(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 368 555**

(51) Int. CI.:

C12P 21/06	(2006.01)
C12N 9/00	(2006.01)
C12N 9/24	(2006.01)
C12N 1/20	(2006.01)
C12N 15/00	(2006.01)
C07H 21/04	(2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Número de solicitud europea: **02706401 .3**

(96) Fecha de presentación: **21.02.2002**

(97) Número de publicación de la solicitud: **1370674**

(97) Fecha de publicación de la solicitud: **17.12.2003**

(54) Título: **ENZIMAS QUE TIENEN ACTIVIDAD DE ALFA-AMILASA Y MÉTODOS DE USO DE LAS MISMAS.**

(30) Prioridad:
21.02.2001 US 270495 P
21.02.2001 US 270496 P
14.05.2001 US 291122 P

(45) Fecha de publicación de la mención BOPI:
18.11.2011

(45) Fecha de la publicación del folleto de la patente:
18.11.2011

(73) Titular/es:
VERENIUM CORPORATION
4955 DIRECTORS PLACE
SAN DIEGO, CA 92121, US

(72) Inventor/es:
CALLEN, Walter;
RICHARDSON, Toby;
FREY, Gerhard;
MILLER, Carl;
KAZAOKA, Martin;
MATHUR, Eric, J. y
SHORT, Jay, M.

(74) Agente: **Mir Plaja, Mireia**

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

Enzimas que tienen actividad de alfa-amilasa y métodos de uso de las mismas

Campo de la Invención

[0001] Esta invención se refiere en general a las enzimas, a los polinucleótidos que codifican las enzimas, al uso de tales polinucleótidos y polipéptidos, y más específicamente, a las enzimas que tienen actividad de alfa-amilasa.

Antecedentes de la Invención

[0002] El almidón es un carbohidrato complejo que se encuentra a menudo en la dieta humana. La estructura del almidón es polímeros de glucosa enlazados por enlaces α-1,4 y α-1,6 glucosídicos. La amilasa es una enzima que cataliza la hidrólisis de los almidones para su transformación en azúcares. Las amilasas hidrolizan enlaces α-1,4-glucosídicos internos en el almidón, en gran medida aleatoriamente, para así producir maltodextrinas de menor peso molecular. La descomposición del almidón es importante en el sistema digestivo y comercialmente. Las amilasas son de considerable valor comercial, siendo usadas en las etapas iniciales (licuefacción) del procesamiento de almidón; en la molienda de maíz en húmedo; en la producción de alcohol; como agentes limpiadores en matrices de detergente; en la industria textil para el desalmidonado; en aplicaciones de horneo; en la industria de las bebidas; en campos petrolíferos en procesos de perforación; en el entintado de papel reciclado y en piensos para animales. Las amilasas son también útiles en el desencolado textil, en procesos de fermentación, en la modificación del almidón en la industria del papel y de la pasta papelera y en otros procesos descritos en la técnica.

[0003] Las amilasas son producidas por una amplia variedad de microorganismos entre los que se incluyen los *Bacillus* y los *Aspergillus*, siendo las de la mayoría de las amilasas comerciales producidas a partir de fuentes bacterianas tales como *Bacillus licheniformis, Bacillus amyloliquefaciens*, *Bacillus subtilis* o *Bacillus stearothermophilus*. En los últimos años, las enzimas que han venido siendo usadas comercialmente han venido siendo las obtenidas de *Bacillus licheniformis* debido a su estabilidad térmica y a su actuación, al menos a pHs neutros y moderadamente alcalinos.

[0004] La FR 2 778 412 A1 y Lévêque et al., FEMS Microbiol Lett. 1 mayo 2000; 186(1):67-71 describen un proceso para producir α-amilasa termofílica que comprende la expresión del gen de α-amilasa de Thermococcus hydrothermalis en E. coli.

[0005] En general, el procesamiento del almidón para su conversión en fructosa consta de cuatro pasos: licuefacción del almidón granular, sacarificación del almidón licuado para su conversión en dextrosa, purificación e isomerización a fructosa. El objeto de un proceso de licuefacción de almidón es el de convertir una suspensión concentrada de gránulos de polímero de almidón en una solución de dextrinas solubles de baja viscosidad y de menor longitud de cadena. Este paso es esencial para la cómoda manipulación con equipos estándar y para una eficiente conversión en glucosa u otros azúcares. Para licuar el almidón granular es necesario gelatinizar los gránulos aumentando la temperatura del almidón granular hasta más de aproximadamente 72ºC. El proceso de calentamiento rompe instantáneamente los gránulos de almidón insolubles para así producir una solución de almidón hidrosoluble. La solución de almidón solubilizada es luego licuada por amilasa. Un gránulo de almidón se compone de: un 69-74% de amilopectina, un 26-31% de amilosa, un 11-14% de agua, un 0,2-0,4% de proteína, un 0,5-0,9% de lípido, un 0,05-0,1% de ceniza, un 0,02-0,03% de fósforo y un 0,1% de pentosan. Aproximadamente un 70% de un gránulo es amorfo y un 30% es cristalino.

[0006] Un proceso de licuefacción enzimática común supone ajustar el pH de una lechada de almidón granular a un valor situado entre 6,0 y 6,5, que es el óptimo valor pH de la alfa-amilasa derivada de Bacillus licheniformis, con adición de hidróxido de calcio, hidróxido sódico o carbonato sódico. La adición de hidróxido de calcio tiene la ventaja de proporcionar también iones de calcio, que se sabe que estabilizan a la alfa-amilasa contra la inactivación. Tras la adición de alfa-amilasa, la suspensión es bombeada a través de un chorro de vapor para aumentar instantáneamente la temperatura hasta un valor situado entre 80 grados y 115 grados C. El almidón es inmediatamente gelatinizado y, debido a la presencia de la alfa-amilasa, despolimerizado por la alfa-amilasa mediante hidrólisis aleatoria de enlaces glicosídicos a-(1-4), siendo así convertido en una masa fluida que se bombea fácilmente.

[0007] En una segunda variación del proceso de licuefacción, la alfa-amilasa es añadida a la suspensión de almidón, la suspensión es mantenida a una temperatura de 80-100 grados C para hidrolizar parcialmente los gránulos de almidón, y la suspensión de almidón parcialmente hidrolizado es bombeada a través de un chorro que está a temperaturas de más de aproximadamente 105 grados C para así gelatinizar a fondo toda estructura granular restante. Tras haber enfriado el almidón gelatinizado, puede hacerse una segunda adición de alfa-amilasa para hidrolizar adicionalmente el almidón.

[0008] Una tercera variación de este proceso es el llamado proceso de molienda en seco. En la molienda en seco se muele grano entero y se le combina con agua. El germen es opcionalmente eliminado mediante separación por flotación o bien mediante técnicas equivalentes. La mezcla resultante, que contiene almidón, fibra, proteína y otros componentes del grano, es licuada usando alfa-amilasa. La práctica general en la técnica es la de realizar la licuefacción enzimática a una temperatura más baja cuando se usa el proceso de molienda en seco. En general se cree que la licuefacción a baja

temperatura es menos eficiente que la licuefacción a alta temperatura para la conversión de almidón en dextrinas solubles.

[0009] Típicamente, tras la gelatinización la solución de almidón es mantenida a una temperatura elevada en presencia de alfa-amilasa hasta ser alcanzado un DE de 10-20, lo cual representa habitualmente un periodo de tiempo de 1-3 horas. El equivalente de dextrosa (DE) es el estándar industrial para medir la concentración de azúcares reductores totales, calculada como D-glucosa sobre la base del peso en seco. El almidón granular no hidrolizado tiene un DE de virtualmente cero, mientras que el DE de la D-glucosa es por definición de 100.

[0010] La molienda de maíz en húmedo es un proceso que produce aceite de maíz, harina de gluten, pienso de gluten y almidón. A menudo se usa amilasa alcalina en la licuefacción de almidón y se usa glucoamilasa en la sacarificación, produciendo glucosa. El maíz, que es un grano que consta de un revestimiento exterior de la semilla (fibra), almidón, una combinación de almidón y glucosa y el germen interior, es sometido a un proceso que se desarrolla en cuatro pasos y redunda en la producción de almidón. El maíz es puesto en remojo, desgerminado y desfibrado, y finalmente se separa el gluten. En el proceso de maceración se quitan los solubles. El producto que queda tras haber quitado los solubles es desgerminado, lo cual redunda en una producción de aceite de maíz y una producción de una torta de aceite que es añadida a los solubles del paso de maceración. El producto que queda es desfibrado y los sólidos de fibra son añadidos a la mezcla de torta de aceite y solubles. La mezcla de sólidos de fibra, torta de aceite y solubles forma un pienso de gluten. Tras el desfibrado, el producto que queda es sometido a separación del gluten. Esta separación redunda en harina de gluten y almidón. El almidón es luego sometido a licuefacción y sacarificación para producir glucosa.

[0011] El enranciamiento de los productos horneados (tal como el pan) ha sido reconocido como un problema que se agrava en la medida en que media un mayor espacio de tiempo entre el momento de preparación del producto de panificación y el momento de su consumo. El vocablo “enranciamiento” se usa para describir cambios indeseables para el consumidor en las propiedades del producto de panificación tras haber salido el mismo del horno, tales como un incremento de la firmeza de la miga, una disminución de la elasticidad de la miga y cambios en la costra, que deviene dura y correosa. La firmeza de la miga del pan sigue aumentando durante el almacenamiento hasta llegar a alcanzar un nivel que se considera negativo. El incremento de la firmeza de la miga, que es considerado como el aspecto más importante del enranciamiento, es reconocido por el consumidor mucho tiempo antes de que el producto de panificación haya llegado a ser de otra manera inadecuado para el consumo.

[0012] El proceso de preparación de jarabe contiene proporciones de polímeros con grados de polimerización (DP) de 4 o más, lo cual podría ser perjudicial. La patente U.S. Nº 5.141.859 propuso un proceso para la preparación de un jarabe con alto contenido de maltosa en el que se emplean dos pasos de sacarificación sucesivos. Este documento propugna de hecho un proceso que comprende un primer paso de sacarificación en presencia de una beta-amilasa y un subsiguiente paso de sacarificación en presencia de una alfa-amilasa maltogénica. Según este documento, la alfa-amilasa maltogénica se usa después del primer paso de sacarificación con beta-amilasa para hidrolizar los oligosacáridos (de DP 3 a DP 7) y en esencia la maltotriosa (trisacárido) para su conversión en maltosa y glucosa.

[0013] Hay por consiguiente necesidad en la industria de una identificación y optimización de amilasas ácidas que sean útiles para procesos comerciales de licuefacción de almidón de maíz. Estas amilasas ácidas de segunda generación ofrecerán unas características de fabricación y/o actuación mejoradas en comparación con las de las enzimas estándar que se usan en la industria.

[0014] Las publicaciones sobre las que aquí se trata se indican solamente en atención a su publicación antes de la fecha de presentación de la presente solicitud. Nada de lo aquí expuesto deberá interpretarse como una admisión de que la invención no tenga derecho a anticiparse a tal publicación en virtud de una invención anterior.

Breve Exposición de la Invención

[0015] La invención aporta un ácido nucleico aislado que tiene una secuencia como la expuesta en la ID SEC Nº: 1 y codifica polipéptidos que tienen actividad de alfa-amilasa.

[0016] Un aspecto de la invención es un ácido nucleico aislado que tiene una secuencia como la expuesta en la ID SEC Nº: 1, secuencias prácticamente idénticas a la misma, y secuencias complementarias de la misma.

[0017] Otro aspecto de la invención es un ácido nucleico aislado que codifica un polipéptido o un fragmento funcional del mismo que tiene una secuencia como la que se expone en la ID SEC Nº: 2 y secuencias prácticamente idénticas a la misma.

[0018] En aun otro aspecto, la invención aporta un polipéptido purificado que tienen una secuencia como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2, y secuencias prácticamente idénticas a la misma.

[0019] Otro aspecto de la invención es un anticuerpo aislado o purificado que se une específicamente a un polipéptido que tiene una secuencia como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2, y secuencias prácticamente idénticas a la misma.

[0020] Otro aspecto de la invención es un método para hacer un polipéptido que tiene una secuencia como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2, y secuencias prácticamente idénticas a la misma. El método incluye el paso de introducir un ácido nucleico que codifica al polipéptido en una célula huésped, en donde el ácido nucleico está operativamente ligado a un promotor, y el paso de cultivar la célula huésped bajo condiciones que permitan la expresión del ácido nucleico.

[0021] El método incluye los pasos de introducir un ácido nucleico que codifica al polipéptido en una célula huésped, en donde el ácido nucleico está operativamente ligado a un promotor, y cultivar la célula huésped bajo condiciones que permitan la expresión del ácido nucleico, produciendo con ello el polipéptido.

[0022] Otro aspecto de la invención es un método que es para generar una variante e incluye los pasos de obtener un ácido nucleico que tenga una secuencia como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1, secuencias prácticamente idénticas a la misma, secuencias complementarias de las secuencias de ácido nucleico de la ID SEC Nº: 1 y cambiar uno o varios nucleótidos en la secuencia por otro nucleótido, delecionar uno o varios nucleótidos en la secuencia, o añadir uno o varios nucleótidos a la secuencia.

[0023] Otro aspecto de la invención es un ensayo para identificar fragmentos o variantes de secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias prácticamente idénticas a la misma que conservan la función enzimática de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias prácticamente idénticas a las mismas. El ensayo incluye los pasos de poner al polipéptido de las secuencias de aminoácidos de la ID SEC Nº: 2, a secuencias prácticamente idénticas a las mismas o a un fragmento o una variante del polipéptido en contacto con una molécula de sustrato bajo condiciones que le permitan funcionar al fragmento o a la variante del polipéptido, y detectar ya sea una disminución del nivel de sustrato o bien un incremento del nivel del específico producto de reacción de la reacción entre el polipéptido y el sustrato, identificando con ello un fragmento o una variante de tales secuencias.

[0024] La invención también aporta un proceso que es para preparar una masa o un producto horneado preparado a partir de la masa y comprende el paso de añadir una amilasa de la invención a la masa en una cantidad que sea eficaz para retardar el enranciamiento del pan. La invención también aporta una masa que comprende dicha amilasa y una premezcla que comprende harina junto con dicha amilasa. Finalmente, la invención aporta un aditivo enzimático de horneo que contiene dicha amilasa.

[0025] El uso de la amilasa según la presente invención proporciona un mejorado efecto anti-enranciamiento, medido p. ej. en forma de una menor afirmación de la miga, del mantenimiento de la elasticidad de la miga, de una rebanabilidad mejorada (p. ej. con menor formación de migajas y con una miga no gomosa) y de un sabor o una apetibilidad mejorados.

