmax) del ejemplo de la enzima de la SEQ ID No: 76 en diferentes reguladores de almacenamiento (como se indica en la figura); 0,5 µg de proteína total por reacción; la Figura 15 ilustra una comparación de la velocidad inicial (usando Vmax) del ejemplo de la enzima de la SEQ ID No: 4 en diferentes reguladores de almacenamiento (como se indica en la Figura); la Figura 16 ilustra una comparación de velocidad inicial (usando Vmax) del ejemplo de la enzima de la SEQ ID NO: 66 en diferentes reguladores de almacenamiento (como se indica en la figura); 0,03 µg de proteína total por reacción; la Figura 17 ilustra una comparación de velocidad inicial (usando Vmax) del ejemplo de la enzima de la SEQ ID NO: 52 en diferentes reguladores de almacenamiento (como se indica en la figura); 0,025 µg de proteína total por reacción; la Figura 18 ilustra una comparación de velocidad inicial (usando nmoles de glucosa por min por µg de enzima) del ejemplo de la enzima de ejemplo SEQ ID No: 28 en diferentes tampones de almacenamiento (como se indica en la figura); la Figura 19 ilustra una comparación de velocidad inicial (usando nmoles de glucosa por min por µg de enzima) de la enzima de la SEQ ID NO: 26 en diferentes reguladores de almacenamiento (como se indica en la figura); la Figura 20 ilustra una comparación de velocidad inicial (usando nmoles de glucosa por min por µg de enzima) del ejemplo de la enzima de la SEQ ID NO: 18 en diferentes reguladores de almacenamiento (como se indica en la figura).

25 Los ensayos Bradford y BCA así como también SDS PAGE y Absorbancia a A280 se usaron para determinar la concentración de las proteínas purificadas. La solución de amilasa estandarizada (Sigma A6211) se usó como referencia en estos ensayos y BSA. La concentración por BCA y A280 fue similar, pero no por Bradford usando BSA como estándar.

30 En estos ensayos, la proteína (enzima) se purificó por precipitación; interacción hidrófoba; exclusión por tamaño; intercambio de iones; afinidad; y en un protocolo de ejemplo, la proteína (enzima) se purificó por precipitación, interacción hidrófoba, intercambio iónico y en una etapa final, afinidad (cromatografía).

35 Las SEQ ID NO: 18, SEQ ID NO: 4 y SEQ ID NO: 26 de ejemplo se purificaron ya sea por precipitación con sulfato de amonio (SEQ ID NO: 26) o etanol (SEQ ID NO: 18 y SEQ ID NO: 4). Los sobrenadantes liofilizados a partir de cultivos que expresan las enzimas de interés se resuspendieron en agua a una concentración de aproximadamente 1 µg / 5ml. Las SEQ ID NO: 18 y SEQ ID NO: 4 se precipitaron por adición de etanol frío (2 volúmenes de solución de proteína). La SEQ ID NO: 26 se precipitó en 80% de solución de sulfato de amonio. Después de la precipitación, las suspensiones de proteína se dializaron una vez contra agua y posteriormente contra regulador de almacenamiento.

40 Los ejemplos de estabilidad de las enzimas purificadas de esta divulgación se analizaron ya sea en PBS a pH 7 o PBS a pH 7,5 (en este caso, los parámetros fueron fosfato de sodio 50 mM; NaCl100 mM) o en acetato de sodio 50 mM, a pH 5,2 con adición de NaCl 150 mM. Se añadió glucosa (20%) o sacarosa (20%) a los reguladores así como un antimicrobiano, 0,1% de metil parabeno (Sigma). Las enzimas se incubaron en sus respectivos reguladores a 37°C durante 1 semana y la actividad de la enzima se analizó y comparó con la actividad de la misma enzima almacenada a +4°C.

45 La concentración de cada enzima purificada se estimó usando 4 métodos diferentes: ensayo Bradford con amilasa Sigma A-6211 como estándar, ensayo BCA con Sigma A-6211 como estándar, absorbancia a 280 nm en urea 8M y medición densitométrica de proteína teñida después de SDS PAGE, usando amilasa Sigma A-6211 como referencia. Los números obtenidos con estos métodos diferentes para cada una de las proteínas purificadas se presentan en la Tabla 3.1. Para la cuantificación final, se usó la concentración obtenida al medir la absorbancia a 280 nm en urea 8M.

50 Un resumen de purificación a gran escala para estos ensayos es:

Resumen de la cantidad recuperada con base en el valor de absorbancia			
	mg/ml	Volumen (ml)	Total en mg
SEQ ID NO:18	53,94	111	5987,34

ES 2 575 912 T3

Resumen de la cantidad recuperada con base en el valor de absorbancia			
	mg/ml	Volumen (ml)	Total en mg
SEQ ID NO:26	39,23	86,8	3405,164
SEQ ID NO:4	70,01	55,4	3878,554

Resumen de pureza				
	Antes de purificación		Después de purificación	
	A260/A280	Relación de expresión	A260/A280	Relación de expresión
SEQ ID NO:18	0,52	91%	0,59	92%
SEQ ID NO:26	0,56	85%	0,65	86%
SEQ ID NO:4	0,67	91%	0,73	85%

Resumen de actividad				
	Almidon sin purificar		Soluble	
	Purificado	Sin purificar	Purificado	Sin purificar
SEQ ID NO:18	10,7228 ± 0,7925	6,4828 ± 0,5681	40,6725 ± 1,7016	24,7354 ± 3,821
SEQ ID NO:26	5,7837 ± 0,2271	4,0691 ± 0,2864	17,4833 ± 0,7831	14,1849 ± 0,3761
SEQ ID NO:4	6,47 ± 0,34	4,074 ± 0,35	127,37 ± 1,78	116,52 ± 1,27

Estimación de la concentración enzimática por diferentes métodos

	Bradford	BCA (A6211)	Absorbancia (Urea)
SEQ ID NO: 18	71,10	57,64	53,94
SEQ ID NO: 26	43,98	57,64	39,29
SEQ ID NO: 4	69,20	57,64	70,01

5

Ejemplo 2: Amilasas termoestables activas a pH alcalino

El siguiente ejemplo describe un ejemplo de método para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, es una amilasa termoestable.

10 Las formulaciones comerciales para lavado automático de vajillas (ADW) se pueden usar para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, es una amilasa termoestable. Los estudios pueden incluir la identificación de amilasas de pH alto; y las enzimas que tienen la habilidad para degradar almidón. Los análisis de secuencia de ADN y bioinformáticos pueden clasificar muchos de estos genes como amilasas, o tener otras especificidades enzimáticas, por ejemplo, neopolulanasas, amilopolulanasas y amilomaltasas.

Estudios bioquímicos

Un ejemplo de un método para determinar si un polipéptido es una amilasa termoestable, es analizar la actividad donde la enzima puede hidrolizar almidón a pH alcalino, por ejemplo, hasta un pH 10 y aproximadamente 50°C.

- 5 La proteína soluble se purifica hasta homogeneidad y se mide la actividad específica (unidades/mg, donde 1 unidad = μmol azúcares reductores/min) a pH 8 y pH 10 (40°C y 50°C) usando 2% de almidón en regulador. La actividad específica se puede determinar removiendo las muestras en diferentes momentos durante una reacción de 30 minutos y analizando los azúcares reductores. La velocidad inicial se puede determinar ajustando las curvas de progreso a una ecuación lineal.

Estudios de estabilidad

- 10 La estabilidad en presencia de la formulación ADW se puede medir mediante análisis bioquímico. El punto de referencia para estos estudios puede ser una enzima comercial en la matriz de formulación. La actividad medida después de la incubación se puede expresar como porcentaje de la actividad original.

Pruebas de lavado

- 15 Las pruebas de lavado usando portaobjetos recubiertos con almidón se pueden llevar a cabo para medir el desempeño de cada una de una enzima purificada de la invención comparada con una amilasa comercial. Los portaobjetos recubiertos con almidón de espagueti se pueden preparar de acuerdo con el siguiente protocolo: se colocan dos portaobjetos previamente pesados recubiertos con almidón uno sobre otro en un tubo cónico de 50 mL y se añaden 25 mL de solución ADW, +/-enzima por tubo. Los tubos se incuban durante 20 minutos a 50°C con rotación suave en un carrusel vertical. Después del período de incubación, se enjuagan inmediatamente los portaobjetos en agua se secan toda la noche en horno. Todas las ensayos se realizan por duplicado y se corre la enzima comercial como control positivo. Los resultados se pueden expresar como % neto de almidón removido, por ejemplo, el % de almidón removido en ADW con enzima, menos el % de almidón removido solo en ADW.

Ejemplo 3: Optimización génica

- 25 El siguiente ejemplo describe un método que sirve de ejemplo para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, evaluar el desempeño de la enzima en presencia del desempeño de ADW.

- 30 Las propiedades de las enzimas pueden ser mejoradas por varias estrategias de evolución, incluyendo tecnologías de Gene Site Saturation Mutagenesis^{MR} (GSSM^{MR}) y GeneReassembly^{MR} (Diversa Corporation, San Diego, CA). Tales técnicas se aplicarán a los ácidos nucleicos de amilasa de la invención con el fin de generar conjuntos de variantes que se pueden cribar para mejorar el desempeño. En un aspecto, las moléculas progenitoras para evolución incluyen cualquier ácido nucleico de la invención, por ejemplo, secuencias que codifican las SEQ ID NO: 2; SEQ ID NO: 4; SEQ ID NO: 6, etc.

Se puede usar un cribado de alto rendimiento (HTS) para evaluar el desempeño de la enzima en presencia del desempeño de ADW. HTS se puede automatizar y ha mostrado resultados consistentes para las amilasas progenitoras.

- 35 **Ejemplo 4:** Caracterización de α -amilasas que tienen actividad a pH alcalino

El siguiente ejemplo describe métodos de ejemplo para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, tiene actividad de alfa-amilasa a pH alcalino.

- 40 Las amilasas de la invención que tienen actividad a pH alcalino pueden ser caracterizadas usando cinéticas sobre 2% de almidón a un pH de 8 y 10 (40°C y 50°C). Se puede definir 1 unidad de actividad como liberación de 1 μmol de azúcares reductores por minuto.

Ejemplo 5: Ensayo de actividad de amilasa; ensayos de extremos reductores de BCA

El siguiente ejemplo describe un método de ejemplo para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, un ensayo de extremos reductores de BCA. La actividad de amilasa de los clones de interés se puede determinar usando la siguiente metodología.

- 45 1. Preparar 2 soluciones de sustrato, de la siguiente manera:

ES 2 575 912 T3

a) 2% de almidón soluble (almidón patata o de maíz granulado) solución a pH 8 por disolución de 2 g de almidón de patata en 100 ml de fosfato de sodio 100 mM, pH 8.

b) 2% de almidón soluble (patata), pH 10, disolviendo 2 g de almidón de patata en 100 ml de carbonato de sodio 100 mM.

5 Calentar ambas soluciones en un baño de agua en ebullición, mientras se mezcla, durante 30 - 40 minutos hasta que se disuelva el almidón.

2. Preparar la Solución A de 64 mg/mL de carbonato de sodio monohidratado, 24 mg/ml de bicarbonato de sodio y 1,95 mg/ml de BCA (sal disódica de 4,4'- dicarboxi-2,2'- biquinolina (Sigma Chemical cat # D-8284)). Agregar dH₂O.

10 3. Preparar una Solución B por combinación de 1,24 mg/ml de sulfato cúprico pentahidratado y 1,26 mg/ml de L-serina. Añadir la mezcla a dH₂O.

4. Preparar un reactivo de trabajo en una proporción 1:1 de soluciones A y B.

15 5. Preparar una solución estándar de maltosa de maltosa 10 mM en dH₂O, en donde la maltosa 10 mM se combina con 2% de almidón soluble a pH deseado a una concentración final de 0, 100, 200, 300, 400, 600 μM. La curva estándar se generará para cada conjunto de puntos de tiempo. Como la curva se determina al agregar 10 μl de los estándares al reactivo de trabajo, elaborará a 0, 1, 2, 3, 4, 6 nmoles de maltosa.

6. Tomar una alícuota de 1 ml de solución de sustrato en tubos de microcentrífuga, equilibrar a la temperatura deseada (5 min) en un bloque de calor o baño de agua caliente. Agregar 50 μl de solución enzimática al interior de la tapa del tubo.

20 7. Mientras la solución se equilibra, mezclar 5 ml tanto de solución A como de solución B. Colocar una alícuota de 100 μl en una placa de PCR de 96 pozos. Colocar la placa en hielo.

8. Después de 5 minutos de equilibrar la temperatura, cerrar la tapa de los tubos, invertir y agitar en forma de vórtice 3 veces. En forma inmediata, colocar una alícuota de 10 μl en la placa a t=0 (punto de tiempo cero). Dejar la mezcla enzimática en un bloque de calor y colocar una alícuota de 10 μl en cada punto tiempo deseado (por ejemplo 0, 5, 10, 15, 20, 30 minutos).

25 9. Asegurar que 12 pozos queden vacíos (únicamente para colocar alícuotas del reactivo de trabajo) para la adición de 10 μl de estándares, para la curva estándar.

10. Cuando todos los puntos tiempo se reúnen y se añaden los estándares, cubrir la placa y calentar a 80°C durante 35 min. Enfriar la placa sobre hielo durante 10 min. Agregar 100 μl de H₂O a todos los pozos. Mezclar y colocar alícuotas de 100 μl en una placa de 96 pozos de fondo plano y leer la absorbancia a 560 nm.

30 11. Dejar los puntos de tiempo de cada muestra en cero contra su propio t=0 (restar el valor promedio de t=0 a A560 de los otros valores promedio de A560). Convertir el A560_(experimental) a μmoles (Dividir A560_(experimental) por la pendiente de la curva estándar (A560/μmoles). Generar una pendiente de los puntos de tiempo y los μmoles (en μmole/min), multiplicar por 100 (ya que el valor en μmoles solamente da cuenta de los 10 μl usados en el ensayo y no de la cantidad elaborada en el rxn de un 1ml). Para obtener la actividad específica, se divide la pendiente (en μmoles/min) por los mg de proteína. Todos los puntos deben medirse mínimo por duplicado, siendo mejor por triplicado.

Dividir la concentración de proteína (mg/ml) por cualquier dilución para obtener mg utilizados en ensayo.

Dividir la pendiente anterior por mg utilizados en el ensayo para obtener la actividad específica.

$$\text{Actividad específica} = 24,93 \text{ } \mu\text{moles/min/mg}$$

40 Véase por ejemplo, Dominic W.S. Wong, Sarah B. Batt y George H. Robertson (2000) J. Agric. Food Chem. 48: 4540 - 4543; Jeffrey D. Fox y John F. Robyt, (1991) Anal. Biochem. 195, 93 - 96.

Ejemplo 6: Cribado para la actividad de α-amilasa

45 El siguiente ejemplo describe un método de ejemplo para determinar si un polipéptido está dentro del alcance de la invención. La actividad de amilasa de los clones se puede evaluar por una cantidad de métodos conocidos en el arte. La siguiente es una metodología de ejemplo. El número de plaquetas cribados, por placa, puede ser de

ES 2 575 912 T3

aproximadamente 10000 pfu. Para cada biblioteca de ADN: se deben cribar al menos 50000 plaquetas por biblioteca aislada y 200000 plaquetas por biblioteca no aislada dependiendo del título pfu para el lisado amplificado λ Zap Express.

Determinación de títulos de la biblioteca lambda

- 5 1) Se añadieron μ L del patrón de la biblioteca amplificada Lambda Zap Express a 600 μ L de células MRF' de *E. coli* ($OD_{600} = 1,0$). Para diluir el patrón de MRF', se usa $MgSO_4$ 10 mM.
2) Incubar a 37°C durante 15 minutos.
3) Transferir la suspensión a 5 - 6mL de agar de superficie NZY a 50°C y mezclar suavemente.
4) Verter inmediatamente la solución de agar sobre una placa grande (150 mm) con medio NZY.
- 10 5) Permitir que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego invertir la placa.
6) Incubar la placa a 39°C durante 8 - 12 hora.
7) El número de plaquetas es aproximado. Se determino que el título de fago produjo 10000 pfu/placa. Diluir una alícuota de la biblioteca fagos con regulador SM de ser necesario.
- 15 Cribado del sustrato
1) Agregar Lambda Zap Express (50000 pfu) de la biblioteca amplificada a 600 μ L de células MRF' de *E. coli* ($OD_{600}=1,0$). Para bibliotecas no ambientales, preparar 4 tubos (50000 pfu por tubo).
2) Incubar a 37°C durante 15 minutos.
3) Mientras se incuban las suspensiones de fago/célula, se añade 1,0 mL de sustrato de almidón rojo (1,2% p/v) a 6,0 mL de agar de superficie NZY a 50°C y se mezcla completamente. Se mantiene la solución a 50°C hasta que se necesite.
4) Transferir 1/5 (10000 pfu) de la suspensión de células a la solución de sustrato/agar de superficie y mezclar suavemente.
5) Se vierte inmediatamente la solución en una placa grande (150 mm) de medio NZY.
- 25 6) Permitir que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego invertir la placa.
7) Repetir los procedimientos 4 - 6 cuatro veces para el resto de la suspensión de células (1/5 de la suspensión cada vez).
8) Incubar las placas a 39°C durante 8 - 12 horas.
- 30 9) Se observan las zonas de aclaramiento (halos) de la placa alrededor de las plaquetas.
10) Las plaquetas con halos se separan del agar y se transfieren a un micro tubo esterilizado. Una punta de pipeta de orificio grande de 200 μ L funciona bien para remover (separar) el tapón de agar que contiene la plaqueta deseada.
11) Los fagos se resuspenden 500 μ l de regulador SM. Se añaden 20 μ L de cloroformo para inhibir cualquier crecimiento adicional de células.
- 35 12) Se incuba la suspensión pura de fagos a temperatura ambiente durante 4 horas o toda la noche antes de la siguiente etapa.

Aislamiento de clones puros

- 1) Se añaden 10uL de la suspensión de fagos resuspendida en 500 μ L de células MRF' de *E. coli* ($OD_{600} = 1,0$).

ES 2 575 912 T3

- 2) Incubar a 37°C durante 15 minutos.
- 3) Mientras se incuban las suspensión de fagos/células, se añade 1,0 mL de sustrato de almidón rojo (1,2% p/v) a 6,0 mL de agar de superficie NZY a 50°C y se mezcla completamente. Se mantiene la solución a 50°C hasta que se necesite.
- 5 4) Se transfiere la suspensión de células a la solución de sustrato/agar de superficie y se mezcla suavemente.
- 5) Se vierte inmediatamente la solución en una placa grande (150 mm) de medio NZY.
- 6) Permitir que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego invertir la placa.
- 7) Incubar las placas a 39°C durante 8 - 12 horas.
- 10 8) Se observan las zonas de aclaramiento (halo) alrededor de una sola plaqueta (clon puro). Si no se puede aislar una sola plaqueta, se ajusta el título y se siembra en placa nuevamente la suspensión del fago
- 9) Las plaquetas unitarias con halo se separan del agar y se transfieren a un micro tubo esterilizado. Una punta de pipeta de orificio grande de 200 µL funciona bien para remover (separar) el tapón de agar que contiene la plaqueta deseada. Para amplificar el título, se separan plaquetas activas individuales en un microtubo.
- 15 10) Los fagos se resuspenden en 500 µL de regulador SM. Se añaden 20 µL de cloroformo para inhibir cualquier crecimiento adicional de células.
- 11) Se incuba la suspensión pura de fagos a temperatura ambiente durante 4 horas o toda la noche antes de la siguiente etapa. Se almacena la suspensión del fago puro a -80°C mediante la adición de DMSO en la suspensión del fago (7% v/v)
- 20 Escisión de un clon puro
- 1) Se añaden 100 µl de la suspensión pura de fagos a 200 µl de células MRF de *E. coli* (OD₆₀₀ = 1,0). A esta suspensión, se le añade 1,0 µL del fago auxiliar EXASSIST (>1 x 10⁶ pfu/mL; Stratagene). Usar tubos Falcon 2059 para la escisión.
- 2) Se incuba la suspensión a 37°C durante 15 minutos.
- 25 3) Se añaden 3,0 mL de medio YT 2 veces a la suspensión de células.
- 4) Se incuba a 30°C durante al menos 6 horas o durante la noche con agitación.
- 5) Se transfiere el tubo a 70°C durante 20 minutos. La suspensión de fagos se puede almacenar a 4°C durante 1 a 2 meses.
- 6) Se transfieren 100 µl de la suspensión del fagémido a un micro tubo que contiene 200 µL de células Exp 505 de *E. coli* (OD₆₀₀ = 1,0).
- 30 7) Se incuba la suspensión a 37°C durante 15 minutos.
- 8) Se añaden 300 µL de SOB a la suspensión.
- 9) Se incuba la suspensión a 37°C durante 30 a 45 minutos.
- 10) Se transfieren 00uL de la suspensión a una placa pequeña (90 mm) de medio LB) que contiene Kanamicina (medio LB con Kanamicina a razón de 50 µg/mL) para bibliotecas de ADN Zap Express o Ampicilina (medio LB con Kanamicina a razón de 100 µg/mL) para bibliotecas de ADN Zap II.
- 35 11) Se transfiere el resto de la suspensión a otra placa pequeña de medio LB.
- 12) Se usan perlas estériles de vidrio para distribuir uniformemente la suspensión sobre la placa.
- 13) Se incuban as placas a 30°C durante 12 a 24 horas,

ES 2 575 912 T3

14) Se observan las colonias en las placas.

15) Se inocula una sola colonia en medio líquido LB que contiene un antibiótico adecuado e se incuba a 30°C durante 12 a 24 horas.

5 16) Se puede preparar un patrón de glicerol mediante la adición de 80% de glicerol en medio de cultivo líquido (15% v/v) y se almacena a -80°C.

Verificación de la actividad

1) Se transfieren 50 µL de cultivo líquido a un microtubo. Se añaden 500 µL de 8% de Amilopectina Azure pH7 en el mismo tubo. Se preparan 2 tubos para cada clon.

2) Se analiza la actividad a 50°C durante 3 horas y durante la noche. Se utiliza regulador de pH 7 como control.

10 3) Se enfría el espécimen de prueba en un baño de agua helada durante 5 minutos.

4) Se añaden 750 µL de etanol y se mezcla completamente.

5) SE centrifuga a 13000 rpm (16000 g) durante 5 minutos.

6) Se mide la OD del sobrenadante a 595 nm.

Análisis RFLP

15 1) Se transfiere un 1,0 mL de cultivo líquido a un microtubo esterilizado.

2) Se centrifuga a 13200 rpm (16000 g) durante 1 minuto.

3) Se descarta el sobrenadante. Se añade otro 1,0 mL de cultivo líquido en el mismo microtubo esterilizado.

4) Se centrifuga a 13200 rpm (16000 g) durante 1 minuto.

5) Se descarta el sobrenadante.

20 6) Se sigue el protocolo del mini kit de centrifugación QIAPREP^{MR} para el aislamiento del plásmido.

7) Se revisa la concentración de ADN usando BioPhotometer.

8) Se usa Sac I y Kpn I para la primera digestión doble. Se incuba a 37°C durante 1 hora.

9) Se usa Pst I y Xho I para la segunda digestión doble. Se incuba a 37°C durante 1 hora.

10) Se añade colorante de carga en la muestra digerida.

25 11) Se corre la muestra digerida en un gel de agarosa al 1,0% durante 1 - 1,5 hora a 120 voltios.

12) Se observa el gel con un visualizador. Se pueden analizar las secuencias de todos los clones con un patrón diferente de digestión.

Ejemplo 7: Ensayo para amilasas

30 El siguiente ejemplo describe un método de ejemplo para determinar si un polipéptido está dentro del alcance de la invención.

Preparación de cultivos huésped

1. Se parte de un cultivo durante la noche de células huésped MRF⁺ XL1-BLUET^{MR}. Se usa una sola colonia de una placa estriada para inocular 10 mL de LB suplementado con 20 µg/mL de tetraciclina. Se desarrolla el cultivo durante la noche agitando a 37°C durante al menos 16 horas.

ES 2 575 912 T3

2. Usando una técnica aséptica, se inoculan 100 mL de cultivo LB_{Tet} fresco el día del cultivo con el huésped MRF' XL1- BLUE^{MR} a partir del cultivo LB_{Tet} durante la noche.

3. Se cultiva en un agitador a 37°C hasta que la OD alcance 0,75 - 1,0.

5 4. Se sedimentan las células huésped 1000 x g durante 10 minutos y se resuspenden suavemente en MgSO₄ 10 mM a OD 5.

5. Se diluye una pequeña cantidad de células huésped a OD 1 para uso en titulación y para pipetear con agujas.

6. Se pueden usar preparaciones huésped hasta para 1 semana cuando se almacenan en hielo o a 4°C.

– Para acortar el tiempo de crecimiento para el día del cultivo, se usa la concentración Tet usual de 1/2X en LB (1/2X =10 µg/mL), o se omite también el antibiótico.

10 – No se utiliza el NZY cuando se selecciona con tetraciclina. La alta concentración de Mg⁺⁺ en medio NZY vuelve a Tet inactivo.

Titulación de bibliotecas lambda

7. Se colocan tres tubos esteriles de microcentrífuga en una gradilla.

15 8. Se toma una alícuota de 995 µL de células huésped preparadas en un tubo y 45 µL de células huésped preparadas a OD 1 en cada uno de los dos tubos restantes.

9. Se añaden 5 µL de una biblioteca lambda al tubo que contiene 995 µL de células huésped y se mezcla mediante agitación tipo vórtice. Esto da como resultado un factor de dilución de 200.

20 10. Se preparan diluciones 1/2000 y 1/20000 mediante la adición consecutiva de 5 µL de una dilución previa a los dos tubos restantes que contienen 45 µL de células huésped preparadas. Se mezcla mediante agitación tipo vórtice después de hacer cada dilución.

11. Se permite que el fago se adsorba al huésped mediante incubación a 37°C durante 15 minutos.

12. Mientras tanto, se pipetea 100 µL de células huésped preparadas OD1 a cada uno de los tres tubos Falcon 2059.

25 13. Se añaden 5 µL de cada dilución a un tubo 2059 separado que contiene células huésped.

14. Se siembra en placa cada una mediante la adición de 3 mL de agar de superficie a cada tubo y se vierte rápidamente sobre placas de 90 mm de NZY. Se asegura una distribución plana y uniforme antes de que el agar de superficie se endurezca.

15. Se invierten las placas y se incuban a 37°C durante la noche.

30 16. Se hace recuento de la placas y se calcula el título del patrón de la biblioteca (en unidades formadores de placas (pfu) por µL).

Cribado de microtítulos lambda para amilasas

Preparación

35 1. Se prepara una cantidad suficiente de cultivo de huésped MRF' XL1-Blue, como se describió anteriormente, para la cantidad de cribado planificado. Un cultivo de 100 mL por lo general es suficiente para cribar 2 - 3 bibliotecas.

2. Se autoclavan varias botellas compatibles con el dispensador QFill2. Estas son las botellas Corning de boca ancha de 250 mL que contiene un anillo sellador alrededor de la tapa.

3. Se asegura que hayan una cantidad suficiente de placas, agar de superficie, almidón BODIPY, solución de almidón rojo, etc., disponible para el cribado.

40 4. Se programa la corrida con un robot el día 2 con un representante de Automation.

ES 2 575 912 T3

Día 1

1. Se etiquetan las placas de 1536 pozos (negras) con cribado de biblioteca y el número de placa. Los adhesivos para tubos Tough-Tags^{MIR}, cortados a la mitad, son ideales para etiquetar las placas de 1536 pozos.
2. Se calculan los volúmenes de biblioteca, células huésped y medio NZY necesarios para el cribado. Esto se hace fácilmente con una hoja de cálculos de Excel.
3. Se combinan los volúmenes calculados de biblioteca lambda y células huésped OD5 en una botella Corning de 250 mL de boca ancha esterilizada (que contiene un anillo sellador).
4. Se permite que se realice la adsorción a 37°C durante 15 minutos.
5. Se añade el volumen calculado de medio NZY y se mezcla bien. Esto se denomina como la suspensión de célula-fago-medio.
6. Se realiza un título en forma concomitante al combinar 50 µL de la suspensión de célula-fago-medio con 250 µL de células huésped OD1 en un tubo Falcon 2059, luego se siembra en placa con 9 mL de agar de superficie sobre una placa de 150 mm de NZY. Se incuba título concomitante de la placa a 37°C durante la noche.
7. Se carga el dispensador con el resto de la suspensión y se arregla cada placa etiquetada de 1536 pozos con 4 µL por pozo. Si el dispensador deja burbujas de aire en algunos pozos, se pueden remover centrifugando las placas a 200 x g durante 1 minuto.
8. Se añaden 0,5 µL de fago de control positivo a la posición del pozo AD46 de al menos dos de las placas de ensayo. Se utilizar un clon lambda fuerte positivo para amilasa para este propósito.
9. Se incuban las placas de ensayo a 37°C durante la noche en una incubadora humidificada (≥95%).

20 Día 2

1. Se recuentan las pfu la placa de título concomitante y se determina la densidad promedio de semillas por pozo (en pfu por pozo).
2. Se pipetea con aguja al menos 2 placas de cada cribado de biblioteca (preferentemente las 2 que contienen controles positivos) de la siguiente manera:
 - a) Se preparan 2 placas base para que actúen como superficie sobre la cual se pueda pipetear con aguja: se combinan 250 µL de células huésped OD1 con 2 mL de almidón rojo al 2% y se siembra en placa con 9 mL de agar de superficie sobre placas de 150 mm de NZY. Se mantiene el nivel de cada placa en la medida de lo posible a medida que el agar de superficie se solidifica a fin de producir un tono rojo uniforme en la placa.
 - b) Se utiliza una pipeta con aguja en la posición 1536 esterilizada dos veces sobre llama, se replican al menos 2 de las placas de cribado sobre las placas base del huésped.
 - c) Se colocan las placas que fueron pipeteadas con aguja en una cabina de flujo laminar sin las tapas durante aproximadamente 15-30 minutos (para ventilar el exceso de humedad).
 - d) Se reemplazan las tapas y se incuban de forma invertida a 37°C toda la noche.
3. Se prepara el regulador de sustrato de almidón 2X BODIPY de la siguiente manera:
 - a) Se calcula el volumen total de solución reguladora de sustrato 2X necesario para todas las placas de cribado a razón de 4 µl por pozo (incluyendo cualquier volumen extra de espacio muerto requerido por el dispensador) y se mide esta cantidad de CAPS 100 mM de pH 10,4 en un recipiente apropiado para el dispensador usado.
 - b) Se retiran suficientes tubos 0,5 mg de almidón BODIPY para producir el volumen requerido de regulador de sustrato 2X [calculado en la etapa a) anterior] a una concentración final de 20 - 30 µg/mL.
 - c) Se disuelve cada tubo de 0,5 mg en 50 µL de DMSO a temperatura ambiente, protegido de la luz, agitación frecuente tipo vórtice. Esto toma más de 15 minutos; algunos lotes de producción de almidón BODIPY disuelven mejor que otros.

ES 2 575 912 T3

d) Se agregan 50 μL de regulador CAPS 100 mM de pH 10,4 a cada tubo y se mezcla mediante agitación tipo vórtice.

e) Se reúnen los contenidos de todos los tubos y se remueve cualquiera de los agregados no disueltos por centrifugación durante 1 minuto a velocidad máxima en una microcentrífuga.

5 f) Se agrega el sobrenadante al resto del regulador CAPS 100 mM medido en la etapa a) anterior.

g) Se protege el regulador del sustrato 2X de la luz envolviéndolo en una lámina de aluminio.

4. Se llevan las placas y el regulador de sustrato a la sala de automatización y se programa el robot con los siguientes parámetros:

a) se dispensan 4 μL de regulador de sustrato por pozo

10 b) Se hace la primera lectura 1 hora después del sustrato, la segunda lectura a las 9 horas y la tercera lectura a las 17 horas; con incubación a 37°C entre lecturas

c) Filtro de excitación: 485 nm; filtro de emisión: 535 nm

d) Se fija la ganancia por Spectrafluor en 70, o se prepara la ganancia óptima para el lote del regulador de sustrato 2X.

15 e) Se asegura que la incubadora utilizada proteja las placas de ensayo de la luz.

Día 3

1. Se revisan las placas pipeteadas con aguja para los aclaramientos en la base bacteriana en todas las posiciones correspondientes a los pozos en la placa de ensayo asociada. También se revisan los aclaramientos en el almidón rojo en cualquiera de las posiciones de la aguja. Si las placas que contienen controles positivos se usaron para pipeteado con agujas, se debe poder observar una gran zona de aclaramiento en el fondo rojo. Se debe tener cuidado con los contaminantes que también forman zonas de aclaramiento en el almidón rojo (véase comentario "Contaminantes que forman zonas de aclaramiento en almidón rojo" al final del Ejemplo 7).

2. Se identifican las coincidencias putativas del archivo de datos producido por ordenador del robot. El programa KANAL producido por Engineering simplifica el análisis de datos. Como regla, un acierto putativo se caracteriza como un pozo que tiene una intensidad de señal que se eleva al menos 1,5 veces por sobre el entorno.

3. Para cada coincidencia putativa, se remueven 2 μL del pozo y se agregan a un tubo que contiene 500 μL del regulador SM y 50 μL de CHCl_3 . Se realiza una agitación tipo vórtice para mezclar y almacenar a 4°C . Esta solución será denominada en lo sucesivo como el patrón $4\text{e-}3$. Las placas de cribado originales deben almacenarse a 4°C , protegidas de la luz, al menos hasta que se completen los desbloques.

30 Este es un ejemplo de un método de desbloquear los aciertos putativos. Es un ensayo de fase líquida que confía en la confirmación de actividad sobre el almidón BODIPY. En forma alternativa, los aciertos putativos pueden ser desarrollados directamente en placas de fase sólida que contienen almidón rojo de manera tal que 2000 - 3000 pfu por acierto son examinadas para zonas de aclaramiento. Sin embargo, la incapacidad para observar zonas de aclaramiento en almidón rojo no es necesariamente un indicio de que un acierto putativo era un falso positivo.

35 Posteriormente sería necesario evaluarlo usando el formato en el cual fue identificado originalmente (es decir, fase líquida usando almidón BODIPY como sustrato). Asimismo, los positivos más débiles son más fácilmente identificados usando el método detallado a continuación.

Día 1

1. En un tubo cónico esterilizado de 50 mL, se combinan 0,5 mL de células huésped OD5 con 45,5 mL de NZY. Esto será denominado la suspensión huésped-medio.

2. Para cada acierto putativo que se analiza, se toma una alícuota de 1 mL de la suspensión huésped-medio en cada uno de los 3 tubos de microcentrífuga estériles.

3. Se fija la pipeta Pipetman de 12 canales en modo multidispensador con un tamaño de alícuota de 20 μL y un número de alícuota de 2x. Se monta la pipeta Pipetman con un conjunto limpio de puntas estériles.

ES 2 575 912 T3

4. Se vierte aproximadamente 1 mL de suspensión huésped-medio en una nueva tina de solución estéril y se carga la pipeta Pipetman multicanal.
5. Se dispensan 20 µL por pozo en la última fila (fila P) de una placa negra de 384 pozos (12 canales x 2 = 24 pozos). Esta fila se usa posteriormente para los controles.
- 5 6. Se expelle el líquido restante en las puntas al tocar las puntas contra la superficie de la tina y se presiona el botón RESETEAR en la pipeta Pipetman. Se deja la pipeta Pipetman hacia abajo a fin de evitar la contaminación de las puntas. No hay necesidad de cambiar las puntas en este momento.
7. Se vierte el resto del fluido en la tina en un contenedor de desechos (como un cubo) teniendo cuidado de evitar contaminación por salpicaduras.
- 10 8. Para que se analice el primer acierto putativo, se toman 111 µL del patrón 4e-3 (véase Día 2 en Cribado de microtítulo lambda para amilasas) y se agregan al primero en un conjunto de tres tubos que contiene 1 mL de suspensión huésped-medio (etapa 2). Se realiza agitación tipo vórtice para mezclar. Esta es la Dilución A.
9. Se toman 111 µL de la Dilución A y se añaden al próximo tubo en el conjunto. Se realiza agitación tipo vórtice para mezclar. Esta es la Dilución B.
- 15 10. Se toman 111 µL de la Dilución B y se añaden al último tubo en el conjunto. Se realiza agitación tipo vórtice para mezclar. Esta es la Dilución C. Se deben tener ahora tres diluciones de fagos, donde las concentraciones de cada una difieren por un factor de 10.
11. Se vierten los contenidos de la Dilución C (la más diluida de las 3 muestras) en la tina de solución y se carga la pipeta Pipetman multicanal.
- 20 12. Se dispensan 20 µL por pozo en la primera fila de la placa de 384 pozos (12 canales x 2 = 24 pozos).
13. Se expelle el líquido restante en las puntas al tocar las puntas contra la superficie de la tina y se presiona el botón RESETEAR en la pipeta Pipetman. Se deja la pipeta Pipetman hacia abajo afin de evitar la contaminación de las puntas. No hay necesidad de cambiar las puntas en este momento
14. Se vacía la tina como se describió anteriormente.
- 25 15. Se vierten los contenidos de la Dilución B en la misma tina y se carga la pipeta Pipetman multicanal.
16. Se dispensan 20 µL por pozo en la segunda fila de la placa de 384 pozos.
17. Se realizan las etapas 13 - 16 en forma similar para dispensar la Dilución A en la tercera fila de la placa.
18. Una vez que todas las tres diluciones han sido arregladas en las primeras 3 filas de la placa, se descartan todas las puntas y la solución de la tina en el contenedor de desechos biotóxicos.
- 30 19. Se monta la pipeta Pipetman con un conjunto limpio de puntas esterilizadas y se abre una nueva tina de solución estéril.
20. Se repiten las etapas 8 -19 para cada acierto putativo restante, usando las filas restantes sobre la placa hasta la fila O. Se pueden analizar cinco aciertos putativos sobre una placa de 384 pozos, guardando la última fila (fila P) para los controles.
- 35 21. Se agregan 0,5 µL de cada control a un pozo separado, Se usan al menos 2 - 3 controles separados, preferiblemente cubriendo un rango de actividad.
22. Se incuban las placas de ensayo a 37°C durante la noche en una incubadora humidificada (≥95%).

Día 2

- 40 1. Se pipetea con aguja todas las placas desbloqueadas sobre una placa base con almidón rojo usando el mismo método descrito para el Día 2 en Cribado de microtítulo lambda para amilasas, salvo que se utilice una pipeta con aguja de 384 posiciones.

ES 2 575 912 T3

2. Se prepara el regulador del sustrato de almidón BODIPY 2X de la siguiente manera:
 - a) Se calcula el volumen total de solución reguladora de sustrato 2X necesario para todas las placas de cribado a razón de 4 μL por pozo (incluyendo cualquier volumen extra de espacio muerto requerido por el dispensador) y se mede esta cantidad de CAPS 100 mM de pH 10,4 en un recipiente apropiado para el dispensador usado.
- 5 b) Se retiran suficientes tubos 0,5 mg de almidón BODIPY para producir el volumen requerido de regulador de sustrato 2X [calculado en la etapa a) anterior] a una concentración final de 20 - 30 $\mu\text{g}/\text{mL}$.
 - c) Se disuelve cada tubo de 0,5 mg en 50 μL de DMSO a temperatura ambiente, protegido de la luz, agitación frecuente tipo vórtice. Esto toma más de 15 minutos; algunos lotes de producción de almidón BODIPY disuelven mejor que otros.
- 10 d) Se agregan 50 μL de regulador CAPS 100 mM de pH 10,4 a cada tubo y se mezcla mediante agitación tipo vórtice.
 - e) Se reúnen los contenidos de todos los tubos y se remueve cualquiera de los agregados no disueltos por centrifugación durante 1 minuto a velocidad máxima en una microcentrífuga.
 - f) Se agrega el sobrenadante al resto del regulador CAPS 100 mM medido en la etapa a) anterior.
- 15 g) Se protege el regulador del sustrato 2X de la luz envolviéndolo en una lámina de aluminio.
3. Se dispensan 20 μL por pozo en todas las placas desbloqueadas.
4. Se envuelven todas las placas en lámina de aluminio y se incuban a temperatura ambiente durante 2-6 horas.
5. Se lee cada placa en el Spectrafluor con la siguiente configuración:
 - a) lectura de fluorescencia (filtro de excitación: 485 nm; filtro de emisión: 535 nm)
- 20 b) definición de placa: 384 pozos, negra
 - c) lectura desde la parte superior
 - d) ganancia óptima
 - e) número de destellos: 3
- 25 6. En las hojas de cálculo de Excel resultantes, se traza un mapa de cada una de las 3 filas putativas en un gráfico separado y se revisa la actividad. Se asegura que todos los controles positivos produzcan señales por sobre el fondo.
7. Para cada acierto putativo que parece tener una señal real entre los pozos, se recoge una muestra de un pozo positivo de la siguiente manera:
 - a) Se selecciona un pozo positivo de una fila que representa la dilución inicial más alta.
- 30 b) Se transfieren 2 μL de ese pozo a un tubo que contiene 500 μL de SM y 50 μL de CHCl_3 . Este se denomina como el patrón de desbloqueo.
 - c) Se almacena a 4°C.
- 35 8. Usando los métodos descritos anteriormente, se siembran en placa aproximadamente 10 μL de cada patrón de desbloqueo sobre placas de 150 mm de NZY usando almidón rojo. El objetivo es obtener varias placas de pozos separados (al menos 20) de las cuales se aísla el núcleo.

Día 3

1. Se revisan las placas pipeteadas con aguja para una incidencia aceptable de los aclaramientos en la base bacteriana correspondiente a los pozos en la placa de ensayo asociada. También se revisan los aclaramientos en el

ES 2 575 912 T3

almidón rojo en los controles positivos y en cualquier acierto putativo evaluado. Debe tenerse cuidado de los contaminantes que también forman zonas de aclaramiento en almidón rojo (véase a continuación).

2. De las placas en fase sólida que contienen diluciones de patrones de desbloqueo, se separan varias placas aisladas, cada una en 500 µl de SM con 50 µl de CHCl₃. Esto se denomina como el patrón aislado.

5 3. Los patrones aislados se pueden evaluar en forma individual en almidón BODIPY usando los métodos descritos anteriormente. Esta etapa puede ser omitida si la plaqueta que fue separada en la etapa 2 produjo una zona de aclaramiento en el fondo de almidón rojo. Los patrones aislados fueron luego analizados individualmente en almidón BODIPY usando los métodos descritos anteriormente. Sin embargo, esta etapa puede ser omitida si la plaqueta que fue separada en la etapa 2 produjo una zona de aclaramiento en el fondo del almidón rojo.

10 Escisiones

Día 1

1. En un tubo Falcon 2059, se mezclan 200 µL de huésped OD1 MRF' XL1-Blue, 100 µl de patrón aislado lambda y 1 µL de patrón de fagos EXASSIST^{MR}.

2. Se incuban en agitador a 37°C durante 15 minutos.

15 3. Se agregan 3 mL de medio NZY.

4. Se incuban en agitador a 30°C durante la noche.

Día 2

1. Se calienta tubo de escisión a 70°C durante 20 minutos.

2. Se centrifuga a 1000 x g durante 10 minutos.

20 3. En un tubo Falcon 2059, se combinan 50 µL de sobrenadante con 200 µL de huésped OD1 EXP505^{MR}.

4. Se incuba en un agitador a 37°C durante 15 minutos.

5. Se añaden 300 µL de medio SOB.

6. Se incuba en agitador a 37°C durante 30 - 45 minutos.

25 7. Se siembran en placa 50 µL sobre una placa grande LB_{Kan50} usando perlas de vidrio estériles. Si las placas están "secas", se puede añadir medio SOB extra para ayudar a separar las células.

8. Se incuba la placa a 30°C durante al menos 24 horas.

9. se cultiva un aislado para secuenciación y/o RFLP.

El cultivo a 30°C reduce el número de copias de plásmido y se usa para mitigar la toxicidad aparente de algunos clones de amilasa

30 Contaminantes que forman zonas de aclaramiento en almidón rojo

Al usar almidón rojo en medio sólido para evaluar fagos en cuanto a su actividad de amilasa, es común ver unidades formadoras de colonias (cfu) contaminantes que forman zonas de aclaramiento en el almidón rojo. Para las placas pipeteadas con aguja, es importante distinguir los clones de fagos positivos para amilasa a partir de estos contaminantes siempre que se alineen con una posición particular de un pozo. La fuente de los microbios contaminantes es presumiblemente del 2% de solución patrón de almidón rojo, que no puede ser esterilizada por autoclavado o por filtración después de la preparación. Se cree que son organismos oportunistas que sobreviven metabolizando el almidón rojo. A fin de reducir estos contaminantes, se pueden usar técnicas estériles al elaborar soluciones de almidón rojo al 2% y almacenar los patrones ya sea a 4°C o sobre hielo.

35

Ejemplo 8: Caracterización de alfa amilasa, pH óptimo y determinación de la actividad específica

El siguiente ejemplo describe un ejemplo de un método para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, mediante determinación de la actividad de alfa amilasa, el pH óptimo y la actividad específica.

5 Las enzimas de esta invención se pueden usar, por ejemplo, tanto para licuefacción de almidón como para molienda húmeda de maíz y para descolado de textiles; por ejemplo, en algunas formas de realización, las enzimas de la invención tienen un pH óptimo de 4,5 a 5,0; a este pH inferior, es posible usar poco o nada de calcio lo cual disminuye los costos totales de operación y la formación de menos subproductos. Además, a este pH bajo, existe un menor uso de productos químicos y de carga de intercambio iónico. La amilasa de *B. licheniformis* estándar de la industria está por debajo del nivel óptimo tanto en la termoestabilidad como en el pH óptimo. En algunas formas de
10 realización, las enzimas de la invención tienen actividad específica de aplicación superior comparada con la amilasa de *B. licheniformis* y, por lo tanto, se requiere mucha menos enzima para hidrolizar una tonelada de almidón (por ejemplo, en algunas formas de realización, se puede usar tanto como 20 veces menos enzima).

El pH óptimo para la hidrólisis de almidón se puede determinar al hacer reaccionar 50 µL de enzima, 0,35 U/ml, con 100 ml de una solución de almidón soluble al 1% (0,0175 U/g de almidón) durante 30 minutos a 95°C. Los extremos
15 freudtores generados en la solución de almidón licuado se pueden medir por ensayo de neocupronina descrito aquí. La hidrólisis porcentual de almidón de maíz se puede determinar al medir el número de extremos reductores del azúcar producidos con el ensayo de neocupronina. Se pesaron 70 gramos de solución regulador (pH 4-7) y se pueden añadir 100 ppm de calcio. Se pueden mezclar 30 gramos de almidón de maíz en la solución reguladora para formar una suspensión de almidón. La enzima se puede agregar y los recipientes se sellan y se incuban a 95°C
20 durante 30 minutos con una velocidad inicial de calentamiento de 6°C por minuto. Una muestra de 1 ml puede ser extraída de los recipientes de reacción y analizada por ensayo de neocupronina. En algunas formas de realización, las enzimas de la invención tienen un pH óptimo entre 4,5 y 5, mientras que la amilasa de *B. licheniformis* comercial actúa en forma óptima hasta aproximadamente pH 6,0.

Ejemplo 9: Ensayos de actividad de amilasa

25 El siguiente ejemplo describe, entre otras cosas, métodos de ejemplo para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, mediante los ensayos descritos a continuación.

Ensayo que usa almidón en RBB

Se agregan 75 µl de sustrato de almidón en RBB (1% de almidón de maíz insoluble en RBB en regulador de NaAc 50 mM, pH = 4,5) en cada pozo de una nueva placa de 96 pozos (fondo V). Se transfieren 5 microlitros de lisado
30 enzimático en cada pozo con sustrato usando BIOMEK^{MR} o ZYMARK^{MR}. Las placas se sellan con cinta de sellado de aluminio y se agitan brevemente en el agitador. Las placas se incuban a 90°C durante 30 minutos, seguido de enfriamiento a temperatura ambiente durante aproximadamente 5 a 10 minutos. Se agregan 100 microlitros de etanol al 100% en cada pozo, las placas se sellan y agitan brevemente en el agitador. Las placas luego se centrifugan a 4000 rpm durante 20 minutos usando una centrífuga de mesa. Se transfieren 100 µl del sobrenadante
35 a una nueva placa de 96 pozos (fondo plano) de BIOMEK^{MR} y se lee la OD₅₉₅. Se deben utilizar controles.

Ensayo que usa almidón con FITC

Se agregan 50 µl de sustrato (almidón con FITC al 0,01% en regulador de NaAc 100 mM, pH = 4.5) en cada pozo de una nueva placa de 384 pozos. Se transfieren 5 µl de lisado enzimático en cada pozo con sustrato y se incuba la
40 placa a temperatura ambiente durante la noche. El cambio de polarización del sustrato, excitación a 485 nm, emisión 535 nm, se lee para cada pozo. Se deben usar controles. Se pueden utilizar placas de 96 pozos para todos los ensayos.

Confirmación de nuevos clones activos

Cada clon positivo del cribado es cultivado e inducido usando un protocolo estándar. Cada clon es examinado para cultivo (es decir, densidad celular con el paso del tiempo), actividad a nivel celular (ensayo de almidón en RBB y
45 ensayo de licuefacción), expresión (gel de proteína) y solubilidad de proteína (por análisis al microscopio). Los nuevos clones elevados confirmados se transfieren para fermentación.

Ejemplo 10: Ejemplo de protocolo para licuefacción de almidón y medición de resultados

El siguiente ejemplo describe y protocolo de ejemplo para la licuefacción de almidón utilizando amilasas de la invención. Un ensayo de ejemplo utiliza almidón licuado a un pH de 4,5 o 6,5 utilizando la reacción en condiciones
50 mostradas a continuación:

Condiciones de reacción: PO₄ 100 mM pH 6,5, almidón licuado al 1% (p/p), DE 12 a 55°C. Se hicieron ensayos tanto de TLC como de HPLC para verificar la actividad. Los perfiles de pH para las amilasas a ser evaluadas se realizan usando regulador de fosfato a un pH de 3,0 - 6,5, a 55°C. A partir de la cantidad de hidrólisis observable, se puede demostrar visualmente que algunos clones son más (o menos) activos a ciertos valores de pH que los otros valores en las condiciones de reacción indicadas anteriormente.

Un ejemplo de un protocolo para la sacarificación de almidón licuado a pH de 6,5:

- Se ajusta el pH del almidón licuado al pH al cual se realiza(n) la(s) sacarificación(es). Se licúa el almidón en regulador de acetato de sodio 100 mM, pH 4,5 con sales de fosfato de sodio 100 mM agregados de manera tal que antes de la sacarificación, el pH se ajusta a pH 6,5.
- Se pesan muestras de 5 gramos de almidón licuado en botellas taradas.
- Se usan 0,04% (p/p) de OPTIDEX L-400^{MR} o aproximadamente 400 mL del patrón diluido 1-10 de OPTIDEX L-400^{MR} por 100 gramos de almidón licuado.
- Se calculan los miligramos de OPTIDEX L-400^{MR} contenidos en los 400 mL del patrón diluido 1-10 de OPTIDEX L-400^{MR}. A continuación, se calcula el volumen de lisados necesarios para brindar la misma concentración de enzima que de OPTIDEX L-400^{MR}.
- Se añaden enzimas a las muestras de almidón licuado e se incuban a temperatura deseada (50°C). Después de 18 horas, se determina DE y se prepara una muestra para análisis de HPLC.

Un ejemplo de una determinación de DE:

Ejemplo de un ensayo de neocuproina:

Se agrega una muestra de 100 ml a 2,0 ml de solución A de neocuproina (40 g/L de carbonato de sodio, 16 g/L de glicina, 0,45 g/L de sulfato de cobre). A esta solución se le agrega 2,0 ml de solución B de neocuproína B (1,2 g/L de clorhidrato de neocuproína Sigma N-1626). Los tubos se mezclan y calientan en un baño de agua en ebullición durante 12 minutos; se enfrían, se diluyen hasta un volumen de 10 ml con agua DI y se lee la OD a 450 nm en el espectrofotómetro. El equivalente de glucosa en la muestra se extrapola a partir de la respuesta de 0,2 mg/ml de glucosa estándar que se corre en forma simultánea.

Ejemplo de análisis por HPLC:

Los perfiles de sacarificación de un carbohidrato se miden por HPLC (Bio-Rad Aminex HPX-87A columna de plata, 80°C) usando un detector de índice de refracción. La fase móvil se filtra con agua Millipore a una velocidad de flujo de 0,7 ml/min. Las muestras de sacarificación se diluyen 1-10 con agua DI acidificada (5 gotas de HCl 6 M en 200 mL de agua DI) luego se filtran a través de un filtro de jeringa de 0,45 µm. El volumen de inyección es 20 µL.

Ejemplo de TLC:

Los productos de reacción fueron w/d en puntos de tiempo cada hora y luego se colocaron y secaron en una placa de TLC. La placa luego se desarrolló en una proporción 10:90 de agua:isopropanol y se visualizó con una mancha de vainilina o una mancha de CAM y luego se calentó para mostrar los azúcares reducibles. El almidón licuado fue parcialmente hidrolizado hasta glucosa en los casos en que se observó actividad.

Ejemplo 11: Licuefacción de almidón usando amilasas y/o glucoamilasas de la invención

Este ejemplo describe un método de ejemplo de la invención para la licuefacción de almidón utilizando amilasas de la invención.

El concentrado de amilasa se puede preparar a partir de los caldos de fermentación por tratamiento térmico, lavado celular, extracción alcalina usando microfiltración y ultrafiltración (48% de rendimiento total). El concentrado de UF puede neutralizarse con ácido acético y formularse con 30% de glicerol a un pH de 4,5. El nivel de actividad de la formulación en la suspensión puede ser representativo de un producto comercial (120 U¹/g 0,5 kg/tonelada de almidón).

Ensayo de actividad de amilasa estándar

ES 2 575 912 T3

- Una cubeta de 1 mL que contiene 950 μ L de MOPS 50 mM pH 7,0 que contiene pNP- α -D-hexa-(1 \rightarrow 4)-glucopiranosido 5 mM se coloca en el controlador de temperatura Peltier del espectrofotómetro Beckman DU- 7400 precalentado a 80°C. El espectrofotómetro queda en cero a 405 nm y se agregan 50 μ L de la solución enzimática a la cubeta, se mezcla bien y se monitorea el aumento en la OD_{405 nm} en un intervalo de un minuto. La velocidad de Δ OD_{405 nm}/min se convierte a una unidad estándar de μ mol/minuto a partir de la respuesta de OD_{405 nm} de 50 μ L de 1 μ mol/mL de PNP en 950 mL de MOPS 50 mM a pH de 7,0, -80°C. Una unidad estándar de alfa amilasa termoestable (DTAA) es igual a la cantidad de enzima que catalizará la liberación de 1 μ mol/mL/minuto de pNP bajo las condiciones definidas del ensayo.

Ensayo de actividad de glucoamilasa estándar

- 10 Una cubeta de 1 mL que contiene 950 μ L de MOPS 50 mM pH 7,0 que contiene pNP- α -D-glucopiranosido 5 mM se coloca en el controlador de temperatura Peltier del espectrofotómetro DU- 7400^{MR} (Beckman) precalentado a 60°C. El espectrofotómetro queda en cero a 405 nm y se agregan 50 μ L de la solución enzimática a la cubeta, se mezcla bien y se monitorea el aumento en la OD_{405 nm} en un intervalo de un minuto. La velocidad de Δ OD_{405 nm}/min se convierte a una unidad estándar de μ mol/minuto a partir de la respuesta de OD_{405 nm} de 50 μ L de 1 μ mol/mL de pNP en 950 mL de MOPS 50 mM a pH 7,0, -60°C. Una unidad estándar de glucoamilasa (DGA) es igual a la cantidad de enzima que catalizará la liberación de 1 μ mol/mL/minuto de pNP bajo las condiciones definidas del ensayo.

Determinación de equivalente de dextrosa

- 20 El método de ensayo de neocuproína se puede usar para medir el DE. Las muestras seleccionadas se pueden medir por un procedimiento que se describe aquí, y/o mediante análisis GPC usando el procedimiento de Fehlings de GPC.

Ensayo de neocuproína

- 25 Se agrega una muestra de 100 ml a 2,0 ml de solución A de neocuproína (40 g/L de carbonato de sodio, 16 g/L de glicina, 0,45 g/L de sulfato de cobre). A esta solución se le agrega 2,0 ml de solución B de neocuproína B (1,2 g/L de clorhidrato de neocuproína Sigma N-1626). Los tubos se mezclan y calientan en un baño de agua en ebullición durante 12 minutos; se enfrían, se diluyen hasta un volumen de 10 ml con agua DI y se lee la OD a 450 nm en el espectrofotómetro. El equivalente de glucosa en la muestra se extrapola a partir de la respuesta de 0,2 mg/ml de glucosa estándar que se corre en forma simultánea.

- 30 La muestra de almidón se diluye ~1 a 16 con agua DI con la dilución exacta registrada. Se agregan 10 mililitros de la muestra diluida a 20 ml de agua DI. Se agregan 10 mililitros de solución A y B de Fehling al almidón diluido. La muestra ebulle durante 3 minutos y se enfría sobre hielo. Se agrega 10 mililitros de KI al 30% y 10 ml de H₂SO₄ 6 N. La solución se titula contra tiosulfato de sodio 0,1 N. El volumen de la solución de titulación se registra y se usa para calcular el DE.

Determinación de almidón residual

- 35 Las muestras de post-sacarificación se pueden chequear para almidón residual usando el procedimiento de yodo de Staley. Se pesan 20 gramos de muestra en un plato de pesaje grande. Se agregan 45 μ L de solución de yodo al plato de pesaje y la solución de almidón se mezcla bien. El azul oscuro indica la presencia de almidón, un azul claro - verde indica almidón leve, verde claro indica un rastro de almidón y amarillo - rojo, ausencia de almidón. La solución de yodo se prepara al disolver 21,25 gramos de yodo y 40,0 gramos de yoduro de potasio en un litro de agua.

- 40 Perfil del oligosacárido

Los perfiles de licuefacción y sacarificación de carbohidratos se pueden medir por HPLC (por ejemplo, una columna AMINEX HPX-87C^{MR} (Bio-Rad) en forma de calcio -80°C) usando un detector de índice de refracción.

Cromatografía de permeación de gel

- 45 La distribución del peso molecular se puede determinar por cromatografía, por ejemplo, en una columna μ L AQUAGEL-OH^{MR} con detección de masas por índice de refracción (Waters Model 2410). Un detector modelo T60^{MR} (Viscotek) se puede usar para mediciones continuas de viscosidad y de dispersión de luz.

Electroforesis capilar

ES 2 575 912 T3

Se puede usar un sistema de glucoproteínas Beckman Coulter P/ACE MDQ^{MR} para separación de oligosacáridos derivados de APTS en un capilar de sílice fundida; en un aspecto, se puede usar detección por fluorescencia inducida por láser.

Licuefacción primaria

- 5 El almidón de línea directamente del proceso GPC es bombeado a un tanque de alimentación de 60 litros donde el pH, DS (sólidos secos) y el nivel de calcio se pueden ajustar antes de la licuefacción. La amilasa se agrega a la suspensión. La suspensión de DS al 32% se bombea a razón de 0,7 litros/minuto mediante una bomba de desplazamiento positivo al chorro - una cámara de mezcla presurizada donde se calienta instantáneamente la suspensión de almidón a más de 100°C mediante inyección de vapor. El almidón licuado parcialmente gelatinizado
- 10 se bombea a través de una red de tuberías (aún bajo presión) para obtener el tiempo de residencia deseado (5 minutos) a temperatura. La presión se libera en un tanque de expansión y se pueden tomar muestras. Las muestras se toman por duplicado.

Licuefacción secundaria

- 15 El almidón licuado se recoje en botellas de vidrio de un litro y se mantiene en un baño de agua a 95°C durante 90 minutos.

Sacarificación

El almidón licuado se enfría a 60°C, el pH se ajusta a 4,5 y las muestras se tratan con glucoamilasa. El progreso de la sacarificación se puede monitorear en el tiempo, por ejemplo, por HPLC.

- 20 Los jarabes licuados producidos con cada amilasa se ajustan a aproximadamente pH 2,5 con HCl 6N en forma inmediata después de licuefacción secundaria de 90 minutos para inactivar cualquier amilasa residual. Los jarabes se ajustan luego a pH 4,5, se colocan en un baño de agua de 60°C y se sacarifican con tres niveles de glucoamilasa.

El grado de sacarificación se monitorea por HPLC en puntos de tiempo de 18 - 88 horas.

Los jarabes licuados se sacarifican con la dosificación estándar - 0,04% de una glucoamilasa de doble fuerza y dos dosis menores (50% y 25%) para monitorear cualquier diferencia en el progreso de sacarificación.

- 25 Los datos del progreso de la sacarificación se pueden analizar por % de desarrollo de dextrosa vs tiempo, por ejemplo, con 0,04% de glucoamilasas; o % de desarrollo de dextrosa vs tiempo, por ejemplo, con 0,02% de glucoamilasas.

Perfil de azúcar después de la sacarificación

Distribución del peso molecular

- 30 La distribución del peso molecular de jarabes licuados a DE de 12 y 18 mediante amilasas de la invención y en algunos aspectos, usando controles, por ejemplo, enzimas comerciales, por ejemplo, *Bacillus licheniformis* o *Bacillus stearothermophilus* comercial, se pueden medir por cromatografía de permeación de gel usando detección por índice de refracción, dispersión de luz y viscosidad. Tanto las amilasas de *B. licheniformis* como de *B. stearothermophilus* generan una distribución bimodal, el pico primario centrado a 2000, un pico secundario en 32000 con un hombro que se extiende más allá del rango de 160000. El pico de peso molecular más bajo representa aproximadamente 60% de la masa total de la muestra. En algunas realizaciones, amilasas de la invención pueden exhibir un solo pico en 2000 con muy poco por encima de 30000.
- 35

HPLC

- 40 Los jarabes de DE 12 y 18 producidos por amilasas de la invención (y las enzimas comerciales de control, por ejemplo, amilasas de *Bacillus licheniformis* y/o amilasas de *Bacillus stearothermophilus* comerciales) se pueden analizar por HPLC. Ambas técnicas producen huellas digitales características de cada clase de amilasa; también se espera que los patrones de oligosacáridos sean diferentes para la amilasa de *B. licheniformis* vs amilasa de *B. stearothermophilus* y amilasas de la invención produzcan huellas digitales características de sus clases enzimáticas. Los jarabes licuados de la invención (por ejemplo, jarabes elaborados por los métodos de la invención y/o elaborados por enzimas de la invención) exhiben evidencia de mayor ramificación en los oligosacáridos. La HPLC solamente resuelve los oligosacáridos en un rango < de DP15 - los fragmentos más grandes no son visibles en estas técnicas. Las amilasas de *Bacillus* son conocidas por licuar almidón de manera tal que la fracción de amilopectina se hidroliza en forma menos extensiva que la fracción de amilosa. Estos fragmentos de amilopectina >DP30 están
- 45

contenidos en la fracción de alto peso molecular centrada en 32000 y en consecuencia, se ve poca evidencia de ramificación en los análisis de HPLC de los jarabes licuados de *Bacillus*. En un aspecto, los oligosacáridos <DP15 en jarabes licuados elaborados usando las amilasas de la invención contienen fragmentos tanto de amilosa como de amilopectina.

5 **Ejemplo 12:** Licuefacción de almidón en condiciones ácidas usando amilasas de la invención

La invención provee métodos para licuefacción de almidón usando amilasas de la invención, incluyendo amilasas activas bajo condiciones ácidas, por ejemplo, entre aproximadamente pH 4,0 y 5,0, por ejemplo, pH 4,5. En una realización, la conversión de almidón en glucosa puede ser catalizada por la acción de la secuencia de dos enzimas: una endoamilasa de la invención, por ejemplo, una amilasa, tal como alfa-amilasas y una exoamilasa, por ejemplo, una glucoamilasa, para licuar el almidón (por ejemplo, la hidrólisis de polímeros de glucosa de alto peso molecular en oligosacáridos que consisten de 2 a 20 unidades de glucosa, por lo general un equivalente de dextrosa de 10 a 12, por una amilasa de la invención), seguido por sacarificación con una exoamilasa, por ejemplo, una glucoamilasa.

En un aspecto, el procesamiento se realiza en una planta de molienda húmeda de maíz que produce una suspensión de almidón que tiene un pH de aproximadamente 4,0 a 4,5. En un aspecto, el pH se eleva, por ejemplo, de 5,8 hasta 6,0 antes de la licuefacción para acomodar una alfa amilasa con una actividad de pH y estabilidad bajas (que puede ser una alfa amilasa de la invención). En un aspecto, las amilasas de la invención pueden licuar almidón a un pH de 4,5 hasta equivalentes de dextrosa que oscilan entre 12 y 18; en un aspecto, usando alfa amilasas de la invención en niveles de aproximadamente 3 a 6 gramos por tonelada de almidón. En este aspecto, el uso de alfa amilasas de la invención permite que la licuefacción de almidón se realice a pH 4,5.

En un aspecto, la licuefacción de almidón se realiza a un pH de 4,5 durante 5 minutos a 105°C hasta 90 minutos a 95°C usando amilasas de la invención. La cantidad de enzima se puede ajustar a fin de ajustar un DE objetivo de 12 a 15 después de la licuefacción. En un aspecto, el almidón licuado se sacarifica luego con una glucoamilasa, por ejemplo una glucoamilasa de *Aspergillus*, durante aproximadamente 48 horas hasta aproximadamente pH 4,5 y 60 °C. Si el jarabe sacarificado no contiene al menos 95% de glucosa, el DE de licuefacción objetivo se eleva y se repite la sacarificación hasta que la licuefacción eventualmente no produzca un jarabe sacarificado que contiene más del 95% de glucosa. La proteína amilasa necesaria para producir una materia prima licuada adecuada para sacarificación se puede determinar, por ejemplo, por PAGE o HPLC.

Ejemplo 13: Licuefacción de almidón usando amilasas de la invención

Este ejemplo describe un ejemplo de un método para licuar almidón usando amilasas y/o glucoamilasas de la invención; y describe uso de amilasas de *Bacillus licheniformis* y *Bacillus stearothermophilus* comerciales como controles. Estos ensayos pueden comparar los niveles del progreso de la sacarificación y los niveles finales de dextrosa a partir de los jarabes generados por las enzimas de la invención y amilasas comerciales.

El equivalente de dextrosa (DE) es el estándar industrial para medir la concentración de azúcares reductores totales, calculado como D-glucosa sobre una base de peso seco. El almidón granuloso no hidrolizado tiene un DE de casi cero, mientras que el DE de la D- glucosa se define como 100. Un ejemplo de un proceso de la invención usa una dosificación de enzimas de aproximadamente 60 a 70 Unidades/kilo de almidón a un pH de 4,5 para alcanzar un DE 19.

Los patrones de oligosacáridos generados por las amilasas de la invención y amilasas comerciales se pueden analizar por distribución de peso molecular (PM) usando, por ejemplo, cromatografía de permeación de gel con detección por dispersión de luz y viscosidad. En un aspecto, las amilasas de la invención pueden generar un DE de 18 y una distribución uniforme de PM de oligosacáridos, no mayor a 20.000. Esto es consistente con la viscosidad menor para los jarabes de la invención (por ejemplo, los jarabes elaborados por métodos de la invención, o elaborados usando las enzimas de la invención). Los perfiles de DP (grados de polimerización) medidos por HPLC también se pueden usar para analizar las diferencias en el patrón de acción.

Los concentrados de amilasa se preparan a partir de caldos de fermentación por tratamiento térmico, lavado celular, extracción alcalina usando microfiltración y ultrafiltración (UF). El concentrado de UF se neutraliza con ácido acético y se formula con 30% de glicerol a pH 4,5. El nivel de actividad de la formulación de la suspensión se puede analizar, por ejemplo, 120 U1/g - 0,5 kg/tonelada de almidón es representativo de un producto comercial.

Ejemplo 14: Amilasas alcalinas para aplicaciones de lavandería y lavado automático de vajillas

En un aspecto, la invención provee detergentes que comprenden las amilasas de la invención, incluyendo amilasas activas bajo condiciones alcalinas, y métodos de fabricación y uso de ellos.

ES 2 575 912 T3

La invención provee enzimas amilasa estables en medio alcalino, que pueden ser comparadas con una enzima(s) comercial de referencia en cuanto a las características importantes en aplicaciones de lavandería y lavado automático de vajillas (ADW):

o Una prueba de lavado ADW en portaobjetos revestidos con almidón.

5 o Ensayo de actividad de la enzima amilasa en presencia de una formulación lavandería / ADW usando un sustrato soluble.

o Ensayo de actividad de la enzima amilasa y/o glucoamilasa, en presencia de agentes quelantes.

o Ensayo de actividad de la enzima amilasa y determinación de rangos óptimos de pH alcalino (por ejemplo, de pH 10 a 11).

10 o Ensayo de actividad de la enzima amilasa de las propiedades termofílicas, por ejemplo, desempeño hasta aproximadamente 65 °C a 70 °C.

La actividad de amilasa se puede medir ya sea mediante un ensayo de reducción de azúcar o al monitorear la fluorescencia a 520 nm (excitación a 485 nm) cuando se usó almidón BODIPY. Las velocidades iniciales se pueden calcular y convertir en un porcentaje de la velocidad máxima.

15 Ensayo de aplicación

Los experimentos pueden ser diseñados para evaluar la actividad y estabilidad de amilasas alcalinas de la invención en formulaciones de lavandería / ADW y con los componentes en forma individual. La actividad de amilasa se puede evaluar bajo condiciones que comprenden el agente quelante EDTA y/o peróxido de hidrógeno; las enzimas de punto de referencia comerciales pueden ser los controles.

20 Por ejemplo, las proteínas purificadas se incuban a 50 °C en presencia o en ausencia de EDTA 5 mM durante un tiempo(s) deseado(s), después del cual se mide la actividad de amilasa residual usando un sustrato soluble. La actividad en presencia de EDTA se expresa como el % de actividad en ausencia de agente quelante. En forma alternativa, la actividad enzimática en presencia de hidróxido de peróxido puede ser evaluada. Las proteínas purificadas se incuban a 50 °C en presencia o en ausencia de H₂O₂ 1 M durante un tiempo deseado, después del

25 cual se mide la actividad de amilasa usando almidón soluble. La actividad en presencia de hidróxido de peróxido se presenta como el % de actividad en ausencia de H₂O₂. La actividad enzimática puede ser analizada en solución de ADW (agua destilada, solución endurecedora, lejía, agentes quelantes, tensioactivos) con sustrato soluble (almidón BODIPY). Las proteínas purificadas pueden reaccionar con almidón soluble a 40 °C en presencia de una formulación de lavandería / ADW. Las velocidades iniciales se calculan sobre 5 minutos y se expresan como unidades
30 fluorescentes (FU)/s por ng de proteína.

Los ensayos de lavado con portaobjetos revestidos de almidón se pueden realizar de la siguiente manera: las proteínas purificadas se incuban con portaobjetos a 50°C durante 30 min en presencia de solución de ADW (agua destilada, solución de endurecimiento acuosa, lejía, agentes quelantes, tensioactivos). La remoción de almidón se mide comparando la pérdida de peso después del tratamiento enzimático con el peso inicial del portaobjetos.

35 Caracterización de ejemplos de amilasas

El gen que codifica la amilasa puede ser modificado para comprender un dominio de enlazamiento de carbohidrato / almidón. Las proteínas se pueden expresar con y sin una etiqueta de histidina en el terminal C y en un huésped de glicosilación y no glicosilación. Las enzimas se pueden expresar en combinaciones de etiqueta His / Huésped y pH y temperatura óptimos determinados. Las enzimas expresadas en un huésped de glicosilación con una etiqueta His se
40 pueden usar para analizar los experimentos. La presencia de la etiqueta His no debe afectar la actividad específica, sin embargo, la glicosilación puede dar como resultado una actividad específica ligeramente inferior que aquella sin glicosilación.

Ejemplo 15: identificación y caracterización de una amilasa termoestable

45 El siguiente ejemplo describe ejemplos de protocolos para la identificación y caracterización de amilasas termoestables.

En un estudio, se cribaron 350 aislados de hongos en medio sólido que contenía almidón granulado como la única fuente de carbono. Las cepas que hidrolizaron completamente el almidón o que mostraron crecimiento significativo se sometieron a aislamiento del ADNc y análisis proteómico de sus proteínas segregadas. Se empleó una

ES 2 575 912 T3

combinación de descubrimiento basado en la secuencia y análisis proteómico para recuperar las secuencias de ADN que codifican amilasas, glucoamilasas y glucosidasas que se demostró por medio de análisis proteómico que eran segregadas durante el crecimiento en almidón granulado.

- 5 Las secuencias recuperadas de ADNc de longitud completa se subclonaron para expresión en *Pichia pastoris* y las proteínas expresadas se caracterizaron en forma adicional con el ensayo BCA (que se usó para determinar el incremento en la concentración de los extremos reductores durante la hidrólisis del almidón por las amilasas) y el ensayo de Glucosa Oxidasa (GO) (que fue usado para detectar la glucosa liberada del almidón por las glucoamilasas); los ejemplos de protocolos para ambos ensayos se describen aquí.

Tabla 1

SEQ ID NOs:	Clase de actividad de la enzima	CBM20	Fuente (omo se determina mediante ARN 18S)
SEQ ID NO: 50 (codificada por, por ejemplo, la SEQ ID NO: 49)	amilasa	no	<i>Aspergillus terreus</i>
SEQ ID NO: 52 (codificada por, por ejemplo, la SEQ ID NO: 51)	amilasa	CBM20	<i>Aspergillus terreus</i>
SEQ ID NO: 54 (codificada por, por ejemplo, la SEQ ID NO: 53)	amilasa	no	<i>Aspergillus terreus</i>
SEQ ID NO: 4 (codificada por, por ejemplo, la SEQ ID NO: 3)	amilasa	CBM20	<i>Cochliobolus heterostrophus</i>
SEQ ID NO: 2 (codificada por, por ejemplo, la SEQ ID NO: 1)	amilasa	CBM20	<i>Cochliobolus heterostrophus</i>
SEQ ID NO: 32 (codificada por, por ejemplo, la SEQ ID NO: 31)	amilasa	no	<i>Penicillium chrysogenum</i> 100%
SEQ ID NO: 46 (codificada por, por ejemplo, la SEQ ID NO: 45)	amilasa	no	<i>Fusarium equiseti</i> 100%
SEQ ID NO: 22 (codificada por, por ejemplo, la SEQ ID NO: 21)	amilasa	no	<i>Penicillium expansum</i> 99%
SEQ ID NO: 24 (codificada por, por ejemplo, la SEQ ID NO: 23)	amilasa	no	<i>Penicillium chrysogenum</i> 100%
SEQ ID NO: 8 (codificada por, por ejemplo, la SEQ ID NO: 7)	glucoamilasa	no	<i>Cochliobolus heterostrophus</i>
SEQ ID NO: 16 (codificada por, por ejemplo, la SEQ ID NO: 15)	glucoamilasa	CBM20	<i>Cochliobolus heterostrophus</i>
SEQ ID NO: 14 (codificada por, por ejemplo, la SEQ ID NO: 13)	glucoamilasa	CBM20	<i>Fusarium verticillioides</i> GZ3639
SEQ ID NO: 18 (codificada por, por ejemplo, la SEQ ID NO: 17)	glucoamilasa	CBM20	<i>Fusarium verticillioides</i> GZ3639
SEQ ID NO: 10 (codificada por, por ejemplo, la SEQ ID NO: 9)	glucoamilasa	no	<i>Penicillium expansum</i> 99%

ES 2 575 912 T3

SEQ ID NOs:	Clase de actividad de la enzima	CBM20	Fuente (omo se determina mediante ARN 18S)
SEQ ID NO: 12 (codificada por, por ejemplo, la SEQ ID NO: 11)	glucoamilasa	CBM20	<i>Fusarium equiseti</i> 100%
SEQ ID NO: 26 (codificada por, por ejemplo, la SEQ ID NO: 25)	glucoamilasa	CBM20	<i>Penicillium verruculosum</i> 100%
SEQ ID NO: 20 (codificada por, por ejemplo, la SEQ ID NO: 19)	glucoamilasa	no	<i>Penicillium chrysogenum</i> 100%
SEQ ID NO: 28 (codificada por, por ejemplo, la SEQ ID NO: 27)	glucoamilasa	CBM20	<i>Fusarium merismoides</i> 99%
SEQ ID NO: 30 (codificada por, por ejemplo, la SEQ ID NO: 29)	glucoamilasa	CBM20	<i>Phoma herbarum</i> 99
SEQ ID NO: 34 (codificada por, por ejemplo, la SEQ ID NO: 33)	glucoamilasa	no	<i>Penicillium herquei</i> 99%
SEQ ID NO: 36 (codificada por, por ejemplo, la SEQ ID NO: 35)	glucoamilasa	CBM20	<i>Fusarium oxysporum</i> 100%
SEQ ID NO: 38 (codificada por, por ejemplo, la SEQ ID NO: 37)	glucoamilasa	CBM20	<i>Cordyceps ophioglossoides</i> 99%
SEQ ID NO: 40 (codificada por, por ejemplo, la SEQ ID NO: 39)	glucoamilasa	CBM20	<i>Penicillium chrysogenum</i> 100%
SEQ ID NO: 42 (codificada por, por ejemplo, la SEQ ID NO: 41)	glucoamilasa	no	<i>Cucurbitaria berberidis</i> 98%
SEQ ID NO: 48 (codificada por, por ejemplo, la SEQ ID NO: 47)	glucoamilasa	CBM20	<i>Aspergillus versicolor</i> 99
SEQ ID NO: 44 (codificada por, por ejemplo, la SEQ ID NO: 43)	α -glucosidasa		<i>Cochliobolus heterostrophus</i>
SEQ ID NO: 6 (codificada por, por ejemplo, la SEQ ID NO: 5)	α -glucosidasa		<i>Cochliobolus heterostrophus</i>

Los aislados fúngicos se criaron sobre medio sólido modificado de acuerdo con Marlida (2000) World J. Microbiol. Biotechnol. 16: 573 - 578. El medio contenía sales Czapek Dox y 1% de almidón resistente MIMAIZE^{260MR} (HiMaize²⁶⁰, National Starch & Chemical, Bridgewater, NJ) como la única fuente de carbono. El almidón MIMAIZE^{260MR} se agregó al agar frío sin esterilización a fin de conservar la estructura granulada el almidón. En forma adicional, se criaron las mismas cepas fúngicas en un medio con 0,5% de almidón rojo (Megazyme, Irlanda) a fin de identificar las cepas que segregaron enzimas amilolíticas.

Después de 5 días de cultivo, los aislados que se crecieron bien en el almidón resistente y segregaron enzimas degradantes de almidón, se visualizaron mediante una zona de aclaramiento en el medio de almidón rojo, se eligieron como "aciertos primarios". En la segunda etapa, se cultivaron los aciertos primarios en medio líquido con sales Czapek Dox y 1% de almidón resistente (MIMAIZE^{260MR}) como la única fuente de carbono durante aproximadamente 2 semanas. Las cepas que aclararon completamente la solución de almidón o que mostraron un crecimiento significativo en este medio se sometieron a aislamiento del ADNc y análisis proteómico de los

sobrenadantes del cultivo. La combinación de descubrimiento basado en la secuencia y análisis proteómico se empleó para recuperar las secuencias que codifican amilasas y glucoamilasas.

5 El descubrimiento basado en secuencia se realizó usando dos enfoques. En uno, los cebadores degenerados universales para la recuperación por PCR de los genes de amilasa y glucoamilasa se diseñaron sobre la base de la alineación de secuencias proteicas conocidas de hongos. En otro enfoque, los cebadores degenerados universales se combinaron en PCR con cebadores basados en las secuencias peptídicas obtenidas del análisis proteómico de los sobrenadantes del cultivo a partir de los aislados de hongos cultivados en almidón granulado. En ambos casos, los cebadores se usaron para PCR con ADN plantilla que comprendía ADNc elaborado de las mismas cepas de hongos que digieren el almidón. Por este medio, las secuencias génicas parciales fueron recuperadas. Los datos de los péptidos también se usaron para facilitar la recuperación de versiones de longitud completa de las secuencias parciales usando PCR RACE 5' y 3'.

15 Para las especies de hongos conocidas con genomas pre-secuenciados, se usó análisis proteómico para determinar las secuencias de péptidos derivadas de amilasas putativas usando el siguiente enfoque: 1) Después del análisis SDS-PAGE de las muestras derivadas de los extractos de hongos cultivados en almidón resistente como la única fuente de carbono, se sometieron las bandas individuales a digestión de proteasa; 2) El análisis de espectroscopia de masas de los péptidos recuperados se llevó a cabo en un instrumento LCQ; 3) Las secuencias peptídicas se determinaron por búsquedas en base de datos SEQUEST^{MR} (Sage-N Research, Inc. y Thermo Scientific).

20 Para las especies de hongos desconocidas, las secuencias peptídicas se determinaron por medio de un método similar al descrito con antelación con excepción de que las masas precisas de los péptidos y sus aminoácidos constituyentes se determinaron usando un instrumento QTOF. Las secuencias de los péptidos recuperados se usaron como plantilla para producir cebadores de oligonucleótidos degenerados para amplificación por PCR de los genes de ADNc elaborados a partir de la especie de hongos relevante cultivada en almidón granulado MIMAIZEZ^{260MR}.

25 Las secuencias de ADNc de longitud completa recuperadas que codifican amilasas y glucoamilasas se subclonaron para expresión en *Pichia pastoris* y las proteínas expresadas fueron caracterizadas adicionalmente usando el ensayo BCA para medir el incremento en la concentración de los extremos reductores durante hidrólisis de almidón por amilasas y el ensayo GO para la detección de glucosa liberada durante la hidrólisis de almidón por glucoamilasas.

30 Para determinar la actividad tanto de las amilasas como de las glucoamilasas, las reacciones de hidrólisis de almidón se realizaron de la siguiente manera: Los ensayos se realizaron por triplicado en una incubadora de mesa Eppendorf con agitación constante (800 rpm), a 37 °C y pH 5,0 en regulador de acetato de sodio 50 mM que contenía 1% de almidón granulado sin purificar. Las reacciones comenzaron al agregar la enzima a la mezcla de reacción. En diferentes puntos de tiempo, se retiraron alícuotas de las reacciones y se neutralizaron ya sea con adición de Tris 1M pH 7,5 (glucoamilasas) o reactivo BCA (amilasas).

35 Ensayo BCA para medir el incremento en la concentración de extremos reductores.

40 Las actividad de las alfa-amilasas se midió por la aparición de grupos reductores formados durante la hidrólisis del almidón. Se utilizó ensayo de ácido bicincoquínico o BCA (Cobre-BCA) de los azúcares reductores realizado de acuerdo con Wong (2000) Microassay for rapid screening of alpha-amylase activity, J. Agric. Food Chem. 48: 4540 - 4543; y, Fox (1991) Miniaturization of three carbohydrate analyses using a micro sample plate reader, Anal. Biochem. 195: 93 - 96.

45 Una alícuota de 10 µl de reacción de hidrólisis de almidón de la amilasa se neutralizó en 100 µl del reactivo BCA (que consiste de 64 mg/mL de carbonato de sodio monohidrato, 24 mg/mL de bicarbonato de sodio, 1,95 mg/mL de BCA, 1,24 mg/mL de sulfafo cúprico pentahidrato, 1,26 mg/mL de L-serina). El desarrollo del color ocurrió durante la incubación de la reacción neutralizada a 80°C durante 35 minutos y fue seguido por la determinación de la absorbancia a 560 nm. Las velocidades iniciales se calcularon en un tiempo de reacción de 50 min. Se construyó una curva estándar usando maltosa para correlacionar $A_{560\text{ nm}}$ con la concentración de los azúcares reductores generados (nmoles). La actividad específica se expresó como nmoles/min/µg de enzima.

Ensayo de glucosa oxidasa/geroxidasa (GO) para la cuantificación de glucosa liberada durante la hidrólisis de almidón:

50 La actividad de las glucoamilasas se midió por el ensayo modificado GO (glucosa oxidasa acoplada) de acuerdo con Bergmeyer, En Determination with Glucose Oxidase and Peroxidase; Bergmeyer, H. U., Ed.; Methods of. Enzymatic Analysis, 2da Ed.; 1974; páginas 1205 - 1212).

Las reacciones GO comenzaron mediante la adición de 10 µl de la reacción de hidrólisis de almidón neutralizada a 90 µl de PBS que contenía glucosa oxidasa (0,1 U/ml), peroxidasa (0,25 U/ml) y Amplex Red 0,05 mM, en placas negras Nunc de 96 pozos. Las placas se mantuvieron a temperatura ambiente en la oscuridad durante 30 min antes de la lectura en un lector de placas de fluorescencia con Ex/Em 545/590 nm. Se usó una curva estándar construida con glucosa para evaluar la cantidad de glucosa producida en las reacciones de hidrólisis. Las velocidades iniciales de hidrólisis de almidón (nmols de glucosa liberada de 1% de almidón granulado / min/ µg de glucoamilasa) se determinaron al graficar la cantidad de glucosa liberada en el tiempo y calcular la pendiente de la línea de mejor ajuste a través de los puntos de datos.

La influencia del pH en el rango de entre aproximadamente pH 3,5 y 6,0 sobre la hidrólisis de almidón por siete (7) glucoamilasas que sirven de ejemplo y dos amilasas que sirven de ejemplo de esta invención, se ilustra en la Figura 10 y en la Figura 11. La Figura 10 ilustra datos que muestran la influencia del pH sobre la hidrólisis de almidón granulado por siete (7) glucoamilasas que sirven de ejemplo de esta invención a 37°C; las velocidades iniciales se calcularon en 15 min y se convirtieron al porcentaje de la máxima velocidad observada. La Figura 11 ilustra los datos que muestran la influencia de pH sobre la hidrólisis de almidón granulado por la SEQ ID NO: 4 (codificada por, por ejemplo, la SEQ ID NO: 3) y la SEQ ID NO: 52 (codificada por, por ejemplo, la SEQ ID NO: 51) amilasas a 37 °C; las velocidades iniciales se calcularon en 50 min y se convirtieron al porcentaje de la máxima velocidad observada.