Breve Descripción de los Dibujos

[0026] Los siguientes dibujos son ilustrativos de realizaciones de la invención y no pretenden limitar el alcance de la invención, que es el abarcado por las reivindicaciones.

La **Figura 1** es un diagrama de bloques de un sistema informático.
La **Figura 2** es un diagrama de flujo que ilustra una realización de un proceso para comparar una nueva secuencia nucleotídica o proteínica con una base de datos de secuencias a fin de determinar los niveles de homología entre la nueva secuencia y las secuencias de la base de datos.
La **Figura 3** es un diagrama de flujo que ilustra una realización de un proceso ejecutado en un ordenador para determinar si dos secuencias son homólogas.
La **Figura 4** es un diagrama de flujo que ilustra una realización de un proceso identificador 300 para detectar la presencia de una característica en una secuencia.
La **Figura 5** ilustra una Curva Estándar de muestra del ensayo del Ejemplo 2.
Las **Figuras 6a y 6b** muestran los fragmentos de peso molecular en jarabes que usan enzimas de la invención y enzimas comerciales.
La **Figura 7** es las secuencias de la invención.
La **Figura 8** muestra una comparación de perfiles de oligosacáridos para la ID SEC Nº: 2 y amilasas comerciales.

Descripción Detallada de la Invención

[0027] La presente invención se refiere a las alfa-amilasas y a los polinucleótidos que las codifican. En el sentido en el que se la utiliza en la presente, la expresión “alfa-amilasa” incluye a las enzimas que tienen actividad de alfa-amilasa, como son por ejemplo las enzimas que son capaces de hidrolizar almidón para su conversión en azúcares. A diferencia de muchas amilasas conocidas, la amilasa del ejemplo de la invención, que se expone en la ID SEC Nº: 2, no es una enzima calcio-dependiente.

[0028] Es altamente deseable poder hacer que disminuya la dependencia del Ca2+ de una alfa-amilasa. En consecuencia, un aspecto de la invención aporta una enzima amilasa que tiene una reducida dependencia del Ca2+ en comparación con las amilasas comerciales o progenitoras. La reducida dependencia del Ca2+ tendrá en general la consecuencia funcional de que la variante presenta una satisfactoria actividad amilolítica en presencia de una concentración de iones de calcio en el medio extraño más baja que la que es necesaria para una enzima comercial o progenitora. Dicha reducida dependencia del Ca2+ tendrá además a menudo la consecuencia de que la variante es menos sensible a las condiciones de depleción de los iones de calcio tales como las que se dan en medios que contienen agentes formadores de complejos de calcio (tales como ciertos formadores de detergente).

[0029] Los polinucleótidos de la invención han sido identificados como aquéllos que codifican polipéptidos que tienen actividad de alfa-amilasa. Un ejemplo de enzima alfa-amilasa de la invención se presenta en la ID SEC Nº: 2, y a dicha enzima se la denomina también aquí ID SEC Nº: 2. Tales amilasas de la invención son particularmente útiles en procesos de molienda de maíz en húmedo, detergentes, procesos de horneo, bebidas y campos petrolíferos (etanol combustible).

[0030] Las alteraciones de las propiedades que pueden lograrse en variantes de la invención son alteraciones p. ej. de la especificidad para el sustrato, de la fijación al sustrato, del modelo de disociación del sustrato, de la estabilidad térmica, del perfil de pH/actividad, del perfil de pH/estabilidad, tal como una incrementada estabilidad a bajos valores pH (como p. ej. a un pH < 6, y en particular a un pH < 5) o a altos valores pH (como p. ej. a un pH > 9), de la estabilidad frente a la oxidación, de la dependencia del Ca2+, de la actividad específica y de otras propiedades de interés. Por ejemplo, la alteración puede redundar en una variante que en comparación con la amilasa progenitora tenga una reducida dependencia del Ca2+ y/o un perfil de pH/actividad alterado.

[0031] La molienda de maíz en húmedo es un proceso que produce aceite de maíz, harina de gluten, pienso de gluten y almidón. Las amilasas de la invención, incluyendo la ID SEC Nº: 2, se usan en la licuefacción de almidón y se usa glucoamilasa en la sacarificación, produciendo glucosa. Las propiedades de las amilasas de la presente invención son singulares por cuanto que permiten la producción de jarabes licuados que pueden ser convertidos para presentar niveles de dextrosa más altos que los de un jarabe licuado por amilasa de Bacillus convencional. Como puede verse en las Figuras 6a-6b y en los Ejemplos, el perfil del peso molecular del almidón licuado producido por amilasas comerciales derivadas de Bacillus licheniformis y Bacillus stearothermophilus presenta una distribución bimodal con un pico primario al nivel de un MW (MW = peso molecular) de 1000-2000, que representa aproximadamente un 60% de la masa, y con un pico secundario al nivel de un MW de 30.000-40.000. Además, hay una considerable fracción al nivel de un MW de más de 100.000. Las amilasas de la invención presentan una homogénea distribución del peso molecular centrada al nivel de un peso molecular de 1000-2000, con menos de un 10% de la masa con un peso molecular de más de 25.000. Los oligosacáridos de peso molecular más alto no son totalmente convertidos en glucosa durante la sacarificación. En consecuencia, las amilasas comerciales contendrán menos dextrosa que los jarabes sacarificados del proceso de licuefacción realizado por las enzimas amilasa de la presente invención, tales como la representada en la ID SEC Nº: 2 y equivalentes funcionales de la misma.

[0032] Las maltodextrinas se utilizan en una amplia variedad de aplicaciones alimentarias y de recubrimiento. Las amilasas de fuentes archaeanas generan una extremadamente uniforme composición de maltodextrina (véase también Leveque et al., Enzyme and Microbial Technology 26:3-14, 2000). El uso de las amilasas de la invención para licuar almidón de maíz redunda en una uniforme composición de maltodextrina. La licuefacción puede ser llevada a cabo a un pH de aproximadamente 4,5-6,5, y preferiblemente a un pH de alrededor de 5,0, y a temperaturas de hasta 105 grados C o más.

[0033] Además de los beneficios demostrados en la sacarificación, los jarabes licuados pueden ser tratados con carbón, secados por pulverización y utilizados como aditivos alimentarios, espesantes, ingredientes de carga poco calóricos, agentes peliculígenos, etc. Se prevé (pero no ha sido aún probado) que las maltodextrinas con un homogéneo perfil del peso molecular presentarán ventajas de actuación en comparación con las maltodextrinas con una distribución bimodal producidas por las enzimas de *Baccillus* convencionales.

[0034] Los vocablos “licuefacción” o “licuar” significan un proceso por medio del cual el almidón es convertido en dextrinas de cadena más corta y menos viscosas. En general, este proceso supone una gelatinización del almidón que se lleva a cabo simultáneamente a o bien va seguida por la adición de alfa-amilasa. En los procesos comerciales se prefiere que el almidón granular se derive de una fuente que comprende a los miembros del grupo que consta de maíz, trigo, milo, sorgo, centeno o bulgur. Sin embargo, la presente invención se aplica a toda fuente de almidón de cereal que sea útil en licuefacción, como p. ej. cualquier otra fuente cereal o vegetal de la que se sepa que produce almidón adecuado para la licuefacción.

[0035] Las expresiones “almidón granular” o “gránulos de almidón” significan un componente hidroinsoluble de granos comestibles que queda después de la eliminación de la cáscara, de la fibra, de la proteína, de la grasa, del germen y de los solubles mediante los pasos de maceración, quebrantamiento mecánico, separación, tamizado, enjuague en contracorriente y centrifugación que son típicos del proceso de molienda de cereales en húmedo. El almidón granular comprende gránulos de almidón intactos que contienen casi exclusivamente moléculas de almidón empacadas (es

decir, amilopectina y amilosa). En el maíz, el componente de almidón granular comprende aproximadamente un 99% de almidón; constando el 1% restante de proteína, grasa, ceniza, fibra y componentes presentes a nivel de trazas y estrechamente asociados a los gránulos. La estructura de empaquetamiento del almidón granular retarda severamente la capacidad de la alfa-amilasa para hidrolizar el almidón. La gelatinización del almidón es utilizada para romper los gránulos para así formar una solución de almidón soluble y facilitar la hidrólisis enzimática.

[0036] La expresión "solución de almidón" significa el almidón gelatinizado hidrosoluble que se obtiene como resultado de calentar almidón granular. Tras un calentamiento de los gránulos hasta una temperatura de más de aproximadamente 72 grados C, el almidón granular se disocia para así formar una mezcla acuosa de moléculas de almidón sueltas. Esta mezcla, que comprende por ejemplo aproximadamente un 75% de amilopectina y un 25% de amilosa en el maíz dentado amarillo, forma una solución viscosa en agua. En procesos comerciales para formar glucosa o fructosa, es la solución de almidón la que es licuada para formar una solución de dextrina soluble. La expresión "actividad de alfa-amilasa" significa una actividad enzimática que parte o hidroliza el enlace glicosídico alfa-(1-4), como p. ej. el de los polímeros de almidón, amilopectina o amilosa. Son adecuadas alfa-amilasas las alfa-amilasas que se dan de manera natural, así como las amilasas recombinantes o mutantes que son útiles en la licuefacción de almidón. Se incluyen en la presente técnicas para producir amilasas variantes que tienen actividad a un pH o a una temperatura que son por ejemplo distintos de los que corresponden a la amilasa de tipo salvaje.

[0037] La gama de temperaturas de licuefacción es en general toda temperatura de licuefacción de la que se sepa que es eficaz en la licuefacción del almidón. Preferiblemente, la temperatura del almidón es de entre aproximadamente 80 grados C y aproximadamente 115 grados C, con mayor preferencia es de aproximadamente 100 grados C a aproximadamente 110 grados C, y con la máxima preferencia es de aproximadamente 105 grados C a aproximadamente 108 grados C.

[0038] En una realización, las secuencias señal de la invención son identificadas a continuación de la identificación de nuevos polipéptidos de amilasa. Las rutas por las cuales las proteínas son clasificadas y transportadas a su correcta ubicación celular se denominan a menudo rutas de direccionamiento de proteínas. Uno de los elementos más importantes en todos estos sistemas de direccionamiento es una corta secuencia de aminoácidos en el terminal amino de un polipéptido de nueva síntesis llamado la secuencia señal. Esta secuencia señal dirige a una proteína a su ubicación apropiada en la célula y es eliminada durante el transporte o cuando la proteína llega a su destino final. Las de la mayoría de las proteínas lisosomales, de membrana o secretadas tienen una secuencia señal amino-terminal que las marca para su translocación al lumen del retículo endoplásmico. Se han determinado más de 100 secuencias señal para proteínas de este grupo. Las secuencias varían en longitud de 13 a 36 residuos de aminoácidos. Son conocidos para los expertos en la materia varios métodos de reconocimiento de secuencias señal. En una realización, los péptidos son identificados por un método denominado SignalP. El método SignalP usa una red neural combinada que reconoce tanto a los péptidos señal como a sus sitios de corte. (Nielsen, H., Engelbrecht, J. Brunalk, S., von Heijne, G., "Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites", Protein Engineering, vol. 10, Nº 1, pp. 1-6 (1997)). Debe entenderse que algunas de las amilasas de la invención pueden no tener secuencias señal. Puede ser deseable incluir una secuencia de ácido nucleico que codifique una secuencia señal de una amilasa operativamente ligada a una secuencia de ácido nucleico de una amilasa distinta, o bien opcionalmente puede desearse una secuencia señal de una proteína no amilasa.

[0039] En el sentido en el que se las utiliza en la presente, las frases "ácido nucleico" o "secuencia de ácido nucleico" se refieren a un oligonucleótido, un nucleótido, un polinucleótido o un fragmento de cualquiera de éstos, a DNA o RNA de origen genómico o sintético que puede ser de una sola hebra o de doble hebra y puede representar una hebra sentido o antisentido, a ácido nucleico peptídico (PNA), o bien a cualquier material tipo DNA o tipo RNA, de origen natural o sintético.

[0040] Una "secuencia codificadora de" o una "secuencia nucleotídica que codifica" un particular polipéptido o proteína es una secuencia de ácido nucleico que es transcrita y traducida a un polipéptido o proteína al ser puesta bajo el control de apropiadas secuencias reguladoras.

[0041] El vocablo "gen" significa el segmento de DNA implicado en la producción de una cadena polipeptídica; e incluye a regiones que preceden y siguen a la región codificadora (conductora y remolque), así como, donde sea aplicable, secuencias interventoras (intrones) entre segmentos codificadores individuales (exones).

[0042] En el sentido en el que se las utiliza en la presente, las expresiones "aminoácido" o "secuencia de aminoácidos" se refieren a un oligopéptido, un péptido, un polipéptido o una secuencia proteica, o a un fragmento, una porción o una subunidad de cualquiera de éstos, y a moléculas sintéticas o que se dan de manera natural.

[0043] En el sentido en el que se le utiliza en la presente, el vocablo "polipéptido" se refiere a aminoácidos unidos entre sí por enlaces peptídicos o enlaces peptídicos modificados, es decir, isoésteres peptídicos, y pueden contener aminoácidos modificados distintos de los 20 aminoácidos codificados por genes. Los polipéptidos pueden ser modificados ya sea mediante procesos naturales, tales como el procesamiento postraslacional, o bien mediante técnicas de modificación química que son perfectamente conocidas en la técnica. Las modificaciones pueden producirse en cualquier sitio en el polipéptido, incluyendo la cadena principal peptídica, las cadenas laterales de aminoácidos y los

terminales amino o carboxilo. Se apreciará que el mismo tipo de modificación puede estar presente en iguales o variables grados en varios sitios en un polipéptido dado. También un polipéptido dado puede tener muchos tipos de modificaciones. Las modificaciones incluyen a los miembros del grupo que consta de acetilización, acilación, ribosilación de ADP, amidación, unión covalente de flavina, unión covalente de una mitad heme, unión covalente de un nucleótido o derivado nucleotídico, unión covalente de un lípido o derivado lipídico, unión covalente de un fosfitidilinositol, ciclización reticulante, formación de enlaces disulfuro, desmetilización, formación de enlaces intermoleculares covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicosilación, formación de anclajes GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, pegilación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfatación y adición mediada por RNA de transferencia de aminoácidos a proteína tal como arginilación. (Véanse Creighton, T.E., Proteins - Structure and Molecular Properties 2ª Ed., W.H. Freeman and Company, Nueva York (1993); Posttranslational Covalent Modification of Proteins, B.C. Johnson, redactor en jefe., Academic Press, Nueva York, pp. 1-12 (1983)).

[0044] En el sentido en el que se le utiliza en la presente, el vocablo “aislado” significa que el material es retirado de su ambiente original (como p. ej. el ambiente natural si se da de manera natural). Por ejemplo, un polinucleótido o polipéptido que se da de manera natural y está presente en un animal viviente no está aislado, pero el mismo polinucleótido o polipéptido separado de algunos de los materiales que coexisten en el sistema natural, o de todos ellos, está aislado. Tales polinucleótidos podrían ser parte de un vector y/o tales polinucleótidos o polipéptidos podrían ser parte de una composición, y seguirían estando aislados si tal vector o composición no es parte de su ambiente natural.

[0045] En el sentido en el que se le utiliza en la presente, el vocablo “purificado” no requiere una pureza absoluta; sino que más bien pretende ser una definición relativa. Los ácidos nucleicos individuales obtenidos de una biblioteca han sido purificados convencionalmente hasta una homogeneidad electroforética. Las secuencias obtenidas de estos clones no podrían ser obtenidas directamente ya sea de la biblioteca o bien de DNA humano total. Los ácidos nucleicos purificados de la invención han sido purificados a partir del resto del DNA genómico en el organismo al menos 104-106 veces. Sin embargo, el vocablo “purificado” también incluye a ácidos nucleicos que han sido purificados a partir del resto del DNA genómico o de otras secuencias en una biblioteca u otro ambiente en al menos un orden de magnitud, típicamente en dos o tres órdenes, y más típicamente en cuatro o cinco órdenes de magnitud.

[0046] En el sentido en el que se le utiliza en la presente, el vocablo “recombinante” significa que el ácido nucleico es adyacente a un ácido nucleico de “cadena principal” al cual no es adyacente en su ambiente natural. Adicionalmente, para estar “enriquecidos”, los ácidos nucleicos representarán un 5% o más del número de insertos de ácido nucleico en una población de moléculas de cadena principal de ácido nucleico. Las moléculas de cadena principal según la invención incluyen a ácidos nucleicos tales como vectores de expresión, ácidos nucleicos autorreplicantes, virus, ácidos nucleicos integrantes y otros vectores o ácidos nucleicos usados para mantener o manipular un inserto de ácido nucleico de interés. Típicamente, los ácidos nucleicos enriquecidos representan un 15% o más del número de insertos de ácido nucleico en la población de moléculas de cadena principal recombinante. Más típicamente, los ácidos nucleicos enriquecidos representan un 50% o más del número de insertos de ácido nucleico en la población de moléculas de cadena principal recombinante. En una realización, los ácidos nucleicos enriquecidos representan un 90% o más del número de insertos de ácido nucleico en la población de moléculas de cadena principal recombinante.

[0047] Las expresiones polipéptidos o proteínas “recombinantes” se refieren a polipéptidos o proteínas producidos por técnicas de DNA recombinante; es decir, producidos a partir de células transformadas por un constructo de DNA exógeno que codifica el polipéptido o la proteína deseado(a). Son polipéptidos o proteínas “sintéticos” los preparados mediante síntesis química. Pueden también usarse métodos de síntesis química de péptidos en fase sólida para sintetizar el polipéptido o los fragmentos de la invención. Tales métodos han sido conocidos en la técnica desde los comienzos de la década de 1960 (Merrifield, R. B., J. Am. Chem. Soc., 85:2149-2154, 1963) (véase también Stewart, J. J. y Young, J. D., Solid Phase Peptide Synthesis, 2ª Ed., Pierce Chemical Co., Rockford, III., pp. 11-12)) y han sido recientemente empleados en kits de diseño y síntesis de péptidos en laboratorio que están disponibles comercialmente (Cambridge Research Biochemicals). Tales kits de laboratorio que están comercialmente disponibles han utilizado en general las enseñanzas de H. M. Geysen et al., Proc. Natl. Acad. Sci., EE.UU., 81:3998 (1984) y permiten la sintetización de péptidos en las puntas de una multitud de “varillas” o “púas” que están todas ellas unidas a un único plato. Cuando se utiliza un sistema de este tipo, un plato de varillas o púas es invertido e insertado en un segundo plato de correspondientes pocillos o depósitos que contienen soluciones para unir o anclar un apropiado aminoácido a las puntas de las varillas o púas. Repitiendo un paso de proceso de este tipo, es decir, invirtiendo e insertando las puntas de las varillas y púas en apropiadas soluciones, los aminoácidos son incorporados a los péptidos deseados. Además está disponible una serie de sistemas de síntesis de péptidos FMOC con los que se cuenta. Por ejemplo, el ensamblaje de un polipéptido o fragmento puede ser realizado sobre un soporte sólido usando un sintetizador de péptidos automatizado Modelo 431 A de la Applied Biosystems, Inc. Tal equipo proporciona un fácil acceso a los péptidos de la invención, ya sea mediante síntesis directa o bien mediante síntesis de una serie de fragmentos que pueden ser acoplados usando otras técnicas conocidas.

[0048] Una secuencia promotora es “ligada operativamente a” una secuencia codificadora cuando la RNA polimerasa que inicia la transcripción en el promotor transcriba la secuencia codificadora en mRNA.

[0049] Los “plásmidos” se designan mediante una “p” minúscula precedida y/o seguida por letras mayúsculas y/o números. Los plásmidos de partida de los que aquí se trata están ya sea comercialmente disponibles, a disposición del público sin limitaciones, o bien pueden construirse a partir de plásmidos disponibles según procedimientos publicados. Además, plásmidos equivalentes a los aquí descritos son conocidos en la técnica y serán obvios para el experto en la materia.

[0050] La expresión “digestión” de DNA se refiere a la fragmentación catalítica del DNA con una enzima de restricción que actúa tan sólo en ciertas secuencias en el DNA. Las distintas enzimas de restricción que aquí se usan están disponibles comercialmente y sus condiciones de reacción, cofactores y otros requisitos fueron usados como sería sabido para el experto en la materia. A efectos analíticos, típicamente se usa 1 µg de plásmido o fragmento de DNA con aproximadamente 2 unidades de enzima en aproximadamente 20 µl de solución tampón. A efectos de aislar los fragmentos de DNA para la construcción de plásmidos, típicamente se digieren de 5 a 50 µg de DNA con 20 a 250 unidades de enzima en un mayor volumen. Los apropiados tampones y las apropiadas cantidades de sustrato para las determinadas enzimas de restricción en particular están especificados por el fabricante. Se usan de ordinario unos tiempos de incubación de aproximadamente 1 hora a 37ºC, pero los mismos pueden variar según las instrucciones del proveedor. Tras la digestión, puede llevarse a cabo electroforesis en gel para aislar el fragmento deseado.

[0051] El vocablo “oligonucleótido” se refiere ya sea a un polidesoxinucleótido de una sola hebra o bien a dos hebras de polidesoxinucleótido complementarias que pueden ser sintetizadas químicamente. Tales oligonucleótidos sintéticos no tienen fosfato del lado 5’, y por consiguiente no se ligarán a otro oligonucleótido sin añadir un fosfato con un ATP (ATP = trifosfato de adenosina) en presencia de una quinasa. Un oligonucleótido sintético se ligará a un fragmento que no haya sido desfosforilado.

[0052] La frase “sustancialmente idénticos” en el contexto de dos ácidos nucleicos o polipéptidos se refiere a dos o más secuencias que tienen al menos una identidad de nucleótidos o de residuos de aminoácidos de un 90-95%, al ser comparados y alineados para una máxima correspondencia, según medición efectuada usando uno de los algoritmos de comparación de secuencias conocidos o bien mediante inspección visual. En algunas realizaciones, las secuencias son sustancialmente idénticas dentro de toda la longitud de las regiones codificantes.

[0053] Adicionalmente, una secuencia de aminoácidos “sustancialmente idéntica” es una secuencia que se diferencia de una secuencia de referencia en una o varias sustituciones, deleciones o inserciones conservadoras o no conservadoras de aminoácidos, particularmente cuando tal sustitución se produce en un sitio que no es el sitio activo de la molécula, y siempre que el polipéptido conserve esencialmente sus propiedades funcionales. Una sustitución de aminoácido conservadora, por ejemplo, sustituye un aminoácido por otro de la misma clase (como p. ej. la sustitución de un aminoácido hidrofóbico tal como isoleucina, valina, leucina o metionina por otro, o la sustitución de un aminoácido polar por otro, tal como la sustitución de arginina por lisina, de ácido glutámico por ácido aspártico o de glutamina por asparagina). Por ejemplo pueden eliminarse uno o varios aminoácidos de un polipéptido alfa-amilasa, lo cual redundará en la modificación de la estructura del polipéptido sin alterar significativamente su actividad biológica. Por ejemplo pueden eliminarse aminoácidos amino- o carboxilo-terminales que no se requieran para la actividad biológica de alfa-amilasa. Las secuencias polipeptídicas modificadas de la invención pueden ser sometidas a ensayo para la determinación de su actividad biológica de alfa-amilasa por una serie de métodos, entre los que se incluye el de poner a la secuencia polipeptídica modificada en contacto con un sustrato para alfa-amilasa y determinar si el polipéptido modificado hace que disminuya la cantidad de sustrato específico en el ensayo o incrementa los bioproductos de la reacción enzimática de un polipéptido alfa-amilasa funcional con el sustrato.

[0054] En el sentido en el que el vocablo es utilizado en la presente, los “fragmentos” son una parte de una proteína que se da de manera natural y puede existir en al menos dos distintas conformaciones. Los fragmentos pueden tener la misma o sustancialmente la misma secuencia de aminoácidos como la proteína que se da de manera natural. La expresión “sustancialmente la misma” significa que una secuencia de aminoácidos es en gran medida pero no enteramente la misma, pero conserva al menos una actividad funcional de la secuencia a la cual es afín. Están también incluidos los fragmentos que tienen estructuras tridimensionales distintas de la de la proteína que se da de manera natural. Es un ejemplo de esto una molécula “proforma”, tal como una proproteína de baja actividad que pueda ser modificada por fragmentación para producir una enzima madura con una actividad significantemente más alta.

[0055] El vocablo “hibridación” se refiere al proceso por medio del cual una hebra de ácido nucleico se une a una hebra complementaria mediante apareamiento de bases. Las reacciones de hibridación pueden ser sensibles y selectivas, de tal manera que una determinada secuencia de interés puede ser identificada incluso en muestras en las cuales esté presente a bajas concentraciones. Pueden definirse condiciones adecuadamente rigurosas por ejemplo por medio de las concentraciones de sal o formamida en las soluciones de prehibridación y de hibridación, o bien por medio de la temperatura de hibridación, y dichas condiciones son perfectamente conocidas en la técnica. La rigurosidad puede ser en particular incrementada reduciendo la concentración de sal, incrementando la concentración de formamida o incrementando la temperatura de hibridación.

[0056] Por ejemplo, la hibridación bajo condiciones de alta rigurosidad podría producirse en formamida aproximadamente al 50% a una temperatura de aproximadamente 37ºC a 42ºC. La hibridación podría producirse bajo condiciones de reducida rigurosidad en formamida a aproximadamente una concentración de un 35% a un 254% a una

temperatura de aproximadamente 30ºC a 35ºC. En particular, la hibridación podría producirse bajo condiciones de alta rigurosidad a 42ºC en formamida al 50%, 5 x SSPE, SDS al 0,3% y 200 n/ml de DNA de esperma de salmón cizallado y desnaturalizado. La hibridación podría producirse bajo condiciones de reducida rigurosidad como las descritas anteriormente, pero en formamida al 35% y a una temperatura reducida de 35ºC. La gama de temperaturas que corresponda a un determinado nivel de rigurosidad puede estrecharse adicionalmente calculando la relación de purina a pirimidina del ácido nucleico de interés y ajustando en consecuencia la temperatura. Son perfectamente conocidas en la técnica variaciones de las gamas de valores y de las condiciones anteriormente mencionadas.

[0057] El vocablo "variante" se refiere a polinucleótidos o polipéptidos de la invención modificados en uno o varios pares de bases, codones, intrones, exones o residuos de aminoácidos (respectivamente), pero que sigan aún conservando la actividad biológica de una alfa-amilasa de la invención. Pueden producirse variantes por cualquier número de medios entre los que se incluyen métodos tales como, por ejemplo, la PCR (PCR = reacción en cadena de la polimerasa) sujeta a error, el reordenamiento, la mutagénesis dirigida por oligonucleótidos, la PCR de ensamblaje, la mutagénesis por PCR sexual, la mutagénesis *in vivo*, la mutagénesis en casete, la mutagénesis de ensamblaje recursivo, la mutagénesis de ensamblaje exponencial, la mutagénesis específica de sitio, el reensamblaje de genes, la GSSM (GSSM = mutagénesis a saturación en sitio génico) y cualquier combinación de dichos métodos. Se incluyen en la presente técnicas para producir amilasas variantes que tienen actividad a un pH o a una temperatura que es por ejemplo distinto(a) de los de la amilasa de tipo salvaje.

[0058] Las enzimas son catalizadores altamente selectivos. Su sello distintivo es su capacidad para catalizar reacciones con exquisitas estereo-, regio- y quimio-selectividades que no tienen paralelo en la química sintética convencional. Además, las enzimas son notablemente versátiles. Pueden ser adaptadas a medida para funcionar en solventes orgánicos, para operar a pHs extremos (como por ejemplo a altos pHs y a bajos pHs), a temperaturas extremas (como por ejemplo a altas temperaturas y a bajas temperaturas) y a extremos niveles de salinidad (como por ejemplo a alto nivel de salinidad y a bajo nivel de salinidad), y para catalizar reacciones con compuestos que no son estructuralmente afines a sus sustratos fisiológicos naturales.

[0059] Las enzimas son reactivas frente a una amplia gama de sustratos naturales y no naturales, permitiendo así la modificación de virtualmente cualquier compuesto orgánico favorito. Además, a diferencia de lo que sucede en el caso de los tradicionales catalizadores químicos, las enzimas son altamente enantioselectivas y regioselectivas. El alto grado de especificidad para grupos funcionales que presentan las enzimas permite hacer un seguimiento de cada reacción en una secuencia sintética que conduce a la obtención de un compuesto con nueva actividad. Las enzimas son también capaces de catalizar muchas reacciones diversas no relacionadas con su función fisiológica en la naturaleza. Por ejemplo, las peroxidasas catalizan la oxidación de fenoles por peróxido de hidrógeno. Las peroxidasas pueden también catalizar reacciones de hidroxilación que no están relacionadas con la función nativa de la enzima. Son otros ejemplos las proteasas que catalizan la descomposición de polipéptidos. En solución orgánica algunas proteasas pueden también acilar azúcares, la cual es una función que está relacionada con la función nativa de estas enzimas.