Otros ensayos también se pueden utilizar para caracterizar una enzima de la invención, y se describen algunos ejemplos de protocolos, a continuación:

Ejemplo de protocolo de extracción de ácido nucleico: los microorganismos, por ejemplo, hongos filamentosos, son cultivados en medio de cultivo líquido. La biomasa se recolecta y el ADN genómico de alto peso molecular se aísla usando DNEASY^{MR} (DNeasy) Plant Maxi Kit (Qiagen, Valencia, CA) que usa protocolos estándares. El ARN total puede ser aislado usando RNEASY^{MR} (RNeasy) Plant Mini Kit (Qiagen) utilizando protocolos estándar.

Ejemplo de protocolo de construcción de biblioteca: El ADN genómico puede ser parcialmente digerido con enzimas de restricción y los fragmentos entre 1-10 kb pueden ser purificados para la construcción de una biblioteca genómica. Los fragmentos pueden ser ligados en el vector Lambda Zap Express^{MR} (Stratagene, San Diego, CA) y empacados en un fago infectable de acuerdo con las instrucciones del fabricante.

Ejemplo de protocolo de cribado de bibliotecas: Las bibliotecas Lambda se pueden usar para infectar células XL1 Blue MRF^{MR} (Stratagene) en agar de superficie. Aproximadamente 50000 pfu de fago se pueden agregar a 600 µl de células OD₆₀₀ = 1. La mezcla se incuba a 37 °C durante 15 minutos en un baño de agua y luego se agrega a 6 ml de 0.7% de agar de superficie fundido y se siembra en placas de agar NZY. La placa se incuba luego durante la noche a 39 °C. Un círculo de nylon (F. Hoffmann-La Roche Ltd., Basel Suiza) puede ser dispuesta encima de la superficie de la plaqueta resultante y se levanta con parte de los fagos adheridos al nylon. El nylon puede ser sumergido en NaCl 1.5 M, NaOH 0,5 M durante 2 minutos, NaCl 1,5 M, Trís 0,5 M pH 7,6 durante 5 minutos y 2X SSC, Tris 0,2 M pH 7,6 durante 30 segundos. El filtro de nylon luego se entrecruza mediante UV, por ejemplo, en un entrecruzador Stratagene.

Los fragmentos de PCR de los genes de amilasa, por ejemplo, glucoamilasas, pueden usar sondas, por ejemplo, usando Expand High Fidelity PCR Kit^{MR} (Roche) que usa 30 ciclos de 95 °C durante 20 segundos, 55 °C durante 30 segundos y 72 °C durante 1 minuto en un termociclador. El fragmento de PCR aislado se puede preparar como una sonda radioactiva usando el Prime It Kit^{MR} (Stratagene) siguiendo las instrucciones del fabricante.

Los elementos de la biblioteca extraídos con el filtro se lavan en una solución de prehibridación (por ejemplo, DIG EASY HYB^{MR}, Roche) durante 2 horas a 42 °C en un horno de hibridación (Robbins). La sonda se puede agregar a 15 ml de DIG EASY HYB^{MR} fresco y se usa para reemplazar la solución de prehibridación. El filtro se lava con sonda durante la noche, por ejemplo a 45 °C. La sonda puede ser removida, y el filtro lavado una vez con 2X SSC, 0,1% de SDS durante 15 minutos y dos veces con 0,1X SSC, 0,1% de SDS durante 15 minutos cada vez. El filtro de nylon puede ser expuesto a película de rayos X durante la noche a -80 °C. Después del desarrollo, las manchas de hibridación sobre la película de rayos X se pueden usar para identificar los clones de la placa original. Un tapón de agar puede ser tomado de la placa donde se alinean las manchas y se suspende en regulador SM para liberar el fago en la solución. Varias plaquetas aisladas correspondientes a los fragmentos genómicos que contienen todo o parte de un gen de amilasa pueden ser aisladas.

100 µl del patrón de fago aislado pueden ser agregados a 200 µl de células XL-1 BLUE MRF^{MR} (Stratagene) y 1 µl de fago auxiliar de EXASSIST^{MR} (Stratagene). La mezcla se puede incubar a 37 °C durante 15 minutos y se pueden añadir 3 ml de medio 2X YT. Esta se puede incubar a 37 °C con agitación durante 2.5 horas. La mezcla se puede calentar durante 20 minutos a 70 °C y se enfría sobre hielo. 100 µl de la mezcla se puede remover y agregar a 200 µl de células SOLR (Stratagene) y se incuban a 37 °C durante 15 minutos. 50 µl se pueden cultivar en placas de LB de kanamicina (50 µg/ml) e incubarse durante la noche a 37 °C. Las colonias resultantes pueden contener fragmentos genómicos clonados en el plásmido pBK-CMV.

Ejemplo de protocolo de secuenciación: El secuenciamiento de ADN sobre los clones candidatos se pueden realizar con el secuenciamiento de ciclo BIGDYE TERMINATOR^{MR} Kit^{MR} VERSION 2.0 (Applied Biosystems, Foster City, CA) y un analizador de ADN 3700^{MR} (Applied Biosystems) usando los protocolos del fabricante. Los intrones potenciales pueden ser identificados al comparar esta secuencia con secuencias consenso para intrones en amilasas conocidas.

Ejemplo de protocolo de síntesis de ADNc: Los cebadores de PCT se usan en una reacción de síntesis de ADNc usando un Kit rtPCR^{MR} THERMOSCRIPT^{MR} (Invitrogen) usando los protocolos del fabricante.

Ejemplo de protocolo de clonación de expresión: Los cebadores PCT se usan para generar un fragmento de PCR usando el clon de ADNc como una plantilla utilizando 30 ciclos de 95 °C durante 20 segundos, 55 °C durante 30 segundos, 72 °C durante 2 minutos, usando el Kit de PCR EXPAND HIGH FIDELITY^{MR} (Roche) y los protocolos del fabricante. Los fragmentos de PCR son digeridos con las enzimas de restricción y ligados en los sitios de restricción correspondientes de un plásmido, por ejemplo, pPIC Z^{MR} alfa (Invitrogen). El constructo puede ser transformado en una levadura, por ejemplo, cepa X-33^{MR} de *Pichia pastoris* (Invitrogen) donde el constructo se integra de manera estable en el cromosoma de *Pichia*. La selección puede basarse en la resistencia a la zeocina. El constructo se puede diseñar de manera tal que el clon de *Pichia* puede ser inducido con metanol para secretar la amilasa madura en el medio. Un cultivo de 1 litro del clon de expresión puede ser inoculado con un cultivo inicial de levadura durante la noche en BMGY y cultivado durante la noche a 30°C en un termo de agitación. Las células de levadura se recolectan por centrifugación al siguiente día y se resuspenden en 1 litro de BMMY. Las células se cultivan a 30 °C en el termo de agitación durante 3 días con metanol agregado a 0,5% final cada 24 horas. El medio que contiene la enzima de glucoamilasa expresada se recolecta luego y se analiza en un ensayo de actividad de glucoamilasa y se somete a electroforesis SDS PAGE usando protocolos estándares para determinar el tamaño de la proteína.

Los cebadores también pueden estar diseñados para la sobreexpresión en *Escherichia coli*. Los cebadores PCR se usan para generar un producto de PCR como antes, a partir de la plantilla de ADNc. El fragmento de PCR puede ser digerido con las enzimas de restricción y ligado en los correspondientes sitios de restricción del plásmido, por ejemplo, un pSE420 (Invitrogen). El constructo puede ser transformado en *Escherichia coli*, por ejemplo, la cepa XL-1 Blue MR (Stratagene). La selección para el plásmido se puede basar en la resistencia a la ampicilina. El gen de amilasa puede estar bajo el control de un promotor lac-z y puede ser inducido con IPTG (isopropil-tio galactopiranosido). El constructo puede ser diseñado de manera tal que el gen de glucoamilasa maduro sea expresado en la célula de *Escherichia* y contendrá un residuo extra de metionina como en el terminal N.

Ejemplo de ensayo "estándar": se pueden agregar alícuotas de enzima a una solución de 5 mM de regulador, 3 mM de malto-oligosacáridos (Sigma, M-3639) en un baño de agua. Las alícuotas de 100 µl se pueden remover en los puntos de tiempo con 200 µl del reactivo glucosa oxidasa (Sigma, GAGO-20) y se incuban a 37 °C, 30 min. La reacción puede ser detenida con adición de ácido sulfúrico 12 N y se determina la absorbancia a 540 nm. La versión de longitud completa de la enzima puede ser analizada para utilización del pH, temperatura y utilización de sustrato.

Ejemplo de ensayo de "actividad": la actividad de las enzimas se puede medir por la liberación de glucosa libre a partir de un sustrato de oligo-dextrina. La glucosa liberada puede ser oxidada en una reacción de acoplamiento que genera un producto coloreado. Una alícuota de enzima se puede agregar a la solución de regulador 5 mM, malto-oligosacáridos 3 mM (Sigma, M-3639) en un baño de agua. 100 µl de alícuotas pueden ser removidos en puntos de tiempo con 200 µl de reactivo de glucosa oxidasa (Sigma, GAGO-20) y se incuban a 37 °C, 30 min. La reacción se detiene con la adición de ácido sulfúrico 12 N y se determina la absorbancia a 540 nm. Los puntos de tiempo son luego graficados para determinar la velocidad relativa para la reacción. Perfil de pH: se puede usar regulador de acetato (pH 4,0, 4,5, 5,0 y 5,4) así como también regulador de fosfato (pH 6,2, 7,0, 8,1) en un ensayo de actividad para determinar la velocidad relativa para la glucoamilasa a cada pH.

Perfil de temperatura: la velocidad relativa de la enzima a varias temperaturas (por ejemplo, 50 °C, 60 °C, 70 °C, 80 °C y 85 °C) se puede determinar en regulador de acetato, pH 5,3.

Datos de estabilidad de temperatura: La enzima puede ser agregada a regulador de acetato 5 mM a una temperatura indicada deseada. Las alícuotas de enzimas se pueden ser removidas colocadas sobre hielo en intervalos de 4 minutos. Las alícuotas luego se analizan en cuanto a la actividad sobre el sustrato durante 20 minutos a 70 °C.

Utilización del sustrato: Las dextrinas maltosa, maltotriosa, panosa, maltotetraosa y maltoheptaosa se pueden sustituir por los malto-oligosacáridos en el ensayo de actividad para evaluar la utilización del sustrato de una amilasa, por ejemplo, una glucoamilasa. La velocidad de liberación de glucosa para varios sustratos se puede analizar en regulador de acetato 5 mM, 70 °C.

Ejemplo 16: ensayo de actividad de amilasa: ensayos de extremos reductores de BCA

ES 2 575 912 T3

El siguiente ejemplo describe un ejemplo de un método para determinar si un polipéptido está dentro del alcance de la invención, por ejemplo, para un ensayo de extremos reductores de BCA. La actividad de amilasa (incluyendo, por ejemplo, glucoamilasa) se puede determinar usando la siguiente metodología.

1. Preparación de 2 soluciones de sustrato de la siguiente manera:

- 5 a) 2% de solución de almidón soluble (almidón de maíz granulado o patata) de pH 8, mediante la disolución de 2 g de almidón de patata en 100 ml de fosfato de sodio 100 mM, pH 8).
- b) 2% de solución de almidón soluble (patata) pH 10, mediante la disolución de 2 g de almidón de patata en 100 ml de carbonato de sodio 100 mM.

10 Se calientan ambas soluciones en un baño de agua en ebullición, mientras se mezcla, durante 30 - 40 minutos hasta que el almidón de disuelva.

2. Se prepara una Solución A de 64 mg/ml de carbonato de sodio monohidrato, 24 mg/ml de bicarbonato de sodio y 1,95 mg/ml de sal disódica de BCA (4,4'-dicarboxi-2,2'-biquinolina) disódica (Sigma Chemical cat. # D-8284). Se añaden por encima a dH₂O.

15 3. Se prepara una Solución B mediante la combinación de 1,24 mg/ml de sulfato cúprico pentahidrato y 1,26 mg/ml de L-serina. Se añade la mezcla a dH₂O.

4. Se prepara un reactivo de trabajo en una proporción 1:1 de soluciones A y B.

20 5. Se prepara una solución estándar de maltosa 10 mM en dH₂O donde la maltosa 10 mM se combina en 2% de almidón soluble a pH deseado hasta una concentración final de 0, 100, 200, 300, 400, 600 µM. Se genera una curva estándar para cada conjunto de puntos de tiempo. Como la curva se determina al agregar 10 µl de los estándares al reactivo de trabajo, que funciona a 0, 1, 2, 3, 4, 6 nmole de maltosa.

6. Se toma una alícuota de 1 ml de la solución del sustrato en tubos de microcentrifuga, equilibrar a temperatura deseada (5 min) en bloque de calor o baño de agua caliente. Se añaden 50 µl de solución enzimática al interior de la tapa del tubo.

25 7. Mientras la solución se equilibra, se mezclan 5 ml de ambas soluciones A & B. Se colocan alícuotas de 100 µl en una placa de PCR de 96 pozos. Se coloca la placa sobre hielo.

8. Después de 5 min de equilibrio de la temperatura, se cierra la tapa de los tubos, invirtiéndolos y agitándolos en forma de vórtice 3 veces. Inmediatamente se coloca una alícuota de 10 µl en la placa que se toma como t=0 (punto de tiempo cero). Se deja la mezcla de la enzima en un bloque de calor y se toman alícuotas de 10 µl en cada punto de tiempo deseado (por ejemplo, 0, 5, 10, 15, 20, 30 minutos).

30 9. Se asegurar que 12 pozos queden vacíos (únicamente para alícuotas del reactivo de trabajo) durante la adición de 10 µl de estándares, para la curva estándar.

35 10. Cuando todos los puntos de tiempo se recolectan y se agregan los estándares, se cubre la placa y se calienta a 80 °C durante 35 min. Se enfría la placa sobre hielo durante 10 min. Se agregan 100 µl de H₂O a todos los pozos. Se mezclan y se colocan alícuotas de 100 µl en una placa de 96 pozos de fondo plano y se lee la absorbancia a 560 nm.

40 11. Fijar en cero los puntos de tiempo de cada muestra contra su propio t=0 (restar el valor promedio t=0 de A560 de los otros valores promedios de A560). Convertir la A560_(experimental) en µmol (Dividir A560_(experimental) por la pendiente de la curva estándar (A560/µmol). Se genera una pendiente de los puntos de tiempo y el µmol (en µmoles/min), se multiplica por 100 (ya que el valor µmol solamente representa 10 µl usados en el ensayo, no la cantidad preparada en 1 ml rxn). Para obtener la actividad específica, se divide la pendiente (en µmol/min) por los mg de proteína. Todos los puntos deben ser hechos como mínimo por duplicado, referiblemente por triplicado. Se divide la concentración de proteína (mg/ml) por cualquier dilución para obtener los mg utilizados en el ensayo. Se divide la pendiente anterior por los mg usados en el ensayo para obtener la actividad específica. Véase, por ejemplo, Wong (2000) J. Agric. Food Chem. 48: 4540 - 4543; Fox (1991) Anal. Biochem. 195, 93 - 96.

45 **Ejemplo17:** Cribado para actividad de amilasa

ES 2 575 912 T3

El siguiente ejemplo describe un método de ejemplo para determinar si un polipéptido está dentro del alcance de la invención. La actividad de amilasa de los clones se puede evaluar mediante una cantidad de métodos conocidos en el arte. El siguiente es un ejemplo de metodología que puede ser utilizada.

- 5 El número de plaquetas cribadas, por placa, puede ser aproximadamente de 10000 pfu. Para cada biblioteca de ADN: aproximadamente 50000 plaquetas por biblioteca aislada y 200000 plaquetas por biblioteca no aislada se pueden cribar dependiendo del título de pfu para el lisado amplificado λ Zap Express.

Determinación del título de biblioteca Lambda

1) Se añadieron μ L del patrón de la biblioteca Lambda Zap Express amplificada a 600 μ L de células MRF' de *E. coli* ($OD_{600} = 1,0$). Para diluir el patrón de MRF', se usó $MgSO_4$ 10 mM.

- 10 2) Se incuba a 37 °C durante 15 minutos.

3) Se transfiere la suspensión a 5 - 6 mL de agar de superficie NZY a 50 °C y se mezcla suavemente.

Se vierte inmediatamente la solución de agar sobre placas grandes (150 mm) de medio NZY.

4) Se permite que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego se invierte la placa.

- 15 5) Se incuba la placa a 39 °C durante 8 - 12 horas.

6) El número de plaquetas es aproximado. Se determina el título de fagos para obtener 10000 pfu/placa.

7) Se diluye una alícuota de fagos de la biblioteca con regulador SM si se requiere.

Cribado del sustrato

- 20 - Se añade Lambda Zap Express (50000 pfu) de la biblioteca amplificada a 600 μ L de células MRF de *E. coli* ($OD_{600} = 1,0$). Para las bibliotecas no ambientales, se preparan 4 tubos (50,000 pfu por tubo).

- Se incuba a 37 °C durante 15 minutos.

- Mientras se incuba la suspensión de fagos/células, se añade 1,0 mL de sustrato de almidón rojo (1,2% p/v) a 6,0 mL de agar de superficie NZY a 50 °C y se mezcla completamente. Se mantiene la solución a 50°C hasta que se necesite.

- 25 - Se transfiere 1/5 (10000 pfu) de la suspensión celular a la solución de sustrato/agar de superficie y se mezcla suavemente.

- Se vierte inmediatamente a solución sobre placa de medio NZY grande (150 mm).

- Se permite que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego se invierte la placa.

- 30 - Se repiten los procedimientos 4 - 6 cuatro veces para el resto de la suspensión celular (1/5 de la suspensión cada vez).

- Se incuban las placas a 39 °C durante 8 - 12 horas.

- Se observa la placa por las zonas de aclaramiento (halos) alrededor de las plaquetas.

- 35 - Las plaquetas con halos se separan del agar y se transfieren a un micro tubo estéril. Una punta de pipeta de 200 μ L de orificio grande funciona bien para remover (separar) el tapón de agar que contiene la plaqueta deseada.

- Se resuspenden los fagos en 500 μ L de regulador SM. Se añaden 20 μ L de cloroformo para inhibir cualquier crecimiento celular adicional.

- Se incuba la suspensión pura de fagos a temperatura ambiente durante 4 horas o durante la noche antes de la siguiente etapa.

ES 2 575 912 T3

Aislamiento de clones puros

- Se añaden 10 µL de la suspensión de fagos re-suspendida en 500 µL de células MRF' de *E. coli* (OD₆₀₀ = 1,0).
- Se incuba a 37 °C durante 15 minutos.
- Mientras se incuba la suspensión de fagos/células, se añade 1 mL de sustrato de almidón rojo (1,2% p/v) a 6,0 mL de agar de superficie NZY a 50 °C y se mezcla completamente. Se mantiene la solución a 50 °C hasta que se necesite.
- Se transfiere la suspensión de células a la solución de sustrato/agar de superficie y se mezcla suavemente.
- Se vierte inmediatamente la solución sobre una placa grande (150 mm) de medio NZY.
- Se permite que el agar de superficie se solidifique completamente (aproximadamente 30 minutos), luego se invierte la placa.
- Se incuba la placa a 39°C durante 8 - 12 horas.
- Se observa la placa para una zona de clarificación (halo) alrededor de a sola plaqueta (clon puro). Si no se puede aislar una sola plaqueta, se ajusta el título y se siembra en placa nuevamente la suspensión de fagos.
- Se separa una sola plaqueta con halo de agar y se la transfiere a un micro tubo estéril. Una punta de pipeta de 200 µl de orificio grande funciona bien para remover (separar) el tapón de agar que contiene la plaqueta deseada. Para amplificar el título, se separan 5 plaquetas activas individuales en un micro tubo.
- Se resuspenden los fagos en 500 µL de regulador SM. Se añaden 20 µL de cloroformo para inhibir cualquier crecimiento celular adicional.
- Se incuba la suspensión pura de fagos a temperatura ambiente durante 4 horas o durante la noche antes de la siguiente etapa. La suspensión de fagos puro se almacena a -80 °C mediante la adición de DMSO en la suspensión de fagos (7% v/v).

Escisión de clones puros

- Se añaden 100 µl de la suspensión pura de fagos a 200 µl de células MRF de *E. coli* (OD₆₀₀ = 1,0). A esta suspensión, se le añade 1,0 µL del fago auxiliar EXASSIST (>1 x 10⁶ pfu/mL; Stratagene). Usar tubos Falcon 2059 para la escisión.
- Se incuba la suspensión a 37°C durante 15 minutos.
- Se añaden 3,0 mL de medio YT 2 veces a la suspensión de células.
- Se incuba a 30°C durante al menos 6 horas o durante la noche con agitación.
- Se transfiere el tubo a 70°C durante 20 minutos. La suspensión de fagos se puede almacenar a 4°C durante 1 a 2 meses.
- Se transfieren 100 µl de la suspensión del fagémido a un micro tubo que contiene 200 µL de células Exp 505 de *E. coli* (OD₆₀₀ = 1,0).
- Se incuba la suspensión a 37°C durante 15 minutos.
- Se añaden 300 µL de SOB a la suspensión.
- Se incuba la suspensión a 37°C durante 30 a 45 minutos.
- Se transfieren 00uL de la suspensión a una placa pequeña (90 mm) de medio LB) que contiene Kanamicina (medio LB con Kanamicina a razón de 50 µg/mL) para bibliotecas de ADN Zap Express o Ampicilina (medio LB con Kanamicina a razón de 100 µg/mL) para bibliotecas de ADN Zap II.
- Se transfiere el resto de la suspensión a otra placa pequeña de medio LB.

ES 2 575 912 T3

- Se usan perlas estériles de vidrio para distribuir uniformemente la suspensión sobre la placa.
 - Se incuban las placas a 30°C durante 12 a 24 horas,
 - Se observan las colonias en las placas.
 - Se inocula una sola colonia en medio líquido LB que contiene un antibiótico adecuado e se incuba a 30°C durante 12 a 24 horas.
- 5
- Se puede preparar un patrón de glicerol mediante la adición de 80% de glicerol en medio de cultivo líquido (15% v/v) y se almacena a -80°C.

Verificación de la actividad

- 10
- Se transfieren 50 µL de cultivo líquido a un microtubo. Se añaden 500 µL de 8% de Amilopectina Azure pH7 en el mismo tubo. Se preparan 2 tubos para cada clon.
 - Se analiza la actividad a 50°C durante 3 horas y durante la noche. Se utiliza regulador de pH 7 como control.
 - Se enfría el espécimen de prueba en un baño de agua helada durante 5 minutos.
 - Se añaden 750 µL de etanol y se mezcla completamente.
 - Se centrifuga a 13000 rpm (16000 g) durante 5 minutos.
- 15
- Se mide la OD del sobrenadante a 595 nm.

Análisis RFLP

- Se transfiere un 1,0 mL de cultivo líquido a un microtubo esterilizado.
 - Se centrifuga a 13200 rpm (16000 g) durante 1 minuto.
 - Se descarta el sobrenadante. Se añade otro 1,0 mL de cultivo líquido en el mismo microtubo esterilizado.
- 20
- Se centrifuga a 13200 rpm (16000 g) durante 1 minuto.
 - Se descarta el sobrenadante.
 - Se sigue el protocolo del mini kit de centrifugación QIAPREP^{MR} para el aislamiento del plásmido.
 - Se revisa la concentración de ADN usando BioPhotometer.
 - Se usa Sac I y Kpn I para la primera digestión doble. Se incuba a 37°C durante 1 hora.
- 25
- Se usa Pst I y Xho I para la segunda digestión doble. Se incuba a 37°C durante 1 hora.
 - Se añade colorante de carga en la muestra digerida.
 - Se corre la muestra digerida en un gel de agarosa al 1,0% durante 1 - 1,5 hora a 120 voltios.
 - Se observa el gel con un visualizador. Se pueden analizar las secuencias de todos los clones con un patrón diferente de digestión.

30 **Ejemplo 18:** Ejemplo de protocolos para purificación de enzimas

Los siguientes ejemplos describen protocolos de ejemplo para la purificación de enzimas de esta divulgación.

SEQ ID NO: 52: 6 g de sobrenadante liofilizado de cultivo de *P. pastoris* (véase discusión en el Ejemplo 25, a continuación) que expresa la SEQ ID No: 52 se suspendió en 24 mL de H₂O y se precipitó con etanol frío. El producto precipitado se re-suspendió en aproximadamente 40 mL de H₂O y se dializó O/N contra agua. Después de

la diálisis, se agregó un regulador de acetato concentrado a pH 6,0 a la muestra para obtener una concentración final de 50 mM. La proteína unida a la columna (Q SEPHAROSE^{MR}; resina Amersham Pharmacia vertida en una columna XK 50^{MR}) en acetato de Na 50 mM pH 6,0 y se eluyó durante un gradiente entre 0 y 400 mM de NaCl en acetato de Na de pH 6,0. Se removieron las proteínas contaminantes de la columna con NaCl 1 M en acetato de Na pH 6,0. La elusión de amilasa de Q SEPHAROSE^{MR} se rastreó con SDS PAGE y los ensayos de actividad usando almidón BODIPY como sustrato. Se realizaron múltiples purificaciones para obtener suficiente proteína para cumplir con el requisito de aproximadamente 1 g. Las fracciones purificadas a partir de estas tandas se reunieron y se concentraron por concentrador de células de agitación.

SEQ ID NO: 48: 24 g de sobrenadante liofilizado de cultivo *P. pastoris* que expresa la SEQ ID NO: 48 se suspendió en 20 mL de regulador de carbonato 50 mM pH 10,0; NaCl 100 mM y se precipitó con etanol frío. El producto precipitado se re-suspendió en aproximadamente 20 mL de regulador de acetato 50 mM pH 5,2; NaCl 500 mM y se dializó O/N contra regulador málico 50 mM pH 3,5; NaCl 500 mM. La proteína se unió a la columna de agarosa -amilosa (NEB) de 100 mL en regulador málico 50 mM pH 3,5; NaCl 500 mM y se eluyó con 0,5% de dextrina de maíz en regulador de carbonato 50 mM pH 10,0; NaCl 50 mM. Las proteínas contaminantes se extrajeron de la columna con regulador de fosfato de sodio 50 mM pH 7,5; NaCl 100 mM. La elusión de glucoamilasa de la agarosa-amilosa se rastreó con SDS PAGE.

Otras enzimas de la invención, pueden ser purificadas usando estos protocolos o variaciones de los mismos o protocolos análogos.

La Figura 26 ilustra tablas que resumen la eficiencia de estos protocolos de purificación para las SEQ ID NO: 4, SEQ ID NO: 18, SEQ ID NO: 26 que sirven de ejemplo y datos de actividad correspondientes, entre otros, almidón sin purificar y almidón soluble comparando la enzima purificada y la no purificada.

Las preparaciones de enzimas purificadas de las SEQ ID NO: 52 y SEQ ID NO: 26 que sirven de ejemplo se analizaron en cuanto a la actividad (la capacidad para hidrolizar) en fermentaciones de almidón sin purificar; estos datos sobre la pureza (determinados mediante análisis densitométrico), concentración, cantidad y regulador de almacenamiento se resumen:

	SEQ ID NO: 52	SEQ ID NO: 48
Clase	Alfa-Amilasa	Glucoamilasa
Regulador de almacenamiento	GMP*	GMA**
Concentración	25 mg/mL	25 mg/mL
Volumen enviado	40 mL	20 mL
Cantidad Total	1,0 g	0,5 g
Pureza	> 90%	> 95%
* GMP - 20% de glucosa; 0,1% de metil parabeno, PBS pH 7,0		
** GMA - 20% de glucosa; 0,1% de metil parabeno; acetato de sodio 50 mM pH. 5,2; NaCl 100 mM.		

Ejemplo 19: Ejemplo de protocolo para licuefacción de almidón y medición de resultados

El siguiente ejemplo describe ejemplos de protocolos para licuefacción de almidón usando, por ejemplo enzimas de esta invención. Condiciones de reacción: PO₄ 100 mM pH 6,5, 1% (p/p) de almidón licuado DE 12 a 55°C. Tanto los ensayos TLC como HPLC se pueden realizar para verificar la actividad.

Un ejemplo de protocolo para la sacarificación de almidón licuado a un pH de 6,5:

- Se ajusta el pH del almidón licuado al pH al cual la(s) sacarificación(es) se va a realizar. Se licua el almidón en regulador de acetato de sodio 100 mM, pH 4,5 con sales de fosfato de sodio 100 mM agregadas de manera tal que, antes de la sacarificación, el pH pudiera ser ajustado a pH 6,5.

- Se pesan muestras de 5 gramos de almidón licuado en botellas taradas.
- Se usa 0,04% (p/p) de OPTIDEX L-400^{MR} o aproximadamente 400 mL de un patrón diluido 1 - 10 veces de OPTIDEX L-400^{MR} por 100 gramos de almidón licuado.
- Se calculan los miligramos de OPTIDEX L-400^{MR} contenidos en 400 mL de reserva diluida 1 - 10 de OPTIDEX L-400^{MR}. A continuación, se calcula el volumen de los lisados requeridos para formar la misma concentración de enzimas que el OPTIDEX L-400^{MR}.
- Se añaden enzimas a las muestras de almidón licuado y se incuban a temperatura deseada (50 °C). Después de 18 horas se determina DE y se prepara una muestra para análisis por HPLC.

Un ejemplo para la determinación de DE:

10 **Ejemplo de ensayo de neocuproína:**

Se agrega una muestra de 100 ml a 2,0 ml de solución A de neocuproína (40 g/L de carbonato de sodio, 16 g/L de glicina, 0,45 g/L de sulfato de cobre). A esta solución se le agrega 2,0 ml de solución B de neocuproína B (1,2 g/L de clorhidrato de neocuproína Sigma N-1626). Los tubos se mezclan y calientan en un baño de agua en ebullición durante 12 minutos; se enfrían, se diluyen hasta un volumen de 10 ml con agua DI y se lee la OD a 450 nm en el espectrofotómetro. El equivalente de glucosa en la muestra se extrapola a partir de la respuesta de 0,2 mg/ml de glucosa estándar que se corre en forma simultánea.

Ejemplo de análisis por HPLC:

Los perfiles de sacarificación de un carbohidrato se miden por HPLC (Bio-Rad Aminex HPX-87A columna de plata, 80°C) usando un detector de índice de refracción. La fase móvil se filtra con agua Millipore a una velocidad de flujo de 0,7 ml/min. Las muestras de sacarificación se diluyen 1 - 10 con agua DI acidificada (5 gotas de HCl 6 M en 200 mL de agua DI) luego se filtran a través de un filtro de jeringa de 0,45 µm. El volumen de inyección es 20 µL.

Ejemplo de TLC:

Los productos de reacción fueron w/d en puntos de tiempo cada hora y luego se colocaron y secaron en una placa de TLC. La placa luego se desarrolló en una proporción 10:90 de agua:isopropanol y se visualizó con una mancha de vanilina o una mancha de CAM y luego se calentó para mostrar los azúcares reducibles. El almidón licuado fue parcialmente hidrolizado hasta glucosa en los casos en que se observó actividad.

Ejemplo 20: Licuefacción de almidón usando glucoamilasas

Este ejemplo describe un método de ejemplo de la invención para licuefacción de almidón usando glucoamilasas. El concentrado de glucoamilasa se puede preparar a partir de los caldos de fermentación por tratamiento térmico, lavado celular, extracción alcalina usando microfiltración y ultrafiltración (48% de rendimiento total). El concentrado de UF puede neutralizarse con ácido acético y formularse con 30% de glicerol a un pH de 4,5. El nivel de actividad de un producto comercial puede ser aproximadamente de 120 U¹/g 0,5 kg/tonelada de almidón.

Ejemplo de un ensayo de actividad de glucoamilasa

Una cubeta de 1 mL que contiene 950 µL de MOPS 50 mM pH 7,0 que contiene pNP-α-D-hexa-(1->4)-glucopiranosido 5 mM se coloca en el controlador de temperatura Peltier del espectrofotómetro Beckman DU- 7400 precalentado a 80°C. El espectrofotómetro queda en cero a 405 nm y se agregan 50 µL de la solución enzimática a la cubeta, se mezcla bien y se monitorea el aumento en la OD_{405 nm} en un intervalo de un minuto. La velocidad de ΔOD_{405 nm}/min se convierte a una unidad estándar de µmol/minuto a partir de la respuesta de OD_{405 nm} de 50 µL de 1 µmol/mL de PNP en 950 mL de MOPS 50 mM a pH de 7,0, -80°C. Una unidad estándar de alfa glucoamilasa termoestable (DTAA) es igual a la cantidad de enzima que catalizará la liberación de 1 µmol/mL/minuto de pNP bajo las condiciones definidas del ensayo.

Ensayo de actividad de glucoamilasa estándar

Una cubeta de 1 mL que contiene 950 µL de MOPS 50 mM pH 7,0 que contiene pNP-α-D-glucopiranosido 5 mM se coloca en el controlador de temperatura Peltier del espectrofotómetro DU- 7400^{MR} (Beckman) precalentado a 60°C. El espectrofotómetro queda en cero a 405 nm y se agregan 50 µL de la solución enzimática a la cubeta, se mezcla bien y se monitorea el aumento en la OD_{405 nm} en un intervalo de un minuto. La velocidad de ΔOD_{405 nm}/min se convierte a una unidad estándar de µmol/minuto a partir de la respuesta de OD_{405 nm} de 50 µL de 1 µmol/mL de pNP

ES 2 575 912 T3

en 950 mL de MOPS 50 mM a pH 7,0, -60°C. Una unidad Diversa estándar de glucoamilasa (DGA) es igual a la cantidad de enzima que catalizará la liberación de 1 $\mu\text{mol/mL/minuto}$ de pNP bajo las condiciones definidas del ensayo.

Determinación del equivalente de dextrosa

- 5 El método de ensayo de neocuproína se puede usar para medir el DE. Las muestras seleccionadas se pueden medir por un procedimiento que se describe aquí, y mediante un análisis GPC usando el procedimiento de Fehlings de GPC.

Ensayo de neocuproína

- 10 Se agrega una muestra de 100 ml a 2,0 ml de solución A de neocuproína (40 g/L de carbonato de sodio, 16 g/L de glicina, 0,45 g/L de sulfato de cobre). A esta solución se le agrega 2,0 ml de solución B de neocuproína B (1,2 g/L de clorhidrato de neocuproína Sigma N-1626). Los tubos se mezclan y calientan en un baño de agua en ebullición durante 12 minutos; se enfrían, se diluyen hasta un volumen de 10 ml con agua DI y se lee la OD a 450 nm en el espectrofotómetro. El equivalente de glucosa en la muestra se extrapola a partir de la respuesta de 0,2 mg/ml de glucosa estándar que se corre en forma simultánea.

- 15 La muestra de almidón se diluye ~1 a 16 con agua DI con la dilución exacta registrada. Se agregan 10 mililitros de la muestra diluida a 20 ml de agua DI. Se agregan 10 mililitros de solución A y B de Fehling al almidón diluido. La muestra ebulle durante 3 minutos y se enfría sobre hielo. Se agrega 10 mililitros de KI al 30% y 10 ml de H_2SO_4 6 N. La solución se titula contra tiosulfato de sodio 0,1 N. El volumen de la solución de titulación se registra y se usa para calcular el DE.

- 20 Determinación de almidón residual

Las muestras de post-sacarificación se pueden chequear para almidón residual usando el procedimiento de yodo de Staley.

- 25 Se pesan 20 gramos de muestra en un plato de pesaje grande. Se agregan 45 μL de solución de yodo al plato de pesaje y la solución de almidón se mezcla bien. El azul oscuro indica la presencia de almidón, un azul claro - verde indica almidón leve, verde claro indica un rastro de almidón y amarillo - rojo, ausencia de almidón. La solución de yodo se prepara al disolver 21,25 gramos de yodo y 40,0 gramos de yoduro de potasio en un litro de agua.

Perfil del oligosacárido

Los perfiles de licuefacción y sacarificación de carbohidratos se pueden medir por HPLC (por ejemplo, una columna AMINEX HPX-87C^{MR} (Bio-Rad) en forma de calcio -80°C) usando un detector de índice de refracción.

- 30 Cromatografía de permeación de gel

La distribución del peso molecular se puede determinar por cromatografía, por ejemplo, en una columna μL AQUAGEL-OH^{MR} con detección de masas por índice de refracción (Waters Model 2410). Un detector Viscotek modelo T60^{MR} se usa para mediciones continuas de viscosidad y de dispersión de luz.

Electroforesis capilar

- 35 Sistema de glucoproteínas Beckman Coulter P/ACE MDQ^{MR} - separación de oligosacáridos derivados de APTS en un capilar de sílice fundida - detección por fluorescencia inducida por láser.

Licuefacción primaria

- 40 El almidón de línea directamente del proceso GPC es bombeado a un tanque de alimentación de 60 litros donde el pH, DS (sólidos secos) y el nivel de calcio se pueden ajustar antes de la licuefacción. La amilasa se agrega a la suspensión. La suspensión de DS al 32% se bombea a razón de 0,7 litros/minuto mediante una bomba de desplazamiento positivo al chorro - una cámara de mezcla presurizado donde se calienta instantáneamente la suspensión de almidón a más de 100°C mediante inyección de vapor. El almidón licuado parcialmente gelatinizado se bombea a través de una red de tuberías (aún bajo presión) para obtener el tiempo de residencia deseado (5 minutos) a temperatura. La presión se libera en un tanque de expansión y se pueden tomar muestras. Las muestras se toman por duplicado.

Licuefacción secundaria

El almidón licuado se recoge en botellas de vidrio de un litro y se mantiene en un baño de agua a 95°C durante 90 minutos.

Sacarificación

5 El almidón licuado se enfría a 60°C, el pH se ajusta a 4,5 y las muestras se tratan con glucoamilasa. El progreso de la sacarificación se monitorea en el tiempo, mediante HPLC.

Sacarificación

10 Los jarabes licuados producidos con cada glucoamilasa se ajustaron hasta aproximadamente pH 2,5 con HCl 6 N en forma inmediata después de licuefacción secundaria de 90 minutos para inactivar cualquier glucoamilasa residual. Los jarabes se ajustan luego a pH 4,5, se colocan en un baño de agua de 60°C y se sacarifican con tres niveles de glucoamilasa. El grado de sacarificación se monitorea por HPLC en puntos de tiempo de 18 - 88 horas.

Los jarabes licuados se sacarifican con la dosificación estándar - 0,04% de una glucoamilasa de doble fuerza y dos dosis menores (50% y 25%) para monitorea cualquier diferencia en el progreso de sacarificación.

Progreso de la sacarificación - desarrollo del % de dextrosa vs tiempo - 0,04% de glucoamilasas.

Ejemplo 21: Ejemplo de un proceso de licuefacción de Almidón

15 Este ejemplo describe un ejemplo de un proceso de licuefacción de almidón de la invención que comprende el uso de enzimas de la invención. La conversión de almidón a glucosa puede ser catalizada por la acción de la secuencia de dos enzimas: amilasas (por ejemplo, alfa-amilasas), que incluye enzimas de la invención para licuar el almidón (por ejemplo, la hidrólisis de polímeros de glucosa de alto peso molecular a oligosacáridos que consisten de 2 a 20 unidades de glucosa, por lo general un equivalente de dextrosa de 10 a 12, por una glucoamilasa de la invención),
 20 seguido de sacarificación con una glucoamilasa. En un aspecto, el procesamiento está en la planta de molinera húmeda de maíz que produce una suspensión de almidón que tiene un pH de alrededor de 4,0 a 4,5. En un aspecto, el pH se eleva, por ejemplo, hasta 5,8, hasta 6,0 antes de licuefacción para acomodar una glucoamilasa con actividad y estabilidad de pH bajo. En un aspecto, las amilasas de la invención pueden licuar almidón a un pH de 4,5 a equivalentes de dextrosa que van de 12 a 18; en un aspecto, usando glucoamilasas a niveles de alrededor de 3 a
 25 6 gramos por tonelada de almidón. En este aspecto, el uso de las glucoamilasas permite que la licuefacción de almidón sea realizada a un pH de 4,5.

En un aspecto, la licuefacción de almidón se realiza a un pH de 4,5 durante 5 minutos a 105 °C hasta 90 minutos a 95°C usando las glucoamilasas. La cantidad de enzima se ajusta a fin de ajustar un DE objetivo de 12 a 15 después de la licuefacción. En un aspecto, el almidón licuado se sacarifica luego con una glucoamilasa, por ejemplo, una
 30 glucoamilasa de *Aspergillus*, durante aproximadamente 48 horas aproximadamente a pH 4,5 y 60 °C. Si el jarabe sacarificado no contiene al menos 95% de glucosa, el DE de la licuefacción objetivo se eleva y se repite la sacarificación hasta que la licuefacción eventualmente produjo un jarabe sacarificado que contenía más de 95% de glucosa. La proteína glucoamilasa requerida para producir un materia prima licuada adecuada para sacarificación se determina por PAGE.

35 **Ejemplo 22:** Identificación de péptidos generados de proteólisis de proteasa en prueba de fluido gástrico simulado (SGF)

Este ejemplo describe la identificación de pequeños péptidos resultantes de la proteólisis con pepsina de la enzima de ejemplo de la invención de la SEQ ID NO: 52 (codificada, por ejemplo, por la SEQ ID NO: 51). Este ejemplo también describe la evaluación de la actividad de los ejemplos de enzimas de la invención en pruebas de "Fluido
 40 gástrico simulado" (SGF) *in vitro*. Las pruebas de SGF mostraron que todas las enzimas fueron digeridas rápidamente por la pepsina proteasa gástrica. En un caso, un fragmento pequeño resistente a pepsina fue observado después de 60 minutos de tratamiento.

La Figura 21 ilustra un SDS PAGE que muestra los resultados de proteólisis (digestibilidad *in vitro*, la prueba de
 45 SGF) de la SEQ ID NO: 52 por pepsina a un pH de 1,3; la pepsina corta en el terminal C de los residuos de Lys (K) y Phe (F) (la flecha superior indica una banda de 36.5 K, la flecha del medio una banda de 5 K (véase discusión, a continuación) y la flecha inferior una banda de 3 K). Se usó estándar preteñido SEEBLUE PLUSZ^{MR} de Invitrogen y MARKER12^{MR} de Invitrogen. Todas las muestras se corrieron sobre geles de Tricina al 16%. Una banda común presente en todas las pruebas es pepsina.

La Figura 22 ilustra la caracterización de los péptidos generados en esta digestión de la SEQ ID No: 52, como se
 50 identifica usando análisis por LC MS/MS (cromatografía líquida / espectrometría de masa / espectrometría de masa); se identifica un sitio de glicosilación ligado a N.

La Figura 23A ilustra el esquema de aislamiento de péptido pequeño usado (para los péptidos generados por la proteólisis de la SEQ ID NO: 52 por pepsina). La Figura 23B se discute a continuación e ilustra un SDS PAGE de los resultados del esquema de aislamiento de péptido pequeño, donde el carril 1 es solamente pepsina, el carril 2 es amilasa no digerida de SEQ ID NO: 52, el carril 3 es la amilasa de la SEQ ID NO: 52 digerida por pepsina, el carril 4 es el corte superior de 30 kd como se ilustra en la Figura 23A, el carril 5 es la que fluye de 30 kd como se ilustra en la Figura 23A, el carril 6 es la muestra capturada por C18 RP (cromatografía en columna de fase reversa C18) como se ilustra en la Figura 23A.

La Figura 23C ilustra el perfil de LC/MS de la fracción eluida por C18 RP y se discute a continuación.

La Figura 23D ilustra la secuencia de los péptidos identificados por el análisis de LC MS/MS; la Figura 23E y la Figura 23F ilustran los sequins "Asn-Xaa-Ser/Thr" en la secuencia producida (resaltada en azul); las asparaginas que se predice son N-glicosiladas se resaltan en rojo.

Ensayos de digestibilidad *in vitro*: No se detectaron versiones de longitud completa de ninguna de las enzimas evaluadas por SDS-PAGE después de tratamiento con SGF, lo que indica que la pepsina había digerido cada una de las proteínas de longitud completa; después de 60 minutos de tratamiento con SGF, los fragmentos proteolíticos pequeños, alrededor de 6 kDa en tamaño, fueron observados en las digestiones de α -amilasa de la SEQ ID NO: 52 (véase discusión en el Ejemplo 22).

Prueba de actividad después del tratamiento con pepsina: Para determinar si existe actividad residual después del tratamiento con pepsina, se modificó la prueba con SGF y se realizaron pruebas de actividad con dextrina como sustrato para las glucoamilasas y almidón BODIPY como sustrato para las α -amilasas. No se observó actividad residual para cualquiera de las enzimas evaluadas, después de tratamiento con SGF durante 60 minutos. La pérdida de actividad fue causada ya sea por digestión con pepsina de las proteínas de prueba (por ejemplo la SEQ ID NO: 4, SEQ ID NO: 76 y SEQ ID No: 2), por la inactivación de la enzima bajo las condiciones de reacción de pH bajo, o ambos.

Para identificar la banda de péptido resistente a pepsina, que migran en un rango de 5 kDa en un SDS-PAGE, se enriqueció con péptidos pequeños mediante filtración de la muestra sobre una membrana filtrante de 30 kDa. El flujo fue purificado de manera adicional usando una columna de extracción de fase sólida C18. Los péptidos recuperados fueron luego analizados por secuenciamiento del terminal N con degradación de Edmond, LC-MS y espectroscopía de masa en tandem LC para determinar la secuencia peptídica. Las muestras analizadas por espectroscopía de masa en tandem LC fueron tratadas primero con proteasa de tripsina.

Purificación de péptidos: Para determinar la secuencia de la banda de péptidos de 5 kDa, se enriqueció el péptido pasando primero la reacción de la SEQ ID NO: 52 tratada de pepsina por una membrana de filtro de 30 kDa. El flujo fue luego purificado usando una columna de extracción de fase sólida C18. Los péptidos recuperados formaron una banda en el rango de 5 kDa en un SDS-PAGE. La Figura 23B ilustra un análisis SDS-PAGE del péptido de alfa-amilasa resistente a pepsina de la SEQ ID NO: 52. Este gel se tiñó con Coomassie. El carril 1 es la pepsina proteasa. El carril 2 es la amilasa de SEQ ID NO: 52. El carril 3 es la amilasa de la SEQ ID NO: 52 tratada con pepsina proteasa. El carril 4 es la misma que la muestra 3 que no fluyó a través de una membrana filtrante de 30 kDa. El carril 5 es el flujo a través de la muestra en el carril 3 una vez que atravesó la membrana filtrante de 30 kDa. El carril 6 es la muestra 5 después de la recuperación de una columna de extracción de fase sólida C18.

Para determinar las masas peptídicas en la muestra recuperada, los péptidos se analizaron mediante análisis LC-MS en una columna de fase reversa C18. El perfil m/z de esta muestra mostró que existen múltiples masas de varios valores m/z que están presentes en esta muestra, como se ilustran en la Figura 23C, un análisis LC-MS de los péptidos pequeños de amilasa de SEQ ID NO: 52. La fase sólida era un material RP C18. Un gradiente de 5% - 80% de ACN (15 - 90 minutos) fue usado como la fase móvil en el experimento.

Secuenciamiento del terminal N por degradación de Edmond: Para determinar la secuencia de los péptidos pequeños de la SEQ ID NO: 52, los péptidos purificados se enviaron a secuenciamiento del terminal N usando el método de degradación de Edmond. Los resultados para este análisis quedaron inconclusos ya que un número de diferentes aminoácidos fue liberado en cada etapa de la escisión. Este resultado fue consistente con el análisis LC/MS que mostró que existe más de una especie prominente de péptidos presentes en esta muestra.

Análisis por espectroscopia de masa en tándem LC: Para determinar la secuencia de péptidos presente en la fracción de péptido purificado de amilasa de la SEQ ID NO: 52, estos fueron tratados con tripsina proteasa y analizados por LC MS/MS. Este análisis resultó en la identificación de los siguientes siete péptidos; estas secuencias peptídicas fueron identificadas por las búsquedas SEQUEST^{MR} del análisis de LC MS/MS de los péptidos trípticos pequeños de amilasa de la SEQ ID NO: 52.

Péptido	Secuencia del péptido	Ubicación del residuo dentro de la SEQ ID NO: 52
1	AGQEQHYSGGSDPANR	349 - 364
2	VFSGDPAYTCPYQN	251 - 264
3	SGDPAYTCPYQN	253 - 264
4	SLLLLLSVFGQATHA	6 - 20
5	YENTGDGTSYHG	90 - 101
6	VYCGGSWQGIINHLD	56 - 70
7	GYSAGATLVETYTCT	448 - 462

5 También se determina que existe una cantidad de picos de péptidos en el espectro de LC-MS que su perfil de MS/MS no coincidió ni con la SEQ ID No: 52 ni con los otros péptidos en la base de datos. Una posible razón para esto podría ser la modificación compleja de glicosilación ligada a N de la proteína progenitora que tiene lugar en el organismo huésped, *Pichia pastoris*. Para probar esta hipótesis, los péptidos extraídos se trataron primero con PNGasa F para remover sus grupos glucosilo ligados a N. Los péptidos resultantes se trataron después con tripsina y se sometieron a análisis de LC MS/MS para determinar su composición peptídica. Este análisis determinó que existían al menos 6 especies de péptidos diferentes presentes en la muestra (véase resumen de la Tabla anterior).
10 La resolución de estos péptidos, después del tratamiento de desglicosilación, sugiere que estos péptidos eran parte de otros péptidos que fueron glicosilados en la muestra.

15 En conclusión, el secuenciamiento del terminal N y el análisis LC-MS mostró que la banda de péptidos de 5 KDa resultante de la digestión con pepsina proteasa de la amilasa de la SEQ ID NO: 52 está compuesta por muchas especies de péptidos diferentes. Esto se debe posiblemente a la digestión incompleta de la amilasa de la SEQ ID NO: 52 por la pepsina proteasa en parte debido al estado de glicosilación de estos péptidos. Por lo tanto, no existe una especie de péptido prominente en los espectros de MS - están presentes múltiples péptidos; múltiples especies de péptidos diferentes fueron secuenciadas en la muestra; el secuenciamiento químico del terminal N también brindó evidencia de la presencia de péptidos múltiples; existe una especie de péptidos que solamente aparece después de la desglicosilación con PNGasa F de la muestra, lo que sugiere un evento relacionado con glicosilación es el responsable por la aparición de este péptido.
20

Esta serie de ejemplos de protocolos se puede usar en cualquier polipéptido de la divulgación, por ejemplo, un ejemplo de secuencia de la divulgación, para determinar la secuencia, motivos, incluyendo motivos de glicosilación, sitios activos y similares.

Ejemplo 23: Amilasas activas a baja temperatura

25 Este ejemplo describe elaboración y caracterización de las enzimas de ejemplo de la divulgación SEQ ID NO: 56, SEQ ID NO: 52; SEQ ID NO: 62; SEQ ID NO: 70, que son activas a bajas temperaturas, incluyendo las que tienen habilidad para hidrolizar almidón a bajas temperaturas. Este ejemplo también describe el desarrollo de un cóctel de enzimas de la invención que puede hidrolizar > 95% del almidón en el maíz molido en azúcares fermentables en no más de 60 horas a alrededor de 30 a 40 °C y alrededor de pH 3,5 a pH 5,5 en presencia de levadura. En un aspecto,
30 la cantidad total de proteína enzimática requerida no es mayor a 50 gramos/toneladas de maíz, por ejemplo, en realizaciones alternativas 0,05% p/p, o 0,005% p/p, o cualquier parte en el rango entre 0,05% p/p a 0,005% p/p.

Las velocidades de reacción iniciales para hidrólisis de almidón fueron determinadas; y se estudiaron la influencia del pH en diferentes rangos, por ejemplos, en el rango de alrededor de pH 3,5 - 7,0 y las temperaturas en diferentes rangos, por ejemplo, en el rango de alrededor de 30 - 40°C, sobre la actividad de las enzimas de la divulgación;
35 también se determinó la especificidad del tipo de enlace de ejemplos de amilasas y/o glucoamilasas de la invención.

Métodos:

1. Determinación de la concentración de proteínas

Los sobrenadantes liofilizados de los cultivos de *P. pastoris* que expresan glucoamilasas y α -amilasas se suspendieron en agua a una concentración de ~10 mg de polvo / ml. Después de la determinación del contenido de proteína por el protocolo Bradford, 5 μ g de una muestra de proteínas y solución BSA estandarizada se corrieron sobre un gel de gradiente de 4 - 20% de Tris-Glicina. El gel se escaneó en un escáner de gel BioRad GS800^{MR} después de tinción con azul Coomassie. El software Bio-Rad QUANTITY ONE^{MR} se usó para la cuantificación de las bandas de BSA y glucoamilasa (o α -amilasas) y se calculó luego la concentración real de enzimas. La concentración de proteínas se ajustó en consecuencia y se confirmó por SDS PAGE adicional.

2. Determinación de velocidades de reacción iniciales.

A menos que se indique lo contrario, los ensayos se realizaron por triplicado a 37 °C y pH 5,0 en regulador (NaCH₃CO₂ 50 mM, CaCl₂ 10 mM; NaN₃ 10 mM y 0,01% de Triton X-100) que contiene 1% de almidón sin purificar, o 0,5% de dextrina o 1% de "almidón de maíz soluble". Los ensayos se realizaron a una escala de 0,5 ml para glucoamilasa y 0,25 ml para alfa-ami-asa (α -amilasa) en una incubadora de mesa Eppendorf con agitación constante (800 rpm).

Para las glucoamilasas, las reacciones comenzaron al agregar la enzima (concentración final de 0,25 μ g/ml) a la mezcla de reacción. A 0, 2,5, 5, 7,5, 10, 15, 20 y 30 min, se retiraron alícuotas de 50 μ l de alícuotas de las reacciones y se neutralizaron por adición a 100 μ l de regulador Tris 1 M, pH 7,5. Para las alfa-amilasas (α -amilasas), las reacciones comenzaron al agregar la enzima (concentración final de 0,4 μ g de proteína total/ml para la SEQ ID No: 56, 13434 y SEQ ID NO: 52; 2 μ g/ml para la SEQ ID NO: 62; 4 μ g/ml para la SEQ ID NO: 70) a la mezcla de reacción y se retiraron alícuotas de 10 μ l de las reacciones y se neutralizaron en reactivo de BCA a 2,5, 10, 15, 20, 25, 30 y 40 min. Para la determinación de los perfiles de temperatura, los ensayos se realizaron a 30, 34, 37 y 40 °C.

El efecto de pH sobre las actividades de glucoamilasa y amilasa se evaluaron a pH de 3,5, 4, 5, 6 y 7, usando el regulador de rango amplio de pH de Britton - Robinson (CH₃COOH 50 mM; H₃PO₄; H₃BO₃). Las reacciones paralelas a un pH de 4, 5 y 6 se realizaron también en presencia de regulador de acetato 50 mM para asegurar que el regulador usado no influyó en los resultados. Para la determinación del perfil de pH de dos alfa-amilasas (α -amilasas) que dependen de calcio de la SEQ ID NO: 56 y SEQ ID NO: 62, se usaron reguladores de ácido málico / acetato / MES en vez del de Britton-Robinson.

Preparación de "almidón de maíz soluble" para la reacción con alfa-amilasas. La dextrina (Sigma D2006) no se pudo usar como sustrato en las reacciones de alfa-amilasa BCA debido al medio con alto contenido de extremos reductores. Por lo tanto, se empleó almidón de maíz calentado (material de Syngenta) como sustrato. Se disolvió 2% de almidón de maíz en agua desionizada y se calentó mezclando en un baño de agua en ebullición durante 30 - 40 minutos, hasta que el almidón se disolvió y la solución tomó aspecto lechoso, pero traslúcido. La solución de almidón calentado se usó durante 2 días, después de lo cual se observaron algunos signos de degradación (aparición de grumos de almidón) y se descartó la solución.

3. Ensayo de glucosa oxidasa / peroxidasa (GO) para la cuantificación de glucosa liberada durante la hidrólisis de almidón.

Se usó un ensayo de glucosa oxidasa/ peroxidasa (GO) acoplada para determinar la cantidad de glucosa liberada por la glucoamilasa durante la hidrólisis de almidón. Las reacciones de GO comenzaron al agregar 10 μ l de la reacción de hidrólisis de almidón neutralizada a 90 μ l de PBS que contenía glucosa oxidasa (0,1 U/ml), peroxidasa (0,25 U/ml) y Amplex Red 0,05 mM, en placas Nunc negras de 96 pozos. Las placas se mantuvieron a temperatura ambiente en oscuridad durante 30 min antes de la lectura en un lector de placas fluorescentes con Ex/Em 545/590 nm. Una curva estándar con concentraciones de glucosa de 0 a 100 μ M se usó para evaluar la cantidad de glucosa producida en las reacciones de hidrólisis. Las velocidades iniciales de hidrólisis de almidón (nmoles de glucosa liberada de 1% de almidón granular / min / μ g de glucoamilasa) se determinaron al graficar la cantidad de glucosa liberada en el tiempo y calculando la pendiente de la línea de mejor ajuste a través de los puntos de datos.

4. Ensayo de BCA para determinar el incremento en la concentración de extremos reductores durante la hidrólisis de almidón.

Una alícuota de 10 μ l de reacción de hidrólisis de almidón de amilasa se neutralizó en 100 μ l de reactivo BCA (que consistía de 64 mg/mL de carbonato de sodio monohidrato, 24 mg/mL de bicarbonato de sodio, 1,95 mg/mL de BCA, 1,24 mg/mL de sulfato cúprico pentahidrato, 1,26 mg/mL de L-serina). El desarrollo de color se produjo durante la incubación de la reacción neutralizada a 80 °C durante 35 minutos y fue seguido por la determinación de la absorbancia a 560 nm. Las velocidades iniciales se calcularon durante un período de reacción de 40 min. Se construyó una curva estándar usando maltosa (0 - 54 μ M) fue construida para correlacionar A_{560 nm} con la concentración de los azúcares reductores generados (nmoles). La actividad específica se expresó como nmoles/min/ μ g de enzima.

5. Especificidad del tipo de enlace de glucoamilasas con maltosa e isomaltosa como sustratos

Las reacciones comenzaron al agregar la enzima (concentración final de 5 µg/ml para maltosa y 30 µg/ml para isomaltosa) a la mezcla de reacción. A los 2, 5, 10, 15, 20, 25, 30 y 40 min, se retiraron alícuotas de 5 µl de las reacciones y se neutralizaron mediante la adición a 10 µl del regulador Tris 1 M, pH 7,5. Se usaron 9 concentraciones de sustrato en los estudios, oscilando de 0 a 12 mM para maltosa y 2,5 a 120 mM para isomaltosa. Las reacciones se realizaron por triplicado a 37 °C y pH 5,0 en regulador (NaCH₃CO₂ 50 mM, CaCl₂ 10 mM), a escala de 50 µl en una incubadora de mesa Eppendorf con agitación constante (800 rpm). La producción de glucosa se midió al final de la reacción usando el ensayo de glucosa oxidasa / peroxidasa (GO).

Resultados

10 1. Caracterización de las glucoamilasas:

1.1 Velocidades de reacción iniciales: Las velocidades iniciales para hidrólisis de almidón granulado y soluble se presentan en la Tabla 2, ilustradas en la Figura 24. Como se puede observar a partir de la Tabla 2, las amilasas y/o glucoamilasas de la divulgación mostraron hasta 3 veces mejor actividad (SEQ ID NO: 48) contra el almidón granulado, con similar o ligeramente mejor actividad sobre el almidón soluble cuando se compara con la enzima de punto de referencia *A. niger*. La SEQ ID NO: 20 no pareció mostrar actividad contra el almidón granulado bajo las condiciones ensayadas (probablemente debido a la falta de un dominio de enlace a almidón). La Tabla 2 (Figura 24) resume los datos que comparan las velocidades iniciales de la hidrólisis de almidón de maíz granulado y soluble (dextrina) por las enzimas que sirven de ejemplo de la divulgación (incluyendo aquellas con actividad de glucoamilasa) y una enzima glucoamilasa de *A. niger* que sirve como referencia a 37 °C, pH 5,0; las velocidades iniciales se expresan como nmoles de glucosa/min/µg de proteína glucoamilasa liberada de 1% de almidón granulado o de 0,5% de dextrina. Cada número es el valor promedio de 6 - 10 puntos de datos.

1.2 Perfil de temperatura: se determinó el efecto de temperatura sobre la actividad de las enzimas que sirven de ejemplo de la divulgación, incluyendo glucoamilasas, sobre almidón granulado como sustrato; se usó una enzima como "punto de referencia", la glucoamilasa de *A. niger* comercialmente disponible. La liberación de glucosa se midió a 30, 32, 34, 36, 38 y 40 °C a un pH 5,0. Las actividades de las glucoamilasas aumentaron con la temperatura; fueron más activas a 40 °C pero retuvieron aproximadamente 50% de actividad pico a 30 °C.

La Figura 25 ilustra el efecto de la temperatura sobre la actividad del ejemplo de la glucoamilasa de la SEQ ID NO: 20 y el punto de referencia" de *A. niger* con almidón soluble (dextrina) como sustrato. La liberación de glucosa se midió a una temperatura indicada a pH 5,0.

1.3 Perfil de pH: La influencia de pH sobre la hidrólisis de almidón se evaluó tanto con almidón granular como con almidón soluble, respectivamente. Todas las glucoamilasas hidrolizaron ambos sustratos mejor a pH inferior, siendo la SEQ ID NO: 26 la de carácter más ácido. Para ambos sustratos, se midió la liberación de glucosa al pH indicado a 37 °C. Las velocidades iniciales se calcularon por encima de los 20 min y se convirtieron a un porcentaje de la velocidad máxima.

1.4 Especificidad de escisión del tipo de enlace: Los parámetros cinéticos para la hidrólisis de maltosa (enlace α-1,4) e isomaltosa (enlace α-1,6) se determinaron para 7 glucoamilasas seleccionadas y la glucoamilasa de *A. niger* como "punto de referencia". Los experimentos se realizaron con lisados de *P. pastoris* liofilizados y las proteínas no se purificaron; por lo tanto, únicamente se reportan aquí los datos independientes de concentración de proteínas. La Tabla 3, a continuación, resume los valores de K_M para maltosa e isomaltosa y la relación de k_{cat}/K_M para maltosa comparado con k_{cat}/K_M para isomaltosa. Estos parámetros determinados para glucoamilasa de *A. niger* están en concordancia con los datos publicados; por ejemplo, K_M para maltosa se reporta que es 1,2 - 2,1 mM; K_M para isomaltosa se reporta que es 19,8 - 42,0 y k_{cat}/K_M para maltosa sobre k_{cat}/K_M para isomaltosa se reporta que está entre 300 - 600, de acuerdo con Frandesen (1995) Biochemistry 34: 10162 - 10169; Sierks y Svensson (1996) Biochemistry 35: 1865 - 1871; Fagerstrom y Kalkkinen (1995) Biotechnol. Appl. Biochem. 21: 223 - 231.

La Tabla 3, resume los parámetros cinéticos para hidrólisis de maltosa e isomaltosa por siete ejemplos de glucoamilasas de esta divulgación y una glucoamilasas de *A. niger* como "punto de referencia" *A. niger*; cada número es el valor promedio de 5 experimentos diferentes:

Tabla 3

	Maltosa	Isomaltosa	k_{cat}/K_M (maltosa) / k_{cat}/K_M (isomaltosa)
	K_M (mM)	K_M (mM)	
SEQ ID NO: 28	0.61 ± 0.06	11.94 ± 4.99	750
SEQ ID NO: 74	1.87 ± 0.17	11.55 ± 2.92	481
SEQ ID NO: 20	2.62 ± 0.19	53.97 ± 23.17	897
SEQ ID NO: 14	2.67 ± 0.15	41.5 ± 5.05	116
SEQ ID NO: 26	0.98 ± 0.33	12.18 ± 0.64	456
SEQ ID NO: 48	2.26 ± 0.12	21.69 ± 3	565
SEQ ID NO: 18	1.01 ± 0.09	11.74 ± 7.74	415
Glucoamilasa de <i>A. niger</i>	0.93 ± 0.1	18.72 ± 3.95	249

5 Como puede observarse a partir de la Tabla 3, el ejemplo de la glucoamilasa de la SEQ ID NO: 20 fue la más fuertemente selectiva para maltosa y tuvo casi 900 veces mayor especificidad hacia los enlaces α -1,4 con relación a los enlaces α -1,6. La glucoamilasa menos selectiva fue la SEQ ID NO: 14 con especificidad ~100 veces mayor hacia los enlaces α -1,4 con relación a los enlaces α -1,6.

Ejemplo 24: Caracterización de alfa-amilasas y/o glucoamilasas de la divulgación

Este ejemplo describe la caracterización de las alfa-amilasas y/o glucoamilasa de la invención.

10 Velocidades iniciales de reacción: Las velocidades iniciales de hidrólisis de almidón granulado y soluble para ocho (8) ejemplos de α -amilasas se compararon con una α -amilasa como "punto de referencia" de *A. oryzae* y los resultados se resumen en la Tabla 4, a continuación. Debido al entorno relativamente alto observado en los ensayos de BCA con α -amilasa de *A. oryzae* de Sigma (A6211), se evaluó también una segunda preparación de la misma amilasa de Megazyme (E-ANAAM^{MR}). Los resultados obtenidos con ambas preparaciones fueron muy similares.

15 Como se puede observar a partir de la Tabla 4, todas las amilasas evaluadas mostraron una actividad significativamente mayor contra almidón soluble cuando se comparó con almidón granulado. Sin embargo, esta diferencia fue menos marcada para las amilasas y/o glucoamilasas de ejemplo de la divulgación que para la enzima de *A. oryzae* que sirve como "punto de referencia".

20 La Tabla 4 muestra los datos que comparan las velocidades iniciales de hidrólisis de almidón de maíz granulado y almidón de maíz soluble para ocho (8) α -amilasas y/o glucoamilasas de ejemplo de la divulgación y la α -amilasa que sirve como "punto de referencia" de *A. oryzae* a 37 °C y pH 5 (nota: *Las velocidades iniciales se expresan como nmoles de los extremos reductores liberados a partir de 1% de almidón/min/ μ g de proteína α -amilasa para las SEQ ID NO: 56, SEQ ID NO: 62, SEQ ID No: 2 y SEQ ID NO: 52; **Las velocidades iniciales para las SEQ ID NO: 70, SEQ ID NO: 66 y SEQ ID NO: 76 se expresan como nmoles ara los extremos ductores liberados a partir de 1% de almidón/min/ μ g de proteína total; cada número es un valor promedio de 5 puntos de datos:

25

Tabla 4

Enzima	Velocidad inicial* ± SD de almidón granulado	Velocidad inicial* ± SD de almidón soluble
SEQ ID NO: 56	15.7 ± 1.67	1607.9 ± 518.22
SEQ ID NO: 70**	2.1 ± 0.24	29.4 ± 3.81

Enzima	Velocidad inicial* \pm SD de almidón granulado	Velocidad inicial* \pm SD de almidón soluble
SEQ ID NO: 62	3.5 \pm 0.37	139.6 \pm 55.96
Enzima	Velocidad inicial* \pm SD de almidón granulado	Velocidad inicial* \pm SD de almidón soluble
SEQ ID NO: 66**	0.8 \pm 0.09	81.3 \pm 31.11
SEQ ID NO: 2	7 \pm 0.75	381.2 \pm 74.15
SEQ ID NO: 52	10.3 \pm 1.75	248.2 \pm 28.46
SEQ ID NO: 76**	2.6 \pm 0.41	160.3 \pm 42.62
Amilasa de <i>A. oryzae</i> (MegazymeE-ANAAM)	0.4 \pm 0.07	498.7 \pm 64.78

Perfil de temperatura: Se determinó el efecto de temperatura sobre hidrólisis de almidón por las α -amilasas caracterizadas. Se midió la actividad a un pH 5,0 durante 40 min de incubación a la temperatura indicada y las velocidades iniciales se calcularon y se graficaron contra el tiempo. Las actividades de las amilasas fueron afectadas por la temperatura en un grado diferente. Cinco de las amilasas y/o glucoamilasas de ejemplo de la divulgación, las SEQ ID NO: 56, SEQ ID NO: 70, SEQ ID No: 76 y SEQ ID NO: 66, fueron más activas a 40 °C y retuvieron aproximadamente 30% de actividad a 30 °C. Las actividades de las SEQ ID NO: 2, SEQ ID NO: 52 y SEQ ID NO: 62 fueron afectadas solamente marginalmente por los cambios en la temperatura en el rango investigado.

Perfil de pH: La influencia de pH sobre las amilasas y/o glucoamilasas de ejemplo de la divulgación sobre la hidrólisis de almidón se evaluó tanto con sustratos de almidón granulado como los de almidón soluble y se determinaron los resultados fueron. El incremento en los extremos reductores se midió en el pH indicado a 37 °C. Las velocidades iniciales se calcularon sobre 40 min y se convirtieron a un porcentaje de la velocidad máxima. La SEQ ID NO: 52 de ejemplo (originalmente de origen fúngico) tuvo el menor pH óptimo (aproximadamente pH 4). Otra enzima de ejemplo, SEQ ID NO: 2 (también originalmente de origen fúngico) también mostró preferencia por pH ácido, con un pH aparente óptimo aproximadamente a pH 4,5 - 5. La amilasa de ejemplo la SEQ ID NO: 66 (originalmente de origen arqueal) tuvo un pH aparente óptimo de aproximadamente pH 5,0, reteniendo aproximadamente 70% de la actividad pico a un pH 4,0. Las enzimas restantes de ejemplo de la divulgación tuvieron un pH óptimo entre 5,0 y 6,0 y fueron casi inactivas a un pH de 4,0 y 3,5.

Ejemplo 25: Expresión de las amilasas y/o glucoamilasas de la divulgación en células huésped

Este ejemplo describe la evaluación de expresión de enzimas de ejemplo de la invención en diferentes células huésped; en particular, la expresión de tres (3) amilasas de ejemplo en diferentes huéspedes de expresión: se estudiaron *Pichia pastoris*, *Hansenula polymorpha* y *Cochilobolus heterostrophus*.

H. polymorpha: se usaron dos vectores de expresión (Artes Biotechnology GmbH, Erkrath, Alemania); pFPMT-Mfa^{MR} con un promotor de formato deshidrogenasa y pTPS1-Mfa^{MR} con un promotor de trehalosa-6-fosfato sintasa. En ambos vectores, la señal de secreción usada fue la pre-prosekuensi del factor de apareamiento MFA1 de *Saccharomyces cerevisiae*.

Se introdujeron plásmidos con genes subclonados de amilasa introducidos en células competentes de *H. polymorpha* por electroporación. Se seleccionaron 96 colonias de cada transformación y se cultivaron durante aproximadamente 30 (3 pasadas) hasta 80 (8 pasadas) generaciones mediante diferentes etapas de sub-cultivo, bajo condiciones selectivas en medio líquido (YNB-Glucosa). Estas etapas permiten incrementar el número de copia y facilitan la integración del plásmido en el cromosoma. Después de estos 3 u 8 de pasadas, los transformantes se cultivaron bajo condiciones no selectivas (medio YPD) para evaluar la estabilización del plásmido (pérdida de cualquier plásmido no integrado). El cribado para clones positivos se realizó sobre placas de agar con almidón rojo. Los clones con una zona de aclaramiento más grande que el control (mismo gen expresado en *P. pastoris*) se seleccionaron como aciertos primarios. La expresión de amilasa por los clones seleccionados se confirmó evaluando los sobrenadantes del cultivo en sustrato de almidón BODIPY y se usó SDS PAGE para visualización de la proteína producida. Se preparó 1 L de cultivos para evaluación adicional.

5 *Cochilobolus heterostrophus*. Se subclonaron ADNc y ADNg de amilasa de *C. heterostrophus* en pCh-ubi (promotor de ubiquitina) y pCh-GPD (promotor de GPD) y dirigieron al locus PKS18 de la cepa C5 de *C. heterostrophus* por transformación de protoplasto mediada por PEG, usando un constructo linealizado de ADN. Uno de los constructos, una versión genómica bajo un promotor ubi (ubiquitina), fue también introducido en la cepa JMD3 a partir de la cual el gen que codifica amilasa había sido suprimido. En forma similar, como en el caso de *P. pastoris*, el cribado de los transformantes se realizó en placas de almidón rojo y se confirmó mediante ensayos de almidón BODIPY y SDS PAGE con sobrenadantes de cultivo. No se observó diferencia en la expresión entre el ADNc y las versiones genómicas. En forma similar, no se observó diferencia significativa con los dos promotores diferentes.

10 Resultados: Se construyeron cepas de: *H. polymorpha* y *C. heterostrophus* que expresan, para comparación, 3 genes que codifican amilasa y genes que codifican glucoamilasa de ejemplo. Además de 3 cepas originales de *P. pastoris*. Todas las cepas usadas para comparación y fermentación preliminar obtenidos con estas cepas se presentan en la Tabla 5, los resultados de expresión estimada (g/L) de las amilasas y/o glucoamilasas de ejemplo de la divulgación producidos en *Pichia*:

Tabla 5

Origen del gen de amilasa	<i>Pichia pastoris</i>
Fúngico	<u>SEQ ID NO: 2</u> 0.01*
bacteriano	<u>SEQ ID NO: 70</u> 0.07*
arqueal	<u>SEQ ID NO: 66</u> 0.13*
*rendimiento de la fermentación	

15 Después de la evaluación de la expresión, se compararon la pureza, actividad específica y desempeño de la proteína en ensayos de hidrólisis de almidón granulado. Estas comparaciones arrojaron las siguientes conclusiones:

- No se observaron diferencias bioquímicas significativas entre las enzimas expresadas en los diferentes huéspedes.
- La expresión en *Hansenula* no pareció generar una mejora en la expresión observada en *P. pastoris*. Además, las proteínas expresadas parecieron ser altamente glicosiladas.

Ejemplo 26: Fermentación de almidón sin purificar usando amilasas de esta divulgación

Este ejemplo describe métodos de ejemplo para fermentación de almidón sin purificar usando amilasas de esta divulgación.

25 Se identificaron inicialmente quince (15) enzimas de ejemplo que consistían de ocho (8) alfa-amilasas y siete (7) glucoamilasas. Los ejemplos de enzimas "principales" se determinaron sobre la base de la actividad a temperaturas de 30° a 40 °C un rango de pH de 3.5 - 7.0 e hidrólisis de 1% de almidón in purificar. No se analizó primero la enzima microbiana purificada para la cantidad de proteínas y proporciones de expresión; y luego se evaluaron en fermentaciones de almidón sin purificar. Se determinaron luego la carga de enzima total y las proporciones de alfa-amilasas con respecto a las glucoamilasas.

30 Nota: para el trabajo de fermentación del almidón sin purificar, se obtuvieron todas las enzimas por expresión en *Pichia pastoris*, excepto por la SEQ ID NO: 78 (codificada por ejemplo, la SEQ ID NO: 77), que se expresó en *Pseudomonas fluorescens* (por ejemplo, como se describe en JBC. 2002, 277(29):26501-26507).

ES 2 575 912 T3

Tabla 1.1. Expresión enzimática y pureza relativa de las alfa-amilasas y glucoamilasas “principales” de ejemplo

Enzima***	proteína/polvo (g/g)*	Relación de expresión**
SEQ ID NO: 56	0,12 ± 0,42	0,6
SEQ ID NO: 52	0,09 ± 0,003	0,6
Enzima***	proteína/polvo (g/g)*	Relación de expresión**
SEQ ID NO: 66	0,05 ± 0,005	0,1 (0,5****)
SEQ ID NO: 66	0,05	0,1 (0,5****)
SEQ ID NO: 66	0,059 ± 0,002	0,5****
SEQ ID NO: 66	0,085	0,5****
SEQ ID NO: 2	0,03 ± 0,005	0,3
SEQ ID NO: 2	0,03 ± 0,002	0,3
SEQ ID NO: 70	0,09 ± 0,003	0,05 (0,25****)
SEQ ID NO: 70	0,126 ± 0,002	0,25****
SEQ ID NO: 70	0,131	0,25****
SEQ ID NO: 62	0,2 ± 0,03	0,75
SEQ ID NO: 78	10 (mgs/ml)	0,25
SEQ ID NO: 76	0,1 ± 0,01	0,2
SEQ ID NO: 26	0,2 ± 0,026	0,8
SEQ ID NO: 28	0,15 ± 0,076	0,45
SEQ ID NO: 28	0,36 ± 0,042	0,45
SEQ ID NO: 18	0,24 ± 0,096	0,8
SEQ ID NO: 74	0,21 ± 0,016	0,8
SEQ ID NO: 48	0,28 ± 0,098	0,5
SEQ ID NO: 20	0,28 ± 0,066	0,7
SEQ ID NO: 14	0,21 ± 0,155	0,2
Control (vector vacío)	0,0063	
<p>* Contenido de proteína se determinó usando protocolo de Bradford. ** expresado como porcentaje de proteína total en lisado clarificado; se determinó por SDS PAGE usando cargas de proteína de 5 µg y un software Quantity One de Bio-Rad. *** donde una enzima aparece varias veces en este listado, se usaron diferentes lotes. **** valor después de la desglucosilación.</p>		

Métodos experimentales

Fermentación de almidón sin purificar

Todas las fermentaciones se realizaron usando maíz Yellow Dent II, molido en el molino Ultra Centrifugo por Glen Mill a 12000 rpm a través de un tamiz de 0,5 mm. La harina de maíz fue analizada en un equilibrador de humedad de Mettler HB-43TM para determinar el contenido de humedad y luego se midieron 10 g en peso seco de harina de maíz en un recipiente de 50 ml con barra de agitación. Se agregó luego agua seguido por la adición de tetraciclina (0,5 mg), H₂SO₄ (175 µl de solución 0,9 M), patrones de enzima purificado/no purificados y levadura para llevar la suspensión total a 33% de sólidos totales. Se colocaron los frascos en una placa de agitación a 30 °C durante 72 horas. Se tomaron muestras a las 50 o 72 horas de fermentación para análisis. Se usó glucoamilasa Sigma de *Aspergillus niger* (A-7095) como punto de referencia añadida a razón de 2,065 Unidades (aproximadamente 0,707 mg de proteína total, o 0,07% de carga de proteína total) /g en peso seco de harina y la carga total de enzima de ensayo (%) se basa en el peso total de la enzima expresada (g) / peso seco de harina de maíz en fermentación (g).

Para el análisis del contenido de etanol y los perfiles de azúcar, se tomaron 1,5 ml de suspensión de la fermentación de cada frasco a las 50 o 72 horas, se colocaron en un tubo de micrófuga de 1,8 ml y se centrifugó a 13000 rpm durante 5 minutos. Se vertió el sobrenadante en una columna spin x de 0,45 µm, se centrifugó a 7000 rpm durante 5 minutos y se tomaron alícuotas de 200 µl del centrifugado en un vial de HPLC y se analizó por HPLC. Se utilizó HPLC Waters equipado con un detector de índice de refracción. La columna utilizada fue la Bio-Rad AMINEX HPX-87H^{MR} (Cat. #2 125-0140). La fase móvil utilizada fue de H₂SO₄ 0,005 M a razón de 0,6 mL/minuto.

Preparación de los estándares

Los estándares usados para el análisis que se describen aquí contienen glucosa, maltosa, maltotriosa, maltodextrina, fructosa, glicerol, ácido láctico, ácido acético, ácido succínico, etanol y metanol. Se prepara un estándar compuesto y posteriormente se lo diluye para producir las diferentes concentraciones necesarias para establecer las curvas estándares para cada compuesto. Todos los reactivos usados para la preparación estándar deben ser almacenados en una cabina de desecación, refrigerada o almacenada de acuerdo con las recomendaciones del fabricante. Se deben emplear los medios más precisos de medición de los componentes. En un frasco volumétrico limpio y seco de 100 ml se añaden todos los reactivos secos.

- Se añaden aproximadamente 35 mL de agua y se forma un remolino para disolver los sólidos.

- Se coloca un tubo Falcon de 50 mL y se equilibra añadiendo 25 mL de agua y luego tara.

- Se añaden al agua agregar glicerol, etanol, metanol, ácidos lácticos y acéticos.

- Se agita suavemente con agitación tipo vórtice para mezclar y se añaden los contenidos al frasco volumétrico con solución de azúcar. Se enjuaga el tubo Falcon de 50 ml 3 veces con aproximadamente 5 mL de agua cada vez y se añade el enjuague al frasco volumétrico.

- Se lleva el volumen del frasco hasta 100 mL. Esta solución representa la mezcla estándar de etanol al 20% La solución es esterilizada usando un filtro de 0,22 o 0,45 µm Millipore STERIFLIP^{MR}.

- Se utilizan diluciones de este estándar para crear los estándares mezclados de 5, 10 y 15% de etanol

- Un estándar mezclado al 5% consiste de 1 parte de 20% de estándar mezclado y 3 partes de agua

- Un estándar mezclado al 10% es una dilución simple 1:1 del estándar mezclado al 20%.

- Una mezcla de 3 partes de 20% de estándar mezclado y 1 parte de agua crearan el estándar mezclado al 15%

- Se debe correr un estándar certificado de etanol al 10,3% v/v (Sigma Cat.# E2385) cada vez que se prepara un nuevo estándar o mensualmente para verificar la precisión del conjunto estándar mezclado.

Resultados

Se analizaron varias alfa-amilasas y glucoamilasas en combinaciones en fermentaciones de almidón sin purificar a diferentes cargas totales de enzima de 0,001, 0,01 y 0,1% (p/p). Las alfa-amilasas analizadas fueron las SEQ ID NO: 56, SEQ ID NO: 70, SEQ ID NO: 52, SEQ ID NO: 2, SEQ ID NO: 62, SEQ ID No: 66 y SEQ ID NO: 76. Las glucoamilasas analizadas fueron las SEQ ID NO: 48, SEQ ID NO: 26, SEQ ID NO: 18, SEQ ID NO: 74, SEQ ID NO: 28, SEQ ID NO: 14 y SEQ ID NO: 20. Todas las combinaciones tenían un número de identificación para ser

ES 2 575 912 T3

5 utilizadas en el proceso experimental. Se calcularon todas las cargas de enzimas con base en la información de la velocidad de expresión ministrada aquí. La relación de alfa-amilasa con respecto a la glucoamilasa se determinó con base en los datos anteriores. El rendimiento de etanol se midió después de 50 o 72 horas de fermentación por análisis de HPLC. Los resultados indicaron que todas las combinaciones enzimáticas produjeron diferentes niveles de etanol y algunas combinaciones superaron al punto de referencia de GA Sigma (Tabla 1.2).

Tabla 1.2. Rendimiento de etanol en fermentaciones de almidón sin purificar de ejemplos de alfa-amilasas y glucoamilasas "principales" después de 50 horas de fermentación^a.

	Alfa-amilasa (AA)	glucoamilasa (GA)	relación AA:GA	Rendimiento de etanol (% v/v)		
				0,001% de carga total de enzima	0,01% de carga total de enzima	0,1% de carga total de enzima
Combo 1	SEQ ID NO: 56	SEQ ID NO: 26	2:1	6,12	11,60	15,56
Combo 2	SEQ ID NO: 56	SEQ ID NO: 18	1:2	6,97	12,53	15,69
Combo 3	SEQ ID NO: 56	SEQ ID NO: 28	1:2	5,96	11,06	16,42
Combo 4	SEQ ID NO: 56	SEQ ID NO: 74	1:2	6,83	11,49	14,29
Combo 5	SEQ ID NO: 56	SEQ ID NO: 20	1:2	4,28	5,61	7,55
Combo 6	SEQ ID NO: 56	SEQ ID NO: 48	1:2	6,40	12,58	15,86
Combo 7	SEQ ID NO: 66	SEQ ID NO: 26	1:2	7,39	15,38	N/A
Combo 8	SEQ ID NO: 66	SEQ ID NO: 18	1:2	5,84	12,81	N/A
Combo 9	SEQ ID NO: 66	SEQ ID NO: 28	1:2	6,12	12,62	N/A
Combo 10	SEQ ID NO: 66	SEQ ID NO: 74	1:2	7,02	12,29	N/A
Combo 11	SEQ ID NO: 66	SEQ ID NO: 20	1:2	4,48	5,72	N/A
Combo 12	SEQ ID NO: 66	SEQ ID NO: 48	1:2	6,54	14,16	N/A
Combo 13	SEQ ID NO: 70	SEQ ID NO: 26	1:2	9,08	15,04	16,76
Combo 14	SEQ ID NO: 70	SEQ ID NO: 18	1:2	7,86	13,00	14,71
Combo 15	SEQ ID NO: 70	SEQ ID NO: 74	1:2	8,18	13,27	16,40
Combo 16	SEQ ID NO: 70	SEQ ID NO: 48	1:2	8,36	13,80	15,90
Combo 17	SEQ ID NO: 52	SEQ ID NO: 26	1:2	7,92	12,67	15,15
Combo 18	SEQ ID NO: 52	SEQ ID NO: 18	2:1	10,66	16,10	14,92
Combo 19	SEQ ID NO: 52	SEQ ID NO: 48	1:2	10,88	15,97	14,87
Combo 20	SEQ ID NO: 62	SEQ ID NO: 26	1:2	7,30	12,67	15,68

ES 2 575 912 T3

	Alfa-amilasa (AA)	glucoamilasa (GA)	relación AA:GA	Rendimiento de etanol (% v/v)		
				0,001% de carga total de enzima	0,01% de carga total de enzima	0,1% de carga total de enzima
Combo 21	SEQ ID NO: 62	SEQ ID NO: 48	1:2	7,24	11,13	14,84
Combo 22	SEQ ID NO: 2	SEQ ID NO: 26	2:1	9,20	14,56	15,36
Combo 23	SEQ ID NO: 2	SEQ ID NO: 48	1:2	9,81	15,42	15,61
Combo 24	SEQ ID NO: 56	SEQ ID NO: 14	1:2	5,46	7,86	13,67
Combo 25	SEQ ID NO: 70	SEQ ID NO: 14	2:1	6,99	10,85	15,72
Combo 26	SEQ ID NO: 78	SEQ ID NO: 48	1:2	N/A	N/A	16,13 ^a
Combo 27	SEQ ID NO: 78	SEQ ID NO: 26	1:2	N/A	N/A	16,88 ^a
Combo 28	SEQ ID NO: 76	SEQ ID NO: 48	1:2	N/A	N/A	17,15 ^a
Control de GA de Sigma				N/A	N/A	16,82 ^b

^a Rendimiento de etanol después de 72 horas de fermentación.