[0060] En un aspecto, la invención incluye un método que es para licuar una composición con contenido de almidón y comprende el paso de poner al almidón en contacto con un polipéptido de la invención (como p. ej. un polipéptido purificado seleccionado de entre los miembros del grupo que consta de polipéptidos que tienen una secuencia de aminoácidos seleccionada de entre los miembros del grupo que consta de la ID SEC Nº: 2 y secuencias complementarias de la ID SEC Nº: 2). En una realización preferida, el polipéptido está expuesto en la ID SEC Nº: 2. El almidón puede ser de un material seleccionado de entre los miembros del grupo que consta de arroz, arroz germinado, maíz, cebada, trigo, legumbres y batata. Se incluye en la presente y se describe en los ejemplos un jarabe de glucosa producido por el método de la invención. Un jarabe de este tipo puede ser un jarabe de maltosa, un jarabe de glucosa, o bien una combinación de los mismos. En particular, en los jarabes producidos usando las amilasas de la invención hay un más alto nivel de fracción de DP2 y un más alto nivel de DP3 (maltotriosa y/o panosa) y menor cantidad de los fragmentos mayores que DP7 en comparación con los jarabes producidos por enzimas comerciales (véase el Ejemplo 5). Esto es consistente con el perfil de licuefacción puesto que hay menor cantidad de los fragmentos grandes en los jarabes licuados según la invención (véase el Ejemplo 5).

[0061] La invención también aporta un método que es para eliminar de un material manchas con contenido de almidón y comprende el paso de poner al material en contacto con un polipéptido de la invención. En un aspecto, la invención aporta un método que es para lavar un objeto y comprende el paso de poner al objeto en contacto con un polipéptido de la invención bajo condiciones suficientes para el lavado. Un polipéptido de la invención puede incluirse como aditivo detergente, por ejemplo. La invención también incluye un método que es para el desencolado textil y comprende el paso de poner al textil en contacto con un polipéptido de la invención bajo condiciones suficientes para el desencolado.

[0062] La invención también aporta un método que es para reducir el enranciamiento de productos de panadería y comprende la adición de un polipéptido de la invención al producto de panadería, antes del horneo.

[0063] La invención también aporta 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 eficaz para mejorar la propiedades de las fibras. La invención incluye un método para el desentintado enzimático de pasta papelera reciclada, en donde el polipéptido es aplicado en una cantidad que es eficaz para un efectivo desentintado de la superficie de las fibras.

[0064] Cualesquiera de los métodos aquí descritos incluyen la posibilidad de adición de una segunda alfa-amilasa o de una beta-amilasa o de una combinación de las mismas. Son conocidas para los expertos en la materia amilasas comerciales u otras enzimas adecuadas para ser usadas en combinación con una enzima de la invención.

[0065] La invención también incluye a un método para incrementar el flujo de fluidos de producción de una formación subterránea eliminando un fluido perjudicial viscoso con contenido de almidón formado durante las operaciones de producción y que se encuentra dentro de la formación subterránea que rodea a una perforación de pozo completada, comprendiendo dicho método los pasos permitir que los fluidos de producción fluyan desde la perforación del pozo; reducir el flujo de fluidos de producción de la formación a niveles inferiores a los caudales previstos; preparar un fluido de tratamiento enzimático mezclando juntamente un fluido acuoso y un polipéptido de la invención; bombear el fluido de tratamiento enzimático a un punto deseado en el interior de la perforación del pozo; permitir que el tratamiento enzimático degrade el fluido perjudicial viscoso con contenido de almidón, con lo cual el fluido puede ser retirado de la formación subterránea y llevado a la superficie del pozo; y en donde el tratamiento enzimático es eficaz para atacar los enlaces alfa-glucosídicos del fluido con contenido de almidón.

[0066] La presente invención explota las singulares propiedades catalíticas de las enzimas. Mientras que el uso de biocatalizadores (es decir, enzimas purificadas o crudas y células no vivientes o vivientes) en las transformaciones químicas normalmente requiere la identificación de un determinado biocatalizador que reaccione con un específico compuesto de partida, la presente invención usa biocatalizadores seleccionados y condiciones de reacción que son específicos para grupos funcionales que están presentes en muchos compuestos de partida.

[0067] Cada biocatalizador es específico para un grupo funcional o para varios grupos funcionales afines, y puede reaccionar con muchos compuestos de partida que contengan este grupo funcional.

[0068] Las reacciones biocatalíticas producen una población de derivados de un único compuesto de partida. Estos derivados pueden ser sometidos a otra serie de reacciones biocatalíticas para así producir una segunda población de compuestos derivados. Pueden producirse miles de variaciones del compuesto original con cada iteración de derivatización biocatalítica.

[0069] Las enzimas reaccionan en sitios específicos de un compuesto de partida sin afectar al resto de la molécula, el cual es un proceso que es muy difícil de lograr usando métodos químicos tradicionales. Este alto grado de especificidad biocatalítica proporciona los medios para identificar un compuesto activo individual dentro de la biblioteca. La biblioteca está caracterizada por la serie de reacciones biocatalíticas usadas para producirla, lo cual constituye el llamado “historial biosintético”. Cribando la biblioteca con respecto a las actividades biológicas y haciendo un seguimiento del historial biosintético se identifica la específica secuencia de reacciones que produce el compuesto activo. Se repite la secuencia de reacciones y se determina la estructura del compuesto sintetizado. Este modo de identificación, a diferencia de otros enfoques de síntesis y de cribaje, no requiere tecnologías de inmovilización, y los compuestos pueden ser sintetizados y sometidos a ensayo libremente en solución usando virtualmente cualquier tipo de ensayo de cribaje. Es importante señalar que el alto grado de especificidad de las reacciones enzimáticas sobre grupos funcionales permite el “seguimiento” de las específicas reacciones enzimáticas que constituyen la biblioteca producida biocatalíticamente.

[0070] Muchos de los pasos procesales son llevados a cabo usando automatización robótica que permite la ejecución de muchos miles de reacciones biocatalíticas y ensayos de cribaje por día, además de asegurar un alto nivel de precisión y reproducibilidad. Como resultado de ello, puede producirse en cuestión de semanas una biblioteca de compuestos derivados que llevaría años producir usando métodos químicos corrientes. (Véase la PCT/US 94/09174 para adicionales enseñanzas sobre la modificación de moléculas, incluyendo a las pequeñas moléculas).

[0071] Puede usarse en el contexto de la invención un método no estocástico llamado reensamblaje de genes sintéticos que está algo relacionado con el reordenamiento estocástico, excepto que los bloques formadores de ácido nucleico no son reordenados o concatenados o quimerizados aleatoriamente, sino que son ensamblados no estocásticamente.

[0072] El método de reensamblaje de genes sintéticos no depende de la presencia de un alto nivel de homología entre los polinucleótidos a reordenar. El método puede usarse para generar no estocásticamente bibliotecas (o conjuntos) de moléculas de progenie que constan de más de 10^{100} distintas quimeras. Concebiblemente, el reensamblaje de genes sintéticos puede incluso usarse para generar bibliotecas que consten de más de 10^{1000} distintas quimeras de progenie.

[0073] Puede usarse en el contexto de la invención un método no estocástico para producir un conjunto de moléculas de ácido nucleico quimérico finalizado que tengan un orden de ensamblaje total que se elige por diseño, constando dicho método de los pasos de generar por diseño una pluralidad de específicos bloques constructores de ácido nucleico que tengan terminales ligables útiles y mutuamente compatibles, y ensamblar estos bloques constructores de ácido nucleico de forma tal que se logre el deseado orden de ensamblaje total.

[0074] Se considera que los terminales ligables mutuamente compatibles de los bloques constructores de ácido nucleico a ensamblar son “útiles” para este tipo de ensamblaje ordenado si permiten que los bloques constructores sean acoplados en órdenes predeterminados. Así, en un ejemplo, el orden de ensamblaje total en el cual los bloques

constructores de ácido nucleico pueden ser acoplados viene especificado por el diseño de los terminales ligables y, si debe usarse más de un paso de ensamblaje, entonces el orden de ensamblaje total en el cual pueden acoplarse los bloques constructores de ácido nucleico viene también especificado por el orden secuencial del (de los) paso(s) de ensamblaje. En un ejemplo, las unidades constructoras hibridadas son tratadas con una enzima, tal como una ligasa (como p. ej. la T4 DNA ligasa), para lograr el enlace covalente de las unidades constructoras.

[0075] En otro ejemplo, el diseño de bloques constructores de ácido nucleico se obtiene tras análisis de las secuencias de un conjunto de plantillas de ácido nucleico progenitor que sirven de base para producir un conjunto de progenie de moléculas de ácido nucleico quimérico finalizado. Estas plantillas de ácido nucleico progenitor sirven así de fuente de información secuencial que ayuda en el diseño de los bloques constructores de ácido nucleico que deben ser mutagenizados, es decir, quimerizados o reordenados.

[0076] En una ejemplificación, la invención usa la quimerización de una familia de genes afines y su familia codificada de productos afines. En una ejemplificación particular, los productos codificados son enzimas. Las alfa-amilasas de la presente invención pueden ser mutagenizadas según los métodos que aquí se describen.

[0077] Las secuencias de una pluralidad de plantillas de ácido nucleico progenitor (como p. ej. los polinucleótidos de las secuencias de ácido nucleico de la ID SEC Nº: 1) pueden ser alineadas a fin de seleccionar uno o varios puntos de demarcación, cuyos puntos de demarcación pueden estar situados en una zona de homología. Los puntos de demarcación pueden ser usados para delinear los límites de los bloques constructores de ácido nucleico a generar. Así, los puntos de demarcación identificados y seleccionados en las moléculas progenitoras sirven de potenciales puntos de quimerización en el ensamblaje de las moléculas de progenie.

[0078] Típicamente, un punto de demarcación útil es una zona de homología (que consta de al menos una base nucleotídica homóloga) compartida por al menos dos plantillas progenitoras, pero el punto de demarcación puede ser una zona de homología que sea compartida por al menos la mitad de las plantillas progenitoras, al menos dos tercios de las plantillas progenitoras, al menos tres cuartos de las plantillas progenitoras, y preferiblemente casi todas las plantillas progenitoras. Aun más preferiblemente, un punto de demarcación útil es una zona de homología que es compartida por todas las plantillas progenitoras.

[0079] En una realización, el proceso de reensamblaje de genes es ejecutado exhaustivamente a fin de generar una biblioteca exhaustiva. En otras palabras, en el conjunto de moléculas de ácido nucleico quimérico finalizado están representadas todas las posibles combinaciones ordenadas de los bloques constructores de ácido nucleico. Al mismo tiempo, es por diseño (o no estocástico) el orden de ensamblaje (es decir, el orden de ensamblaje de cada bloque constructor en la secuencia de 5' a 3 de cada ácido nucleico quimérico finalizado). Debido a la naturaleza no estocástica del método, queda reducida en gran medida la posibilidad de productos secundarios indeseados.

[0080] En otra realización, el método prevé que el proceso de reensamblaje de genes sea ejecutado sistemáticamente, por ejemplo para generar una biblioteca sistemáticamente compartimentalizada, con compartimentos que pueden ser cribados sistemáticamente, como p. ej. uno a uno. En otras palabras, el método prevé que, por medio del uso selectivo y juicioso de específicos bloques constructores de ácido nucleico, combinado con el uso selectivo y juicioso de reacciones de ensamblaje secuencialmente escalonadas, puede lograrse un diseño experimental donde se hagan específicos conjuntos de productos de progenie en cada uno de varios recipientes de reacción. Esto permite llevar a cabo un procedimiento de examen y cribaje sistemático. Así, esto permite examinar sistemáticamente en grupos más pequeños un número potencialmente muy grande de moléculas de progenie.

[0081] Debido a su capacidad para realizar quimerizaciones de una manera que es altamente flexible pero asimismo exhaustiva y sistemática, particularmente cuando hay un bajo nivel de homología entre las moléculas progenitoras, el método permite la generación de una biblioteca (o de un conjunto) que consta de un gran número de moléculas de progenie. Debido a la naturaleza no estocástica del presente método de reensamblaje de genes, las moléculas de progenie generadas preferiblemente comprenden una biblioteca de moléculas de ácido nucleico quimérico finalizado que tienen un orden de ensamblaje total que es elegido por diseño. En una realización en particular, una biblioteca generada de este tipo consta de un número de distintas especies moleculares de progenie que va desde más de 10^3 hasta más de 10^{1000}.

[0082] En un ejemplo, un conjunto de moléculas de ácido nucleico quimérico finalizado producidas como se ha descrito consta de un polinucleótido que codifica un polipéptido. Según una realización, este polinucleótido es un gen que puede ser un gen artificial. Según otra realización, este polinucleótido es una ruta génica, que puede ser una ruta génica artificial. Uno o varios genes artificiales generados puede(n) ser incorporado(s) a una ruta génica artificial, tal como una ruta susceptible de operar en un organismo eucariótico (incluyendo una planta).

[0083] En otra ejemplificación, la naturaleza sintética del paso en el cual son generados los bloques constructores permite el diseño y la introducción de nucleótidos (como p. ej. uno o varios nucleótidos que pueden ser, por ejemplo, codones o intrones o secuencias reguladoras) que pueden más tarde ser opcionalmente eliminados en un proceso *in vitro* (p. ej. por mutagénesis) o en un proceso *in vivo* (p. ej. utilizando la capacidad de corte y empalme de genes de un

organismo huésped). Se aprecia que en muchos casos la introducción de estos nucleótidos puede también ser deseable por muchas otras razones además del potencial beneficio de crear un punto de demarcación útil.

[0084] Así, según otro ejemplo, puede usarse un bloque constructor de ácido nucleico para introducir un intrón. Así, pueden introducirse intrones funcionales en un gen artificial de la invención. La invención también prevé que puedan ser introducidos intrones funcionales en una ruta génica artificial de la invención. En consecuencia, puede generarse un polinucleótido quimérico que sea un gen artificial que contenga uno o varios intrones introducidos artificialmente.

[0085] En consecuencia, puede generarse un polinucleótido quimérico que sea una ruta génica artificial que contenga uno o varios de los intrones introducidos artificialmente. Preferiblemente, el (los) intrón (intrones) que se introduce(n) artificialmente es (son) funcional(es) en una o varias células huésped para el corte y empalme de genes muy de la misma manera como los intrones que se dan de manera natural sirven funcionalmente en el corte y empalme de genes. Puede usarse en el contexto de la invención un proceso para producir polinucleótidos artificiales con contenido de intrones a introducir en organismos huésped para recombinación y/o corte y empalme.

[0086] Un gen artificial producido en el contexto de la invención puede también servir de sustrato para su recombinación con otro ácido nucleico. Análogamente, una ruta génica artificial producida en el contexto de la invención puede también servir de sustrato para su recombinación con otro ácido nucleico. En un caso preferido, la recombinación se ve facilitada por o se produce en zonas de homología entre el gen artificial con contenido de intrones y un ácido nucleico, que sirve de contraparte de recombinación. En un caso particularmente preferido, la contraparte de recombinación puede también ser un ácido nucleico generado mediante la invención, incluyendo a un gen artificial o una ruta génica artificial. La recombinación puede ser facilitada por o producirse en zonas de homología que existan en el intrón o en los varios intrones introducidos artificialmente en el gen artificial.

[0087] El método de reensamblaje de genes sintéticos que es útil en el contexto de la invención utiliza una pluralidad de bloques constructores de ácido nucleico, cada uno de los cuales preferiblemente tiene dos terminales ligables. Los dos terminales ligables de cada bloque constructor de ácido nucleico pueden ser dos terminales romos (es decir, que tengan cada uno un saliente de cero nucleótidos), o bien preferiblemente un terminal romo y un saliente, o aun más preferiblemente dos salientes.

[0088] Un saliente útil con esta finalidad puede ser un saliente del lado 3' o un saliente del lado 5'. Así, un bloque constructor de ácido nucleico puede tener un saliente del lado 3' o bien y como alternativa un saliente del lado 5' o bien y como alternativa dos salientes del lado 3' o bien y como alternativa dos salientes del lado 5'. El orden total en el cual los bloques constructores de ácido nucleico se ensamblan para formar una molécula de ácido nucleico quimérico finalizado se determina mediante diseño experimental a propósito y no es aleatorio.

[0089] Según un ejemplo preferido, un bloque constructor de ácido nucleico es generado mediante síntesis química de dos ácidos nucleicos de una sola hebra (también denominados oligos de una sola hebra) y poniéndolos en contacto para así permitirles hibridarse para formar un bloque constructor de ácido nucleico de doble hebra.

[0090] Un bloque constructor de ácido nucleico de doble hebra puede ser de tamaño variable. Los tamaños de estos bloques constructores pueden ser pequeños o grandes. Los tamaños preferidos para el bloque constructor van desde 1 par de bases (sin incluir salientes de tipo alguno) hasta 100.000 pares de bases (sin incluir salientes de tipo alguno). Se prevén también otras preferidas gamas de tamaños que tienen límites inferiores que van desde 1 bp (bp = par(es) de bases) hasta 10.000 bp (incluyendo todo valor entero entre ambos) y límites superiores que van desde 2 bp hasta 100.000 bp (incluyendo todo valor entero entre ambos).

[0091] Existen muchos métodos por medio de los cuales puede generarse un bloque constructor de ácido nucleico de doble hebra que sea útil para la invención; y dichos métodos son conocidos en la técnica y pueden ser fácilmente ejecutados por el experto en la materia.

[0092] Según un ejemplo, se genera un bloque constructor de ácido nucleico de doble hebra generando en primer lugar dos ácidos nucleicos de una sola hebra y permitiéndoles hibridarse para así formar un bloque constructor de ácido nucleico de doble hebra. Las dos hebras de un bloque constructor de ácido nucleico de doble hebra pueden ser complementarias en cada nucleótido aparte de cualesquiera que formen un saliente; no conteniendo así emparejamientos incorrectos, aparte de cualesquiera salientes. Según otro ejemplo, las dos hebras de un bloque constructor de ácido nucleico de doble hebra son complementarias en menos sitios que en cada nucleótido aparte de cualesquiera que formen un saliente. Así, según este ejemplo, un bloque constructor de ácido nucleico de doble hebra puede ser usado para introducir degeneración de codones. Preferiblemente, la degeneración de codones se introduce usando la mutagénesis a saturación en sitio que aquí se describe, usando uno o varios casetes N,N,G/T o bien y como alternativa usando uno o varios casetes N,N,N.

[0093] El método de recombinación *in vivo* que es útil en el contexto de la invención puede ser ejecutado a ciegas en un conjunto de híbridos desconocidos o alelos de un polinucleótido o de una secuencia específico(a). Sin embargo, no es necesario conocer la secuencia real de DNA o de RNA del polinucleótido específico.

[0094] El enfoque de usar recombinación dentro de una población mixta de genes puede ser útil para la generación de cualesquiera proteínas útiles, como por ejemplo interleuquina I, anticuerpos, tPA y hormona del crecimiento. Este enfoque puede usarse para generar proteínas que tengan una especificidad o actividad alterada. El enfoque puede también ser útil para la generación de secuencias de ácido nucleico híbrido, tales como por ejemplo regiones promotoras, intrones, exones, secuencias potenciadoras, regiones no traducidas 31 o regiones no traducidas 51 de genes. Así, este enfoque puede ser usado para generar genes que tengan incrementadas tasas de expresión. Este enfoque puede también ser útil en el estudio de secuencias de DNA repetitivo. Finalmente, este enfoque puede ser útil para mutar ribozimas o aptámeros.

[0095] En un ejemplo la invención que aquí se describe está dirigida al uso de ciclos repetidos de selección, recombinación y redistribución reductiva que permiten el desarrollo molecular dirigido de secuencias lineales altamente complejas, tales como DNA, RNA o proteínas mediante recombinación.

[0096] El reordenamiento *in vivo* de moléculas es útil para proporcionar variantes y puede ser llevado a cabo utilizando la propiedad natural de las células de recombinar multímeros. Mientras que la recombinación *in vivo* ha proporcionado la principal ruta natural hacia la diversidad molecular, la recombinación genética sigue siendo un proceso relativamente complejo que supone 1) el reconocimiento de homologías; 2) el corte de hebras, la invasión de hebras y pasos metabólicos que conducen a la producción de quiasma recombinante; y finalmente 3) la resolución de quiasma para su conversión en moléculas recombinadas discretas. La formación del quiasma requiere el reconocimiento de secuencias homólogas.

[0097] En otra realización, la invención usa un método para producir un polinucleótido híbrido a partir de al menos un primer polinucleótido y un segundo polinucleótido. El método puede usarse para producir un polinucleótido híbrido introduciendo en una adecuada célula huésped al menos un primer polinucleótido y un segundo polinucleótido que comparten al menos una región de homología secuencial parcial. Las regiones de homología secuencial parcial promueven procesos que redundan en una reorganización secuencial que produce un polinucleótido híbrido. En el sentido en el que se la utiliza en la presente, la expresión “polinucleótido híbrido” es toda secuencia nucleotídica que se obtiene como resultado del método de la presente invención y contiene secuencia de al menos dos secuencias polinucleotídicas originales. Tales polinucleótidos híbridos pueden resultar de eventos de recombinación intermolecular que promueven una integración secuencial entre moléculas de DNA. Por añadidura, tales polinucleótidos híbridos pueden resultar de procesos de redistribución reductiva intramolecular que utilizan secuencias repetidas para alterar una secuencia nucleotídica dentro de una molécula de DNA.

[0098] La invención usa unos medios para generar polinucleótidos híbridos que pueden codificar polipéptidos híbridos biológicamente activos (como p. ej. alfa-amilasas híbridas). En un aspecto, los polinucleótidos originales codifican polipéptidos biológicamente activos. El método que se usa en el contexto de la invención produce nuevos polipéptidos híbridos utilizando procesos celulares que integran la secuencia de los polinucleótidos originales de forma tal que el polinucleótido híbrido resultante codifica un polipéptido que presenta actividades derivadas de los polipéptidos biológicamente activos originales. Por ejemplo, los polinucleótidos originales pueden codificar una enzima determinada a partir de diferentes microorganismos. Una enzima codificada por un primer polinucleótido desde un organismo o una variante puede por ejemplo funcionar eficazmente bajo una determinada condición ambiental, como p. ej. la de alta salinidad. Una enzima codificada por un segundo polinucleótido desde un distinto organismo o variante puede funcionar eficazmente bajo una distinta condición ambiental, tal como la de temperaturas extremadamente altas. Un polinucleótido híbrido que contenga secuencias de los polinucleótidos originales primero y segundo puede codificar una enzima que presente características de ambas enzimas codificadas por los polinucleótidos originales. Así, la enzima codificada por el polinucleótido híbrido puede funcionar eficazmente bajo condiciones ambientales compartidas por cada una de las enzimas codificadas por los polinucleótidos primero y segundo, tales como p. ej. las de alta salinidad y de temperaturas extremas.

[0099] Las enzimas codificadas por los polinucleótidos de la invención incluyen a hidrolasas, tales como alfa-amilasas. Un polipéptido híbrido obtenido como resultado del método que se usa en el contexto de la invención puede presentar una actividad enzimática especializada que no presenten las enzimas originales. Por ejemplo, a continuación de la recombinación y/o de la redistribución reductiva de polinucleótidos que codifiquen actividades de hidrolasa, el polipéptido híbrido resultante codificado por un polinucleótido híbrido puede ser cribado con respecto a las actividades de hidrolasa especializadas que se obtienen de cada una de las enzimas originales, es decir, con respecto al tipo de enlace en el cual actúa la hidrolasa y a la temperatura a la cual funciona la hidrolasa. Así por ejemplo, la hidrolasa puede ser cribada para determinar aquellas funcionalidades químicas que distinguen a la hidrolasa híbrida de las hidrolasas originales, tales como: (a) enlaces amida (enlaces peptídicos), es decir, proteasas; (b) enlaces éster, es decir, amilasas y lipasas; (c) acetales, es decir, glicosidasas, y por ejemplo la temperatura, el pH o la concentración salina a los cuales funciona el polipéptido híbrido.

[0100] Pueden aislarse fuentes de los polinucleótidos originales de organismos individuales (“aislados”), de colecciones de organismos que hayan sido cultivados en medios definidos (“cultivos de enriquecimiento”), o de organismos no cultivados (“muestras ambientales”). El uso de un enfoque independiente del cultivo para obtener polinucleótidos que codifiquen nuevas bioactividades a partir de muestras ambientales es preferible en grado sumo puesto que permite acceder a recursos de biodiversidad no explotados.

[0101] Las “bibliotecas ambientales” se generan a partir de muestras ambientales y representan los genomas colectivos de organismos que se dan de manera natural y están archivados en vectores de clonación que pueden ser propagados en adecuados huéspedes procarióticos. Debido al hecho de que el DNA clonado es inicialmente extraído directamente de muestras ambientales, las bibliotecas no quedan limitadas a la pequeña fracción de procariotas que pueden cultivarse en cultivo puro. Adicionalmente, una normalización del DNA ambiental presente en estas muestras podría permitir una más igual representación del DNA de todas las especies presentes en la muestra original. Esto puede incrementar espectacularmente la eficacia de la operación de encontrar genes interesantes de entre los constituyentes secundarios de la muestra, que pueden estar subrepresentados en varios órdenes de magnitud en comparación con las especies dominantes.

[0102] Por ejemplo, las bibliotecas de genes generadas a partir de uno o varios microorganismos no cultivados son cribadas para determinar una actividad de interés. Las potenciales rutas que codifican moléculas bioactivas de interés son primeramente capturadas en células procarióticas en forma de bibliotecas de expresión génica. Los polinucleótidos que codifican actividades de interés son aislados de tales bibliotecas e introducidos en una célula huésped. La célula huésped es cultivada bajo condiciones que promueven la recombinación y/o la redistribución reductiva creando biomoléculas potencialmente activas con nuevas o acrecentadas actividades.

[0103] Los microorganismos a partir de los cuales puede prepararse el polinucleótido incluyen a los miembros del grupo que consta de microorganismos procarióticos, tales como Eubacterias y Archaebacterias, y microorganismos eucarióticos inferiores tales como hongos, algunas algas y protozoos. Los polinucleótidos pueden ser aislados de muestras ambientales, en cuyo caso el ácido nucleico puede ser recuperado sin cultivo de un organismo o recuperado a partir de uno o varios organismos cultivados. Tales microorganismos pueden ser extremófilos, tal como hipertermófilos, psicrófilos, psicrotrofos, halófilos, barófilos y acidófilos. Son particularmente preferidos los polinucleótidos que codifican enzimas aisladas a partir de microorganismos extremofílicos. Tales enzimas pueden funcionar a temperaturas de más de 100ºC en fuentes termales terrestres y en chimeneas termales en el fondo del mar, a temperaturas de menos de 0ºC en aguas árticas, en el ambiente salino saturado del Mar Muerto, a valores pH de alrededor de 0 en yacimientos hulleros y en fuentes geotermales ricas en azufre, o a valores pH de más de 11 en lodos de aguas cloacales. Por ejemplo, varias amilasas y lipasas clonadas y expresadas desde organismos extremofílicos presentan alta actividad dentro de toda una amplia gama de temperaturas y valores pH.

[0104] Los polinucleótidos seleccionados y aislados como se ha descrito anteriormente son introducidos en una adecuada célula huésped. Es una adecuada célula huésped toda célula que sea capaz de promover la recombinación y/o la redistribución reductiva. Los polinucleótidos seleccionados están ya preferiblemente en un vector que incluye apropiadas secuencias de control. La célula huésped puede ser una célula eucariótica superior, tal como una célula de mamífero, o bien una célula eucariótica inferior, tal como una célula de levadura, o bien y preferiblemente la célula huésped puede ser una célula procariótica, tal como una célula bacteriana. La introducción del constructo en la célula huésped puede efectuarse mediante transfección por fosfato cálcico o transfección mediada por DEAE-Dextrano (DEAE = dietilaminoetanol), o mediante electroporación (Davis et al., 1986).

[0105] Como ejemplos representativos de huéspedes apropiados, pueden mencionarse los siguientes: células bacterianas, tales como *E. coli, Streptomyces* y *Salmonella typhimurium*; células fungales, tales como levadura; células de insecto tales como Drosophila S2 y Spodoptera Sf9; células animales tales como CHO, COS o melanoma de Bowes; adenovirus y células vegetales. Se considera que a partir de las enseñanzas que aquí se dan está al alcance de los expertos en la materia la selección de un huésped apropiado.

[0106] Con referencias particulares a varios sistemas de cultivo de células de mamífero que pueden emplearse para expresar proteína recombinante, los ejemplos de sistemas de expresión de mamífero incluyen las líneas COS-7 de fibroblastos de riñón de mono, que se describen en “SV40-transformed simian cells support the replication of early SV40 mutants” (Gluzman, 1981), y otras líneas celulares capaces de expresar un vector compatible, como por ejemplo las líneas celulares C127, 3T3, CHO, HeLa y BHK. Los vectores de expresión de mamífero comprenderán un origen de replicación, un adecuado potenciador y promotor, y también cualesquiera necesarios sitios de unión de ribosomas, sitios de poliadenilación, sitios dadores y aceptores de corte y empalme, secuencias de terminación transcripcional y secuencias no transcritas que flanquean al extremo 5’. Secuencias de DNA derivadas de los sitios de poliadenilación y de corte y empalme de SV40 pueden ser usadas para proporcionar los requeridos elementos genéticos no transcritos.