^b Carga total de enzimas se estimó que era aproximadamente de 0,07%.

Ejemplo 27: Expresión de enzimas en plantas; fermentación de almidón sin purificar

5 Este ejemplo describe la expresión en una planta de nueve ejemplos de enzimas de esta divulgación (enzimas que no tiene optimización de codón) y fermentación de almidón sin purificar utilizando material vegetal.

10 Se escogieron cinco (5) alfa-amilasas, las SEQ ID NO: 52, SEQ ID NO: 4, SEQ ID NO: 70, SEQ ID NO: 66 y SEQ ID NO: 76 y 4 glucoamilasas, las SEQ ID NO: 48, SEQ ID NO: 18, SEQ ID NO: 26 y SEQ ID NO: 28 para la expresión con base en un rendimiento alto de etanol de las enzimas microbianas en fermentación de almidón sin purificar. Esta expresión en maíz se llevó a cabo con las secuencias génicas sin optimización del codón. Los objetivos de este estudio fueron evaluar los niveles potenciales de expresión de estas enzimas y evaluar si la expresión de estas enzimas tendría efectos nocivos sobre las características agronómicas del maíz a través de sus etapas de desarrollo.

Métodos experimentales

Construcción de vector

15 La Tabla 2.1 resume las abreviatura utilizadas para los elementos de ADN en los mapas que se muestran continuación.

Tabla 2.1. Abreviaturas usadas en los mapas del constructo

Nombre	Función	Descripción
cAMY(SEQ ID NO: 66)-02	CDS	Alfa-amilasa, SEQ ID NO: 66, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
cAmy(SEQ ID NO: 18)-02	CDS	Glucoamilasa, SEQ ID NO: 18, menos la secuencia guía nativa, unida a

ES 2 575 912 T3

		una secuencia señal de Gamma Zeina dirigida al apoplasto
Nombre	Función	Descripción
cAmy(SEQ ID NO: 52)-02	CDS	Alfa-amilasa, SEQ ID NO: 52, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
cAmy(SEQ ID NO: 28)-02	CDS	Glucamilasa, SEQ ID NO: 28, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
Nombre	Función	Descripción
cAmy(SEQ ID NO: 26)-02	CDS	Glucamilasa, SEQ ID NO: 26, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
cAmy(SEQ ID NO: 48)-02	CDS	Glucamilasa, SEQ ID NO: 48, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
Nombre	Función	Descripción
cAmy(SEQ ID NO: 76)-02	CDS	Alfa-amilasa, SEQ ID NO: 76, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto
cAmy(SEQ ID NO: 70)-02	CDS	Alfa-amilasa, SEQ ID NO: 70, menos la secuencia guía nativa, unida a una secuencia señal de Gamma Zeina dirigida al apoplasto

El ejemplo de vector denominado como "15745", a continuación, es un vector binario que contiene una alfa-amilasa SEQ ID NO: 70 de ejemplo menos su secuencia guía nativa. La secuencia guía nativa fue reemplazada por la secuencia señal de Gamma Zeina para dirigirla al apoplasto. La fusión de Gamma Zeina-SEQ ID NO: 70 se expresa utilizando el promotor de Gamma Zeina. El vector binario también contiene un casete Ubi-PMI-Nos para la selección de manosa.

Un fragmento de cAmy(SEQ ID NO: 70)-01 que contiene aminoácidos 86 - 525 (más el codón de terminación), se amplificó por PCR a partir del plásmido 15649. La amplificación por PCR introdujo un sitio Afel en el extremo 5' y un sitio BgIII en el extremo 3' del fragmento. La adición del sitio Afel agregó un aminoácido de alanina en frente del fragmento de cAmy(SEQ ID NO: 70) (SY1709: 92 - 93). Este amplicón de PCR se clonó en pCR2,1-TOPO y se secuenció para verificar la presencia de los nuevos sitios de enzimas de restricción (SY17092111-112,122-124). Una vez que se verificó, se digirió cAmy(SEQ ID NO: 70) fuera de la estructura de TOPO a través de Afel/BgIII. En forma similar, el constructo 15460ZeinAmyVN fue digerido con Afel/BgIII. Esta estructura y cAmy(SEQ ID NO: 70) se ligaron en forma conjunta. Cuando se ligaron, cAmy(SEQ ID NO: 70) permaneció en la estructura con el péptido señal xGZein27ss-01, solamente con la adición de un residuo de alanina en la unión de clonación. Los transformantes se cribaron para la presencia del gen de amilasa usando cebadores ZeinAmy1199F y prGTL-03R. Los clones positivos produjeron un amplicón de aproximadamente 1,6 kB (SY1709:130). Cinco clones fueron seleccionados y fueron digeridos con SanDI y RsrII para remover prGZein-01:cAmy(SEQ ID No: 70):t35s-08. El constructo 15468 fue alineado con RsrII. La estructura binaria y el casete digerido de uno de los clones positivos (SanDI/RsrII) fueron ligados en forma conjunta en el sitio de restricción de RsrII. Después de la transformación en células Top10, se cribaron los transformantes por la presencia del casete usando cebadores ZeinAmy1199F y Mubi-5(SY1709:146,150). Tres transformantes positivos fueron seleccionados y chequeados a través de una digestión diagnóstica con BamHI. Estos clones fueron secuenciados (SY1709:162; ambas uniones de clonación así como también toda la secuencia codificadora cAmy(SEQ ID NO: 70)), se confirmó que todos los tres fueron correctos en la secuencia.

El vector de ejemplo designado como "15750", como se ilustra en la Figura 28, a continuación, es un vector binario que contiene una glucamilasa, la SEQ ID NO: 18 de ejemplo, menos su secuencia guía nativa. La secuencia guía nativa se reemplazó con la secuencia señal de Gamma Zeina para dirigirse al apoplasto. La fusión de Gamma Zeina-SEG ID NO: 18 es expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete de Ubi-PMI-Nos para selección de manosa.

Un fragmento de cAmy(SEQ ID NO: 18)-01, que contiene los aminoácidos 88 - 712 (más el codón de terminación), fue amplificado por PCR a partir del plásmido 15652. La amplificación por PCR introdujo un sitio Afel en el extremo 5' y un sitio BgIII en el extremo 3' del fragmento. La adición del sitio Afel resultó en el cambio del aminoácido 88 de

- una serina por una alanina (SY1709: 113-114,116). Este amplicón de PCR fue clonado en pCR2.1-TOPO y secuenciado para verificar la presencia de los nuevos sitios de la enzima de restricción (SY1709:116,131). Una vez verificado, el fragmento de cAmy(SEQ ID NO: 18) fue digerido fuera de la estructura TOPO a través de Afel/BgIII. En forma similar, el constructo 154602einAmyVN fue digerido con Afel/BgIII. Esta estructura y el fragmento de cAmy(SEQ ID NO: 18) fueron ligados en forma conjunta. Cuando se ligaron, cAmy(SEQ ID NO: 18) permaneció en la estructura con el péptido señal xGZein27ss-O1, con la adición de un residuo de alanina en la unión de clonación. Seis clones fueron seleccionados y fueron digeridos con SanDI y RsrII para remover prGZein-01:cAmy(SEQ ID NO: 18):t35s-08. El constructo 15468 fue alineado con RsrII. La estructura binaria y el casete digerido de uno de los clones positivos (SanDI/RsrII) fueron ligados en forma conjunta en el sitio de restricción de RsrII. Después de la transformación en células Top10, los transformantes se cribaron por la presencia del casete usando cebadores ZeinAmy1199F y Mubi-5 (SY17092181). Adicionalmente, todos los 6 transformantes fueron seleccionados y chequeados a través de una digestión diagnóstica con SacI. Tres de los clones positivos fueron secuenciados (SY17092189-190) (ambas uniones de clonación así como también toda la secuencia codificadora cAmy(SEQ ID NO: 18)); se confirmó que todos los tres eran correctos en secuencia.
- 15 El vector de ejemplo designado como "15751", como se ilustra en la Figura 29, a continuación, es un vector binario que contiene una glucoamilasa, la SEQ ID NO: 48 de ejemplo, menos su secuencia guía nativa. La secuencia guía nativa se reemplazó con la secuencia señal de Gamma Zeina para dirigirse al apoplasto. La fusión de Gamma Zeina-SEQ ID NO: 48 es expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete de Ubi-PMI-Nos para selección de manosa.
- 20 Usando 15651 como plantilla, se usaron los cebadores de mutagenesis en una reacción de mutagénesis de sitios múltiples (QUICKCHANGE^{MR}, Stratagene) para remover un sitio Afel interno. introducir un sitio Afel en el extremo 5' (la mutación estaba en la secuencia señal, sin cambio de péptido) e introducir un sitio BgIII en el extremo 3' externo para la clonación. Se llevo a cabo el análisis de la enzima de restricción y se seleccionaron los clones que fueron positivos por tener el sitio Afel interno removido y el sitio de BgIII introducido. Un clon positivo fue enviado para análisis de secuencia que fue confirmada. Este clon de ADN fue luego usado como plantilla para amplificación por PCR para introducir un sitio Afel en el extremo 5' del dominio de glucoamilasa. La mutación estaba en la secuencia señal nativa y no cambió la secuencia del péptido glucoamilasa. El producto de PCR fue digerido de manera secuencial con Afel y luego con BgIII. El producto digerido fue purificado por gel y ligado a la estructura de 15460AmyZeinVN, creando cAmy(SEQ ID NO: 48) con una secuencia señal de gamma zeina que se dirige al apoplasto. Se verificó la secuencia de este vector intermedio y luego se digirió con SanDI y RsrII en forma secuencial. Este fragmento fue purificado por gel y luego ligado en el corte binario 15468 con RsrII y tratado por CIP. Los clones positivos se cribaron por tamaño por PCR y luego uno de los clones positivos se envió para análisis de secuencia del fragmento entero ligado en el binario. Se confirmó la secuencia.
- 25
- 30
- 35 El vector de ejemplo designado como "15761", como se ilustra en la Figura 30, a continuación, es un vector binario que contiene una alfa amilasa, la SEQ ID NO: 66 de ejemplo, menos su secuencia guía nativa. La secuencia guía nativa fue reemplazada con la secuencia señal de Gamma Zeina para dirigirse al apoplasto. La fusión de Gamma Zeina-SEQ ID NO: 66 fue expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete de Ubi-PMI-Nos para selección de manosa.
- 40 Un fragmento de cAmy(SEQ ID NO: 66)-01, que contiene los aminoácidos 86 - 521 fue mutagenizado usando mutagénesis dirigida a sitio de acuerdo con el protocolo Quick Change de Stratagene. La mutagenesis requirió 3 cebadores individuales (todos amplificándose en la dirección hacia adelante) e introdujeron tres mutaciones: un sitio Afel en el extremo 5', un sitio BgIII en el extremo 3' (externo al gen) y la remoción de un sitio Afel interno de (SY1709:174). El constructo mutagenizado fue transformado en células Top10 por cribado para determinar si alguna/todas las mutaciones deseadas estaban presentes. Quince transformantes fueron seleccionados para cribado y fueron digeridos en forma separada con Afel y BgIII (SY17092185-186). Se determinó a partir de las digestiones de restricción que existían 6 clones posibles con todas las tres mutaciones deseadas; tres de estos clones se enviaron a secuenciamiento (SY1709:190; SY1777:11). Para todos los tres clones, un sitio de Afel se había introducido en el extremo 5' y el sitio interno había sido removido. Además, el sitio BgIII en el extremo 3' había sido introducido. Después de la confirmación, la alfa amilasa fue digerida fuera del constructo mutagenizado con Afel/BgIII y ligada a la estructura de 154602einAmyVN que había sido digerida con Afel y BgIII. Cuando se ligaron, cAmy(SEQ ID NO: 66) permaneció en la estructura con el péptido señal xGZein27ss sin cambios de aminoácidos en la unión de clonación. Los transformantes se cribaron para la presencia del gen de amilasa usando cebadores ZeinAmy1199F y prGTL-03R. Los clones positivos produjeron un amplicón de aproximadamente 1,6 kB (SY1777:24). Tres clones fueron seleccionados y digeridos con SanDI/RsrII para remover prGZein-01:cAmy(SEQ ID NO: 66):t35s-08. El constructo 15468 fue alineado con RsrII. La estructura binaria y el casete digerido de uno de los clones positivos (SanDI/RsrII) fueron ligados en forma conjunta en el sitio de restricción de RsrII. Después de la transformación en células Top10, los transformantes se cribaron para la presencia del casete usando cebadores ZeinAmy1199F y prGTL-03R (SY1777:47-48). Seis transformantes positivos fueron seleccionados y chequeados a través de una digestión diagnóstica con HindIII. Tres de estos clones fueron secuenciados (SY1777:56, 62) tanto en las uniones de clonación como también la secuencia codificadora completa; se confirmó que todos los tres eran correctos en secuencia.
- 45
- 50
- 55
- 60

El vector de ejemplo designado como "15756", como se ilustra en la Figura 31, a continuación, es un vector binario que contiene una glucoamilasa, la SEQ ID NO: 26 de ejemplo, menos su secuencia guía nativa. La secuencia guía nativa fue reemplazada con la secuencia señal Gamma Zeina para dirigirse al apoplasto. La fusión de Gamma Zeina-SEQ ID NO: 26 es expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete Ubi-PMI-Nos para selección de manosa.

Usando el vector de ejemplo 15653 como plantilla, los cebadores de mutagenesis se usaron en una reacción de mutagenesis de sitios múltiples (QUICKCHANGE^{MR}, Stratagene) para remover un sitio Afel interno, introducir un sitio Afel en el extremo 5' (la mutación estaba en la secuencia señal, sin cambio de péptido) e introducir un sitio SacI en el extremo 3' externo para clonación. La reacción se transformó en las células competentes Top10. Las colonias fueron seleccionadas y cribadas por análisis de enzimas de restricción. Los clones que fueron positivos por tener un sitio interno Afel removido, el sitio Afel en el extremo 5' insertado y el sitio SacI introducido, fueron seleccionados. Se envió un clon positivo para análisis de secuencia, que fue confirmada. Este clon de ADN fue luego digerido con SacI y Afel. La digestión fue luego purificada por gel y ligada a la estructura de 15460AmyZeinVN creando cAmy(SEQ ID NO: 26) con una secuencia señal de gamma zeina que se dirige al apoplasto. Este vector intermedio fue verificado en secuencia, luego digerido con SanDI y RsrII. Este fragmento fue purificado en gel y ligado en el corte del vector binario 15468 con RsrII y tratado con CIP. Los clones positivos fueron cribados por tamaño por PCR y luego uno de los clones positivos fue enviado a análisis de secuencia de todo el fragmento ligado en el binario. Se confirmó la secuencia.

El vector de ejemplo designado como "15742", como se ilustra en la Figura 32, a continuación, es un vector binario que aloja (que comprende) los 616 aminoácidos del terminal C de una glucoamilasa, la SEQ ID NO: 28 de ejemplo (cAmy(SEQ ID No: 28)-01) fusionada en la estructura con la secuencia señal de Gamma Zeina (xGZein27ss-01), que dirige la glucoamilasa al apoplasto para crear (cAmy(SEQ ID NO: 28)-02). La expresión es dirigida por la región 5' de Gamma Zeina A (prGZein-01), que es un promotor específico de la semilla.

El vector 15654 que aloja cAmy(SEQ ID NO: 28)-01 sirvió como plantilla de ADN por PCR modificado por la adición de un sitio NcoI (extremo 5' prima) y un sitio BglII (extremo 3' prima). El producto de PCR fue TOPO-clonado y el Clon #1 fue secuenciado para validación (SY1533:181). El nuevo componente, alias "cAmy(SEQ ID NO: 28)" fue digerido con NcoI/BglII y purificado en gel. El fragmento fue ligado en a la estructura al sitio NcoI/BglII de un vector de clonación, alias "154602einAmyVN" de esta manera generando un casete que aloja prGZein-01:xGZein27ss-01:cAmy(SEQ ID NO: 28):iPEPC9-01:t35s-08 (SY1533:183). El casete génico fue digerido de manera secuencial con SanDI/RsrII, purificado en gel y ligado en el sitio de RsrII del vector binario 15468. El ligamiento exitoso fue confirmado por PCR y la secuencia de ADN (SY1533:185-187).

El vector de ejemplo designado como "15749", como se ilustra en la Figura 33, a continuación, es un vector binario que contiene una alfa amilasa, la SEQ ID NO: 52 de ejemplo, menos la secuencia guía nativa. La secuencia guía nativa fue reemplazada con la secuencia señal de Gamma Zeina para dirigirse al apoplasto (cAmy(SEQ ID NO: 52)-02). La fusión de Gamma zeina-SEQ ID NO: 52 es expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete de Ubi-PMI-Nos para selección de manosa.

Un fragmento de cAmy(SEQ ID NO: 52)-01, que contiene los aminoácidos 88 - 674 (más el codón de terminación), fue amplificado por PCR a partir del plásmido 15648. La amplificación por PCR introdujo un sitio Afel en el extremo 5' y un sitio SacI en el extremo 3' del fragmento. La adición del sitio Afel provocó que el aminoácido 88 cambiara de serina a una alanina (SY1709:95-96). El amplicón de PCR fue clonado en pCR2.1-TOPO y secuenciado para verificar la presencia de los nuevos sitios de la enzima de restricción (SY1709:103,111-112). Una vez verificado, se digirió cAmy(SEQ ID NO: 52) fuera de la estructura TOPO a través de Afel/SacI. En forma similar se digirió el constructo 15460ZeinAmyVN con Afel/SacI. Esta estructura y cAmy(SEQ ID NO: 52) fueron ligados en forma conjunta. Cuando se ligaron cAmy(SEQ ID NO: 52) permaneció en el marco con el péptido señal xGZein27ss-01, con únicamente el primer residuo de aminoácido cambiado de una Leucina por una Alanina. Los transformantes se cribaron para la presencia del gen amilasa usando cebadores ZeinAmy1199F y prGTL-03R. Los clones positivos produjeron un amplicón de aproximadamente 2,1kB (SY1709:130). Cinco clones fueron seleccionados y fueron digeridos con SanDI y RsrII para remover prGZein-01:cAmy(SEQ ID NO: 52):t35s08. El constructo 15468 fue alineado con RsrII. La estructura binaria y casete digerido de uno de los clones positivos (SanDI/RsrII) fueron ligados en forma conjunta en el sitio de restricción de RsrII. Después de la transformación en las células Top10, los transformantes se cribaron para la presencia del casete usando cebadores ZeinAmy1199F y Mubi-5 (SY1709: 146, 150). Tres transformantes positivos fueron seleccionados y chequeados por medio de digestión diagnóstica con BamHI. Estos clones fueron secuenciados (SY1709:162; ambas uniones de clonación así como la secuencia codificadora cAmy(SEQ ID NO: 52) entera) y se determinó que había una mutación puntual en la secuencia de prGZein. Esta mutación es un cambio de un par de bases individuales, T por C en el nucleótido 340. El promotor en la estructura de 15460ZeinAmy original fue secuenciado nuevamente para chequear la mutación y no estaba presente en la fuente estructural. Sin embargo, estaba presente cuando los clones positivos fueron secuenciados una segunda vez (SY1709:182,189). Esta mutación pudo haber ocurrido durante la replicación bacteriana.

ES 2 575 912 T3

El vector de ejemplo designado como "15718", como se ilustra en la Figura 34, a continuación, es un vector binario que contiene una alfa amilasa, la SEQ ID NO: 76 de ejemplo, menos su secuencia guía nativa. La secuencia guía nativa fue reemplazada con la secuencia señal de Gamma Zeina para dirigirse al apoplasto. La fusión de Gamma Zeina-SEQ ID NO: 76 (cAmy(SEQ ID NO: 76)-02) es expresada usando el promotor de Gamma Zeina. El vector binario contiene también un casete de Ubi-PMI-Nos para selección de manosa.

Un fragmento de cAmy(SEQ ID NO: 76)-01, que contiene los aminoácidos 87-508 (más el codón de terminación), fue amplificado por PCR a partir del plásmido 15650. La amplificación por PCR introdujo un sitio de Afel en el extremo 5' y un sitio de BgIII en el extremo 3' del fragmento. La adición del sitio Afel cambió un aminoácido de glicina por una alanina. El producto de PCR fue digerido con Afel y BgIII y clonado en el corte de la estructura de 15460AmyZeinVN con Afel y BgIII, creando cAmy(SEQ ID NO: 76) fusionada con la secuencia señal de gamma zeina. Este vector intermedio fue luego digerido con SmaI y RsrII en forma secuencial. Este fragmento fue purificado en gel y luego ligado en el corte de vector binario 15468 con RsrII y tratado por CIP. Los clones positivos fueron cribados por tamaño por PCR y luego uno de los clones positivos fue enviado para análisis de secuencia del inserto entero. Se confirmó la secuencia.

15 Análisis de actividad enzimática

Veinte semillas de los eventos seleccionados fueron seleccionadas al azar y reunidas. Las semillas fueron luego trituradas y evaluadas usando el ensayo Megazyme CERALPHA HR^{MR} para alfa-amilasas o el ensayo de glucoamilasa de Megazyme para glucoamilasas.

Cálculos de actividad

20 Una unidad CERALPHA^{MR} de actividad de α -amilasa se define como la cantidad de enzima, en presencia de exceso de α -glucosidasa termoestable, requerida para liberar un micromol de p-nitrofenol en un minuto bajo las condiciones de ensayo definidas. Nuestra condición de ensayo es un pH de 5,5 (NaOAc 100 mM, regulador de pH 5,5 con 1mg/ml de BSA), 60 °C. Nótese que la actividad calculada es una unidad de CERALPHA^{MR}.

25 De acuerdo con Megazyme, PNP 1 mM en 1% de fosfato trisódico produce una absorbancia de 18,1 a 400 nm. Por lo tanto, 1 M de PNP en fosfato trisódico produce una absorbancia de 18.100 a 400 nm.

$$\Delta A_{400} / 18,100 = [\text{PNP}] \text{ (mol/l o M) en la placa final.}$$

$$[\text{PNP}] \text{ en placa (mol/l} \times 8 \times 100 \times 10^{-6} \times 1 \times 1 \times 10^6 \text{ (}\mu\text{mol/mol) = PNP (}\mu\text{mol/rxn)}$$

$$\text{PNP (}\mu\text{mol/rxn) / (0,05 ml} \times 20 \text{ min) = unidades / ml de } \alpha\text{-amilasa diluida}$$

$$\text{unidades / ml en enzima diluida} \times \text{dilución} = \text{Unidades Ceralfa / ml en muestra de } \alpha\text{-amilasa.}$$

30 ó

$$\Delta A_{400} \times 0,0442 \times \text{dilución de enzima} = \text{Unidades Ceralfa / ml en muestra de } \alpha\text{-amilasa.}$$

Este SOP es adaptado de un método de ensayo SOP CERALPHA HR^{MR} y un procedimiento de ensayo de α -amilasa Megazyme (método CERALPHA^{MR}) usando Reactivo HR^{MR} de amilasa, ICC Estándar No. 303.

35 La Figura 27 ilustra la base teórica del procedimiento de ensayo de α -amilasa CERALPHA^{MR}, a partir del procedimiento de ensayo de alfa-amilasa de Megazyme (Método CERALPHA^{MR}) usando reactivo HR^{MR} de amilasa: la figura ilustra el esquema de reacción total: una vez que la alfa-amilasa escinde un enlace en el sustrato p-nitrofenil maltosacárido bloqueado, el producto de reacción no bloqueado que contiene el constituyente p-nitrofenilo es instantáneamente escindido en glucosa y p-nitrofenilo libres por las cantidades en exceso de alfa-glucosidasa termoestable, que son partes integrales de la mezcla de sustrato y se libera el p-nitrofenilo libre. Se termina la
40 reacción y se desarrolla el color del fenolato por la adición de fosfato trisódico, pH alrededor de 11,0.

Extracción de glucoamilasa a partir de harina de maíz y ensayo de actividad para la hidrólisis de almidón: Extracción de glucoamilasa y ensayo de semilla individual de maíz

Para describir el procedimiento para la extracción de glucoamilasa y el ensayo de harina de maíz:

Preparación del regulador de extracción:

45 Borato de sodio 25 mM, 01% de Tween 20, regulador de pH 10

ES 2 575 912 T3

Procedimiento:

Trituración de semillas reunida:

- 5 Se reúnen 20 semillas. Se agregan las semillas a un vaso de trituración Kelco. Se coloca una bola de acero en cada vaso por encima de las semillas. Se coloca un anillo de goma alrededor del vaso y se coloca encima la tapa. Se tritura en el Kelco durante 45 segundos. Usando una espátula, se transfiere la harina del vaso a la plataforma de pesaje y luego se vierte en un tubo cónico de 15 mL. Se lava el vaso, anillo y la bola de acero con agua y jabón.

Extracción

1. Se añaden 3 ml del regulador de extracción a las muestras.
2. Se sellan los tubos con las tapas usando un martillo.
- 10 3. Se agita vigorosamente hasta que la harina se suspende en el regulador.
4. Se agita por rotación a temperatura ambiente en un rotor Rugged al 70% durante 5 minutos.
5. Se incuba a 45 °C en un baño de agua durante 30 minutos
6. Se agitan por rotación las muestras durante otros 5 minutos a temperatura ambiente después de la incubación.
7. Se centrifugan las muestras a 3000 rpm durante 5 minutos.
- 15 Ensayo GA (la información del reactivo viene después del procedimiento de ensayo)
1. Se diluyen muestras en el regulador de dilución para obtener un ensayo en el rango de 0,5 a 1,5 de OD₄₀₀.
2. Se enciende la máquina de PCR. Se fija el programa para incubar a 40 °C durante 10 min y luego una rampa descendente hasta a 4 °C.
3. Sobre hielo, se agregan 50 µl de cada muestra de glucoamilasa diluida a una placa de PCR.
- 20 4. Se agregan 150 µl de solución de terminación a cada pozo de una placa de microtitulación. Deben haber dos pozos de solución de terminación para cada muestra analizada.
5. Se mezclan 50 µl de reactivo de sustrato con la muestra. Se pipetea dos veces. Se prepara el sustrato como se indica en las instrucciones de Megazyme (10 ml de agua / botella).
- 25 6. Inmediatamente, se extraen 20 µl de la mezcla y se mezclan con 150 µl de solución de terminación. Se pipetea tres veces formando remolino. Este es el punto de tiempo 0.
7. Se coloca la placa en la incubadora y se presiona el botón de inicio. Asegurarse de que las muestras estén en la máquina cuando se inicia el ciclo ascendente. Se incuba la placa de PCR durante 10 min a 40 °C.
8. Permitir que la máquina realice el ciclo descendente hasta 4 °C mientras las muestras están en la máquina.
9. Esperar 2 minutos.
- 30 10. Se remueven 20 µl del producto y se mezcla con 150 µl de solución de terminación en una placa de microtitulación. Este es el punto de tiempo de 10 min.
11. Se deja reposar la placa durante 15 min.
12. Se lee la OD₄₀₀ en el lector de placas. Fijar Pathcheck en ON (encendido) (esto normalizará la lectura de la absorbancia de la muestra en una longitud de trayectoria de 1 cm).
- 35 13. Se calcula ΔA_{400} , el cambio en A_{400} entre el punto de tiempo a 0 y 10 min.

Preparación de soluciones patrón:

ES 2 575 912 T3

Substrato (Megazyme catalog #: R-AMGR3)

p-nitrofenil- α -maltósido (4 mM)

α -glucosidasa termoestable (5 U/ml)

5 Se disuelven los contenidos totales de un vial en 10,0 ml de agua MilliQ. Se divide en alícuotas y se almacenan congeladas. Se almacena sobre hielo durante el uso.

Solución de terminación: 2% de solución base trizma

2,0 g de base trizma (Sigma T-1503)

Se añade agua MilliQ hasta un volumen final de 100 ml

Regulador de dilución: regulador de acetato de sodio 200 mM (pH 4,5)

10 Cálculos de actividad

Se define una unidad de actividad de α -glucoamilasa como la cantidad de enzima, en presencia de un exceso de α -glucosidasa termoestable, necesaria para liberar un micromol de p-nitrofenol en un minuto bajo las condiciones de ensayo definidas.

De acuerdo con Megazyme, PNP 1 mM en 2% de base trizma produce una absorbancia de 18,1 a 400 nm.

$$15 \quad \text{Actividad de GA en muestras (U/ml)} = (A400/10) \times (170/10) \times (1/18,1) \times \text{Dilución} \\ = A400 \times 0,0939 \times \text{Dilución}$$

Este SOP es adaptado del ensayo de Megazyme de amiloglucosidasa usando p-nitrofenil- α -maltósido más α -glucosidasa termoestable.

Fermentación de almidón sin purificar: llevada a cabo como se describió anteriormente.

20 Resultados

Ensayo TAQMAN^{MR} y ensayo de actividad enzimática

25 Los números de copia de los transgenes se determinaron mediante ensayos TAQMAN^{MR} primarios y secundarios. Los cebadores específicos para el gen marcador seleccionable usados en todos los vectores de transformación de maíz descritos anteriormente, pmi, se usaron en los ensayos TAQMAN^{MR} primarios. Los cebadores específicos tanto para el gen pmi como para el gen marcador seleccionable bacteriano, spec, se usaron en los ensayos TAQMAN^{MR} secundarios.

30 La expresión de alfa-amilasas o glucoamilasas en semillas transgénicas T1 maduras secadas de constructos seleccionados también se analizó mediante ensayos de actividad enzimática. Los resultados de los números de copias génicas en eventos generados a partir del constructo 15749 y la expresión de la alfa-amilasa SEQ ID NO: 52 de ejemplo en estos eventos, fueron resumidos en la Tabla 2.2.

Tabla 2.2. Números de copias transgénicas y actividad de α -amilasa en eventos generados a partir del constructo 15749

Número de la planta	Actividad de α -amilasa en semillas reunidas (U/g)
1	25,6
2	23,5
3	21,3
4	21,2

ES 2 575 912 T3

Número de la planta	Actividad de α -amilasa en semillas reunidas (U/g)
5	20,2
6	18,4
7	15,0
8	14,3
9	14,0
10	12,7
11	12,2
12	11,1
13	10,2
14	9,9
15	9,6
16	7,6
17	6,9
18	6,8
19	6,4
20	6,4
21	6,0
22	5,4
23	5,3
24	5,3
25	4,7
26	4,7
27	4,5
28	4,2
29	4,0
30	3,9
31	3,7
32	3,6

ES 2 575 912 T3

Número de la planta	Actividad de α -amilasa en semillas reunidas (U/g)
34	3,1
35	3,0
36	2,8
37	2,5
38	2,4
39	2,0
40	1,5
41	0,4
42	0,4
43	0,3
44	0,2
45	0,1

Los resultados de los números de copias génicas en los eventos generados a partir del constructo 15751 y la expresión de la glucoamilasa SEQ ID NO: 48 en estos eventos se resumen en la Tabla 2.3:

Tabla 2.3. Números de copias transgénicas y actividad de glucoamilasa en eventos generados a partir del constructo 15751

5

Planta	Actividad de glucoamilasa en semillas reunidas (U/g)
1	4,86
2	4,48
3	4,47
4	4,22
5	4,20
6	4,08
7	4,05
8	3,94
9	3,88
10	3,84
11	3,82
12	3,68

ES 2 575 912 T3

Planta	Actividad de glucoamilasa en semillas reunidas (U/g)
13	3,65
14	3,57
15	3,50
16	3,47
17	3,43
18	3,33
19	2,91
20	1,13
21	0,73

La Tabla 2.4 resume los resultados de los números de copias génicas en eventos generados a partir del constructo 15756 y la expresión de la glucoamilasa SEQ ID NO: 26 en estos eventos. Las semillas de los múltiples eventos de los constructos 15761 y 15718 que expresan las alfa-amilasas de las SEQ ID NO: 66 y SEQ ID NO: 76, respectivamente, se reunieron y trituraron para formar muestras compuestas. Las dos muestras compuestas se evaluaron para las actividades de α -amilasa. En forma similar, las semillas de los múltiples eventos de los constructos 15742 y 15750 que expresan las glucoamilasas de las SEQ ID NO: 28 y SEQ ID NO: 18, respectivamente, también se reunieron y trituraron para formar muestras compuestas. Las dos muestras compuestas fueron evaluadas para las actividades de glucoamilasa. Los resultados se resumen en la Tabla 2.5.

- 5
- 10 Tabla 2.4 Números de copias transgénicas y actividad de glucoamilasa en eventos generados a partir del constructo 15756

Número de la planta	Actividad de glucoamilasa en semillas reunidas (U/g)
1	4,41
2	4,31
3	4,18
4	4,15
5	3,95
6	3,94
7	3,88
8	3,83
9	3,78
10	3,68
11	3,67
12	3,64

ES 2 575 912 T3

Número de la planta	Actividad de glucoamilasa en semillas reunidas (U/g)
13	3,60
14	3,58
15	3,45
16	3,27
17	3,18
18	3,16
19	3,14
20	3,10
21	3,07
22	3,00
23	2,75

Tabla 2.5. Actividad de α -amilasa o glucoamilasa en semillas generadas a partir de los constructos 15718, 15761, 15742 y 15750

Constructo	Variedad	Enzima expresada	Tipo de enzima	Actividad de la semilla reunida (U/g)
15718	JHAX707	SEQ ID NO: 76	α -amilasa	0,033
15761	JHAX708	SEQ ID NO: 66	α -amilasa	0,085
15742	JHAX709	SEQ ID NO: 28	Glucoamilasa	3,399
15750	JHAX710	SEQ ID NO: 18	Glucoamilasa	3,55

5 Fermentación de almidón sin purificar usando enzimas expresadas en maíz

Las semillas de los múltiples eventos de constructos seleccionados que expresan α -amilasas o glucoamilasas se reunieron y trituraron para formar muestras compuestas para fermentación de almidón sin purificar. En todos los experimentos, se usó la harina de maíz que contiene α -amilasas a la tasa de inclusión de 20% (p/p) y harina de maíz que contenía glucoamilasas a la tasa de inclusión de 50% (p/p). La harina de maíz diente amarillo II constituyó 30% (p/p) restante de la harina de maíz. Las fermentaciones de almidón sin purificar se llevaron a cabo siguiendo el SOP estándar descrito anteriormente (Sección I). La capacidad de estas combinaciones de α -amilasa y glucoamilasa expresadas en maíz para producir etanol en la fermentación de almidón sin purificar se resume en la Tabla 2.6.

10

Tabla 2.6. Rendimiento de etanol después de 72 horas de fermentación de almidón sin purificar usando enzimas de ejemplo “principales” expresadas en maíz sin optimización de codones*

Constructo (α-amilasa)	Constructo (Glucoamilasa)	Tasa de inclusión (α-amilasa) (% p/p)	Tasa de inclusión (Glucoamilasa) (% p/p)	Tasa de inclusión (Producto) (% p/p)	Rendimiento de etanol (% v/v)
15749 (SEQ ID NO: 52)	15756 (SEQ ID NO: 26)	20	50	30	6.96
15749 (SEQ ID NO: 52)	15742 (SEQ ID NO: 28)	20	50	30	8.90
15761 (SEQ ID NO: 66)	15751 (SEQ ID NO: 48)	20	50	30	7.30
15718 (SEQ ID NO: 76)	15750 (SEQ ID NO: 18)	20	50	30	8.33
N/A	N/A	0	0	100	2.25

5 **Ejemplo 28:** Expresión en plantas de ejemplos de enzimas usando genes optimizados de codones de maíz sintético y fermentación de almidón sin purificar usando material vegetal

10 Los genes sintéticos con codones optimizados para expresión en maíz fueron generados para 2 α-amilasas, la SEQ ID NO: 52 (codificada, por ejemplo, por la SEQ ID NO: 51) y la SEQ ID NO: 4 (codificada, por ejemplo, por la SEQ ID NO: 3) y 2 glucoamilasas, las SEQ ID NO: 48 (codificada, por ejemplo, por la SEQ ID NO: 47) y la SEQ ID NO: 26 (codificada, por ejemplo, por la SEQ ID NO: 25). La Tabla a continuación correlaciona la SEQ ID NO: de tipo silvestre con la SEQ ID NO: asignada a la secuencia optimizada por codones. Las enzimas escogidas para optimización de codones se seleccionaron con el fin de obtener el mayor rendimiento de etanol en la fermentación de almidón sin purificar. Los objetivos de este estudio eran generar eventos comerciales para el desarrollo del producto y generar suficiente material para evaluar el desempeño de la fermentación de almidón sin purificar a partir de las enzimas expresadas en maíz.

Enzima	SEQ ID NO: de tipo silvestre	SEQ ID NO: de codón optimizado
Amilasa	51	79
Amilasa	3	80
Glucoamilasa	47	81
Glucoamilasa	25	82

15

Métodos experimentales

Genes sintéticos

ES 2 575 912 T3

Los genes sintéticos fueron elaborados por GENEART^{MR}. Las secuencias fueron optimizadas por codones usando tecnología GENEOPTIMIZER^{MR} de propiedad de GENEART^{MR}. Las secuencias de los genes sintéticos se enlistan a continuación. El uso de codones se adaptó al codón de tendencia de genes de *Zea mays*.

Construcción del vector

- 5 La Tabla 3.1 resume las abreviaturas usadas para los elementos de ADN en los mapas mostrados a continuación.

Tabla 3.1. Abreviaturas usadas en los mapas de constructos comerciales

Nombre	Función	Descripción
cAmy (SEQ ID NO: 79)Apo-01	CDS	Gen que codifica la enzima de la SEQ ID NO: 52 de alfa-amilasa sintética de codón optimizado de maíz dirigido al apoplasto
cAmy (SEQ ID NO: 79)ER-01	CDS	Gen de alfa-amilasa de codón optimizado de maíz sintético dirigido al RE con secuencia señal de Gamma Zeina y secuencia KDEL que codifica la enzima de la SEQ ID NO: 52
cAmy (SEQ ID NO: 80)Apo-01	CDS	Gen de alfa-amilasa de codón optimizado de maíz sintético dirigido al apoplasto que codifica la enzima de la SEQ ID NO: 4
cAmy (SEQ ID NO: 80)ER-01	CDS	Gen de alfa-amilasa de codón optimizado de maíz sintético dirigido al RE con secuencia señal de Gamma Zeina y secuencia KDEL que codifica la enzima de la SEQ ID NO: 4
cGAmy (SEQ ID NO: 82)Apo-01	CDS	Gen de glucoamilasa de codón optimizado de maíz sintético dirigido al apoplasto que codifica la enzima de la SEQ ID NO: 26
cGAmy (SEQ ID NO: 82)ER-01	CDS	Gen de glucoamilasa de codón optimizado de maíz sintético dirigido al RE con secuencia señal de Gamma Zeina y secuencia KDEL que codifica la enzima de la SEQ ID NO: 26
cGAmy (SEQ ID NO: 81)Apo-01	CDS	Gen de glucoamilasa de codón optimizado de maíz sintético dirigido al apoplasto que codifica la enzima de la SEQ ID NO: 48
cGAmy (SEQ ID NO: 81)ER-01	CDS	Gen de glucoamilasa de codón optimizado de maíz sintético (con adición de retención en el RE de KDEL) dirigido al RE que codifica la enzima de la SEQ ID NO: 48

- 10 El vector de ejemplo designado como "15740", a continuación, es un vector binario para transformación de maíz que aloja al promotor específico de semillas prGTL-03 que dirige la versión optimizada de maíz sintético dirigida al apoplasto de la alfa-amilasa de la SEQ ID NO: 51 de ejemplo, la versión optimizada es la SEQ ID NO: 79 de ejemplo o, cAmy(SEQ ID NO: 79)Apo-01. Este binario contiene también un casete de Ubi-PMI-Nos para selección.

- 15 Vector de clonación: El constructo alias 15460ZeinAmyVN fue digerido con SacI y RsrII para extraer iPEPC9-01 y t35s-08. Asimismo, el constructo 15460 fue digerido con SacI y RsrII para remover tNOS-03-01. El terminador t35s-08/iPEPC9-01 se ligó en 15460 usando estos sitios (SY1709:6). La reacción de ligación fue transformada en células DH5-alfa; los transformantes se cribaron con PCR de colonia. Cuatro clones positivos fueron secuenciados y a partir de estas secuencias tanto las uniones de clonación como los sitios de enzimas de restricción fueron determinados como correctos (SY1709215-16). Para insertar el promotor prGTL-03, 15460 que contenía t35s-08/iPEPC9-01 fue digerido con BamHI y HindIII y se purificó en gel (SY1709:14). El constructo 11267 que contenía prGTL-03 también fue digerido con BamHI y HindIII y el promotor fue purificado en gel (SY171023-4). Luego fue ligado en la estructura de 15460 usando estos sitios y transformado en células competentes DH5-alfa. Diez transformantes se cribaron usando HindIII y BamHI para determinar si estaba presente prGTL-03 en el vector; todos los diez tenían patrón de banda correcto (SY1709225). Dos clones fueron secuenciados para confirmar uniones de clonación y sitios de enzimas de restricción y ambos fueron correctos sin diferencias en las secuencias (SY1709:34-35). Los patrones de glicerol fueron preparados a partir del clon #1 y se almacenaron a -80 grados Celsius. El vector de clonación fue digerido con BamHI/SacI y tratado con CIP. La versión de codón optimizado de maíz sintético de alfa-amilasa de SEQ ID NO: 51, conocida como SEQ ID NO: 79 (también conocida como cAmy(SEQ ID NO: 79)) fue digerida con BamHI y SacI y ligada en el sitio BamHI/SacI de vector de clonación prGTL-03 para crear el alias GTL+SYN(SEQ ID
- 20
- 25

NO: 79) (SY1773:48). Vector binario: un fragmento de Ascl/BamHI del vector 15468 fue ligado en el sitio Ascl/BamHI del vector binario 12678 para crear el alias "12678 Rsrll" (SY1533:189). Este clon fue luego digerido con Rsrll y tratado con CIP. El vector de clonación "GTL+SYN(SEQ ID NO: 79)" se digirió con SanDI/Rsrll y se ligó en el sitio de Rsrll para crear B-prGTL:(SEQ ID NO: 79):t35S:PMI (SY1773:56). Los clones positivos se identificaron por PCR y confirmaron por secuenciación de ADN.

El vector de ejemplo designado como "15741", como se ilustra en la Figura 37, a continuación, es un vector binario para transformación de maíz que aloja al promotor específico de semilla prGTL-03 que dirige la versión optimizada de maíz sintético dirigida al apoplasto de la glucoamilasa de SEQ ID NO: 25 de ejemplo, la versión optimizada es la SEQ ID NO: 82 de ejemplo o cGAmY(SEQ ID NO: 82)Apo-01. Este binario también contiene un casete de Ubi-PMI-Nos para selección.

Vector de clonación: El constructo, alias 15460ZeinAmyVN fue digerido con SacI y Rsrll para remover iPEPC9-01 y t35s-08. Igualmente, el constructo 15460 fue digerido con SacI y Rsrll para extraer tNOS-03-01. El terminador t35s-08/iPEPC9-01 se ligó en 15460 usando estos sitios (SY1709:6). La reacción de ligamiento fue transformada en células DH5-alfa; los transformantes se cribaron con PCR de colonias. Se secuenciaron cuatro clones positivos y a partir de estas secuencias se determinó que tanto las uniones de clonación como los sitios de las enzimas de restricción eran correctos (SY1709:15-16). Para insertar al promotor prGTL-03, se digirió 15460 que contenía t35s-08/iPEPC9-01 se digirió con BamHI y HindIII y se purificó en gel (SY1709:14). El constructo 11267 que contenía prGTL-03 también se digirió con BamHI y HindIII y el promotor se purificó en gel (SY1710:3-4). Luego se lo ligó en la estructura 15460 usando estos sitios y se transformó en células DH5-alfa competentes. Diez transformantes fueron cribados usando HindIII y BamHI para determinar si prGTL-03 estaba presente en el vector; todos los diez tuvieron patrón de banda correcto (SY1709:25). Dos clones se secuenciaron para confirmar las uniones de clonación y los sitios de enzimas de restricción y ambos eran correctos sin ninguna diferencia de secuencia (SY1709:34-35). Los patrones de glicerol fueron preparados a partir del clon #1 y almacenados a -80 grados Celsius. El vector de clonación se digirió con BamHI/SacI y tratado con CIP. La versión de codón optimizado de maíz sintético de glucoamilasa de la SEQ ID NO: 25, conocida como SEQ ID NO: 82 (también conocida como cAmy(SEQ ID NO: 82 o SYN(SEQ ID NO: 82)) se digirió con BamHI y SacI y se ligó en el sitio de BamHI/SacI del vector de clonación prGTL-03 para crear el alias GTL+SYN(SEQ ID NO: 82) (SY1773:48). Vector binario: Un fragmento de Ascl/BamHI del vector 15468 se ligó en el sitio de Ascl/BamHI del vector binario 12678 para crear el alias "12678 Rsrll" (SY1533:189). Este clon fue luego digerido con Rsrll y tratado con CIP. El vector de clonación "GTL+SYN(SEQ ID NO: 82)" se digirió con SanDI/Rsrll y se ligó en el sitio Rsrll para crear B-prGTL:(SEQ ID NO: 82):t35S:PMI (SY1773:66). Los clones positivos fueron identificados por PCR y confirmados por ADN.

El vector de ejemplo designado como "15742", como se ilustra en la Figura 38, a continuación, es un vector binario que aloja (que comprende) al promotor específico de semillas prGTL-03 que dirige la versión optimizada de maíz sintético dirigida al apoplasto de la glucoamilasa de SEQ ID NO: 47 de ejemplo, la versión optimizada es la SEQ ID NO: 81 de ejemplo. Este binario también contiene un casete de Ubi:PMI:NOS para selección.

Vector de clonación: El constructo alias15460ZeinAmyVN se digirió con SacI y Rsrll para remover iPEPC9-01 y t35s-08. Igualmente, el constructo 15460 se digirió con SacI y Rsrll para remover tNOS-03-01. El terminador t35s-08/iPEPC9-01 se ligó en 15460 usando estos sitios (SY1709:6). La reacción de ligación fue transformada en células DH5-alfa; los transformantes se cribaron con PCR de colonias. Se secuenciaron cuatro clones positivos y a partir de estas secuencias se determinó que tanto las uniones de clonación como los sitios de enzimas de restricción eran correctos (SY1709:15-16). Para insertar al promotor prGTL-03, 15460 que contiene t35s-08/iPEPC9-01 se digirió con BamHI y HindIII y se purificó en gel (SY1709:14). El constructo 11267 que contenía prGTL-03 también se digirió con BamHI y HindIII y el promotor fue purificado en gel (SY1710:3-4). Luego se ligó a la estructura 15460 usando estos sitios y transformado en las células DH5-alfa competentes. Se cribaron diez transformantes usando HindIII y BamHI para determinar si prGTL-03 estaba presente en el vector; todos los diez tuvieron patrón de banda correcto (SY1709:25). Se secuenciaron dos clones para confirmar las uniones de clonación y los sitios de enzimas de **restricción** y ambos fueron correctos sin ninguna diferencia de secuencias (SY1709:34-35). Las patrones de glicerol se prepararon a partir del clon #1 y se almacenaron a -80 grados Celsius. El vector de clonación se digirió con BamHI/SacI y tratado por CIP. La versión de codón optimizado de maíz sintético de glucoamilasa de la SEQ ID NO: 47, conocida como SEQ ID NO: 81 (también conocida como cAmy(SEQ ID NO: 81) se digirió con BamHI y SacI y se ligó en el sitio de BamHI/SacI del vector de clonación prGTL-03 para crear el alias GTL+SYN(SEQ ID NO: 81)) (SY1773:48). Vector binario: Un fragmento de Ascl/BamHI del vector 15468 se ligó en el sitio de Ascl/BamHI del vector binario 12678 para crear el alias "12678 Rsrll" (SY1533:189). Este clon luego fue digerido con Rsrll y tratado con CIP. El vector de clonación "GTL+SYN(SEQ ID NO: 81)" se digirió con SanDI/Rsrll y se ligó en el sitio de Rsrll para crear B-prGTL:(SEQ ID NO: 81):t35S:PMI (SY1773:56). Los clones positivos se identificaron por PCR y se confirmaron mediante secuenciación de ADN (SY1773:58).

El vector de ejemplo designado como "15743", como se ilustra en Figura 39 a continuación, es un vector binario que aloja (que comprende) al promotor específico de la semilla prGTL-03 que dirige la versión optimizada de maíz sintético dirigida al apoplasto del ejemplo de la alfa-amilasa de la SEQ ID NO: 3 de ejemplo, la versión optimizada es la SEQ ID NO: 80. Este binario también contiene un casete de Ubi:PMI:NOS para selección.

Vector de clonación: El constructo alia515460ZeinAmyVN se digirió con SacI y RsrII para remover iPEPC9-01 y t35s-08. Igualmente, el constructo 15460 se digirió con SacI y RsrII para remover tNOS-03-01. El terminador t35s-08/iPEPC9-01 se ligó en 15460 usando estos sitios (SY1709:6). La reacción de ligación fue transformada en células DH5-alfa; los transformantes se cribaron con PCR de colonias. Se secuenciaron cuatro clones positivos y a partir de estas secuencias se determinó que tanto las uniones de clonación como los sitios de enzimas de restricción eran correctos (SY1709:15-16). Para insertar el promotor prGTL-03, se digirió 15460 que contiene t35s-08/iPEPC9-01 con BamHI y HindIII y se purificó en gel (SY1709:14). El constructo 11267 que contenía prGTL-03 también se digirió con BamHI y HindIII y el promotor se purificó en gel (SY1710:3-4). Luego se lo ligó en la estructura de 15460 usando estos sitios y se lo transformó en células DH5-alfa competentes. Se cribaron diez transformantes usando HindIII y BamHI para determinar si prGTL-03 estaba presente en el vector; todos los diez tenían un patrón de bandas correcto (SY1709:25). Se secuenciaron dos clones para confirmar uniones de clonación como los sitios de enzimas de restricción y ambos eran correctos sin diferencias de secuencia (SY1709:34-35). Los patrones de glicerol se prepararon a **partir del clon #1 y se almacenaron a -80 grados Celsius. El vector de clonación se digirió con BamHI/SacI y se lo trató con CIP.** La versión de codón optimizado de maíz sintético de alfa-amilasa SEQ ID NO: 3, conocida como SEQ ID NO: 80 (también conocida como cAmy(SEQ ID NO: 80) se digirió con BamHI y SacI y se ligó en el sitio de BamHI/SacI del vector de clonación prGTL-03 para crear el alias GTL+SYN(SEQ ID NO: 80) (SY1773:48). Vector binario: Un fragmento de AscI/BamHI del vector 15468 se ligó en el sitio de AscI/BamHI del vector binario 12678 para crear el alias "12678 RsrII" (SY1533:189). Este clon luego fue digerido con RsrII y tratado por CIP. El vector de clonación "GTL+SYN(SEQ ID NO: 80)" se digirió con SanDI/RsrII y se ligó en el sitio de RsrII para crear B-prGTL:(SEQ ID NO: 80):t35S:PMI (SY1773:56). Los clones positivos se identificaron por PCR y confirmaron por ADN secuenciamiento.

El vector de ejemplo designado "15862", como se ilustra en la Figura 40 a continuación, es un vector binario para transformación de maíz que aloja (que comprende) el promotor específico de semillas prGTL-03 que maneja la versión optimizada por maíz sintético de la glucoamilasa de SEQ ID NO: 47 de ejemplo, la versión optimizada es la SEQ ID NO: 81 de ejemplo o cG Amy(SEQ ID NO: 81)ER-01 que es dirigida al ER por la señal de retención en ER KDEL. Este binario también contiene un casete Ubi-PMI-Nos para selección.

La versión de codón optimizado de maíz de la SEQ ID NO: 47, conocida como SYN(SEQ ID NO: 81) (cAMY(SEQ ID NO: 81)-03) fue replicada por PCR y clonada con TOPO con cebadores de manera tal que se agregó una señal de retención en ER KDEL al extremo 3 prima (SY1773:253). La presencia de KDEL se confirmó por secuenciamiento de ADN (SY1773:254 & 60). Los datos de la secuencia de ADN no confirmaron discrepancias (SY1773:63). El clon #1 se digirió con BamHI/BglII, purificado por gel de agarosa y se ligó en un vector de clonación que aloja el promotor prGTL-03 y t35s-08 para crear el alias "GTL+SYN(SEQ ID NO: 81) kdel v2" (SY1773:266). Un casete génico prGTL-03:cAmy(SEQ ID NO: 81)kdel:t353-08 fue ligado como un fragmento de SanDI/RsrII en el sitio de RsrII de una versión modificada de vector binario 12678 para crear B-prGTL-(SEQ ID NO: 81)KDEL:PMI (SY1773:83). La integridad del vector binario se confirmó por PCR y secuenciamiento de ADN (SY1773:88).

El vector ejemplificativo designado "15880", como se ilustra en la Figura 41 a continuación, es un vector binario que contiene (que comprende) la versión de codón optimizado de maíz de la alfa amilasa de SEQ ID No: 51 de ejemplo, la versión de codón optimizado es la SEQ ID NO: 79 de ejemplo, se expresó usando el promotor de glutelina de arroz (prGTL-03). La alfa amilasa también contiene una secuencia señal de Gamma Zeina y secuencia de KDEL para retención en ER. Este vector contiene un casete de Ubi-PMI-Nos para selección de manosa.

La alfa amilasa de SEQ ID NO: 51 optimizada por codones de maíz, conocida como SEQ ID NO: 79 o cAmy(SEQ ID NO: 79) fue replicada por PCR con cebadores de manera tal que una señal de retención en ER (KDEL) se agregó al extremo 3 prima (SY1777:180). El producto de PCR fue luego purificado por gel, clonado por TOPO y transformado en células competentes TOP10. La presencia del KDEL sin discrepancias se confirmó por secuenciamiento (SY1773:173). Los clones TOPO se digirieron con BamHI y BglII para obtener el gen de amilasa modificado que fue luego purificado por gel. El producto purificado por gel se ligó en 15460 que contiene prGTL-03 (promotor de glutelina de arroz) y t35s-08 para crear el alias "GTL + syn(SEQ ID NO: 79) KDEL v2" (SY1773:280; SY1777:180). Después de la transformación en células competentes DHSa, los clones se cribaron para orientación de inserto por PCR. Asimismo, tanto CDS como también las uniones de clonación se secuenciaron antes de la construcción de vector binario (SY1777:189, 191). El casete génico prGTL:cAmy(SEQ ID NO: 79):t35s-08 se digirió con SanDI/RsrII y se ligó en una versión modificada del constructo 12678 en el sitio de RsrII (SY1818:112). Este constructo se modificó ligando un fragmento de AscI/BamHI del vector 15468 en el sitio de AscI/BamHI del vector binario 12678 para crear el alias 12678 RsrII" (SY1533:189). La reacción de ligamiento se transformó en células competentes TOP10 y los transformantes se cribaron en forma posterior por PCR de colonias, Dos de los transformantes cribados fueron positivos para el casete génico y se digirieron con BamHI para confirmar su presencia. Después de la confirmación de digestión, el casete completo a partir del clon # 5 fue secuenciado (SY1818:34). Los datos confirmaron que todo el casete, inclusive todas las uniones, está presente sin discrepancias en la secuencia (SY1818:37-38).

El vector ejemplificativo designado "15884", como se ilustra en la Figura 42 a continuación, es un vector binario que contiene (que comprende) la versión optimizada por codones de maíz de la SEQ ID NO: 25 de ejemplo, la versión de codón optimizado de maíz es la SEQ ID NO: 82 de ejemplo, se expresó usando el promotor glutelina de arroz

(prGTL-OS). La glucoamilasa también contiene una secuencia señal de Gamma Zeina y señal de retención en ER (KDEL). Este vector contiene un casete de Ubi-PMI-Nos para selección de manosa.

La versión de codón optimizado de maíz de glucoamilasa de SEQ ID NO: 25, conocida como SEQ ID NO: 82 o cAmy(SEQ ID NO: 82) fue replicada por PCR con cebadores de manera tal que una señal de retención en ER (KDEL) se agregó al extremo 3 prima (SY1777:184). El producto de PCR fue luego purificado por gel, clonado por TOPO y transformado en células competentes TOP10. La presencia del KDEL sin discrepancias se confirmó por secuenciamiento (SY1777:190). Los clones TOPO se digirieron con BamHI y BglII para obtener el gen de amilasa modificado que fue luego purificado por gel. El producto purificado por gel se ligó en 15460 que contenía prGTL-03 (promotor de glutelina de arroz) y t355-08 para crear el alias "GTL syn(SEQ ID NO: 82) KDEL v2" (SY1818; 1-3). Después de la transformación en las células competentes Top10, los clones se cribaron para la orientación del inserto por PCR. Asimismo, tanto CDS como también las uniones de clonación se secuenciaron antes de la construcción del vector binario (SY1818229-30). Se determinó a partir del secuenciamiento que existe un cambio en el par de bases simple en la secuencia codificadora de cAmy(SEQ ID NO: 82) en el par de bases 156. Es un cambio de T a C que no cambia la secuencia de aminoácido de la proteína (SY1818: 39-40). El casete génico prGTL: syn(SEQ ID NO: 82) KDEL: t35s-08 se digirió con SanDI/RsrII y se ligó en una versión modificada del constructo 12678 en el sitio de RsrII (SY1818:31). Este constructo se modificó ligando un fragmento de AscI/BamHI del vector 15468 en el sitio de AscI/BamHI del vector binario 12678 para crear el alias "12678 RsrII" (SY1533:189). La reacción de ligamiento se transformó en células competentes TOP10 y los transformantes se cribaron en forma posterior usando digestiones de enzimas de restricción. Dos de los transformantes se digirieron con varias combinaciones de las enzimas de restricción inclusive NotI, SacI, EcoRV y PstI para confirmar la presencia - del casete génico (SY1818:45). Después de la confirmación de digestión, todo el casete a partir del clon #7 fue secuenciado (SY1818246-47). Los datos confirmaron que todo el casete, inclusive todas las uniones, está presente y solamente contiene el cambio de par de bases simple descrito anteriormente (SY1818:50).

El vector ejemplificativo designado "15890", como se ilustra en la Figura 43 a continuación, es un vector binario para transformación de maíz que contiene la versión optimizada por codones de maíz de la alfa-amilasa de SEQ ID NO: 3 de ejemplo, la versión de codón optimizado de maíz es la SEQ ID NO: 80 de ejemplo o cAmy(SEQ ID NO: 80) se expresó usando el promotor de glutelina de arroz (prGTL-03). La glucoamilasa también contiene una secuencia señal de Gamma Zeina y señal de retención en ER (KDEL). Este vector contiene un casete de Ubi-PMI-Nos para selección de manosa.

La versión de codón optimizado de maíz de la alfa-amilasa de SEQ ID NO: 3 de ejemplo, conocida como cAmy(SEQ ID NO: 80) fue replicada por PCR con cebadores de manera tal que una señal de retención en ER (KDEL) se agregó al extremo 3 prima (SY1777:184). El producto de PCR fue luego purificado por gel, clonado por TOPO y transformado en células competentes TOP10. La presencia del KDEL sin discrepancias se confirmó por secuenciamiento (SY1777:190). Los clones TOPO se digirieron con BamHI y BglII para obtener el gen de amilasa modificado que fue luego purificado por gel. El producto purificado por gel se ligó en 15460 que contiene prGTL-03 (promotor de glutelina de arroz) y t355-08 para crear el alias "GTL + syn(SEQ ID NO: 80) KDEL v2" (SY1818; 1-3). Después de la transformación en células competentes TOP10, los clones se cribaron para la orientación del inserto por PCR. Asimismo, tanto CDS como también las uniones de clonación se secuenciaron antes de la construcción del vector binario (SY1818:51). El casete génico prGTL: syn(SEQ ID NO: 80) KDEL: t35s-08 se digirió con SanDI/RsrII y se ligó en una versión modificada del constructo 12678 en el sitio de RsrII (SY1818:52). Este constructo se modificó ligando un fragmento de AscI/BamHI del vector 15468 en el sitio de AscI/BamHI del vector binario 12678 para crear el alias 12678 RsrII" (SY1533:189). La reacción de ligamiento se transformó en células competentes TOP10 y los transformantes se cribaron en forma posterior usando PCR y digestiones de enzimas de restricción. Dos de los transformantes que tuvieron resultados de PCR positivos se digirieron con varias combinaciones de enzimas de restricción inclusive NcoI, KpnI, EcoRV, XbaI y BglII para confirmar la presencia del casete génico (SY1818:65-66). Después de la confirmación de digestión, todo el casete a partir del clon #2 fue secuenciado (SY1818:67). Los datos confirmaron que todo el casete, inclusive todas las uniones, está presente sin diferencias de secuencia (SY1818:69-70).

El vector ejemplificativo designado "15889", como se ilustra en la Figura 44 a continuación, es un vector binario para transformación de maíz que contiene una pila (stack) molecular de la versión optimizada por codones de glucoamilasa la de SEQ ID NO: 25 de ejemplo, la versión de codón optimizado es la SEQ ID NO: 82 o la cGAmy(SEQ ID NO: 82) de ejemplo fue expresada por dos promotores diferentes. En el primer casete, el cAmy(SEQ ID NO: 82) es dirigido por el promotor de glutelina de arroz y tiene una secuencia señal de Gamma Zeina y señal KDEL para retención en el ER. En el segundo casete, la glucoamilasa es manejada por el promotor de inhibidor de alfa tripsina y es dirigida al apoplasto con la secuencia señal de Gamma Zeina. Este constructo también contiene un casete de Ubi-PMI-Nos para selección de manosa.

El promotor de arroz, prATI-01 fue mutagenizado a través de mutagénesis dirigida a sitio para extraer un sitio de RsrII interno y fue sometido como un nuevo componente en el constructo #15882. Después de la confirmación en secuencia de la extracción de este sitio, el vector de clonación "15460+prATI+t35508" se digirió con BamHI/SacI, tratado por CIP y se extrajo con gel (SY1777:162). La glucoamilasa sintética optimizada por codones Syn(SEQ ID

NO: 82) (BamHI/SacI) se ligó en la estructura digerida y el constructo resultante se conoce como "15460prATIm.syn(SEQ ID NO: 82)" (SY1777:170). Los transformantes se cribaron para la presencia del gen, usando tanto PCR de colonias como una digestión de diagnóstico de NcoI/BglII. Tres clones positivos se secuenciaron; se confirmaron las uniones de clonación para los tres clones (SY1777:177-179). Después de la confirmación de secuencia. "15460prATIm.syn(SEQ ID NO: 82)" se digirió con SanDI/RsrII y el casete de ATI: syn(SEQ ID NO: 82);t35s-08 se ligó en la estructura binaria y se transformó en células competentes TOP10 (SY1818:53). Diez transformantes fueron seleccionados para cribado, por PCR, uno de los cuales confirmó tener el casete que contiene el promotor de ATI (SY1818:56). El clon positivo fue luego cribado con digestiones de enzima de restricción de diagnóstico usando las siguientes enzimas: EcoRI, BglII, NcoI y HindIII (SY1818:260, 66). Se determinó que el clon #9 era correcto analizando el patrón de banda de varias digestiones. Asimismo, las uniones de clonación del clon 9 se secuenciaron y también se confirmaron como correctas (SY1818:65).

El vector ejemplificativo designado "15934", como se ilustra en la Figura 45 a continuación. es un vector binario para transformación de planta que contiene una pila molecular de la versión de codón optimizado de la glucoamilasa de SEQ ID NO: 47 de ejemplo. la versión optimizada por codones es la SEQ ID NO: 81 o la cGAmy(SEQ ID NO: 81) de ejemplo fue expresada por dos promotores diferentes. En el primer casete. El GAmy(SEQ ID NO: 81) es manejado por el promotor de glutelina de arroz y tiene una señal KDEL para retención en ER (el componente es cGAmy(SEQ ID NO: 81)ER-01). En el segundo casete, el GAmy(SEQ ID NO: 81) es manejado por el promotor de inhibidor de alfa tripsina y es dirigido al apoplasto (el componente es cGAmy(SEQ ID NO: 81)Apo-01). Este constructo también contiene un casete de Ubi-PMI-Nos para de selección manosa. NOTA: Se determinó por digestión de restricción y secuenciamiento que una de las inserciones de limite izquierdo había desaparecido. El proyecto eligió proceder con el vector.

El vector de clonación "15460+prATI+t35508" se digirió con BamHI/SacI. se trató por CIP y se extrajo con gel (SY1777:162). La glucoamilasa sintética optimizada por codones cGAmy(SEQ ID NO: 81)Apo-01 (BamHI/SacI) se ligó la estructura digerida y el constructo resultante se conoce como "15460 prATIm.syn(SEQ ID NO: 81)" (SY1777:170). Los transformantes se cribaron para la presencia del gen, usando tanto PCR de colonias como una digestión de NcoI/BglII de diagnóstico. Tres clones positivos se secuenciaron; y se confirmaron las uniones de clonación para los tres clones (SY1777:177-179). Después de la confirmación de secuencia, 15460prATIm.syn(SEQ ID NO: 81) se digirió con SanDI/RsrII y el casete de ATI: cGAmy(SEQ ID NO: 81)Apo-01;t35s-08 se purificó con gel (SY1773:72). El vector binario "B-prGTL: (SEQ ID NO: 81)KDEL: PMI" fue construido y etiquetado como constructo #15862 (SY1773:92). Fue alineado con RsrII. tratado por CIP y extraído con gel. Después de la purificación, el casete de ATI: cGAmy(SEQ ID NO: 81)Apo-01;t35s-08 se ligó en la estructura binaria y se transformó en células competentes TOP10 (SY1818:22-23). Los transformantes se cribaron inicialmente con PCR de colonias usando cebadores que solamente se anillaban en el promotor de ATI. Se seleccionaron diez transformantes y se confirmó que tenían el casete que contenía el promotor de ATI. El clon 8 fue seleccionado para análisis de restricción para confirmar de manera adicional la presencia del casete en el vector binario. Dos enzimas que se cortan en el promotor de ATI en forma específica así como también la estructura fueron seleccionadas para descartar la posibilidad de un casete doble. Cuando el clon #8 se digirió con AvrII y NcoI, se determinó que solamente existía una copia de cada casete presente (SY1818: 93 - 94).

Análisis de actividad enzimática

Se usó análisis de semillas individuales o análisis de semillas reunidas para medir la actividad de alfa-amilasa o glucoamilasa en semillas transgénicas. Para el análisis de semillas individuales, se seleccionaron al azar 12 semillas de cada evento y se las trituró en forma individual. Para el análisis de semillas reunidas, se seleccionaron al azar 20 semillas de eventos seleccionados, se las acopió y molió. Las harinas fueron luego evaluadas usando el ensayo CERALPHA HR^{MR} de Megazyme para las alfa-amilasas o el ensayo de glucoamilasa de Megazyme para las glucoamilasas. Los procedimientos de operación estándares del ensayo (SOP) se describieron con antelación, inclusive el "análisis de actividad enzimática", la extracción de alfa-amilasa de la harina de maíz y el ensayo de actividad, la extracción de glucoamilasa de la harina de maíz y el ensayo de actividad y el protocolo de fermentación de almidón sin purificar.

Resultados

Ensayo Taqman y ensayo de actividad enzimática

Los números de copias de los transgenes se determinaron por ensayos primarios y secundarios de Taqman. Los cebadores específicos para el gen marcador seleccionable usados en los vectores de transformación de maíz descritos con antelación, pmi, se usaron en los ensayos primarios de Taqman. Los cebadores específicos para el gen pmi, el gen marcador seleccionable bacteriano, spec y genes que codifican las alfa-amilasas o glucoamilasas se usaron en los ensayos secundarios de Taqman.

La expresión de alfa-amilasas o glucoamilasas en las semillas transgénicas T1 secas maduras de constructos seleccionados también fue analizada mediante ensayos de actividad enzimática. Los resultados de los números de

ES 2 575 912 T3

copias de genes en los eventos seleccionados generados a partir del constructo 15840 se resumieron en la Tabla 3.2. La expresión de la alfa-amilasa de codón optimizado de la SEQ ID NO: 79 de ejemplo fue medida en estos eventos mediante análisis de actividad enzimática en 12 semillas seleccionadas al azar en forma individual. La actividad promedio de las 12 semillas para cada evento también se muestra en la Tabla 3.2, a continuación.

5 En forma similar, la Tabla 3.3 muestra los números de copias de genes de los eventos seleccionados generados del constructo 15841. La expresión de la glucoamilasa de codón optimizado de la SEQ ID NO: 82 de ejemplo fue también medida en estos eventos mediante análisis de actividad enzimática en 12 semillas seleccionadas al azar en forma individual, y la actividad promedio de las 12 semillas para cada evento se resume en la Tabla 3.3, a continuación.

10 Veinte (20) semillas de cada uno de los eventos seleccionados generados a partir de los constructos 15842 y 15843, por otro lado, fueron reunidas para ensayos enzimáticos para determinar los niveles de expresión de la glucoamilasa de la SEQ ID NO: 81 de codón optimizado y la alfa-amilasa de codón optimizado de la SEQ ID NO: 80 de ejemplos en estos eventos, respectivamente. Los resultados se resumen en las Tablas 3.4 y 3.5, respectivamente, a continuación. Los números de copias transgénicas también son ilustrados.

15 Tabla 3.2. Números de copias transgénicas y actividad de alfa-amilasa en los eventos seleccionados generados a partir del constructo 15840

Número de la planta	Actividad promedio de 12 semillas (U/g)
1	377,3
2	291,9
3	95,1
4	119,4
5	155,4
6	129,4
7	125,3
8	539,4
9	1362,4
10	195,1
11	88,4
12	458,2
13	722,9
14	151,7
15	164,8

Tabla 3.3. Números de copias transgénicas y actividad de glucoamilasa en eventos seleccionados generados a partir del constructo 15841

Número de la planta	Actividad promedio de 12 semillas (U/g)
1	2,5

ES 2 575 912 T3

Número de la planta	Actividad promedio de 12 semillas (U/g)
2	3,5
3	2,6
4	2,8
5	2,4
6	3,0
7	2,6
8	2,5
9	2,6
10	2,8
11	3,0
12	3,0
13	2,9
14	3,4
15	2,3
16	2,5
17	2,6
18	3,3
19	2,4
20	2,9
21	3,5
22	2,3
23	2,5
24	2,3

Tabla 3.4. Números de copias transgénicas y actividad de glucoamilasa en eventos seleccionados generados a partir del constructo 15842

Número de la planta	Actividad de las semillas reunidas (U/g)
1	1,5
2	1,3

ES 2 575 912 T3

Número de la planta	Actividad de las semillas reunidas (U/g)
3	1,3
4	1,2
5	1,3
6	1,6
7	1,4
8	1,4
9	1,3
10	1,2
11	1,4
12	1,6
13	1,7
14	1,2
15	1,3
16	1,6
17	1,4
18	1,6
19	1,6
20	1,6

Tabla 3.5. Números de copias transgénicas y actividad de alfa-amilasa en eventos seleccionados generados a partir del constructo 15843

Número de la planta	Actividad de las semillas reunidas (U/g)
1	116,14
2	107,07
3	105,02
4	277,65
5	240,90
6	165,25
7	425,17

ES 2 575 912 T3

Número de la planta	Actividad de las semillas reunidas (U/g)
8	151,02
9	139,74
10	245,27
11	197,68
12	179,77
13	368,59
14	350,91
15	225,03
16	233,18
17	366,24
18	220,36

Tabla 3.6 Producción de etanol después de 72 horas de fermentación de almidón crudo utilizando enzimas "principales" de ejemplo expresadas en maíz utilizando genes sintéticos de codón optimizado de maíz

Constructo (α -amilasa)	Constructo (Glucoamilasa)	Tasa de inclusión (α -amilasa) (% p/p)	Tasa de inclusión (Glucoamilasa) (% p/p)	Tasa de inclusión (Producto) (% p/p)	Rendimiento de etanol (% v/v)
15840 (SEQ ID NO: 79)	15841 (SEQ ID NO: 82)	20	50	30	16,77
15840 (SEQ ID NO: 79)	15842 (SEQ ID NO: 81)	20	50	30	17,08
15843 (SEQ ID NO: 80)	15841 (SEQ ID NO: 82)	20	50	30	7,30
15843 (SEQ ID NO: 80)	15842 (SEQ ID NO: 81)	20	50	30	10,60
N/A	N/A	0	0	100	2,25

5 Fermentación de almidón sin purificar usando enzimas expresadas en maíz

Las semillas de los eventos mostradas anteriormente que expresan alfa-amilasas y glucoamilasas fueron reunidas y trituradas para elaboración de muestras compuestas para fermentaciones de almidón sin purificar. En todos los experimentos, la harina de maíz que contenía alfa-amilasas fue usada en la tasa de inclusión de 20% (p/p) y la harina de maíz que contenía glucoamilasas a la velocidad de inclusión de 50% (p/p). La harina de maíz del producto Yellow Dent II laborada a partir del 30% (p/p) restante de la harina de maíz. Las fermentaciones de almidón sin procesar se llevaron a cabo siguiendo el SOP estándar descrito anteriormente (Sección I). La capacidad de estas combinaciones de alfa-amilasa y glucoamilasa expresadas en maíz que utiliza genes sintéticos de codón optimizado de maíz para producir etanol en una fermentación de almidón sin purificar se resume en la Tabla 3.6, a continuación.

Ejemplo 29: Constructos de expresión de *Pichia pastoris*

Este ejemplo describe la expresión de enzimas en levadura utilizando, por ejemplo, sistemas de expresión de *Pichia* de ejemplo.

5 Para la construcción de constructos de expresión en vectores pPICZalfa y pAO815 (ambos de Invitrogen, Carlsbad, CA), se utilizó tecnología de clonación Xi. Se digirió pPICZalfa con EcoRI, y se lo trató con cócteles de clonación Xi (de acuerdo con el protocolo del fabricante de Genlantis, una división del Gene Therapy Systems, Inc., San Diego, CA). Los genes fueron amplificados por reacciones PCR haciendo coincidir el extremo con las secuencias de vector. Para la transformación de huéspedes *Pichia* se digirió el ADN del plásmido verificado para volver al ADN lineal y transformarlo en huéspedes *Pichia*. Los transformantes se seleccionaron bajo placas deficientes en Zeocina (pPICZalfa) o en histidina.

Todas las enzimas se obtuvieron por expresión en *Pichia pastoris*, excepto para la SEQ ID NO: 78 (codificada por SEQ ID No: 77), que fue expresada en *Pseudomonas fluorescens* (véase JBC, 2002, 277(29): 26501 - 26507).

I. Caracterización de glucoamilasas y amilasas.

Métodos:

15 6. Determinación de la concentración de proteínas

Los sobrenadantes liofilizados de los cultivos de *P. pastoris* que expresan glucoamilasas y α -amilasas se suspendieron en agua a una concentración de ~10 mg de polvo / ml. Después de la determinación del contenido de proteína por el protocolo Bradford, 5 μ g de una muestra de proteína y solución BSA estandarizada se corrieron sobre un gel de gradiente de 4 - 20% de Tris-Glicina. El gel se escaneó en un escáner de gel BioRad GS800 después de tinción con azul de Coomassie. El software Bio-Rad QUANTITY ONE se usó para la cuantificación de las bandas de BSA y glucoamilasa (o α -amilasas) y se calculó luego la concentración real de enzimas. La concentración de proteínas se ajustó en consecuencia y se confirmó por SDS PAGE adicional.

7. Determinación de velocidades de reacción iniciales.

25 A menos que se indique lo contrario, los ensayos se realizaron por triplicado a 37 °C y pH 5,0 en regulador (NaCH₃CO₂ 50 mM, CaCl₂ 10 mM; NaN₃ 10 mM y 0,01% de Triton X-100) que contiene 1% de almidón sin purificar, o 0,5% de dextrina o 1% de "almidón de maíz soluble" (véase la nota más abajo sobre la preparación de "almidón soluble"). Los ensayos se realizaron a una escala de 0,5 ml para glucoamilasa y a escala de 0,25 ml para alfa-amilasa en una incubadora de mesa Eppendorf con agitación constante (800 rpm).

30 Para las glucoamilasas, las reacciones comenzaron al agregar la enzima (concentración final de 0,25 μ g/ml) a la mezcla de reacción. A 0, 2,5, 5, 7,5, 10, 15, 20 y 30 min, se retiraron alícuotas de 50 μ l de alícuotas de las reacciones y se neutralizaron por adición a 100 μ l de regulador Tris 1 M, pH 7,5.

35 Para las alfa-amilasas, las reacciones comenzaron al agregar la enzima (concentración final de 0,4 μ g de proteína total/ml para la SEQ ID No: 56, SEQ ID NO: 2 y SEQ ID NO: 52; 2 μ g/ml para la SEQ ID NO: 62; 4 μ g/ml para la SEQ ID NO: 70 y la SEQ ID NO: 66) a la mezcla de reacción y se retiraron alícuotas de 10 μ l de las reacciones y se neutralizaron en reactivo de BCA a 2,5, 10, 15, 20, 25, 30 y 40 min.

Para la determinación de los perfiles de temperatura, los ensayos se realizaron a 30, 34, 37 y 40 °C.

40 El efecto de pH sobre las actividades de glucoamilasa y amilasa se evaluaron a pH de 3,5, 4, 5, 6 y 7, usando el regulador de rango amplio de pH de Britton - Robinson (CH₃COOH 50 mM; H₃PO₄; H₃Bo₃). Las reacciones paralelas a un pH de 4, 5 y 6 se realizaron también en presencia de regulador de acetato 50 mM para asegurar que el regulador usado no influyó en los resultados. Para la determinación del perfil de pH de dos alfa-amilasas que dependen de calcio (SEQ ID NO: 56 y SEQ ID NO: 62), se usaron reguladores de ácido málico / acetato / MES en vez del de Britton-Robinson.

45 Preparación de "almidón de maíz soluble" para la reacción con alfa-amilasas. La dextrina (Sigma D2006) no se pudo usar como sustrato en las reacciones de alfa-amilasa BCA debido al medio con alto contenido de extremos reductores. Por lo tanto, se empleó almidón de maíz calentado como sustrato. Específicamente, se disolvió 2% de almidón de maíz en agua desionizada y se calentó mezclando en un baño de agua en ebullición durante 30 - 40 minutos, hasta que el almidón se disolvió y la solución tomó aspecto lechoso, pero traslúcido. La solución de almidón calentado se usó durante 2 días, después de lo cual se observaron algunos signos de degradación (aparición de grumos de almidón) y se descartó la solución.

8. Ensayo de glucosa oxidasa / peroxidasa (GO) para la cuantificación de glucosa liberada durante la hidrólisis de almidón.

Se usó un ensayo de glucosa oxidasa/ peroxidasa (GO) acoplada para determinar la cantidad de glucosa liberada por la glucoamilasa durante la hidrólisis de almidón. Las reacciones de GO comenzaron al agregar 10 µl de la reacción de hidrólisis de almidón neutralizada a 90 µl de PBS que contenía glucosa oxidasa (0,1 U/ml), peroxidasa (0,25 U/ml) y Amplex Red 0,05 mM, en placas Nunc negras de 96 pozos. Las placas se mantuvieron a temperatura ambiente en oscuridad durante 30 min antes de la lectura en un lector de placas fluorescentes con Ex/Em 545/590 nm. Una curva estándar con concentraciones de glucosa de 0 a 100 µM se usó para evaluar la cantidad de glucosa producida en las reacciones de hidrólisis. Las velocidades iniciales de hidrólisis de almidón (nmoles de glucosa liberada de 1% de almidón granulado / min / µg de glucoamilasa) se determinaron al graficar la cantidad de glucosa liberada en el tiempo y calculando la pendiente de la línea de mejor ajuste a través de los puntos de datos.

9. Ensayo de BCA para determinar el incremento en la concentración de extremos reductores durante la hidrólisis de almidón.

Una alícuota de 10 µl de reacción de hidrólisis de almidón de amilasa se neutralizó en 100 µl de reactivo BCA (que consistía de 64 mg/mL de carbonato de sodio monohidrato, 24 mg/mL de bicarbonato de sodio, 1,95 mg/mL de BCA, 1,24 mg/mL de sulfato cúprico pentahidrato, 1,26 mg/mL de L-serina). El desarrollo de color se produjo durante la incubación de la reacción neutralizada a 80 °C durante 35 minutos y fue seguido por la determinación de la absorbancia a 560 nm. Las velocidades iniciales se calcularon durante un periodo de reacción de 40 min. Se construyó una curva estándar usando maltosa (0 - 54 µM) fue construida para correlacionar $A_{560\text{ nm}}$ con la concentración de los azúcares reductores generados (nmoles). La actividad específica se expresó como nmoles/min/µg de enzima.

10. Especificidad del tipo de enlace de glucoamilasas con maltosa e isomaltosa como sustratos

Las reacciones comenzaron al agregar la enzima (concentración final de 5 µg/ml para maltosa y 30 µg/ml para isomaltosa) a la mezcla de reacción. A los 2, 5, 10, 15, 20, 25, 30 y 40 min, se retiraron alícuotas de 5 µl de las reacciones y se neutralizaron mediante la adición a 10 µl del regulador Tris 1 M, pH 7,5. Se usaron 9 concentraciones de sustrato en los estudios, oscilando de 0 a 12 mM para maltosa y 2,5 a 120 mM para isomaltosa. Las reacciones se realizaron por triplicado a 37 °C y pH 5,0 en regulador (NaCH₃CO₂ 50 mM, CaCl₂ 10 mM), a escala de 50 µl en una incubadora de mesa Eppendorf con agitación constante (800 rpm). La producción de glucosa se midió al final de la reacción usando el ensayo de glucosa oxidasa / peroxidasa (GO).

30 Resultados

2. Caracterización de glucoamilasas

1.1 Velocidades iniciales de reacción:

Las velocidades iniciales para hidrólisis de almidón granulado y soluble se presentan en la Tabla 1. Como se puede observar a partir de la Tabla 1, las glucoamilasas de ejemplo de la invención mostraron hasta 3 veces mejor actividad (SEQ ID NO: 48) contra el almidón granulado, con similar o ligeramente mejor actividad en el almidón soluble al compararse con la enzima de punto de referencia de *A. niger*. La SEQ ID NO: 48 también fue expresada en *E. coli* (véase la sección anterior sobre los constructos de expresión de *Picha*) - esta enzima expresada en *E. coli* se etiqueta como SEQ ID NO: 48 (Ec) en la Tabla a continuación. La SEQ ID NO: 48 (Ec) no mostró actividad contra almidón granulado bajo las condiciones evaluadas (probablemente debido a la falta de un dominio de enlace a almidón).

Tabla 1: Comparación de las velocidades iniciales de hidrólisis de almidón de maíz granulado y almidón soluble (dextrina) por amilasas y/o glucoamilasas de ejemplo "principales" y una enzima glucoamilasa de punto de referencia de *A. niger* (Sigma A7095) a 37 °C. pH 5,0:

Tabla 1

Enzima	Velocidad inicial* ± SD del almidón granulado	Velocidad inicial* ± SD del almidón soluble
SEQ ID NO: 48	35,6 ± 3,8	60,7 ± 5,1
SEQ ID NO: 26	28,8 ± 3,4	51,8 ± 7,4

Enzima	Velocidad inicial* \pm SD del almidón granulado	Velocidad inicial* \pm SD del almidón soluble
SEQ ID NO: 74	25,1 \pm 2,5	84,3 \pm 3,5
SEQ ID NO: 18	24,3 \pm 4,3	58,3 \pm 3,0
SEQ ID NO: 28	17,8 \pm 4,3	33,8 \pm 4,0
SEQ ID NO: 14	6,6 \pm 1,2	53,9 \pm 4,1
SEQ ID NO: 48(Ec)	0	59,3 \pm 8,5
Glucoamilas de <i>A. niger</i> (Sigma A7095)	11,3 \pm 2,7	43,3 \pm 7,4
*Las velocidades iniciales se expresan como nmoles de de glucosa/min/ μ g de proteína glucoamilasa libreada del 1% de almidón granular o a partir de 0,5% de dextrina. Cada número es el valor promedio de 6 - 10 puntos de datos.		

1.2 Perfil de temperatura:

El efecto de la temperatura (30 °C - 40 °C) sobre la hidrólisis de almidón por las glucoamilasas caracterizadas se presenta en las Figuras 1A y 1B. Las actividades de las glucoamilasas se incrementaron con la temperatura; fueron más activas a 40 °C pero retuvieron aproximadamente 50% de la actividad pico a 30°C.

Figura 28A: Efecto de la temperatura sobre la actividad de las amilasas y/o glucoamilasas de ejemplo y la glucoamilasa de *A. niger* con almidón granulado como sustrato (glucoamilasa de *A. niger* como punto de referencia (Sigma A7095)). La liberación de glucosa se midió a la temperatura indicada a un pH de 5,0. La SEQ ID NO: 20 de ejemplo o está incluido en el gráfico ya que no mostró ninguna actividad contra almidón granulado bajo estas condiciones particulares de ensayo.

Figura 28B: Efecto de la temperatura sobre la actividad de la glucoamilasa de la SEQ ID NO: 20 de ejemplo y la glucoamilasa de *A. niger* (Sigma A7095) con almidón soluble (Dextrina) como sustrato. La liberación de glucosa se midió a la temperatura indicada a pH de 5,0.

1.3 Perfil de pH:

La influencia del pH sobre la hidrólisis de almidón se evaluó tanto en el almidón granulado como en el soluble y los resultados se presentan en la Figura 2A y en la Figura 2B, respectivamente. Todas las glucoamilasas hidrolizaron mejor ambos sustratos a pH inferior, siendo la SEQ ID NO: 26 de carácter más ácido.

Figura 29A: Efecto del pH sobre la actividad de las glucoamilasas con almidón granulado como sustrato. La liberación de glucosa se midió en el pH indicado a 37 °C. Las velocidades iniciales se calcularon sobre 20 min y se convirtieron a un porcentaje de la velocidad máxima. La SEQ ID NO: 20 de ejemplo no se incluyó en el gráfico porque no mostró actividad contra almidón granulado, bajo las condiciones particulares aquí evaluadas.