[0107] Células huésped que contengan los polinucleótidos de interés pueden ser cultivadas en medios nutrientes convencionales modificados según sea apropiado para activar promotores, seleccionar transformantes o amplificar genes. Las condiciones de cultivo, tales como la temperaturas, el pH y cosas similares son las anteriormente usadas con la célula huésped seleccionada para expresión, y le resultarán obvias al experto en la materia. Los clones que se identifiquen como aquéllos que tengan la actividad enzimática especificada pueden ser luego secuenciados para identificar la secuencia polinucleotídica que codifique una enzima que tenga la actividad acrecentada.

[0108] Se imagina que el método usado en el contexto de la presente invención puede ser usado para generar nuevos polinucleótidos que codifiquen rutas bioquímicas a partir de uno o varios operones o agrupaciones de genes o partes de los mismos. Por ejemplo, las bacterias y muchos eucariotas tienen un mecanismo coordinado para regular genes cuyos

productos están implicados en procesos afines. Los genes son agrupados en estructuras denominadas “agrupaciones de genes” en un único cromosoma y son transcritos juntamente bajo el control de una única secuencia reguladora que incluye un único promotor que inicia la transcripción de toda la agrupación. Así, una agrupación de genes es un grupo de genes adyacentes que son idénticos o afines habitualmente con respecto a su función. Un ejemplo de una ruta bioquímica codificada por agrupaciones de genes son las poliquetidas. Las poliquetidas son moléculas que son una extremadamente rica fuente de bioactividades, e incluyen a antibióticos (tales como tetraciclinas y eritromicina), agentes anti-cáncer (daunomicina), inmunosupresores (FK506 y rapamicina) y productos veterinarios (monensina). Muchas poliquetidas (producidas por poliquétido sintasas) son valiosas como agentes terapéuticos. Las poliquétido sintasas son enzimas multifuncionales que catalizan la biosíntesis de una enorme variedad de cadenas de carbono que se diferencian en cuanto a la longitud y a los patrones de funcionalidad y ciclización. Los genes de poliquétido sintasas caen en agrupaciones de genes y al menos un tipo (denominado tipo I) de poliquétido sintasas tiene enzimas y genes de gran tamaño, complicando la manipulación genética y los estudios *in vitro* de estos genes/proteínas.

[0109] DNA de agrupación génica puede ser aislado de distintos organismos y ligado en vectores, y en particular en vectores que contengan secuencias reguladoras de expresión que puedan controlar y regular la producción de una proteína detectable o actividad de conjunto a fin a proteína desde las agrupaciones génicas ligadas. El uso de vectores que tengan una excepcionalmente gran capacidad para la introducción de DNA exógeno es particularmente apropiado para el uso con tales agrupaciones de genes, y los mismos se describen a modo de ejemplo en la presente para incluir el factor F (o factor de fertilidad) de E. coli. Este factor F de E. coli es un plásmido que afecta la transferencia a alta frecuencia de sí mismo durante la conjugación y es ideal para lograr y propagar de manera estable grandes fragmentos de DNA, tales como agrupaciones de genes de muestras microbianas mixtas. Una realización particularmente preferida es la de usar vectores de clonación denominados “fósmidos” o vectores cromosómicos artificiales bacterianos (BAC). Éstos se derivan del factor F de E. coli que es capaz de integrar de manera estable grandes segmentos de DNA genómico. Cuando se efectúa integración con DNA de una muestra ambiental no cultivada mixta, esto hace que sea posible lograr grandes fragmentos genómicos en forma de una “biblioteca de DNA ambiental” estable. Otro tipo de vector que es para usar en la presente invención es un vector cósmido. Los vectores cósmidos fueron originalmente diseñados para clonar y propagar grandes segmentos de DNA genómico. La clonación en vectores cósmidos está descrita en detalle en Sambrook et al., Molecular Cloning: A Laboratory Manual, 2ª Ed., Cold Spring Harbor Laboratory Press (1989). Una vez ligados en un vector apropiado, pueden introducirse en una adecuada célula huésped dos o más vectores que contengan distintas agrupaciones de genes de poliquétido sintasas. Las regiones de homología secuencial parcial compartidas por las agrupaciones de genes promoverán procesos que redundan en una reorganización secuencial que redunda en una agrupación de genes híbridos. La nueva agrupación de genes híbridos puede luego ser cribada para localizar actividades acrecentadas que no se encuentran en las agrupaciones de genes originales.

[0110] Son conocidos para los expertos en la materia y se exponen en toda la presente memoria descriptiva métodos de cribaje para la localización de varias actividades enzimáticas. Tales métodos pueden ser empleados al aislar los polipéptidos y polinucleótidos de la invención.

[0111] Como ejemplos representativos de vectores de expresión que pueden ser usados, pueden mencionarse partículas virales, baculovirus, fagos, plásmidos, fagémidos, cósmidos, fósmidos, cromosomas artificiales bacterianos, DNA viral (como p. ej. vacunas, adenovirus, virus de la viruela aviar, pseudorrabia y derivados de SV40), cromosomas artificiales basados en P1, plásmidos de levadura, cromosomas artificiales de levadura y cualesquiera otros vectores específicos de específicos huéspedes de interés (tales como bacillus, aspergillus y levadura). Así por ejemplo, el DNA puede estar incluido en cualquiera de los de una variedad de vectores de expresión para expresar un polipéptido. Tales vectores incluyen secuencias de DNA cromosomal, no cromosomal y sintético. Son conocidos para los expertos en la materia y están disponibles comercialmente grandes cantidades de vectores adecuados. Se indican los siguientes vectores a modo de ejemplo: bacterianos: vectores pQE (Qiagen), plásmidos pBluescript, vectores pNH, vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucarióticos: pXT1, pSEG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Sin embargo puede usarse cualquier otro plásmido u otro vector, siempre que los mismos sean replicables y viables en el huésped. Pueden emplearse con la presente invención vectores de bajo número de copias o de alto número de copias.

[0112] La secuencia de DNA del vector de expresión es ligada operativamente a una apropiada secuencia (promotor) de control de expresión para dirigir la síntesis de RNA. Particulares promotores bacterianos que se nombran incluyen a los miembros del grupo que consta de lacl, lacZ, T3, T7, gpt, lambda PR, PL y trp. Los promotores eucarióticos incluyen a los miembros del grupo que consta de CMV inmediato temprano, HSV timidina quinasa, SV40 temprano y tardío, LTRs de retrovirus y metalotioneína-I de ratón. La selección del apropiado vector y promotor queda perfectamente dentro del nivel de conocimientos de los expertos en la materia. El vector de expresión también contiene un sitio de unión de ribosomas para la iniciación de la traducción y un terminador de transcripción. El vector puede también incluir apropiadas secuencias para amplificar la expresión. Pueden seleccionarse regiones promotoras a partir de cualquier gen deseado usando vectores de cloranfenicol transferasa (CAT) u otros vectores con marcadores seleccionables. Por añadidura, los vectores de expresión preferiblemente contienen uno o varios genes marcadores seleccionables para proporcionar un rasgo fenotípico para la selección de células huésped transformadas, tal como la resistencia a la dihidrofolato reductasa o a la neomicina para el cultivo de células eucarióticas, o tal como la resistencia a la tetraciclina o a la ampicilina en E. coli.

[0113] La redistribución *in vivo* se centra en procesos "intermoleculares" que se denominan colectivamente "recombinación", que en las bacterias se considera en general como un fenómeno "dependiente de RecA". La invención puede basarse en procesos de recombinación de una célula huésped para recombinar y redistribuir secuencias, o en la capacidad de la célula para mediar procesos reductivos para reducir la complejidad de secuencias cuasi repetidas en la célula por deleción. Este proceso de "redistribución reductiva" se produce en virtud de un proceso "intramolecular" independiente de RecA.

[0114] Por consiguiente, pueden generarse por medio del proceso de redistribución reductiva nuevos polinucleótidos. El método supone la generación de constructos que contengan secuencias consecutivas (secuencias codificadoras originales), su inserción en un vector apropiado, y su subsiguiente introducción en una apropiada célula huésped. La redistribución de las identidades moleculares individuales se produce en virtud de procesos combinatorios entre las secuencias consecutivas en el constructo que posee regiones de homología, o entre unidades cuasi repetidas. El proceso de redistribución recombina y/o reduce la complejidad y la extensión de las secuencias repetidas, y redunda en la producción de una nueva especie molecular. Pueden aplicarse varios tratamientos para acrecentar la tasa de redistribución. Éstos podrían incluir el tratamiento con luz ultravioleta o con sustancias químicas dañadoras del DNA, y/o el uso de líneas de células huésped que presenten acrecentados niveles de "inestabilidad genética". Así, el proceso de redistribución puede suponer una recombinación homóloga o la propiedad natural de las secuencias cuasi repetidas para dirigir su propia evolución.

[0115] Las secuencias repetidas o "cuasi repetidas" desempeñan un papel en la estabilidad genética. En la presente invención, las "cuasi repeticiones" son repeticiones que no quedan limitadas a su estructura unitaria original. Las unidades cuasi repetidas pueden presentarse como un conjunto de secuencias en un constructo; como unidades consecutivas de secuencias similares. Una vez ligadas, las uniones entre las secuencias consecutivas devienen esencialmente invisibles y la naturaleza cuasi repetitiva del constructo resultante es ahora continua a nivel molecular. El proceso de deleción que la célula lleva a cabo para reducir la complejidad del constructo resultante opera entre las secuencias cuasi repetidas. Las unidades cuasi repetidas proporcionan un prácticamente ilimitado repertorio de plantillas sobre las cuales pueden producirse eventos de deslizamiento. Los constructos que contienen las cuasi repeticiones proporcionan así efectivamente una elasticidad molecular suficiente como para que los eventos de deleción (y potencialmente de inserción) puedan producirse virtualmente en cualquier sitio dentro de las unidades cuasi repetitivas.

[0116] Cuando las secuencias cuasi repetidas están todas ellas ligadas en la misma orientación, como por ejemplo cabeza con cola o viceversa, la célula no puede distinguir unidades individuales. En consecuencia, el proceso reductivo puede producirse en todas las partes de las secuencias. En contraste con ello, cuando por ejemplo las unidades son presentadas cabeza con cabeza, en lugar de cabeza con cola, la inversión delinea los puntos finales de la unidad adyacente, con lo cual la formación de deleción favorecerá la pérdida de unidades discretas. Así, con el presente método es preferible que las secuencias estén en la misma orientación. Una orientación aleatoria de secuencias cuasi repetidas redundará en la pérdida de eficiencia de redistribución, mientras que una orientación consistente de las secuencias ofrecerá la más alta eficiencia. Sin embargo, mientras que el tener una menor cantidad de las secuencias contiguas en la misma orientación reduce la eficiencia, ello puede seguir proporcionando elasticidad suficiente para la efectiva recuperación de nuevas moléculas. Pueden hacerse constructos con las secuencias cuasi repetidas en la misma orientación para permitir la obtención de una más alta eficiencia.

[0117] Pueden ensamblarse secuencias en una orientación de cabeza con cola usando cualesquiera de los de una variedad de métodos, entre los que se incluyen los siguientes:

a) Pueden utilizarse cebadores que incluyan una cabeza de poli-A y una cola de poli-T que al hacerse de una sola hebra proporcionarían orientación. Esto se lleva a cabo haciendo que las pocas primeras bases de los cebadores estén hechas de RNA y puedan ser por consiguiente fácilmente eliminadas por RNAsa H.
b) Pueden utilizarse cebadores que incluyan sitios singulares de corte de restricción. Se requerirían sitios múltiples, una batería de secuencias singulares y repetidos pasos de síntesis y ligación.
c) Las pocas bases interiores del cebador podrían ser tioladas y podría usarse una exonucleasa para producir moléculas correctamente provistas de colas.

[0118] La recuperación de las secuencias redistribuidas se basa en la identificación de vectores de clonación con un reducido índice repetitivo (RI). Las secuencias codificadoras redistribuidas pueden luego ser recuperadas por amplificación. Los productos son reclonados y expresados. La recuperación de vectores de clonación con reducido RI puede verse afectada por:

1) El uso de vectores que se mantengan de manera estable tan sólo cuando el constructo es reducido en su complejidad.
2) La recuperación física de vectores acortados por procedimientos físicos. En este caso, el vector de clonación sería recuperado usando procedimientos de aislamiento de plásmidos estándar y fraccionado en tamaño en un gel de agarosa o en una columna con corte de bajo peso molecular utilizando procedimientos estándar.
3) La recuperación de vectores que contengan genes interrumpidos que puedan ser seleccionados al disminuir el tamaño del inserto.

4) El uso de técnicas de selección directa con un vector de expresión y la apropiada selección.

[0119] Las secuencias codificadoras (como por ejemplo genes) de organismos afines pueden mostrar un alto grado de homología y codificar productos proteicos del todo diversos. Estos tipos de secuencias son particularmente útiles en la presente invención como cuasi repeticiones. Sin embargo, mientras que los ejemplos que se ilustran a continuación demuestran la redistribución de secuencias codificadoras originales casi idénticas (cuasi repeticiones), este proceso no queda limitado a tales repeticiones casi idénticas.

[0120] El ejemplo siguiente muestra un método que es útil para la invención. Se describen secuencias (cuasi repeticiones) de ácido nucleico codificante derivadas de tres (3) especies singulares. Cada secuencia codifica una proteína con un conjunto singular de propiedades. Cada una de las secuencias se diferencia en un solo par de bases o en unos pocos pares de bases en una posición singular de la secuencia. Las secuencias cuasi repetidas son amplificadas y ligadas por separado o colectivamente en forma de ensamblajes aleatorios, de forma tal que en la población de moléculas ligadas están disponibles todas las posibles permutaciones y combinaciones. El número de unidades cuasi repetidas puede ser controlado por medio de las condiciones de ensamblaje. El número medio de unidades cuasi repetidas en un constructo está definido como el índice repetitivo (RI).

[0121] Una vez formados, los constructos pueden ser o pueden no ser fraccionados en tamaño en un gel de agarosa según protocolos publicados, insertados en un vector de clonación y transfectados en una apropiada célula huésped. Las células son luego propagadas y se efectúa “redistribución reductiva”. La tasa del proceso de distribución reductiva puede ser estimulada mediante la introducción de daño del DNA, si se desea. Carece de importancia el hecho de si la reducción del RI es mediada por formación de deleción entre secuencias repetidas por medio de un mecanismo “intramolecular”, o mediada por eventos tipo recombinación por medio de mecanismos “intermoleculares”. El resultado final es una redistribución de las moléculas en forma de todas las combinaciones posibles.