Figura 29B: Efecto del pH sobre la actividad de de las glucoamilasas con almidón soluble como sustrato. La liberación de glucosa se midió en el pH indicado a 37 °C. Las velocidades iniciales se calcularon sobre 20 min y se convirtieron a un porcentaje de la velocidad máxima.

1.4 Especificidad de escisión del tipo de enlace:

Los parámetros cinéticos para la hidrólisis de maltosa (enlace α -1,4) (la maltosa es 2 alfa -D-glucosas) e isomaltosa (enlace α -1,6) se determinaron para 7 glucoamilasas seleccionadas y la glucoamilasa de *A. niger* como punto de referencia (Sigma A7095). Los experimentos se realizaron con lisados de *P. pastoris* liofilizados y las proteínas no se purificaron; por lo tanto, únicamente se reportaron en este documento los datos independientes de concentración de proteínas. La Tabla 2, a continuación, resume los valores de K_M para maltosa e isomaltosa y la relación de k_{cat}/K_M o para maltosa comparado con k_{cat}/K_M para isomaltosa. Estos parámetros determinados para glucoamilasa de *A. niger* (Sigma A7095) están en concordancia con los datos publicados (K_M para maltosa se reporta que es 1,2 - 2,1 mM; K_M para isomaltosa se reporta que es 19,8 - 42,0 y k_{cat}/K_M para maltosa sobre k_{cat}/K_M para isomaltosa se

reporta que está entre 300 - 600, de acuerdo con Frandesen y colaboradores. (1995); Sierks y Svensson;1996; Fagerstrom y Kalkkinen; 1995).

5 Como puede observarse a partir de la Tabla 2, el ejemplo de la glucoamilasa de la SEQ ID NO: 20 fue la más fuertemente selectiva para maltosa y tuvo casi 900 veces mayor especificidad hacia los enlaces α -1,4 con relación a los enlaces α -1,6. La glucoamilasa menos selectiva fue la SEQ ID NO: 14 con especificidad \sim 100 veces mayor superior hacia los enlaces α -1,4 con relación a los enlaces α -1,6.

Tabla 2. Parámetros cinéticos para hidrólisis de maltosa e isomaltosa mediante 7 glucoamilasas de ejemplo de esta invención y una como punto de referencia (glucoamilasa de *A. niger* (Sigma A7095)).

Enzima	Maltosa	Isomaltosa	K_{cat}/K_M (maltosa) / K_{cat}/K_M (isomaltosa)
	K_M (mM)	K_M (mM)	
SEQ ID NO: 28	0,61 \pm 0,06	11,94 \pm 4,99	750
SEQ ID NO: 74	1,87 \pm 0,17	11,55 \pm 2,92	481
SEQ ID NO: 20	2,62 \pm 0,19	53,97 \pm 23,17	897
SEQ ID NO: 14	2,67 \pm 0,15	41,5 \pm 5,05	116
SEQ ID NO: 26	0,98 \pm 0,33	12,18 \pm 0,64	456
SEQ ID NO: 48	2,26 \pm 0,12	21,69 \pm 3	565
SEQ ID NO: 18	1,01 \pm 0,09	11,74 \pm 7,74	415
Glucoamilasa <i>A. niger</i> (Sigma A7095)	0,93 \pm 0,1	18,72 \pm 3,95	249
Cada número es un valor promedio de 5 experimentos diferentes			

10 Caracterización de amilasas

Velocidades iniciales de reacción: Las velocidades iniciales de hidrólisis de almidón granulado y almidón soluble se presentan en la Tabla 3. Se compararon ocho alfa-amilasas de ejemplos con una alfa -amilasa de punto de referencia de *A. oryzae*. Como se puede observar a partir de la Tabla 3 a continuación todas las amilasas evaluadas mostraron actividad significativamente superior contra almidón soluble cuando se compara con almidón granulado.

15 Sin embargo, esta diferencia fue menos marcada para las amilasas y/o glucoamilasas que para la enzima como punto de referencia.

Tabla 3: Comparación de las velocidades iniciales de hidrólisis del almidón de maíz de granulado y el almidón de maíz soluble por medio de 8 α -amilasas y una alfa-amilasa como punto de referencia de *A. oryzae* a 37 °C y pH 5:

Tabla 3

Ejemplo de enzima	Velocidad inicial* \pm SD de almidón granulado	Velocidad inicial* \pm SD de almidón soluble
SEQ ID NO: 56	15,7 \pm 1,67	1607,9 \pm 518,22
SEQ ID NO: 70**	20,5	109,1
SEQ ID NO: 62	3,5 \pm 0,37	139,6 \pm 55,96
SEQ ID NO: 66**	2,1	70,8

ES 2 575 912 T3

Ejemplo de enzima	Velocidad inicial* \pm SD de almidón granulado	Velocidad inicial* \pm SD de almidón soluble
SEQ ID NO: 2	7 \pm 0,75	381,2 \pm 74,15
SEQ ID NO: 52	10,3 \pm 1,75	248,2 \pm 28,46
SEQ ID NO: 78	0,4 \pm 0,06	232 \pm 52,63
SEQ ID NO: 76**	25,2	809,1
Amilasa de <i>A. oryzae</i> (MegazymeE-ANAAM)	0,4 \pm 0,07	498,7 \pm 64,78

*Las velocidades iniciales se expresan como nmoles de los extremos reductores liberados a partir de 1% de almidón /min/ μ g de proteína alfa-amilasa.

**Datos obtenidos utilizando enzima purificada.

-Cada número es un valor promedio de 5 puntos de datos.

Perfil de temperatura: El efecto de la temperatura sobre la hidrólisis de almidón por las α -amilasas caracterizadas se presenta en la Figura 30. Las actividades de las amilasas fueron afectadas por la temperatura en un grado diferente. Cinco de las amilasas de ejemplo (SEQ ID NO: 56, SEQ ID NO: 70, SEQ ID NO: 78, SEQ ID NO: 76 y SEQ ID NO: 66) fueron más activas a 40 °C y retuvieron aproximadamente 30% de la actividad a 30 °C. Las actividades de la SEQ ID NO: 2, SEQ ID No: 52 y SEQ ID NO: 62 de ejemplo fueron afectadas sólo marginalmente por los cambios en temperatura en el rango investigado.

La Figura 30 ilustra: la influencia de la temperatura sobre la hidrólisis de almidón por 9 α -amilasas. La actividad se midió a un pH de 5,0 durante 40 min de incubación a la temperatura indicada y las velocidades iniciales fueron calculadas y graficadas contra tiempo. En la Figura 30:

* Las velocidades para las SEQ ID NO: 56, SEQ ID NO: 2 y SEQ ID NO: 52 de ejemplos se presentan en el eje izquierdo; las velocidades para el resto de las enzimas se presentan en el eje derecho.

* Las velocidades iniciales se expresan como nmoles de extremos reductores liberados a partir de 1% de almidón de maíz granulado/min/ μ g de enzima en 250 μ l de reacción.

- Las actividades de las SEQ ID NO: 70, SEQ ID NO: 66 y SEQ ID NO: 76 de ejemplo se expresan en nmoles/min/ μ g de proteína total en sobrenadante de *Pichia pastoris*.

Perfil de pH: La influencia del pH sobre la hidrólisis de almidón se evaluó con sustratos tanto de almidón granulado como de almidón soluble y los resultados se presentan en la Figura 31A y en la Figura 31B, respectivamente. La SEQ ID NO: 52 de ejemplo (de origen fúngico) tuvo el pH menos óptimo (\sim pH 4). Otra enzima de ejemplo (SEQ ID NO: 2) (de origen fúngico) también mostró preferencia por el pH ácido, con un pH óptimo aparente de \sim pH 4,5 - 5. La amilasa de la SEQ ID NO: 66 de ejemplo (de origen arqueal) tuvo un pH óptimo aparente de \sim pH 5,0, reteniendo \sim 70% de actividad pico a un pH de 4,0. Las enzimas restantes tuvieron un pH óptimo aparente entre 5,0 y 6,0 y fueron casi inactivas a un pH de 4,0 y 3,5.

Figura 31A: Efecto del pH sobre las actividades de α -amilasas con almidón granulado como sustrato. El incremento en los extremos reductores se midió al pH indicado a 37 °C. Las velocidades iniciales se calcularon sobre 40 min y se convirtieron a un porcentaje de la velocidad máxima.

Figura 31B: Efecto del pH sobre las actividades de α -amilasas con almidón soluble como sustrato. El incremento en los extremos reductores se midió al pH indicado a 37 °C. Las velocidades iniciales se calcularon sobre 40 min y se convirtieron a un porcentaje de la velocidad máxima.

Listado de secuencias

<110> VERENIUM CORPORATION SYNGENTA PARTICIPATIONS AG

ES 2 575 912 T3

<120> Amilasas y glucoamilasas, ácidos nucleicos que las codifican y métodos para formarlas y utilizarlas

<130> V 8573/MH

<140> 12 152 656.0

<141> 21.12.2007

5 <150> US 60/892,823

<151> 2007-03-02

<150> US 60/877,068

<151> 2006-12-21

<160> 82

10 <170> PatentIn 3.3

<210> 1

<211> 1710

<212> ADN

<213> *Cochliobolus heterostrophus* ATCC 48331

15 <400> 1

ES 2 575 912 T3

gccgacacca atgcttggaa gtcccgcagc atctactttg tcctgacgga tcgtattgcc 60
 cgcaacagca gcgacacggg cggtcagcg tgcagcgacc tcggcaacta ctgCGgtgga 120
 actttccagg gcctcgagtc taagctcgac tacatcaagg gacttggatt cgatgccatt 180
 tggatcacc ccgtcgtctc aaacaaggct gctggatacc atggctactg ggccgaggac 240
 ttgtatgccg tcaactcaaa ctacggcact gctgcccact tgaagagctt ggttgccgct 300
 gcccatgcc aaggcatcta catgatggtc gacgttgtcg caaaccacat gggcctgga 360
 gcaatcacia acaaccgccc tgaacctctc aaccaggctt catcatacca ccctccttgc 420
 aacatcgact acaacaacca aaccagtgtc gaggtatgtc aaatagccgg actccccgac 480
 atctacacca ccaagagcga gatccgcagc ctctcaaca cctgggtcaa ctggctcgta 540
 aacgagtaca gcttcgacgg tgtccgcacg gacaccgtca agcacgtcga aaaggacttt 600
 tggcctggct tctctgcccg taccggtgtc tacaacattg gcgagggtgt tgacggagac 660
 ccagcctacc ttgccccgta cgccaagctt atgcccggcc tcctcaacta cgagctctac 720
 taccgatga acaactttta ccagcaaagc ggctcttccc aggcgcttgt agacatgatg 780
 aacactgtca gcaaacctt ccctgaccca tctgccttgg gaaccttctt cgacaaccac 840
 gacaacaagc gctgggtgaa cgtaagaac gaccagactc tgctcaagaa cgctcttgtc 900
 tatgtcatcc tcgcacgtgg tatccccatc ttgtactatg gtaccgagca gggatagct 960
 ggtggtgacg acccagctaa ccgagaggat ctgtggcgca gtggcttcaa caccaatgcc 1020
 aacctctacc aagccatcaa gaaactgacc gccgcccgac aggctgcccg tggctctgca 1080
 ggaaacgacc acgtccacct gtacgtcgcc gacacggctt acgcctggag ccgtgccaac 1140
 ggcaacctga ttgtcctcac caccaacgct ggcggaact ccaacacca gcactgcttc 1200

 aacacgcaaa aggcaaacgg ccgctggacc aacgtctacg gcaacggcgc caccgtctct 1260
 gccgatagca acggccaaat ctgCGtctcc gtcacaaacg gcgagcccgt tgcctcctc 1320
 gccggctccg ctacccccac cactggcact accctctcca cccgaccgc cactgccacc 1380
 gccacaccaa ccgcatgcc caccgcccgc tccgtctctt taccaccgg cgtcaccact 1440
 gttcccgtg acaccatcaa aatcactggc aacacggccc agctaggtaa ctggactccc 1500
 gccaacggtc ttgccttgtc cgcagctagc tacacatcca gcaaccctat ctggaccatt 1560
 accgtgcccc tggccgctgg atcctccatc tcgtacaagt ttgtcaagat tgacagtgga 1620
 ggaactgtca cctgggagag tgaccccaac aggtcataca ctgCGccgag ctgccaggcg 1680
 agtgccggtg tgaacagctc atggcaatag 1710

<210> 2

<211> 569

<212> PRT

5 <213> Cochliobolus heterostrophus ATCC 48331

<220>

<221> DOMINIO

<222> (13)...(353)

<223> Alfa amilasa, dominio catalítico

<220>

5 <221> DOMINIO

<222> (469)...(564)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

10 <222> (22)...(25)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (148)...(151)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (558)...(561)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

20 <220>

<221> SITIO

<222> (573)...(576)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 2

ES 2 575 912 T3

Ala Asp Thr Asn Ala Trp Lys Ser Arg Ser Ile Tyr Phe Val Leu Thr
1 5 10 15

Asp Arg Ile Ala Arg Asn Ser Ser Asp Thr Gly Gly Ser Ala Cys Ser
20 25 30

Asp Leu Gly Asn Tyr Cys Gly Gly Thr Phe Gln Gly Leu Glu Ser Lys
35 40 45

Leu Asp Tyr Ile Lys Gly Leu Gly Phe Asp Ala Ile Trp Ile Thr Pro
50 55 60

ES 2 575 912 T3

Val Val Ser Asn Lys Ala Ala Gly Tyr His Gly Tyr Trp Ala Glu Asp
 65 70 75 80
 Leu Tyr Ala Val Asn Ser Asn Tyr Gly Thr Ala Ala Asp Leu Lys Ser
 85 90 95
 Leu Val Ala Ala Ala His Ala Lys Gly Ile Tyr Met Met Val Asp Val
 100 105 110
 Val Ala Asn His Met Gly Pro Gly Ala Ile Thr Asn Asn Arg Pro Glu
 115 120 125
 Pro Leu Asn Gln Ala Ser Ser Tyr His Pro Pro Cys Asn Ile Asp Tyr
 130 135 140
 Asn Asn Gln Thr Ser Val Glu Val Cys Gln Ile Ala Gly Leu Pro Asp
 145 150 155 160
 Ile Tyr Thr Thr Lys Ser Glu Ile Arg Thr Leu Leu Asn Thr Trp Val
 165 170 175
 Asn Trp Leu Val Asn Glu Tyr Ser Phe Asp Gly Val Arg Ile Asp Thr
 180 185 190
 Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Phe Ser Ala Ala Thr
 195 200 205
 Gly Val Tyr Asn Ile Gly Glu Val Phe Asp Gly Asp Pro Ala Tyr Leu
 210 215 220
 Ala Pro Tyr Ala Lys Leu Met Pro Gly Leu Leu Asn Tyr Ala Val Tyr
 225 230 235 240
 Tyr Pro Met Asn Asn Phe Tyr Gln Gln Thr Gly Ser Ser Gln Ala Leu
 245 250 255
 Val Asp Met Met Asn Thr Val Ser Asn Thr Phe Pro Asp Pro Ser Ala
 260 265 270
 Leu Gly Thr Phe Leu Asp Asn His Asp Asn Lys Arg Trp Leu Asn Val
 275 280 285
 Lys Asn Asp Gln Thr Leu Leu Lys Asn Ala Leu Ala Tyr Val Ile Leu
 290 295 300
 Ala Arg Gly Ile Pro Ile Leu Tyr Tyr Gly Thr Glu Gln Gly Tyr Ala
 305 310 315 320
 Gly Gly Asp Asp Pro Ala Asn Arg Glu Asp Leu Trp Arg Ser Gly Phe
 325 330 335

ES 2 575 912 T3

Asn Thr Asn Ala Asn Leu Tyr Gln Ala Ile Lys Lys Leu Thr Ala Ala
 340 345 350

Arg Gln Ala Ala Gly Gly Leu Ala Gly Asn Asp His Val His Leu Tyr
 355 360 365

Val Ala Asp Thr Ala Tyr Ala Trp Ser Arg Ala Asn Gly Asn Leu Ile
 370 375 380

Val Leu Thr Thr Asn Ala Gly Gly Asn Ser Asn Thr Gln His Cys Phe
 385 390 395 400

Asn Thr Gln Lys Ala Asn Gly Arg Trp Thr Asn Val Tyr Gly Asn Gly
 405 410 415

Ala Thr Val Ser Ala Asp Ser Asn Gly Gln Ile Cys Val Ser Val Thr
 420 425 430

Asn Gly Glu Pro Val Val Leu Leu Ala Gly Ser Ala Thr Pro Thr Thr
 435 440 445

Gly Thr Thr Leu Ser Thr Arg Thr Ala Thr Ala Thr Ala Thr Pro Thr
 450 455 460

Ala Cys Pro Thr Ala Val Ser Val Ser Phe Thr His Arg Val Thr Thr
 465 470 475 480

Val Pro Gly Asp Thr Ile Lys Ile Thr Gly Asn Thr Ala Gln Leu Gly
 485 490 495

Asn Trp Thr Pro Ala Asn Gly Leu Ala Leu Ser Ala Ala Ser Tyr Thr
 500 505 510

Ser Ser Asn Pro Ile Trp Thr Ile Thr Val Pro Leu Ala Ala Gly Ser
 515 520 525

Ser Ile Ser Tyr Lys Phe Val Lys Ile Asp Ser Gly Gly Thr Val Thr
 530 535 540

Trp Glu Ser Asp Pro Asn Arg Ser Tyr Thr Ala Pro Ser Cys Gln Ala
 545 550 555 560

Ser Ala Gly Val Asn Ser Ser Trp Gln
 565

<210> 3

<211> 1770

<212> ADN

5 <213> Cochliobolus heterostrophus ATCC 48331

ES 2 575 912 T3

<400> 3

atgttggtgc	tcaacatctt	caccaccctc	ttcttctaca	tcacctgcat	cgtctccgcc	60
gccgacacca	atgcttggaa	gtcccgcagc	atctactttg	tcctgacgga	tcgtattgcc	120
cgcaacagca	gcgacacggg	cggctcagcg	tgcagcgacc	tcggcaacta	ctgcggtgga	180
actttccagg	gcctcgagtc	taagctcgac	tacatcaagg	gacttggatt	cgatgccatt	240
tggattaccc	ccgtcgtctc	aaacaaggct	gctggatacc	atggctactg	ggccgaggac	300
ttgtatgccg	tcaactcaaa	ctacggcact	gctgccgact	tgaagagctt	ggttgccgct	360
gcccattgcca	agggcatcta	catgatggtc	gacgttgtcg	caaaccacat	gggtcctgga	420
gcaatcacia	acaaccgccc	tgaacctctc	aaccaggctt	catcatacca	ccctccttgc	480
aacatcgact	acaacaacca	aaccagtgtc	gaggtatgtc	aatagccgg	actccccgac	540
atctacacca	ccaagagcga	gatccgcacg	ctcctcaaca	cctgggtcaa	ctggctcgta	600
aacgagtaca	gcttcgacgg	tgtccgcata	gacaccgtca	agcacgtcga	aaaggacttt	660
tggcctggct	tctctgccgc	taccggtgtc	tacaacattg	gcgaggtggt	tgacggagac	720
ccagcctacc	ttgccccgta	cgccaagctt	atgccccgcc	tcctcaacta	cgcagtctac	780
taccgatga	acaactttta	ccagcaaacg	ggctcttccc	aggcgcttgt	agacatgatg	840
aacactgtca	gcaacacctt	ccctgacca	tctgccttgg	gaaccttctt	cgacaaccac	900
gacaacaagc	gctggttgaa	cgtcaagaac	gaccagactc	tgctcaagaa	cgctcttgct	960
tatgtcatcc	tcgcacgtgg	tatccccatc	ttgtactatg	gtaccgagca	gggatacgtt	1020
ggtgggtgacg	accagctaa	ccgagaggat	ctgtggcgca	gtggcttcaa	caccaatgcc	1080
aacctctacc	aagccatcaa	gaaactgacc	gccgccccgac	aggctgccgg	tggctctgca	1140
ggaaacgacc	acgtccacct	gtacgtcgcc	gacacggctt	acgcctggag	ccgtgccaac	1200
ggcaacctga	ttgtcctcac	caccaacgct	ggcggcaact	ccaacacca	gcactgcttc	1260
aacacgcaaa	aggcaaacgg	ccgctggacc	aacgtctacg	gcaacggcgc	caccgtctct	1320
gccgatagca	acggccaaat	ctgcgtctcc	gtcaciaaacg	gcgagcccgt	tgctcctctc	1380
gccggctccg	ctacccccac	cactggcact	accctctcca	cccgcaccgc	cactgccacc	1440
gccacacca	ccgcatgccc	caccgccgtc	tccgtctcct	tcaccaccg	cgtcaccact	1500
gttcccgggtg	acaccatcaa	aatcactggc	aacacggccc	agctaggtaa	ctggactccc	1560
gccaacggcc	ttgccttgtc	cgcagctagc	tacacatcca	gcaaccctat	ctggaccatt	1620
accgtgcccc	tggccgctgg	atcctccatc	tcgtacaagt	ttgtcaagat	tgacagtgga	1680
ggaactgtca	cctgggagag	tgaccccaac	aggtcataca	ctgcgccgag	ctgccaggcg	1740
agtgccagtg	tgaacagctc	atggcaatag				1770

<210> 4

5 <211> 589

ES 2 575 912 T3

<212> PRT

<213> Cochliobolus heterostrophus ATCC 48331

<220>

<221> SEÑAL

5 <222> (1)...(20)

<220>

<221> DOMINIO

<222> (33)...(373)

<223> Alfa amilasa, dominio catalítico

10 <220>

<221> DOMINIO

<222> (489)...(586)

<223> dominio de enlazamiento de almidón

<220>

15 <221> SITIO

<222> (42)...(45)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

20 <222> (168)...(171)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (578)...(581)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

ES 2 575 912 T3

<220>

<221> SITIO

<222> (593)...(596)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

5 <400> 4

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

ES 2 575 912 T3

Asn Ile Asp Tyr Asn 165 Asn Gln Thr Ser Val 170 Glu Val Cys Gln Ile 175 Ala
 Gly Leu Pro Asp 180 Ile Tyr Thr Thr Lys 185 Ser Glu Ile Arg Thr 190 Leu Leu
 Asn Thr Trp 195 Val Asn Trp Leu Val 200 Asn Glu Tyr Ser Phe 205 Asp Gly Val
 Arg Ile 210 Asp Thr Val Lys His 215 Val Glu Lys Asp Phe 220 Trp Pro Gly Phe
 Ser Ala Ala Thr Gly Val 230 Tyr Asn Ile Gly Glu 235 Val Phe Asp Gly Asp 240
 Pro Ala Tyr Leu Ala 245 Pro Tyr Ala Lys Leu 250 Met Pro Gly Leu Leu 255 Asn
 Tyr Ala Val Tyr 260 Tyr Pro Met Asn Asn 265 Phe Tyr Gln Gln Thr 270 Gly Ser
 Ser Gln Ala 275 Leu Val Asp Met Met 280 Asn Thr Val Ser Asn 285 Thr Phe Pro
 Asp Pro 290 Ser Ala Leu Gly Thr 295 Phe Leu Asp Asn His 300 Asp Asn Lys Arg
 Trp Leu Asn Val Lys Asn 310 Asp Gln Thr Leu Leu 315 Lys Asn Ala Leu Ala 320
 Tyr Val Ile Leu Ala 325 Arg Gly Ile Pro Ile 330 Leu Tyr Tyr Gly Thr 335 Glu
 Gln Gly Tyr Ala 340 Gly Gly Asp Asp Pro Ala Asn Arg Glu Asp 350 Leu Trp
 Arg Ser Gly 355 Phe Asn Thr Asn Ala 360 Asn Leu Tyr Gln Ala 365 Ile Lys Lys
 Leu Thr Ala Ala Arg Gln Ala 375 Ala Gly Gly Leu Ala 380 Gly Asn Asp His
 Val His Leu Tyr Val Ala 390 Asp Thr Ala Tyr Ala 395 Trp Ser Arg Ala Asn 400
 Gly Asn Leu Ile Val 405 Leu Thr Thr Asn Ala 410 Gly Gly Asn Ser Asn 415 Thr
 Gln His Cys Phe 420 Asn Thr Gln Lys Ala 425 Asn Gly Arg Trp Thr 430 Asn Val
 Tyr Gly Asn Gly Ala Thr Val Ser Ala Asp Ser Asn Gly Gln Ile Cys

ES 2 575 912 T3

gtgaagatct acgacgcggc tgaacaggtc tttcagatcc aggaatccgt ctggcctcgc	60
cctgctgatg acgagggcac cgaccctgaa aagtcggctt tgaccttcac ttggaccgat	120
agcccgtttt cttttgccat taaacgtagg gccaccaatg agacgctctt cgatacttcg	180
gcagcttctc ttgttttcga gacgcagtac cttcgcttaa gaaccgctct accacctttg	240
ccaaacctgt acggtcttgg tgaatcaacg gatgctttcc atctcaacac caccaactat	300
acgcgaactc tttggaatcg agatgcctat ggcacgccac caggatccaa cctctacgga	360
gctcacccga tctactttga tcaccgcggt gagaatggta ctcacgggtg tttcttggct	420
agctctgagg gtatggacat caagattgat gatactgacg gccagttcct tgagtacaac	480
actctcggtg gtgttcttga cttctacttc cttgctggac ccggccctaa ggaagttgca	540
actcaatact cagccctttc tggcttgcct gctatgatgc cttactgggg tttcggctca	600
caccaatgca agtatggcta ccgtgatgtt tgggaggttg ccgaggttgt agcaaaactac	660
tctgctgctg atatcccact tgagaccatg tggaccgata tcgactacat ggagcttcgc	720

ES 2 575 912 T3

cgcttggtca ccctggatcc cgagcgttac ccgcttgagc ttgttcgcca gcttgtagac 780
 tatttgcatt ctcaccaaca gcaactacatt ctgatgggtca actccgctgt ctggagcggg 840
 gactatgatg cctacaacaa cgggtgcaaag cttgaagtct tccagaagaa gagcaacggc 900
 tccttcgaac aggggtgctgt ctggccaggc cctactgtct tccagattg gttccatccc 960
 aacaccaga aatactggga tgaggaattc gctcgcttct tcgaccccg c tactgggtgtt 1020
 gatatcgatg gactttggaa tgacatgaat gaaccgcga acttctgccc ataccctgc 1080
 tcagaccctg aagcctattc tgaggagtcc aagaaccac ccgagccacc ggcggtgctc 1140
 acttctgcag gtcgcaaat ccctggcttc ccagcagggt tccagccaca gtctaactct 1200
 agcactgcaa ggcgttctgt tgtaaagga ccatctagca tgcgtccttc caagcgcaa 1260
 ggcgctaaca gcgctggaga tgctaagcac ctcggtcttc ccggtcgtga tctgatcaac 1320
 cccaagtacc agatccacaa cgaagccggt tcaatcagca acaggaccct ggctacggat 1380
 atcaagaatt acgatggctc ttatcactac gatacgcaca acttctgggg ctcgatgatg 1440
 agcattacct ctcacaagtc tatgcaagct cgccgtccc aaagacggcc attcattatc 1500
 actaggatcat cttccctgg cctcggttct tatctcgaa agtggcttgg tgacaacgtc 1560
 tccgagtggg cacaataccg cttctcgatt gccggcatct tgaacttcaa caccatcttc 1620
 cagatcccca tggctcggctc agatatttgc ggtttcggcc gaaacacgac cgagactctc 1680
 tgcgcccgt ggaccactct tgggtgcttc taccggttca tgaggaacca cgccggcgac 1740
 acttccatca gccaaagaata ctatcgctgg cctctcacca gggccgcagc caagaacgcc 1800
 atcgcagtca ggtacaggct cttggactac ttctacacgg ccttccaccg ccaggccacc 1860
 accggtctac ccagcttga cccctcttc ttccactacc ccaccgacgc caaaccttc 1920
 ggcattgagc accagttctt ctacggagac agcatcctcg tctcgcccgt cctcgaagaa 1980
 aactccacct cagtctccat ctacctccc aaagatgtct tctacgacta ctggaccggc 2040
 gagcgcattc aaggaaacgg cgagaacatt aacctcactg acgtaggatt cgacaccatc 2100
 cccctccag tcaaaggtgg atccatcctc cctctccg cgaatccgc aaacacaacc 2160
 accgagctcc gcaaacaaaa ctttgcctc tggatcgac caaatgctac caaccaagcc 2220
 tctggctcgc tctacctga tgatggagat tccctcgagc agaagtctac ttcgctcatt 2280
 aacttctct tcaacaacgg cgccttcagc atgagcggcg atttcggatt cgagactgag 2340
 cttgtcattc agaatatcac catcctgggt acctcgcaga gcgtacaggg ccctggtgcg 2400
 cttaccaagg gctgggaaca taactttggg gctggtgctg gtatgccgca gtttgatagt 2460
 ggtgcagccg agtcgcgtcg agttgttct gcagtttggg gtttggttc gggggctgtc 2520
 agtgtttggg tcagcttgta a 2541

<210> 6

<211> 846

<212> PRT

5 <213> Cochliobolus heterostrophus ATCC 48331

<220>

<221> DOMINIO

<222> (169)...(710)

<223> Familia de Glicosil hidrolasas 31

<220>

<221> SITIO

5 <222> (53)...(56)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (97)...(100)

10 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (134)...(137)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

15 <220>

<221> SITIO

<222> (222)...(225)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

20 <221> SITIO

<222> (303)...(306)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

25 <222> (349)...(356)

ES 2 575 912 T3

<223> Sitio activo de la familia de glicosil hidrolasas 31. Id del prositio = PS00129

<220>

<221> SITIO

<222> (405)...(408)

5 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (554)...(584)

<223> Firma 2 de la familia de Glicosil hidrolasas 31. Id del prositio = PS00707

10 <220>

<221> SITIO

<222> (563)...(566)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (671)...(674)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

20 <222> (701)...(704)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (729)...(732)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (746)...(749)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

5 <220>

<221> SITIO

<222> (772)...(775)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

10 <221> SITIO

<222> (797)...(800)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 6

ES 2 575 912 T3

Met Lys Ile Tyr Asp Ala Ala Glu Gln Val Phe Gln Ile Gln Glu Ser
1 5 10 15

Val Trp Pro Arg Pro Ala Asp Asp Glu Gly Thr Asp Pro Glu Lys Ser
20 25 30

Ala Leu Thr Phe Thr Trp Thr Asp Ser Pro Phe Ser Phe Ala Ile Lys
35 40 45

Arg Arg Ala Thr Asn Glu Thr Leu Phe Asp Thr Ser Ala Ala Ser Leu
50 55 60

Val Phe Glu Thr Gln Tyr Leu Arg Leu Arg Thr Ala Leu Pro Pro Leu
65 70 75 80

Pro Asn Leu Tyr Gly Leu Gly Glu Ser Thr Asp Ala Phe His Leu Asn
85 90 95

Thr Thr Asn Tyr Thr Arg Thr Leu Trp Asn Arg Asp Ala Tyr Gly Thr
100 105 110

Pro Pro Gly Ser Asn Leu Tyr Gly Ala His Pro Ile Tyr Phe Asp His
115 120 125

Arg Gly Glu Asn Gly Thr His Gly Val Phe Leu Ala Ser Ser Glu Gly
130 135 140

Met Asp Ile Lys Ile Asp Asp Thr Asp Gly Gln Phe Leu Glu Tyr Asn
145 150 155 160

Thr Leu Gly Gly Val Leu Asp Phe Tyr Phe Leu Ala Gly Pro Gly Pro
165 170 175

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

Met Pro Tyr Trp Gly Phe Gly Ser His Gln Cys Lys Tyr Gly Tyr Arg
195 200 205

Asp Val Trp Glu Val Ala Glu Val Val Ala Asn Tyr Ser Ala Ala Asp
210 215 220

Ile Pro Leu Glu Thr Met Trp Thr Asp Ile Asp Tyr Met Glu Leu Arg
225 230 235 240

ES 2 575 912 T3

Arg Leu Phe Thr Leu Asp Pro Glu Arg Tyr Pro Leu Glu Leu Val Arg
 245 250 255
 Gln Leu Val Asp Tyr Leu His Ala His Gln Gln His Tyr Ile Leu Met
 260 265 270
 Val Asn Ser Ala Val Trp Ser Gly Asp Tyr Asp Ala Tyr Asn Asn Gly
 275 280 285
 Ala Lys Leu Glu Val Phe Gln Lys Lys Ser Asn Gly Ser Phe Glu Gln
 290 295 300
 Gly Ala Val Trp Pro Gly Pro Thr Val Phe Pro Asp Trp Phe His Pro
 305 310 315 320
 Asn Thr Gln Lys Tyr Trp Asp Glu Glu Phe Ala Arg Phe Phe Asp Pro
 325 330 335
 Ala Thr Gly Val Asp Ile Asp Gly Leu Trp Asn Asp Met Asn Glu Pro
 340 345 350
 Ala Asn Phe Cys Pro Tyr Pro Cys Ser Asp Pro Glu Ala Tyr Ser Glu
 355 360 365
 Glu Ser Lys Asn Pro Pro Glu Pro Pro Ala Val Arg Thr Ser Ala Gly
 370 375 380
 Arg Gln Ile Pro Gly Phe Pro Ala Gly Phe Gln Pro Gln Ser Asn Ser
 385 390 395 400
 Ser Thr Ala Arg Arg Ser Val Val Lys Gly Pro Ser Ser Met Arg Pro
 405 410 415
 Ser Lys Arg Gln Ala Pro Asn Ser Ala Gly Asp Ala Lys His Leu Gly
 420 425 430
 Leu Pro Gly Arg Asp Leu Ile Asn Pro Lys Tyr Gln Ile His Asn Glu
 435 440 445
 Ala Gly Ser Ile Ser Asn Arg Thr Leu Ala Thr Asp Ile Lys Asn Tyr
 450 455 460
 Asp Gly Ser Tyr His Tyr Asp Thr His Asn Phe Trp Gly Ser Met Met
 465 470 475 480
 Ser Ile Thr Ser His Lys Ser Met Gln Ala Arg Arg Pro Glu Arg Arg
 485 490 495
 Pro Phe Ile Ile Thr Arg Ser Ser Phe Pro Gly Leu Gly Ser Tyr Leu
 500 505 510

ES 2 575 912 T3

Gly Lys Trp Leu Gly Asp Asn Val Ser Glu Trp Ala Gln Tyr Arg Phe
 515 520 525
 Ser Ile Ala Gly Ile Leu Asn Phe Asn Thr Ile Phe Gln Ile Pro Met
 530 535 540
 Val Gly Pro Asp Ile Cys Gly Phe Ala Gly Asn Thr Thr Glu Thr Leu
 545 550 555 560
 Cys Ala Arg Trp Thr Thr Leu Gly Ala Phe Tyr Pro Phe Met Arg Asn
 565 570 575
 His Ala Gly Asp Thr Ser Ile Ser Gln Glu Tyr Tyr Arg Trp Pro Leu
 580 585 590
 Thr Arg Ala Ala Ala Lys Asn Ala Ile Ala Val Arg Tyr Arg Leu Leu
 595 600 605
 Asp Tyr Phe Tyr Thr Ala Phe His Arg Gln Ala Thr Thr Gly Leu Pro
 610 615 620
 Ser Leu Asn Pro Leu Phe Phe His Tyr Pro Thr Asp Ala Lys Thr Phe
 625 630 635 640
 Gly Ile Glu His Gln Phe Phe Tyr Gly Asp Ser Ile Leu Val Ser Pro
 645 650 655
 Val Leu Glu Glu Asn Ser Thr Ser Val Ser Ile Tyr Leu Pro Lys Asp
 660 665 670
 Val Phe Tyr Asp Tyr Trp Thr Gly Glu Arg Ile Gln Gly Asn Gly Glu
 675 680 685
 Asn Ile Asn Leu Thr Asp Val Gly Phe Asp Thr Ile Pro Leu His Val
 690 695 700
 Lys Gly Gly Ser Ile Leu Pro Leu Arg Ala Glu Ser Ala Asn Thr Thr
 705 710 715 720
 Thr Glu Leu Arg Lys Gln Asn Phe Val Leu Trp Ile Ala Pro Asn Ala
 725 730 735
 Thr Asn Gln Ala Ser Gly Ser Leu Tyr Leu Asp Asp Gly Asp Ser Leu
 740 745 750
 Glu Gln Lys Ser Thr Ser Leu Ile Asn Phe Ser Phe Asn Asn Gly Ala
 755 760 765
 Phe Ser Met Ser Gly Asp Phe Gly Phe Glu Thr Glu Leu Val Ile Gln
 770 775 780
 Asn Ile Thr Ile Leu Gly Thr Ser Gln Ser Val Gln Gly Pro Val Ala

ES 2 575 912 T3

```

atgctccctc gcgtgttcct cctcacaggc ttcgtccacc atgcctacac attatcaata    60
cctcgtcttc cctccctcgc cacttttgct gcttccagcc gtgtgcagca acagcaaccg    120
cttcaagaca cacttgatgc ctggataaag catgaagagc gtattgcact cgataaactg    180
ctcgccaaca tcgcgcttgg aggtagcaat gtccaaggaa aagggtgtggc tgagggcacc    240
gtcatcgcca gccctagtca agacgggccg gactactggt tccaatgggt ccgtgacgct    300
gctatcacca tggataccct cgtaacatc tatgccgatg acccctcgtc ctcgctgctg    360
tcttctctat ccaccatcct agacgcatac acctccctcc aagggtgacat tcagcgcact    420
tcaaaccctg ctggcacatt tgacgacctt tccggactag gtgagcccaa gttccaagtc    480
gatggcaagc catttaccgg ctcgtgggga cgacctcagc gcgatgggcc ggcccttcgc    540
gcactgacgc ttatgcatta tctccgagag tacaatgcat cccatccctc actatggagc    600
tctcccaact cgggaagactt ttttggtctg ttctacaccg ctgaaatgcc ccctcgtagc    660
atcatcaaag cagatctcga gtatgtcagc catttctgga accagtcagg gttcgatctc    720
tgggaagagg tcgagggcct gcacttcttc accctcatgg tcagcgcgag gagtctgagg    780
gaaggcagtc acctggcaag agtctttgga gacgtcgggt cggcggattg gtaccaaag    840
caagctggct acattgagaa cctgttgggc aaattctgga atgcgcaaaa aggacatctt    900
gttgagacgc tttggagcaa gaggagtggc ctcgactgct gactgcttct cggctctttg    960
cacgcccttc ctaagagcgg atctgaagat gatgatgttg tgtaccctcc gtggtctgac   1020
gagattcttg tttctctgct tgctttgact cgagaccagc gcgaccgttt ccccatcaac   1080
agcaatcctt ctggccagga tgacgacgat gacgacgtag acgagtccac attccagggg   1140
acaggccttg gtcgttatcc cgaggatgta tacgacggct atggaaactc gaaccgtggg   1200
ggaaaccctt ggttcctttg cacctcttct gccgccgaga ttctttaccg caccgcctcg   1260
cacatctctg caactggcaa cttgactctt accgatgttg gccttccctt ttacgagtct   1320
ctcctcggta gtcctcctt agatgtcagc gttggcacgt ttggcccccac agacgccctt   1380
ttcactccg tcatcgagcg ccttcagagc accggcgacg agtttctgca agtcgtcaag   1440

acgcatgtgg atgccgaggg cagcatgagt gagcagtttg accgtgtaac tgggtacatg   1500
cgtgggtgcc agcatttgac ttggagttat ggtgcctttt tgcaggctgc tagagcaaga   1560
aagatttctg tg                                               1572

```

<210> 8

<211> 524

<212> PRT

5 <213> Cochliobolus heterostrophus ATCC 48331

<220>

<221> DOMINIO

<222> (54)...(518)

ES 2 575 912 T3

<223> Familia de glicosil hidrolasas 15

<220>

<221> SITIO

<222> (194)...(197)

5 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (237)...(240)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <220>

<221> SITIO

<222> (239)...(249)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<220>

15 <221> SITIO

<222> (433)...(436)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 8

ES 2 575 912 T3

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

ES 2 575 912 T3

Ala Tyr Thr Ser Leu Gln Gly Asp Ile Gln Arg Thr Ser Asn Pro Ser
130 135 140

Gly Thr Phe Asp Asp Leu Ser Gly Leu Gly Glu Pro Lys Phe Gln Val
145 150 155 160

Asp Gly Lys Pro Phe Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly
165 170 175

Pro Ala Leu Arg Ala Leu Thr Leu Met His Tyr Leu Arg Glu Tyr Asn
180 185 190

Ala Ser His Pro Ser Leu Trp Ser Ser Pro Asn Ser Glu Asp Phe Phe
195 200 205

Gly Ser Phe Tyr Thr Ala Glu Met Pro Pro Arg Ser Ile Ile Lys Ala
210 215 220

Asp Leu Glu Tyr Val Ser His Phe Trp Asn Gln Ser Gly Phe Asp Leu
225 230 240

Trp Glu Glu Val Glu Gly Leu His Phe Phe Thr Leu Met Val Ser Ala
245 250 255

Arg Ser Leu Arg Glu Gly Ser His Leu Ala Arg Val Phe Gly Asp Val
260 265 270

Gly Ala Ala Asp Trp Tyr Gln Lys Gln Ala Gly Tyr Ile Glu Asn Leu
275 280 285

Leu Gly Lys Phe Trp Asn Ala Gln Lys Gly His Leu Val Glu Thr Leu
290 295 300

Trp Ser Lys Arg Ser Gly Leu Asp Cys Gly Leu Leu Leu Gly Ser Leu
305 310 315 320

His Ala Leu Pro Lys Ser Gly Ser Glu Asp Asp Asp Val Val Tyr Pro
325 330 335

Pro Trp Ser Asp Glu Ile Leu Val Ser Leu Leu Ala Leu Thr Arg Asp
340 345 350

Gln Arg Asp Arg Phe Pro Ile Asn Ser Asn Pro Ser Gly Gln Asp Asp
355 360 365

Asp Asp Asp Asp Val Asp Glu Ser Thr Phe Gln Gly Thr Gly Leu Gly
370 375 380

Arg Tyr Pro Glu Asp Val Tyr Asp Gly Tyr Gly Asn Ser Asn Arg Gly
385 390 395 400

Gly Asn Pro Trp Phe Leu Cys Thr Ser Ser Ala Ala Glu Ile Leu Tyr

ES 2 575 912 T3

405 410 415
 Arg Thr Ala Ser His Ile Ser Ala Thr Gly Asn Leu Thr Leu Thr Asp
 420 425 430
 Val Gly Leu Pro Phe Tyr Glu Ser Leu Leu Gly Ser Ser Ser Leu Asp
 435 440 445
 Val Asp Val Gly Thr Phe Gly Pro Thr Asp Ala Leu Phe His Ser Val
 450 455 460
 Ile Glu Arg Leu Gln Ser Thr Gly Asp Glu Phe Leu Gln Val Val Lys
 465 470 475
 Thr His Val Asp Ala Glu Gly Ser Met Ser Glu Gln Phe Asp Arg Val
 485 490 495
 Thr Gly Tyr Met Arg Gly Ala Gln His Leu Thr Trp Ser Tyr Gly Ala
 500 505 510
 Phe Leu Gln Ala Ala Arg Ala Arg Lys Ile Ser Val
 515 520

<210> 9

<211> 1467

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 9

ES 2 575 912 T3

```

atgttgacat tgaatgtttt gacggcactg ttggcacctg gtgtgctatc atctgccttg      60
cctgctcgcg atctcaatgt gcgtgccgat tcaaagcctc ctgctcttca agccatcctg      120
aataacattg ggcgccgatgg atctgctggt tcgggagcat ctgccggcgt tgtggtagca      180
tccccatcta aatcggatcc cgattacttt tacacctgga cgcgatgatgc agccttgaca      240
tacaaggtct tgattgatga gctcattgcc ggagacacct cgctggaatc taccatccag      300
gactacatct ctgctcaagc aaagctgcag gccgtgtcta acccatctgg tgacttgctc      360
gatggatcag gtcttgacaga gcccaagtac cacgtggact tgactgcctt cacggaggcc      420
tggggtcgtc ctcagcgcga tggacccgct ctgcgagcga cagctctgat tacctatggt      480
aactatattga tttcgaagga gaagacatct gtcgtcaagt cgaacatctg gccattgtg      540
caaaatgatc tgaactatgt tgcgcaatac tggaaacaaa ccggattcga tctgtgggag      600
gagggtccagg gctcctcatt cttcaccatt gctgcacaac accgcgcatt ggtggaaggg      660
agcgcatttg ccaaggcgtt gggagagtcc tgcgagggat gcgattccca ggcacctcag      720
gtcttatgct tccagcaatc tttctggaat ggcaaggctg ttgtttccaa ctttgccaac      780
aatggtcgaa ccggtcttga tgccaattct gtacttactt cgattgtaa ctttgacccc      840
aaagctccat gcgatgacgt taccttccag ccttgctctg ctcgtgccct gtcgaaccac      900

aagctatacg ttgactcgtt ccgtaagatc tatcccgtga acagcggcaa ggaagctgga      960
actgccgttg ctgttggacg ttatgccgag gacacttaca tgggcggcaa cccatgggtac     1020
ttgaccacc c tggccgcggc cgagcagttg tatgatgctc tttaccagtg gaagcaactg     1080
ggatcgctgg aatcaccga ggccagtctt cttttcttca aggatcttgt ttcgtctgcc     1140
gccgctggga agtacccctag ctctcggaa acttacacgt ccatcactgc cgcagtcaag     1200
aaatatgccg atggatttat ggctgttgtt aaggagcaca cgcctagcga tggatctcta     1260
tccgaacaat tcaactcggga caacggcagc ccggcctctg ctaaggactt gacctggctc     1320
tacgcggcgt tgctgtctgc tactcgacgt gaggctggaa cagtgcctcc tagctggggc     1380
gcgtcgactg ccaataaagt gccctcaaag tgtgagggaa gctcggccaa gggaaagctac     1440
acgactccat cggctcggcaa gtggtaa                                     1467

```

<210> 10

<211> 488

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

ES 2 575 912 T3

<222> (1)...(22)

<220>

<221> DOMINIO

<222> (33)...(449)

5 <223> Familia de glicosil hidrolasas 15

<220>

<221> SITIO

<222> (194)...(197)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <400> 10

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

ES 2 575 912 T3

100					105					110					
Ser	Asn	Pro	Ser	Gly	Asp	Leu	Ser	Asp	Gly	Ser	Gly	Leu	Ala	Glu	Pro
		115					120					125			
Lys	Tyr	His	Val	Asp	Leu	Thr	Ala	Phe	Thr	Glu	Ala	Trp	Gly	Arg	Pro
	130					135					140				
Gln	Arg	Asp	Gly	Pro	Ala	Leu	Arg	Ala	Thr	Ala	Leu	Ile	Thr	Tyr	Gly
145					150					155					160
Asn	Tyr	Leu	Ile	Ser	Lys	Glu	Lys	Thr	Ser	Val	Val	Lys	Ser	Asn	Ile
				165					170					175	
Trp	Pro	Ile	Val	Gln	Asn	Asp	Leu	Asn	Tyr	Val	Ala	Gln	Tyr	Trp	Asn
			180					185					190		
Gln	Thr	Gly	Phe	Asp	Leu	Trp	Glu	Glu	Val	Gln	Gly	Ser	Ser	Phe	Phe
		195					200					205			
Thr	Ile	Ala	Ala	Gln	His	Arg	Ala	Leu	Val	Glu	Gly	Ser	Ala	Phe	Ala
	210					215					220				
Lys	Ala	Leu	Gly	Glu	Ser	Cys	Glu	Gly	Cys	Asp	Ser	Gln	Ala	Pro	Gln
225					230					235					240
Val	Leu	Cys	Phe	Gln	Gln	Ser	Phe	Trp	Asn	Gly	Lys	Ala	Val	Val	Ser
				245					250					255	
Asn	Phe	Ala	Asn	Asn	Gly	Arg	Thr	Gly	Leu	Asp	Ala	Asn	Ser	Val	Leu
			260					265					270		
Thr	Ser	Ile	Val	Asn	Phe	Asp	Pro	Lys	Ala	Pro	Cys	Asp	Asp	Val	Thr
		275					280					285			
Phe	Gln	Pro	Cys	Ser	Ala	Arg	Ala	Leu	Ser	Asn	His	Lys	Leu	Tyr	Val
	290					295					300				
Asp	Ser	Phe	Arg	Lys	Ile	Tyr	Pro	Val	Asn	Ser	Gly	Lys	Glu	Ala	Gly
305					310					315					320
Thr	Ala	Val	Ala	Val	Gly	Arg	Tyr	Ala	Glu	Asp	Thr	Tyr	Met	Gly	Gly
				325					330					335	
Asn	Pro	Trp	Tyr	Leu	Thr	Thr	Leu	Ala	Ala	Ala	Glu	Gln	Leu	Tyr	Asp
			340					345					350		
Ala	Leu	Tyr	Gln	Trp	Lys	Gln	Leu	Gly	Ser	Leu	Glu	Ile	Thr	Glu	Ala
		355					360					365			
Ser	Leu	Pro	Phe	Phe	Lys	Asp	Leu	Val	Ser	Ser	Ala	Ala	Ala	Gly	Lys
	370					375					380				

ES 2 575 912 T3

Tyr Pro Ser Ser Ser Glu Thr Tyr Thr Ser Ile Thr Ala Ala Val Lys
385 390 395 400

Lys Tyr Ala Asp Gly Phe Met Ala Val Val Lys Glu His Thr Pro Ser
405 410 415

Asp Gly Ser Leu Ser Glu Gln Phe Thr Arg Asp Asn Gly Ser Pro Ala
420 425 430

Ser Ala Lys Asp Leu Thr Trp Ser Tyr Ala Ala Leu Leu Ser Ala Thr
435 440 445

Arg Arg Glu Ala Gly Thr Val Pro Pro Ser Trp Gly Ala Ser Thr Ala
450 455 460

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

Thr Thr Pro Ser Val Gly Lys Trp
485

<210> 11

<211> 1746

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 11

ES 2 575 912 T3

atgtttaatc aagttctcta tggcttggct gccactgcac tatggcaggg ccaagtcgtc 60
 gcttcaccaa gcaaagacaa ctccctcgag aagttcatca ccaagcaagc cgatatttcc 120
 atcaaagggtg tcctcgccaa tatcggctct gatggtaaac gtgcacaagg tgcagcgctt 180
 ggtgccgtcg tcgcaagccc atcacgagaa gaccctgact attggtacac ttggacgctg 240
 gactcagcct tgacttataa ggtcctcgtc gagcgtctca tccacggcga aaagtctctc 300
 cagcgggaaga tagacgaata cgtctccgct caagccaaac ttcagaaaac cacaaccca 360
 tccggtagtc cagagtccgg tggctcggc gagccaaagt tccacgtgaa cctcactgcc 420
 ttcaccggat cttggggccg cccacagcgc gatggtcctc ctctgcgtgc tacagccttg 480
 accctatacg cagagtggct catctcccac ggagacaagt ccaaggctgt gaacaaagtg 540
 tggccagtga ttgagaagga tcttgcgtat actaccaagt tctggaatcg cactgggtat 600
 gatctctggg aagaagtcaa tggatcttct ttctttacgc tatcggcttc gcatcgggct 660
 ctcataggag gagcggcttt ggcgaagaag ttgggcaagt cttgccctga ctgtgctgcc 720
 aatgccctc gtgtcctttg cttcatgcag agcttttggg ctggtgggta cattgactcg 780
 aacatcaatg tgaaggatgg tcgcaagggc ttggatgcta actccattct gtcatcgatc 840
 cacacatttg atcccaactc caaatgcacc gactcgacat tccagccttg ctccctcacga 900

 gcgctcgcca accacaaggc ggtggtcgac tctttcaggt ccatctacat tgtcaacaag 960
 aacagaggca aaggtaaagc cgcagctgtt ggcagataca gtgaggatgt gtactacaac 1020
 ggcaaccctt ggtacttgac taccctgcc gctgctgagc aattgtacgc cgcgctgtac 1080
 caatggaaca aggttggggc tgtctcaatt gacgacgttt ctttaccttt cttccgggac 1140
 cttgtgccta agggatccaa gggcacgtat aagaagaaca gcaagacgta caaggagatc 1200
 gtcaaggctg tcaaggccta cgctgacggg tttgtcgctg ttgttcaaac ctacacacc 1260
 aaagatgggt cgctagccga gcagttcgac cgagccactg gtactcccaa atcggcggtc 1320
 catctcactt ggtcttacgc ctccctcgtc agcgcgaccg aacgtcgctc cagcatcgtc 1380
 tctccctctt ggggcgagag cagcgccaat aaggttcctg cagtatgcga agcagccccg 1440
 gcttgtgaca cgacgataac cttcaacgtg aagaatgtgg aagtgtcttc tgaccagaag 1500
 gtttacgtcg ttggctccgt gactgaactt tccaactggt cacctgatga gggcatacca 1560
 ctacaggagg gtaccaaggg gttgtggagt gccaaaggta agattccttc tgatacaagc 1620
 tttgagtata agtacatcaa gaagacaagt ggcggagatg ttacgtgggt gagcgacccg 1680
 aacaatcggg ctgttacggg cagtaagtgt ggaagcgcga gtactcttga taatgagtgg 1740
 aggtag 1746

<210> 12

<211> 581

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(21)

<220>

<221> DOMINIO

5 <222> (38)...(454)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

<222> (483)...(576)

10 <223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (139)...(142)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

15 <220>

<221> SITIO

<222> (199)...(202)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

20 <221> SITIO

<222> (201)...(211)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<220>

<221> SITIO

25 <222> (210)...(213)

ES 2 575 912 T3

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 12

ES 2 575 912 T3

Met Phe Asn Gln Val Leu Tyr Gly Leu Ala Ala Thr Ala Leu Trp Gln
1 5 10 15

Gly Gln Val Val Ala Ser Pro Ser Lys Asp Asn Ser Leu Glu Lys Phe
20 25 30

Ile Thr Lys Gln Ala Asp Ile Ser Ile Lys Gly Val Leu Ala Asn Ile
35 40 45

Gly Ser Asp Gly Lys Arg Ala Gln Gly Ala Ala Pro Gly Ala Val Val
50 55 60

Ala Ser Pro Ser Arg Glu Asp Pro Asp Tyr Trp Tyr Thr Trp Thr Arg
65 70 75 80

Asp Ser Ala Leu Thr Tyr Lys Val Leu Val Glu Arg Leu Ile His Gly
85 90 95

Glu Lys Ser Leu Gln Arg Lys Ile Asp Glu Tyr Val Ser Ala Gln Ala
100 105 110

Lys Leu Gln Lys Thr Thr Asn Pro Ser Gly Ser Pro Glu Ser Gly Gly
115 120 125

Leu Gly Glu Pro Lys Phe His Val Asn Leu Thr Ala Phe Thr Gly Ser
130 135 140

Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro Leu Arg Ala Thr Ala Leu
145 150 155 160

Thr Leu Tyr Ala Glu Trp Leu Ile Ser His Gly Asp Lys Ser Lys Ala
165 170 175

Val Asn Lys Val Trp Pro Val Ile Glu Lys Asp Leu Ala Tyr Thr Thr
180 185 190

Lys Phe Trp Asn Arg Thr Gly Tyr Asp Leu Trp Glu Glu Val Asn Gly
195 200 205

Ser Ser Phe Phe Thr Leu Ser Ala Ser His Arg Ala Leu Ile Glu Gly
210 215 220

Ala Ala Leu Ala Lys Lys Leu Gly Lys Ser Cys Pro Asp Cys Ala Ala
225 230 235 240

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

Tyr Ile Asp Ser Asn Ile Asn Val Lys Asp Gly Arg Lys Gly Leu Asp

ES 2 575 912 T3

	260							265							270
Ala	Asn	Ser 275	Ile	Leu	Ser	Ser	Ile 280	His	Thr	Phe	Asp	Pro 285	Asn	Ser	Lys
Cys	Thr 290	Asp	Ser	Thr	Phe	Gln 295	Pro	Cys	Ser	Ser	Arg 300	Ala	Leu	Ala	Asn
His 305	Lys	Ala	Val	Val	Asp 310	Ser	Phe	Arg	Ser	Ile 315	Tyr	Ile	Val	Asn	Lys 320
Asn	Arg	Gly	Lys	Gly 325	Lys	Ala	Ala	Ala	Val 330	Gly	Arg	Tyr	Ser	Glu 335	Asp
Val	Tyr	Tyr	Asn 340	Gly	Asn	Pro	Trp	Tyr 345	Leu	Thr	Thr	Leu	Ala 350	Ala	Ala
Glu	Gln	Leu 355	Tyr	Ala	Ala	Leu	Tyr 360	Gln	Trp	Asn	Lys	Val 365	Gly	Ala	Val
Ser	Ile 370	Asp	Asp	Val	Ser	Leu 375	Pro	Phe	Phe	Arg	Asp 380	Leu	Val	Pro	Lys
Gly 385	Ser	Lys	Gly	Thr	Tyr 390	Lys	Lys	Asn	Ser	Lys 395	Thr	Tyr	Lys	Glu	Ile 400
Val	Lys	Ala	Val	Lys 405	Ala	Tyr	Ala	Asp	Gly 410	Phe	Val	Ala	Val	Val 415	Gln
Thr	Tyr	Thr	Pro 420	Lys	Asp	Gly	Ser	Leu 425	Ala	Glu	Gln	Phe	Asp 430	Arg	Ala
Thr	Gly	Thr 435	Pro	Lys	Ser	Ala	Val 440	His	Leu	Thr	Trp	Ser 445	Tyr	Ala	Ser
Phe	Val 450	Ser	Ala	Thr	Glu	Arg 455	Arg	Ser	Ser	Ile	Val 460	Ser	Pro	Ser	Trp
Gly 465	Glu	Ser	Ser	Ala	Asn 470	Lys	Val	Pro	Ala	Val 475	Cys	Glu	Ala	Ala	Pro 480
Ala	Cys	Asp	Thr	Thr 485	Ile	Thr	Phe	Asn	Val 490	Lys	Asn	Val	Glu	Val 495	Ser
Ser	Asp	Gln	Lys 500	Val	Tyr	Val	Val	Gly 505	Ser	Val	Thr	Glu	Leu 510	Ser	Asn
Trp	Ser	Pro 515	Asp	Glu	Gly	Ile	Pro 520	Leu	Thr	Glu	Gly	Thr 525	Lys	Gly	Leu
Trp	Ser 530	Ala	Lys	Val	Lys	Ile 535	Pro	Ser	Asp	Thr	Ser 540	Phe	Glu	Tyr	Lys

ES 2 575 912 T3

Tyr Ile Lys Lys Thr Ser Gly Gly Asp Val Thr Trp Leu Ser Asp Pro
545 550 555 560

Asn Asn Arg Ala Val Thr Gly Ser Lys Cys Gly Ser Ala Ser Thr Leu
565 570 575

Asp Asn Glu Trp Arg
580

<210> 13

<211> 1749

<212> ADN

5 <213> Fusarium verticillioides GZ3639

<400> 13

ES 2 575 912 T3

atgtttactc	agatcttgta	tggcctcacg	gccttatctg	ccctccaagg	tcaggtcact	60
gcatcaccag	gcggttccag	ccttgatcgc	ttcatttcca	aagaggccga	catctccatc	120
aagggcgtgc	ttgccaatat	tggcgccgat	ggcaagcgag	cccaaggcgc	tgcacctggc	180
gccgtttag	caagtccatc	gagaacagat	ccagactact	ggtacacatg	gacccgagac	240
tccgccttga	catacaaagt	ccttgttgag	cgcttcattc	acggagacaa	gtctctccag	300
cgcaagatcg	acgaatatgt	ctctgcccga	gccaagctcc	agggcgtcac	caacccatct	360
ggcggccccg	agtcaggcgg	ccttggggaa	cccaagtttc	acgtcaacct	cacagctttc	420
acaggatcct	ggggctgctc	tcaacgagat	ggcctctctc	tgagagctac	tgcattgacg	480
ctgtacgcca	actggcttgt	ttctcacggc	gaccgctcca	aggccgtcaa	caaggtttgg	540
cctgtcattg	agaaggatct	tgcatacacc	gtcaagttct	ggaacagaac	cggttacgat	600
ctttgggagg	aggttaacgg	atcttcgctc	ttcacctctc	ctgcttcaca	tcgtgctctg	660
gttgagggag	ctgctcttgc	taagaagctt	ggcaagtctt	gctccgactg	cgcaaccaac	720
gccccccgtg	ttctctgctt	catgcaaagc	ttctggaccg	gcagctacat	cgactcgaac	780
atcaatgtca	acgatggccg	caaggtctt	gatgccaaact	ccattctgtc	ttctattcac	840
acctttgatc	cttcttcgaa	gtgcacagac	tctaccttcc	agccttgctc	ttcaagggct	900
cttgcaacc	acaaggaagt	agtggactct	ttccgctcca	tctatgggtg	caacaaaaac	960
agaggtaaag	gtaaagctgc	tgctgtcggg	cgatacagtg	aggatgtgta	ctacgacggg	1020
aacccttggg	acttggctac	tcttgcgctc	gccgagcaac	tgtacgctgc	tgtctatcaa	1080
tggaacaaga	tcggttccat	cacagttgat	agtgtgtcgc	tccccttttt	cagtgacctt	1140
gtaccaaaag	tttccaaggg	aacctatcgc	aagaacagca	agacatacaa	ggctattatc	1200
aaggctgtca	cttcatacgc	tgatggcttt	gtcgccgctg	tgcagacctc	tactcccaaa	1260
gatggctccc	tcgagagca	gttcgataag	tctactggaa	ctccaagtc	agctgttcac	1320
ctaacctggg	cctacgcctc	ctttgtcggg	gctgccgagc	gtcgactggg	cgctcttctt	1380
ccagcttggg	gagagagcaa	cgccaacaag	gtgcctgctg	tttgcaagc	agctccagcc	1440
tgcgacacca	ccatcacggt	caatgtgaag	aacgttgatg	tcacgtcggg	ccagaagggt	1500
tacattggtg	gcgggatcac	tcaactttcc	aactgggcc	ctgctgacgg	cattgctgctt	1560
gaggaatcca	cgagaccaa	gggcttgtgg	actgtgaagg	tcaagattcc	atctgatacc	1620
agctttgagt	ataagtatat	aaagaagacg	agtgatggaa	ctgttacatg	ggagagtgc	1680
cccaataaca	gtgcggcgac	gggcagcaag	tgcggaagca	gcagtacat	caacgatgag	1740
tgaggtag						1749

<210> 14

<211> 582

<212> PRT

5 <213> *Fusarium verticillioides* GZ3639

<220>

<221> SEÑAL

<222> (1)...(21)

<220>

5 <221> DOMINIO

<222> (37)...(453)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

10 <222> (482)...(577)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (138)...(141)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (200)...(210)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

20 <220>

<221> SITIO

<222> (209)...(212)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

25 <221> SITIO

ES 2 575 912 T3

<222> (570)...(573)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 14

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

ES 2 575 912 T3

Ser Ala Leu Thr Tyr₈₅ Lys Val Leu Val Glu₉₀ Arg Phe Ile His Gly₉₅ Asp
Lys Ser Leu Gln₁₀₀ Arg Lys Ile Asp Glu₁₀₅ Tyr Val Ser Ala Gln₁₁₀ Ala Lys
Leu Gln Gly₁₁₅ Val Thr Asn Pro Ser₁₂₀ Gly Gly Pro Glu Ser₁₂₅ Gly Gly Leu
Gly Glu₁₃₀ Pro Lys Phe His Val₁₃₅ Asn Leu Thr Ala Phe₁₄₀ Thr Gly Ser Trp
Gly Arg₁₄₅ Pro Gln Arg Asp₁₅₀ Gly Pro Pro Leu Arg₁₅₅ Ala Thr Ala Leu Thr₁₆₀
Leu Tyr Ala Asn Trp₁₆₅ Leu Val Ser His Gly₁₇₀ Asp Arg Ser Lys Ala₁₇₅ Val
Asn Lys Val Trp₁₈₀ Pro Val Ile Glu₁₈₅ Lys Asp Leu Ala Tyr Thr₁₉₀ Val Lys
Phe Trp Asn₁₉₅ Arg Thr Gly Tyr Asp₂₀₀ Leu Trp Glu Glu Val₂₀₅ Asn Gly Ser
Ser Phe₂₁₀ Phe Thr Leu Ser Ala₂₁₅ Ser His Arg Ala Leu₂₂₀ Val Glu Gly Ala
Ala Leu₂₂₅ Ala Lys Lys Leu₂₃₀ Gly Lys Ser Cys Ser₂₃₅ Asp Cys Ala Thr Asn₂₄₀
Ala Pro Arg Val₂₄₅ Leu Cys Phe Met Gln Ser₂₅₀ Phe Trp Thr Gly Ser₂₅₅ Tyr
Ile Asp Ser Asn₂₆₀ Ile Asn Val Asn Asp₂₆₅ Gly Arg Lys Gly Leu₂₇₀ Asp Ala
Asn Ser Ile₂₇₅ Leu Ser Ser Ile His₂₈₀ Thr Phe Asp Pro Ser₂₈₅ Ser Lys Cys
Thr Asp₂₉₀ Ser Thr Phe Gln Pro₂₉₅ Cys Ser Ser Arg Ala₃₀₀ Leu Ala Asn His
Lys Glu Val Val Asp Ser₃₁₀ Phe Arg Ser Ile Tyr₃₁₅ Gly Val Asn Lys Asn₃₂₀
Arg Gly Lys Gly Lys₃₂₅ Ala Ala Ala Val Gly₃₃₀ Arg Tyr Ser Glu Asp₃₃₅ Val
Tyr Tyr Asp Gly₃₄₀ Asn Pro Trp Tyr Leu₃₄₅ Ala Thr Leu Ala Ala₃₅₀ Ala Glu

ES 2 575 912 T3

Gln Leu Tyr Ala Ala Val Tyr Gln Trp Asn Lys Ile Gly Ser Ile Thr
 355 360 365
 Val Asp Ser Val Ser Leu Pro Phe Phe Ser Asp Leu Val Pro Lys Val
 370 375 380
 Ser Lys Gly Thr Tyr Arg Lys Asn Ser Lys Thr Tyr Lys Ala Ile Ile
 385 390 395 400
 Lys Ala Val Thr Ser Tyr Ala Asp Gly Phe Val Ala Val Val Gln Thr
 405 410 415
 Tyr Thr Pro Lys Asp Gly Ser Leu Ala Glu Gln Phe Asp Lys Ser Thr
 420 425 430
 Gly Thr Pro Lys Ser Ala Val His Leu Thr Trp Ser Tyr Ala Ser Phe
 435 440 445
 Val Gly Ala Ala Glu Arg Arg Thr Gly Val Val Pro Pro Ala Trp Gly
 450 455 460
 Glu Ser Asn Ala Asn Lys Val Pro Ala Val Cys Glu Ala Ala Pro Ala
 465 470 475 480
 Cys Asp Thr Thr Ile Thr Phe Asn Val Lys Asn Val Asp Val Thr Ser
 485 490 495
 Asp Gln Lys Val Tyr Ile Val Gly Gly Ile Thr Gln Leu Ser Asn Trp
 500 505 510
 Ala Pro Ala Asp Gly Ile Ala Leu Glu Glu Ser Thr Ser Thr Lys Gly
 515 520 525
 Leu Trp Thr Val Lys Val Lys Ile Pro Ser Asp Thr Ser Phe Glu Tyr
 530 535 540
 Lys Tyr Ile Lys Lys Thr Ser Asp Gly Thr Val Thr Trp Glu Ser Asp
 545 550 555 560
 Pro Asn Asn Ser Ala Ala Thr Gly Ser Lys Cys Gly Ser Ser Ser Thr
 565 570 575
 Ile Asn Asp Glu Trp Arg
 580

<210> 15

<211> 1188

<212> ADN

5 <213> Cochliobolus heterostrophus ATCC 48332

<400> 15

atgttgcca agatcttgct acccgttggt gcacttgcag ccagtgccaa tgcgcatggc 60

ES 2 575 912 T3

tacctactt	cacccatgag	ccgaacaggg	ctgaatgcac	agtctggggc	tgatacgtgc	120
cccgaatgca	caattctgga	gccggtaaca	gcatggcctg	atcttgacgc	cgcagctgtc	180
ggccgctcgg	gtccttgagg	gtacaatgct	cgtgtgtcag	ttgactacaa	ccagcccggg	240
cctcgtgagg	gttctgagcc	cgtcattacc	tacaaggctg	gggatgtcgt	cgatgtacag	300
tggtgtgtag	atgccaacgg	tgaccatggt	ggcatgttca	cctaccgcat	ctgccaaaac	360
caagctcttg	tcgacaagct	cttgaccctt	ggatacctcc	ctactgaagc	agagaagcaa	420
gcagcagagg	attgcttcca	agctggagag	ttgaagtgca	cagatgtccc	tggacagaca	480
tgcggattca	actctgactg	ccagcaaggc	caagcctgct	ggaggaatga	ctggttcaca	540
tgcggtggct	tcaacgacaa	cctgaaatgc	aggagcgtcg	acaacgcccc	cctaaattca	600
tgctacacca	gcatcgcagg	cggtacacc	gtcagctcaa	agatcaaaaat	cccaaactac	660
accagcaacc	acaccctcct	ctccttcaa	tggaactctt	tccaaacccc	gcaggtctac	720
ctcacctgcg	ccgacatcaa	aatcaccggc	agcagctccg	gcacctcccc	acccccaaact	780
tcgagcaaac	cccctacgtc	ctcctccaag	cccacttcca	cctccacctc	cacctccgca	840
ccctccgcaa	ccccaccgc	ctgcccacc	cccgtctcca	ccgtcgcctg	aaccttcaac	900
tccaaaacca	ccacctcctt	tggccaaacc	gtcaaactcg	caggttccat	ctcccagctc	960
ggcagctgga	acactgcaa	tgcgccctgc	ctgtccgccc	cccagtacac	ctcgtccaac	1020
ccgctctgga	caaccacctt	caacctacct	gcaggcacca	gcttcgagta	caagtacatc	1080
aaagttgaca	gtagtggcgc	cgtcacgtac	gagagtgggg	ccaatagaca	gtacactgtg	1140
cctaacgggt	gcgctaacac	tgctactgta	gaagggactt	ggaaataa		1188

<210> 16

<211> 395

<212> PRT

5 <213> cochliobolus heterostrophus ATCC 48332

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

10 <221> DOMINIO

<222> (294)...(390)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

ES 2 575 912 T3

<222> (222)...(225)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (226)...(229)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 16

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

Asn Ala His Gly Tyr Leu Thr Ser Pro Met Ser Arg Thr Gly Leu Asn
20 25 30

;

ES 2 575 912 T3

Ala Gln Ser Gly Ala Asp Thr Cys Pro Glu Cys Thr Ile Leu Glu Pro
35 40 45

Val Thr Ala Trp Pro Asp Leu Asp Ala Ala Ala Val Gly Arg Ser Gly
50 55 60

Pro Cys Gly Tyr Asn Ala Arg Val Ser Val Asp Tyr Asn Gln Pro Gly
65 70 75 80

Pro Arg Trp Gly Ser Glu Pro Val Ile Thr Tyr Lys Ala Gly Asp Val
85 90 95

Val Asp Val Gln Trp Cys Val Asp Ala Asn Gly Asp His Gly Gly Met
100 105 110

Phe Thr Tyr Arg Ile Cys Gln Asn Gln Ala Leu Val Asp Lys Leu Leu
115 120 125

Thr Pro Gly Tyr Leu Pro Thr Glu Ala Glu Lys Gln Ala Ala Glu Asp
130 135 140

Cys Phe Gln Ala Gly Glu Leu Lys Cys Thr Asp Val Pro Gly Gln Thr
145 150 155 160

Cys Gly Phe Asn Ser Asp Cys Gln Gln Gly Gln Ala Cys Trp Arg Asn
165 170 175

Asp Trp Phe Thr Cys Gly Gly Phe Asn Asp Asn Leu Lys Cys Arg Ser
180 185 190

Val Asp Asn Ala Pro Leu Asn Ser Cys Tyr Thr Ser Ile Ala Gly Gly
195 200 205

Tyr Thr Val Ser Ser Lys Ile Lys Ile Pro Asn Tyr Thr Ser Asn His
210 215 220

Thr Leu Leu Ser Phe Lys Trp Asn Ser Phe Gln Thr Pro Gln Val Tyr
225 230 235 240

Leu Thr Cys Ala Asp Ile Lys Ile Thr Gly Ser Ser Ser Gly Thr Ser
245 250 255

Pro Pro Pro Thr Ser Ser Lys Pro Pro Thr Ser Ser Ser Lys Pro Thr
260 265 270

Ser Thr Ser Thr Ser Thr Ser Ala Pro Ser Ala Thr Pro Thr Ala Cys
275 280 285

Ala Thr Pro Val Ser Thr Val Ala Val Thr Phe Asn Ser Lys Thr Thr
290 295 300

ES 2 575 912 T3

Thr Ser Phe Gly Gln Thr Val Lys Leu Ala Gly Ser Ile Ser Gln Leu
305 310 315 320

Gly Ser Trp Asn Thr Ala Asn Ala Pro Ala Leu Ser Ala Ala Gln Tyr
325 330 335

Thr Ser Ser Asn Pro Leu Trp Thr Thr Thr Leu Asn Leu Pro Ala Gly
340 345 350

Thr Ser Phe Glu Tyr Lys Tyr Ile Lys Val Asp Ser Ser Gly Ala Val
355 360 365

Thr Tyr Glu Ser Gly Ala Asn Arg Gln Tyr Thr Val Pro Asn Gly Cys
370 375 380

Ala Asn Thr Val Thr Val Glu Gly Thr Trp Lys
385 390 395

<210> 17

<211> 1938

<212> ADN

5 <213> Fusarium verticillioides GZ3639

<400> 17

ES 2 575 912 T3

atgtactttg tgccttctgc ctttcttctc ggctcattcg ttcttcagaa cgtcttgggc	60
cgaccaacct tcgatgagag gagtctctta caagagagac agtcttcggt cgactccttt	120
atcaagtctg agagctcaat tgctattgag caactcctct gcaatatcgg ttccgatggc	180
tgcaactcca agaacgtagc cactgggtatt gtcattgctt ctccagacac acaggatcct	240
gactactttt acacctggac tcgagatgct gctctagtct tcaagtacgt cgttgatagg	300
ttcatcaacc agtatgatgc tggcttgtag aggaagatcc aggagtacat cgcctcccaa	360
gccaagctcc aggggtgttc caacccttct gggctgcttt cggatggctc aggtctggga	420
gaggccaagt tcaacgtcga catgagtgcc ttcactgggtg gttggggctg acctcagcga	480
gatgggtccag ctctgctgct gactgctatg atcacctatg ccaactggct gattgccaac	540
ggctacacct ccacagccaa tgacattgtg tggcctgttg ttcgcaacga ccttaactat	600
gtggctcagt attggaatca aaccggattt gacttgtggg aggaggtcaa gggtagttcg	660
ttcttcacaa ctggttccca gtatcgagct ctctattgaag gcgccgctct ggccaagaag	720
ctcggcaagt cgggagacaa ctactccaac atcgcctctc aggctctctg cttcttgtag	780
acttactgga tctcttctgg caatacgtc gactctaaca tcaatgtcaa tgacggccgc	840
actggcaagg acgccaacag tatcctgtcg tccatacaca atttcgacct tgctctgaac	900
tgtgatcccc ccacctcca gccatgcagt gacaaggccc tcgccaacca caaggcggtt	960
actgactctt tccgctcatg gaacatcaac aagggtatct ctcaaggctc agctgtcgcc	1020
gttggacgat acgtcgaaga tgtctactac aatggcaacc cctggtacct cgctacgctt	1080
gccgcggcag agcagctcta cgatgccatt tatgtctgga agcagcaggg atccatcact	1140
gtgtcggacg tctctctttc gttcttcaag gacctcgtct ctttggctct taccggaaca	1200
tacgccagtg actctgccac cttcaagagc atcactgacg ccgtctccaa gtatgctgat	1260
gggtatgttg ccatcgttgc aaagtatgtc ggcacagatg gccacctcgc agagcagttt	1320
gacaagaacg acggccatcc tctttctgcc acagacttga cttggtcata tgccgattc	1380
ctctcagctg ctgatcgtcg agctgggtgtt attcctcctt cttgggctgg aagcgtggct	1440
gctgtcccca accaatgctg taccaatact gttgctggat cctactcatc ggctactgca	1500
acttcgttcc cagcatcgca aacacccaag ggtgggtgtg ccaactcaac tggcgcccag	1560
acttccactt ccaactccac ttccacttcc agctcgtcca ctggtaccag ttgcctact	1620
gcaacctctg tggctgtcac tttccaagaa gttgtcacca ccaactttgg tgataccatc	1680
aagatcgttg gcaacatcgc tgctctcggc aactgggaca catcaaaggc tgttgccctg	1740
agcgcctccg actataccgc ctcgaaccct gtgtggaagg ccaccatttc cctaactgca	1800
ggacagtcca tccagtataa gtacatcaat gttagaagg acggctctct tacctgggag	1860
aaggaccca accgcacctc cgctgttctt aagacatgtg ccacaacggc taccaagtct	1920
gacaagtggc agtcttga	1938

<210> 18

<211> 645

<212> PRT

<213> Fusarium verticillioides GZ3639

<220>

<221> DOMINIO

<222> (46)...(465)

5 <223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

<222> (543)...(639)

<223> Dominio de enlazamiento de almidón

10 <220>

<221> SITIO

<222> (209)...(212)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (211)...(221)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<220>

<221> SITIO

20 <222> (250)...(253)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (633)...(636)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

ES 2 575 912 T3

<400> 18

Met Tyr Phe Val Ser Ser Ala Phe Leu Leu Gly Ser Phe Val Leu Gln
1 5 10 15

Asn Val Leu Gly Arg Pro Thr Phe Asp Glu Arg Ser Leu Leu Gln Glu

ES 2 575 912 T3

Thr Phe Gln Pro Cys Ser Asp Lys Ala Leu Ala Asn His Lys Ala Val
 305 310 315 320

Thr Asp Ser Phe Arg Ser Trp Asn Ile Asn Lys Gly Ile Ser Gln Gly
 325 330 335

Ser Ala Val Ala Val Gly Arg Tyr Val Glu Asp Val Tyr Tyr Asn Gly
 340 345 350

Asn Pro Trp Tyr Leu Ala Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp
 355 360 365

Ala Ile Tyr Val Trp Lys Gln Gln Gly Ser Ile Thr Val Ser Asp Val
 370 375 380

Ser Leu Ser Phe Phe Lys Asp Leu Val Ser Leu Val Ser Thr Gly Thr
 385 390 400

Tyr Ala Ser Asp Ser Ala Thr Phe Lys Ser Ile Thr Asp Ala Val Ser
 405 410 415

Lys Tyr Ala Asp Gly Tyr Val Ala Ile Val Ala Lys Tyr Val Gly Thr
 420 425 430

Asp Gly His Leu Ala Glu Gln Phe Asp Lys Asn Asp Gly His Pro Leu
 435 440 445

Ser Ala Thr Asp Leu Thr Trp Ser Tyr Ala Ala Phe Leu Ser Ala Ala
 450 455 460

Asp Arg Arg Ala Gly Val Ile Pro Pro Ser Trp Ala Gly Ser Val Ala
 465 470 475 480

Ala Val Pro Asn Gln Cys Gly Thr Asn Thr Val Ala Gly Ser Tyr Ser
 485 490 495

Ser Ala Thr Ala Thr Ser Phe Pro Ala Ser Gln Thr Pro Lys Gly Gly
 500 505 510

Val Pro Thr Pro Thr Gly Ala Gln Thr Ser Thr Ser Thr Ser Thr Ser
 515 520

Thr Ser Ser Ser Thr Gly Thr Ser Cys Pro Thr Ala Thr Ser Val
 530 535 540

Ala Val Thr Phe Gln Glu Val Val Thr Thr Asn Phe Gly Asp Thr Ile
 545 550 555 560

Lys Ile Val Gly Asn Ile Ala Ala Leu Gly Asn Trp Asp Thr Ser Lys
 565 570 575

ES 2 575 912 T3

Ala Val Ala Leu Ser Ala Ser Asp Tyr Thr Ala Ser Asn Pro Val Trp
580 585 590

Lys Ala Thr Ile Ser Leu Thr Ala Gly Gln Ser Ile Gln Tyr Lys Tyr
595 600 605

Ile Asn Val Lys Lys Asp Gly Ser Leu Thr Trp Glu Lys Asp Pro Asn
610 615 620

Arg Thr Tyr Ala Val Pro Lys Thr Cys Ala Thr Thr Ala Thr Lys Ser
625 630 635 640

Asp Lys Trp Gln Ser
645

<210> 19

<211> 1467

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 19

ES 2 575 912 T3

```

atgttgacat tgaatgtttt gacagcactg ttggctccta ttgtgctgtc atctgccttg      60
cctgctcgcg atctcaatgc gcgtgccgat tcgaagcctg ctgctcttca agccatcctg      120
aacaacattg gcgccgatgg atctgctgtt tcgggagcat ctgcaggcgt tgtggtagca      180
tccccatcca aatcggatcc cgattacttt tacacctgga ctcgtgatgc agccttgaca      240
tacaaggtct tgattgatga gttcattgcc ggagacacct cgctggaacc taccatccag      300
gactacgtct ctgctcaagc aaagctgcag gccgtgtcta acccatctgg tgacttgctc      360
gatggatcag gacttgcaga gcctaagtac cacgtggact tgactgcctt cacggaggcc      420
tggggctcgtc ctcagcgcga cggaccgcct ctgagagcca cagcactgat tacctacggt      480
aactatattga tttcgaagga gagaacatcc gtcgtcaagt cgaacatctg gccattgtg      540
caaatgatc tcaactatgt tgcgcaatac tggaaacaaa ccggattcga tctgtgggag      600
gaggctcagg gctcctcgtt ttttaccatt gctgcacaac accgcgcatt ggtggagggc      660
agcgcatttg ccaaggcgcct gggagagtcc tgcgagggat gcgattccca ggcacctcag      720
gtcctatgct tccagcagtc cttctgggat ggcaaggcta ttgtttctaa cttcgccaac      780
aatggccgaa ccggtctcga tgccaattcg gtgcttacct cgatcggaaa cttcgacccc      840
aaggctccct gcgatgacgt gaccttccag ccttgctctg ctcgtgccct gtcgaatcac      900
aagctatacg ttgactcgtt ccgtaagatc taccctgtga acagtggtaa ggaagctgga      960
actgccgttg ctgttgacg ttattccgaa gacacttaca tgggcgggaa cccatggtat     1020
ttgaccacc c tggccgcggc cgagcagctg tacgatgctc tttaccagtg gaagcaactg     1080
ggatcgctag aaatcaccga gatcagtctt cttttcttca aggaccttgt ttcgtctgcc     1140
gccgccggaa agtatcccag ctctcggaa acttacacgt ccatcactgc cgcagtcaaa     1200

aagtatgccg atggatttgt ggctgttggt aaggagcaca cccccagcga tgggtctctg     1260
gctgaacaat aactcgggga caacggcagc ccggcctctg ctaaggactt gacctggtcc     1320
tacggcggcg tcctgtctgc tacccgccgt gaggctggaa cagtgcctcc tacctggggg     1380
gcgtcgactg ccaataaggt accctcaaag tgtgagggta gctcggccaa gggaagctac     1440
acgacacccat cagtcggcaa gtggtaa                                     1467

```

<210> 20

<211> 488

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

ES 2 575 912 T3

<222> (1)...(22)

<220>

<221> DOMINIO

<222> (33)...(449)

5 <223> Familia de glicosil hidrolasas 15

<220>

<221> SITIO

<222> (194)...(197)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <400> 20

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

ES 2 575 912 T3

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

Asn Tyr Leu Ile Ser Lys Glu Arg Thr Ser Val Val Lys Ser Asn Ile
165 170 175

Trp Pro Ile Val Gln Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn
180 185 190

Gln Thr Gly Phe Asp Leu Trp Glu Glu Val Glu Gly Ser Ser Phe Phe
195 200 205

Thr Ile Ala Ala Gln His Arg Ala Leu Val Glu Gly Ser Ala Phe Ala
210 215 220

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

Val Leu Cys Phe Gln Gln Ser Phe Trp Asp Gly Lys Ala Ile Val Ser
245 250 255

Asn Phe Ala Asn Asn Gly Arg Thr Gly Leu Asp Ala Asn Ser Val Leu
260 265 270

Thr Ser Ile Gly Asn Phe Asp Pro Lys Ala Pro Cys Asp Asp Val Thr
275 280 285

Phe Gln Pro Cys Ser Ala Arg Ala Leu Ser Asn His Lys Leu Tyr Val
290 295 300

Asp Ser Phe Arg Lys Ile Tyr Pro Val Asn Ser Gly Lys Glu Ala Gly
305 310 315 320

Thr Ala Val Ala Val Gly Arg Tyr Ser Glu Asp Thr Tyr Met Gly Gly
325 330 335

Asn Pro Trp Tyr Leu Thr Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp
340 345 350