[0122] Opcionalmente, el método comprende el paso adicional de cribar los miembros de la biblioteca del conjunto reordenado para identificar miembros reordenados individuales de la biblioteca que tengan la capacidad de unirse o de otra manera interactuar o catalizar una reacción determinada (p. ej. tal como el dominio catalítico de una enzima) con una macromolécula predeterminada, tal como por ejemplo un receptor proteináceo, un oligosacárido, un virón u otro predeterminado compuesto o estructura.

[0123] Los polipéptidos que se identifican a partir de tales biblioteca pueden ser usados a efectos terapéuticos, diagnósticos, de investigación y afines (como p. ej. catalizadores, solutos para incrementar la osmolaridad de una solución acuosa, y cosas similares), y/o pueden ser sometidos a uno o varios ciclos adicionales de reordenamiento y/o selección.

[0124] En otro ejemplo se contempla que antes de la recombinación o redistribución o durante las mismas los polinucleótidos generados por medio del método que se usa en el contexto de la invención pueden ser sometidos a agentes o procesos que promuevan la introducción de mutaciones en los polinucleótidos originales. La introducción de tales mutaciones incrementaría la diversidad de los polinucleótidos y polipéptidos híbridos resultantes codificados a partir de las mismas. Los agentes o procesos que promueven la mutagénesis pueden incluir, aunque sin carácter limitativo, a los miembros del grupo que consta de: (+)-CC-1065 o un análogo sintético tal como (+)-CC-1065-(N3-adenina (véase Sun y Hurley, (1992); un aducto de 4’-fluoro-4-aminobifenilo N-acetilado o desacetilado capaz de inhibir la síntesis de DNA (véase, por ejemplo, van de Poll et al. (1992)); o un aducto de 4-aminobifenilo N-acetilado o desacetilado capaz de inhibir la síntesis de DNA (véase también van de Poll et al. (1992), pp. 751-758); cromo trivalente, una sal de cromo trivalente, un aducto de DNA-hidrocarburo aromático policíclico (PAH) capaz de inhibir la replicación de DNA, tal como 7-bromometil-benz[a]antraceno (“BMA”), tris(2,3-dibromopropil)fosfato (“Tris-BP”), 1,2-dibromo-3-cloropropano (“DBCP”), 2-bromoacroleína (2BA), benzo[a]pireno-7,8-dihidrodiol-9-10-epóxido (“BPDE”), una sal halógena de platino(II), N-hidroxi-2-amino-3-metilimidazol[4,5-f]-quinolina (“N-hidroxi-IQ”) y N-hidroxi-2-amino-1-metil-6-fenilimidazo[4,5-f]-piridina (“N-hidroxi-PhIP”). Medios especialmente preferidos para enlentecer o detener la amplificación por PCR constan de los miembros del grupo que consta de luz UV, (+)-CC-1065 y (+)-CC-1065-(N3-adenina). Quedan particularmente incluidos en éstos medios que son aductos de DNA o polinucleótidos que comprenden los aductos de DNA de los polinucleótidos o del conjunto de polinucleótidos y pueden ser liberados o eliminados por medio de un proceso que incluye el paso de calentar la solución que comprende a los polinucleótidos antes de su adicional procesamiento.

[0125] En otro ejemplo puede usarse un método para producir proteínas recombinantes que tengan actividad biológica tratando una muestra que comprenda polinucleótidos plantilla de doble hebra que codifiquen una proteína de tipo salvaje bajo condiciones según la invención, que permiten la producción de polinucleótidos híbridos o redistribuidos.

[0126] Es también útil en el contexto de la invención el uso de cebadores codón específicos (que contengan una secuencia N,N,N degenerada) para introducir mutaciones puntuales en un polinucleótido, para así generar un conjunto de polipéptidos de progenie en el cual esté representada toda una gama completa de sustituciones de aminoácidos individuales en cada posición de aminoácido (mutagénesis a saturación en sitio génico (GSSM)). Los oligos que se usan constan contiguamente de una primera secuencia homóloga, una secuencia N,N,N degenerada, y preferible pero no necesariamente una segunda secuencia homóloga. Los productos translacionales de la progenie del lado de cola

obtenidos mediante el uso de tales oligos incluyen a todos los posibles cambios de aminoácidos en cada sitio de aminoácidos a lo largo del polipéptido, porque la degeneración de la secuencia N,N,N incluye codones para todos los 20 aminoácidos.

[0127] En un ejemplo se usa un oligo degenerado de este tipo (que consta de un casete N,N,N degenerado) para someter a cada codón original en una plantilla polinucleotídica progenitora a toda una gama de sustituciones de codones. En otro ejemplo se usan ya sea en el mismo oligo o no al menos dos casetes N,N,N degenerados para someter a al menos dos codones originales en una plantilla polinucleotídica progenitora a toda una gama de sustituciones de codones. Así, puede estar contenida en un oligo más de una secuencia N,N,N para introducir mutaciones de aminoácidos en más de un sitio. Las de esta pluralidad de secuencias N,N,N pueden ser directamente contiguas, o bien pueden estar separadas por una o varias secuencias nucleotídicas adicionales. En otro ejemplo, los oligos que son útiles para introducir adiciones y deleciones pueden ser usados ya sea en solitario o bien en combinación con los codones que contienen una secuencia N,N,N para introducir cualquier combinación o permutación de adiciones/deleciones y/o sustituciones de aminoácidos.

[0128] En una ejemplificación particular, es posible mutagenizar simultáneamente dos o más posiciones de aminoácidos contiguas usando un oligo que contenga tripletes N,N,N contiguos, es decir, una secuencia (N,N,N)n degenerada.

[0129] En otro ejemplo son útiles en el contexto de la invención casetes degenerados que tengan menos degeneración que la secuencia N,N,N. Por ejemplo, puede ser deseable en algunos casos usar (p. ej. en un oligo) una secuencia triplete degenerada que conste de tan sólo un N, donde dicho N puede estar en la primera, en la segunda o en la tercera posición del triplete. En las otras dos posiciones del triplete pueden usarse cualesquiera otras bases, incluyendo cualesquiera combinaciones y permutaciones de las mismas. Como alternativa, puede ser deseable en algunos casos usar (p. ej. en un oligo) una secuencia triplete N,N,N degenerada, una secuencia triplete N,N,G/T o una secuencia triplete N,N,G/C.

[0130] Se aprecia sin embargo que el uso de un triplete degenerado (tal como una secuencia triplete N,N,G/T o una secuencia triplete N,N,G/C) como se describe en la presente invención es ventajoso por varias razones. En un ejemplo esta invención usa unos medios para generar sistemáticamente y con bastante facilidad sustitución de toda la gama de posibles aminoácidos (para un total de 20 aminoácidos) en todas y cada una de las posiciones de aminoácidos de un polipéptido. Así, para un polipéptido de 100 aminoácidos la invención aporta una manera de generar sistemáticamente y con bastante facilidad 2000 especies distintas (es decir, 20 posibles aminoácidos por posición multiplicado por 100 posiciones de aminoácidos). Se aprecia que mediante el uso de un oligo que contenga una secuencia triplete N,N,G/C o N,N,G/T degenerada se aportan 32 secuencias individuales que codifican para 20 posibles aminoácidos. Así, en un recipiente de reacción en el cual una secuencia polinucleotídica progenitora es sometida a mutagénesis a saturación usando un oligo de este tipo se generan 32 distintos polinucleótidos de progenie que codifican 20 polipéptidos distintos. En contraste con ello, el uso de un oligo no degenerado en mutagénesis dirigida por sitio conduce a tan sólo un producto polipeptídico de progenie por cada recipiente de reacción.

[0131] Pueden también usarse en el contexto de la invención oligos no degenerados, que pueden ser opcionalmente usados en combinación con cebadores degenerados de los que se dan a conocer. Se aprecia que en algunas situaciones es ventajoso usar oligos no degenerados para generar específicas mutaciones puntuales en un polinucleótido de trabajo. Esto proporciona unos medios para generar específicas mutaciones de punto silencioso, mutaciones puntuales que conduzcan a correspondientes cambios de aminoácidos y mutaciones puntuales que ocasionen la generación de codones de parada y la correspondiente expresión de fragmentos polipeptídicos.

[0132] Así, cada recipiente de reacción de mutagénesis a saturación puede contener polinucleótidos que codifiquen al menos 20 moléculas polipeptídicas de progenie de forma tal que todos los 20 aminoácidos estén representados en la específica posición de aminoácidos que corresponde a la posición codón mutagenizada en el polinucleótido progenitor. Los polipéptidos de progenie degenerados 32 veces generados a partir de cada recipiente de reacción de mutagénesis a saturación pueden ser sometidos a amplificación clonal (siendo p. ej. clonados en un adecuado huésped E. coli usando un vector de expresión) y sometidos a cribaje de expresión. Cuando se identifica por cribaje a un polipéptido de progenie individual que presenta un favorable cambio de propiedad (es comparación con el polipéptido progenitor), el mismo puede ser secuenciado para identificar la correspondiente sustitución de aminoácidos favorable contenida en el mismo.

[0133] Se aprecia que tras haber mutagenizado todas y cada una de las posiciones de aminoácidos de un polipéptido progenitor usando mutagénesis a saturación como aquí se describe, pueden identificarse en más de una posición de aminoácidos favorables cambios de aminoácidos. Pueden ser generadas una o varias nuevas moléculas de progenie que contengan una combinación de la totalidad o parte de estas favorables sustituciones de aminoácidos. Por ejemplo, si se identifican en cada una de 3 posiciones de aminoácidos en un polipéptido 2 específicos cambios favorables de aminoácidos, las permutaciones incluyen 3 posibilidades en cada posición (sin cambio con respecto al aminoácido original, y cada uno de dos cambios favorables) y 3 posiciones. Así, hay en total 3 x 3 x 3 o 27 posibilidades, que incluyen 7 que fueron examinadas anteriormente - 6 mutaciones puntuales individuales (es decir, 2 en cada una de tres posiciones) y ausencia de cambio en toda posición.

[0134] Puede usarse mutagénesis a saturación en sitio junto con reordenamiento, quimerización, recombinación y otros procesos de mutagénesis, además del cribaje. En el contexto de esta invención pueden usarse de manera iterativa cualesquiera procesos de mutagénesis, incluyendo la mutagénesis a saturación. En una ejemplificación se usa en combinación con el cribaje el uso iterativo de cualesquiera procesos de mutagénesis.

[0135] Así, en una ejemplificación no limitativa esta invención usa mutagénesis a saturación en combinación con adicionales procesos de mutagénesis, tales como un proceso en el que dos o más polinucleótidos afines son introducidos en una adecuada célula huésped de forma tal que es generado un polinucleótido híbrido por recombinación y redistribución reductiva.

[0136] Además de llevar a cabo mutagénesis a lo largo de toda la secuencia de un gen, en el contexto de la presente invención puede usarse mutagénesis para sustituir cada y cualquier número de bases en una secuencia polinucleotídica, en donde el número de bases a mutagenizar es preferiblemente cualquier entero de 15 a 100.000. Así, en lugar de mutagenizar cada posición a lo largo de una molécula, puede someterse a mutagénesis todo número o un número discreto de bases (preferiblemente un subconjunto que totalice de 15 a 100.000). Preferiblemente se usa para mutagenizar cada posición o grupo de posiciones a lo largo de una secuencia polinucleotídica un nucleótido aparte. Un grupo de 3 posiciones a mutagenizar puede ser un codón. Las mutaciones son preferiblemente introducidas usando un cebador mutagénico que contenga un casete heterólogo, que también recibe el nombre de casete mutagénico. Los casetes preferidos pueden tener de 1 a 500 bases. Cada posición nucleotídica en tales casetes heterólogos puede ser N, A, C, G, T, A/C, A/G, A/T, C/G, C/T. G/T, C/GT, A/G/T, A/C/T, A/C/G o E, donde E es cualquier base que no sea A, C, G o T (puede llamarse a E oligo diseñador).

[0137] En un sentido general, la mutagénesis a saturación consta de mutagenizar un conjunto completo de casetes mutagénicos (en donde cada casete es de una longitud de preferiblemente poco más o menos 1-500 bases) en la definida secuencia polinucleotídica a mutagenizar (en donde la secuencia a mutagenizar es preferiblemente de una longitud de aproximadamente 15 a 100.000 bases). Así, se introduce en cada casete en mutagenizar un grupo de mutaciones (que van de 1 a 100 mutaciones). Una agrupación de mutaciones a introducir en un casete puede ser distinta de o igual a una segunda agrupación de mutaciones a introducir en un segundo casete durante la aplicación de un ciclo de mutagénesis a saturación. Tales agrupaciones están ejemplificaciones por deleciones, adiciones, agrupamientos de codones particulares y agrupamientos de casetes nucleotídicos particulares.

[0138] Las secuencias definidas a mutagenizar incluyen a los miembros del grupo que consta de un gen completo, una ruta, cDNA, un marco de lectura abierto entero (ORF), un represor/transactivador, potenciador o promotor entero, un origen de replicación, un intrón, un operador o cualquier grupo funcional polinucleotídico. Generalmente, a estos efectos una “secuencia definida” puede ser cualquier polinucleótido que sea una secuencia polinucleotídica de 15 bases, y secuencias polinucleotídicas de longitudes de entre 15 bases y 15.000 bases (esta invención menciona específicamente todo entero entre ambos números). Las consideraciones para elegir agrupaciones de codones incluyen los tipos de aminoácidos codificados por un casete mutagénico degenerado.

[0139] En una ejemplificación particularmente preferida de una agrupación de mutaciones que pueden introducirse en un casete mutagénico, esta invención prevé específicamente sustituciones de codones degenerados (usando oligos degenerados) que codifican para 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 y 20 aminoácidos en cada posición, y una biblioteca de polipéptidos codificados por los mismos.

[0140] En un aspecto, la presente invención se refiere a un ácido nucleico aislado, sintético o recombinante que comprende una secuencia seleccionada de entre los miembros del grupo que consta de:

(a) la ID SEC Nº: 1,
(b) que codifica un polipéptido que tiene una secuencia como la expuesta en la ID SEC Nº: 2, o un fragmento enzimáticamente activo que tiene actividad de amilasa de la ID SEC Nº: 2 de una longitud de al menos 75, 100 o 150 residuos de aminoácidos;
(c) que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y (I) tiene al menos una identidad secuencial de un 90% en toda la longitud de la ID SEC Nº: 1, o (II) tiene una longitud de al menos 200 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 1, o (III) tiene una longitud de al menos 150 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 97% con la ID SEC Nº: 1, o (IV) tiene una longitud de al menos 75 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 99% con la ID SEC Nº: 1;
(d) que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y (I) tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o (II) tiene un fragmento enzimáticamente activo de (I) y al menos una longitud de 150 residuos de aminoácidos y tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o un fragmento enzimáticamente activo de (I) una longitud de al menos 75 residuos de aminoácidos y tiene una identidad secuencial de al menos un 98% con la ID SEC Nº: 2; o bien
(e) secuencias plenamente complementarias de (a), (b), (c) o (d),
en donde la identidad secuencial se determina usando un algoritmo de comparación de secuencias que consta del algoritmo FASTA, versión 3.0t78, con los parámetros por defecto.