Ala Leu Tyr Gln Trp Lys Gln Leu Gly Ser Leu Glu Ile Thr Glu Ile
355 360 365

Ser Leu Pro Phe Phe Lys Asp Leu Val Ser Ser Ala Ala Ala Gly Lys
370 375 380

Tyr Pro Ser Ser Ser Glu Thr Tyr Thr Ser Ile Thr Ala Ala Val Lys
385 390 395 400

Lys Tyr Ala Asp Gly Phe Val Ala Val Val Lys Glu His Thr Pro Ser
405 410 415

ES 2 575 912 T3

Asp Gly Ser Leu Ala Glu Gln Tyr Thr Arg Asp Asn Gly Ser Pro Ala
420 425 430

Ser Ala Lys Asp Leu Thr Trp Ser Tyr Ala Ala Leu Leu Ser Ala Thr
435 440 445

Arg Arg Glu Ala Gly Thr Val Pro Pro Thr Trp Gly Ala Ser Thr Ala
450 455 460

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

Thr Thr Pro Ser Val Gly Lys Trp
485

<210> 21

<211> 1488

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 21

ES 2 575 912 T3

```

atggttctag ctcggcttgc ctggcttgcc ggcctggta gcactgccgt tgctgcgacc 60
ccagcggaaat ggcgctcgca gtccatctac ttcattgctca cggatcgttt tgcccggacg 120
gatggttcaa ctactgctgc ctgcgatacc gctgacagaa aatactgtgg cggaacctgg 180
cagggaaatca ttgacaagct ggactatata caaggaatgg gctttacagc catttgatc 240
actcctgtga ccggtcaatt gagcggggaa accgcgtacg gagatgccta tcacggatac 300
tggcagcagg atatctattc tctcgattcc aactatggaa ccgcagacga tctcaaggcc 360
ctcgtgcggt ctttgcacga acgcaacatg tatctcatgg tcgatgtcgt agctaaccac 420
atgggctaca atggcccagg tactgacgtg gactacacca aattcaacc cttcaacgat 480
gcaaagtatt tccactcgta ctgcccatac accgattaca acgacgacac catgtcgcaa 540
aactgctggc ttggcgataa caaggctctg ctaccggacc tgaatacaca gagcaaggag 600
gtccaggatc tatggtatga ctgggttggg tctttggtct ccaactactc tatcgacggt 660
ctccgtgtcg acacagtcaa acatgtccag aaagatttct ggcccggcta caacaaagcc 720
gcgggcgtct actgcgtagg cgagatcctt gatggtgacc cagattacac ctgtccatac 780
caggaggtaa tggacggagt gctcaactac ccgatttatt acccactcct caaggccttc 840
caatcgacct cgggaagcat gaccgatcta tacaacatga tcaacacggg gaaatcgacc 900
tgcaaggact cgacccttct tggaaatttc ttggagaacc acgataaacc ccgttttgcc 960
cacgtcaccg acgacattgc cctcgccaag aacgcagcta cgtttaccat tatggcagac 1020
ggcattccta ttgtctatgc aggacaggag cagcactaca gtggtggcga ggaccggct 1080
aatcggaag ctctgtggtt gtcaggatac aacacggaca gtgagctgta caagctcata 1140
gccaaaggcca atggtgctag aagccaggcc attgctaagg gtaccaacta tacgattac 1200

cagaaccaac caatctacaa agatgagagc accatcgcca tgcggaaggg cttcgacggt 1260
ggacagacaa tcaactgctc gacgaatctc ggcgcagggg gtaaagagga ctctgtttcg 1320
attcctgata ccggattcaa ggctggtgca aagttgactg aggtcgtctc ctgctgtagt 1380
gttactgtcg gtgacaatgg ggaggtgtct gtcctatgg cggctggagc gccgaggatt 1440
ttgctcccta ctttttctg tggggctcg actctgtgtt catcatag 1488

```

<210> 22

<211> 495

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

10 <222> (1)...(18)

<220>

<221> DOMINIO

<222> (31)...(387)

<223> Alfa amilasa, dominio catalítico

<220>

5 <221> SITIO

<222> (197)...(200)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

10 <222> (218)...(221)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (402)...(405)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 22

ES 2 575 912 T3

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

ES 2 575 912 T3

Gly Thr Ala Asp Asp Leu Lys Ala Leu Ala Ala Ala Leu His Glu Arg
 115 120 125

Asn Met Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Tyr Asn
 130 135 140

Gly Pro Gly Thr Asp Val Asp Tyr Thr Lys Phe Asn Pro Phe Asn Asp
 145 150 155 160

Ala Lys Tyr Phe His Ser Tyr Cys Pro Ile Thr Asp Tyr Asn Asp Asp
 165 170 175

Thr Met Ser Gln Asn Cys Trp Leu Gly Asp Asn Lys Val Ser Leu Pro
 180 185 190

Asp Leu Asn Thr Gln Ser Lys Glu Val Gln Asp Leu Trp Tyr Asp Trp
 195 200 205

Val Gly Ser Leu Val Ser Asn Tyr Ser Ile Asp Gly Leu Arg Val Asp
 210 215 220

Thr Val Lys His Val Gln Lys Asp Phe Trp Pro Gly Tyr Asn Lys Ala
 225 230 235 240

Ala Gly Val Tyr Cys Val Gly Glu Ile Leu Asp Gly Asp Pro Asp Tyr
 245 250 255

Thr Cys Pro Tyr Gln Glu Val Met Asp Gly Val Leu Asn Tyr Pro Ile
 260 265 270

Tyr Tyr Pro Leu Leu Lys Ala Phe Gln Ser Thr Ser Gly Ser Met Thr
 275 280 285

Asp Leu Tyr Asn Met Ile Asn Thr Val Lys Ser Thr Cys Lys Asp Ser
 290 295 300

Thr Leu Leu Gly Asn Phe Leu Glu Asn His Asp Asn Pro Arg Phe Ala
 305 310 315 320

His Val Thr Asp Asp Ile Ala Leu Ala Lys Asn Ala Ala Thr Phe Thr
 325 330 335

Ile Met Ala Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln Glu Gln His
 340 345 350

Tyr Ser Gly Gly Glu Asp Pro Ala Asn Arg Glu Ala Leu Trp Leu Ser
 355 360 365

Gly Tyr Asn Thr Asp Ser Glu Leu Tyr Lys Leu Ile Ala Lys Ala Asn
 370 375 380

ES 2 575 912 T3

Gly Ala Arg Ser Gln Ala Ile Ala Lys Gly Thr Asn Tyr Thr Ile Tyr
385 390 395 400

Gln Asn Gln Pro Ile Tyr Lys Asp Glu Ser Thr Ile Ala Met Arg Lys
405 410 415

Gly Phe Asp Gly Gly Gln Thr Ile Thr Val Leu Thr Asn Leu Gly Ala
420 425 430

Gly Gly Lys Glu Asp Ser Val Ser Ile Pro Asp Thr Gly Phe Lys Ala
435 440 445

Gly Ala Lys Leu Thr Glu Val Val Ser Cys Ala Ser Val Thr Val Gly
450 455 460

Asp Asn Gly Glu Val Ser Val Pro Met Ala Ala Gly Ala Pro Arg Ile
465 470 475 480

Leu Leu Pro Thr Phe Leu Leu Glu Gly Ser Thr Leu Cys Ser Ser
485 490 495

<210> 23

<211> 1389

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 23

ES 2 575 912 T3

atgaaactgt cccacaccct gacagctctt ctactccctc tgatatgcac tgtatcagca 60
 gcggatgtga aagcatggaa gtcccgcaat atttactttg ccctgactga tcgcattgcc 120
 cgtggaagcg atgatacagg tggagatgcc tgcggttaacc ttggaaacta ctgcggtgga 180
 actttccaag gcttggagtc taagcttgac tatattaagg ggatgggctt tgatgcaatc 240
 tggatcacc cttgttatcgc taatgcgcca ggtggatc atggatactg ggcgcgggat 300
 ctctatagca tcaacgagaa ctatggcacc gctgatgatc tcaaaagcct agttgatgct 360
 gctcataaga aggggatcta tgtcatggca gacgtagttg ccaaccacat gggaggcccc 420
 atcagcgata acaagccgga gccattgaac caggagagct cctaccattc cacctgcaca 480
 atcgactact caagccagga tagtgttgaa aactgccgta tcacagcaga cctacctgat 540
 gtgaacacac agagtcccga gatccgtgcc ctcttcaga aatgggtcaa atggctcgtc 600
 acggaatagc ggttcgacgg cttgcgcat gacactgtta agcacgttga aaaggacttc 660
 tggtcggcct tctcctccgc tgccggggtt tacaccatcg gcgaagtctg ggacggtgac 720
 ccagcttacc ttgccgata tgcacaggac atggacggcc tgctcaacta tgcagtctac 780
 tatcccgtga acaactttta ccagcaaaag ggctcttccc aggacatagt tgacatgcat 840
 gataaaatcg acaccgcttt ccctgatccc agcgccttg gcacattcat agacaatcac 900
 gacaatgcgc gatggctgag cgtaaagaat gacaagtcgc ttttgaaaaa cgcactcgcc 960

 tacgtgattc tcgcccgggg catccccatt gtctactacg gcacggaaca gggctacgcg 1020
 ggcggaacg accccgcaa cgcgaggac ctatggcgca gcaaattcag caccgacgcg 1080
 gacttgatac aggccatc cctgctctca gcggcgagaa acgcttctgg tggcctcgct 1140
 gacaacgacc atgtccatct gtacgtcgcg gagtcggcat atgctgagg cagggcgggt 1200
 ggaaacctca tcgtccttac gtccaatggt ggatctgggt ctgaggctaa tcaactgcttc 1260
 gactctaaga agcctggtgg gacatggaag aataccttg gagaggggac aatcaccgcc 1320
 gatgaaggcg gaaaaatttg catctctatc tctaattggt agcctgagg gctggttgca 1380
 agcacttga 1389

<210> 24

<211> 462

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

ES 2 575 912 T3

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

5 <221> DOMINIO

<222> (33)...(373)

<223> Alfa amilasa, dominio catalítico

<220>

<221> SITIO

10 <222> (379)...(382)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 24

```

Met Lys Leu Ser His Thr Leu Thr Ala Leu Leu Leu Pro Leu Ile Cys
 1          5          10          15

Thr Val Ser Ala Ala Asp Val Lys Ala Trp Lys Ser Arg Asn Ile Tyr
 20          25          30

Phe Ala Leu Thr Asp Arg Ile Ala Arg Gly Ser Asp Asp Thr Gly Gly
 35          40          45

Asp Ala Cys Gly Asn Leu Gly Asn Tyr Cys Gly Gly Thr Phe Gln Gly
 50          55          60

Leu Glu Ser Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile
 65          70          75          80

Trp Ile Thr Pro Val Ile Ala Asn Ala Pro Gly Gly Tyr His Gly Tyr
 85          90          95

Trp Ala Arg Asp Leu Tyr Ser Ile Asn Glu Asn Tyr Gly Thr Ala Asp
 100         105         110

Asp Leu Lys Ser Leu Val Asp Ala Ala His Lys Lys Gly Ile Tyr Val
    
```

ES 2 575 912 T3

115 120 125

Met Ala Asp Val Val Ala Asn His Met Gly Gly Pro Ile Ser Asp Asn
130 135 140

Lys Pro Glu Pro Leu Asn Gln Glu Ser Ser Tyr His Ser Thr Cys Thr
145 150 155 160

Ile Asp Tyr Ser Ser Gln Asp Ser Val Glu Asn Cys Arg Ile Thr Ala
165 170 175

Asp Leu Pro Asp Val Asn Thr Gln Ser Pro Glu Ile Arg Ala Leu Phe
180 185 190

Gln Lys Trp Val Lys Trp Leu Val Thr Glu Tyr Gly Phe Asp Gly Leu
195 200 205

Arg Ile Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Ser Ala Phe
210 215 220

Ser Ser Ala Ala Gly Val Tyr Thr Ile Gly Glu Val Trp Asp Gly Asp
225 230 235 240

Pro Ala Tyr Leu Ala Gly Tyr Ala Gln Asp Met Asp Gly Leu Leu Asn
245 250 255

Tyr Ala Val Tyr Tyr Pro Val Asn Asn Phe Tyr Gln Gln Lys Gly Ser
260 265 270

Ser Gln Asp Ile Val Asp Met His Asp Lys Ile Asp Thr Ala Phe Pro
275 280 285

Asp Pro Ser Ala Leu Gly Thr Phe Ile Asp Asn His Asp Asn Ala Arg
290 295 300

Trp Leu Ser Val Lys Asn Asp Lys Ser Leu Leu Lys Asn Ala Leu Ala
305 310 315 320

Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr Glu
325 330 335

Gln Gly Tyr Ala Gly Gly Asn Asp Pro Ala Asn Arg Glu Asp Leu Trp
340 345 350

Arg Ser Lys Phe Ser Thr Asp Ala Asp Leu Tyr Lys Ala Ile Ser Leu
355 360 365

Leu Ser Ala Ala Arg Asn Ala Ser Gly Gly Leu Ala Asp Asn Asp His
370 375 380

Val His Leu Tyr Val Ala Glu Ser Ala Tyr Ala Trp Ser Arg Ala Gly
385 390 395 400

ES 2 575 912 T3

Gly Asn Leu Ile Val Leu Thr Ser Asn Gly Gly Ser Gly Ser Glu Ala
405 410 415

Asn His Cys Phe Asp Ser Lys Lys Pro Gly Gly Thr Trp Lys Asn Thr
420 425 430

Phe Gly Glu Gly Thr Ile Thr Ala Asp Glu Gly Gly Lys Ile Cys Ile
435 440 445

Ser Ile Ser Asn Gly Glu Pro Ala Val Leu Val Ala Ser Thr
450 455 460

<210> 25

<211> 1923

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 25

ES 2 575 912 T3

atgactatTT cTgcctctc ctccgtgctt tttgactgg ctctaggTca aagtgcacta 60
 gcagcccctc agctgtctcc tcgtgctaca actagctTgg acgcctggct tgccagtgaa 120
 actactgtct ctctcaatgg aatctTggac aacatcggtg ccagcggTgc ctatgctcag 180
 agtgccaagg ctggTgtTgt tattgccagt cccagcaca a gcagcccgga ttattactac 240
 acgtggacca gagactctgc cctTactgTta aaagTcctta tcgatctctt ccgtaatggg 300
 aacgtagatc tgcagaccgt cattgaggaa tacatcactg cacaagctta tctacagacc 360
 gtctccaacc cctccggaga tctctcaagc ggtgctggTc tcgctgagcc caagTtcaac 420
 gttgacatga gcgctTcac tggcgctTgg ggacgtccgc aacgcgacgg acctgctctc 480
 cgtgctattg cctTgatcga tttTggcaac tggTtgattg acaatggata ctctagctat 540
 gccgTtagca acgtctggcc cattgtacgc aacgatctgt cctacgtTgc ccagTactgg 600
 agccaaagcg gttatgatct ctgggaggag gtcaacagca tgTccttctt caccattgcc 660
 aaccagcacc gtgccctcgt cgagggaaagc acctTggccg gccgagTtg tgcacctgc 720
 tcgtggTgtg actctcaagc cccccagatc ctctgctaca tgcagaactt ctggaccgga 780
 tcttacatca atgccaacac tggcggtggT cgTtccggca aggacgcca cactgttctg 840
 gctagtatca gcactTttga tcccgaggct acctgtgacg acgtcactTt ccagccctgc 900
 tcctcccgcg ccctggccaa ccacaaagTc tacactgact cgTtccgatc tgtctattct 960
 ctgactccg gcattgctga gggcgTtgcT gTcgcggtag gacgTaccTc tgaggattca 1020
 tactataacg gcaatccgTg gTttTgacc actctcgctg ctgcggaaca gTtgtacgat 1080
 gccatataTc aatggaacaa gatcggctcc atcacaatca ccagcacctc tttggctTtc 1140
 tttaatgacg tctacagctc tgccgctgtt ggcactTacg cctctggcag cactgcctac 1200
 acagccattg tcagcgcagt caagacctat gccgacggat atgtcagcat cgtgcaggct 1260

 catgccatga ccaacggctc tctctccgag cagTttgaca aggcctccgg taccagctc 1320
 tctgcccgtg atcttacatg gTcgtacgct gcctTgctca ctgctaacat gcgcccgaac 1380
 ggaattgtgc ctcttctctg gggagctgcc tccgccaact caatccccag ttctTgctct 1440
 acgggctctg caactggcac ttacagcact cctaccgga cctctggcc cagtacactg 1500
 accagcggca ccgaggtac cactaccact tctgctacta ctaccacctc cacctctgtc 1560
 tccaaaacta ccaTaccac caccagtact acctctTgca ccaccccgac ctccgtagcc 1620
 gTcacctTtg acgagattgc aaccacctac tacggcgaga acgTttacat ttctggctcg 1680
 atctcgcagc tcggtagctg ggataccagc agtgccattg cgctcagcgc cagccagTac 1740
 acttccagca acaacctctg gTtctgtgact atcaacctgc ccgctggaac aacctTtcag 1800
 tacaagtaca ttcgcaagga gTcggatggT tcgattgttT gggagagtga tcctaaccgc 1860
 tcgtacactg tgccTtccgg ttgtggTgta agcacagcta ctgagagtga tactTggcga 1920
 tag 1923

<210> 26

<211> 640

<212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

5 <221> SEÑAL

<222> (1)...(21)

<220>

<221> DOMINIO

<222> (42)...(457)

10 <223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

<222> (538)...(634)

<223> Dominio de enlazamiento de almidón

15 <220>

<221> SITIO

<222> (206)...(216)

<223> Firma de la región del sitio activo de glucoamilasa. Id del propositio = PS00820

<220>

20 <221> SITIO

<222> (431)...(434)

<223> Sitio de N-glicosilación. Id del propositio = PS00001

<220>

<221> SITIO

25 <222> (628)...(631)

ES 2 575 912 T3

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 26

Met Thr Ile Ser Arg Leu Ser Ser Val Leu Phe Ala Leu Ala Leu Gly
1 5 10 15
Gln Ser Ala Leu Ala Ala Pro Gln Leu Ser Pro Arg Ala Thr Thr Ser
20 25 30

ES 2 575 912 T3

Leu Asp Ala Trp Leu Ala Ser Glu Thr Thr Val Ser Leu Asn Gly Ile
 35 40 45
 Leu Asp Asn Ile Gly Ala Ser Gly Ala Tyr Ala Gln Ser Ala Lys Ala
 50 55 60
 Gly Val Val Ile Ala Ser Pro Ser Thr Ser Ser Pro Asp Tyr Tyr Tyr
 65 70 75 80
 Thr Trp Thr Arg Asp Ser Ala Leu Thr Leu Lys Val Leu Ile Asp Leu
 85 90
 Phe Arg Asn Gly Asn Val Asp Leu Gln Thr Val Ile Glu Glu Tyr Ile
 100 105 110
 Thr Ala Gln Ala Tyr Leu Gln Thr Val Ser Asn Pro Ser Gly Asp Leu
 115 120 125
 Ser Ser Gly Ala Gly Leu Ala Glu Pro Lys Phe Asn Val Asp Met Ser
 130 135 140
 Ala Tyr Thr Gly Ala Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu
 145 150 155 160
 Arg Ala Ile Ala Leu Ile Asp Phe Gly Asn Trp Leu Ile Asp Asn Gly
 165 170 175
 Tyr Ser Ser Tyr Ala Val Ser Asn Val Trp Pro Ile Val Arg Asn Asp
 180 185 190
 Leu Ser Tyr Val Ala Gln Tyr Trp Ser Gln Ser Gly Tyr Asp Leu Trp
 195 200 205
 Glu Glu Val Asn Ser Met Ser Phe Phe Thr Ile Ala Asn Gln His Arg
 210 215 220
 Ala Leu Val Glu Gly Ser Thr Phe Ala Gly Arg Val Gly Ala Ser Cys
 225 230 235 240
 Ser Trp Cys Asp Ser Gln Ala Pro Gln Ile Leu Cys Tyr Met Gln Asn
 245 250 255
 Phe Trp Thr Gly Ser Tyr Ile Asn Ala Asn Thr Gly Gly Gly Arg Ser
 260 265 270
 Gly Lys Asp Ala Asn Thr Val Leu Ala Ser Ile Ser Thr Phe Asp Pro
 275 280 285
 Glu Ala Thr Cys Asp Asp Val Thr Phe Gln Pro Cys Ser Ser Arg Ala
 290 295 300

ES 2 575 912 T3

Leu Ala Asn His Lys Val Tyr Thr Asp Ser Phe Arg Ser Val Tyr Ser
 305 310 315 320
 Leu Asp Ser Gly Ile Ala Glu Gly Val Ala Val Ala Val Gly Arg Tyr
 325 330 335
 Pro Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Thr Thr Leu
 340 345 350
 Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile Tyr Gln Trp Asn Lys Ile
 355 360 365
 Gly Ser Ile Thr Ile Thr Ser Thr Ser Leu Ala Phe Phe Asn Asp Val
 370 375 380
 Tyr Ser Ser Ala Ala Val Gly Thr Tyr Ala Ser Gly Ser Thr Ala Tyr
 385 390 395 400
 Thr Ala Ile Val Ser Ala Val Lys Thr Tyr Ala Asp Gly Tyr Val Ser
 405 410 415
 Ile Val Gln Ala His Ala Met Thr Asn Gly Ser Leu Ser Glu Gln Phe
 420 425 430
 Asp Lys Ala Ser Gly Thr Gln Leu Ser Ala Arg Asp Leu Thr Trp Ser
 435 440 445
 Tyr Ala Ala Leu Leu Thr Ala Asn Met Arg Arg Asn Gly Ile Val Pro
 450 455 460
 Pro Ser Trp Gly Ala Ala Ser Ala Asn Ser Ile Pro Ser Ser Cys Ser
 465 470 475 480
 Thr Gly Ser Ala Thr Gly Thr Tyr Ser Thr Pro Thr Gly Thr Ser Trp
 485 490 495
 Pro Ser Thr Leu Thr Ser Gly Thr Ala Gly Thr Thr Thr Thr Ser Ala
 500 505 510
 Thr Thr Thr Thr Ser Thr Ser Val Ser Lys Thr Thr Thr Thr Thr
 515 520 525
 Ser Thr Thr Ser Cys Thr Thr Pro Thr Ser Val Ala Val Thr Phe Asp
 530 535 540
 Glu Ile Ala Thr Thr Tyr Tyr Gly Glu Asn Val Tyr Ile Ser Gly Ser
 545 550 555 560
 Ile Ser Gln Leu Gly Ser Trp Asp Thr Ser Ser Ala Ile Ala Leu Ser
 565 570 575
 Ala Ser Gln Tyr Thr Ser Ser Asn Asn Leu Trp Phe Val Thr Ile Asn

ES 2 575 912 T3

580

585

590

Leu Pro Ala Gly Thr Thr Phe Gln Tyr Lys Tyr Ile Arg Lys Glu Ser
595 600 605

Asp Gly Ser Ile Val Trp Glu Ser Asp Pro Asn Arg Ser Tyr Thr Val
610 615 620

Pro Ser Gly Cys Gly Val Ser Thr Ala Thr Glu Ser Asp Thr Trp Arg
625 630 635 640

<210> 27

<211> 1911

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 27

ES 2 575 912 T3

atgtatatcc	tgtcttctgc	ttttctcctc	ggctccctcg	ctcttcagag	cgttctgggt	60
agaccggcgg	ataagatcca	agtgagacaa	tctggtatcg	aggactttat	caagtccgag	120
agcccgattg	ccatcgagca	gcttctgtgc	aacattgggt	ccgaagggtg	ccatgccaag	180
aatgttcccc	ctggcatcgt	cattgcttct	ccagacaccc	aagatcccga	ctacttttat	240
acctggactc	gagatgctgc	tcttgttttc	aagtatatg	tcgacagggt	cattcaccag	300
tacgatgcta	ctctccagaa	gcgcattcag	gagtacattg	cctctcaagc	caagctccag	360
ggcgtctcca	accctctcgg	atccctctcg	gatggctctg	gccttggtga	agccaagtac	420
tacgtcaacc	tgagccccta	cactgggtggc	tggggtcgac	ctcaacgtga	tggccctgcc	480
cttcgagcca	ccgccatgat	cacttatgcc	aactggctga	tcgctaacgg	ttacacctct	540
acggctaagt	acattgtgtg	gcccgtgatt	cgcaacgatc	tcaactacgt	tgctcagtat	600
tggactcaaa	ctggctttga	cttggtggaa	gaggtcaggg	gtagctcatt	cttcacgact	660
gctgctcaat	accgagctct	cgtcgaaggt	gctgcccttg	ctaaggctct	tggcaagtct	720
ggcgatacct	attctaacat	tgcgccccta	acactctgct	tcttgagac	ctactgggtc	780
tcgaacggta	gatacgtcga	ctccaacatt	aatgtgaacg	atggccgcac	cggaaggac	840
gccaacagca	tcctcgcac	tatccacaac	tttgacccaa	gcattggctg	tgatgctgca	900
acattccaac	cctgcagtga	caaggccctc	gccaacccca	aggcagtcac	cgacaccttc	960
cgctcctaca	acctcaacaa	gggcatcgca	cagggaaactg	ctgtggctat	cggaagatac	1020
attgaagatg	tctactacaa	cggcaacccc	tggtacctca	ccacactcgc	cgctgccgag	1080
cagctatatg	atgccgtcta	cgctctggaag	cagaagggat	ccatcactgt	gactgataca	1140
tctctgtcgt	tcttcaagga	tcttggtgctg	agtgtttcca	ctggcacata	tgccagcggc	1200
tcgaccactt	tccagcagat	tatcgacgcc	gtgtcgacct	acgccgacgg	atacgttgcc	1260
attgttcgca	agtacgtggg	tccgaacggc	gccctggctg	agcagttctc	caaggacaat	1320
ggcactccca	tgtccgctga	tgatcttacg	tggtcatatg	ctgctttcct	ctcagccact	1380
gaacgccgag	ctggcattgt	tcctcctact	tggcaaaaga	gtgttcctgc	agttcccaac	1440
agctgtggat	caagcacagt	cgttggatca	tacacatcgg	ccacgcagac	ctcgttcctt	1500
ccttcgcaaa	ccccccagga	cggcgtgcca	actcccacag	gaccaactcc	caccgacgga	1560
ggacctactt	cctctcctac	aagctgtgcc	atcgccacat	ctgtcgacgt	cacctttaat	1620
gaggttgtca	agaccgagta	cggcgcacacc	atcaagattg	tcggcagcat	cgctgccctg	1680
ggcagctggg	acaccacgaa	ggccatctcc	ctgagcgcct	ccgactacac	ggcctcgaac	1740
cccctgtgga	agacgacaat	ctccctcacg	gctggtcagt	cttttgagta	caagtatatc	1800
aacatcaaga	aggacggctc	gctgggtgtg	gagcgtgacc	ccaaccgttc	ttatactggt	1860
cccaagactt	gtgagaccaa	ggctaccaag	tcggatagtt	ggcaaggata	a	1911

<210> 28

<211> 636

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

10 <222> (1)...(20)

<220>

ES 2 575 912 T3

<221> DOMINIO

<222> (42)...(461)

<223> Familia de glicosil hidrolasas 15

<220>

5 <221> DOMINIO

<222> (534)...(630)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

10 <222> (207)...(217)

<223> Firma de la región del sitio activo de glucoamilasa. Id del pro sitio = PS00820

<220>

<221> SITIO

<222> (624)...(627)

15 <223> Sitio de N-glicosilación. Id del pro sitio = PS00001

<400> 28

Met Tyr Ile Leu Ser Ser Ala Phe Leu Leu Gly Ser Leu Ala Leu Gln
1 5 10 15

Ser Val Leu Gly Arg Pro Ala Asp Lys Ile Gln Val Arg Gln Ser Gly
20 25 30

Ile Glu Asp Phe Ile Lys Ser Glu Ser Pro Ile Ala Ile Glu Gln Leu
35 40 45

Leu Cys Asn Ile Gly Ser Glu Gly Cys His Ala Lys Asn Val Pro Ala
50 55 60

Gly Ile Val Ile Ala Ser Pro Asp Thr Gln Asp Pro Asp Tyr Phe Tyr

ES 2 575 912 T3

Leu Thr Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Val Tyr Val
 355 360 365
 Trp Lys Gln Lys Gly Ser Ile Thr Val Thr Asp Thr Ser Leu Ser Phe
 370 375 380
 Phe Lys Asp Leu Val Ser Ser Val Ser Thr Gly Thr Tyr Ala Ser Gly
 385 390 395 400
 Ser Thr Thr Phe Gln Gln Ile Ile Asp Ala Val Ser Thr Tyr Ala Asp
 405 410 415
 Gly Tyr Val Ala Ile Val Arg Lys Tyr Val Gly Pro Asn Gly Ala Leu
 420 425 430
 Ala Glu Gln Phe Ser Lys Asp Asn Gly Thr Pro Met Ser Ala Asp Asp
 435 440 445
 Leu Thr Trp Ser Tyr Ala Ala Phe Leu Ser Ala Thr Glu Arg Arg Ala
 450 455 460
 Gly Ile Val Pro Pro Thr Trp Gln Lys Ser Val Pro Ala Val Pro Asn
 465 470 475 480
 Ser Cys Gly Ser Ser Thr Val Val Gly Ser Tyr Thr Ser Ala Thr Gln
 485 490 495
 Thr Ser Phe Pro Pro Ser Gln Thr Pro Gln Asp Gly Val Pro Thr Pro
 500 505 510
 Thr Gly Pro Thr Pro Thr Asp Gly Gly Pro Thr Ser Ser Pro Thr Ser
 515 520 525
 Cys Ala Ile Ala Thr Ser Val Asp Val Thr Phe Asn Glu Val Val Lys
 530 535 540
 Thr Glu Tyr Gly Asp Thr Ile Lys Ile Val Gly Ser Ile Ala Ala Leu
 545 550 555 560
 Gly Ser Trp Asp Thr Thr Lys Ala Ile Ser Leu Ser Ala Ser Asp Tyr
 565 570 575
 Thr Ala Ser Asn Pro Leu Trp Lys Thr Thr Ile Ser Leu Thr Ala Gly
 580 585 590
 Gln Ser Phe Glu Tyr Lys Tyr Ile Asn Ile Lys Lys Asp Gly Ser Leu
 595 600 605
 Val Trp Glu Arg Asp Pro Asn Arg Ser Tyr Thr Val Pro Lys Thr Cys
 610 615 620

ES 2 575 912 T3

Glu Thr Lys Ala Thr Lys Ser Asp Ser Trp Gln Gly
625 630 635

<210> 29

<211> 1767

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 29

ES 2 575 912 T3

atgcttttct cgctgcttct tcgtgccttg tcagcctcgc ttttggcggg cgctgtccag 60
 ggctcaccgg ccaatctcca tgcgcgcaa acctctgtcg accaattcat tcaaaccag 120
 gccgaaatat ccatcaaggg ggttctagcg aacatcggtg cagatggatc caaagcccaa 180
 ggagcccctg ccgggatcgt cgttgcaagt ccatctcgtt ctaatccaga ctattggtat 240
 acctggacca gagacgcagc attgacatac aaggccctca ttgagcgctt tgtc gatggt 300
 gatacttctc tccgcacaaa ggtcgcagag tatgtctcag cgcaagcata cctgcaaggc 360
 gtgtccaatc catcaggggg gcccgactct ggaggcctag gagagcccaa attcaacgta 420
 gaccgcaccg ctttcacagg agcatggggg cgaccacagc gtgatgggccc tccgcttcgt 480
 gctacagctc tgatcatcta cgcaaactgg ctggttgcaa acgggcagca gagtcaagca 540
 ctcaaacctg tgtggccggg cattgcaaag gacttggcgt acaccgttcg atactggaac 600
 cagacaggct ttgatctgtg ggaagaagta aacggatcat ctttcttcac gctgtctcgc 660
 tcccacagag cgcttgtcga gggcaatgca cttgctcaga aactcggaca aacttgacc 720
 ggatgcgcgga acgctgcacc ccaggttctc tgctttgtgc aaagcttttg gacaggcagc 780
 tatattgatt ctaatatcaa tgtaacgat gggcgtacgg gcaaggatgc aaactcaatt 840
 ctgtcatcta tccatacgtt tgacccggag gctggatgca ctgatgccac attccaacct 900
 tgttcttccc gtgtcttgc aaaccacaag gcagtcactg actctttccg atccgtgtat 960
 ggcatcaacc gtggaattgc tcaaggcagt gcagtggccg tgggcaggta ctctgaagat 1020
 gtgtattaca acggcaacc atggtatctt accactctcg ccgctgccga gcaactatac 1080
 tctgcgatct atcagtggaa taagaaaggc tccatcaccg tagatgcagt atccctcccc 1140
 ttcttccgcy atctagtacc atccatcgcg acaggaacat atgccagcag ctcgtcaacc 1200
 tacacatcca tcgtctcagc agtaaagacg tacgcagacg gttacatcag cgtgatccag 1260
 aaatacacgc cctcgaacgg cgcgctcgcg gagcaattcg aaaagagcaa tggatcacca 1320
 ctgtccgccc cagacctcac ctggtcctac gccgcgttcc tcaccgcaac agagcggcgc 1380
 gcaggcgtcg taagtccac ctggggcgaa cccaccaaca acgttcccc atcaacctgc 1440
 actggcacac ccgcctgcaa cgcgcgcac acgttcaacg tgcgcgccac caccacgttc 1500
 ggcgacaaca tcttcatcgt ggggcagctg acacagctgg gtaattggga tcctgcgagc 1560
 gccgtgccgc tcagtgccag caagtatact agcagcgcac cgttgtggta tgccgatatc 1620
 aatctcccgg ccgagacaac gtttgagtac aagtatattc gcaagacgag tgcagggcag 1680

gtggtgtggg agagtgatcc gaatagacgg tatacgcgt ctgcgggggtg tggaagtagt 1740
 gcgacggtga atgattcttg gaggtag 1767

<210> 30

<211> 588

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(21)

<220>

<221> DOMINIO

5 <222> (42)...(458)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

<222> (487)...(583)

10 <223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (203)...(206)

<223> Sitio de N-glicosilación. Id del propositio = PS00001

15 <220>

<221> SITIO

<222> (205)...(215)

<223> Firma de la región del sitio activo de glucoamilasa. Id del propositio = PS00820

<220>

20 <221> SITIO

<222> (214)...(217)

<223> Sitio de N-glicosilación. Id del propositio = PS00001

<400> 30

ES 2 575 912 T3

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

ES 2 575 912 T3

Ser Ala Gln Ala Tyr Leu Gln Gly Val Ser Asn Pro Ser Gly Gly Pro
 115 120 125

Asp Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Arg Thr Ala
 130 135 140

Phe Thr Gly Ala Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro Leu Arg
 145 150 155 160

Ala Thr Ala Leu Ile Ile Tyr Ala Asn Trp Leu Val Ala Asn Gly Gln
 165 170 175

Gln Ser Gln Ala Leu Asn Thr Val Trp Pro Val Ile Ala Lys Asp Leu
 180 185 190

Ala Tyr Thr Val Arg Tyr Trp Asn Gln Thr Gly Phe Asp Leu Trp Glu
 195 200 205

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

Leu Val Glu Gly Asn Ala Leu Ala Gln Lys Leu Gly Gln Thr Cys Thr
 225 230 235 240

Gly Cys Ala Asn Ala Ala Pro Gln Val Leu Cys Phe Val Gln Ser Phe
 245 250 255

Trp Thr Gly Ser Tyr Ile Asp Ser Asn Ile Asn Val Asn Asp Gly Arg
 260 265 270

Thr Gly Lys Asp Ala Asn Ser Ile Leu Ser Ser Ile His Thr Phe Asp
 275 280 285

Pro Glu Ala Gly Cys Thr Asp Ala Thr Phe Gln Pro Cys Ser Ser Arg
 290 295 300

Ala Leu Ala Asn His Lys Ala Val Thr Asp Ser Phe Arg Ser Val Tyr
 305 310 315 320

Gly Ile Asn Arg Gly Ile Ala Gln Gly Ser Ala Val Ala Val Gly Arg
 325 330 335

Tyr Ser Glu Asp Val Tyr Tyr Asn Gly Asn Pro Trp Tyr Leu Thr Thr
 340 345 350

Leu Ala Ala Ala Glu Gln Leu Tyr Ser Ala Ile Tyr Gln Trp Asn Lys
 355 360 365

Lys Gly Ser Ile Thr Val Asp Ala Val Ser Leu Pro Phe Phe Arg Asp
 370 375 380

ES 2 575 912 T3

Leu Val Pro Ser Ile Ala Thr Gly Thr Tyr Ala Ser Ser Ser Thr
385 390 395 400

Tyr Thr Ser Ile Val Ser Ala Val Lys Thr Tyr Ala Asp Gly Tyr Ile
405 410 415

Ser Val Ile Gln Lys Tyr Thr Pro Ser Asn Gly Ala Leu Ala Glu Gln
420 425 430

Phe Glu Lys Ser Asn Gly Ser Pro Leu Ser Ala Ala Asp Leu Thr Trp
435 440 445

Ser Tyr Ala Ala Phe Leu Thr Ala Thr Glu Arg Arg Ala Gly Val Val
450 455 460

Ser Pro Thr Trp Gly Glu Pro Thr Asn Asn Val Pro Pro Ser Thr Cys
465 470 475 480

Thr Gly Thr Pro Ala Cys Asn Ala Arg Ile Thr Phe Asn Val Arg Ala
485 490 495

Thr Thr Thr Phe Gly Asp Asn Ile Phe Ile Val Gly Gln Leu Thr Gln
500 505 510

Leu Gly Asn Trp Asp Pro Ala Ser Ala Val Pro Leu Ser Ala Ser Lys
515 520 525

Tyr Thr Ser Ser Asp Pro Leu Trp Tyr Ala Asp Ile Asn Leu Pro Ala
530 535 540

Glu Thr Thr Phe Glu Tyr Lys Tyr Ile Arg Lys Thr Ser Ala Gly Gln
545 550 555 560

Val Val Trp Glu Ser Asp Pro Asn Arg Arg Tyr Thr Thr Ser Ala Gly
565 570 575

Cys Gly Ser Ser Ala Thr Val Asn Asp Ser Trp Arg
580 585

<210> 31

<211> 1479

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

ES 2 575 912 T3

<400> 31

```

atggttctag ctcgcttgc ctggcttgc ggcctggtta gcactgccat tgctgcgacc 60
ccagcggaat ggcgctcga gtccatctac ttcatgctca cggatcgttt tgcccggacc 120
gatggttcaa ctactgctgc ctgtgatacc gctgacagaa aatactgtgg cggaacctgg 180
cagggatca ttgacaagct ggactatata caaggaatgg gctttacggc catttggatc 240

actccggtga ccggtcaatt gagcgggaa accgctacg gagatcccta tcacggatac 300
tggcagcagg atatctactc tctcgattcc aactatggaa ccgcagacga tctcaaggcc 360
ctcgtgcgg ctttgcaaa acgcgacatg tatctcatgg tcgatgtcgt agcaaaccac 420
atgggctaca atggcgagg tgctgacgtg gactacacca aattcaacc cttcaacgat 480
gcaaagtatt tccactccta ctgcccattc accgattaca acgacgacac catgtcgcga 540
aactgctggc ttggcgataa caaggctctg ctaccggatc tgaatacaca gagcaaggag 600
gtgcaggatc tatggtatga ctgggttga tctttggtct ccaactactc catcgatgga 660
cttcgctgc acacagtcaa acatgtccag aaagatttct ggcccggcta caacaaagcc 720
gccccgctc actgcgtagg cgaaatcctt gatggtgacc cagattacac ctatccatac 780
caggaggtaa tggacggagt gctcaactac ccgatttact acccactcct caaggccttc 840
cagtcgagct cgggaagcat gaccgatctg tacaacatga tcaacacggt gaaatcgacc 900
tgcaaggact caacccttct tggaaattc ttggagaacc acgataacc acgttttgcc 960
catgccaccg atgacattgc cctcgccaag aacgcagcca catttaccat tatggcagat 1020
ggcattccta ttgtctatgc aggacaggag cagcactaca gtggtggcga ggaccggct 1080
aatcgcgagg ctctgtggtt atccggatac aacaccgaca gcgagctgta caagctcatt 1140
gccaggcca atggtgctag aaaccaggcc attgctaaga gtaccaatta tactatttac 1200
cagaaccacc caatctaaa agacgagagc gccatcgcca tgcggaagg cttcgtcggc 1260
ggacagacaa tcaactgtct gacgaatctc ggtgcagggg gtaaagagta ttcagtttca 1320
attcctgata ctggattcaa ggctggtgag aagttgactg aggttgtgct ctgcactagt 1380
gttactgttg gtgatagtg ggaggtgtct gttcctatgg cgagtggagc gccgaggatc 1440
ttgctcccca cgtctttgct cgagggctcg gctctgtgc 1479

```

<210> 32

<211> 493

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

10 <221> SEÑAL

<222> (1)...(18)

<220>

<221> DOMINIO

<222> (31)...(387)

<223> Alfa amilasa, dominio catalítico

5 <220>

<221> SITIO

<222> (197)...(200)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

10 <221> SITIO

<222> (218)...(221)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

15 <222> (402)...(405)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 32

ES 2 575 912 T3

Met Val Leu Ala Arg Leu Ala Trp Leu Ala Gly Leu Val Ser Thr Ala
1 5 10 15

Ile Ala Ala Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile Tyr Phe Met
20 25 30

Leu Thr Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Ala Ala Cys
35 40 45

Asp Thr Ala Asp Arg Lys Tyr Cys Gly Gly Thr Trp Gln Gly Ile Ile
50 55 60

Asp Lys Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile Trp Ile
65 70 75 80

Thr Pro Val Thr Gly Gln Leu Ser Gly Glu Thr Ala Tyr Gly Asp Pro
85 90 95

Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Ser Leu Asp Ser Asn Tyr
100 105 110

Gly Thr Ala Asp Asp Leu Lys Ala Leu Ala Ala Ala Leu His Lys Arg
115 120 125

Asp Met Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly Tyr Asn
130 135 140

Gly Ala Gly Ala Asp Val Asp Tyr Thr Lys Phe Asn Pro Phe Asn Asp
145 150 155 160

Ala Lys Tyr Phe His Ser Tyr Cys Pro Ile Thr Asp Tyr Asn Asp Asp
165 170 175

Thr Met Ser Gln Asn Cys Trp Leu Gly Asp Asn Lys Val Ser Leu Pro
180 185 190

Asp Leu Asn Thr Gln Ser Lys Glu Val Gln Asp Leu Trp Tyr Asp Trp
195 200 205

Val Gly Ser Leu Val Ser Asn Tyr Ser Ile Asp Gly Leu Arg Val Asp
210 215 220

Thr Val Lys His Val Gln Lys Asp Phe Trp Pro Gly Tyr Asn Lys Ala
225 230 235 240

Ala Gly Val Tyr Cys Val Gly Glu Ile Leu Asp Gly Asp Pro Asp Tyr
245 250 255

ES 2 575 912 T3

Thr Tyr Pro Tyr₂₆₀ Gln Glu Val Met Asp₂₆₅ Gly Val Leu Asn Tyr₂₇₀ Pro Ile
 Tyr Tyr Pro₂₇₅ Leu Leu Lys Ala Phe₂₈₀ Gln Ser Ser Ser Gly₂₈₅ Ser Met Thr
 Asp Leu₂₉₀ Tyr Asn Met Ile Asn₂₉₅ Thr Val Lys Ser Thr₃₀₀ Cys Lys Asp Ser
 Thr₃₀₅ Leu Leu Gly Asn Phe₃₁₀ Leu Glu Asn His Asp₃₁₅ Asn Pro Arg Phe Ala₃₂₀
 His Ala Thr Asp Asp₃₂₅ Ile Ala Leu Ala Lys₃₃₀ Asn Ala Ala Thr Phe₃₃₅ Thr
 Ile Met Ala Asp₃₄₀ Gly Ile Pro Ile Val₃₄₅ Tyr Ala Gly Gln Glu₃₅₀ Gln His
 Tyr Ser Gly₃₅₅ Gly Glu Asp Pro Ala₃₆₀ Asn Arg Glu Ala Leu₃₆₅ Trp Leu Ser
 Gly Tyr₃₇₀ Asn Thr Asp Ser Glu₃₇₅ Leu Tyr Lys Leu Ile₃₈₀ Ala Lys Ala Asn
 Gly Ala Arg Asn Gln Ala₃₉₀ Ile Ala Lys Ser Thr₃₉₅ Asn Tyr Thr Ile Tyr₄₀₀
 Gln Asn His Pro Ile₄₀₅ Tyr Lys Asp Glu Ser₄₁₀ Ala Ile Ala Met Arg₄₁₅ Lys
 Gly Phe Val Gly₄₂₀ Gly Gln Thr Ile Thr₄₂₅ Val Leu Thr Asn Leu₄₃₀ Gly Ala
 Gly Gly Lys₄₃₅ Glu Tyr Ser Val Ser₄₄₀ Ile Pro Asp Thr Gly₄₄₅ Phe Lys Ala
 Gly Ala₄₅₀ Lys Leu Thr Glu Val₄₅₅ Val Ser Cys Thr Ser₄₆₀ Val Thr Val Gly
 Asp₄₆₅ Ser Gly Glu Val Ser₄₇₀ Val Pro Met Ala Ser₄₇₅ Gly Ala Pro Arg Ile₄₈₀
 Leu Leu Pro Thr Ser₄₈₅ Leu Leu Glu Gly Ser₄₉₀ Ala Leu Cys

<210> 33

<211> 1485

<212> ADN

5 <213> Desconocido

ES 2 575 912 T3

<220>

<223> obtenido a partir de una muestra ambiental

<400> 33

```

atggtaggct tcaatatctt gactctcgcc ttgctggccc cagctgcact ctcatctgct      60
gtccaccctc accggcgcca gtctggactt gatgccttta ttcagtccga gtcacgggtt      120
tcaactccaag gcattctgaa caatatcgga gccaatggct ctgctgtttc aggcgcatcc      180
gccgggtgtag ttgttgcgtc accatccaag tcagaccccg attatTTTTta cacttggact      240
cgcgatgcbg ccctgacact cgcggtactg attgaccaat tcatcgcbgg agaaagctcc      300
ctagagacgg tcatccagca gtacatctcc gcccaagcaa aactgcagac agtgtccaac      360
ccatcaggag acctgtccga cggttcaggt ctggcagagc ccaaattcca aaccgatctc      420
agcgccttca ctagatcctg gggctgcacct caacgcbgatg gacctgctct ccgtgcbgaca      480
gcbctcatcg tgtacggaaa tcacctctc tcttccggaa agcagtcctg cgtcaaatcc      540
aacatctggc cgatcgtgca gaacgatctc aactacgtgg cacagtactg gaaccagact      600
ggtttcgatc tgtgggaaga ggtccagggc tcgtcctttt tcaccatcgc agcbgaacac      660
cgtgcbcttg tggagggagc tgcattcgcc aaatcgbctg gagaggbctg tgatggctgt      720
gattcacagg ctccgcaggt tctttgcttc ctccaggact tctggaatgg aaatgctgtt      780
gtttctaatc tggcaaatga tggctgctca gggttggatg ccaattctat cctcagctcg      840
atccagacct tcgatcctag tgccacttgc gatgatagca cttccagcc atgctctggg      900
cgtgcactgc taaaccacaa ggcggtgggt gattctttcc gatccatcta caacattaac      960
agtggtaaag aagctggcaa agccgtggct gttggctgct atgccgaaga cacctaccag     1020
ggaggcaacc cttggtatct cgctactgct gctgcbgcbg agcagctgta tgatgctctc     1080
taccagtgga agaaacaggg atcgbctggcg attaccaga cbagctcttc cttcttccag     1140
gatttggact cgactgctcg tgttggcaat tactccagct cttcctcaac ctacacctct     1200
ttgactggcg ctgtcaagac ctatgcbgat ggtttcctat cbattgttca gcagtatacb     1260
cccagcaatg gagctctggc cbagcbgttc acgcbgagaca atggcacccc cgtctctgca     1320
cacgacctga cctggtctta tgcacatctt ttgactgctg cbgagcbctg caatggaatt     1380
gtccctgcta gctggggtgc atctaaggcc aaccaagtgc cbagcbagtg ccagggcagtg     1440
tcagcbagcbg ggtcttatac tacgccaact gtggggctct ggtga                       1485

```

5 <210> 34

<211> 494

<212> PRT

<213> Desconocido

<220>

10 <223> Obtenido a partir de una muestra ambiental

<220>

ES 2 575 912 T3

<221> DOMINIO

<222> (39)...(455)

<223> Familia de glicosil hidrolasas 15

<220>

5 <221> SITIO

<222> (52)...(55)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

10 <222> (201)...(204)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (203)...(213)

15 <223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<400> 34

ES 2 575 912 T3

Met Val Gly Phe Asn Ile Leu Thr Leu Ala Leu Leu Ala Pro Ala Ala
 1 5 10 15
 Leu Ser Ser Ala Val His Pro His Arg Arg Gln Ser Gly Leu Asp Ala
 20 25 30
 Phe Ile Gln Ser Glu Ser Ser Val Ser Leu Gln Gly Ile Leu Asn Asn
 35 40 45
 Ile Gly Ala Asn Gly Ser Ala Val Ser Gly Ala Ser Ala Gly Val Val
 50 55 60
 Val Ala Ser Pro Ser Lys Ser Asp Pro Asp Tyr Phe Tyr Thr Trp Thr
 65 70 75 80
 Arg Asp Ala Ala Leu Thr Leu Ala Val Leu Ile Asp Gln Phe Ile Ala
 85 90 95
 Gly Glu Ser Ser Leu Glu Thr Val Ile Gln Gln Tyr Ile Ser Ala Gln
 100 105 110
 Ala Lys Leu Gln Thr Val Ser Asn Pro Ser Gly Asp Leu Ser Asp Gly
 115 120 125
 Ser Gly Leu Ala Glu Pro Lys Phe Gln Thr Asp Leu Ser Ala Phe Thr
 130 135 140
 Arg Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala Thr
 145 150 155 160
 Ala Leu Ile Val Tyr Gly Asn His Leu Leu Ser Ser Gly Lys Gln Ser
 165 170 175
 Val Val Lys Ser Asn Ile Trp Pro Ile Val Gln Asn Asp Leu Asn Tyr
 180 185 190
 Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe Asp Leu Trp Glu Glu Val
 195 200 205
 Gln Gly Ser Ser Phe Phe Thr Ile Ala Ala Gln His Arg Ala Leu Val
 210 215 220
 Glu Gly Ala Ala Phe Ala Lys Ser Leu Gly Glu Ala Cys Asp Gly Cys

ES 2 575 912 T3

225					230						235				240
Asp	Ser	Gln	Ala	Pro	Gln	Val	Leu	Cys	Phe	Leu	Gln	Asp	Phe	Trp	Asn
				245					250					255	
Gly	Asn	Ala	Val	Val	Ser	Asn	Leu	Ala	Asn	Asp	Gly	Arg	Ser	Gly	Leu
			260					265					270		
Asp	Ala	Asn	Ser	Ile	Leu	Ser	Ser	Ile	Gln	Thr	Phe	Asp	Pro	Ser	Ala
		275					280					285			
Thr	Cys	Asp	Asp	Ser	Thr	Phe	Gln	Pro	Cys	Ser	Gly	Arg	Ala	Leu	Leu
	290					295					300				
Asn	His	Lys	Ala	Val	Val	Asp	Ser	Phe	Arg	Ser	Ile	Tyr	Asn	Ile	Asn
305					310					315					320
Ser	Gly	Lys	Glu	Ala	Gly	Lys	Ala	Val	Ala	Val	Gly	Arg	Tyr	Ala	Glu
				325					330					335	
Asp	Thr	Tyr	Gln	Gly	Gly	Asn	Pro	Trp	Tyr	Leu	Ala	Thr	Ala	Ala	Ala
			340					345					350		
Ala	Glu	Gln	Leu	Tyr	Asp	Ala	Leu	Tyr	Gln	Trp	Lys	Lys	Gln	Gly	Ser
		355					360					365			
Leu	Ala	Ile	Thr	Gln	Thr	Ser	Leu	Pro	Phe	Phe	Gln	Asp	Leu	Asp	Ser
	370					375					380				
Thr	Ala	Arg	Val	Gly	Asn	Tyr	Ser	Ser	Ser	Ser	Ser	Thr	Tyr	Thr	Ser
385					390					395					400
Leu	Thr	Gly	Ala	Val	Lys	Thr	Tyr	Ala	Asp	Gly	Phe	Leu	Ser	Ile	Val
				405					410					415	
Gln	Gln	Tyr	Thr	Pro	Ser	Asn	Gly	Ala	Leu	Ala	Glu	Gln	Phe	Thr	Arg
			420					425					430		
Asp	Asn	Gly	Thr	Pro	Val	Ser	Ala	His	Asp	Leu	Thr	Trp	Ser	Tyr	Ala
		435					440					445			
Ser	Phe	Leu	Thr	Ala	Ala	Asp	Arg	Arg	Asn	Gly	Ile	Val	Pro	Ala	Ser
	450					455					460				
Trp	Gly	Ala	Ser	Lys	Ala	Asn	Gln	Val	Pro	Thr	Gln	Cys	Gln	Gly	Ser
465					470					475					480
Ser	Ala	Thr	Gly	Ser	Tyr	Thr	Thr	Pro	Thr	Val	Gly	Ser	Trp		
				485					490						

ES 2 575 912 T3

<210> 35

<211> 1914

<212> ADN

<213> Desconocido

5 <220>

<223> Obtenido a partir de una muestra ambiental

<400> 35

ES 2 575 912 T3

atggctctc gtttttgac tacgctgtgt gcgctgactc ttggttcagc tgcgctggcc 60
 gcgccacagt tggcaccgcg tgcaactggt agcttggaca cttggctggc atccgagaaa 120
 gctgttgcaa gggagggtat tctcgataat atcgggctcg atggggcgta tgctaagagc 180
 tcgaagtcgg ggattgtcat tgccagtccg agtacagaca gtcccgatta ttactacact 240
 tggactcgtg actcggccct ggtcatgaag accttgggtg acctgttcaa gaatggcgat 300
 accgaccttc tcaccgttat cgaggagtac atcagctccc aggcctatat ccagaccggt 360
 tccaacccat ctggaagtct ttcgagtggg ggactcgggtg aaccaagt caatgttgat 420
 gaaacatcgt ataccggatc ttggggtcga cctcagcgcg atggctctgc actccgagct 480
 actgcactgg ttgccttcgg acagtgggtg attgacaacg gatatacatc taccgccacc 540
 gatattgtgt ggccaattgt tcgcaatgat ctttcctacg tcgcacaata ctggaacagc 600
 tctggatacg acctatggga agagggtcaac ggtgaatcat tcttcaccac cgccgttcaa 660
 caccgtgctt tggctcgggg cagcaagttc gccagccagg tgggatcatc ttgctcctac 720
 tgcgactctc aggacccca ggttctctgc ttccttcaat catactggac cggctcgtac 780
 actctagcca atttcggtag cagccgcaca ggcaaagacg cgaacaccct cctgggcagc 840
 attcacacat tcgaccccg ggcaggatgc gatgacacga ccttcagcc ttgctcggcc 900
 cgcgcgctgg caaatcacia ggtcgtcact gactcgttcc gctctattta taccgtcaat 960
 tcaggcaaga gtgcgggaca agcagtggca gttggctgat accccgagga ttcttactac 1020
 aatggtaacc catggtactt gtgcaactatg gctgctgccg agcttctcta cgatgctttg 1080
 taccaatgga ataaagctgg ctctttgacc atcagcagcg tctcgtgag cttcttcacc 1140
 gacctgtaca gctcagctgc aactgggtacc tattcatcat ccagtctac gttctcatcg 1200
 atagtcagcg cggggaagac ctacgcggat ggctacatga gcattgccga gcaatatgcc 1260
 tatacgaacg gctctatgtc tgaacaattc tccaagtctg atggaactcc agagtctgcc 1320
 cgtgacctga cttggtcata tgcagccctg cttaccgcaa acatgcgtcg caactcgggt 1380
 gttcccccaa cctggggcga gacatctgca agcagcgtcc cgggaacatg ctccgcgacg 1440
 tcggcgacgg gaacctacag cactgcaacc aacacgaact ggcccgtcac attgactggt 1500
 ggatcgggct ctactactac tgggtggaaca acgaccagca aaaccagcac caccacaact 1560
 tccaagacga gcactaccac tacatcatgc acgactcaa ctagtggtggc agttacattt 1620
 gatgttatcg cgaccacttc atacggcgag aacatcaagt tggctgggtc aattgctgcc 1680
 cttggtagct gggacaccag tagcgctatt gcactaagtg cagataaata tactagctcg 1740
 aacaacctat ggtatgtaac tgtgaatctg gctgctggtc aggtcattca gtacaagtat 1800
 atccggggtg aaagtgatag tacgattgag tgggagagtg atccgaaccg ctcttatact 1860

 gtgccagcag cctgtgccac .aactgccgtg acaatcagcg acacttggcg gtaa . 1914

<210> 36

<211> 637

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(20)

5 <220>

<221> DOMINIO

<222> (41)...(455)

<223> Familia de glicosil hidrolasas 15

<220>

10 <221> DOMINIO

<222> (535)...(631)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

15 <222> (202)...(205)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (204)...(214)

20 <223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<220>

<221> SITIO

<222> (429)...(432)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

25 <220>

ES 2 575 912 T3

<221> SITIO

<222> (625)...(628)

<223> Sitio de N-glicosilación. Id del prosito = PS00001

<400> 36

Met Ala Pro Arg Phe Trp Thr Thr Leu Cys Ala Leu Thr Leu Gly Ser
1 5 10 15

Ala Ala Leu Ala Ala Pro Gln Leu Ala Pro Arg Ala Thr Gly Ser Leu
20 25 30

Asp Thr Trp Leu Ala Ser Glu Lys Ala Val Ala Arg Glu Gly Ile Leu
35 40 45

Asp Asn Ile Gly Ser Asp Gly Ala Tyr Ala Lys Ser Ser Lys Ser Gly
50 55 60

Ile Val Ile Ala Ser Pro Ser Thr Asp Ser Pro Asp Tyr Tyr Tyr Thr
65 70 75 80

Trp Thr Arg Asp Ser Ala Leu Val Met Lys Thr Leu Val Asp Leu Phe
85 90 95

5

ES 2 575 912 T3

Lys Asn Gly Asp Thr Asp Leu Leu Thr Val Ile Glu Glu Tyr Ile Ser.
 100 105 110
 Ser Gln Ala Tyr Ile Gln Thr Val Ser Asn Pro Ser Gly Ser Leu Ser
 115 120 125
 Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu Thr Ser Tyr
 130 135 140
 Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg Ala
 145 150 155 160
 Thr Ala Leu Val Ala Phe Gly Gln Trp Leu Ile Asp Asn Gly Tyr Thr
 165 170 175
 Ser Thr Ala Thr Asp Ile Val Trp Pro Ile Val Arg Asn Asp Leu Ser
 180 185 190
 Tyr Val Ala Gln Tyr Trp Asn Ser Ser Gly Tyr Asp Leu Trp Glu Glu
 195 200 205
 Val Asn Gly Glu Ser Phe Phe Thr Thr Ala Val Gln His Arg Ala Leu
 210 215 220
 Val Glu Gly Ser Lys Phe Ala Ser Gln Val Gly Ser Ser Cys Ser Tyr
 225 230 235 240
 Cys Asp Ser Gln Ala Pro Gln Val Leu Cys Phe Leu Gln Ser Tyr Trp
 245 250 255
 Thr Gly Ser Tyr Thr Leu Ala Asn Phe Gly Ser Ser Arg Thr Gly Lys
 260 265 270
 Asp Ala Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp Pro Glu Ala
 275 280 285
 Gly Cys Asp Asp Thr Thr Phe Gln Pro Cys Ser Ala Arg Ala Leu Ala
 290 295 300
 Asn His Lys Val Val Thr Asp Ser Phe Arg Ser Ile Tyr Thr Val Asn
 305 310 315 320
 Ser Gly Lys Ser Ala Gly Gln Ala Val Ala Val Gly Arg Tyr Pro Glu
 325 330 335
 Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Tyr Leu Cys Thr Met Ala Ala
 340 345 350
 Ala Glu Leu Leu Tyr Asp Ala Leu Tyr Gln Trp Asn Lys Ala Gly Ser
 355 360 365
 Leu Thr Ile Ser Ser Val Ser Leu Ser Phe Phe Thr Asp Leu Tyr Ser

ES 2 575 912 T3

370 . 375 380 .

Ser 385 Ala Ala Thr Gly Thr 390 Tyr Ser Ser Ser Ser 395 Ser Thr Phe Ser Ser 400

Ile Val Ser Ala Val 405 Lys Thr Tyr Ala Asp 410 Gly Tyr Met Ser Ile 415 Ala

Glu Gln Tyr Ala 420 Thr Asn Gly Ser 425 Met Ser Glu Gln Phe Ser Lys 430

Ser Asp Gly 435 Thr Pro Glu Ser Ala 440 Arg Asp Leu Thr Trp 445 Ser Tyr Ala

Ala Leu 450 Leu Thr Ala Asn Met 455 Arg Arg Asn Ser Val 460 Val Pro Pro Thr

Trp 465 Gly Glu Thr Ser Ala 470 Ser Ser Val Pro Gly 475 Thr Cys Ser Ala Thr 480

Ser Ala Thr Gly 485 Thr Tyr Ser Thr Ala Thr 490 Asn Thr Asn Trp Pro Val 495

Thr Leu Thr Gly 500 Gly Ser Gly Ser Thr 505 Thr Thr Gly Gly Thr 510 Thr Thr

Ser Lys Thr 515 Ser Thr Thr Thr Thr 520 Ser Lys Thr Ser Thr 525 Thr Thr Thr

Ser Cys 530 Thr Thr Pro Thr Ser 535 Val Ala Val Thr Phe 540 Asp Val Ile Ala

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

Leu Gly Ser Trp Asp 565 Thr Ser Ser Ala Ile 570 Ala Leu Ser Ala Asp 575 Lys

Tyr Thr Ser Ser 580 Asn Asn Leu Trp Tyr 585 Val Thr Val Asn Leu Ala Ala 590

Gly Gln Val 595 Ile Gln Tyr Lys Tyr 600 Ile Arg Val Glu Ser 605 Asp Ser Thr

Ile Glu 610 Trp Glu Ser Asp Pro 615 Asn Arg Ser Tyr Thr 620 Val Pro Ala Ala

Cys 625 Ala Thr Thr Ala Val 630 Thr Ile Ser Asp Thr 635 Trp Arg

<210> 37

<211> 1923

<212> ADN

5 <213> Desconocido

ES 2 575 912 T3

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 37

```

atggctcctc gattctggat tgcgcttgg gcgcttactt tcggtcaggc gatcgctgcg      60
ccgcagatcg ccttggcacc tcgcgcgaca ggtagcttgg acacctggtt ggcacatcgag    120
accacggtgg ccagacaagg taccctgaac aatatcggtt ccgctggcgc gtattctgcg    180
agcgcgaagc cgggaatcct gattgctagt cctagcactt ccagtcccga ttattactac    240
acatggaccc gtgactctgc tttggctctc aaggccttgg tggatatggt caagaatggt    300
gacactgccc ttctcactgt gatcgaagaa tatatcagcg cacaggccta tatccagaca    360
gtctccaacc cttctggtgg tctttccagc ggaggactcg gagagcccaa attcaatggt    420
gatgagactg ccttcaccgg gtcttggggc cggccccagc gtgatggtc ggctctccgc    480
gcaactgcct tgatctcatt cggacagtgg ctcccttgaca acggatactc cacatacgcg    540
accaacattg tttggccggt tgtacgaaac gatctctctt acgttgctca atactggaac    600
caaaactggat acgatctctg ggaagagggt tccggctcct ccttcttcac gattgccgcc    660
cagcaccgcg ccttgggtga aggtagcact ttcgccacca gggtcggtgc ttcgtgctca    720
tactgtgatt ctccaggcacc ccagggtctg tgcctctctc aatccttctg gactggctca    780
tacacactgg ccaactttgg cggtgccgc tctggcaaag acgcgaatac tcttctcgga    840
agcattcaca catttgaccc cgaggccggc tgcgatgaca ccacctcca accatgctcg    900
gctcgagcat tggccaacca caaggtggtg actgactctt tccgatcggg ttacacaatc    960
aactctggga ttgcggcagg caaggctgtc tctgttgcc gtactcggg ggattcttac   1020
tataatggaa acccttggtg tctgtgtacc ctggctgcag cggagcagtt atatgacgct   1080
ctttacacgt ggaatcgggt tggtctttg actatcactt ctgtctcgt gagtttcttc   1140
aaagatttgt acagttctgc tgctactggt acctactcat cgtccagtgc tacatactct   1200
tcgatcgta gcgcagtcaa gacttacgcc gatggatatg tcagtattgt ggagaactac   1260
gctctgacca atggctctat gtccgagcaa tttcaaagt ccgatggtc tcaactgtcc   1320
gctcgtgact tgacttggtc atacgcagct cttctcactg ccaatgagcg ccgaaatgcc   1380
gtcgtcccgg cgccatggg cgagacggcg gccagcagtg tacctgcca atgcagctcc   1440
acttctgcca ctggctactt cagtacggcc accaacaccg catggccatc gacctcact   1500
agcggaaacc gaagtggaac caccacgacc ggcactggga cgaccacaaa ggcaaccacc   1560
acaacttcta ccaagactac ctccacgacc acatcatgta ccacaccgac gtcggtggcg   1620
gtgacctttg atgtgattgc caccactgtg tatggagaga atatcaagct tgccggctct   1680
atctctcaac ttggctcgtg ggatacaagc agtgcgattg ctctcagcgc tcttcttac   1740
acttcaagca accatctttg gtatgtgacg gtcacactgc ccgctgggtc taccttcacc   1800
tataagtata ttcgggtcga gagtgatggc tcaatccagt gggagagtga tccaaacttg   1860

tcatacaccg tccctcaagc ttgcggcact tcagctgtca caattagcga tacttggagg   1920
tga                                                                                   1923

```

5 <210> 38

<211> 640

<212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

5 <220>

<221> SEÑAL

<222> (1)...(19)

<220>

<221> DOMINIO

10 <222> (42)...(456)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

<222> (538)...(634)

15 <223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (203)...(206)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

20 <220>

<221> SITIO

<222> (205)...(215)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

<220>

25 <221> SITIO

ES 2 575 912 T3

<222> (430)...(433)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (628)...(631)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 38

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

ES 2 575 912 T3

Phe Lys Asn Gly Asp Thr Ala Leu Leu Thr Val Ile Glu Glu Tyr Ile
 100 105 110
 Ser Ala Gln Ala Tyr Ile Gln Thr Val Ser Asn Pro Ser Gly Gly Leu
 115 120 125
 Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu Thr Ala
 130 135 140
 Phe Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu Arg
 145 150 155 160
 Ala Thr Ala Leu Ile Ser Phe Gly Gln Trp Leu Leu Asp Asn Gly Tyr
 165 170 175
 Ser Thr Tyr Ala Thr Asn Ile Val Trp Pro Val Val Arg Asn Asp Leu
 180 185 190
 Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp Leu Trp Glu
 195 200 205
 Glu Val Ser Gly Ser Ser Phe Phe Thr Ile Ala Ala Gln His Arg Ala
 210 215 220
 Leu Val Glu Gly Ser Thr Phe Ala Thr Arg Val Gly Ala Ser Cys Ser
 225 230 235 240
 Tyr Cys Asp Ser Gln Ala Pro Gln Val Leu Cys Phe Leu Gln Ser Phe
 245 250 255
 Trp Thr Gly Ser Tyr Thr Leu Ala Asn Phe Gly Gly Gly Arg Ser Gly
 260 265 270
 Lys Asp Ala Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp Pro Glu
 275 280 285
 Ala Gly Cys Asp Asp Thr Thr Phe Gln Pro Cys Ser Ala Arg Ala Leu
 290 295 300
 Ala Asn His Lys Val Val Thr Asp Ser Phe Arg Ser Val Tyr Thr Ile
 305 310 315 320
 Asn Ser Gly Ile Ala Ala Gly Lys Ala Val Ser Val Gly Arg Tyr Ser
 325 330 335
 Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Tyr Leu Cys Thr Leu Ala
 340 345 350
 Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Thr Trp Asn Arg Val Gly
 355 360 365

ES 2 575 912 T3

Ser Leu Thr Ile Thr Ser Val Ser Leu Ser Phe Phe Lys Asp Leu Tyr
 370 375 380
 Ser Ser Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Ala Thr Tyr Ser
 385 390 395
 Ser Ile Val Ser Ala Val Lys Thr Tyr Ala Asp Gly Tyr Val Ser Ile
 405 410 415
 Val Glu Asn Tyr Ala Leu Thr Asn Gly Ser Met Ser Glu Gln Phe Ser
 420 425
 Lys Ser Asp Gly Ser Gln Leu Ser Ala Arg Asp Leu Thr Trp Ser Tyr
 435 440 445
 Ala Ala Leu Leu Thr Ala Asn Glu Arg Arg Asn Ala Val Val Pro Ala
 450 455 460
 Pro Trp Gly Glu Thr Ala Ala Ser Ser Val Pro Ala Gln Cys Ser Ser
 465 470 475 480
 Thr Ser Ala Thr Gly Thr Phe Ser Thr Ala Thr Asn Thr Ala Trp Pro
 485 490 495
 Ser Thr Leu Thr Ser Gly Thr Gly Ser Gly Thr Thr Thr Thr Gly Thr
 500 505 510
 Gly Thr Thr Thr Lys Ala Thr Thr Thr Ser Thr Lys Thr Thr Ser
 515 520 525
 Thr Thr Thr Ser Cys Thr Thr Pro Thr Ser Val Ala Val Thr Phe Asp
 530 535 540
 Val Ile Ala Thr Thr Val Tyr Gly Glu Asn Ile Lys Leu Ala Gly Ser
 545 550 555 560
 Ile Ser Gln Leu Gly Ser Trp Asp Thr Ser Ser Ala Ile Ala Leu Ser
 565 570 575
 Ala Ser Ser Tyr Thr Ser Ser Asn His Leu Trp Tyr Val Thr Val Thr
 580 585 590
 Leu Pro Ala Gly Ser Thr Phe Thr Tyr Lys Tyr Ile Arg Val Glu Ser
 595 600 605
 Asp Gly Ser Ile Gln Trp Glu Ser Asp Pro Asn Leu Ser Tyr Thr Val
 610 615 620
 Pro Gln Ala Cys Gly Thr Ser Ala Val Thr Ile Ser Asp Thr Trp Arg
 625 630 635 640

<210> 39

<211> 1896

<212> ADN

5 <213> Desconocido

ES 2 575 912 T3

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 39

```

atggctcctc gtttttggac tgcgctttgg gcgcttacgc ttggccatgc cgttgtggcc 60
acgccacaga tattggctcc tcgtgcgact ggcagtttgg atacctgggt ggcgtctgag 120
actgtggtgg caagacaggg tatcttggac aacatcggat cggctggtgc ccatgctgcc 180
aatgcaaagc cgggcgttgt actcgccagc ccgagtacct cggaccctga ttattattat 240
acttggactc gtgactcggc cctagtgttc aaaaatctgg tcgacatggt caggagcggc 300
gacagcgtt tgttggaggt tatcgaagaa tacatcagct cgcaggccta tatccagacg 360
gtgtcaaadc catctggagg tctttccggc ggcggtggat tgggagagcc caagttcaac 420
gccgatgaga cagctttcac tggctcttgg ggacgccctc agcgtgatgg accagccttg 480
cgagccactg ccttgatctc atttggccaa tggcttattg acaatgggta caccacctac 540
gcgaccgaca ttgtctggcc tgcgtgagc aatgatctct cctatgttcc ccagtactgg 600
aatcagactg gatttgatct ctgggaagaa gtttctggct catcattctt caccgtcgcg 660
gctcagcacc gcgctttggt ggagggaaagc acattcgcga gtcaggtcgg ttcttcatgt 720
ttgtactgtg actcccaggc tccgcagggt ctgtgcttcc tacagtcctt ctggactgga 780
tcttacattc tggccaactt tgggtggtggc cgctctggga aggacgcaa tacactgctc 840
ggcagcattc acacattcga cccagaggca ggatgcgatg acacgacggt ccagccttgc 900
tcagcacgag cactcgcaaa ccacaaagtt gtaactgact cgttccgatc gatctactct 960
gtcaactctg gtatcgctgc gggtaaggct gtttctgttg gtcgataccc agaggactca 1020
tactacaatg gtaacccttg gtatctgtgc actttggctg cagctgagca gttgtatgat 1080
gctatctaca catggaaccg tattggttct ttgaccatta cctctgtctc tttgagcttc 1140
ttcaaagacc taticagctc tgctgcgacc ggcacctact cctcatccag tgafacatac 1200
tcctcgatcg tggccgctgt aaaggaatat gcagatggat atgtagcat cgtggaaaaa 1260
tacgctgcat caagtgggtc cttgtccgag caattctcca agtcagatgg ctcgcagctg 1320
tcagctcgtg acctgacttg gtcctacgca gccctactta ctgccaatga gcgtcgaaac 1380
gccatagtcc ctgcaccatg gggggagaca tctgctagca gtgtccctgg gcagtgtcaa 1440
tatacttcgg ccattggtac ttacagcagc gcaacaaaca ccgectggcc taccactttg 1500
actagtggat cgggcagtgt aactaccacc aagacgacta ccactacctc gaagccgaca 1560
acaacatcat gcactactcc gacctgtt gcagtgacgt tcaatgtgat tgctactact 1620
gaatatggcc agaacatcaa actcgtgga tctatctctc aacttggctc atggtcgcca 1680
agcagtgtg tcgcattgag cgcttccaaa tacaccacga gcaaccacct gtggtttgtg 1740
accgtgacac ttccagtagg cactagcttc agctacaagt acatccaggt ggcgagcgat 1800

ggcactatca agtgggaaag tgacccaaac cagtcgtata ccgttcccgc cacgtgtggt 1860
actactgctg tcaccgtcag tgatacatgg aggtag 1896

```

5 <210> 40

ES 2 575 912 T3

<211> 631

<212> PRT

<213> Desconocido

<220>

5 <223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

10 <221> DOMINIO

<222> (42)...(457)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

15 <222> (529)...(625)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (204)...(207)

20 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (206)...(216)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

25 <220>

ES 2 575 912 T3

<221> SITIO

<222> (619)...(622)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 40

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

Ala Val Val Ala Thr Pro Gln Ile Leu Ala Pro Arg Ala Thr Gly Ser
20 25 30

Leu Asp Thr Trp Leu Ala Ser Glu Thr Val Val Ala Arg Gln Gly Ile
35 40 45

Leu Asp Asn Ile Gly Ser Ala Gly Ala His Ala Ala Asn Ala Lys Pro
50 55 60

Gly Val Val Leu Ala Ser Pro Ser Thr Ser Asp Pro Asp Tyr Tyr Tyr
65 70 75 80

Thr Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Leu Val Asp Met
85 90 95

Phe Arg Ser Gly Asp Ser Ala Leu Leu Glu Val Ile Glu Glu Tyr Ile
100 105 110

5

ES 2 575 912 T3

Ser Ser Gln Ala Tyr Ile Gln Thr Val Ser Asn Pro Ser Gly Gly Leu
115 120 125

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

Ala Phe Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala Leu
145 150 155 160

Arg Ala Thr Ala Leu Ile Ser Phe Gly Gln Trp Leu Ile Asp Asn Gly
165 170 175

Tyr Thr Thr Tyr Ala Thr Asp Ile Val Trp Pro Val Val Arg Asn Asp
180 185 190

Leu Ser Tyr Val Ser Gln Tyr Trp Asn Gln Thr Gly Phe Asp Leu Trp
195 200 205

Glu Glu Val Ser Gly Ser Ser Phe Phe Thr Val Ala Ala Gln His Arg
210 215 220

Ala Leu Val Glu Gly Ser Thr Phe Ala Ser Gln Val Gly Ser Ser Cys
225 230 235 240

Leu Tyr Cys Asp Ser Gln Ala Pro Gln Val Leu Cys Phe Leu Gln Ser
245 250 255

Phe Trp Thr Gly Ser Tyr Ile Leu Ala Asn Phe Gly Gly Gly Arg Ser
260 265 270

Gly Lys Asp Ala Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp Pro
275 280 285

Glu Ala Gly Cys Asp Asp Thr Thr Phe Gln Pro Cys Ser Ala Arg Ala
290 295 300

Leu Ala Asn His Lys Val Val Thr Asp Ser Phe Arg Ser Ile Tyr Ser
305 310 315 320

Val Asn Ser Gly Ile Ala Ala Gly Lys Ala Val Ser Val Gly Arg Tyr
325 330 335

Pro Glu Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Tyr Leu Cys Thr Leu
340 345 350

Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile Tyr Thr Trp Asn Arg Ile
355 360 365

Gly Ser Leu Thr Ile Thr Ser Val Ser Leu Ser Phe Phe Lys Asp Leu
370 375 380

ES 2 575 912 T3

Tyr Ser Ser Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Asp Thr Tyr
 385 390 395 400
 Ser Ser Ile Val Ala Ala Val Lys Glu Tyr Ala Asp Gly Tyr Val Ser
 405 410 415
 Ile Val Glu Lys Tyr Ala Ala Ser Ser Gly Ser Leu Ser Glu Gln Phe
 420 425 430
 Ser Lys Ser Asp Gly Ser Gln Leu Ser Ala Arg Asp Leu Thr Trp Ser
 435 440 445
 Tyr Ala Ala Leu Leu Thr Ala Asn Glu Arg Arg Asn Ala Ile Val Pro
 450 455 460
 Ala Pro Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Gln Cys Gln
 465 470 475 480
 Tyr Thr Ser Ala Ile Gly Thr Tyr Ser Ser Ala Thr Asn Thr Ala Trp
 485 490 495
 Pro Thr Thr Leu Thr Ser Gly Ser Gly Ser Val Thr Thr Thr Lys Thr
 500 505 510 515
 Thr Thr Thr Thr Ser Lys Pro Thr Thr Thr Ser Cys Thr Thr Pro Thr
 515 520 525
 Thr Val Ala Val Thr Phe Asn Val Ile Ala Thr Thr Glu Tyr Gly Gln
 530 535 540
 Asn Ile Lys Leu Ala Gly Ser Ile Ser Gln Leu Gly Ser Trp Ser Pro
 545 550 555 560
 Ser Ser Ala Val Ala Leu Ser Ala Ser Lys Tyr Thr Thr Ser Asn His
 565 570 575
 Leu Trp Phe Val Thr Val Thr Leu Pro Val Gly Thr Ser Phe Ser Tyr
 580 585 590
 Lys Tyr Ile Gln Val Ala Ser Asp Gly Thr Ile Lys Trp Glu Ser Asp
 595 600 605
 Pro Asn Gln Ser Tyr Thr Val Pro Ala Thr Cys Gly Thr Thr Ala Val
 610 615 620
 Thr Val Ser Asp Thr Trp Arg
 625 630

<210> 41

<211> 1563

<212> ADN

5 <213> Desconocido

ES 2 575 912 T3

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 41

```

atgctgggtc ctgtggatcat tctcacacat tttctccact atgcttgccg tttctcgata      60
ccccatttgc acttgaatatt tgcgagtgcc agtgacgcac aaataccaca gaagccgctg      120
cgggatacgc tgggaagtctg gctcgaagct gaagagcgca tagctctaaa taaattactc      180
gccaacgtag ctccaggcgg tagcaatgtc gagggcaaag gcgtggcacc cggaactgtc      240
attgagagtc cgagtcagga cgggccagat tattggtatc agtgggttcg agatgccgct      300
atccaccatga atacgctcgt agatttgtat gccgagaacc cttcatccgc cttatcatct      360
cgtctttcga cgacctaga tgcatacgtc tccctacaac gtgatctgca gcatacatca      420
aatccctcgg gctcatttga tgattcgtcc ggccttggtg aaccgaaatt tgaagtcgac      480
ggaacgccat ttacaggctc atggggaagg ccacaacgcg atggaccggc gttaagagcg      540
ctgactctta tgcggtatct tcgagaatac aacgctagcc atccttcgct ttggtcttca      600
aatgaggcta ccgacttcta tagtctttat tacgaagcag agatgccacc gcggagcgtg      660
atcaaggcag accttgaata cgtgagccat ttctggaacg aatccagttt tgatctttgg      720
gaagaaacag aaggtcttca cttctttaac ttgatggtca gcgcacgaag tctaaggag      780
ggcagtgagc tagcaagagc atttggagat attggtgcag cagaatggta tatcgaacaa      840
gctggctaca tcgaaaagct tctgagcaaa ttctggaacc cgaacaaagg ccatctcgtg      900
gagaccttgt ggagtaagcg ctcaggtcta gattgtggtc ttcttctcgg gtctctacat      960
gctcttctcgt ctcagggtt ggaagaagag gcagtttcc cggcctgggc cgacgagatc     1020
ctcgtgagct tgcttgctct gacagaagat caaagagacc gcttccaat caacagcaac     1080
ccatagtagg atcaagatgg atcccatcca gcgctcgagg gtacaggagt tggacgttat     1140
cccgaagacg tctacgacgg ctatggcacc agtaatcgcg gaggcaacc atggttccta     1200
tgcacatcaa gcgcagccga aatactgtac cgcagtgctt cttacttcta tacagcctcc     1260
aacctcacca tctccaccgc atctcttccg ttctacacct cctgcttgc aacttcaagc     1320
ctcgatgtcc aagtcggaac atttggacc tcagatacgc tttccactc tgtcatcgag     1380
catctcaagt ccgagggcga ttcgttcctt gaagttgtga aaacgcatgt agatgatgag     1440
ggacgaatga gcgagcaatt tgacagagtg acgggttata tgcgcggtgc gcgggatttg     1500
acctggagtt acggtgcatt tctacaggcg gtgaaggcgc gaaggagtat tcaggaggtt     1560
tga                                                                                   1563

```

5 <210> 42

<211> 520

<212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> DOMINIO

<222> (53)...(512)

5 <223> Familia de glicosil hidrolasas 15

<220>

<221> SITIO

<222> (193)...(196)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <220>

<221> SITIO

<222> (236)...(239)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (427)...(430)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 42

ES 2 575 912 T3

Met Leu Val Pro Val Val Ile Leu Thr His Phe Leu His Tyr Ala Cys
 1 5 10 15
 Ala Phe Ser Ile Pro His Leu His Leu Lys Phe Ala Ser Ala Ser Asp
 20 25 30
 Ala Gln Ile Pro Gln Lys Pro Leu Arg Asp Thr Leu Glu Val Trp Leu
 35 40 45
 Glu Ala Glu Glu Arg Ile Ala Leu Asn Lys Leu Leu Ala Asn Val Ala
 50 55 60
 Pro Gly Gly Ser Asn Val Glu Gly Lys Gly Val Ala Pro Gly Thr Val
 65 70 75 80
 Ile Ala Ser Pro Ser Gln Asp Gly Pro Asp Tyr Trp Tyr Gln Trp Val
 85 90 95
 Arg Asp Ala Ala Ile Thr Met Asn Thr Leu Val Asp Leu Tyr Ala Glu
 100 105 110
 Asn Pro Ser Ser Ala Leu Ser Ser Arg Leu Ser Thr Thr Leu Asp Ala
 115 120 125
 Tyr Ala Ser Leu Gln Arg Asp Leu Gln His Thr Ser Asn Pro Ser Gly
 130 135 140
 Ser Phe Asp Asp Ser Ser Gly Leu Gly Glu Pro Lys Phe Glu Val Asp
 145 150 155 160
 Gly Thr Pro Phe Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro
 165 170 175
 Ala Leu Arg Ala Leu Thr Leu Met Arg Tyr Leu Arg Glu Tyr Asn Ala
 180 185 190
 Ser His Pro Ser Leu Trp Ser Ser Asn Glu Ala Thr Asp Phe Tyr Ser
 195 200 205

ES 2 575 912 T3

Leu Tyr Tyr Glu Ala Glu Met Pro Pro Arg Ser Val Ile Lys Ala Asp
 210 215 220
 Leu Glu Tyr Val Ser His Phe Trp Asn Glu Ser Ser Phe Asp Leu Trp
 225 230 235 240
 Glu Glu Thr Glu Gly Leu His Phe Phe Asn Leu Met Val Ser Ala Arg
 245 250 255
 Ser Leu Arg Glu Gly Ser Glu Leu Ala Arg Ala Phe Gly Asp Ile Gly
 260 265 270
 Ala Ala Glu Trp Tyr Ile Glu Gln Ala Gly Tyr Ile Glu Lys Leu Leu
 275 280 285
 Ser Lys Phe Trp Asn Pro Asn Lys Gly His Leu Val Glu Thr Leu Trp
 290 295 300
 Ser Lys Arg Ser Gly Leu Asp Cys Gly Leu Leu Leu Gly Ser Leu His
 305 310 315 320
 Ala Leu Pro Ala Gln Gly Leu Glu Glu Glu Ala Val Phe Pro Pro Trp
 325 330 335
 Ser Asp Glu Ile Leu Val Ser Leu Leu Ala Leu Thr Glu Asp Gln Arg
 340 345 350
 Asp Arg Phe Pro Ile Asn Ser Asn Pro Tyr Glu Asp Gln Asp Gly Ser
 355 360 365
 His Pro Ala Phe Glu Gly Thr Gly Val Gly Arg Tyr Pro Glu Asp Val
 370 375 380
 Tyr Asp Gly Tyr Gly Thr Ser Asn Arg Gly Gly Asn Pro Trp Phe Leu
 385 390 395 400
 Cys Thr Ser Ser Ala Ala Glu Ile Leu Tyr Arg Ser Ala Ser Tyr Phe
 405 410 415
 Tyr Thr Ala Ser Asn Leu Thr Ile Ser Thr Ala Ser Leu Pro Phe Tyr
 420 425 430
 Thr Ser Leu Leu Ala Thr Ser Ser Leu Asp Val Gln Val Gly Thr Phe
 435 440 445
 Gly Pro Ser Asp Thr Leu Phe His Ser Val Ile Glu His Leu Lys Ser
 450 455 460
 Ala Gly Asp Ser Phe Leu Glu Val Val Lys Thr His Val Asp Asp Glu
 465 470 475 480

ES 2 575 912 T3

Gly Arg Met Ser Glu Gln Phe Asp Arg Val Thr Gly Tyr Met Arg Gly
485 490 495

Ala Arg Asp Leu Thr Trp Ser Tyr Gly Ala Phe Leu Gln Ala Val Lys
500 505 510

Ala Arg Arg Ser Ile Gln Glu Val
515 520

<210> 43

<211> 3030

<212> ADN

5 <213> *Cochliobolus heterostrophus* ATCC 48332

<400> 43

ES 2 575 912 T3

atgacgcaca caagtttcgt ccaggcttct actgtgctgt cctctcttct cgcactcaca 60
 gctggccaag cgcccaccac ttcttcaaat ggtggatgga gtactacgct cgctggcacg 120
 cccacctact tcaatcccat ctttaccatc cctccatctg cggatgaggg cgttcaacag 180
 attcccaaca tctatgacct acaagccgtg gatgcgcaag acgtttgccc tggttacaca 240
 gcatcaaacc ctacgagag tgatcgtggc ctgaccgcca cgttgaccct ggctggaaag 300
 ccatgcaatg tttacggcac cgatatccag gaactagact tgaaagtcga gtaccaggcc 360
 aagggaagac tgtcagtcaa cattgtacca aagtacactg gtgcaagcaa ccagtcgcat 420
 tggattgttc ccgaagatct gattccgctg cctcaggtcg aggaatcatc cgagcagacc 480
 gatctcaagt tcaactgggg aaatgagcca tctttttggt tcaatgtcga gcgctcctcg 540
 actggagaca tcatcttcac aactcaaggc acccatctca tttacgaaaa tcaatttggt 600
 gaatttgtaa acagcctgcc agaggactac aacctttacg gtctgggtga acgaattcac 660
 ggtcttcgtc ttaacaacaa cttcactgcg accatctacg ctgctgacgt aggtgatcct 720
 atcgaccgca acctgtacgg tagccaccct ttttacctag agaccggta ctttgagaag 780
 ggcgaaaact gtagaccaa gcccttgacg cagtccgaga tcggccagaa gaaggatacc 840
 aaaacaaatg gttcgcccta cgagtcgctg tcccacggcg tttattaccg caacactcat 900
 ggtatggacg tgggtgctgaa acctgatcac ctaacttggg ggaccctggg aggtgctatt 960
 gatctgtact tcttcgatgg accttcacag ccggacgtta ctaaagcgta tcaaaaggcg 1020
 gctattggcc tgccagcaat gcagcaatac tggacctttg gctttcatca atgccgttgg 1080
 ggctaccgca actggaccga gactcgagaa attgttgaga ccatgagagc tttcaacatc 1140
 ccaatggaaa ccatctggct cgacattgac tacatggacc agtaccgtga ctttacgctg 1200
 gatcccgtgg ctttccctcc ctctgaagtg gccgacttct tcggatggct gcatggaaac 1260
 aaccagcatt ttgtaccat cgttgatgca gctatttaca tccccaaacc gcaaaacgca 1320
 agcgatgcat atgacacata tacgcgcggc aatgaatctg gtgtgttctt gtccaacca 1380
 gacggcagcc aatatacgg tgctgtctgg cctgggtaca ccgttttccc tgattggttg 1440
 tcgcccacg gtgtatcttg gtgggtaaag gagatggtcg agtgggtacaa agaggtacca 1500
 tatagcgggt tctgggttga catgactgaa gtctcttctt tctgcggttg atcctgtggt 1560

ES 2 575 912 T3

actggcaacg tcaccctgaa cccggctcac cgcaccttct cgcttcccgg tgaggtgaac 1620
aacgtcatct atgattaccc agagagcttc aacatcacia atggtaccga ggcggcgctc 1680
gcttctgcag cagcttccgc ccaagcatcc aggaaggcaa ctgacgacggc gacggtgacc 1740
gacgaagtga cttctacttc gacaagctac ttccgctcga cccctaccgc tggagagcga 1800
aacatcaact atcctcctta tgtcatcaac cacgtacaag atggcgctga tcttgctggt 1860
catgctgtca gcccgaatgc aactcatgca aatggtgtgg aggaatacga tgtccacaac 1920
ctctttggtc accagatcat taacgctact tatcacggtc tcctttcggg gttccctgga 1980
aagcggccct ttatcattgg acgctccaca tttgctggca gcggcaagtg ggctggctcat 2040
tgggggtggcg acaacgcttc taaatgggccc tacatgttct tctcgattcc acaagcactt 2100
tccttttcac tatttggtat tccgatgttt ggtgtcgata cctgaggatt caacggaaac 2160
actgacatgg agctttgctc tcgatggatg cagctttcgg cctttttccc cttttaccgc 2220
aaccacaacg tactctccgc cattcccaaa gagccctacc gatgggaggc cgtggcttct 2280
gcttcgagga ccgcaatgca catccgatac tctctgttgc catacatgta cacgctattc 2340
aatgacgctc acaagactgg gtcaactgtg atgcggggcgc tggcatggga attcccaaac 2400
gagcccaaac tcgctgggtg agatacacia ttcttgcttg ggccaaacat cctagtcacg 2460
cctgtcctcg agcctcaagt tgatactgtc aagggagtgt tccctggtat tgcgacggc 2520
gagacatggt ttgattggta ttctggcgag cgtgtccaag ccgaggctgg cgtaaacacg 2580
accatctcag cacctttggg ccacatcccg gtatacattc gcggaggttc ggtcctacct 2640
attcaggagc ctggatacac gactaccgag tcccgcggga acccctgggg tctcattgtc 2700
gcactctcca gtgaaggcac tgcacgggt cacctatacg ttgatgatgg tgagtccatc 2760
gagccagact cgtgcctgaa tggtgcattt gctgctacga gcggaaagt ggaggttgat 2820
gttcagggcg agttcaagga cacgaacgca cttgccaacg tgacagtttt gggagctcct 2880
gcggtccaaa acgtcaagct gaatggtgag gcgattgatg ctagcaatgt tgattacaac 2940
aagaccagta gcgttttgaa gctgacggga ctcaatgaac tgacaagtcc tggagcttgg 3000
cagggtagct ggacgctgac ttggaataa 3030

<210> 44

<211> 1009

<212> PRT

5 <213> Cochliobolus heterostrophus ATCC 48332

<220>

<221> DOMINIO

<222> (323)...(879)

<223> Familia de Glicosil hidrolasas 31

10 <220>

<221> SITIO

<222> (139)...(142)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (230)...(233)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (267)...(270)

10 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (369)...(372)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

15 <220>

<221> SITIO

<222> (445)...(448)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

20 <221> SITIO

<222> (457)...(460)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

25 <222> (531)...(534)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (559)...(562)

5 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (635)...(638)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <220>

<221> SITIO

<222> (657)...(660)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (695)...(698)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

20 <222> (872)...(875)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (967)...(970)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

ES 2 575 912 T3

<220>

<221> SITIO

<222> (995)...(998)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

5 <400> 44

Met Thr His Thr Ser Phe Val Gln Ala Ser Thr Val Leu Ser Ser Leu
1 5 10 15
Leu Ala Leu Thr Ala Gly Gln Ala Pro Thr Thr Ser Ser Asn Gly Gly
20 25 30

ES 2 575 912 T3

Trp Ser Thr Thr Leu Ala Gly Thr Pro Thr Tyr Phe Asn Pro Ile Phe
 35 40 45
 Thr Ile Pro Pro Ser Ala Asp Glu Gly Val Gln Gln Ile Pro Asn Ile
 50 55 60
 Tyr Asp Pro Gln Ala Val Asp Ala Gln Asp Val Cys Pro Gly Tyr Thr
 65 70 75 80
 Ala Ser Asn Pro Gln Gln Ser Asp Arg Gly Leu Thr Ala Thr Leu Thr
 85 90
 Leu Ala Gly Lys Pro Cys Asn Val Tyr Gly Thr Asp Ile Gln Glu Leu
 100 105
 Asp Leu Lys Val Glu Tyr Gln Ala Lys Gly Arg Leu Ser Val Asn Ile
 115 120 125
 Val Pro Lys Tyr Thr Gly Ala Ser Asn Gln Ser His Trp Ile Val Pro
 130 135 140
 Glu Asp Leu Ile Pro Arg Pro Gln Val Glu Glu Ser Ser Glu Gln Thr
 145 150 155 160
 Asp Leu Lys Phe Asn Trp Gly Asn Glu Pro Ser Phe Trp Phe Asn Val
 165 170 175
 Glu Arg Ser Ser Thr Gly Asp Ile Ile Phe Thr Thr Gln Gly Thr His
 180 185 190
 Leu Ile Tyr Glu Asn Gln Phe Val Glu Phe Val Asn Ser Leu Pro Glu
 195 200 205
 Asp Tyr Asn Leu Tyr Gly Leu Gly Glu Arg Ile His Gly Leu Arg Leu
 210 215 220
 Asn Asn Asn Phe Thr Ala Thr Ile Tyr Ala Ala Asp Val Gly Asp Pro
 225 230 235 240
 Ile Asp Arg Asn Leu Tyr Gly Ser His Pro Phe Tyr Leu Glu Thr Arg
 245 250 255
 Tyr Phe Glu Lys Gly Glu Asn Cys Ser Thr Lys Pro Leu Thr Gln Ser
 260 265 270
 Glu Ile Gly Gln Lys Lys Asp Thr Lys Thr Asn Gly Ser Pro Tyr Glu
 275 280 285
 Ser Arg Ser His Gly Val Tyr Tyr Arg Asn Thr His Gly Met Asp Val
 290 295 300

ES 2 575 912 T3

Val Leu Lys Pro Asp His Leu Thr Trp Arg Thr Leu Gly Gly Ala Ile
 305 310 315 320
 Asp Leu Tyr Phe Phe Asp Gly Pro Ser Gln Pro Asp Val Thr Lys Ala
 325 330 335
 Tyr Gln Lys Ala Ala Ile Gly Leu Pro Ala Met Gln Gln Tyr Trp Thr
 340 345 350
 Phe Gly Phe His Gln Cys Arg Trp Gly Tyr Arg Asn Trp Thr Glu Thr
 355 360 365
 Arg Glu Ile Val Glu Thr Met Arg Ala Phe Asn Ile Pro Met Glu Thr
 370 375 380
 Ile Trp Leu Asp Ile Asp Tyr Met Asp Gln Tyr Arg Asp Phe Thr Leu
 385 390 395 400
 Asp Pro Val Ala Phe Pro Pro Ser Glu Val Ala Asp Phe Phe Gly Trp
 405 410 415
 Leu His Gly Asn Asn Gln His Phe Val Pro Ile Val Asp Ala Ala Ile
 420 425 430
 Tyr Ile Pro Asn Pro Gln Asn Ala Ser Asp Ala Tyr Asp Thr Tyr Thr
 435 440 445
 Arg Gly Asn Glu Ser Gly Val Phe Leu Ser Asn Pro Asp Gly Ser Gln
 450 455 460
 Tyr Ile Gly Ala Val Trp Pro Gly Tyr Thr Val Phe Pro Asp Trp Leu
 465 470 475 480
 Ser Pro Asn Gly Val Ser Trp Trp Val Lys Glu Met Val Glu Trp Tyr
 485 490 495
 Lys Glu Val Pro Tyr Ser Gly Phe Trp Val Asp Met Thr Glu Val Ser
 500 505 510
 Ser Phe Cys Val Gly Ser Cys Gly Thr Gly Asn Val Thr Leu Asn Pro
 515 520 525
 Ala His Pro Pro Phe Ser Leu Pro Gly Glu Val Asn Asn Val Ile Tyr
 530 535 540
 Asp Tyr Pro Glu Ser Phe Asn Ile Thr Asn Gly Thr Glu Ala Ala Ser
 545 550 555 560
 Ala Ser Ala Ala Ala Ser Ala Gln Ala Ser Arg Lys Ala Thr Ala Thr
 565 570 575

ES 2 575 912 T3

Ala Thr Val Thr Asp Glu Val Thr Ser Thr Ser Thr Ser Tyr Phe Arg
580 585 590

Ser Thr Pro Thr Ala Gly Glu Arg Asn Ile Asn Tyr Pro Pro Tyr Val
595 600 605

Ile Asn His Val Gln Asp Gly Ala Asp Leu Ala Val His Ala Val Ser
610 615 620

Pro Asn Ala Thr His Ala Asn Gly Val Glu Glu Tyr Asp Val His Asn
625 630 635 640

Leu Phe Gly His Gln Ile Ile Asn Ala Thr Tyr His Gly Leu Leu Ser
645 650 655

Val Phe Pro Gly Lys Arg Pro Phe Ile Ile Gly Arg Ser Thr Phe Ala
660 665 670

Gly Ser Gly Lys Trp Ala Gly His Trp Gly Gly Asp Asn Ala Ser Lys
675 680 685

Trp Ala Tyr Met Phe Phe Ser Ile Pro Gln Ala Leu Ser Phe Ser Leu
690 695 700

Phe Gly Ile Pro Met Phe Gly Val Asp Thr Cys Gly Phe Asn Gly Asn
705 710 715 720

Thr Asp Met Glu Leu Cys Ser Arg Trp Met Gln Leu Ser Ala Phe Phe
725 730 735

Pro Phe Tyr Arg Asn His Asn Val Leu Ser Ala Ile Pro Gln Glu Pro
740 745 750

Tyr Arg Trp Glu Ala Val Ala Ser Ala Ser Arg Thr Ala Met His Ile
755 760 765

Arg Tyr Ser Leu Leu Pro Tyr Met Tyr Thr Leu Phe Asn Asp Ala His
770 775 780

Lys Thr Gly Ser Thr Val Met Arg Ala Leu Ala Trp Glu Phe Pro Asn
785 790 795 800

Glu Pro Gln Leu Ala Gly Val Asp Thr Gln Phe Leu Leu Gly Pro Asn
805 810 815

Ile Leu Val Thr Pro Val Leu Glu Pro Gln Val Asp Thr Val Lys Gly
820 825 830

Val Phe Pro Gly Ile Val Asp Gly Glu Thr Trp Phe Asp Trp Tyr Ser
835 840 845

Gly Glu Arg Val Gln Ala Glu Ala Gly Val Asn Thr Thr Ile Ser Ala

ES 2 575 912 T3

atgaagcttc ttcaactcgc cgccctggtg gcgctccctca gccccttcac caacgctgcc 60
gacgcaaacg cctggaagtc gcgaaacatt tactttgctc ttacagaccg cgttgctcgc 120
agcggtagcg ataacggtgg caatgcctgt ggcaatcttg gaaattattg tggtggaacg 180
ttcaagggtc ttgaggctaa gcttgattac atcaagggca tgggctttga tgctatctgg 240
attactcctg ttgttgagaa caccgatgga ggataccatg gatactgggc taaagacctg 300
tacgcggtta attccaagta tggtaaccaag gatgacttga agaactcttg caaggctgcc 360
cacggcaaga acatgtacgt catggccgac gtcgctcga accacatggg caagggcatc 420
caaaaccaca agccggagcc cctcaaccag caaagctcct accacagctc ctgcccac 480
gactacaaca accaaaacag catcgagcag tgcgaaatcg ccggtctgcc cgatctcaac 540

accggcaagg cagaagtcaa gaaggtcctc aacgactgga tcaagtggct cgtctccgag 600
tacagcttcg acggtatccg catcgacaca gtcaagcacg tgcgagaagag cttctggcct 660
gatttccaga aggcagctgg cgttttcgcc atcggtgagg tttgggatgg aagccctgat 720
taccttgctg gttactccaa ggtcatgcct ggtctgctga actatgctat gtactatccc 780
atgaaccgat tctatcagca gaagggatgat cttctgctg tggtgatata gtacaacgag 840
atcagccaga agttcgatga tctacgcag cttggtacct tcatcgacaa ccacgacaat 900
gcgcgctggt tgagccaaaa gaacgacaag gccctcctca agaacgccct cgcattcacc 960
atcctcgcgc gcggtattcc catcgtgtac tacggcaccg aacaaggcta cgcaggaggc 1020
aacgaccccg ccaaccgcga agatctctgg cgcagcaact tcagcaccga ctccgacctg 1080
taccaaacca tttccaagct cggcaaggct cgctccgccg tcggtggtct tgccggcaac 1140
gaccagaaat tctcaagtc caatgacagc gcactcatct ggagccgcgc agacggcgat 1200
ctgatcgttg ttacgcttaa ccgtgaaaag ggatattctg gagagtactg cttcaacact 1260
ggcaagaaca acaagacttg ggatcgtgtt ttgggatctg gaagtgttaa gtctgatggt 1320
agtggtaagc tttgtgttag ctacactaat ggtgagcctg aggttcttgt tgctgcttag 1380

<210> 46

<211> 459

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

10 <222> (1)...(19)

<220>

ES 2 575 912 T3

<221> DOMINIO

<222> (32)...(371)

<223> Alfa amilasa, dominio catalítico

<220>

5 <221> SITIO

<222> (358)...(361)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

10 <222> (393)...(396)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (430)...(433)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 46

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

Thr Asn Ala Ala Asp Ala Asn Ala Trp Lys Ser Arg Asn Ile Tyr Phe
20 25 30

ES 2 575 912 T3

Ala Leu Thr Asp Arg Val Ala Arg Ser Gly Ser Asp Asn Gly Gly Asn
 35 40 45

Ala Cys Gly Asn Leu Gly Asn Tyr Cys Gly Gly Thr Phe Lys Gly Leu
 50 55 60

Glu Ala Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile Trp
 65 70 75 80

Ile Thr Pro Val Val Glu Asn Thr Asp Gly Gly Tyr His Gly Tyr Trp
 85 90 95

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

Leu Lys Asn Leu Val Lys Ala Ala His Gly Lys Asn Met Tyr Val Met
 115 120 125

Ala Asp Val Val Ala Asn His Met Gly Lys Gly Ile Gln Asn His Lys
 130 135 140

Pro Glu Pro Leu Asn Gln Gln Ser Ser Tyr His Ser Ser Cys Ala Ile
 145 150 155 160

Asp Tyr Asn Asn Gln Asn Ser Ile Glu Gln Cys Glu Ile Ala Gly Leu
 165 170 175

Pro Asp Leu Asn Thr Gly Lys Ala Glu Val Lys Lys Val Leu Asn Asp
 180 185 190

Trp Ile Lys Trp Leu Val Ser Glu Tyr Ser Phe Asp Gly Ile Arg Ile
 195 200 205

Asp Thr Val Lys His Val Glu Lys Ser Phe Trp Pro Asp Phe Gln Lys
 210 215 220

Ala Ala Gly Val Phe Ala Ile Gly Glu Val Trp Asp Gly Ser Pro Asp
 225 230 235 240

Tyr Leu Ala Gly Tyr Ser Lys Val Met Pro Gly Leu Leu Asn Tyr Ala
 245 250 255

Met Tyr Tyr Pro Met Asn Arg Phe Tyr Gln Gln Lys Gly Asp Pro Ser
 260 265 270

Ala Val Val Asp Met Tyr Asn Glu Ile Ser Gln Lys Phe Asp Asp Pro
 275 280 285

Thr Gln Leu Gly Thr Phe Ile Asp Asn His Asp Asn Ala Arg Trp Leu
 290 295 300

ES 2 575 912 T3

Ser Gln Lys Asn Asp Lys Ala Leu Leu Lys Asn Ala Leu Ala Phe Thr
305 310 315 320

Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr Glu Gln Gly
325 330 335

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

Asn Phe Ser Thr Asp Ser Asp Leu Tyr Gln Thr Ile Ser Lys Leu Gly
355 360 365

Lys Ala Arg Ser Ala Val Gly Gly Leu Ala Gly Asn Asp Gln Lys Phe
370 375 380

Leu Lys Ser Asn Asp Ser Ala Leu Ile Trp Ser Arg Ala Asp Gly Asp
385 390 395 400

Leu Ile Val Val Thr Leu Asn Arg Gly Lys Gly Tyr Ser Gly Glu Tyr
405 410 415

Cys Phe Asn Thr Gly Lys Asn Asn Lys Thr Trp Asp Arg Val Leu Gly
420 425 430

Ser Gly Ser Val Lys Ser Asp Gly Ser Gly Lys Leu Cys Val Ser Tyr
435 440 445

Thr Asn Gly Glu Pro Glu Val Leu Val Ala Ala
450 455

<210> 47

<211> 1905

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 47

ES 2 575 912 T3

atgacacgca ttctcacct cgcccttcat gggctggctc ttgtccaaag tgttgttggg 60
gctcccaat tggcccccag agcgacaacc agtctggatg catggttggc atcggagacg 120
accgttgccg tggatgggat ccttgacaac gtgggttcta gtggagccta cgccaaaagt 180
gcgaagtccg gtatcgtgat tgccagtcca agcaccgaca acccagacta ctactacag 240
tggactcgcg atgctgcggt gaccgtcaag gccctgatcg atcttttccg caatggcgag 300
acaagccttc agaccgtgat catggagtac attagctctc aggcgtacct ccagaccgta 360
tccaaccctc cggggctcct gtccaccggg ggtctggcag aaccaaata ttatgtcgat 420
gagactgcct acacgggcag ctggggtcgt cccagcggg atggtcctgc cctgagagcg 480
acggcgatga tcgactttgg taactggctc attgacaacg gatattcaac ctacgcctct 540
aacattgtgt ggccgatcgt ccgcaacgat ctgtcgtacg ttgcccaata ttggaaccaa 600

accggatatg acctctggga agaagtgaac ggatcctcct tcttcacat tgccgtgcag 660
caccgggctc tgggtgaagg cagcaccttc gcctccaaag ttggcgctc atgctcgtgg 720
tgcgactcgc aggcgccgca ggtgctttgc ttctgcaga gattctggac aggctcgtac 780
atcatggcca attttggcgg cgggcgatcg ggcaaagatg ccaacaccgt cctgggaagc 840
atccatacct tcgacccgaa tgccgggttc gacgacacca cgttccagcc atgctctccg 900
cgagcgtgga cgaaccacaa ggtctacacc gattcgttcc gttctatcta ctctatcaac 960
tcgggaatta gccagggcaa ggctgttgcc gtgggtcgtc accccgagga ctcttactac 1020
aatggaaacc cgtggttcct tacgacgttg gctgccgcag agcagttgta cgatgccatt 1080
taccagtggc agaagattgg gtctattacc atcacggacg tctcgtggc cttcttcaag 1140
gatctttaca gctctgcggc cgttgggacg tacgcctctt cgagctcggc cttcacctcc 1200
atcgtgaatg ccgtgaagac gtacgctgat ggatatatga gtattgtgca aacctatgcg 1260
atgacgaatg gctccctttc tgagcagttc ggcaagtccg acggcttctc cctgtccgcc 1320
cgcgatctca cttggtcgta tgcagccctc ctacagcca acttgagaag aaactcggtc 1380
gttccccat cctgggggtga gacgaccgca actagcgtac cctccgtctg ctccgcgacc 1440
tctgccactg gcacctacag taccgccacc aacaccgctg ggcccagcac tctgaccagc 1500
gggactggcg ctacaacgac cacaagcaag gcgacttcta ctactactac ctcgtcggcc 1560
tcgacgacca cagctggatg tgtcgttccg accgcgggtg cagtcacctt tgatgagatt 1620
gctaccacaa cctatggcga gaatgtctac gtggtgggtt ccatctcgca actgggcagc 1680
tgggacacca gcaaggccgt ggccctgagt gccagcaaat acacctccag caacaatctc 1740
tggatgcca cggtcaccct tcccgtggg acgaccttc agtacaagtt catccgggtt 1800
tcgagcagcg ggactgttac atgggagagt gaccgcaacc gttcatacac ggttccgtcc 1860
gcttgtggga cgtctactgc ggtggtgaac accacttggc gctag 1905

<210> 48

<211> 634

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

5 <221> DOMINIO

<222> (41)...(455)

<223> Familia de glicosil hidrolasas 15

<220>

<221> DOMINIO

10 <222> (532)...(628)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (202)...(205)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (204)...(214)

<223> Firma de la región del sitio activo de glucoamilasa. Id del prositio = PS00820

20 <220>

<221> SITIO

<222> (213)...(216)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

25 <221> SITIO

<222> (429)...(432)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (622)...(625)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (639)...(642)

10 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 48

ES 2 575 912 T3

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

ES 2 575 912 T3

Ala Leu Leu Thr Ala Asn Leu Arg Arg Asn Ser Val Val Pro Pro Ser
 450 455 460

Trp Gly Glu Thr Thr Ala Thr Ser Val Pro Ser Val Cys Ser Ala Thr
 465 470 475 480

Ser Ala Thr Gly Thr Tyr Ser Thr Ala Thr Asn Thr Ala Trp Pro Ser
 485 490 495

Thr Leu Thr Ser Gly Thr Gly Ala Thr Thr Thr Thr Ser Lys Ala Thr
 500 505 510

Ser Thr Thr Thr Thr Ser Ser Ala Ser Thr Thr Thr Ala Gly Cys Val
 515 520 525

Val Pro Thr Ala Val Ala Val Thr Phe Asp Glu Ile Ala Thr Thr Thr
 530 535 540

Tyr Gly Glu Asn Val Tyr Val Val Gly Ser Ile Ser Gln Leu Gly Ser
 545 550 555 560

Trp Asp Thr Ser Lys Ala Val Ala Leu Ser Ala Ser Lys Tyr Thr Ser
 565 570 575

Ser Asn Asn Leu Trp Tyr Ala Thr Val Thr Leu Pro Ala Gly Thr Thr
 580 585 590

Phe Gln Tyr Lys Phe Ile Arg Val Ser Ser Ser Gly Thr Val Thr Trp
 595 600 605

Glu Ser Asp Pro Asn Arg Ser Tyr Thr Val Pro Ser Ala Cys Gly Thr
 610 615 620

Ser Thr Ala Val Val Asn Thr Thr Trp Arg
 625 630

<210> 49

<211> 1620

<212> ADN

5 <213> Aspergillus terreus

<400> 49

ES 2 575 912 T3

atgtcctttt tcctgtcctg cctctatctg agcctgtgtg gctcagcact ggcagccaat 60
ctcacgtcat ggaaatccag atccatctac cagacgatga cggacagggt cgcgcgacc 120
gacggctcga ccaccgccgc gtgcaacacc accgccggtt tatactgcgg agggacgtgg 180
cgcggcacga tcaatcatct cgactacatc caggggatgg gcttcgatgc ggtcatgatc 240
tccccatca tcgagaatat cgatggccgc gtctcctacg gtgaagccta ccatgggtac 300
tggccgctgg acctggagtc tttgaacaca cgcttcggca cgaaacagga cctcctcgat 360
ttgagtaatg cgctccactc ccgcgggatg tacctgatga tggataccgt gattaataac 420

atggcatata tcacgcgtgg gcaggaccct gcgacggcaa ttgactactc ggtgttcacg 480
ccgttcaaca acgcggatta ttttcatccc tactgcaaga tcacggattg gaacaacctc 540
accgacgcta gtctgtgtca gacgggtgat ctggaggctg cattgcccga cctgttcaca 600
gagcatacgg acgtgcagga tcgtctcata agctgggcca aggaaatgat ccaaacatat 660
tccatcgacg gacttcgat tgacgccgc aaacacgtcg atcctgagtt tctcgccaga 720
ttcgtgaacg aagtcgatgt attcaccacc ggagaggctc tccagcgcga agtcgacatc 780
atctgcgact accacaacaa atacatcacc agcatgccca attaccgat ctacttctcc 840
atgctggatg ctttcacaga gggcaacacg tcgtcattga tgagccaagt ggaagcaatg 900
aagggtcctt gtcaagcgt taccgccctt gtgtctttt ccgagaacca tgaccagccg 960
cggattccca gcatgaacaa agacatcgcg ctgcacaaa atgttctcac tttcaccatt 1020
ctttttgatg gcattcccat ggtctatcaa ggtcaagagc aacacttggc cggatctggg 1080
acaccgaaaa accgtgaagc cctgtggttg tccaagtatg ataccgaagc cgaattatac 1140
caactgctcg ccaaactcaa cgcgatccgc aaacatgcca cctccctggg cagcgactac 1200
ctctatgcc aaaccagacc tatctaccgg ggcggaagtg agctcgcatt ctacaaaggc 1260
attgagggcc gacaggtgat tacggttcta tcatcgcaag gcgctcaggg aaatccatac 1320
gatctgtatc tgcccgtgtc atataatccg ggaacagcgg tgatggaggt cctcaactgt 1380
gtgaactcca cgggtgggtga cgacggtcag ctcaaggtgc ccatggagaa gggagagccg 1440
cgggtgttct tcccgattga gctgatgggc ggaagcggac tgtgtggata ttccaaggat 1500
aatgttacgg tttctcagtt gaaaacaggg cacgactcga catcccgggg aagcaagatg 1560
acgggcagtg ctgcacttct tatgatgctg tcgttgggag cgagcctggt gttatggtga 1620

<210> 50

<211> 539

<212> PRT

5 <213> *Aspergillus terreus*

<220>

<221> SEÑAL

<222> (1)...(18)

<220>

<221> DOMINIO

<222> (31)...(390)

5 <223> Alfa amilasa, dominio catalítico

<220>

<221> SITIO

<222> (20)...(23)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <220>

<221> SITIO

<222> (49)...(52)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (181)...(184)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

20 <222> (293)...(296)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (469)...(472)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

ES 2 575 912 T3

<220>

<221> SITIO

<222> (508)...(511)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

5 <400> 50

Met Ser Phe Phe Leu Ser Cys Leu Tyr Leu Ser Leu Cys Gly Ser Ala
1 5 10 15

Leu Ala Ala Asn Leu Thr Ser Trp Lys Ser Arg Ser Ile Tyr Gln Thr
20 25 30

Met Thr Asp Arg Phe Ala Arg Thr Asp Gly Ser Thr Thr Ala Ala Cys
35 40 45

Asn Thr Thr Ala Gly Leu Tyr Cys Gly Gly Thr Trp Arg Gly Thr Ile
50 55 60

Asn His Leu Asp Tyr Ile Gln Gly Met Gly Phe Asp Ala Val Met Ile
65 70 75 80

Ser Pro Ile Ile Glu Asn Ile Asp Gly Arg Val Ser Tyr Gly Glu Ala
85 90 95

Tyr His Gly Tyr Trp Pro Leu Asp Leu Glu Ser Leu Asn Thr Arg Phe
100 105 110

Gly Thr Lys Gln Asp Leu Leu Asp Leu Ser Asn Ala Leu His Ser Arg
115 120 125

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

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

Pro Phe Asn Asn Ala Asp Tyr Phe His Pro Tyr Cys Lys Ile Thr Asp
165 170 175

Trp Asn Asn Leu Thr Asp Ala Ser Leu Cys Gln Thr Gly Asp Leu Glu
180 185 190

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

ES 2 575 912 T3

Leu Ile Ser Trp Ala Lys Glu Met Ile Gln Thr Tyr Ser Ile Asp Gly
 210 215 220
 Leu Arg Ile Asp Ala Ala Lys His Val Asp Pro Glu Phe Leu Ala Arg
 225 230 235 240
 Phe Val Asn Glu Val Asp Val Phe Thr Thr Gly Glu Val Leu Gln Arg
 245 250 255
 Glu Val Asp Ile Ile Cys Asp Tyr His Asn Lys Tyr Ile Thr Ser Met
 260 265 270
 Pro Asn Tyr Pro Ile Tyr Phe Ser Met Leu Asp Ala Phe Thr Glu Gly
 275 280 285
 Asn Thr Ser Ser Leu Met Ser Gln Val Glu Ala Met Lys Gly Pro Cys
 290 295 300
 His Asp Val Thr Ala Leu Val Ser Phe Ser Glu Asn His Asp Gln Pro
 305 310 315 320
 Arg Ile Pro Ser Met Asn Lys Asp Ile Ala Leu Ala Lys Asn Val Leu
 325 330 335
 Thr Phe Thr Ile Leu Phe Asp Gly Ile Pro Met Val Tyr Gln Gly Gln
 340 345 350
 Glu Gln His Leu Asp Gly Ser Gly Thr Pro Lys Asn Arg Glu Ala Leu
 355 360 365
 Trp Leu Ser Lys Tyr Asp Thr Gln Ala Glu Leu Tyr Gln Leu Leu Ala
 370 375 380
 Lys Leu Asn Ala Ile Arg Lys His Ala Thr Ser Leu Gly Ser Asp Tyr
 385 390 395 400
 Leu Tyr Ala Gln Thr Arg Pro Ile Tyr Arg Gly Gly Ser Glu Leu Ala
 405 410 415
 Phe Tyr Lys Gly Ile Glu Gly Arg Gln Val Ile Thr Val Leu Ser Ser
 420 425 430
 Gln Gly Ala Gln Gly Asn Pro Tyr Asp Leu Tyr Leu Pro Val Ser Tyr
 435 440 445
 Asn Pro Gly Thr Ala Val Met Glu Val Leu Asn Cys Val Asn Ser Thr
 450 455 460
 Val Gly Asp Asp Gly Gln Leu Lys Val Pro Met Glu Lys Gly Glu Pro
 465 470 475 480
 Arg Val Phe Phe Pro Ile Glu Leu Met Gly Gly Ser Gly Leu Cys Gly

ES 2 575 912 T3

485 490 495
Tyr Ser Lys Asp Asn Val Thr Val Ser Gln Leu Lys Thr Gly His Asp
500 505 510
Ser Thr Ser Arg Gly Ser Lys Met Thr Gly Ser Ala Ala Leu Leu Met
515 520 525
Met Leu Ser Leu Gly Ala Ser Leu Val Leu Trp
530 535

<210> 51

<211> 1824

<212> ADN

5 <213> Aspergillus terreus

<400> 51

ES 2 575 912 T3

```

atgaagtgga ccttctcgct cctcctctta ctgtccgtgt tcggtcaggc tactcatgcc      60
ctgaccccgag cagaatggcg cagccagtca atctacttcc tgttgaccga tcgctttggt      120
cgaacagaca attccacaac tgctgcctgt gacaccaccg acagagtata ctgagggtgt      180
agctggcagg gaatcatcaa ccatctcgat tacatccaag ggatgggatt cactgccatc      240
tggatcaccg cggtcactgg acagtcttac gaaaacacgg gcgacggcac ctctaccat      300
ggatactggc agcaggacat ctacgacctc aactacaact acggaacggc ccaagacctc      360
aagaacctag ccagtgcttt gcacgagcgc ggcatgtatt tgatggttga tgtggttgcc      420
aatcacatgg gctatgacgg agcgggaaac accgtggact acagtgtttt caacccttc      480
tcctcctcca gctactttca ccatactgc ctcatctcca attacgacaa ccagaccaat      540
gttgaagact gctggctggg tgataccacc gtttcgctgc cagatctcga cacgacaagc      600
acagccgtgc gggacatctg gtacgactgg gtggcagact tggtcgcaa ctattccatc      660
gacggctctg gtgtcgacac tgtaaaacac gtcgaaaaag acttttgcc cgactacaac      720
agcgcagcag gcgtctactg tgtcggcgag gtcttttcag gcgatcctgc atacacatgc      780
ccgtaccaga actacatgga cggcgtgctc aactatccaa tctactacca gcttctctat      840
gcgtttgagt cgtccagcgg cagcatcagc gatctctata acatgatcag ctccgttgcc      900
tccagctgca aggatcccac actcctgggt aatttcatcg agaaccacga taacccccgc      960
tttgcttctc acacgagcga ctactcgag gctaagaacg tgatcacctt catcttctg      1020
agcgacggta ttcccatcgt ctacgccgga caggaacagc actacagcgg aggcagcgac      1080
ccagctaacc gcgaggccac ctggctgtct ggatactcca cgagcgccac gctgtacacc      1140
tggatgcct ctacaaacca gatccgcagc ctggcgatct ccaaggacgc gggatacgtg      1200
caggccaaga acaaccctt ctactccgat tccaacacca tcgccatgcg caagggcacg      1260
acagccggcg cgcaagtcac caccgtctc agcaacaagg gcgcctccgg cagctctac      1320
accctctctt tgagcggcac aggctactcc gccggcgcga ccctggtcga gacgtacacc      1380
tgcactacgg tgactgtaga ctcgagcggc aacctgcccg tcccaatgac atccggcttg      1440

ccgcgagtgt ttgtcccgtc gtcctgggtg aatgggagcg cgttttgcaa cacagaatgc      1500
acggccgcca cgtccctccc ggttctcttc gaggaactgg ttacgacgac ctacggggag      1560
aacatttatc tgagcggctc gatcagccag ctgggcagtt ggaatacggc ctcggtgtt      1620
gctctgtcgg cgagtcaata tacctcgtcc aacccgaaat ggtatgtgag tgtgacgttg      1680
cctgtgggca cgctggtcca gtacaagttt atcaagaagg ggtcggacgg gagtggtgtc      1740
tgggagagtg atccgaaccg gtcgtatacc gttccggctg ggtgcgaggg cgcgacgggtg      1800
acagttgctg atacttgag gtga      1824

```

<210> 52

<211> 607

<212> PRT

5 <213> *Aspergillus terreus*

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

5 <221> DOMINIO

<222> (33)...(389)

<223> Alfa amilasa, dominio catalítico

<220>

<221> DOMINIO

10 <222> (505)...(602)

<223> Dominio de enlazamiento de almidón

<220>

<221> SITIO

<222> (44)...(47)

15 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (179)...(182)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

20 <220>

<221> SITIO

<222> (220)...(223)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

25 <221> SITIO

ES 2 575 912 T3

<222> (498)...(501)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (595)...(598)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 52

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

Ala Thr His Ala Leu Thr Pro Ala Glu Trp Arg Ser Gln Ser Ile Tyr
20 25 30

Phe Leu Leu Thr Asp Arg Phe Gly Arg Thr Asp Asn Ser Thr Thr Ala
35 40 45

ES 2 575 912 T3

Ala Cys Asp Thr Thr Asp Arg Val Tyr Cys Gly Gly Ser Trp Gln Gly
50 55 60

Ile Ile Asn His Leu Asp Tyr Ile Gln Gly Met Gly Phe Thr Ala Ile
65 70 75 80

Trp Ile Thr Pro Val Thr Gly Gln Phe Tyr Glu Asn Thr Gly Asp Gly
85 90 95

Thr Ser Tyr His Gly Tyr Trp Gln Gln Asp Ile Tyr Asp Leu Asn Tyr
100 105 110

Asn Tyr Gly Thr Ala Gln Asp Leu Lys Asn Leu Ala Ser Ala Leu His
115 120 125

Glu Arg Gly Met Tyr Leu Met Val Asp Val Val Ala Asn His Met Gly
130 135 140

Tyr Asp Gly Ala Gly Asn Thr Val Asp Tyr Ser Val Phe Asn Pro Phe
145 150 155 160

Ser Ser Ser Ser Tyr Phe His Pro Tyr Cys Leu Ile Ser Asn Tyr Asp
165 170 175

Asn Gln Thr Asn Val Glu Asp Cys Trp Leu Gly Asp Thr Thr Val Ser
180 185 190

Leu Pro Asp Leu Asp Thr Thr Ser Thr Ala Val Arg Asp Ile Trp Tyr
195 200 205

Asp Trp Val Ala Asp Leu Val Ala Asn Tyr Ser Ile Asp Gly Leu Arg
210 215 220

Val Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Asp Tyr Asn
225 230 235 240

Ser Ala Ala Gly Val Tyr Cys Val Gly Glu Val Phe Ser Gly Asp Pro
245 250 255

Ala Tyr Thr Cys Pro Tyr Gln Asn Tyr Met Asp Gly Val Leu Asn Tyr
260 265 270

Pro Ile Tyr Tyr Gln Leu Leu Tyr Ala Phe Glu Ser Ser Ser Gly Ser
275 280 285

Ile Ser Asp Leu Tyr Asn Met Ile Ser Ser Val Ala Ser Ser Cys Lys
290 295 300

Asp Pro Thr Leu Leu Gly Asn Phe Ile Glu Asn His Asp Asn Pro Arg
305 310 315 320

ES 2 575 912 T3

Phe Ala Ser Tyr Thr Ser Asp Tyr Ser Gln Ala Lys Asn Val Ile Thr
 325 330 335
 Phe Ile Phe Leu Ser Asp Gly Ile Pro Ile Val Tyr Ala Gly Gln Glu
 340 345 350
 Gln His Tyr Ser Gly Gly Ser Asp Pro Ala Asn Arg Glu Ala Thr Trp
 355 360 365
 Leu Ser Gly Tyr Ser Thr Ser Ala Thr Leu Tyr Thr Trp Ile Ala Ser
 370 375 380
 Thr Asn Gln Ile Arg Ser Leu Ala Ile Ser Lys Asp Ala Gly Tyr Val
 385 390 395 400
 Gln Ala Lys Asn Asn Pro Phe Tyr Ser Asp Ser Asn Thr Ile Ala Met
 405 410 415
 Arg Lys Gly Thr Thr Ala Gly Ala Gln Val Ile Thr Val Leu Ser Asn
 420 425 430
 Lys Gly Ala Ser Gly Ser Ser Tyr Thr Leu Ser Leu Ser Gly Thr Gly
 435 440 445
 Tyr Ser Ala Gly Ala Thr Leu Val Glu Thr Tyr Thr Cys Thr Thr Val
 450 455 460
 Thr Val Asp Ser Ser Gly Asn Leu Pro Val Pro Met Thr Ser Gly Leu
 465 470 475 480
 Pro Arg Val Phe Val Pro Ser Ser Trp Val Asn Gly Ser Ala Leu Cys
 485 490 495
 Asn Thr Glu Cys Thr Ala Ala Thr Ser Leu Pro Val Leu Phe Glu Glu
 500 505
 Leu Val Thr Thr Thr Tyr Gly Glu Asn Ile Tyr Leu Ser Gly Ser Ile
 515 520 525
 Ser Gln Leu Gly Ser Trp Asn Thr Ala Ser Ala Val Ala Leu Ser Ala
 530 535 540
 Ser Gln Tyr Thr Ser Ser Asn Pro Lys Trp Tyr Val Ser Val Thr Leu
 545 550 555 560
 Pro Val Gly Thr Ser Phe Gln Tyr Lys Phe Ile Lys Lys Gly Ser Asp
 565 570 575
 Gly Ser Val Val Trp Glu Ser Asp Pro Asn Arg Ser Tyr Thr Val Pro
 580 585 590

ES 2 575 912 T3

Ala Gly Cys Glu Gly Ala Thr Val Thr Val Ala Asp Thr Trp Arg
 595 600 605

<210> 53

<211> 1389

<212> ADN

5 <213> Aspergillus terreus

<400> 53

```

atgaagctat cccgcgccct gacagtcttc cttctccacc ttacatccac tgcattggca      60
gcggatgtca atgcgtggaa gtctcgcaat atttactttg ctttgacgga tcgcattgcc      120
cgcagtagcg atgatactgg tggaagtgct tgcggtaacc tgggagacta ctgcggcgga      180
accttccaag gcttgagtc caagcttgac tacatcaagg ggatgggctt tgatgctatc      240
tggatcactc ctgtggtagc taatgctccg ggtggttata atggttattg ggacacaggat      300
ctctatagta tcaactccaa ctatggcacc gctgatgacc tcaaaagcct ggtcaatgct      360
gctcatgaga agggaatgta tgttatggcg gacgtcgtgg ccaaccacat ggggaagtccc      420
atcagcgata acaagcccga gccaatgaac caggagagct cttaccattc agcctgcaca      480
attgactact cagaccagag tagcattgaa gattgccgta tcgcatcaga cctgcccgat      540
gtcaacacag aaagttccga aattcggacc ctcttcagg agtggatcag ctggctggtg      600
aaggaatacc agtttgacgg gctgctgata gacacggta agcatggtga aaaagacttc      660
tggccgggct tctgctctgc cgccggcgtc tacaccatcg gcgaagtctg ggacggcgac      720
ccaaactacc ttgctggata cgcaaacagt atggacgctg tgctcaacta cgcaatctac      780
tatcccatga accgattcta ccagcaacag ggtcctcct cgcacatcgt cagcatgcac      840
gatcaaatca gttccctggt cccaaccca accgccctcg gcacgctcct ggacaaccac      900
gacaacgccc gttggctgag ccagaagaac gacgcctctc tgctgaaaaa cgccctcgcc      960
tatgtcattc tcgcccgcgg catccccatc gtgtactacg gcacggagca gggctacgcc     1020
ggcggcaatg accccgcaa ccgcgaggac ctatggcgca gcaacttcga caccgacgcc     1080
gacttgtagc aggcctatc ccggctctcg gcggcgaggg catcatttgg cgggctcgct     1140
gatgacgacc atgtccatct gtatgtcgct gatacggcgt acgcgtggag cagggccggt     1200
ggggacctga ttgtcctcac gtctaatagt ggatctggct ctgagtctaa gtactgcttc     1260
gattcaaaga agcctgggtg atcatggaac aacacctttg gaacgggaac atatactgct     1320
gatggtgacg ggcaactctg tgtcaccacc tcaaatggtg aaccctggtg gctggttgca     1380
gatgtgtag                                     1389
    
```

<210> 54

<211> 462

10 <212> PRT

<213> Aspergillus terreus

<220>

<221> SEÑAL

<222> (1)...(20)

<220>

5 <221> DOMINIO

<222> (33)...(373)

<223> Alfa amilasa, dominio catalítico

<220>

<221> SITIO

10 <222> (436)...(439)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 54

ES 2 575 912 T3

Met Lys Leu Ser Arg Ala Leu Thr Val Phe Leu Leu His Leu Thr Ser
1 5 10 15

Thr Ala Leu Ala Ala Asp Val Asn Ala Trp Lys Ser Arg Asn Ile Tyr
20 25 30

Phe Ala Leu Thr Asp Arg Ile Ala Arg Ser Ser Asp Asp Thr Gly Gly
35 40 45

Ser Ala Cys Gly Asn Leu Gly Asp Tyr Cys Gly Gly Thr Phe Gln Gly
50 55 60

Leu Gln Ser Lys Leu Asp Tyr Ile Lys Gly Met Gly Phe Asp Ala Ile
65 70 75 80

Trp Ile Thr Pro Val Val Ala Asn Ala Pro Gly Gly Tyr His Gly Tyr
85 90 95

Trp Ala Gln Asp Leu Tyr Ser Ile Asn Ser Asn Tyr Gly Thr Ala Asp
100 105 110

Asp Leu Lys Ser Leu Val Asn Ala Ala His Glu Lys Gly Met Tyr Val
115 120 125

Met Ala Asp Val Val Ala Asn His Met Gly Ser Pro Ile Ser Asp Asn
130 135 140

Lys Pro Glu Pro Met Asn Gln Glu Ser Ser Tyr His Ser Ala Cys Thr
145 150 155 160

Ile Asp Tyr Ser Asp Gln Ser Ser Ile Glu Asp Cys Arg Ile Ala Ser
165 170 175

Asp Leu Pro Asp Val Asn Thr Glu Ser Ser Glu Ile Arg Thr Leu Phe
180 185 190

Gln Glu Trp Ile Ser Trp Leu Val Lys Glu Tyr Gln Phe Asp Gly Leu
195 200 205

Arg Ile Asp Thr Val Lys His Val Glu Lys Asp Phe Trp Pro Gly Phe
210 215 220

Cys Ser Ala Ala Gly Val Tyr Thr Ile Gly Glu Val Trp Asp Gly Asp
225 230 235 240

ES 2 575 912 T3

Pro Asn Tyr Leu Ala Gly Tyr Ala Asn Ser Met Asp Ala Val Leu Asn
 245 250 255
 Tyr Ala Ile Tyr Tyr Pro Met Asn Arg Phe Tyr Gln Gln Gln Gly Ser
 260 265 270
 Ser Ser Asp Ile Val Ser Met His Asp Gln Ile Ser Ser Leu Phe Pro
 275 280 285
 Asn Pro Thr Ala Leu Gly Thr Phe Leu Asp Asn His Asp Asn Ala Arg
 290 295 300
 Trp Leu Ser Gln Lys Asn Asp Ala Ser Leu Leu Lys Asn Ala Leu Ala
 305 310 315 320
 Tyr Val Ile Leu Ala Arg Gly Ile Pro Ile Val Tyr Tyr Gly Thr Glu
 325 330 335
 Gln Gly Tyr Ala Gly Gly Asn Asp Pro Ala Asn Arg Glu Asp Leu Trp
 340 345 350
 Arg Ser Asn Phe Asp Thr Asp Ala Asp Leu Tyr Gln Ala Ile Ser Arg
 355 360 365
 Leu Ser Ala Ala Arg Ala Ser Phe Gly Gly Leu Ala Asp Asp Asp His
 370 375 380
 Val His Leu Tyr Val Ala Asp Thr Ala Tyr Ala Trp Ser Arg Ala Gly
 385 390 395 400
 Gly Asp Leu Ile Val Leu Thr Ser Asn Ser Gly Ser Gly Ser Glu Ser
 405 410 415
 Lys Tyr Cys Phe Asp Ser Lys Lys Pro Gly Gly Ser Trp Asn Asn Thr
 420 425 430
 Phe Gly Thr Gly Thr Tyr Thr Ala Asp Gly Asp Gly Gln Leu Cys Val
 435 440 445
 Thr Thr Ser Asn Gly Glu Pro Val Val Leu Val Ala Asp Val
 450 455 460

<210> 55

<211> 1269

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 55

ES 2 575 912 T3

atgttgaagc agttcacgaa ggcctgatc accctgacga gcctgctggc gctcgtcctc 60
 gtcgcaccgt tggccagcgc gggcccgtg gatggcaaca gcagcgacgt catgttgtag 120
 ggcttcatt ggtactcgtta ccagtcgttt ccgtgggtggg gcgtcatcaa gaacaacgcg 180
 gcgagcatca aggccgacgg cttcaccatg gtgtggctgc cgccgcccag cgacgcggcc 240
 tccaacgagg gctacctgcc gcgccggctc gagctgctgg acagcaagta tggcaccggg 300
 acggacctgg tcaacgccct gtccgcgctg aatgccaatg gtgtgaagcc cattgcggac 360
 atcgtcatca accaccgctg gggcaccacg ggctgggagg acttcacgct gcctccgtgg 420
 ggctcgaacg cgggtgtgccg cggcgacgag tggagcgggg ccacgggcaa cgcggtatcg 480
 ggcgatggct tcaacgccgg gcgcgacatc gatcacacgc agaccttcgt gcaggacggc 540
 atcgtcacct ggatgaacaa ctcgctgaag agcgtcgggt tcgctgggtg gcggtatgac 600
 tacgtgaagg gctacagcgg ctctacgctc ggctcgtaca acaccgcac gacgccgtac 660
 ttctccgtgg gcgagctgtg gacggacctg gacctgaaca accccaacc ccaccgccag 720
 ctgatcatga attggatcga cgcgacgggt ggccgggtccg cggcgttcga cttcacgacc 780
 aagggcctgc tgcagcaggc ggtgcagtac aacgagttct ggctggctgaa ggatgcggcg 840
 ggcgcgccag cgggtgccat tggttggtgg gcagcgaagt ccgtgacctt catcgacaat 900
 cacgacacgg gcccgagcta tccgagcggc ggccagaacc actggccggt ccctgggtgac 960
 aagatcctcc aggggtacgc ctacatcctg actcactctg gcatccccctg cgtgtactgg 1020
 gtgcactaca aggactgggg ccaggcgaac acggacgcca tcaagaagct gatcagcatc 1080
 cgcaagtcca agggcatcac cagcacctcc tcggtgagca tccaggccgc ggacagctcg 1140
 aagtagccg ccatcatcac cggcaacaac ggcaagggtg ccgtgaagat cggcttcggc 1200
 gcctggtctc cgccggggcac ctggacgctg gccacctccg gcaacaacta cgccgtctgg 1260
 acgcagtaa 1269

<210> 56

<211> 422

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<220>

10 <221> SEÑAL

<222> (1)...(27)

<220>

<221> DOMINIO

<222> (36)...(361)

<223> Alfa amilasa, dominio catalítico

<220>

<221> DOMINIO

5 <222> (362)...(422)

<223> Dominio de lámina beta del terminal C de alfa amilasa

<220>

<221> SITIO

<222> (33)...(36)

10 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (188)...(191)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

15 <400> 56

ES 2 575 912 T3

Met Leu Lys Gln Phe Thr Lys Arg Leu Ile Thr Leu Thr Ser Leu Leu
1 5 10 15

Ala Leu Val Leu Val Ala Pro Leu Ala Ser Ala Gly Pro Leu Asp Gly
20 25 30

Asn Ser Ser Asp Val Met Leu Gln Gly Phe His Trp Tyr Ser Tyr Gln
35 40 45

Ser Phe Pro Trp Trp Gly Val Ile Lys Asn Asn Ala Ala Ser Ile Lys
50 55 60

Ala Asp Gly Phe Thr Met Val Trp Leu Pro Pro Pro Ser Asp Ala Ala
65 70 75 80

Ser Asn Glu Gly Tyr Leu Pro Arg Arg Leu Glu Leu Leu Asp Ser Lys
85 90 95

Tyr Gly Thr Arg Thr Asp Leu Val Asn Ala Leu Ser Ala Leu Asn Ala
100 105 110

Asn Gly Val Lys Pro Ile Ala Asp Ile Val Ile Asn His Arg Val Gly
115 120 125

Thr Thr Gly Trp Ala Asp Phe Thr Leu Pro Pro Trp Gly Ser Asn Ala
130 135 140

Val Cys Arg Gly Asp Glu Trp Ser Gly Ala Thr Gly Asn Ala Asp Thr
145 150 155 160

Gly Asp Gly Phe Asn Ala Gly Arg Asp Ile Asp His Thr Gln Thr Phe
165 170 175

Val Gln Asp Gly Ile Val Thr Trp Met Asn Asn Ser Leu Lys Ser Val
180 185 190

Gly Phe Ala Gly Trp Arg Tyr Asp Tyr Val Lys Gly Tyr Ser Gly Ser
195 200 205

Tyr Val Gly Ser Tyr Asn Thr Arg Thr Thr Pro Tyr Phe Ser Val Gly
210 215 220

Glu Leu Trp Thr Asp Leu Asp Leu Asn Asn Pro Asn Pro His Arg Gln
225 230 235 240

Leu Ile Met Asn Trp Ile Asp Ala Thr Gly Gly Arg Ser Ala Ala Phe
245 250 255

Asp Phe Thr Thr Lys Gly Leu Leu Gln Gln Ala Val Gln Tyr Asn Glu

ES 2 575 912 T3

			260					265					270		
Phe	Trp	Arg	Leu	Lys	Asp	Ala	Ala	Gly	Ala	Pro	Ala	Gly	Ala	Ile	Gly
		275					280					285			
Trp	Trp	Ala	Ala	Lys	Ser	Val	Thr	Phe	Ile	Asp	Asn	His	Asp	Thr	Gly
	290					295					300				
Pro	Ser	Tyr	Pro	Ser	Gly	Gly	Gln	Asn	His	Trp	Pro	Phe	Pro	Gly	Asp
305					310					315					320
Lys	Ile	Leu	Gln	Gly	Tyr	Ala	Tyr	Ile	Leu	Thr	His	Ser	Gly	Ile	Pro
				325					330					335	
Cys	Val	Tyr	Trp	Val	His	Tyr	Lys	Asp	Trp	Gly	Gln	Ala	Asn	Thr	Asp
			340					345					350		
Ala	Ile	Lys	Lys	Leu	Ile	Ser	Ile	Arg	Lys	Ser	Lys	Gly	Ile	Thr	Ser
		355					360					365			
Thr	Ser	Ser	Val	Ser	Ile	Gln	Ala	Ala	Asp	Ser	Ser	Lys	Tyr	Ala	Ala
	370					375					380				
Ile	Ile	Thr	Gly	Asn	Asn	Gly	Lys	Val	Ala	Val	Lys	Ile	Gly	Phe	Gly
385					390					395					400
Ala	Trp	Ser	Pro	Pro	Gly	Thr	Trp	Thr	Leu	Ala	Thr	Ser	Gly	Asn	Asn
				405					410					415	
Tyr	Ala	Val	Trp	Thr	Gln										
			420												

<210> 57

<211> 1581

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 57

ES 2 575 912 T3

```

atgccaaact ggtaaaga cgctgttttc tacgagattt accctcagtc cttcaaggac      60
accaactgcg acggtatcgg cgacatcaat ggaatcattg agaagctgga ttatgtgaag    120
gagctgggat gcaatgccct ttggatcaat ccttgtttcg attcaccatt taaggatgcc    180
ggctacgacg tcagagacta caagaaggta gcacccagat atggcacaaa tgctgacctt    240
tacagattat tcggcgaagc ccacaacaga ggcattcacg tactccttga cctggttccg    300
ggccacactt ccgaggaaca tgcgtggttc caggagagca aaaaggctga aaagaatgaa    360
tacacagacc gctatgtatg gaccaactgc tggatccatg gcattgctgg tcatccctat    420
attggcggcg aggtgacag agacggctgc tacatgctga acttctttaa gtgccagcct    480
gcccttaact acggattcct taaccgcacc gatgattggc agtctgcacc tgacgcacct    540

gagtgcacg caacaagaga ggccctcaag gacatcatgc gtttctggct tgaccacggc    600
tgcgatggtt tccgtgttga tatggcagat tctctggtaa aagaggatga tgagaacaag    660
tccgctaccg gcgcaatctg gagaaatatt agagaaatgc ttgataagga ctatcctgaa    720
gctgccattg tttcagagtg gagcaatcct cagcaggccc ttaagagcgg cttccacgct    780
gacttctatc ttgatcacca cggaaacggc tataatacc ttatccgca caatgagacc    840
ccgggcgcg accacagttt ccttaagaag gatggaaacg gcgatatcat gcgtttcctc    900
ctggattacc tgccaagta tgacagcacc aagaatgatg gctacataag ctttatcacc    960
tgcaatcacg aactccaag agcccgcaga acactgggct acgatgagct taagattgca   1020
tgggctcttt tcctaact tcccggagtt ccgttcatct actacggtga tgagattgga   1080
atgagatadc tggacattc tacgaaggaa ggcggataca caagaaccgg caccagaact   1140
ccaatgcagt gggacaactc caagaaccac ggtttctcag atgcaggtgc tgatgcctc   1200
tatcttcac aggatccgtc tggcgtgct cctacagtcg aagatcagga aaaagatcct   1260
tcttactcc ttaacgtgac aaaagagctg acagctcttc gccataaata cactgacctg   1320
caggccgatg gctcttttga cgtcatctat gcagaaaaag agcagttccc atttatctac   1380
aaacgcgga atctgctcct tgcgatcaac ccatcagaaa gcaagtccag tgctgccctt   1440
ccggatgagg catatataaa gaaggatgat acaaaagccg gcaagcttac acctgtatac   1500
tcaatcggtg agttcaagca ggaggatcac agcctcacgc ttcagggcca gtccttcggt   1560
gttttctcac tegaatcata a                                             1581

```

<210> 58

<211> 526

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

ES 2 575 912 T3

<220>

<221> DOMINIO

<222> (12)...(434)

<223> Alfa amilasa, dominio catalítico

5 <220>

<221> SITIO

<222> (170)...(173)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

10 <221> SITIO

<222> (222)...(225)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

15 <222> (431)...(434)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 58

Met Pro Asn Trp Leu Lys Asp Ala Val Phe Tyr Glu Ile Tyr Pro Gln
1 5 10 15

ES 2 575 912 T3

Ser Phe Lys Asp Thr Asn Cys Asp Gly Ile Gly Asp Ile Asn Gly Ile
 20 25 30

Ile Glu Lys Leu Asp Tyr Val Lys Glu Leu Gly Cys Asn Ala Leu Trp
 35 40 45

Ile Asn Pro Cys Phe Asp Ser Pro Phe Lys Asp Ala Gly Tyr Asp Val
 50 55 60

Arg Asp Tyr Lys Lys Val Ala Pro Arg Tyr Gly Thr Asn Ala Asp Leu
 65 70 75 80

Tyr Arg Leu Phe Gly Glu Ala His Asn Arg Gly Ile His Val Leu Leu
 85 90 95

Asp Leu Val Pro Gly His Thr Ser Glu Glu His Ala Trp Phe Gln Glu
 100 105 110

Ser Lys Lys Ala Glu Lys Asn Glu Tyr Thr Asp Arg Tyr Val Trp Thr
 115 120 125

Asn Cys Trp Ile His Gly Ile Ala Gly His Pro Tyr Ile Gly Gly Glu
 130 135 140

Ala Asp Arg Asp Gly Cys Tyr Met Leu Asn Phe Phe Lys Cys Gln Pro
 145 150 155 160

Ala Leu Asn Tyr Gly Phe Leu Asn Arg Thr Asp Asp Trp Gln Ser Ala
 165 170 175

Pro Asp Ala Pro Glu Cys Ile Ala Thr Arg Glu Ala Leu Lys Asp Ile
 180 185 190

Met Arg Phe Trp Leu Asp His Gly Cys Asp Gly Phe Arg Val Asp Met
 195 200 205

Ala Asp Ser Leu Val Lys Glu Asp Asp Glu Asn Lys Ser Ala Thr Gly
 210 215 220

Ala Ile Trp Arg Asn Ile Arg Glu Met Leu Asp Lys Asp Tyr Pro Glu
 225 230 235 240

Ala Ala Ile Val Ser Glu Trp Ser Asn Pro Gln Gln Ala Leu Lys Ser
 245 250 255

Gly Phe His Ala Asp Phe Tyr Leu Asp His His Gly Asn Gly Tyr Asn
 260 265 270

Thr Leu Ile Arg Asp Asn Glu Thr Pro Gly Gly Asp His Ser Phe Leu
 275 280 285

ES 2 575 912 T3

Lys Lys Asp Gly Asn Gly Asp Ile Met Arg Phe Leu Leu Asp Tyr Leu
 290 295 300

Pro Lys Tyr Asp Ser Thr Lys Asn Asp Gly Tyr Ile Ser Phe Ile Thr
 305 310 315 320

Cys Asn His Asp Thr Pro Arg Ala Arg Arg Thr Leu Gly Tyr Asp Glu
 325 330 335

Leu Lys Ile Ala Trp Ala Leu Phe Leu Thr Leu Pro Gly Val Pro Phe
 340 345 350

Ile Tyr Tyr Gly Asp Glu Ile Gly Met Arg Tyr Leu Asp Ile Pro Thr
 355 360 365

Lys Glu Gly Gly Tyr Thr Arg Thr Gly Thr Arg Thr Pro Met Gln Trp
 370 375 380

Asp Asn Ser Lys Asn His Gly Phe Ser Asp Ala Gly Ala Asp Val Leu
 385 390 400

Tyr Leu Pro Gln Asp Pro Ser Gly Asp Ala Pro Thr Val Glu Asp Gln
 405 410 415

Glu Lys Asp Pro Ser Ser Leu Leu Asn Val Thr Lys Glu Leu Thr Ala
 420 425 430

Leu Arg His Lys Tyr Thr Asp Leu Gln Ala Asp Gly Ser Phe Asp Val
 435 440 445

Ile Tyr Ala Glu Lys Glu Gln Phe Pro Phe Ile Tyr Lys Arg Gly Asn
 450 455 460

Leu Leu Leu Ala Ile Asn Pro Ser Glu Ser Lys Ser Ser Ala Ala Leu
 465 470 475 480

Pro Asp Glu Ala Tyr Ile Lys Lys Asp Asp Thr Lys Ala Gly Lys Leu
 485 490 495

Thr Pro Val Tyr Ser Ile Gly Glu Phe Lys Gln Glu Asp His Ser Leu
 500 505 510

Thr Leu Gln Gly Gln Ser Phe Val Val Phe Ser Leu Glu Ser
 515 520 525

<210> 59

<211> 1572

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 59

ES 2 575 912 T3

```

atgcccggat ggttgaagga tgcagtatTT tacgaaatct acccccagtc ctttgccgat      60
tcaaacggcg acgggatcgg tgatctgcag gggatcatcg gaaagcttga ctatgtgaag     120
gagcttggat gcaacgcgct ctggatcaat ccctgttacg attcgccctt taaggatgcc     180
ggctacgacg tccgcgacta caaaaaggTg gccgaaaggT acggcacaaa cgacgatctg     240
aaggagctct tttcaaaggc acatgagaag ggcattccacg tccttctgga cctggTcccc     300
ggccacacct ccgaggaaca cgcttggttt aaggagagtt caaaagccga aaagaacgag     360
ttttccggcc gctatatctg gaccggacac tggctcgagg gcgtaccagg tcatccctgg     420
atcgcgggag aatgcgaaag accggcctgc tatatgctca acttcttcaa atgccagccg     480
gccctgaatt acggtttttt acatcccga aagccctggc agtcctccac ggattcccct     540
gaggcgagag cgacacatga agccatgaag gatgtcatgc gcttctggct ggacgccggc     600
tgcgacggct tccgcgtgga catggcagac agccttgtaa aggatgatga cgagaagaaa     660
tcctgtacct gcgcgctgtg gcgcgaggTg aggcagatgc tcgatgacga ttatccggag     720
gccgccatca tctccgagTg gtccaatccc gaattggcca ttaattccgc acattttcac     780
atggacttct gtctcgacca cgcgagcaac ggttatcata cacttctcag ggattactgg     840
ggtgacgacg gcgacaacag ctttttcaaa gccggcggtc acggcgatgt gatgcgtttc     900
cttgatgact atctgcccag atatgaggcg acaaaggata acggctttat aagccttata     960
tcctgcaatc acgacaccag aaggcctgct tattctctct cagagcgcgga gatcaaggTg    1020
gcctatgcct ttatctatac gatgcccggc gtcccctttg tctactacgg ggacgagata    1080
gcgatgagat atctggatct tcccaccaaa gaaggcggat atgacaggac cggcagcagg    1140
acgccgatgc agtgggatgg cggcgccaac ttcggtttct cctccgctcc cgcggacaaa    1200
ctctacctgc cgcaggatcc ctcaaaggat gctcccacgg cagagaaggc catggccgat    1260
cccgcctctg tatatcacgt gataaaagcc cttctgacca taagacatga gcacaaagca    1320
ttgcaggcgg acggcggatt ggaagtacta tacgccgaaa aagaaaagct tccctttgTt    1380
tacgcaagaa gaagcggcga cgacggcgtg ctggtggcgc tcaatccgtc attaaatgag    1440
gtctcagccc gggTgaacgc cgcagTggat gacgagctcc tctccctcgg tgagacctcc    1500
gtatcctgtg aaggcagcga gagcaccata accatggggc ctcagtcatt catcatatTT    1560
aagctggcct ga                                                              1572

```

<210> 60

<211> 523

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<220>

<221> DOMINIO

ES 2 575 912 T3

<222> (12)...(435)

<223> Alfa amilasa, dominio catalítico

<400> 60

Met Pro Gly Trp Leu Lys Asp Ala Val Phe Tyr Glu Ile Tyr Pro Gln

ES 2 575 912 T3

1 5 10 15
 Ser Phe Ala Asp₂₀ Ser Asn Gly Asp₂₅ Gly₂₅ Ile Gly Asp Leu₃₀ Gln₃₀ Gly Ile
 Ile Gly Lys₃₅ Leu Asp Tyr Val₄₀ Lys₄₀ Glu Leu Gly Cys₄₅ Asn₄₅ Ala Leu Trp
 Ile Asn₅₀ Pro Cys Tyr Asp₅₅ Ser₅₅ Pro Phe Lys Asp₆₀ Ala₆₀ Gly Tyr Asp Val
 Arg Asp Tyr Lys Lys Val₇₀ Ala Glu Arg Tyr₇₅ Gly₇₅ Thr Asn Asp Asp₈₀ Leu₈₀
 Lys Glu Leu Phe Ser₈₅ Lys Ala His Glu₉₀ Lys₉₀ Gly Ile His Val₉₅ Leu₉₅ Leu
 Asp Leu Val₁₀₀ Pro Gly His Thr Ser₁₀₅ Glu₁₀₅ Glu His Ala Trp₁₁₀ Phe₁₁₀ Lys Glu
 Ser Ser Lys₁₁₅ Ala Glu Lys Asn₁₂₀ Glu₁₂₀ Phe Ser Gly Arg Tyr₁₂₅ Ile Trp Thr
 Gly His₁₃₀ Trp Leu Glu Gly Val₁₃₅ Pro Gly His Pro₁₄₀ Trp₁₄₀ Ile Ala Gly Glu
 Cys₁₄₅ Glu Arg Pro Ala Cys₁₅₀ Tyr Met Leu Asn₁₅₅ Phe₁₅₅ Phe Lys Cys Gln₁₆₀ Pro₁₆₀
 Ala Leu Asn Tyr Gly₁₆₅ Phe Leu His Pro₁₇₀ Glu₁₇₀ Lys Pro Trp Gln₁₇₅ Ser₁₇₅ Ser
 Thr Asp Ser₁₈₀ Pro Glu Ala Arg Ala₁₈₅ Thr₁₈₅ His Glu Ala Met₁₉₀ Lys₁₉₀ Asp Val
 Met Arg Phe₁₉₅ Trp Leu Asp Ala Gly₂₀₀ Cys Asp Gly Phe₂₀₅ Arg₂₀₅ Val Asp Met
 Ala Asp₂₁₀ Ser Leu Val Lys Asp₂₁₅ Asp Asp Glu Lys Lys₂₂₀ Ser Cys Thr Cys
 Ala Leu Trp Arg Glu Val₂₃₀ Arg Gln Met Leu Asp₂₃₅ Asp Asp Tyr Pro Glu₂₄₀
 Ala Ala Ile Ile Ser₂₄₅ Glu Trp Ser Asn₂₅₀ Pro Glu Leu Ala Ile Asn₂₅₅ Ser
 Ala His Phe His₂₆₀ Met Asp Phe Cys Leu₂₆₅ Asp His Ala Ser Asn₂₇₀ Gly Tyr
 His Thr Leu₂₇₅ Leu Arg Asp Tyr Trp₂₈₀ Gly Asp Asp Gly Asp₂₈₅ Asn Ser Phe

ES 2 575 912 T3

Phe Lys Ala Gly Gly His Gly Asp Val Met Arg Phe Leu Asp Asp Tyr
 290 295 300

Leu Pro Arg Tyr Glu Ala Thr Lys Asp Asn Gly Phe Ile Ser Leu Ile
 305 310 315 320

Ser Cys Asn His Asp Thr Arg Arg Pro Ala Tyr Ser Leu Ser Glu Arg
 325 330 335

Glu Ile Lys Val Ala Tyr Ala Phe Ile Tyr Thr Met Pro Gly Val Pro
 340 345 350

Phe Val Tyr Tyr Gly Asp Glu Ile Ala Met Arg Tyr Leu Asp Leu Pro
 355 360 365

Thr Lys Glu Gly Gly Tyr Asp Arg Thr Gly Ser Arg Thr Pro Met Gln
 370 375 380

Trp Asp Gly Gly Ala Asn Phe Gly Phe Ser Ser Ala Pro Ala Asp Lys
 385 390 395 400

Leu Tyr Leu Pro Gln Asp Pro Ser Lys Asp Ala Pro Thr Ala Glu Lys
 405 410 415

Ala Met Ala Asp Pro Ala Ser Val Tyr His Val Ile Lys Ala Leu Leu
 420 425 430

Thr Ile Arg His Glu His Lys Ala Leu Gln Ala Asp Gly Gly Leu Glu
 435 440 445

Val Leu Tyr Ala Glu Lys Glu Lys Leu Pro Phe Val Tyr Ala Arg Arg
 450 455 460

Ser Gly Asp Asp Gly Val Leu Val Ala Leu Asn Pro Ser Leu Asn Glu
 465 470 475 480

Val Ser Ala Arg Val Asn Ala Ala Val Asp Asp Glu Leu Leu Ser Leu
 485 490 495

Gly Glu Thr Ser Val Ser Cys Glu Gly Ser Glu Ser Thr Ile Thr Met
 500 505 510

Gly Pro Gln Ser Phe Ile Ile Phe Lys Leu Ala
 515 520

<210> 61

<211> 2061

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 61

ES 2 575 912 T3

atggtagcag gctttgggct ctacggagct gccttgctca cccaatggc ggcacaggcc 60
gaaacgggca gagaagtcac gctgcagggc ttaactgga attccacctc gcaagcgggt 120
gggcactata atgagctggc aaatcgtgct ggtgaaatcg ctggggctgg tatcgacatt 180
gtttggttcc cacctccatc cagagccgca gaccgggttg gatacctgcc caacgagtgg 240
tataatatga actccaatta tggaaatcgg acgactctac aggcggtat cggaatcctt 300
cgcaacaacg gtgtcaaac cgctcgggat attgtcgta accaccgtgt gggcacaacg 360
aattgggagg acttcacgaa tccttctttc ggtgataaca accgtgcat cacacgcgat 420
gatgaatggc atcagtcgtc gggtaactgg gataccggtg aggcgtatag tgccgccgt 480
gacttgacc acacctacgg ccccggtgca aatgagatta aaaactggct gaattggttg 540
aagagtgaca tcggctttga tggttggcgc tatgatatgg tcaagggtt cagcgggtac 600
tatgtgggag agtacaacac agcaactagc ccctacattt ccgttgggtga gttcttcgac 660
tatgatcgcc agaaggtggc cgctggatc aacgccacca atgctcgtc tcgcccctc 720
gattttcaa cagcaacct cctctatgtg gctgtcacc aaaacaacta cggagttctt 780
cgtgatggtg aaggtaaggc caatggactc atcggttgg ggcacaacg ggctatcaca 840
ttcattgaga accatgatac ggaagaagcc cgcaacggg agtacacccc cgccttcccc 900
caatgggcta ccatgcaggg ttatgcctac attttgacgc acccggcat tccatgcgtt 960
ttctggaatg attggcggtg ggacttccgt tcggaaatca atcaactgat cgccatccgc 1020
agggcccaag gcatcaacga tggcagcagc ctacagatcc aggttgccga tggcagccgc 1080
tacggggcga tcatcaatgg aacacggcg gtcaagattg gtcctggaaa ctggagccca 1140
agtggtagtt ggacgctggc tgccgccgga accaattacg ccgtctggac aaacgggtggc 1200
ggaacaccaa cgccaacacc cacaccaacc ccaactcagg gcccacaac ggttacttgg 1260
aatccatcca ctccgacagc cggacaaaat gtgacgatca cttatcctc cggccgctcc 1320
cttgccagtt cctccaacgt caacctctac tggggtgtca atggatggac caatgtcca 1380
accaaagcga tgaccaagaa cagttccaat gactggacaa ccacgatcac acttccctcc 1440
aacacaacac gtttgaactt tgtgttcaat aacggctcca gttgggacaa caacagcagt 1500
caggattgga atgtcaacgt gaccgctgta acgcccacac caacaccgac accgactccc 1560
acacctacgg caactccgac accgacacc acccggcgc cactgccac acccaccaca 1620
actccgactc ctctgatcat ttggtatcgc atcgaagccc gtcacagtgg taaagttcta 1680
gacgttgcca gtgcatcaac ttcaaacggg ggcaatgtgc atcagtggc ttatgctgg 1740
ggacagaacc aacaatggcg tgtgtcgcac gctgggaacg ggtttgttta catcctaac 1800
cggaacagtg gaaaagcact tgaagtggc aacttctcca ccagtaacgg cgggaacgctc 1860
cagcaatggg attatgccgg tggttccagt caacagtgga agctcattga gacgaccaac 1920
ggatatgtgc aaatccagaa tcggaacagc ggcaaagcca ttgatgtttc cgctgcttcg 1980
actaccaatg gtgccaacat ccatcagtgg acctacggcg gcggcaacaa tcagcagtgg 2040
aagctgatcc caatcaatta a 2061

<210> 62

<211> 686

<212> PRT

<213> Desconocido

5 <220>

<223> Obtenido a partir de muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(20)

10 <220>

<221> DOMINIO

<222> (25)...(340)

<223> Alfa amilasa, dominio catalítico

<220>

15 <221> DOMINIO

<222> (341)...(398)

<223> Dominio de hoja beta del terminal C de alfa amilasa

<220>

<221> DOMINIO

20 <222> (547)...(680)

<223> Dominio de lectina de trébol beta tipo ricina

<220>

<221> SITIO

<222> (34)...(37)

25 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (90)...(93)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

5 <220>

<221> SITIO

<222> (234)...(237)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

10 <221> SITIO

<222> (436)...(439)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

15 <222> (474)...(477)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (488)...(491)

20 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (498)...(501)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

25 <220>

<221> SITIO

<222> (504)...(507)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

5 <221> SITIO

<222> (513)...(516)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

10 <222> (620)...(623)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 62

ES 2 575 912 T3

Met Val Ala Gly Phe Gly Leu Tyr Gly Ala Ala Leu Leu Thr Pro Met
1 5 10 15

Ala Ala Gln Ala Glu Thr Gly Arg Glu Val Met Leu Gln Gly Phe Asn
20 25 30

Trp Asn Ser Thr Ser Gln Ala Gly Gly His Tyr Asn Glu Leu Ala Asn
35 40 45

Arg Ala Gly Glu Ile Ala Gly Ala Gly Ile Asp Ile Val Trp Phe Pro
50 55 60

Pro Pro Ser Arg Ala Ala Asp Arg Val Gly Tyr Leu Pro Asn Glu Trp
65 70 75 80

Tyr Asn Met Asn Ser Asn Tyr Gly Asn Arg Thr Thr Leu Gln Ala Ala
85 90 95

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

Val Asn His Arg Val Gly Thr Thr Asn Trp Ala Asp Phe Thr Asn Pro
115 120 125

Ser Phe Gly Asp Asn Asn Arg Ala Ile Thr Arg Asp Asp Glu Trp His
130 135 140

Gln Ser Ser Gly Asn Trp Asp Thr Gly Glu Ala Tyr Ser Ala Ala Arg
145 150 155 160

Asp Leu Asp His Thr Tyr Gly Pro Val Gln Asn Glu Ile Lys Asn Trp
165 170 175

Leu Asn Trp Leu Lys Ser Asp Ile Gly Phe Asp Gly Trp Arg Tyr Asp
180 185 190

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

Thr Ser Pro Tyr Ile Ser Val Gly Glu Phe Phe Asp Tyr Asp Arg Gln
210 215 220

Lys Val Val Gly Trp Ile Asn Ala Thr Asn Ala Arg Ser Arg Ala Phe

ES 2 575 912 T3

225					230						235	.			240
Asp	Phe	Pro	Thr	Arg	Asn	Leu	Leu	Tyr	Val	Ala	Val	Thr	Gln	Asn	Asn
				245					250					255	
Tyr	Gly	Val	Leu	Arg	Asp	Gly	Glu	Gly	Lys	Ala	Asn	Gly	Leu	Ile	Gly
			260					265					270		
Trp	Trp	Pro	Gln	Arg	Ala	Ile	Thr	Phe	Ile	Glu	Asn	His	Asp	Thr	Glu
		275					280					285			
Glu	Ala	Arg	Asn	Gly	Glu	Tyr	Thr	Pro	Ala	Phe	Pro	Gln	Trp	Ala	Thr
	290					295					300				
Met	Gln	Gly	Tyr	Ala	Tyr	Ile	Leu	Thr	His	Pro	Gly	Ile	Pro	Cys	Val
305					310					315					320
Phe	Trp	Asn	Asp	Trp	Arg	Trp	Asp	Phe	Arg	Ser	Glu	Ile	Asn	Gln	Leu
				325					330					335	
Ile	Ala	Ile	Arg	Arg	Ala	Gln	Gly	Ile	Asn	Asp	Gly	Ser	Ser	Leu	Ser
			340					345					350		
Ile	Gln	Val	Ala	Asp	Gly	Ser	Arg	Tyr	Gly	Ala	Ile	Ile	Asn	Gly	Asn
		355					360					365			
Thr	Ala	Val	Lys	Ile	Gly	Pro	Gly	Asn	Trp	Ser	Pro	Ser	Gly	Ser	Trp
	370					375					380				
Thr	Leu	Ala	Ala	Ala	Gly	Thr	Asn	Tyr	Ala	Val	Trp	Thr	Asn	Gly	Gly
385					390					395					400
Gly	Thr	Pro	Thr	Gln	Gly	Pro	Thr								
				405					410					415	
Thr	Val	Thr	Trp	Asn	Pro	Ser	Thr	Pro	Thr	Ala	Gly	Gln	Asn	Val	Thr
			420					425					430		
Ile	Thr	Tyr	Pro	Ser	Gly	Arg	Ser	Leu	Ala	Ser	Ser	Ser	Asn	Val	Asn
		435					440						445		
Leu	Tyr	Trp	Gly	Val	Asn	Gly	Trp	Thr	Asn	Val	Gln	Thr	Lys	Ala	Met
	450					455					460				
Thr	Lys	Asn	Ser	Ser	Asn	Asp	Trp	Thr	Thr	Thr	Ile	Thr	Leu	Pro	Ser
465					470					475					480
Asn	Thr	Thr	Arg	Leu	Asn	Phe	Val	Phe	Asn	Asn	Gly	Ser	Ser	Trp	Asp
				485					490					495	
Asn	Asn	Ser	Ser	Gln	Asp	Trp	Asn	Val	Asn	Val	Thr	Ala	Val	Thr	Pro
			500					505					510		

ES 2 575 912 T3

Thr Pro Thr Pro Thr Pro Thr Pro Thr Pro Thr Ala Thr Pro Thr Pro
 515 520 525

Thr Pro Thr Pro Thr Pro Thr Ala Thr Pro Thr Pro Thr Pro Thr Pro
 530 535 540

Ser Ile Ile Trp Tyr Arg Ile Glu Ala Arg His Ser Gly Lys Val Leu
 545 550 555 560

Asp Val Ala Ser Ala Ser Thr Ser Asn Gly Gly Asn Val His Gln Trp
 565 570 575

Ser Tyr Ala Gly Gly Gln Asn Gln Gln Trp Arg Val Val Asp Ala Gly
 580 585 590

Asn Gly Phe Val Tyr Ile Leu Asn Arg Asn Ser Gly Lys Ala Leu Glu
 595 600 605

Val Gly Asn Phe Ser Thr Ser Asn Gly Gly Asn Val Gln Gln Trp Asp
 610 615 620

Tyr Ala Gly Gly Ser Ser Gln Gln Trp Lys Leu Ile Glu Thr Thr Asn
 625 630 635

Gly Tyr Val Gln Ile Gln Asn Arg Asn Ser Gly Lys Ala Ile Asp Val
 645 650 655

Ser Ala Ala Ser Thr Thr Asn Gly Ala Asn Ile His Gln Trp Thr Tyr
 660 665 670

Gly Gly Gly Asn Asn Gln Gln Trp Lys Leu Ile Pro Ile Asn
 675 680 685

<210> 63

<211> 1611

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 63

ES 2 575 912 T3

```

atgacagagt ggtggcgtgg tgcagtgacc tatcaagtct atccaaggtc gtttcaggac      60
agcaacggcg acggcatcgg cgacctgccc ggcatacccg cccggcttga gtatctggcc      120
gatcttggcg tggacgcggt ctggctgtca ccgttcttca aaagcccgat gaaggacatg      180
ggctatgacg tcagcgacta ttgcgatgtc gatccggctct tcggcaccct cgccgatttt      240
gacgcccctgc tggcccgcgc gcatgagctg gggctcaagg tgatcatcga ccaggtcctt      300
agccacagtt ccgacctgca ccttgctttt gtgaccagtc gcagcgaccg cgtgaacccg      360
aaggcggact ggtatgtctg ggccgatccc aagcccgcgc gcagcccgcc caacaactgg      420

ctgtcgggtg tcggtggctc ggcattggcc tgggacgcgc gcagaaaaca gtattacctg      480
cacaatttcc tgaccagcca gccggacctg aactaccaca acccgaaggc gcaggactgg      540
gcgctggaca acatgcgttt ctggctggac cggggcgtgg acgggttccg ctttgacacc      600
gtcaactact tcttccacga tcccttggtg cgcagcaacc ctgccgatca ccgcaacaag      660
cctgaggctg acggcaatcc ctacggcatg cagtaccacc tgcatgacaa gaaccagccc      720
gagaacctga tctggatgga gcggatacgg gtgcttctgg accaatacgg tgccgcaagc      780
gtcggcgaga tgggcgaaaag tcaccacgcc atccggatga tgggcgacta caccgctccg      840
gggcggtctc atcaatgcta cagctttgaa ttcatggggg atgaatacac cgcaaacctg      900
ttccgggacc ggatagaaaag ctttttcaag ggtgccccta aaggctggcc gatgtgggcg      960
ttttcaaacc acgatgtcgt ccgccatgtc agtcgctggg caaaacatgg cctcaccccc     1020
gaggcgggtg ccaagcagac aggtgcggtg cttctgtcgc ttgagggtc gatctgcctg     1080
tgggagggcg aggagctggg ccagaccgat accgaactgg ccttgatga gttgaccgat     1140
ccgcagggca tegtctttt gcccgaaacc atcggccgcg acaatactcg gacgccaatg     1200
gtttgggacg catcgccgca tggcgggttt tcgaccgtca caccctggct gccggtgaaa     1260
ccggaacagg ccgcgcgtca tgtggccggg caaacgggtg atgccgcctc ggtgctggaa     1320
agctaccggg cgatgctggc cttccggcgc gctgaaccgg cccttaggac cgggcggacg     1380
cggtttctgg atctggccga accggttctg ggctttgtgc gcggcgaagg ggaggggtcg     1440
atcctgtgcc tgttcaatct gtcgcctggt gcgcgggggg ttgcggtcga aggcgtgggc     1500
ccgccgatcg gcccgggcca gcaggctatc ctttcgggcg gacggctagg ccttgccccg     1560
aacggcgccg ccttctctgc ggtgaccgga acagtccgcg ttctggacta a              1611

```

<210> 64

<211> 536

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<220>

ES 2 575 912 T3

<221> DOMINIO

<222> (12)...(449)

<223> Alfa amilasa, dominio catalítico

<400> 64

Met Thr Glu Trp Trp Arg Gly Ala Val Thr Tyr Gln Val Tyr Pro Arg
1 5 10 15

Ser Phe Gln Asp Ser Asn Gly Asp Gly Ile Gly Asp Leu Pro Gly Ile
20 25 30

Thr Ala Arg Leu Glu Tyr Leu Ala Asp Leu Gly Val Asp Ala Val Trp
35 40 45

5 Leu Ser Pro Phe Phe Lys Ser Pro Met Lys Asp Met Gly Tyr Asp Val
50 55 60

ES 2 575 912 T3

Ser Asp Tyr Cys Asp Val Asp Pro Val Phe Gly Thr Leu Ala Asp Phe
 65 70 75 80
 Asp Ala Leu Leu Ala Arg Ala His Glu Leu Gly Leu Lys Val Ile Ile
 85 90 95
 Asp Gln Val Leu Ser His Ser Ser Asp Leu His Pro Ala Phe Val Thr
 100 105 110
 Ser Arg Ser Asp Arg Val Asn Pro Lys Ala Asp Trp Tyr Val Trp Ala
 115 120 125
 Asp Pro Lys Pro Asp Gly Ser Pro Pro Asn Asn Trp Leu Ser Val Phe
 130 135 140
 Gly Gly Ser Ala Trp Ala Trp Asp Ala Arg Arg Lys Gln Tyr Tyr Leu
 145 150 155 160
 His Asn Phe Leu Thr Ser Gln Pro Asp Leu Asn Tyr His Asn Pro Lys
 165 170 175
 Val Gln Asp Trp Ala Leu Asp Asn Met Arg Phe Trp Leu Asp Arg Gly
 180 185 190
 Val Asp Gly Phe Arg Phe Asp Thr Val Asn Tyr Phe Phe His Asp Pro
 195 200 205
 Leu Leu Arg Ser Asn Pro Ala Asp His Arg Asn Lys Pro Glu Ala Asp
 210 215 220
 Gly Asn Pro Tyr Gly Met Gln Tyr His Leu His Asp Lys Asn Gln Pro
 225 230 235 240
 Glu Asn Leu Ile Trp Met Glu Arg Ile Arg Val Leu Leu Asp Gln Tyr
 245 250 255
 Gly Ala Ala Ser Val Gly Glu Met Gly Glu Ser His His Ala Ile Arg
 260 265 270
 Met Met Gly Asp Tyr Thr Ala Pro Gly Arg Leu His Gln Cys Tyr Ser
 275 280 285
 Phe Glu Phe Met Gly Tyr Glu Tyr Thr Ala Asn Leu Phe Arg Asp Arg
 290 295 300
 Ile Glu Ser Phe Phe Lys Gly Ala Pro Lys Gly Trp Pro Met Trp Ala
 305 310 315 320
 Phe Ser Asn His Asp Val Val Arg His Val Ser Arg Trp Ala Lys His
 325 330 335

ES 2 575 912 T3

Gly Leu Thr Pro Glu Ala Val Ala Lys Gln Thr Gly Ala Leu Leu Leu
 340 345 350

Ser Leu Glu Gly Ser Ile Cys Leu Trp Glu Gly Glu Glu Leu Gly Gln
 355 360 365

Thr Asp Thr Glu Leu Ala Leu Asp Glu Leu Thr Asp Pro Gln Gly Ile
 370 375 380

Val Phe Trp Pro Glu Pro Ile Gly Arg Asp Asn Thr Arg Thr Pro Met
 385 390 395 400

Val Trp Asp Ala Ser Pro His Gly Gly Phe Ser Thr Val Thr Pro Trp
 405 410 415

Leu Pro Val Lys Pro Glu Gln Ala Ala Arg His Val Ala Gly Gln Thr
 420 425 430

Gly Asp Ala Ala Ser Val Leu Glu Ser Tyr Arg Ala Met Leu Ala Phe
 435 440 445

Arg Arg Ala Glu Pro Ala Leu Arg Thr Gly Arg Thr Arg Phe Leu Asp
 450 455 460

Leu Ala Glu Pro Val Leu Gly Phe Val Arg Gly Glu Gly Glu Gly Ala
 465 470 475 480

Ile Leu Cys Leu Phe Asn Leu Ser Pro Val Ala Arg Gly Val Ala Val
 485 490 495

Glu Gly Val Gly Pro Pro Ile Gly Pro Gly Gln Gln Ala Ile Leu Ser
 500 505 510

Gly Gly Arg Leu Gly Leu Gly Pro Asn Gly Ala Ala Phe Leu Arg Val
 515 520 525

Thr Gly Thr Val Arg Val Leu Asp
 530 535

<210> 65

<211> 1386

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 65

ES 2 575 912 T3

atgaagttga agtaccttgc cttagttttg ttggctgtgg cttcgatagg cctactctcg 60
actccagtgg gtgctgccaa gtactccgaa ctccaagagg gcggtgttat aatgcaggcc 120
ttctactggg atgttcccgg agggggaatc tgggtgggaca ccataagaca gaaaatcccg 180
gagtggtagc acgctggaat ctcggcgata tggattcctc cagctagcaa agggatgggc 240

ggtggttat ccatgggcta cgatccctac gatttctttg acctcggcga gtactatcag 300
aagggaacag ttgagacgcg cttcggctca aaggaggaac tggatgaacat gataaacacc 360
gcacactcct atggcataaa ggtgatagcg gacatagtca taaaccaccg cgccggtgga 420
gaccttgagt ggaacccctt tgtaaacaac tatacttggga cagacttctc caaggctgcc 480
tccggtaaat acacggccaa ctaccttgac ttccacccaa acgaggtcaa gtgctgcat 540
gagggtacat ttggtgactt tccggacatc gcccacgaga agagctggga tcagtactgg 600
ctctgggcaa gcaatgagag ctacgccgca tatctccgga gcatagggat cgatgcatgg 660
cgtttcgact acgtcaaagg ttacggagcg tgggttgta atgactggct cagctggtgg 720
ggaggctggg ccggttgaga gtactgggac acgaacgttg atgactcct taactgggca 780
tacgacagcg gtgccaaggt ctttgacttc ccgctctact acaagatgga cgaagccttt 840
gacaacacca acatccccgc tttggtttac gccctccaga acggaggaac agtcgtttcc 900
cgcgatccct tcaaggcagt aactttcgtt gcccaaccacg atacagatat aatctggaac 960
aagtatccgg cttatgcggt catccttacc tatgaggac agcctgttat attttaccgc 1020
gactacgagg agtggctcaa caaggataag cttacaacc ttatctggat acacgagcac 1080
cttgccggag gaagtaccaa gatcctctac tacgataacg atgagctaat attcagagg 1140
gagggtacg ggagcaagcc gggcctcata acctacataa acctcggaaa cgactgggcc 1200
gagcgtggg tgaacgtcgg ctcaaagtt gccggctaca caatccatga atacacaggc 1260
aatctcgtg gctgggttga cagggtgggt cagtacgat gatgggttaa actgacggca 1320
cctcctcatg atccagccaa cggatattac ggctactcag tctggagcta cgcaggcgtc 1380
ggatga 1386

<210> 66

<211> 461

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<220>

<221> SEÑAL

10 <222> (1)...(25)

<220>

ES 2 575 912 T3

<221> DOMINIO

<222> (35)...(368)

<223> Alfa amilasa, dominio catalítico

<220>

5 <221> SITIO

<222> (152)...(155)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

10 <222> (208)...(211)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 66

Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val Ala Ser Ile

ES 2 575 912 T3

1		5						10						15	
Gly	Leu	Leu	Ser ₂₀	Thr	Pro	Val	Gly	Ala ₂₅	Ala	Lys	Tyr	Ser	Glu ₃₀	Leu	Glu
Glu	Gly	Gly ₃₅	Val	Ile	Met	Gln	Ala ₄₀	Phe	Tyr	Trp	Asp	Val ₄₅	Pro	Gly	Gly
Gly	Ile ₅₀	Trp	Trp	Asp	Thr	Ile ₅₅	Arg	Gln	Lys	Ile	Pro ₆₀	Glu	Trp	Tyr	Asp
Ala ₆₅	Gly	Ile	Ser	Ala	Ile ₇₀	Trp	Ile	Pro	Pro	Ala ₇₅	Ser	Lys	Gly	Met	Gly ₈₀
Gly	Gly	Tyr	Ser	Met ₈₅	Gly	Tyr	Asp	Pro	Tyr ₉₀	Asp	Phe	Phe	Asp	Leu ₉₅	Gly
Glu	Tyr	Tyr	Gln ₁₀₀	Lys	Gly	Thr	Val	Glu ₁₀₅	Thr	Arg	Phe	Gly	Ser ₁₁₀	Lys	Glu
Glu	Leu	Val ₁₁₅	Asn	Met	Ile	Asn	Thr ₁₂₀	Ala	His	Ser	Tyr	Gly ₁₂₅	Ile	Lys	Val
Ile	Ala ₁₃₀	Asp	Ile	Val	Ile	Asn ₁₃₅	His	Arg	Ala	Gly	Gly ₁₄₀	Asp	Leu	Glu	Trp
Asn ₁₄₅	Pro	Phe	Val	Asn	Asn ₁₅₀	Tyr	Thr	Trp	Thr	Asp ₁₅₅	Phe	Ser	Lys	Val	Ala ₁₆₀
Ser	Gly	Lys	Tyr	Thr ₁₆₅	Ala	Asn	Tyr	Leu	Asp ₁₇₀	Phe	His	Pro	Asn	Glu ₁₇₅	Val
Lys	Cys	Cys	Asp ₁₈₀	Glu	Gly	Thr	Phe	Gly ₁₈₅	Asp	Phe	Pro	Asp	Ile ₁₉₀	Ala	His
Glu	Lys	Ser ₁₉₅	Trp	Asp	Gln	Tyr	Trp ₂₀₀	Leu	Trp	Ala	Ser	Asn ₂₀₅	Glu	Ser	Tyr
Ala	Ala ₂₁₀	Tyr	Leu	Arg	Ser	Ile ₂₁₅	Gly	Ile	Asp	Ala	Trp ₂₂₀	Arg	Phe	Asp	Tyr
Val ₂₂₅	Lys	Gly	Tyr	Gly	Ala ₂₃₀	Trp	Val	Val	Asn	Asp ₂₃₅	Trp	Leu	Ser	Trp	Trp ₂₄₀
Gly	Gly	Trp	Ala	Val ₂₄₅	Gly	Glu	Tyr	Trp	Asp ₂₅₀	Thr	Asn	Val	Asp	Ala ₂₅₅	Leu
Leu	Asn	Trp	Ala ₂₆₀	Tyr	Asp	Ser	Gly	Ala ₂₆₅	Lys	Val	Phe	Asp	Phe ₂₇₀	Pro	Leu
Tyr	Tyr	Lys ₂₇₅	Met	Asp	Glu	Ala	Phe ₂₈₀	Asp	Asn	Thr	Asn	Ile ₂₈₅	Pro	Ala	Leu

ES 2 575 912 T3

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

Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn
 305 310 315 320

Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val
 325 330 335

Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn
 340 345 350

Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr Lys Ile
 355 360 365

Leu Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Met Arg Glu Gly Tyr Gly
 370 375 380

Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Asn Asp Trp Ala
 385 390 400

Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr Thr Ile His
 405 410 415

Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Trp Val Gln Tyr
 420 425 430

Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala Asn Gly
 435 440 445

Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly
 450 455 460

<210> 67

<211> 2022

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 67

ES 2 575 912 T3

ttgaaaaaaa acaccattag cgccctggtc gcaggatgg tattaggctt tgcaccaac 60
 gcaatggcgg ttcctagaac cgcttttgta cacctctttg aatggaaatg ggaagatggt 120
 gcacaggagt gtgaaacatt tctcggacct aaaggctttg ccgcagtgca agtctctccg 180
 ccaactaaat ctcaaacac ggatgcatgg tggggccggt atcaaccgt tagttatgct 240
 tttgaaggac gcagcggtaa tcgcagccaa tttaaaaata tggtgcaacg ttgtaaagct 300
 gtaggcgtcg atatatacgt agatgcagtg attaaccaca tggcagccta cgacagaaat 360
 ttcctgatg taccctatag cagtaatgac tttaaactct gtacaggaga tattgactat 420

 aataaccgtt ggcaaacaca gcattgtgat ttagtccggtc ttaatgatct aaaaacagga. 480
 tctgactacg tccgcaaaa aatagcggat tatatgaacg acgcaatcag tatgggtgta 540
 gctggtttcc gtattgatgc agccaaacat ataccagcag gtgatatagc tgccattaaa 600
 ggtaaatata atggtaatcc atacatcttc caagaggtaa ttgggtgatc cggcgaacct 660
 gttcgaccga ctgaatacac ctttatcggg ggtgtcacgg aatttcaatt tgctcgaaaa 720
 ttgggtccag cttccgcaa tagtaatatt gcttggttaa aagacattgg cagtcaaatg 780
 gaattatcca gtgctgatgc cgtaacattt gtaacgaatc atgatgaaga gcgtcataac 840
 ccgaatggtc ctatttggca cggcgttcaa ggtaatgggt atgcattagc aaatattttc 900
 accttagctt acccttacgg ctatcaaaa atcatgtcag gatacttctt ccacggtgac 960
 tttaacgcag ctccaccaag cagtggata cacacaggaa atgctgtggt ttttgatggc 1020
 ggagactggg tatgcgaaca caaatggcgc ggtattgcta acatggttgc cttccgcaac 1080
 tatacagcaa gcgaatggcg tatcagtaat tgggtggcaaa acagtaacga ccaaattgct 1140
 tttggtcgcg gtggtttagg ttttggtgtt attaataaac gtgctaattg tagcattaat 1200
 caaagttttg atacgggaat gcctgatggc caatactgta acataataga agctaacttt 1260
 gatgaaagca ccggccaatg tagtgcagct acagattcca acggtaacg cgttattacc 1320
 gtcagtggtg ggcaagctaa ctttaatgta gcaggcgatc atgctgctgc aattcatggt 1380
 ggcgcaaaaa ttggtgatca atgtagtggg gatgattgcc catgtacagg atccgattgt 1440
 aataatgatc ctaaacctga ttttgagta ccagcaacat caattgtac atcagaaaat 1500
 ttacctacgc tatattactg gggagcacag cctacagata gcttagcgaa tgcagcttgg 1560
 ccagggtgctg caatgcaaac aaatggcgac tttaaagtgc atgatttagg tgcgaacta 1620
 accaaaatta acgcatctt tagtgacaat ggtgcaata aaacagctga tctaactggt 1680
 actggtgcag gttgttataa agacgggact tggagcacct taaaaattg tggctttgaa 1740
 attaccggtg cacaaaccaa tccagtcggt ggcgacgaag tctggtactt ccgaggtact 1800
 gctaattgact ggggtaaagc acaattagat tatgacgcaa ctagcggttt gtattacaca 1860
 atacaaagct ttaatgggta agaagcacct gcgcgtttta aaattgataa tggtagttgg 1920
 actgaagctt atccaacagc tgattaccaa gttacagata acaattcata ccgattaac 1980
 tttaatagcg atagcaaagc gattacagta aacgcacaat aa 2022

<210> 68

<211> 673

<212> PRT

<213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

5 <220>

<221> SEÑAL

<222> (1)...(23)

<220>

<221> DOMINIO

10 <222> (27)...(359)

<223> Alfa amilasa, dominio catalítico

<220>

<221> DOMINIO

<222> (372)...(463)

15 <223> Alfa amilasa, dominio beta de todos los terminales C

<220>

<221> SITIO

<222> (88)...(91)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

20 <220>

<221> SITIO

<222> (365)...(368)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

25 <221> SITIO

<222> (402)...(405)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

5 <222> (406)...(409)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (561)...(564)

10 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (646)...(649)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

15 <220>

<221> SITIO

<222> (664)...(667)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 68

ES 2 575 912 T3

Met Lys Lys Asn Thr Ile Ser Ala Leu Val Ala Gly Met Val Leu Gly
 1 5 10 15

Phe Ala Ser Asn Ala Met Ala Val Pro Arg Thr Ala Phe Val His Leu
 20 25 30

Phe Glu Trp Lys Trp Glu Asp Val Ala Gln Glu Cys Glu Thr Phe Leu
 35 40 45

Gly Pro Lys Gly Phe Ala Ala Val Gln Val Ser Pro Pro Thr Lys Ser
 50 55 60

His Asn Thr Asp Ala Trp Trp Gly Arg Tyr Gln Pro Val Ser Tyr Ala
 65 70 75 80

Phe Glu Gly Arg Ser Gly Asn Arg Ser Gln Phe Lys Asn Met Val Gln
 85 90 95

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

ES 2 575 912 T3

His Met Ala Ala Tyr Asp Arg Asn Phe Pro Asp Val Pro Tyr Ser Ser
 115 120 125
 Asn Asp Phe Asn Ser Cys Thr Gly Asp Ile Asp Tyr Asn Asn Arg Trp
 130 135 140
 Gln Thr Gln His Cys Asp Leu Val Gly Leu Asn Asp Leu Lys Thr Gly
 145 150 155 160
 Ser Asp Tyr Val Arg Gln Lys Ile Ala Asp Tyr Met Asn Asp Ala Ile
 165 170 175
 Ser Met Gly Val Ala Gly Phe Arg Ile Asp Ala Ala Lys His Ile Pro
 180 185 190
 Ala Gly Asp Ile Ala Ala Ile Lys Gly Lys Leu Asn Gly Asn Pro Tyr
 195 200 205
 Ile Phe Gln Glu Val Ile Gly Ala Ser Gly Glu Pro Val Arg Pro Thr
 210 215 220
 Glu Tyr Thr Phe Ile Gly Gly Val Thr Glu Phe Gln Phe Ala Arg Lys
 225 230 235 240
 Leu Gly Pro Ala Phe Arg Asn Ser Asn Ile Ala Trp Leu Lys Asp Ile
 245 250 255
 Gly Ser Gln Met Glu Leu Ser Ser Ala Asp Ala Val Thr Phe Val Thr
 260 265 270
 Asn His Asp Glu Glu Arg His Asn Pro Asn Gly Pro Ile Trp His Gly
 275 280 285
 Val Gln Gly Asn Gly Tyr Ala Leu Ala Asn Ile Phe Thr Leu Ala Tyr
 290 295 300
 Pro Tyr Gly Tyr Pro Lys Ile Met Ser Gly Tyr Phe Phe His Gly Asp
 305 310 315 320
 Phe Asn Ala Ala Pro Pro Ser Ser Gly Ile His Thr Gly Asn Ala Cys
 325 330 335
 Gly Phe Asp Gly Gly Asp Trp Val Cys Glu His Lys Trp Arg Gly Ile
 340 345 350
 Ala Asn Met Val Ala Phe Arg Asn Tyr Thr Ala Ser Glu Trp Arg Ile
 355 360 365
 Ser Asn Trp Trp Gln Asn Ser Asn Asp Gln Ile Ala Phe Gly Arg Gly
 370 375 380

ES 2 575 912 T3

Gly Leu Gly Phe Val Val Ile Asn Lys Arg Ala Asn Gly Ser Ile Asn
 385 390 395 400
 Gln Ser Phe Asp Thr Gly Met Pro Asp Gly Gln Tyr Cys Asn Ile Ile
 405 410 415
 Glu Ala Asn Phe Asp Glu Ser Thr Gly Gln Cys Ser Ala Ala Thr Asp
 420 425 430
 Ser Asn Gly Gln Ala Val Ile Thr Val Ser Gly Gly Gln Ala Asn Phe
 435 440 445
 Asn Val Ala Gly Asp His Ala Ala Ala Ile His Val Gly Ala Lys Ile
 450 455 460
 Gly Asp Gln Cys Ser Gly Asp Asp Cys Pro Cys Thr Gly Ser Asp Cys
 465 470 475 480
 Asn Asn Asp Pro Lys Pro Asp Phe Ala Val Pro Ala Thr Ser Ile Cys
 485 490 495
 Thr Ser Glu Asn Leu Pro Thr Leu Tyr Tyr Trp Gly Ala Gln Pro Thr
 500 505 510
 Asp Ser Leu Ala Asn Ala Ala Trp Pro Gly Val Ala Met Gln Thr Asn
 515 520 525
 Gly Asp Phe Lys Cys His Asp Leu Gly Val Glu Leu Thr Lys Ile Asn
 530 535 540
 Ala Ile Phe Ser Asp Asn Gly Ala Asn Lys Thr Ala Asp Leu Thr Val
 545 550 555 560
 Thr Gly Ala Gly Cys Tyr Lys Asp Gly Thr Trp Ser Thr Leu Gln Asn
 565 570 575
 Cys Gly Phe Glu Ile Thr Gly Ala Gln Thr Asn Pro Val Gly Gly Asp
 580 585 590
 Glu Val Trp Tyr Phe Arg Gly Thr Ala Asn Asp Trp Gly Lys Ala Gln
 595 600 605
 Leu Asp Tyr Asp Ala Thr Ser Gly Leu Tyr Tyr Thr Ile Gln Ser Phe
 610 615 620
 Asn Gly Glu Glu Ala Pro Ala Arg Phe Lys Ile Asp Asn Gly Ser Trp
 625 630 635 640
 Thr Glu Ala Tyr Pro Thr Ala Asp Tyr Gln Val Thr Asp Asn Asn Ser
 645 650 655

ES 2 575 912 T3

Tyr Arg Ile Asn Phe Asn Ser Asp Ser Lys Ala Ile Thr Val Asn Ala
660 665 670

Gln

<210> 69

<211> 1455

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 69

ES 2 575 912 T3

atgaaccgtc caggcacggg cgcctcgggg cgcccacaat ctcgctccgc cacgtcgtgg 60
 caatcccgca acggcggctg gctgctcgcc tcgctgctgg ccgtgtgttt cgcaacggcg 120
 cccgtgcgcg ccgatgtcat cctgcatgca ttcaactggc cgtatgcgac ggtcgaagcc 180
 cgcgccaacg agctgcgaga cctgggctat cgcgccgtac tggtcgcacc gccggtgaag 240
 tccgaaggca atgctgtggg ggcgcgctac cagccgcagg actaccgggt catcgagcat 300
 ccgctcggca atcgcgagtc gttcgtgctc atgtccagtg ctctgctgct ccgcggcatt 360
 cgcgtctacg ccgacatcgt gctgaatcac atggccaatg aagccccgca gcgaccggat 420
 ctgaactacc ccggtcaacg ggtcctggat cagtatgcag gcaataaccgc gtacttcgctg 480
 cagcaacgac tctacggcga tctccgctac aacttcatgt cggcctggga ctttggcccc 540
 gcccactgca tcggcaacta ccacgacgtc tggcaggtgc agaactggcg gctgtgcagc 600
 ggcgctggcg atgccggcct gcccgacctg ctggccagcg attacatcgt cggacagcag 660
 cgcacctatc tccaggcact gaagaacctc ggcgtctccg ggttgctcat cgacgcggcc 720
 aagcacatgc cgctcagcca catcaaccgc gtactgactg ccgacctcaa ggccggcatg 780
 catgtgttcg gcgaagtcac caccacgggt ggcgtcgggt atcccgaata cgatctgttt 840
 ctccggccct atctggatgg caccgatcac ggtgcctacg acttcccgtt cttcgaggcg 900
 atccgccgcg cattcggctt cgggtggcagc atgagcacc tggtcgatcc tggggccgct 960
 ggactggccc tgccaatgc ccggtcgatc accttcaccg tcaccacga catcccgaac 1020
 aatggtgtgt tccggcatct gctgctggat gccggcgacg aaaccctggc gtacgcctac 1080
 atcctgggtc gtgacggcgg cagtccactg ctgtactccg atcacaacga gagcggcgac 1140
 aaccgctggg ttcatgccta tcggcgcaac gacctcggcg ccatgatccg ctttcacaat 1200
 gccaatcatg gcaatgacat gcaggtgctc gcacacggga actgccatct gctctttcgc 1260
 cgcggcaatc gcggcatcgt ggcgatcaac aagtgcgggc atacgggtcaa cgccacggtg 1320
 aacatgaaca acagtgtgct gtgggtggcat acccctgata gcgatgtcct tgatgccggc 1380
 agcgtggtgc agattggcag cgccagccac acgttcagcc ttccgccgcg gcgctcgg 1440
 atgtggctgc gctga 1455

<210> 70

<211> 484

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<220>

<221> SEÑAL

<222> (1)...(44)

<220>

<221> DOMINIO

<222> (45)...(399)

5 <223> Alfa amilasa, dominio catalítico

<220>

<221> SITIO

<222> (381)...(384)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

10 <220>

<221> SITIO

<222> (443)...(446)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

15 <221> SITIO

<222> (449)...(452)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 70

ES 2 575 912 T3

Met Asn Arg Pro Gly Thr Gly Ala Ser Gly Arg Pro Gln Ser Arg Ser
 1 5 10 15

Ala Thr Ser Trp Gln Ser Arg Asn Gly Gly Trp Leu Leu Ala Ser Leu
 20 25 30

Leu Ala Val Cys Phe Ala Thr Ala Pro Val Arg Ala Asp Val Ile Leu
 35 40 45

His Ala Phe Asn Trp Pro Tyr Ala Thr Val Glu Ala Arg Ala Asn Glu
 50 55 60

Leu Arg Asp Leu Gly Tyr Arg Ala Val Leu Val Ala Pro Pro Val Lys
 65 70 75 80

Ser Glu Gly Asn Ala Trp Trp Ala Arg Tyr Gln Pro Gln Asp Tyr Arg
 85 90 95

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

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

Asn His Met Ala Asn Glu Ala Pro Gln Arg Pro Asp Leu Asn Tyr Pro
 130 135 140

Gly Gln Arg Val Leu Asp Gln Tyr Ala Gly Asn Thr Ala Tyr Phe Ala
 145 150 155 160

ES 2 575 912 T3

Gln Gln Arg Leu Tyr Gly Asp Leu Arg Tyr Asn Phe Met Ser Ala Trp
 165 170 175

Asp Phe Gly Pro Ala His Cys Ile Gly Asn Tyr His Asp Val Trp Gln
 180 185 190

Val Gln Asn Trp Arg Leu Cys Ser Gly Ala Gly Asp Ala Gly Leu Pro
 195 200 205

Asp Leu Leu Ala Ser Asp Tyr Ile Val Gly Gln Gln Arg Thr Tyr Leu
 210 215 220

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

Lys His Met Pro Leu Ser His Ile Asn Arg Val Leu Thr Ala Asp Leu
 245 250 255

Lys Ala Gly Met His Val Phe Gly Glu Val Ile Thr His Gly Gly Val
 260 265 270

Gly Asp Pro Glu Tyr Asp Leu Phe Leu Arg Pro Tyr Leu Asp Gly Thr
 275 280 285

Asp His Gly Ala Tyr Asp Phe Pro Leu Phe Glu Ala Ile Arg Arg Ala
 290 295 300

Phe Gly Phe Gly Gly Ser Met Ser Thr Leu Val Asp Pro Gly Ala Val
 305 310 315 320

Gly Leu Ala Leu Pro Asn Ala Arg Ser Ile Thr Phe Thr Val Thr His
 325 330 335

Asp Ile Pro Asn Asn Gly Val Phe Arg His Leu Leu Leu Asp Ala Gly
 340 345 350

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

Pro Leu Leu Tyr Ser Asp His Asn Glu Ser Gly Asp Asn Arg Trp Val
 370 375 380

His Ala Tyr Arg Arg Asn Asp Leu Ala Ala Met Ile Arg Phe His Asn
 385 390 395 400

Ala Asn His Gly Asn Asp Met Gln Val Leu Ala His Gly Asn Cys His
 405 410 415

Leu Leu Phe Arg Arg Gly Asn Arg Gly Ile Val Ala Ile Asn Lys Cys
 420 425 430

ES 2 575 912 T3

Gly His Thr Val Asn Ala Thr Val Asn Met Asn Asn Ser Val Leu Trp
435 440 445

Trp His Thr Pro Tyr Arg Asp Val Leu Asp Ala Gly Ser Val Val Gln
450 455 460

Ile Gly Ser Ala Ser His Thr Phe Ser Leu Pro Pro Arg Arg Ala Arg
465 470 475 480

Met Trp Leu Arg

<210> 71

<211> 1719

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de muestra ambiental

<400> 71

ES 2 575 912 T3

gtgcccgaac tgaaatggg gacagacagct atatTTTTatc aaatctaccc gcgctctttc 60
 gccgacggca acggcgatgg catcggcgat ttcaaaggca tcatcggcaa actcgattat 120
 ctacaaaatc ttggcataga tgcgctctgg ctctcgccctc acttcccctc ccccaactgg 180
 gattgctggct acgatatcag cgattaccgc aacgTtgcgc cggaatacgg cacgctggac 240
 gatttcaaaa ctttcttgag cgaatcgcac aaacgcggta tccgcgtcat tctcgacctc 300
 gtgctgaatc acacctccga tgaacatccg tggttcatcg aatcgaaatc cagcccgcat 360
 aatcccaaat ccgattggta tgtgtgggtc gatacgccgc ccaacaattg gcagtcctgc 420
 ttcgatggcg atgcctggac atacgtccct gaacgcggcc aatattatta tcaactctc 480
 atgaaacagc agcccgatct caactggcat aatccgcagg tcaaacaggc catgtggggag 540
 gcggtgcgct tctggctcga tctcggcggtg gacggcttcc gcctggacgc catcggcacg 600
 atctacgaag aCCcaaatct cacgccgcat aatgtcccga tgaatttggc tgagctgcgt 660
 cacttcacag atgtcgcaa aacgccggaa gagatcaagc tcaaagaaaa atactggcac 720
 gacatgttca agcatcaatg gggtcagccc ggcgttcatg acctgatgaa agaactgcgc 780
 gccatcctcg atgaatatga tggcgaccgc atgctggtcg gcgaagatga caacatcgat 840
 tacatgggca acggagacga cgaattgcag ctggtcttca acttcccgtt gatgcgcgcc 900
 gatcgtctca cccccacca tattcggcgc aacaaaaag agcgtttgac tcgtctgaat 960
 gctttaccgg ttaaaggctg ggcttgcaac acgctcggca accatgatag ttcacgcgtc 1020
 tacaccaaT tcggtgaccg gatccacggc gcggaccatg cacgtctcaa cctggcgctt 1080
 ttgctcacc tgacggcac gccgttctta tacaacggcg aagagatcgg catgaccgac 1140
 cacatcatta ccgatcccac caaactgcgc gacacatgg caacctggta ttacaacagc 1200
 cttgtcaacg aatgaaggT cgagccagcg gaggccgcc ttcgcgccgg acagatgacg 1260
 cgcgacaaaa accgtacccc catgcaatgg gacaataagc ccaatgccgg tttttgccca 1320

gataaagccg aaccctggtt gccagtcaac cccaattacc gcgcaggcat taacgtccgc 1380
 gagcaaacat cgaacccgaa ctgctgctc aattactata aacgtctcat ccacttgccg 1440
 cgggaaacgc ctgccttgat cgctggagat tacgttccgc ttcaccagac atccaaagat 1500
 catctggcct tctgcgcaa aacagattca caaacgatcc tggtcgtttt gaattactcc 1560
 cccaataaat tggaattgga tttctcgcgc accgtcgaaa tgaaaggccg cccgctgatc 1620
 gcaattttct ccagcgcaga tgaccgccg caggcggcac aaagcccaaa gaaagtatcg 1680
 gtcggcgctt acggagtctt gctggcagaa gtaaaatag 1719

<210> 72

<211> 572

5 <212> PRT

<213> Desconocido

<220>

<223> obtenido a partir de una muestra ambiental

ES 2 575 912 T3

<220>

<221> DOMINIO

<222> (14)...(405)

<223> Alfa amilasa, dominio catalítico

5 <220>

<221> SITIO

<222> (104)...(107)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

10 <221> SITIO

<222> (525)...(528)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<400> 72

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

15

ES 2 575 912 T3

Trp Val Asp Thr Pro Pro Asn Asn Trp Gln Ser Cys Phe Asp Gly Asp
 130 135 140
 Ala Trp Thr Tyr Val Pro Glu Arg Gly Gln Tyr Tyr Tyr His Tyr Phe
 145 150 160
 Met Lys Gln Gln Pro Asp Leu Asn Trp His Asn Pro Gln Val Lys Gln
 165 170 175
 Ala Met Trp Glu Ala Val Arg Phe Trp Leu Asp Leu Gly Val Asp Gly
 180 185 190
 Phe Arg Leu Asp Ala Ile Gly Thr Ile Tyr Glu Asp Pro Asn Leu Thr
 195 200 205
 Pro His Asn Val Pro Met Asn Leu Ala Glu Leu Arg His Phe Thr Asp
 210 215 220
 Val Ala Lys Thr Pro Glu Glu Ile Lys Leu Lys Glu Lys Tyr Trp His
 225 230 235 240
 Asp Met Phe Lys His Gln Trp Gly Gln Pro Gly Val His Asp Leu Met
 245 250 255
 Lys Glu Leu Arg Ala Ile Leu Asp Glu Tyr Asp Gly Asp Arg Met Leu
 260 265 270
 Val Gly Glu Asp Asp Asn Ile Asp Tyr Met Gly Asn Gly Asp Asp Glu
 275 280 285
 Leu Gln Leu Val Phe Asn Phe Pro Leu Met Arg Ala Asp Arg Leu Thr
 290 295 300
 Pro Asp His Ile Arg Arg Asn Gln Lys Glu Arg Leu Thr Arg Leu Asn
 305 310 315 320
 Ala Leu Pro Val Lys Gly Trp Ala Cys Asn Thr Leu Gly Asn His Asp
 325 330 335
 Ser Ser Arg Val Tyr Thr Lys Phe Gly Asp Arg Ile His Gly Ala Asp
 340 345 350
 His Ala Arg Leu Asn Leu Ala Leu Leu Leu Thr Leu His Gly Thr Pro
 355 360 365
 Phe Leu Tyr Asn Gly Glu Glu Ile Gly Met Thr Asp His Ile Ile Thr
 370 375 380
 Asp Pro Thr Lys Leu Arg Asp Thr Met Ala Thr Trp Tyr Tyr Asn Ser
 385 390 395 400

ES 2 575 912 T3

Leu Val Asn Glu Met Lys Val Glu Pro Ala Glu Ala Ala Leu Arg Ala
 405 410 415

Gly Gln Met Thr Arg Asp Lys Asn Arg Thr Pro Met Gln Trp Asp Asn
 420 425 430

Lys Pro Asn Ala Gly Phe Cys Pro Asp Lys Ala Glu Pro Trp Leu Pro
 435 440 445

Val Asn Pro Asn Tyr Arg Ala Gly Ile Asn Val Arg Glu Gln Thr Ser
 450 455 460

Asn Pro Asn Ser Leu Leu Asn Tyr Tyr Lys Arg Leu Ile His Leu Arg
 465 470 475 480

Arg Glu Thr Pro Ala Leu Ile Ala Gly Asp Tyr Val Pro Leu His Gln
 485 490 495

Thr Ser Lys Asp His Leu Ala Phe Leu Arg Lys Thr Asp Ser Gln Thr
 500 505 510

Ile Leu Val Val Leu Asn Tyr Ser Pro Asn Lys Leu Glu Leu Asp Phe
 515 520 525

Ser Arg Thr Val Glu Met Lys Gly Arg Pro Leu Ile Ala Ile Phe Ser
 530 535 540

Ser Ala Asp Asp Arg Pro Gln Ala Ala Gln Ser Pro Lys Lys Val Ser
 545 550 555 560

Val Gly Ala Tyr Gly Val Leu Leu Ala Glu Val Lys
 565 570

<210> 73

<211> 1836

<212> ADN

5 <213> Thermomyces lanuginosus ATCC 200065

<400> 73

ES 2 575 912 T3

```

ttgtgcgcg cccttggact cgccgccttg atcgtccaag gcggaagaag cagacctgaa    60
acaaccgtcc cacatgcaac gggctcgctc gacgacttcc tcgccgcaca gagtccgatt    120
gctttccaag gcacatcctgaa caatatcggg cctagcggag cgtactcggg aggtgtcaat    180
ccgggtgtgg tcattgagag tccaagtaaa caagatcccg actactttta cacctgggtg    240
cgcgacgctg ctctcaactgt ccaatatctg gtggaggagc tggttgcagg aaatgccagt    300
cttcagttcc tcattcagga ctacatcagc tcccaggcac gactgcagac ggtggaaaat    360
ccatccggct cctctcctgc ggggtggtcta ggagagccca agtttcatgt cgacgagacc    420
gcctttacgg actcctgggg ccgaccacag cgggacggcc cgcctctccg cgccattgcc    480
atgatttcgt ttgccaatta cctgattgac aacggtcata aatcgactgt ggaggacatc    540

atctggccga ttgttcgcaa tgacttgtcc tatgtctcgc agcattggaa cgaaacaact    600
tttgacatct gggaggaagt ccatagctca tcgtttttca ccacggctgt ccagtaccgt    660
gctctggtcc aaggcagtg cttggctagc aagctcggcc atacctgcga caactgcggg    720
tcccaagcac cgcagatcct ttgcttctg cagtcgtatt ggaccgggtc gcacatctta    780
gccaacaccg gtggcggccg ctccggaaaag gacgtcagca cgatcctcgg cgtcattggc    840
tcgtttgatc cgaacccga ctgtgatgac gttaccttcc agccctgctc ggccccggct    900
cttgcaaatc acaagcaggt cgttgacagc ttccgcagta tctatgcat caacgctggc    960
atcccgtcag ggtcggctgt tgcggttggg cgttatcccg aggatgtcta tcagggtgga    1020
caccctgtgt acctaacaac ggtcgcggcg gcggagcagc tttacgacgc catttaccag    1080
tggaaccatg tagggcacat cgacatcaat gctgtcaatc tggacttctt caagagcatt    1140
tatccgtcag ccgccgaggg cacatacaca tcagactctt caacatttca agacattata    1200
tctgctgtac ggacctatgc ggacgggttt ctacgcgtaa ttgagaaata cactccgccg    1260
gataacttgc ttgccgagca gttccaccgg gagacgggca ttccactatc ggcagcttct    1320
ctgacatggt cctacgccgc gctcaacacg gccgcgcagc ggcgagcgtc aatcgtgcc    1380
tcaccgtgga actctaacag cacagatctc ccggacaaat gctcggcaac ctccggcaaca    1440
gggccgtatg ccacgcccac aaacacggca tggccaacca ctacgcagcc accggagcgg    1500
ccggcatgca caccgccgtc ggaagtaaca ctacacttca acgcgctcgt cgacaccgcg    1560
tttggccaga atatttatct cgtgggtctc attccggagc tcggatcgtg ggatccggcc    1620
aacgccctct tgatgagcgc aaagagctgg actagcggaa atccggctctg gacgctatcc    1680
atttcccttc cagcaggaac ctcttttgag tacaagttca ttcgaaagga tgatggttcc    1740
tcggatgttg tctgggaaaag tgaccggaat cgttcgtaca acgtgccgaa ggattgcggt    1800
gccaacacgg ccaccgtgaa ttcttggtgg cgatga                                1836

```

<210> 74

<211> 611

<212> PRT

5 <213> Thermomyces lanuginosus ATCC 200065

<220>

<221> DOMINIO

<222> (39)...(453)

<223> Familia de glicosil hidrolasas 15

5 <220>

<221> DOMINIO

<222> (508)...(605)

<223> Dominio de enlazamiento de almidón

<220>

10 <221> SITIO

<222> (99)...(102)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

15 <222> (200)...(203)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (473)...(476)

20 <223> Sitio de N-glicosilación. Id del prositio = PS00001

<220>

<221> SITIO

<222> (599)...(602)

<223> Sitio de N-glicosilación. Id del prositio = PS00001

25 <400> 74

ES 2 575 912 T3

Met Cys Ala Ala Leu Gly Leu Ala Ala Leu Ile Val Gln Gly Gly Glu
 1 5 10 15
 Ala Arg Pro Glu Thr Thr Val Pro His Ala Thr Gly Ser Leu Asp Asp
 20 25 30
 Phe Leu Ala Ala Gln Ser Pro Ile Ala Phe Gln Gly Ile Leu Asn Asn
 35 40 45
 Ile Gly Pro Ser Gly Ala Tyr Ser Glu Gly Val Asn Pro Gly Val Val
 50 55 60
 Ile Ala Ser Pro Ser Lys Gln Asp Pro Asp Tyr Phe Tyr Thr Trp Val
 65 70 75 80
 Arg Asp Ala Ala Leu Thr Val Gln Tyr Leu Val Glu Glu Leu Val Ala
 85 90 95
 Gly Asn Ala Ser Leu Gln Phe Leu Ile Gln Asp Tyr Ile Ser Ser Gln
 100 105 110
 Ala Arg Leu Gln Thr Val Glu Asn Pro Ser Gly Ser Leu Ser Ser Gly
 115 120 125
 Gly Leu Gly Glu Pro Lys Phe His Val Asp Glu Thr Ala Phe Thr Asp
 130 135 140
 Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Pro Leu Arg Ala Ile Ala
 145 150 155 160
 Met Ile Ser Phe Ala Asn Tyr Leu Ile Asp Asn Gly His Gln Ser Thr
 165 170 175
 Val Glu Asp Ile Ile Trp Pro Ile Val Arg Asn Asp Leu Ser Tyr Val
 180 185 190
 Ser Gln His Trp Asn Glu Thr Thr Phe Asp Ile Trp Glu Glu Val His
 195 200 205
 Ser Ser Ser Phe Phe Thr Thr Ala Val Gln Tyr Arg Ala Leu Val Gln
 210 215 220
 Gly Ser Ala Leu Ala Ser Lys Leu Gly His Thr Cys Asp Asn Cys Gly
 225 230 235 240

ES 2 575 912 T3

Ser Gln Ala Pro Gln Ile Leu Cys Phe Leu Gln Ser Tyr Trp Thr Gly
 245 250 255

Ser His Ile Leu Ala Asn Thr Gly Gly Gly Arg Ser Gly Lys Asp Val
 260 265 270

Ser Thr Ile Leu Gly Val Ile Gly Ser Phe Asp Pro Asn Ala Asp Cys
 275 280 285

Asp Asp Val Thr Phe Gln Pro Cys Ser Ala Arg Ala Leu Ala Asn His
 290 295 300

Lys Gln Val Val Asp Ser Phe Arg Ser Ile Tyr Ala Ile Asn Ala Gly
 305 310 315 320

Ile Pro Ser Gly Ser Ala Val Ala Val Gly Arg Tyr Pro Glu Asp Val
 325 330 335

Tyr Gln Gly Gly His Pro Trp Tyr Leu Thr Thr Ala Ala Ala Ala Glu
 340 345 350

Gln Leu Tyr Asp Ala Ile Tyr Gln Trp Asn His Val Gly His Ile Asp
 355 360 365

Ile Asn Ala Val Asn Leu Asp Phe Phe Lys Ser Ile Tyr Pro Ser Ala
 370 375 380

Ala Glu Gly Thr Tyr Thr Ser Asp Ser Ser Thr Phe Gln Asp Ile Ile
 385 390 395 400

Ser Ala Val Arg Thr Tyr Ala Asp Gly Phe Leu Ser Val Ile Glu Lys
 405 410 415

Tyr Thr Pro Pro Asp Asn Leu Leu Ala Glu Gln Phe His Arg Glu Thr
 420 425 430

Gly Ile Pro Leu Ser Ala Ala Ser Leu Thr Trp Ser Tyr Ala Ala Leu
 435 440 445

Asn Thr Ala Ala Gln Arg Arg Ala Ser Ile Val Pro Ser Pro Trp Asn
 450 455 460

Ser Asn Ser Thr Asp Leu Pro Asp Lys Cys Ser Ala Thr Ser Ala Thr
 465 470 475 480

Gly Pro Tyr Ala Thr Pro Thr Asn Thr Ala Trp Pro Thr Thr Thr Gln
 485 490 495

Pro Pro Glu Arg Pro Ala Cys Thr Pro Pro Ser Glu Val Thr Leu Thr
 500 505 510

ES 2 575 912 T3

Phe Asn Ala Leu Val Asp Thr Ala Phe Gly Gln Asn Ile Tyr Leu Val
515 520 525

Gly Ser Ile Pro Glu Leu Gly Ser Trp Asp Pro Ala Asn Ala Leu Leu
530 535 540

Met Ser Ala Lys Ser Trp Thr Ser Gly Asn Pro Val Trp Thr Leu Ser
545 550 555 560

Ile Ser Leu Pro Ala Gly Thr Ser Phe Glu Tyr Lys Phe Ile Arg Lys
565 570 575

Asp Asp Gly Ser Ser Asp Val Val Trp Glu Ser Asp Pro Asn Arg Ser
580 585 590

Tyr Asn Val Pro Lys Asp Cys Gly Ala Asn Thr Ala Thr Val Asn Ser
595 600 605

Trp Trp Arg
610

<210> 75

<211> 1344

<212> ADN

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<400> 75

ES 2 575 912 T3

atgaaccggg ttaagagact ttcgattttg gtcgttcttt tccttaccgc cttcattccg	60
actgtttttg ccggcgagca gcctctcgcc atctttcacg cctttaacga ccccttact	120
cttgttgaat cctatgtctg cgaactcgcc gggcagggat actcacatgt ccagatatct	180
ccagcgcaga agtcgaaccc tgctcgagcc tggatgccc ggtatcaacc cgtagatfff	240
actgtcatcg aagggatggg cactgagagc gatctgagga agctcacgga taaggccac	300
gcgtgtggaa taaaggtgat cgccgatgtg gtcttcaacc acatgtcgag catggacgag	360
tacaaggggc ttgacaagtt tccgggactt gtcctgctg atttccaccg gcagtgcggc	420
atcgattatt caaaacgaga ttcggtgctg aactgttggc tcggaggcga cttgcccgat	480
ctggaccagt cccggccgag ggtacaggat gttcagagag cccacataag gaagctcctt	540
tccctcggca tagacggctt ccgcttcgat gcggctaac acatcgacc cattgttgtg	600
aaagactaca tcgatctcat cgacagggag agcaacggca ggacctggaa ctacctcgag	660
gtcatcgagg atgacggcac tcaggccacg gactacaact ggatagcggc agtgaccgat	720
ttcgtcctct acaaggagtc gttgaggaag gccttcagtc tcggcgggga cctgcatcg	780
ctcaagatgc ctgtggctgt caatgattcg cggagtatcg tcttcgggag aaatcacgac	840
accgtgccgg agaataacca gaactgcatc gtcggctgct acgacagccg ggaggactcc	900
tatcttgcca cggcatacgt cctggccccg gaatcgggag tcccgctggt cctcaactgg	960
gacaactacg acgccccta catcagcacc ggcgtgaagt tccgccagat catgacgcag	1020
cgaggacgat cggccatgaa cgtgaaggag aatgtgctgg gcgtcatcga cagtctgtc	1080
gtcatgatga tggagcgcgg gagtgaaggc tttttcgtcc tcaacaagag cgccgaccgg	1140
ttcgatatcc cagttctgga tctgacactg accaatctcg agggatgtta tcgggagctg	1200
agaagaaaat tcaccgtcgc catcgagaga aagtacggta agaaatttgt caccgggtgg	1260
ggacgatggg accggggggg cctcgaaatc tacggccgcg acgctctcta cttcatcgg	1320
gaaccctggg agcagtgacg gtaa	1344

<210> 76

<211> 447

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido a partir de una muestra ambiental

<220>

<221> SEÑAL

10 <222> (1)...(24)

ES 2 575 912 T3

<400> 76

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

ES 2 575 912 T3

Arg Lys Leu Leu Ser Leu Gly Ile Asp Gly Phe Arg Phe Asp Ala Ala
 180 185 190

Lys His Ile Asp Pro Ile Val Val Lys Asp Tyr Ile Asp Leu Ile Asp
 195 200 205

Arg Glu Ser Asn Gly Arg Thr Trp Asn Tyr Leu Glu Val Ile Glu Asp
 210 215 220

Asp Gly Thr Gln Ala Thr Asp Tyr Asn Trp Ile Ala Ala Val Thr Asp
 225 230 235 240

Phe Val Leu Tyr Lys Glu Ser Leu Arg Lys Ala Phe Ser Leu Gly Gly
 245 250 255

Asp Leu Arg Ser Leu Lys Met Pro Val Ala Val Asn Asp Ser Arg Ser
 260 265 270

Ile Val Phe Gly Arg Asn His Asp Thr Val Pro Glu Asn Asn Gln Asn
 275 280 285

Cys Ile Val Gly Cys Tyr Asp Ser Arg Glu Asp Ser Tyr Leu Ala Thr
 290 295 300

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

Asp Asn Tyr Asp Ala Pro Tyr Ile Ser Thr Gly Val Lys Phe Arg Gln
 325 330 335

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

Leu Gly Val Ile Asp Ser Pro Val Val Met Met Met Glu Arg Gly Ser
 355 360 365

Glu Gly Phe Phe Val Leu Asn Lys Ser Ala Asp Arg Phe Asp Ile Pro
 370 375 380

Val Leu Asp Leu Thr Leu Thr Asn Leu Glu Gly Cys Tyr Arg Glu Leu
 385 390 395 400

Arg Arg Lys Phe Thr Val Ala Ile Glu Arg Lys Tyr Gly Lys Lys Phe
 405 410 415

Val Thr Arg Trp Gly Arg Trp Asp Arg Gly Gly Leu Glu Ile Tyr Gly
 420 425 430 435

Arg Asp Ala Leu Tyr Phe Ile Arg Glu Pro Trp Glu Gln Cys Arg
 435 440 445

ES 2 575 912 T3

<210> 77

<211> 1311

<212> ADN

<213> Secuencia Artificial

5 <220>

<223> Polinucleótido generado sintéticamente

<400> 77

atggccaagt	acctggagct	cgaagagggc	ggggtcataa	tgcagggcgtt	ctactgggac	60
gtgccttcag	gaggaatatg	gtgggacaca	atcgggcaga	agataccgga	gtggtacgat	120
gccggaatct	ccgcaatatg	gattcctccc	gcgagcaagg	gtatgagcgg	cggctattcg	180
atgggctacg	acccctacga	ttattttgac	ctcgggtgagt	actaccagaa	gggaacggtg	240
gaaacgaggt	tcggctcaaa	gcaggagctc	ataaacatga	taaacacggc	ccatgcctac	300
ggcataaagg	tcatagcgga	catcgtcata	aaccaccgcg	caggcggaga	cctcagtggtg	360
aacccttcg	ttggggacta	cacctggacg	gacttctcaa	aggtggcctc	gggcaaatat	420
actgccaaact	acctcgactt	ccaccggaac	gagctccatg	cgggcgattc	cggaacattt	480
ggaggctatc	ccgacatatg	ccacgacaag	agctgggacc	agtactggct	ctgggcccagc	540
caggagagct	acgcggcata	tctcaggagc	atcggcatcg	atgcctggcg	ctttgactac	600
gtgaagggct	acggagcgtg	ggtcgtcaag	gactggctca	actgggtggg	cggctgggcc	660
gttggcgagt	actgggacac	caacgttgat	gcactcctca	actgggccta	ctcgagcggc	720
gccaaggctc	tcgacttccc	gctctactac	aagatggacg	cggcctttga	caacaagaac	780
attcccgcac	tcgtcgaggc	cctcaagaac	gggggcacag	tcgtcagccg	cgaccctgtt	840
aaggccgtaa	ccttcgttgc	aaaccacgac	accgatataa	tctggaacaa	gtatccagcc	900
tacgcgttca	tcctcaccta	cgaggggccag	ccgacaatat	tctaccgcga	ctacgaggag	960
tggctcaaca	aggataagct	caagaacctc	atctggatac	atgacaacct	cgccggagga	1020
agcacgagca	tagtttacta	cgacagcgac	gagatgatct	tcgtgaggaa	cggctatgga	1080
agcaagcctg	gccttataac	ttacatcaac	ctcggctcga	gcaaggtttg	aaggtggggt	1140
tatgtgccga	agttcgcggg	cgcgtgcatc	cacgagtata	ctggtaacct	cggaggctgg	1200
gtagacaagt	acgtctactc	aagcggctgg	gtctatctcg	aagctccagc	ttacgaccct	1260
gccaacgggc	agtatggcta	ctccgtgtgg	agctactgcg	gtgttgggtg	a	1311

<210> 78

10 <211> 436

<212> PRT

<213> Secuencia Artificial

<220>

<223> Polipéptido generado sintéticamente

15 <400> 78

ES 2 575 912 T3

Met Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala
 1 5 10 15
 Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg
 20 25 30
 Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile
 35 40 45
 Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp
 50 55 60
 Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val
 65 70 75 80
 Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr
 85 90 95
 Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His
 100 105 110
 Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr
 115 120 125
 Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr
 130 135 140
 Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe
 145 150 155 160
 Gly Gly Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp
 165 170 175
 Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly
 180 185 190
 Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val
 195 200 205
 Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr
 210 215 220
 Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly
 225 230 235 240
 Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe
 245 250 255
 Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly
 260 265 270
 Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn
 275 280 285
 His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
 290 295 300
 Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu
 305 310 315 320
 Trp Leu Asn Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Asp Asn
 325 330 335
 Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met
 340 345 350
 Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr
 355 360 365
 Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys
 370 375 380
 Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp
 385 390 395 400
 Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro
 405 410 415
 Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr
 420 425 430
 Cys Gly Val Gly 435

<210> 79

<211> 1764

<212> ADN

5 <213> Secuencia artificial

<220>

ES 2 575 912 T3

<223> Polinucleótido generado sintéticamente

<400> 79

```

cttactccag ctgagtggag gtctcagagc atctacttcc tcctcactga tcgcttcggc      60
aggaccgata actctactac cgctgcttgc gataccactg atagggttta ctgcggcggc      120
tcttggcagg gcatcatcaa ccacctggat tacatccagg gcatgggctt cactgctatc      180
tggatcactc cagtgactgg ccagttctac gagaacactg gcgatggcac ttcataccac      240
ggctactggc agcaggatat ctacgacctg aactacaact acggcactgc tcaggatctt      300
aagaacctgg cttctgctct tcacgagagg ggcatgtacc tgatgggtga tgtgggtggc      360
aaccacatgg gctacgatgg cgctggcaac actgtggatt actccgtgtt caaccattc      420
agcagcagct cttacttcca cccgtactgc ctgatcagca actacgacaa ccagaccaac      480
gttgaggatt gctggcttgg cgatacaacc gtgtctctgc cagacctgga taccacttct      540
actgctgtga gggatatctg gtatgattgg gtggccgatc ttgtggctaa ctactctatc      600
gatggcctga ggggtggacac tgtaagcac gtggagaagg atttctggcc agactacaac      660
tctgctgctg gcgtttactg tgttggggag gttttctctg gcgatccagc ttacacttgc      720
ccgtaccaga actacatgga tggcgtgctg aactaccaa tctactacca gctcctgtac      780
gctttcgagt caagctctgg ctctatcagc gacctgtaca acatgatcag cagcgtggct      840
tcttcttgca aggacccaac ctttcttggc aactttatcg agaaccacga caaccacgc      900

ttcgtagct acaccagcga ttactctcag gccaaagaacg tgatcacctt catcttctc      960
tcagacggca tcccaatcgt ttacgctggc caggagcagc attactctgg cggctctgat      1020
ccagetaaac gcgaggctac ttggctttct ggctactcta cctctgctac cctgtacact      1080
tggatcgcta gactaacca gatccgctca ctggctatct ctaaggatgc tggctacgtc      1140
caggetaaga acaaccatt ctactccgac tctaacacca tcgctatgag gaagggcact      1200
actgctggcg ctcaggtgat cactgtgctg tctaacaagg gcgcttctgg ctcttcttac      1260
acccttctc tgtctggcac tggctactct gctggcgcta ctcttggtga gacctacacc      1320
tgcaccactg tgaccgttga ttcttctggc aacctgccag tgccaatgac ttctggcctt      1380
ccaaggggtg tcgttccaag ctcttgggtt aacggctctg ctctgtgcaa cactgagtgc      1440
actgctgcta cttctctgcc agtgctgttc gaggagctgg tgactactac ttacggcgag      1500
aacatctacc ttagcggctc tatctctcag cttggctctt ggaacactgc ttctgctgtg      1560
gctctttctg ctagccagta caccagctct aacccaaagt ggtacgtgtc tgtgactctt      1620
ccagtgggca ctagcttcca gtacaagttt atcaagaagg gctctgacgg ctctgttgtt      1680
tgggagagcg acccaaacag gtcttacact gtccagctg gctgcgaggg cgctactgtt      1740
actgttgccg atacctggcg ctga      1764

```

<210> 80

5 <211> 1710

<212> ADN

ES 2 575 912 T3

<213> Secuencia artificial

<220>

<223> Polinucleótido generado sintéticamente

<400> 80

```

gccgatacaa acgcttgaa gagccgctct atctacttcg tgctgaccga taggatcgct    60
aggaactctt ctgatactgg cggctctgct tgctctgac ttggcaacta ctgcggcggc    120
actttccagg gccttgagag caagctggat tacatcaagg gccttggctt cgatgctatc    180
tggatcactc cagtgggtgc taacaaggct gctggctacc acggctactg ggctgaggat    240
ctgtacgctg tgaactctaa ctacggcact gctgctgac tgaagtctct ggtggctgct    300
gctcacgcta agggcatcta catgatggg gagctgggtg ctaaccacat gggcccaggc    360
gctatcacta acaacaggcc agagccactt aaccaggctt caagctacca tccaccgtgc    420
aacatcgact acaacaacca gacttctgtg gaggtttgcc agatcgctgg cttccagat    480
atctacacca ccaagagcga gatcaggact ctgctgaaca cttgggtgaa ctggcttgtg    540
aacgagtact ctttcgatgg cgtcaggatc gataccgta agcacgtgga gaaggattc    600
tggccaggct tctctgctgc tactggcgtt tacaacatcg gcgagggtgt cgatggcgat    660
ccagcttacc ttgctccata cgccaagctg atgccaggcc ttcttaacta cgccgtgtac    720
taccgatga acaacttcta ccagcagact ggctcttctc aggctctggt ggacatgatg    780
aacaccgtgt ccaacacttt cccagatccc tctgctcttg gcaccttctt ggacaaccac    840
gataacaagc ggtggctgaa cgtgaagaac gatcagacc tgctgaagaa cgctctggct    900

tacgttatcc ttgctagggg catcccaatc ctttactacg gcacagagca gggctacgct    960
ggcggcgatg atccagctaa ccgcgaggat ctttggcgtc ctggcttcaa cactaacgcc   1020
aacctgtacc aggctatcaa gaagctgact gctgctaggc aggctgctgg cggccttgct   1080
ggcaacgatc acgtccacct gtacgtggct gatactgctt acgcttggtc tagggctaac   1140
ggcaacctga tcgtgcttac tactaacgct ggcggcaact ctaacactca gcactgcttc   1200
aacaccaga aggctaacgg ccgctggact aacgtttacg gcaacggcgc tactgtttct   1260
gctgatagca acggccaaat ctgcgtgtct gttactaacg gcgagccagt ggttcttctt   1320
gctggctctg ctactccaac tactggcacc accctgtcta ctaggactgc taccgctact   1380
gctacaccaa ctgcttggcc aaccgctgtg tctgtgtctt tcaccacag ggtgacaact   1440
gttccaggcg acacgatcaa gatcactggc aacctgccc agcttggcaa ctggactcca   1500
gctaacggcc ttgctcttct tgctgctagc tacaccagct ctaaccacat ctggaccatc   1560
actgtgccac ttgctgctgg ctctagcatc agctacaagt tcgtgaagat cgattctggc   1620
ggcactgtta cttgggagag cgacccaac aggtcttaca ctgctccatc ttgccaggct   1680
tcagctagcg ttaacagcag ctggcagtaa                                1710

```

5

<210> 81

ES 2 575 912 T3

<211> 1851

<212> ADN

<213> Secuencia artificial

<220>

5 <223> polinucleótido generado sintéticamente

<400> 81

```
gaggctgctc cacagcttgc tccaagggct actacctctc tggatgcttg gcttgcttct    60
gagaccactg ttgctctgga tgggattctt gataatgttg gctcttctgg cgcttacgct    120
aagtctgcta agtcaggcat cgtgatcgct tctccatcta cggacaacc ggactactac    180
tacacttggg ccagggatgc tgctcttact gtgaaggccc tgatcgatct ttcaggaac    240
ggcgagacct ctcttcagac tgtgatcatg gagtacatca gctctcaggc ttacctccag    300
actgtgtcta acccatctgg ctcactttct actggcggcc ttgccgagcc aaagtactac    360
gtggatgaga ctgcttacac tggctcttgg ggcaggccac agagggatgg cccagctctt    420
agggctactg ccatgatcga tttcggcaac tggcttatcg ataacggcta ctctacctac    480
gctagcaaca ttgtgtggcc aattgtgagg aacgatctgt cttacgtggc tcagtactgg    540
aaccagactg gctacgatct ttgggaggag gttaacggct ctagcttctt cactatcgcc    600
gttcagcaca gggctcttgt tgagggctct accttcgctt ctaaggttgg ggcttcttgc    660
tcttggtgcy attctcaggc tccacagggt ctgtgcttcc ttcagagggt ctggactggc    720
tcttacatca tggctaactt cggcggcggc cgctctggca aggacgctaa cactgtgctg    780
ggctctatcc acactttcga tccaaacgct ggctgcgacg atacaacttt ccagccatgc    840
tctccaaggg ctctggctaa ccacaagggt tacaccgata gcttccgctc tatctactct    900
```

ES 2 575 912 T3

. atcaacagcg gcatctctca gggcaaggct gttgctggtg gccgctaccc agaggattct 960
 tactacaacg gcaacccgtg gttccttact actcttgctg ctgccgagca gctttacgat 1020
 gctatctacc agtggcagaa gatcggctct atcaccatca ctgacgtgtc tctggccttc 1080
 ttcaaggacc tgtactcttc tgctgctggt ggcacctacg cttctagctc ttctgccttc 1140
 acctctatcg tgaacgccgt taagacttac gctgacggct acatgtctat cgttcagacc 1200
 cacgctatga ctaacggctc tctgtctgag cagttcggca agtctgatgg cttctctctg 1260
 tctgctaggg atctgacttg gtcttacgct gctctgctga ctgctaacct taggcgcaac 1320
 tctgttggtc caccatcttg gggcgagact actgctactt ctgtgccatc tgtgtgctct 1380
 gctacttcag ctaccggcac ctactctacc gctactaaca ctgcttgcc atctactctg 1440
 acatctggca ctggcgctac cactactacc tctaaggcta cctctaccac cactaccagc 1500
 tctgcttcta ctactactgc tggctgcgtt gttccaactg ctgtggctgt gaccttcgac 1560
 gagatcgcta ctactactta cggcgagaac gtgtacgttg tgggctctat ctctcagctt 1620
 ggcagctggg atacttctaa ggctgtggcc ctgtctgctt ctaagtacac cagcagcaac 1680
 aacctttggt acgctaccgt tactcttcca gctggcacca ctttcagta caagtttacc 1740
 aggggtcat cttcaggcac tgtgacctgg gagtctgac caaacaggtc ttacaccgtg 1800
 ccatctgctt gcggcacttc tactgctgtg gtgaacacca cttggcgcta a 1851

<210> 82

<211> 1860

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Polinucleótido generado sintéticamente

<400> 82

ES 2 575 912 T3

```

gccccacagc tttctccaag ggctaccacc tctcttgatg cttggctggc ttctgagacc 60
actgttagcc tgaacgggat tcttgataat atcggcgctt caggcgctta cgctcagtct 120
gctaaggctg gcgttggtat cgctagccca agcacctcta gccagacta ctaçtacacc 180
tggaccaggg attctgctct gacccttaag gtgctgatcg accttttcag gaacggcaac 240
gttgatcttc agaccgtgat cgaggagtac atcactgctc aggcttacct tcagactgtg 300
agcaacccat ctggcgatct ttcttcaggc gctggccttg cçgagccaaa gttcaacgtg 360
gacatgtctg cttacactgg cgcttggggc aggccacaga gggatggccc agctcttagg 420
gctatcgccc ttatcgattt cggcaactgg cttatcgata acggctactc tagctacgct 480
gtgtçtaatg tgtggccaat tgtgaggaac gatctgtctt acgtggctca gtactggctc 540
cagtctggct acgatctttg ggaggagggtg aacagcatgt ctttcttcac catcgçtaac 600
cagcacaggg ctcttggtga gggctctact ttcgctggca gagttggcgc ttcttgctct 660
tggtgçgatt ctçaggctcc acagatcctt tgctacatgc agaatttctg gaccggctct 720
tacatcaacg ctaacactgg cggçggccgc tctggcaagg atgçtaacac cgtgctggct 780
agcatctçta ctttcgaccc agaggctact tgçgatgatg tgacçttcca gçcatgctct 840

tctagggctc tggçtaacca caagggtgac accgattctt tccgctctgt gtactctctt 900
gactctggga tcgctgaggg cgttgctgtt gctgttgçcc gctaccçaga ggattcttac 960
tacaacggca acccggtggtt cçttactact cttgctgctg cçgagçagct ttacgagçct 1020
atçtaccagt ggaacaagat cggctctatc accatçacçt ctacçtçact ggçtttcttc 1080
aacgacgtgt actcttctgc tgctgttgçc acttacgctt caggctçtac çgçttacact 1140
gctatçgtgt ccgctgttaa gacttacgct gacggctacg tgtçtatçgt tcaggçtçac 1200
gctatgacta acggctctct gtctgagçag ttcgataagg cttctggçac tcagçtttct 1260
gctagggçatc tgacçtggçc ttacgctgct ctgctgactg ctaacatgag gçgcaacggc 1320
atçgttccac catçttgggg cgctgçttca gçtaactçta tcccaagçtc ttgctçtact 1380
ggctctgçta ctggçactta ctçtaccçca actggçactt cttggçcçatc tactctgact 1440
tcaggçactg ctggçaccac tactacçtct gctaccacca ccacçtçtac tagçgtgtct 1500
aagaccacca cçactaccac ttçtactacc agctgçacca ctccacttc agtggçtgtg 1560
acçttçgatg agatçgçtac cacttactac ggçgagaacg tgtacatçag çggctçtact 1620
tctçagçttg gçagçtggga tacttcttct gccatçgçcc tttctgçttc tcagçtacc 1680
agçagçaaça acçtgtgggt çgtgacçatc aacçttccag ctggçactac ttçcagçtac 1740
aagçtatatca ggaaggagtc tgacggçtct atçgtttggg agagçgaccc aaacaggtct 1800
tacactgtgc catctggçtg çggçgtttct actgçtaccg agagçgatac ttggçgçtaa 1860

```

Referencias

1. Frandsen TP, Christensen T. Stoffer B, Lehmbeck J, Dupont C, Honzatko R B, Svensson B (1995) *Biochemistry*. 34:10162-9.
- 5 2. Sierks M R y Svensson B. (1996) *Biochemistry*; 35: 1865 - 71.
3. Fagerstrom R y Kalkkinen N. (1995) *Biotechnol Appl Biochem*. 21: 223 - 31.

REIVINDICACIONES

1. Un ácido nucleico aislado, sintético o recombinante que consiste de:

(a) una secuencia de ácido nucleico que tiene al menos 98%, 99% o 100% de identidad de secuencia con la SEQ ID NO: 51, en donde el ácido nucleico codifica al menos un polipéptido que tiene una actividad de alfa-amilasa;

5 (b) una secuencia de ácido nucleico que codifica un polipéptido que tiene la secuencia de SEQ ID NO: 52;

(c) una secuencia de ácido nucleico completamente complementaria a (a), o (b); o,

(d) una secuencia de ácido nucleico como se expone en SEQ ID NO: 51.

10 2. Una secuencia de ácido nucleico aislado, sintético, o recombinante como se expone en la SEQ ID NO: 51, en donde el ácido nucleico codifica un polipéptido que tiene una actividad de alfa-amilasa y una secuencia de aminoácidos como se expone en la SEQ ID NO: 52.

3. Un vector, un casete de expresión o un vehículo de clonación que comprende un ácido nucleico que comprende la secuencia de ácido nucleico de la reivindicación 1;

15 en donde el ácido nucleico está opcionalmente enlazado operativamente a un promotor; en donde opcionalmente el promotor es un promotor de planta o un promotor viral; en donde opcionalmente el promotor es un promotor inducible o un promotor específico de tejido; en donde opcionalmente el promotor específico de tejido es opcionalmente específico de la semilla, un promotor específico del endospermo, un promotor específico de hoja, específico de la raíz, específico del tallo o inducido por abscisión.

20 4. Una célula transformada que comprende un ácido nucleico que comprende la secuencia de ácido nucleico de la reivindicación 1, o un vector, casete de expresión, o vehículo de clonación como se establece en la reivindicación 3 en donde la célula es opcionalmente una célula bacteriana, una célula de mamífero, una célula fúngica, una célula de levadura, una célula de insecto o una célula de planta, en donde la célula de planta es opcionalmente una célula de patata, arroz, maíz, trigo, tabaco o cebada.

5. Un polipéptido aislado, sintético o recombinante que tiene actividad de alfa-amilasa que comprende:

25 (i) una secuencia de aminoácidos que tiene al menos 99% o identidad de secuencia completa con la secuencia de SEQ ID NO: 52 sobre la longitud completa del polipéptido;

(ii) una secuencia de aminoácidos codificada por el ácido nucleico de la reivindicación 1;

(iii) el polipéptido de cualquiera de (i) a (ii) en donde el polipéptido tiene una secuencia como se establece en la SEQ ID NO: 52 o fragmentos enzimáticamente activos del mismo que tienen actividad de alfa-amilasa.

30 6. El polipéptido aislado, sintético o recombinante de la reivindicación 5 en donde la actividad de alfa-amilasa comprende hidrolizar enlaces glucosídicos, y, opcionalmente, los enlaces glucosídicos comprenden un enlace α -1,4-glucosídico o un enlace α -1,6-glucosídico, y/o en donde opcionalmente la actividad de alfa-amilasa comprende hidrolizar enlaces glucosídicos en un polisacárido, oligosacárido o almidón, o y/o la actividad de alfa-amilasa comprende escindir una unidad de maltosa o de D-glucosa de extremo no reductor de un polisacárido, oligosacárido o almidón.

35 7. Un método para producir un polipéptido recombinante que comprende las etapas de: (a) proveer un ácido nucleico enlazado operativamente a un promotor, en donde el ácido nucleico comprende la secuencia de ácido nucleico de la reivindicación 1; y (b) expresar el ácido nucleico de la etapa (a) bajo condiciones que permiten la expresión del polipéptido, produciendo de ese modo un polipéptido recombinante, que comprende además transformar una célula huésped con el ácido nucleico de la etapa (a) seguido por la expresión del ácido nucleico de la etapa (a),
40 produciendo de ese modo un polipéptido recombinante en una célula transformada.

8. Un método para hidrolizar un polisacárido, oligosacárido o almidón que comprende las siguientes etapas:

(a) proveer el polipéptido de la reivindicación 5;

(b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; en donde la composición comprende opcionalmente un enlace α -1,4-glucosídico o un enlace α -1,6-glucosídico y

(c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde los polipéptidos hidroliza el polisacárido, oligosacárido o almidón.

9. Un método para la licuefacción o la eliminación de un polisacárido, oligosacárido o almidón que comprende las siguientes etapas:

5 (a) proveer el polipéptido de la reivindicación 5;

(b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; y

(c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde los polipéptidos licuan o elimina el polisacárido, oligosacárido o almidón.

10. Un método para lavar un objeto que comprende las siguientes etapas:

10 (a) proveer una composición que comprende el polipéptido de la reivindicación 5;

(b) proveer un objeto; y

(c) poner en contacto la composición de la etapa (a) y el objeto de la etapa (b) bajo condiciones en donde la composición puede lavar el objeto.

11. Un método para hidrolizar un polisacárido, oligosacárido o almidón en un pienso o un alimento antes del consumo por un animal que comprende las siguientes etapas:

15 (a) ontemer un material alimenticio que comprende un almidón, donde el polisacárido, oligosacárido o almidón en el polisacárido, oligosacárido o almidón puede ser hidrolizado por el polipéptido de la reivindicación 5; y

(b) adicionar el polipéptido de la reivindicación 5, al material del pienso o alimento en cantidades suficientes y durante un periodo de tiempo suficiente para causar la hidrólisis del polisacárido, oligosacárido o almidón y la formación de un alimento o pienso tratado, hidrolizando de ese modo el polisacárido, oligosacárido o almidón en el alimento o el pienso antes del consumo por el animal, en donde el alimento o el pienso opcionalmente comprende arroz, maíz, cebada, trigo, legumbres, o patata.

20 (c) poner en contacto la composición de la etapa (a) y el objeto de la etapa (b) bajo condiciones en donde la composición puede lavar el objeto.

12. Un método para el desencolado textil que comprende las siguientes etapas:

(a) proveer el polipéptido de la reivindicación 5;

25 (b) proveer una tela; y

(c) poner en contacto el polipéptido de la etapa (a) y la tela de la etapa (b) bajo condiciones en donde la alfa-amilasa puede desencolar la tela.

13. Un método para el tratamiento de fibras lignocelulósicas que comprende las siguientes etapas:

(a) proveer el polipéptido de la reivindicación 5;

30 (b) proveer una fibra lignocelulósica; y

(c) poner en contacto el polipéptido de la etapa (a) y la fibra de la etapa (b) bajo condiciones en donde el polipéptido puede tratar la fibra mejorando así las propiedades de la fibra.

14. Un método para producir un jarabe con alto contenido de maltosa o con alto contenido de glucosa que comprende las siguientes etapas:

35 (a) proveer el polipéptido de la reivindicación 5;

(b) proveer una composición que comprende un polisacárido, oligosacárido o almidón; opcionalmente en donde el polisacárido, oligosacárido, o almidón es de arroz, maíz, cebada, trigo, legumbres, patata o batata y

(c) poner en contacto el polipéptido de la etapa (a) y la composición de la etapa (b) bajo condiciones en donde el polipéptido de la etapa (a) puede hidrolizar la composición de la etapa (b), produciendo de este modo un jarabe con alto contenido de maltosa o alto contenido de glucosa.

5 15. Un método para mejorar el flujo de los fluidos de producción que contienen al polisacárido, oligosacárido, o almidón, que comprende las siguientes etapas:

(a) proveer el polipéptido de la reivindicación 5;

(b) proveer la producción del fluido que comprende un polisacárido, oligosacárido, o almidón; opcionalmente en donde el fluido de producción es de una formación subterránea; y

10 (c) poner en contacto el polipéptido de la etapa (a) y el fluido de producción de la etapa (b) bajo condiciones en donde el polipéptido puede hidrolizar el polisacárido, oligosacárido o almidón en el fluido de producción, mejorando así su flujo por la disminución de su densidad.

16. Un método para cambiar la viscosidad de una composición, que comprende las siguientes etapas:

(i)

(a) proveer una composición y el polipéptido de la reivindicación 5; y

15 (b) tratar la composición con el polipéptido de la reivindicación 5; o (ii) el método de (i), en donde la composición comprende un suelo o un lodo de perforación.

17. Un método para el uso de una alfa-amilasa en la elaboración de cerveza o en la producción de alcohol que comprende las siguientes etapas:

(a) proveer el polipéptido de la reivindicación 5;

20 (b) proveer una composición utilizada para la elaboración de cerveza o en la producción de alcohol que comprende un polisacárido, oligosacárido, o almidón; en donde la composición que comprende un polisacárido, oligosacárido, o almidón es opcionalmente una cerveza;

25 (c) combinar el polipéptido de la etapa (a) con la composición de la etapa (b) bajo condiciones en donde el polipéptido puede hidrolizar el polisacárido, oligosacárido, o almidón en la composición utilizada para la elaboración de cerveza o la producción de alcohol.

18. Un método para bioblanquear una composición que comprende:

(i) proveer una composición en donde la composición es un papel o un producto de pulpa, y el polipéptido de la reivindicación 5;

(ii) tratar la composición con el polipéptido de la reivindicación 5.

30 19. Un método para el destintado de papel y fibras que comprende las siguientes etapas:

(a) proveer el polipéptido como se establece en la reivindicación 5;

(b) proveer una composición que comprende papel o fibra;

(c) poner en contacto el polipéptido de la etapa (a) y la composición de la etapa (b) bajo condiciones en donde el polipéptido puede destintar el papel o la fibra.

35 20. Un método para la fabricación de un combustible que comprende:

(i) (a) proveer el polipéptido de la reivindicación 5;

(b) proveer una composición que comprende un polisacárido, oligosacárido, o almidón; y

(c) poner en contacto el polipéptido de (a) con la composición de (b) bajo condiciones en donde el polipéptido hidroliza el polisacárido, oligosacárido, o almidón; o

(ii) el método de (i) en donde el polipéptido es una enzima termoestable;

(iii) el método de (i) en donde el combustible es a base de etanol.

21. Un método para procesar un material de biomasa que comprende lignocelulosa que comprende

5 (i) (a) proveer una composición que comprende el polipéptido de la reivindicación 5, y, proveer un material de biomasa; y

(b) poner en contacto la composición con el material de biomasa;

(ii) el método de (i), en donde el material de biomasa comprende o se deriva de un cultivo agrícola, o es un subproducto de un alimento o una producción de piensos, o es un producto de desecho lignocelulósico, o es un residuo de una planta o un residuo de papel o un producto de residuo de papel,

10 y opcionalmente el material de biomasa comprende patatas, soja, cebada, centeno, maíz, avena, trigo, remolacha, o caña de azúcar;

(iii) el método de (i) o (ii), en el que el polipéptido tiene la actividad que comprende una actividad de alfa-amilasa;

15 (iv) el método de cualquiera de (ii), en donde el residuo de la planta comprende tallos, hojas, cáscaras, vainas, mazorcas, madera, astillas de madera, pulpa de madera y aserrín y, o, el residuo de papel comprende papel de fotocopiadora usado o descartado, papel de la impresora de ordenador, papel de cuaderno, papel de notas, papel de máquina de escribir, periódicos, revistas, cartón y materiales de embalaje a base de papel; o

(v) el método de (i) a (iv), en donde el procesamiento del material de biomasa genera un bioetanol, biopropanol, biobutanol, o un biodiesel;

(vi) el método de (i) a (v), en donde el material de biomasa comprende una lignocelulosa.

20 22. Un método para la fabricación de bioetanol, biopropanol, biobutanol, o un biodiesel, que comprende

(i)

(a) proveer el polipéptido de la reivindicación 5, y, proveer una composición que comprende un polisacárido u oligosacárido; y

25 (b) poner en contacto la composición que comprende un polisacárido u oligosacárido con el polipéptido de la reivindicación 5;

(ii) el método de (i), en donde la composición que comprende un polisacárido u oligosacárido comprende una planta, producto de planta o derivado de la planta;

(iii) el método de (ii), en donde la planta o producto de la planta comprende las plantas de caña de azúcar o productos de plantas, remolacha o remolacha de azúcar, trigo, maíz, soja, patata, arroz o cebada;

30 (iv) el método de cualquiera de (i) a (iii), en donde el polipéptido tiene actividad que comprende una actividad de alfa-amilasa; o

(v) el método de cualquiera de (i) a (iv), en donde el polisacárido u oligosacárido comprende un azúcar fermentable.

23. Un método para fabricar un combustible que comprende

(i)

35 (a) proveer el polipéptido de la reivindicación 5, y, proveer una composición que comprende un azúcar fermentable; y

(b) poner en contacto la composición que comprende un azúcar fermentable con el polipéptido de la reivindicación 5;

(ii) el método de (i), en donde la composición que comprende un azúcar fermentable comprende una planta, producto de la planta o derivado de la planta;

(iii) el método de (ii), en donde la planta o producto de la planta comprende las plantas de caña de azúcar o productos de plantas, remolacha o remolacha de azúcar, trigo, maíz, soja, patata, arroz o cebada;

(iv) el método de cualquiera de (i) a (iii), en donde el polipéptido tiene actividad que comprende una actividad de alfa-amilasa; o

5 (v) el método de cualquiera de (i) a (iv), en donde el combustible comprende un bioetanol o una mezcla de gasolina-etanol, o comprende un bioetanol, biopropanol, biobutanol, o un biodiesel.

24. Un método para producir un azúcar, comprendiendo el método:

(i)

(a) proveer el polipéptido de la reivindicación 5;

10 (b) proveer una composición que comprende un polisacárido o un oligosacárido; y

(c) poner en contacto la composición de la etapa (b) con el polipéptido de la etapa (a), generando así azúcares;

(ii) el método de (i), en donde la composición que comprende un polisacárido comprende un almidón;

(iii) el método de cualquiera de (i) o (ii), en donde el azúcar es un azúcar fermentable;

25. Una composición que comprende el polipéptido de la reivindicación 5,

15 en donde la composición es seleccionada de: un detergente, un pienso o un alimento, un material textil o tela, un papel o un producto de papel o pulpa de papel, un líquido o jarabe con alto contenido de maltosa o con alto contenido de glucosa, una composición antiarrancamiento, una bebida alcohólica, una cerveza, una composición farmacéutica, un producto para el cuidado oral, una pasta de dientes, una crema dental, un gel o un polvo de dientes, un odontic, un enjuague bucal, una formulación de enjuague para antes o después del cepillado, una goma
20 de mascar, un comprimido para deshacer en la boca o un caramelo, un fluido para perforación de pozos de petróleo, una solución de bioblanqueamiento, un desinfectante, una biodefensa o un agente biodesintoxicante, un producto lácteo, o un material de biomasa.

26. Una composición de liberación controlada o liberación retardada que comprende el polipéptido de la reivindicación 5, que comprende

25 (a) un ingrediente deseado recubierto por un revestimiento de polímero de látex;

(b) la composición de liberación controlada o liberación retardada de (a), en donde el ingrediente deseado comprende una enzima;

30 (c) la composición de liberación controlada o liberación retardada de (a) o (b), en donde el ingrediente deseado comprende una molécula pequeña, un fármaco, un polisacárido, un lípido, un ácido nucleico, una vitamina, un antibiótico, o un insecticida;

(d) la composición de liberación controlada o liberación retardada de (a), (b), o (c), en donde el ingrediente deseado comprende una pella o una matriz;

(e) la composición de liberación controlada o liberación retardada de cualquiera de (a) a (d), en donde la pella o matriz comprende material comestible;

35 (f) la composición de liberación controlada o liberación retardada de cualquiera de (a) a (e), en donde el revestimiento de polímero de látex comprende una pintura de látex; o

40 (g) la composición de liberación controlada o liberación retardada de cualquiera de (a) a (f), en donde el revestimiento de polímero de látex comprende un (met)acrilato, un acetato de vinilo, un estireno, un etileno, un cloruro de vinilo, un butadieno, un cloruro de vinilideno, un versatato de vinilo, un propionato de vinilo, un acrilato de t-butilo, un acrilonitrilo, un neopreno, un maleato, un fumarato o una combinación de los mismos o un derivado de los mismos.

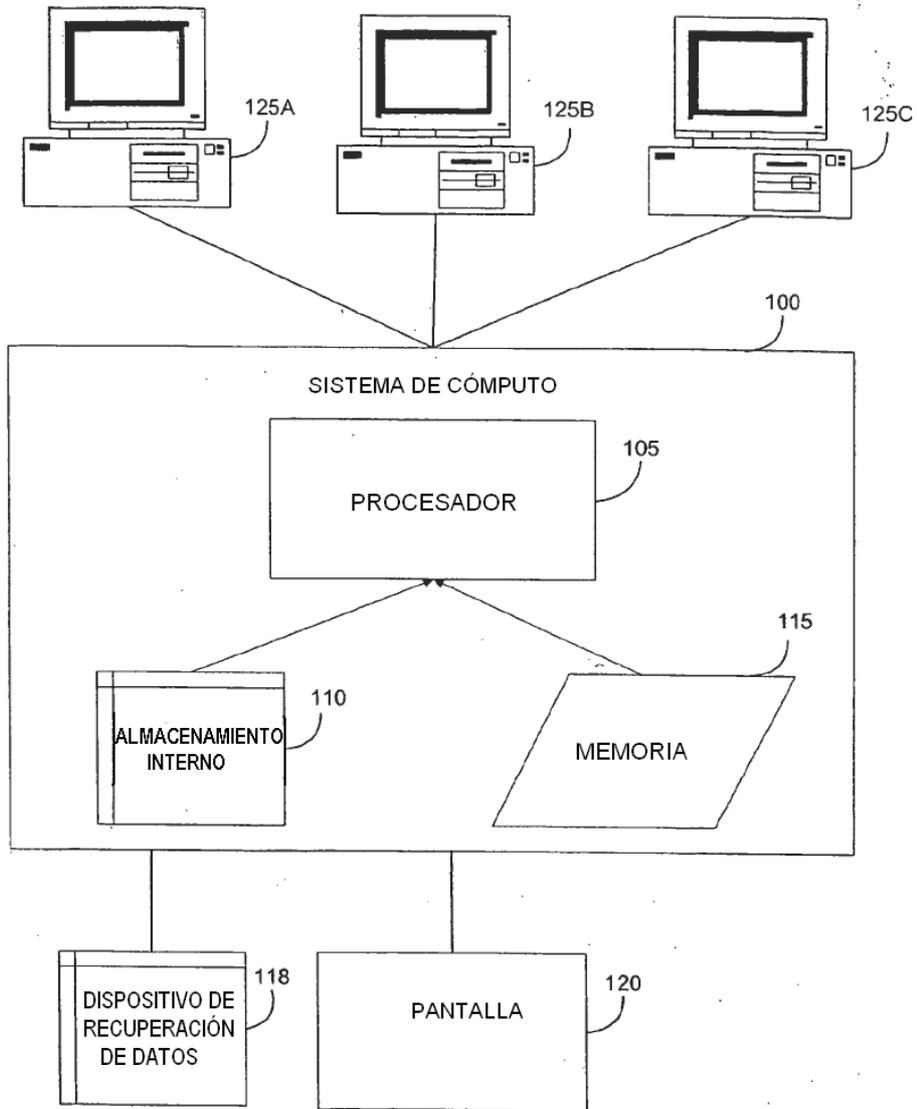


FIGURA 1

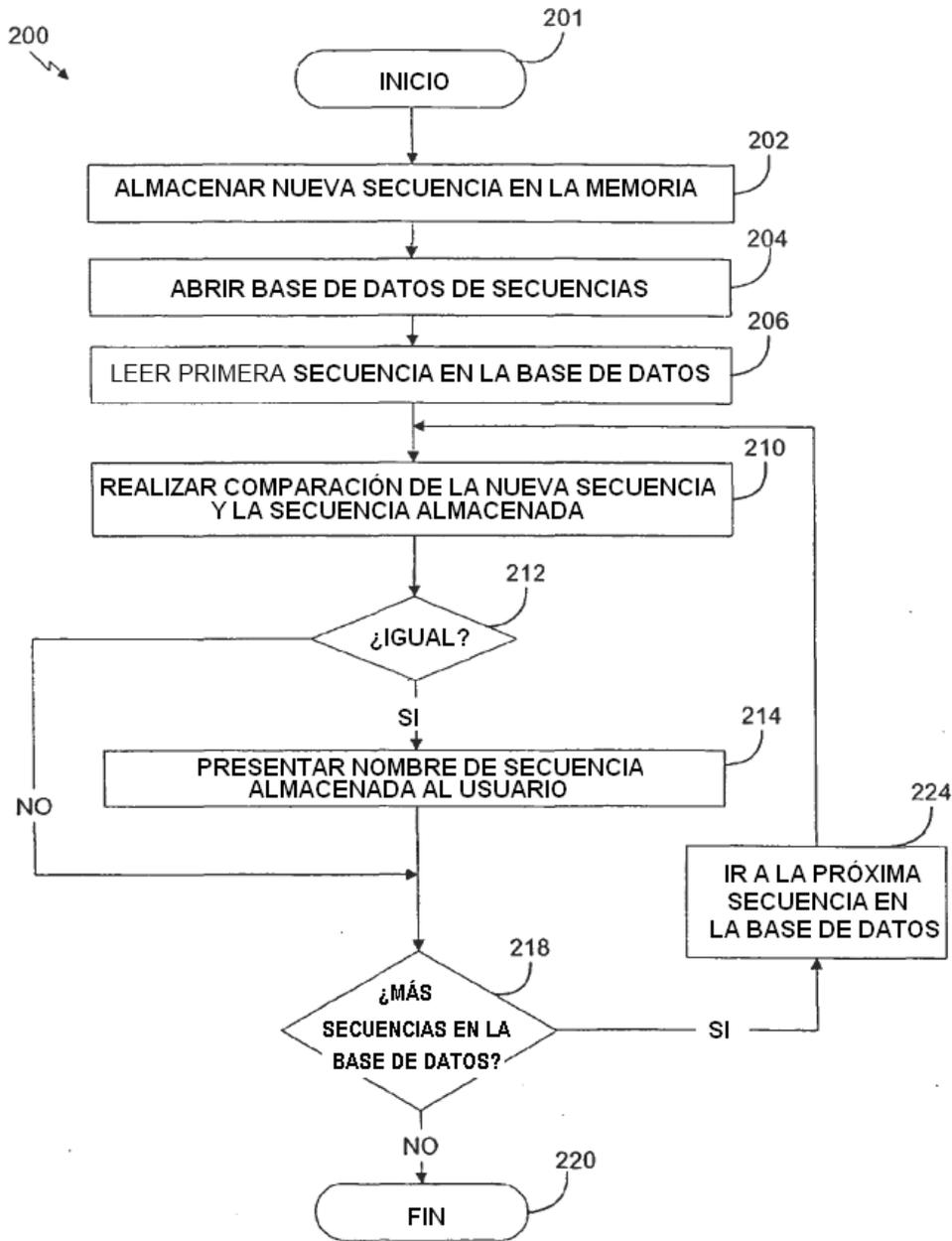


FIGURA 2

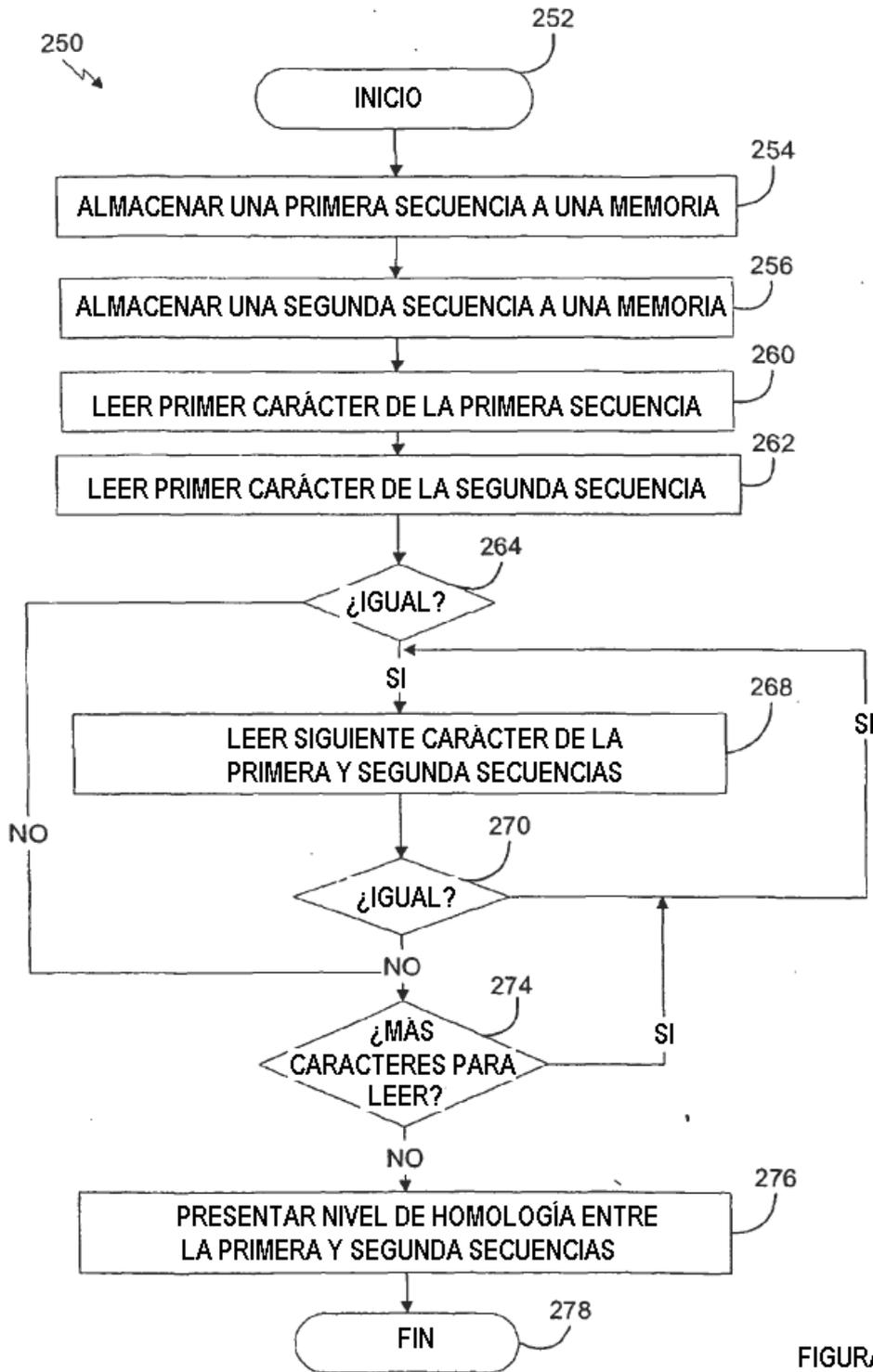


FIGURA 3

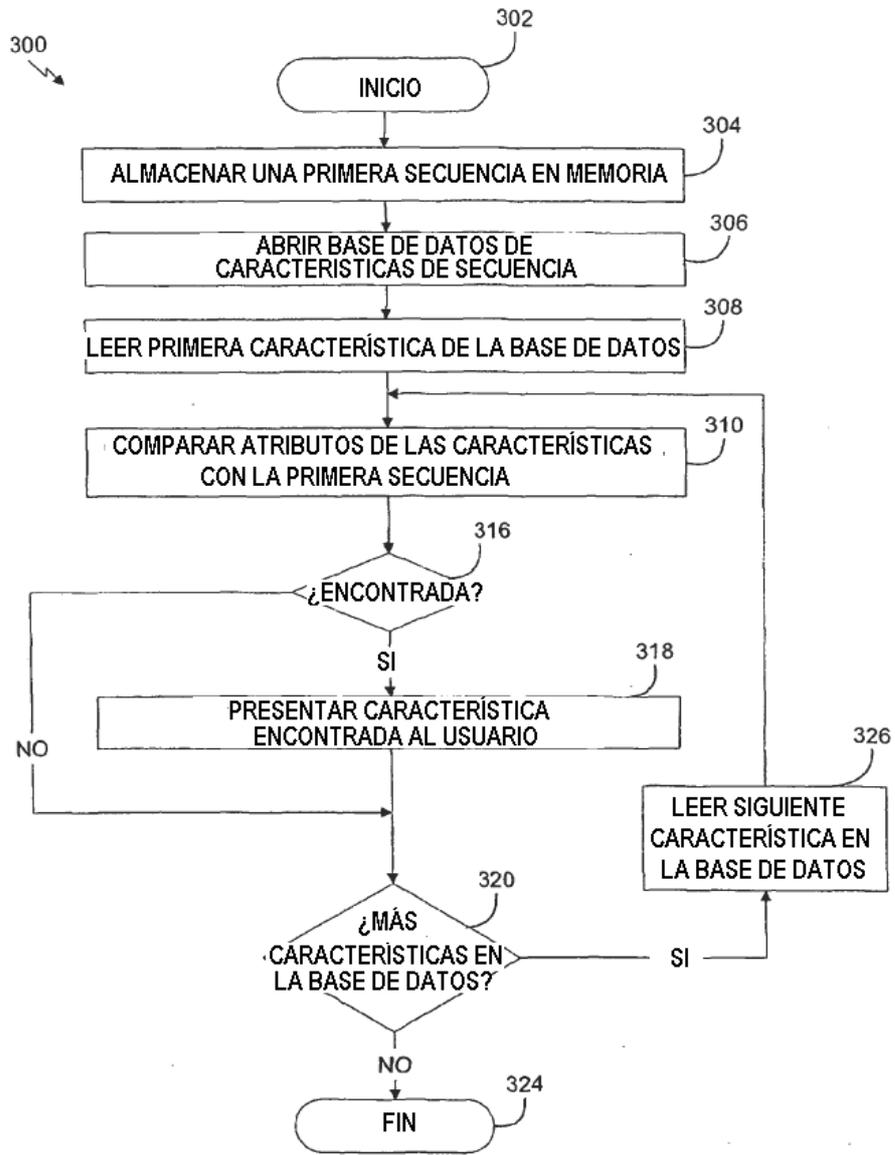


FIGURA 4

Figura 5

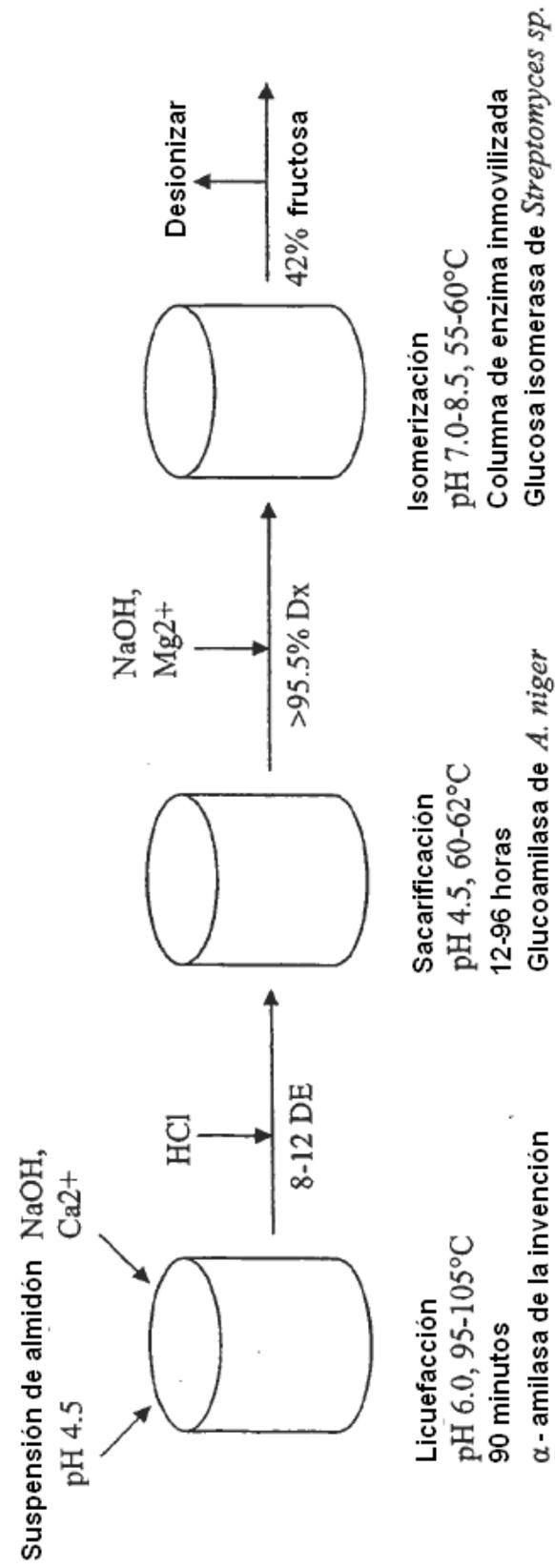


FIGURA 6

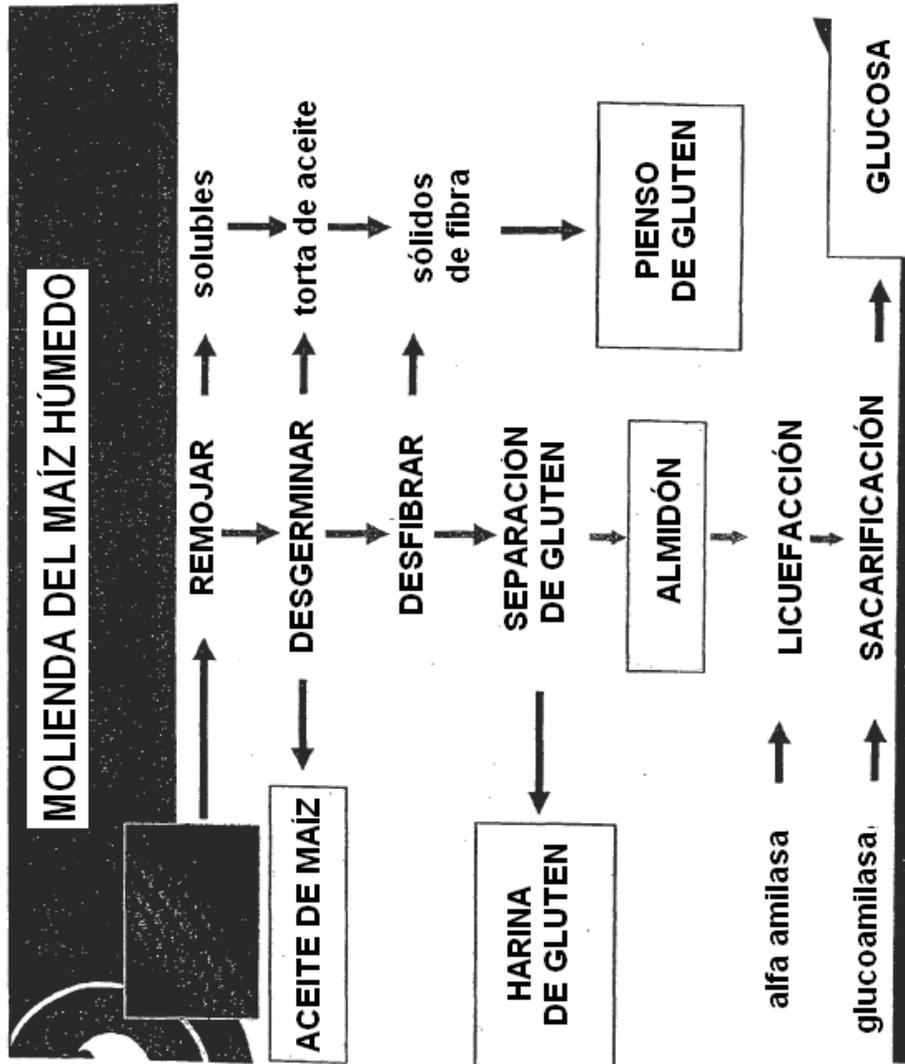


Figura 7

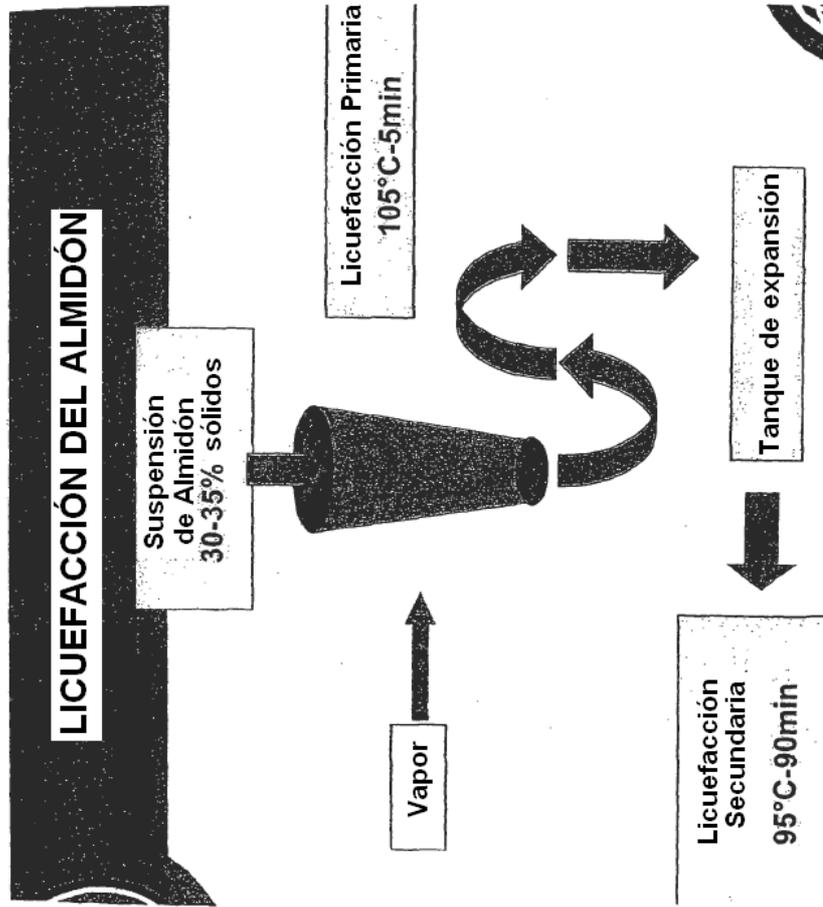


FIGURA 8

Licuefacción de almidón Proceso subóptimo

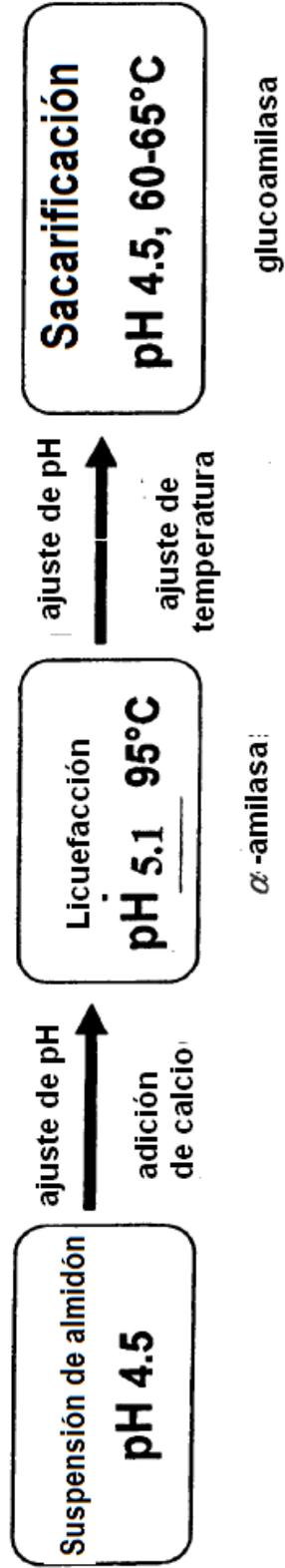


FIGURA 9

Oportunidades del Procesamiento de Almidón

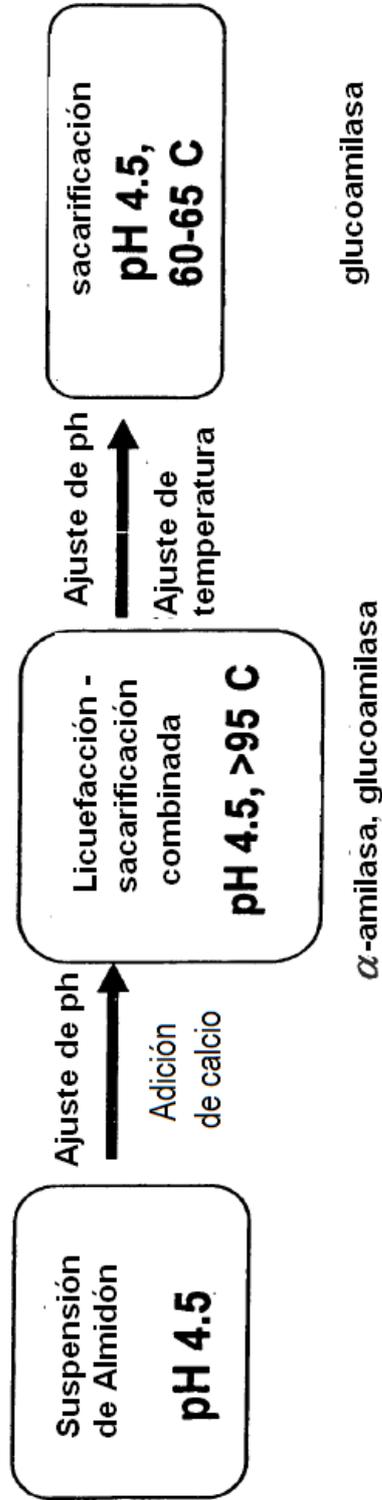


Figura 10

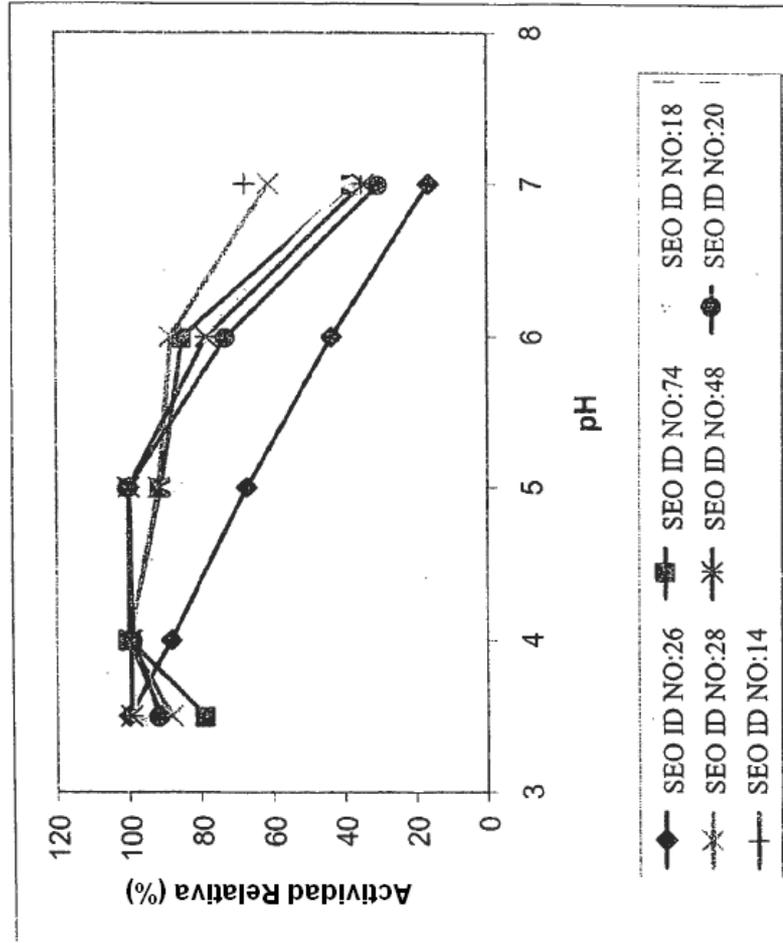


Figura 11

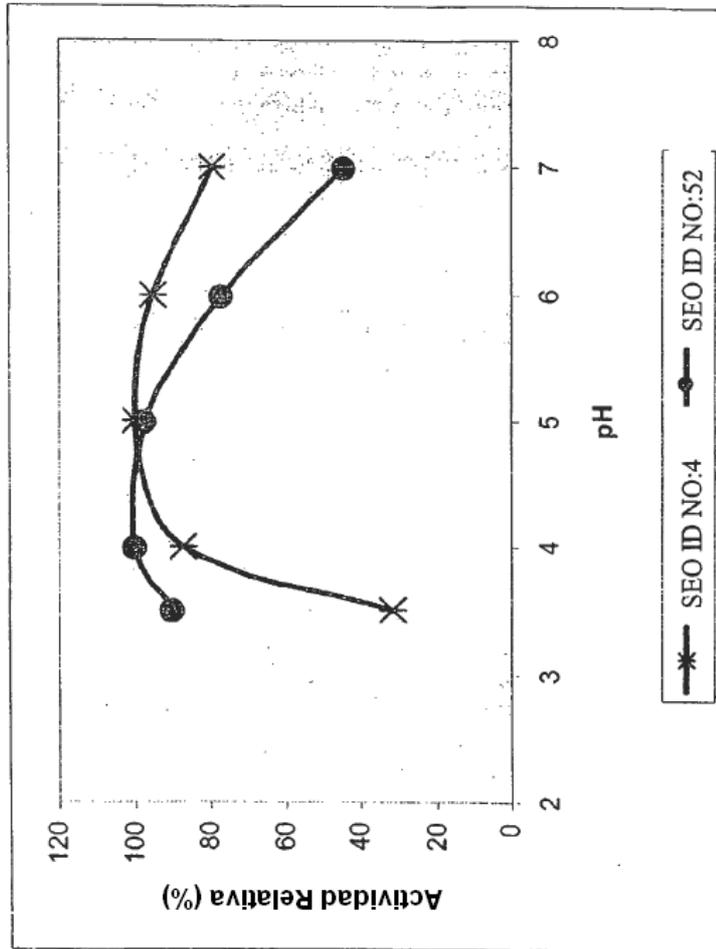


Figura 12

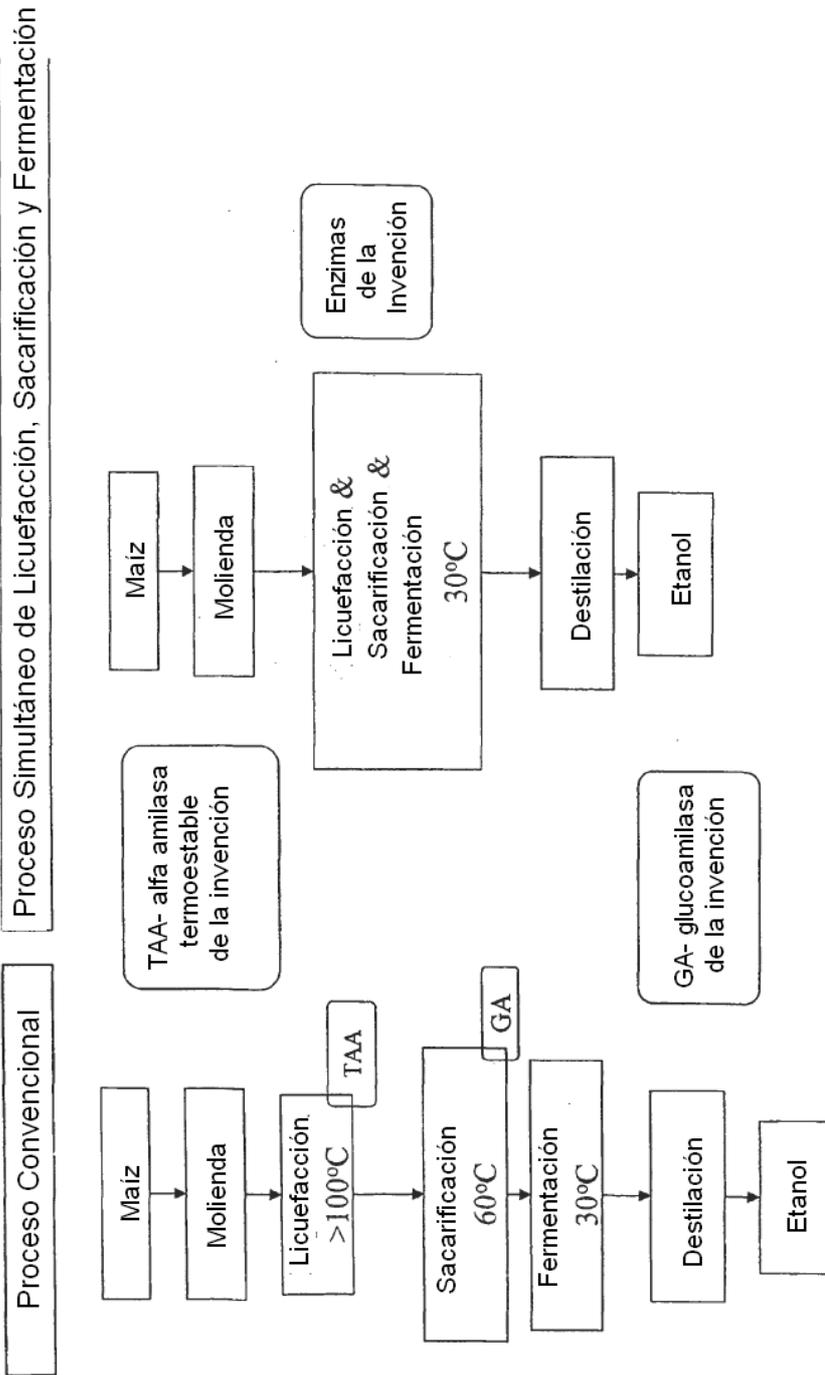


Figura 13

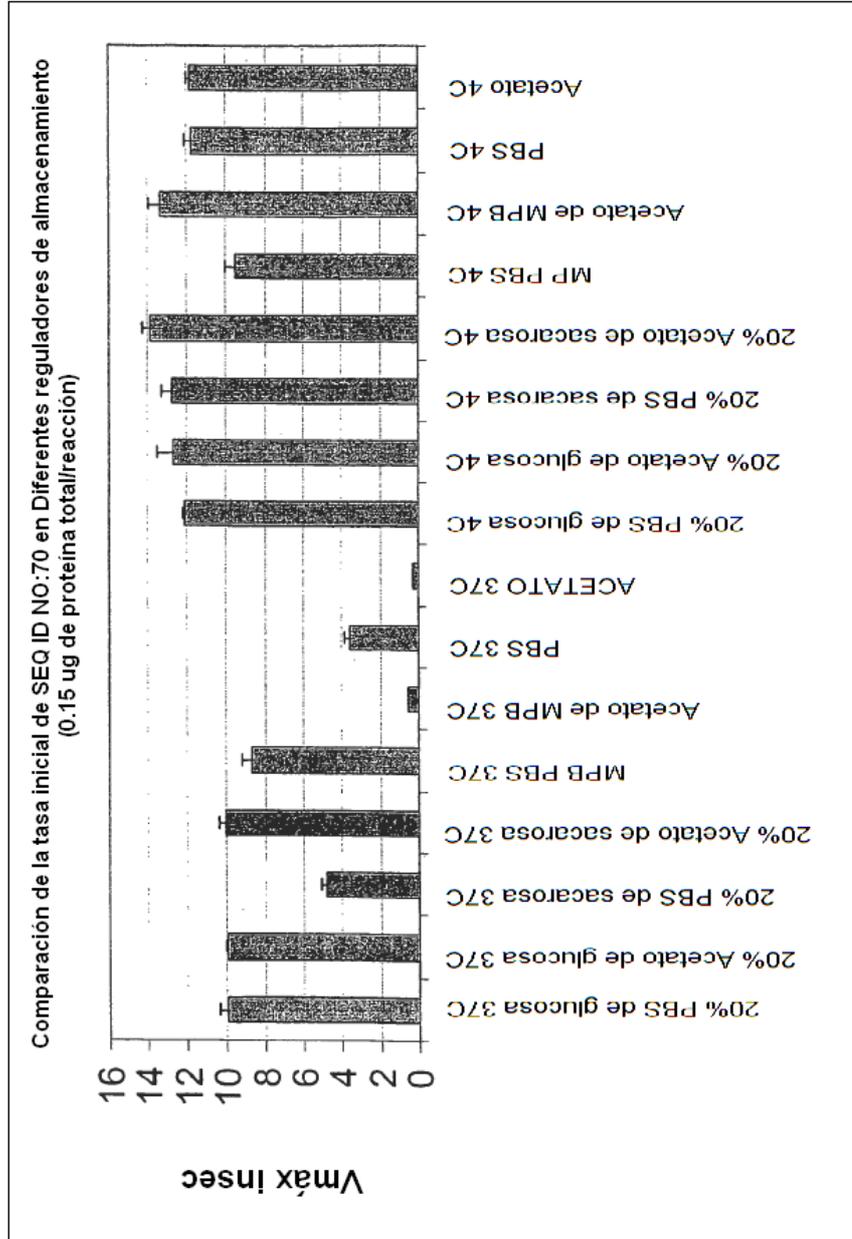


Figura 14

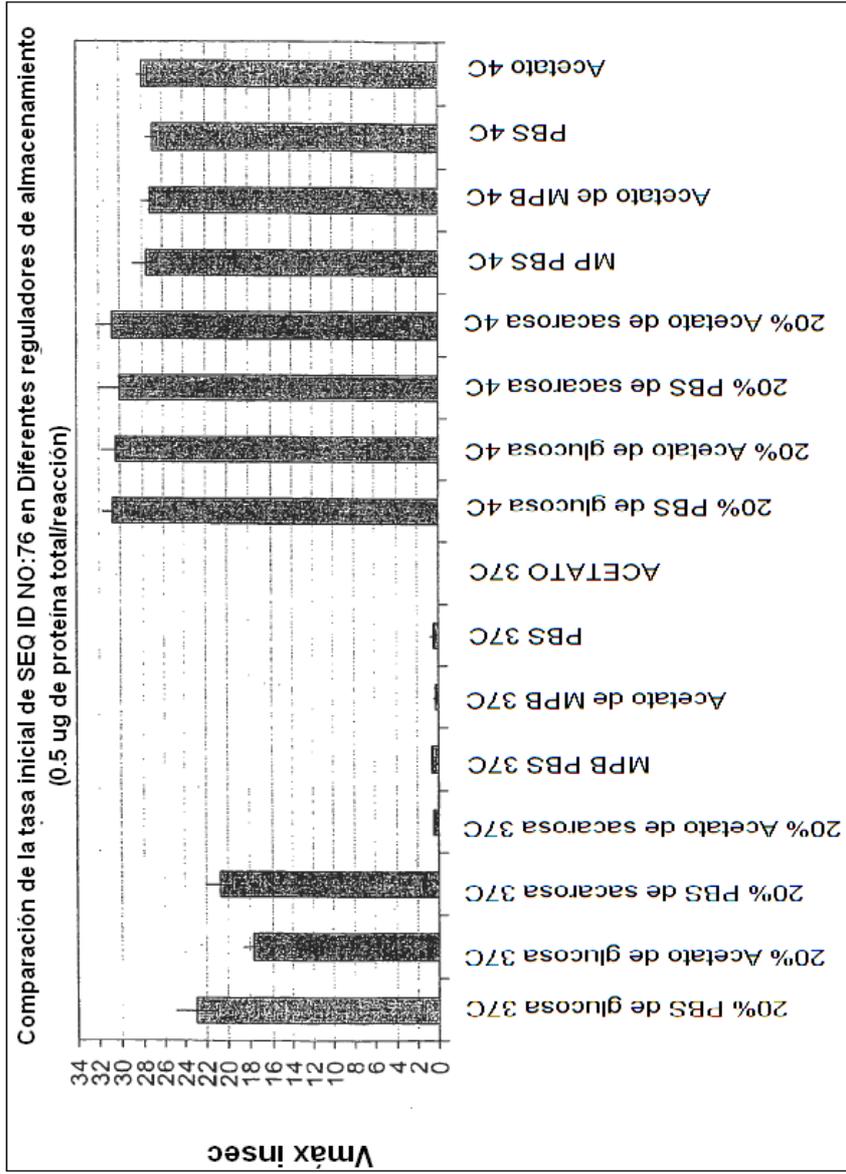
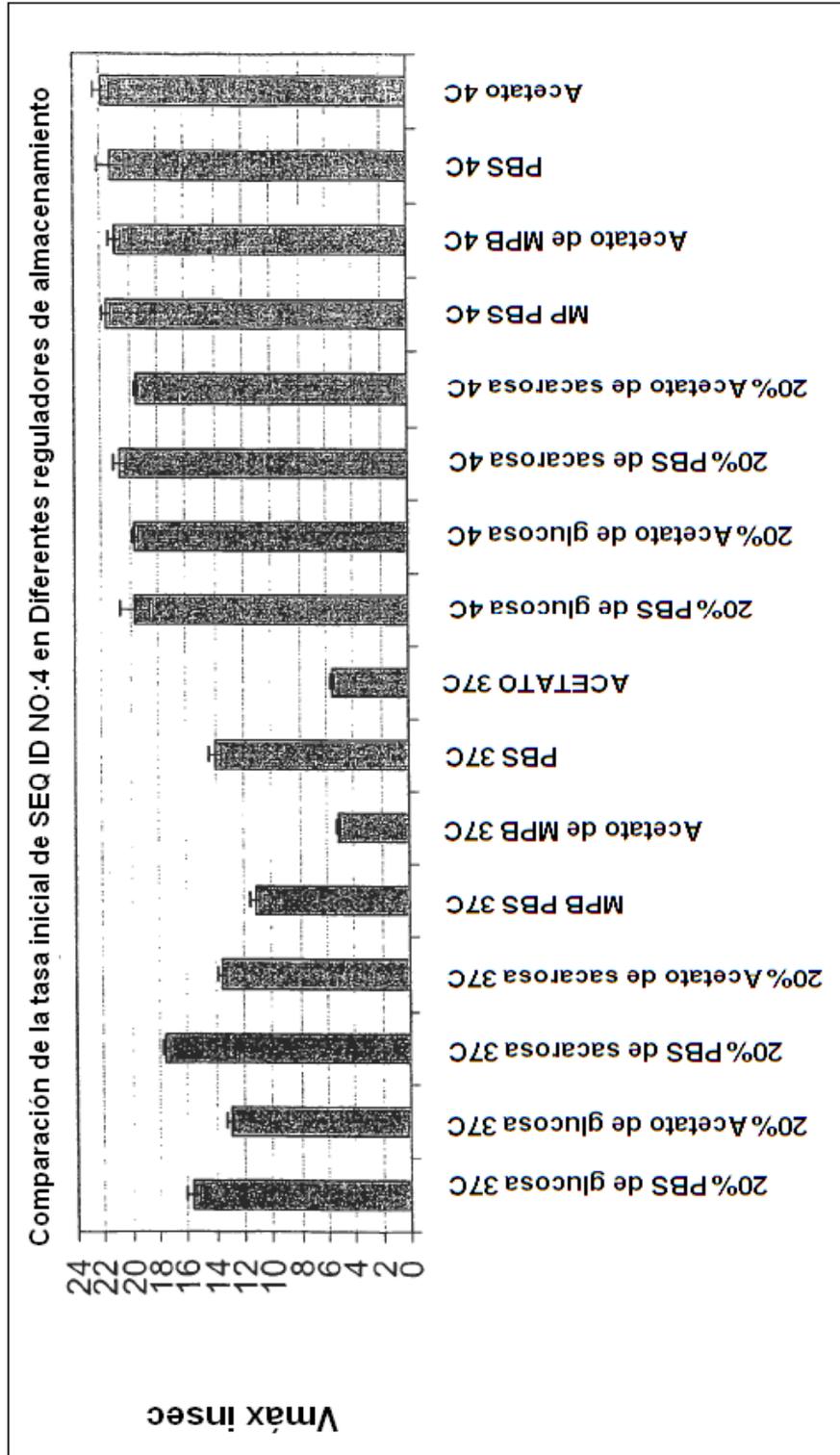


Figura 15



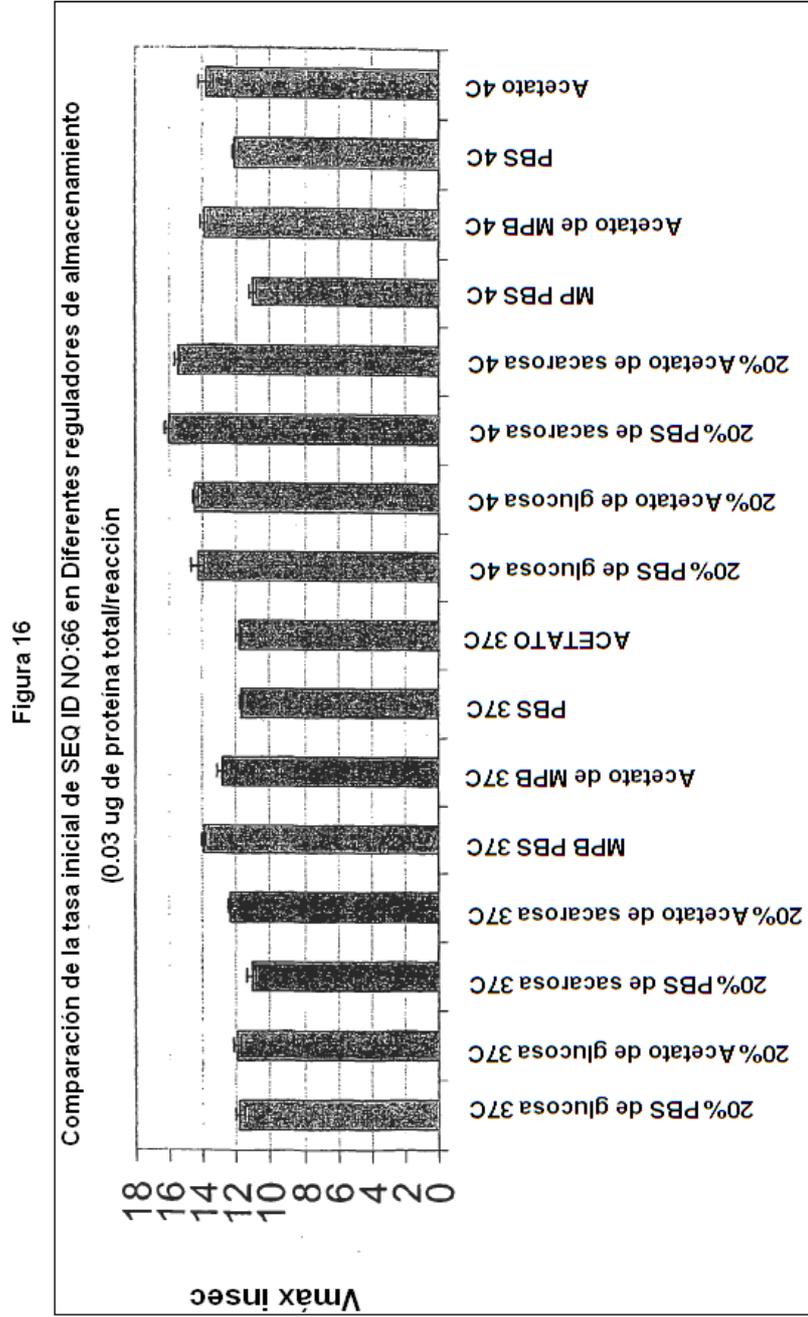
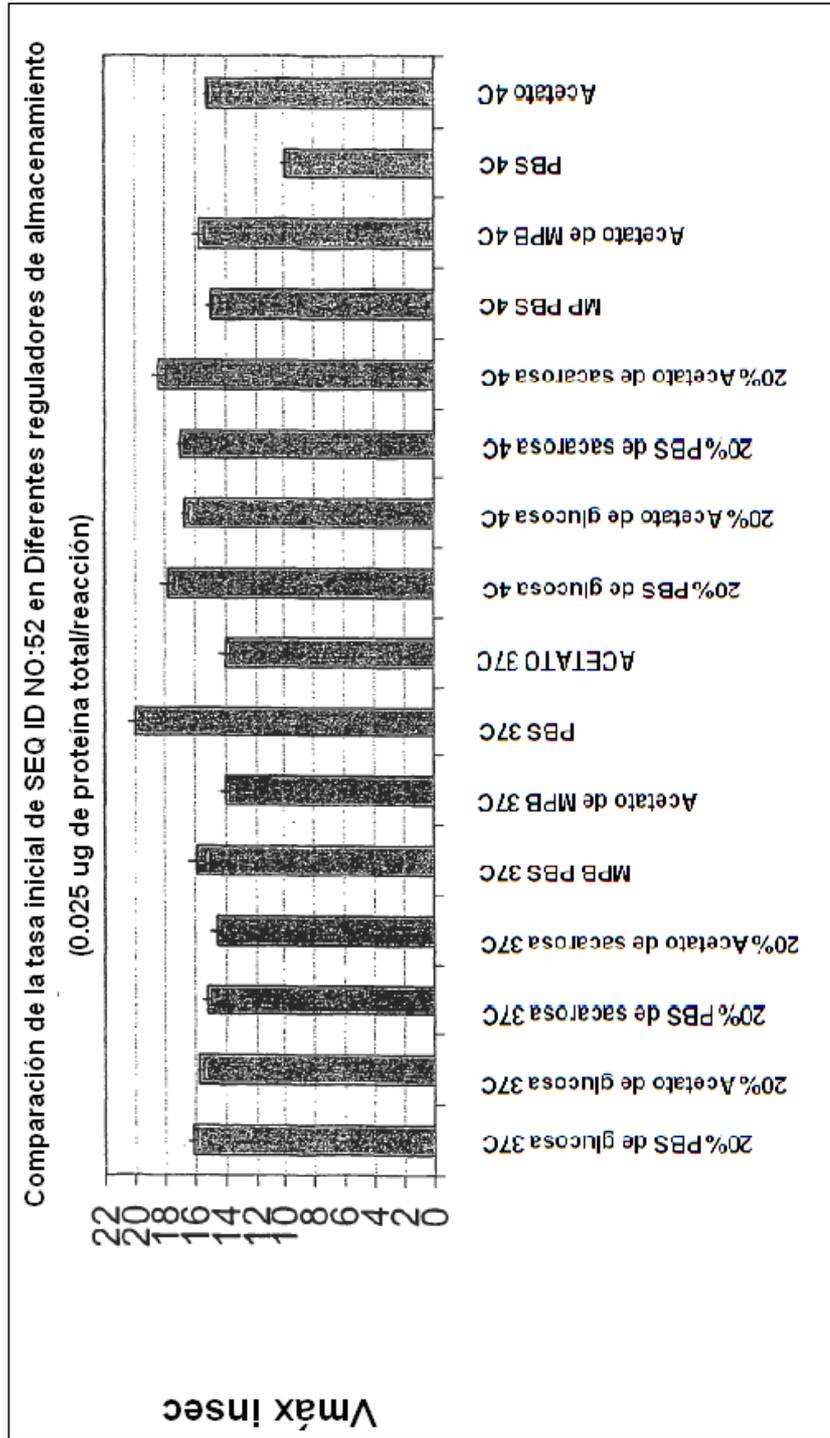


Figura 17



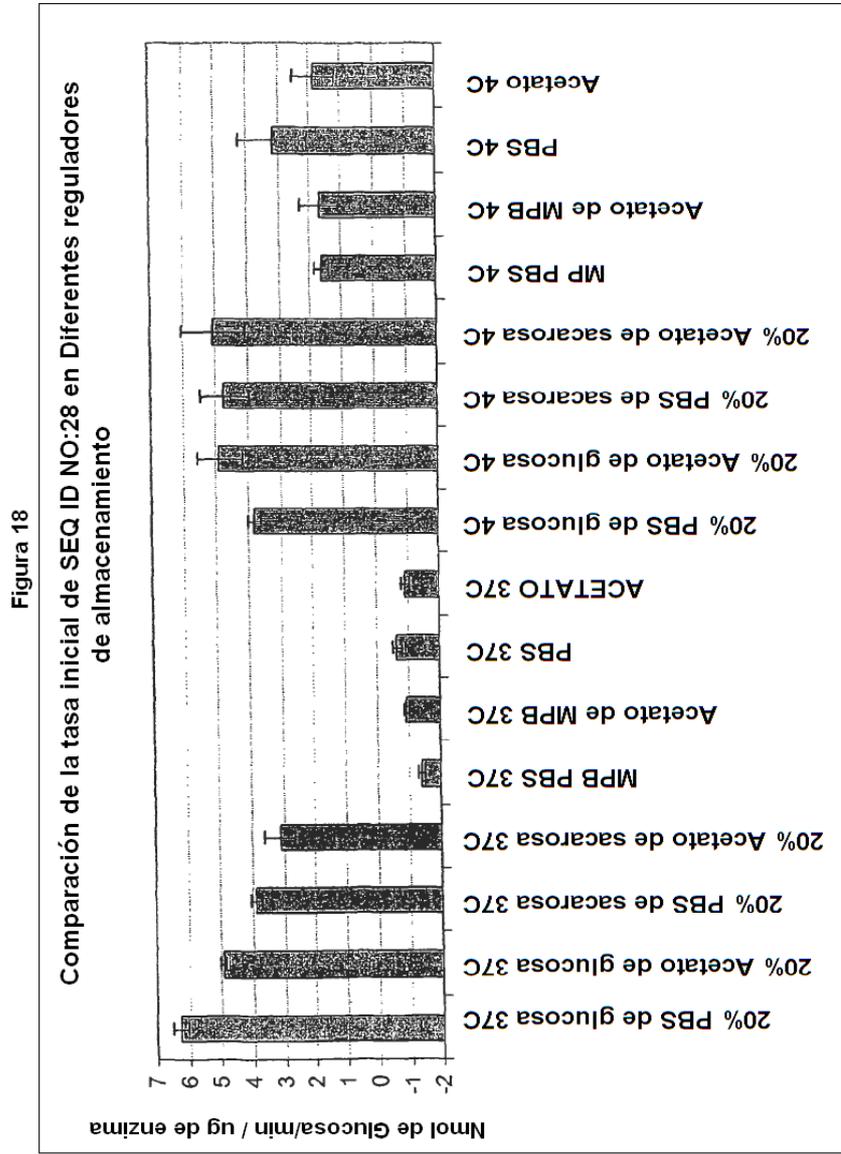


Figura 19

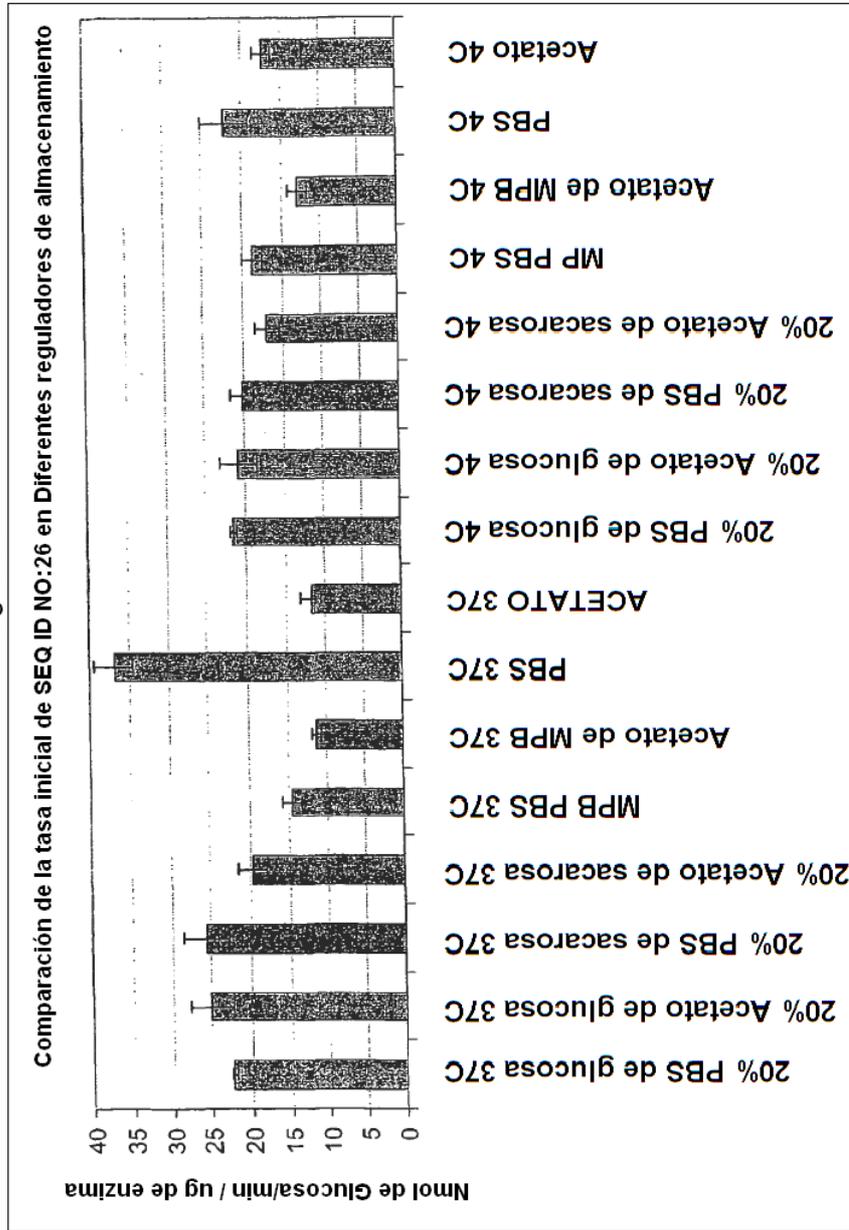


Figura 20

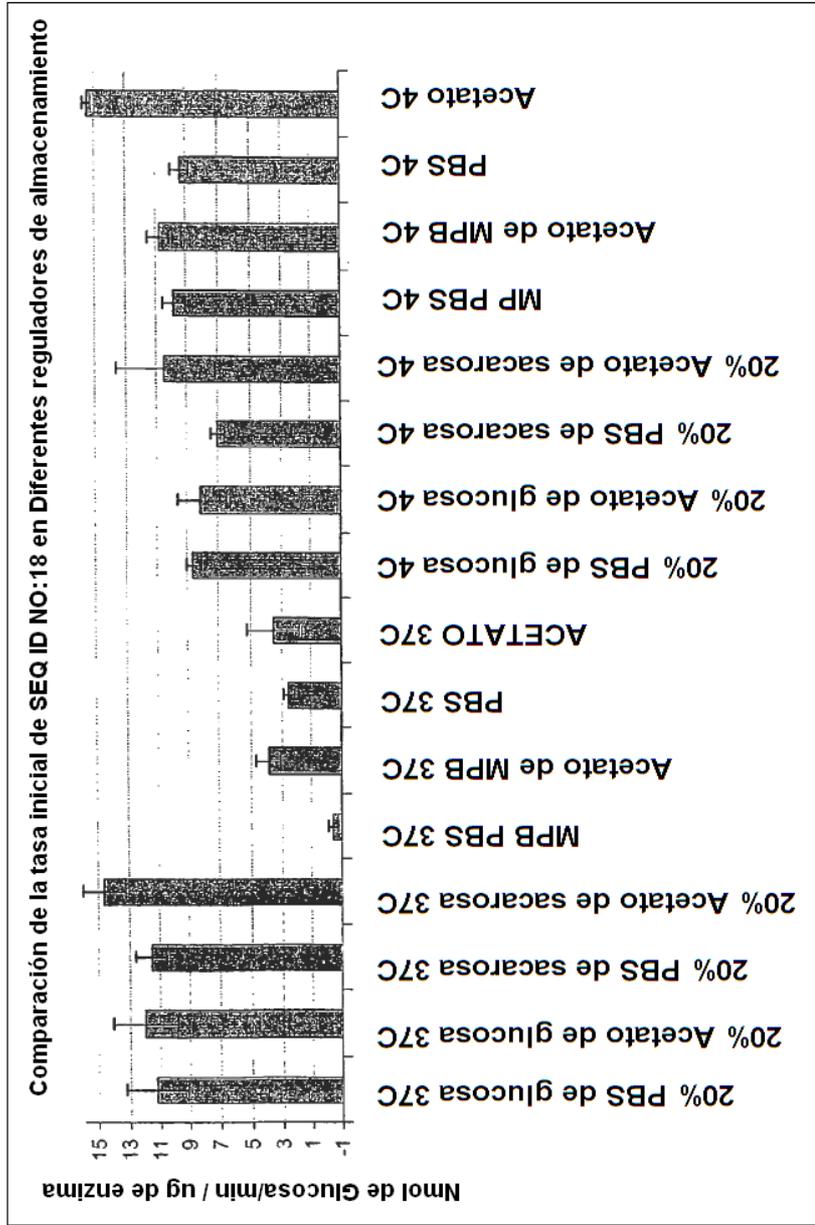


Figura 21

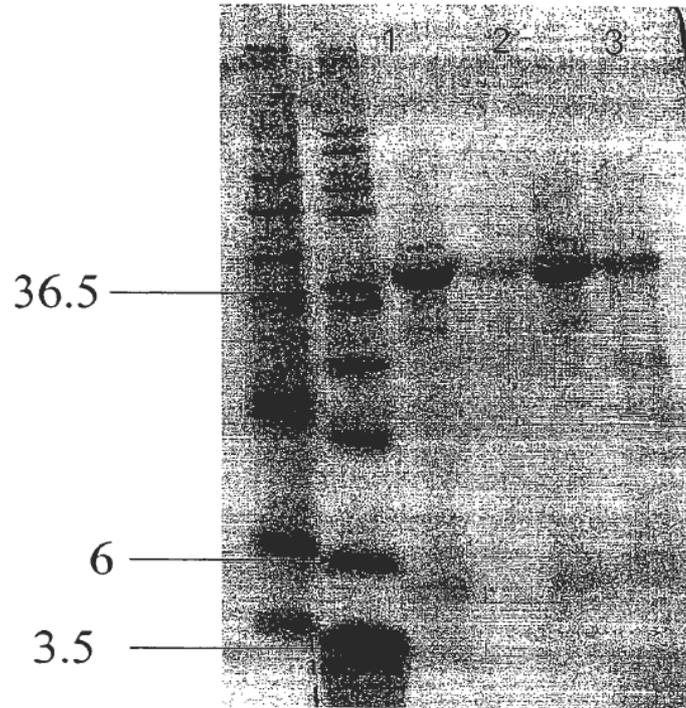


Figura 22

Identificación de péptido por LC MS/MS digerido con tripsina

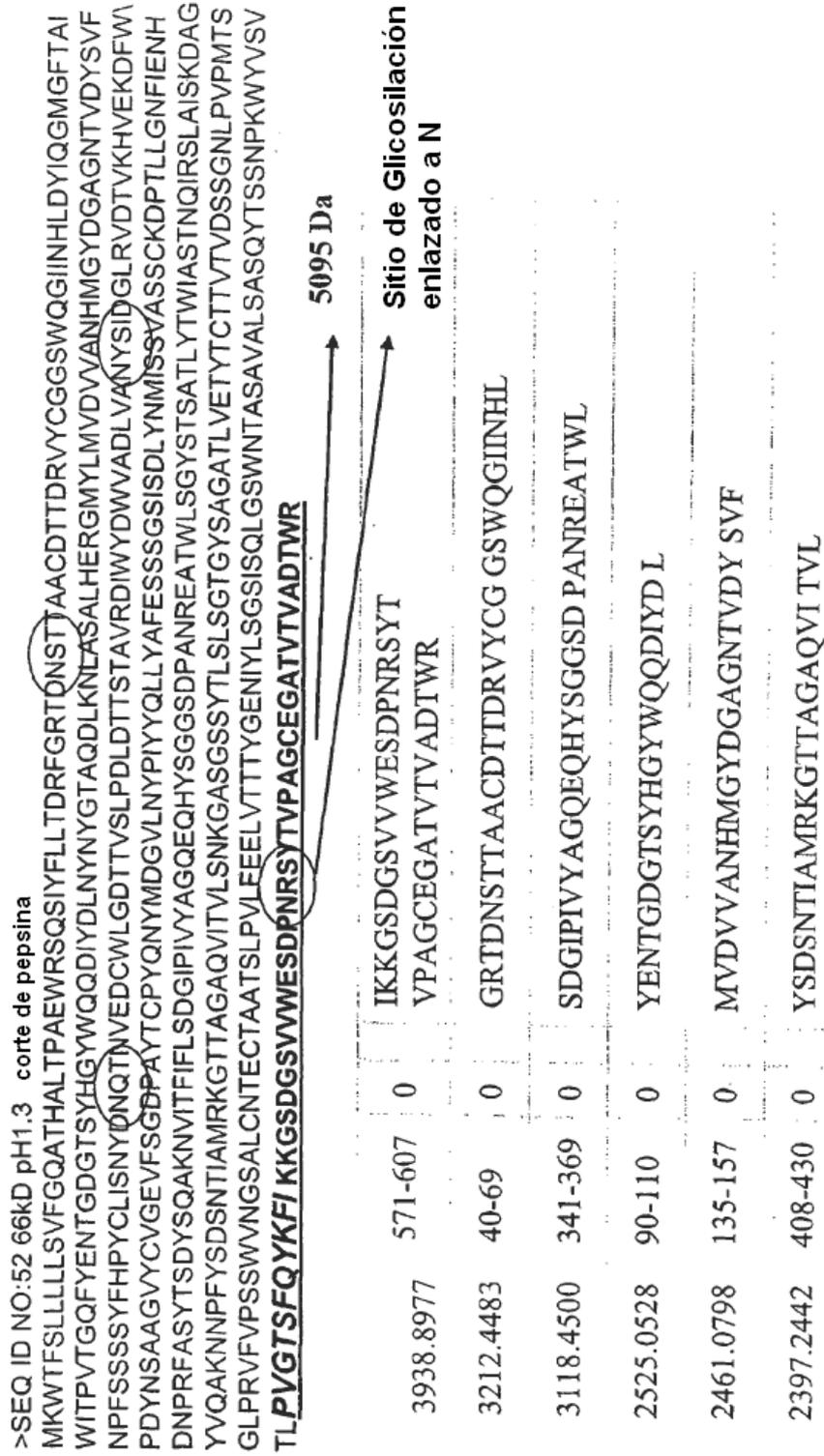


Figura 23A

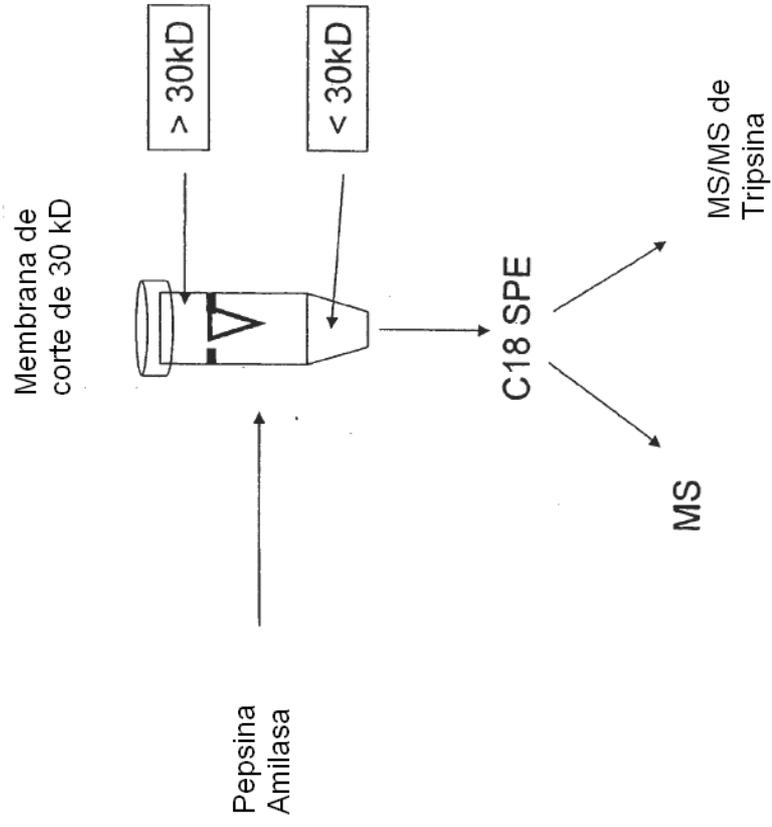


Figura 23B

- 1) Pepsina
- 2) Amilasa SEQ ID NO:52
- 3) SEQ ID NO:52 Pepsina
- 4) Corte superior de 30kd
- 5) Flujo a través de 30kd
- 6) C18 RP capturado

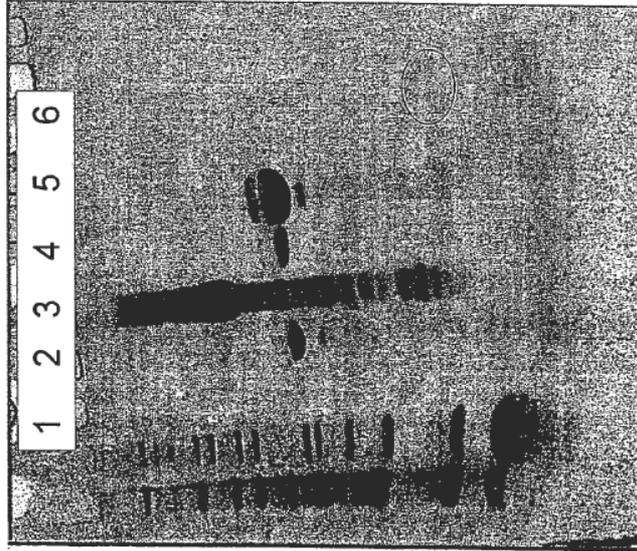


Figura 23C

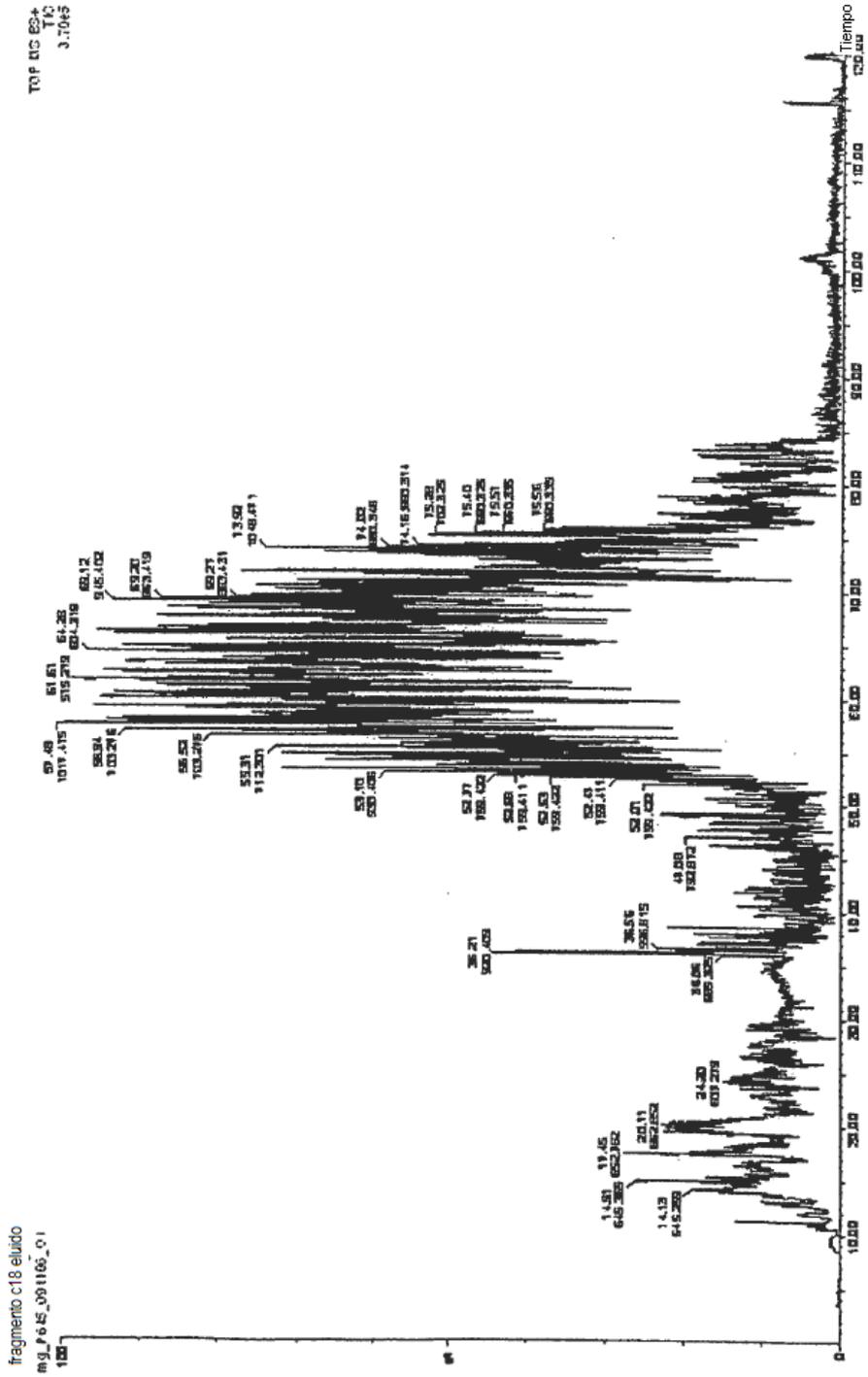


Figura 23D

>SEQ ID NO:52 66kD pH1.3 corte de pepsina
 MKWTFSLLLSVFGQATHALTPAEWRSSQSIYFLLTDRFGRTDNSTTAACDITDRVCGGSWQQIINHLDYIQGMGFTAI
 WITPVTGQFYENTGDGTSYHGYWQQDIYDLNINYGTAQDLKNLASALHERGMYLMVDVVANHMGYDGGAGNTVDYSVF
 NPFSSSYFHPYCLISNYDNQTNVEDCWLGDITVSLPDLTSTAVRDIWYDADLVANYSIDGLRVDTVKHVEKDFW
 PDYNSAAGVYCVGEVFGDPAYTCYQNYMDGVLNYPYQLLYAFESSGSDLYNMISSVASSCKOPTLLGNFIENH
 DNPRFASYSYDYSQAKNVITFIFLSDGIPIVYAGQEQHYSGGSDPANREATWLSGYSTSATLYTWIASTNQIRSLAISKDAG
 YVQAKNPFYSDSNTIAMRKGTAGAQVITVLSNKGAGSSYTLSSLSGTGYAGATLVETVYCTTVVDSGGLPVPMTS
 GLPRVFPSSWVNGSALCNTECTAATSLPVLFEELVTTTYGENIYLSGISQLGSWNTASAVALSASQYTSNPKWVYVS
 TLPVGTSFQYKFIKKGSDGVSVVWESDPNRSYTPAGCEGATVTVADTWR

3938.8977	571-607	0	IKKGS DGSVVWESDPNRSYT VPAGCEGATVTVADTWR
3212.4483	40-69	0	GRITD NSTTAACDITDRVYCG GSWQQIINH L
3118.4500	341-369	0	SDGIPIVYAGQEQHYSGGSD PANREATWL
2525.0528	90-110	0	YENTGDGTSYHGYWQQDIYD L
2461.0798	135-157	0	MVDV VANHMGYDGGAGNTVDY SVF
2397.2442	408-430	0	YSDSN TIAMRKGT TAGA QVI TVL

Figura 23E

Asn-Xaa-Ser/Thr sequins en la producción de la secuencia a continuación están resaltadas en azul.
 Las asparquinas que se predice que están N-glicosiladas están resaltadas en rojo.

Sitio glicosilado putativo

Producción para la "Secuencia"

SEQ ID NO:52

Nombre: Secuencia Longitud: 607

MKWTESLLLLLVFGQATHALTPAEWRSSQSYFLLTDRFRGRIDNSTTAACDITDRVYCGGSWQGIINHLDYIQGMGF
 TAIWITPVTGQFYENTGDCGTSYHGYYWQDDIYDLNXYGTAQDLKNLASALHERGMVLMVDVVAHMGYDAGNTVDYSVFNPF
 SSSSYFHPYCLISNYDQTNVEDCWLGDFTVSLPDLDTTSTAVRDINWYDVAADLVANYSIDGLRVDTVKHVEKDEWPF
 DYMSAAGVYCVGEVFSGDFAYTCEYQNYDGVANYPYYQLLYAFESSGSIISDLXNMISVASSCKDPTLLGNFIEHNDNPR
 FASYTSDYSQARRVITFIFLSDGIFIVYAGQEQHYSGGSDPANREATWLSGYSTSATLYTWIASINQIRSLAISKDG
 YVQAKNPFYSDSNTIAMRKGTTAGAQVITVLSNKGAGSSYTLSLSGTGYAGATLVEITYTCTTIVTVDSSGNLFPVPT
 SGLPRVFPSSWVNGSALCNTECTAATSLEVLFEELVTTYGENIYLSGSIQLGSWNTASAVALSASQYTS
 SNPKWYVSVTL
 PVGTSEQYKEIKKGS
 DGSVVWESDENRSYTVRAGCEGATTVVADTWR

Figura 23F

SEQ ID NO:52

Nombre: Secuencia Longitud: 607

MKWTFSLLLLSVFGQATHALTPAERRSQSIYFLLTDRFGRIDNSTTACDITDRVYCGGSAQGIIMHDDYIQGMGFTAI
 WITPVTGQFYENTGDGTSYHGYNQDIYDLNRYNGTAQDLKNLASALHERGNLYLMVDVVANHMGYDAGNTVDYSVENPE
 SSSSYFHPYCLISNYDNQTNVEDCWLGDITVSLPDLDTTSTAVRDIWYDWNVADLVANYSIDGLRVDVTKHVEKDFWPDYD
 SAAGVYCVGEVFSGDPAYICPYONXMDGVLNYP IYYQLLYAFESSGSSIDLYNMISSVASSCKDPTLLGNF IENHDNPR
 FASYTSDYSQAKNVITFIFLSDGIP IVYAGQEQHYSGGSDRANREATWLSGYSTSATLYTWIASTNQIRSLAISK DAGYV
 QAKNPFYSDSNTIAMRKGTAGAQVITVLSNKGASGSSYTLSSLGTGYSAGATLVEYYTCTTIVTVDSSGNLFPVMTSGL
 PRFVPSWVNGSALCHNTECTAATSLEPLFEELVTTTYGENIYLSGSIQGLGSWNTASAVALSASQYTSNPKWYVSVTL
 PVGTSEQYKFIKKGSDGSVVWESDPRNSYTVPAGCEGATVTVADINR

Figure 24

Enzima	Velocidad Inicial* \pm SD de Almidón Granulado	Velocidad Inicial* \pm SD de Almidón Soluble
SEQ ID NO:48	35.6 \pm 3.8	60.7 \pm 5.1
SEQ ID NO:26	28.8 \pm 3.4	51.8 \pm 7.4
SEQ ID NO:74	25.1 \pm 2.5	84.3 \pm 3.5
SEQ ID NO:18	24.3 \pm 4.3	58.3 \pm 3.0
SEQ ID NO:28	17.8 \pm 4.3	33.8 \pm 4.0
SEQ ID NO:14	6.6 \pm 1.2	53.9 \pm 4.1
SEQ ID NO:20	0	59.3 \pm 8.5
<i>A.niger</i>	11.3 \pm 2.7	43.3 \pm 7.4

Figura 25

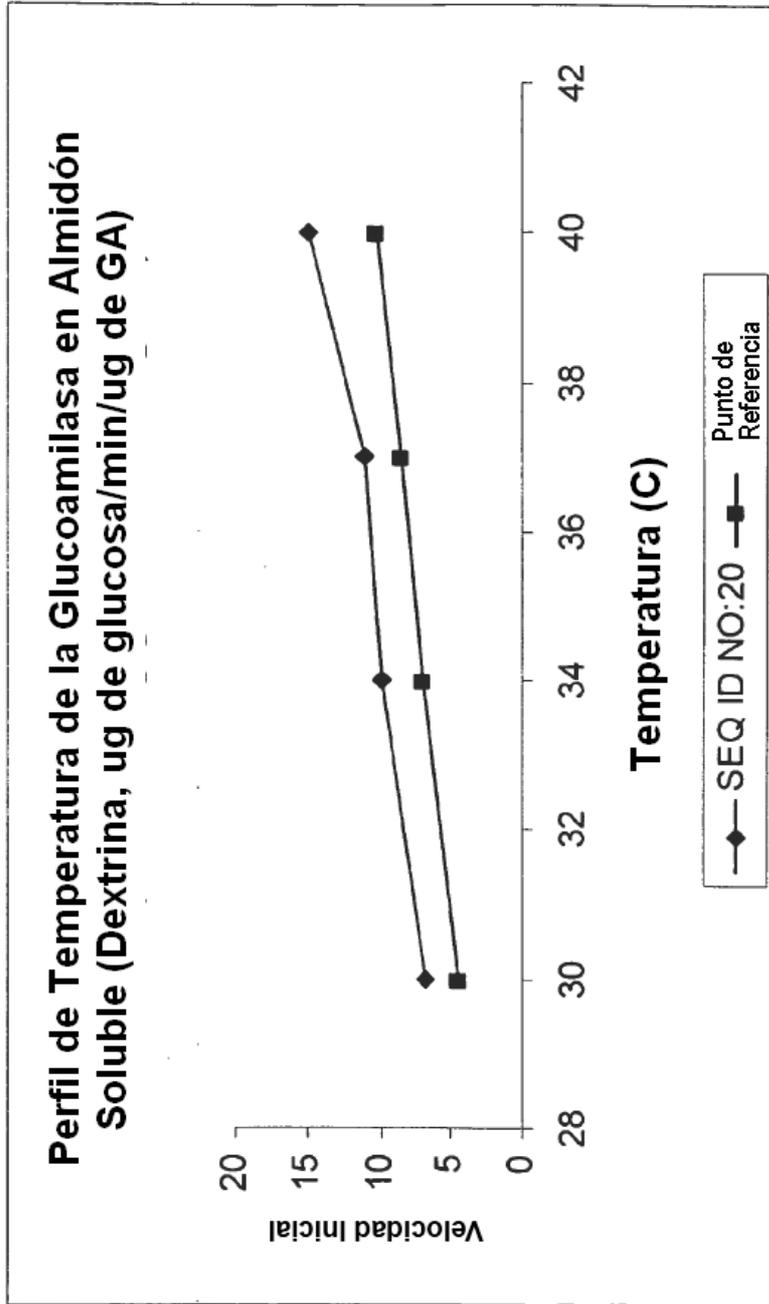


Figura 26

Resumen de la pureza				
	Antes de la purificación		Después de la Purificación	
	A260/A280	Relación de la expresión	A260/A280	Relación de la expresión
SEQ ID NO:18	0.52	91%	0.59	92%
SEQ ID NO:26	0.56	85%	0.65	86%
SEQ ID NO:4	0.67	91%	0.73	85%

Resumen de la actividad				
	Almidón crudo		Soluble	
	Purificado	Sin purificar	Purificado	Sin purificar
SEQ ID NO:18	10.7228 ± 0.7925	6.4828 ± 0.5681	40.6725 ± 1.7016	24.7354 ± 3.821
SEQ ID NO:26	5.7837 ± 0.2271	4.0691 ± 0.2864	17.4833 ± 0.7831	14.1849 ± 0.3761
SEQ ID NO:4	6.47 ± 0.34	4.074 ± 0.35	127.37 ± 1.78	116.52 ± 1.27

Figura 27

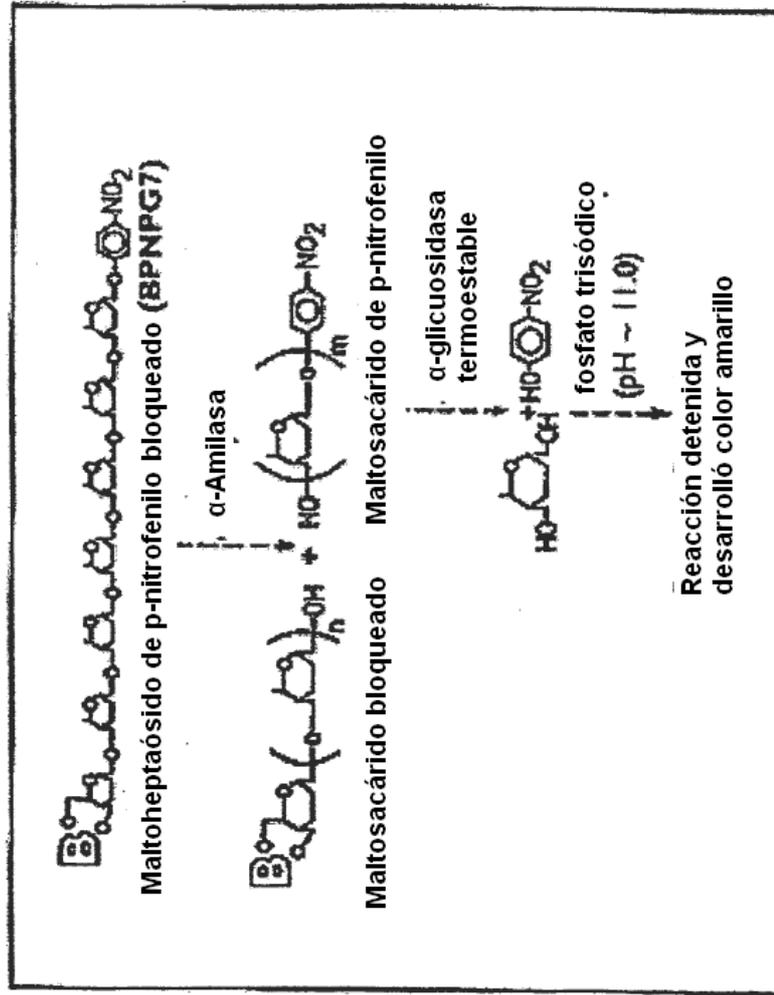


Figura 28A

**Perfil de Temperatura de Glucoamilasas en Almidón Granulado -
Velocidad Inicial (ug de glucosa/min/ug de glucoamilasa)**

Temperatura (°C)	SEQ ID NO:74	SEQ ID NO:28	SEQ ID NO:26	Punto de referencia	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:48	SEQ ID NO:20
30	3.04	1.64	3.12	1.4	0.96	3.92	4.6	inactiva
34	3.92	2.24	4.36	1.56	1.08	5.2	5.64	inactiva
37	4.36	2.76	5.16	1.84	1.36	5.24	6.84	inactiva
40	5.12	3.32	5.92	1.84	1.44	5.72	7.84	inactiva

Figura 28B

**Perfil de Temperatura de Glucoamilasas en Almidón Soluble -
(Dextrina) Velocidad Inicial (ug de glucosa/min/ug de glucoamilasa)**

Temperatura (°C)	SEQ ID NO:20	Punto de referencia
30	6.840	4.380
34	9.920	6.980
37	11.080	8.600
40	15.040	10.380

Figura 29

Figura 29A

Perfil de pH de Glucoamilasas en Almidón Granular - Porcentaje de Actividad Relativa

pH	SEQ ID NO:26	SEQ ID NO:74	SEQ ID NO:18	SEQ ID NO:28	SEQ ID NO:14	SEQ ID NO:48	Punto de referencia	SEQ ID NO:20
3.5	100.0000	78.6207	98.1308	88.1720	60.6061	99.3151	100.0000	inactiva
4.0	88.1081	100.0000	100.0000	100.0000	66.6667	99.3151	100.0000	inactiva
5.0	67.0270	91.0345	87.8505	91.3978	66.6667	100.0000	100.0000	inactiva
6.0	43.2432	84.8276	79.4393	88.1720	100.0000	78.0822	40.0000	inactiva
7.0	16.2162	37.2414	38.3178	61.2903	45.4545	34.2466	18.6667	inactiva

Figura 29B

Perfil de pH de Glucoamilasas en Almidón Soluble - (Dextrina)

Porcentaje de Actividad Relativa

pH	SEQ ID NO:74	SEQ ID NO:14	SEQ ID NO:18	SEQ ID NO:28	SEQ ID NO:20	SEQ ID NO:26	SEQ ID NO:48	Punto de referencia
3.5	83.4997	72.1981	81.9000	79.4340	91.6548	98.2773	91.3527	91.3255
4.0	85.6287	76.7159	88.3000	91.1321	98.7874	100.0000	89.8116	100.0000
5.0	90.3526	83.0582	100.0000	100.0000	100.0000	85.4799	100.0000	80.8967
6.0	100.0000	100.0000	80.9000	98.6792	72.9672	46.4315	78.5103	67.9337
7.0	75.5156	81.2337	60.6922	68.8679	29.8859	11.8130	29.1952	18.3236

Figura 30

Perfil de Temperatura de Amilasas en Almidón Granular . Porcentaje de Actividad relativa

Temp. (°C)	SEQ ID NO:56	SEQ ID NO:70	SEQ ID NO:62	SEQ ID NO:66	SEQ ID NO:2	SEQ ID NO:52	SEQ ID NO:78	SEQ ID NO:76	Sigma-A6211	Megazyme E-ANAAM
30	6.44	0.82	2.24	0.26	6.05	8.99	0.14	1.26	0.42	0.167432
34	10.98	1.29	3.26	0.47	6.97	11.21	0.27	2.20	0.49	0.229423
37	14.61	1.88	3.25	0.75	7.46	11.53	0.36	2.85	0.55	0.307667
40	17.66	2.36	3.11	1.27	6.93	11.40	0.51	3.26	0.54	0.29692

Figura 31

Figura 31A

Perfil de pH de Amilasas en Almidón Granular - Porcentaje de Actividad Relativa

pH	SEQ ID NO:56	SEQ ID NO:70	SEQ ID NO:62	SEQ ID NO:66	SEQ ID NO:2	SEQ ID NO:52	SEQ ID NO:78	SEQ ID NO:76	Sigma-A6211	Megazyme E-ANAAM
3.5	0.00	3.69	0.00	24.31	31.59	89.70	6.89	6.87	72.06	51.25
4.0	0.00	3.88	0.00	70.83	86.95	100.00	9.85	9.33	94.34	74.69
5.0	93.07	45.60	50.20	100.00	100.00	96.97	87.70	72.80	100.00	100.00
6.0	100.00	100.00	100.00	97.10	95.08	76.59	100.00	100.00	78.16	62.43
7.0	92.52	99.12	85.70	64.29	79.10	43.67	61.30	71.84	71.04	47.70

Figura 31B

Perfil de pH de Amilasas en Almidón Soluble - Porcentaje de Actividad Relativa

pH	SEQ ID NO:86	SEQ ID NO:70	SEQ ID NO:62	SEQ ID NO:66	SEQ ID NO:2	SEQ ID NO:52	SEQ ID NO:78	SEQ ID NO:76	Sigma-A6211	Megazyme E-ANAAM
3.5	0.00	2.23	0.00	11.00	51.35	100.00	1.22	1.25	38.55	46.04
4.0	6.94	2.57	0.00	47.94	89.60	96.63	18.89	3.62	63.72	60.97
5.0	74.57	37.01	35.90	100.00	100.00	83.16	84.63	83.03	100.00	100.00
6.0	97.83	65.88	93.03	86.33	94.27	60.31	100.00	100.00	57.80	50.62
7.0	100.00	100.00	100.00	56.96	87.91	52.17	66.50	27.79	11.53	13.24