[0141] Un aspecto de la invención es un ácido nucleico aislado que comprende una de las secuencias de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, las secuencias complementarias de la misma, 75, 100, 150, 200, 300, 400 o 500 bases consecutivas de una de las secuencias de una secuencia de ácido nucleico de la ID SEC Nº: 1 (o las secuencias complementarias de la misma). Los ácidos nucleicos aislados pueden comprender DNA, incluyendo cDNA, DNA genómico y DNA sintético. El DNA puede ser de doble hebra o de una sola hebra, y si es de una sola hebra puede ser la hebra codificadora o la hebra no codificadora (antisentido). Como alternativa, los ácidos nucleicos aislados pueden comprender RNA.

[0142] Como se expone más detalladamente más adelante, los ácidos nucleicos aislados de una de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma pueden usarse para preparar uno de los polipéptidos de una secuencia de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma, o fragmentos enzimáticamente activos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2, y secuencias sustancialmente idénticas a las mismas.

[0143] En consecuencia, otro aspecto de la invención es un ácido nucleico aislado que codifica uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas, o fragmentos enzimáticamente activos que comprenden al menos 75, 100 o 150 aminoácidos consecutivos de uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2. Las secuencias codificadoras de estos ácidos nucleicos pueden ser idénticas a una de las secuencias codificadoras de uno de los ácidos nucleicos de las secuencias de ácido nucleico de la ID SEC Nº: 1, o un fragmento de las mismas, o bien pueden ser distintas secuencias codificadoras que codifiquen uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2, secuencias sustancialmente idénticas a las mismas, y fragmentos que tengan al menos 75, 100 o 150 aminoácidos consecutivos de uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2, como resultado de la redundancia o degeneración del código genético. El código genético es perfectamente conocido para los expertos en la materia y puede obtenerse, por ejemplo, en la página 241 de B. Lewin, Genes VI, Oxford University Press, 1997.

[0144] El ácido nucleico aislado que codifica uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas puede incluir, aunque sin carácter limitativo: tan sólo la secuencia codificadora de una de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias prácticamente idénticas a la misma, y adicionales secuencias codificadoras tales como secuencias conductoras o secuencias proproteína y secuencias no codificadoras, tales como intrones o secuencias no codificadoras del lado 5' y/o del lado 3' de la secuencia codificadora. Así, en el sentido en el que se la utiliza en la presente, la expresión "polinucleótido que codifica un polipéptido" incluye a un polinucleótido que incluye tan sólo la secuencia codificadora para el polipéptido así como un polinucleótido que incluye una adicional secuencia codificadora y/o no codificadora.

[0145] Como alternativa, las secuencias de ácido nucleico del grupo de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas pueden ser mutagenizadas usando técnicas convencionales, tales como la mutagénesis dirigida por sitio u otras técnicas con las que están familiarizados los expertos en la materia, para introducir cambios silenciosos en los polinucleótidos de las secuencias de ácido nucleico de la ID SEC Nº: 1, y secuencias sustancialmente idénticas a las mismas. En el sentido en el que se la utiliza en la presente, la expresión "cambios silenciosos" incluye, por ejemplo, cambios que no alteran la secuencia de aminoácidos codificada por el polinucleótido. Tales cambios pueden ser deseables a fin de incrementar el nivel del polipéptido producido por células huésped que contienen un vector que codifica el polipéptido introduciendo codones o pares de codones que se den frecuentemente en el organismo huésped.

[0146] La invención también se refiere a polinucleótidos que tienen cambios de nucleótidos que redundan en sustituciones, adiciones, deleciones, fusiones y truncaciones de aminoácidos en los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y en secuencias sustancialmente idénticas a las mismas. Tales cambios de nucleótidos pueden introducirse usando técnicas tales como la mutagénesis dirigida por sitio, la mutagénesis química aleatoria, la deleción por exonucleasa III y otras técnicas de DNA recombinante. Como alternativa, tales cambios de nucleótidos pueden ser variantes alélicas que se den de manera natural y que se aíslan identificando ácidos nucleicos que se hibriden específicamente con sondas que comprendan al menos 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 300, 400 o 500 bases consecutivas de una de las secuencias de los miembros del grupo que consta de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas (o las secuencias complementarias de las mismas) bajo condiciones de alta, moderada o baja rigurosidad como las que aquí se prevén.

[0147] Los ácidos nucleicos aislados de las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a las mismas, las secuencias complementarias de las mismas, o un fragmento que comprenda al menos 75, 100, 150, 200, 300, 400 o 500 bases consecutivas de una de las secuencias del grupo de secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a la misma, o las secuencias complementarias de la misma pueden también usarse como sondas para determinar si una muestra biológica, tal como una muestra de tierra, contiene un organismo que tenga una secuencia de ácido nucleico de la invención o un organismo del cual el ácido nucleico fue obtenido. En tales procedimientos se obtiene una muestra biológica que potencialmente albergue al organismo del cual fue aislado el ácido nucleico, y se obtienen ácidos

nucleicos de la muestra. Los ácidos nucleicos son puestos en contacto con la muestra bajo condiciones que permitan a la muestra hibridarse específicamente con cualesquiera secuencias complementarias de las que están aquí presentes.

[0148] Cuando sea necesario pueden determinarse condiciones que le permitan a la sonda hibridarse específicamente con secuencias complementarias poniendo a la sonda en contacto con secuencias complementarias de muestras de las que se sepa contienen la secuencia complementaria, así como con secuencias de control que no contengan la secuencia complementaria. Las condiciones de hibridación tales como la concentración salina del tampón de hibridación, la concentración de formamida del tampón de hibridación o la temperatura de hibridación pueden variarse para identificar condiciones que le permitan a la sonda hibridarse específicamente con ácidos nucleicos complementarios.

[0149] Si la muestra contiene el organismo del cual fue aislado el ácido nucleico, se detecta entonces la hibridación específica de la sonda. La hibridación puede ser detectada etiquetando la sonda con un agente detectable tal como un isótopo radiactivo, un colorante fluorescente o una enzima capaz de catalizar la formación de un producto detectable.

[0150] Los expertos en la materia están familiarizados con muchos métodos para usar las sondas etiquetadas para detectar la presencia de ácidos nucleicos complementarios en una muestra. Éstos incluyen a los miembros del grupo que consta de Southern Blots, Northern Blots, procedimientos de hibridación de colonias y blots de puntos. Se aportan protocolos para cada uno de estos procedimientos en Ausubel et al. Current Protocls in Molecular Biology, John Wiley & Sons, Inc. (1997) y en Sambrook et al., Molecular Cloning: A Laboratory Manual 2ª Ed., Cold Spring Harbor Laboratory Press (1989).

[0151] Como alternativa, puede usarse más de una sonda (al menos una de las que sean capaces de hibridarse específicamente con cualesquiera secuencias complementarias de las que estén presentes en la muestra de ácido nucleico) en una reacción de amplificación para determinar si la muestra contiene un organismo que contenga una secuencia de ácido nucleico de la invención (como p. ej. un organismo del cual fue aislado el ácido nucleico). Típicamente, las sondas comprenden oligonucleótidos. En una realización, la reacción de amplificación puede comprender una reacción PCR (PCR = reacción en cadena de la polimerasa). Se describen protocolos de PCR en Ausubel y Sambrook, supra. Como alternativa, la amplificación puede comprender una reacción en cadena de la ligasa, 3SR, o una reacción de desplazamiento de hebra. (Véanse Barany, F., "The Ligase Chain Reaction in a PCR World", PCR Methods and Applications 1:5-16, 1991; E. Fahy et al., "Self sustained Sequence Replication (3SR): An Isothermal Transcription-based Amplification System Alternative to PCR", PCR Methods and Applications 1:25-33, 1991; y Walker G.T. et al., "Strand Displacement Amplification - an Isothermal in vitro DNA Amplification Technique", Nucleic Acid Research 20:1691-1696, 1992). En tales procedimientos, los ácidos nucleicos de la muestra son puestos en contacto con las sondas, se lleva a cabo la reacción de amplificación, y se detecta todo producto de amplificación resultante. El producto de amplificación puede detectarse llevando a cabo electroforesis en gel de los productos de reacción y coloreando el gel con un intercalador tal como bromuro de etidio. Como alternativa, una o varias de las sondas pueden ser etiquetadas con un isótopo radiactivo y la presencia de un producto de amplificación radiactivo puede detectarse mediante autorradiografía tras electroforesis en gel.

[0152] Sondas derivadas de secuencias cercanas a los extremos de las secuencias del grupo de secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas pueden también usarse en procedimientos de desplazamiento sobre el cromosoma para identificar clones que contengan secuencias genómicas situadas junto a las secuencias del grupo de secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas. Tales métodos permiten el aislamiento de genes que codifiquen proteínas adicionales del organismo huésped.

[0153] Los ácidos nucleicos aislados de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas, las secuencias complementarias de las mismas, o un fragmento que comprenda al menos 75, 100, 150, 200, 300, 400 o 500 bases consecutivas de una de las secuencias del grupo de secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o las secuencias complementarias de la misma pueden usarse como sondas para identificar y aislar ácidos nucleicos afines. En algunas realizaciones, los ácidos nucleicos afines pueden ser cDNAs o bien DNAs genómicos de organismos distintos de aquél del cual fue aislado el ácido nucleico. Por ejemplo, los otros organismos pueden ser organismos afines. En tales procedimientos, se pone a una muestra de ácido nucleico en contacto con la sonda bajo condiciones que le permitan a la sonda hibridarse específicamente con secuencias afines. La hibridación de la sonda con ácidos nucleicos del organismo afín es entonces detectada usando cualesquiera de los métodos que se han descrito anteriormente.

[0154] En las reacciones de hibridación de ácidos nucleicos las condiciones que se usen para alcanzar un determinado nivel de rigurosidad variarán en dependencia de la naturaleza de los ácidos nucleicos que se hibriden. Por ejemplo pueden considerarse al seleccionar las condiciones de hibridación la longitud, el grado de complementariedad, la composición de la secuencia nucleotídica (como p. ej. el contenido de GC frente al contenido de AT) y el tipo de ácido nucleico (como p. ej. RNA frente a DNA) de las regiones de hibridación de los ácidos nucleicos. Una consideración adicional es la de si se inmoviliza a uno de los ácidos nucleicos por ejemplo sobre un filtro.

[0155] La hibridación puede realizarse bajo condiciones de baja rigurosidad, moderada rigurosidad o alta rigurosidad. Como ejemplo de hibridación de ácido nucleico, una membrana de polímero que contiene ácidos nucleicos desnaturalizados inmovilizados es primeramente prehibridada por espacio de 30 minutos a 45ºC en una solución que consta de NaCl 0,9M, NaH2PO4 50mM, pH 7,0, Na2EDTA 5,9mM, SDS al 0,5%, 10 x reactivo de Denhardt y 0,5 mg/ml de ácido polirriboadenílico. Se añaden entonces a la solución aproximadamente 2 x 107 cpm (actividad específica 4-9 x 108 cpm/ µg) de sonda oligonucleotídica marcada en el extremo con 32P. Tras 12 - 16 horas de incubación, la membrana es lavada por espacio de 30 minutos a temperatura ambiente en 1 x SET (NaCl 150mM, Tris hidrocloruro 20mM, pH 7,8, Na2EDTA 1mM) que contiene un 0,5% de SDS, efectuándose a continuación un lavado por espacio de 30 minutos en 1 x SET sin usar a Tm -10ºC para la sonda oligonucleotídica. La membrana es entonces expuesta contra una película autorradiográfica para la detección de señales de hibridación.

[0156] Variando la rigurosidad de las condiciones de hibridación que se usan para identificar ácidos nucleicos tales como cDNAs o DNAs genómicos que se hibridan con la sonda detectable, pueden identificarse y aislarse ácidos nucleicos que tengan distintos niveles de homología con la sonda. La rigurosidad puede variarse llevando a cabo la hibridación a temperaturas variables inferiores a las temperaturas de fusión de las sondas. La temperatura de fusión, o Tm, es la temperatura (bajo definidas condiciones de fuerza iónica y pH) a la cual el 50% de la secuencia diana se hibrida con una sonda perfectamente complementaria. Se seleccionan condiciones muy rigurosas de manera que sean iguales o aproximadamente 5ºC inferiores a la Tm para una sonda particular. La temperatura de fusión de la sonda puede calcularse usando las fórmulas siguientes:

[0157] Para muestras de entre 14 y 70 nucleótidos de longitud la temperatura de fusión (Tm) se calcula usando la fórmula: Tm = 81,5 +16,6(log [Na+] + 0,41 (fracción G+C) - (600/N), donde N es la longitud de la sonda.

[0158] Si la hibridación se realiza en una solución que contenga formamida, la temperatura de fusión puede calcularse usando la ecuación: Tm = 81,5 + 16,6(log [Na+] + 0,41 (fracción G+C) - (0,63% de formamida) - (600/N), donde N es la longitud de la sonda.

[0159] La prehibridación puede realizarse en 6 x SSC, 5 x reactivo de Denhardt, SDS al 0,5%, 100 µg de DNA de esperma de salmón fragmentado desnaturalizado o 6 x SSC, 5 x reactivo de Denhardt, SDS al 0,5%, 100 µg de DNA de esperma de salmón fragmentado desnaturalizado, formamida al 50%. Las fórmulas para las soluciones de SSC y de Denhardt están enumeradas en Sambrook et al., supra.

[0160] La hibridación se lleva a cabo añadiendo la sonda detectable a las soluciones de prehibridación enumeradas anteriormente. Cuando la sonda comprenda DNA de doble hebra, se la desnaturaliza antes de la adición a la solución de hibridación. Se pone al filtro en contacto con la solución de hibridación por espacio de un periodo de tiempo suficiente para permitir que la sonda se hibride con cDNAs o DNAs genómicos que contengan secuencias complementarias de la misma u homólogas de la misma. Para sondas de una longitud de más de 200 nucleótidos, la hibridación puede realizarse a temperaturas de 15 a 25ºC más bajas que la Tm. Para sondas más cortas, tales como sondas oligonucleotídicas, la hibridación puede llevarse a cabo a temperaturas de 5 a 10ºC más bajas que la Tm. Típicamente, para hibridaciones en 6 x SSC, la hibridación es llevada a cabo a aproximadamente 68ºC. Habitualmente, para hibridaciones en soluciones que contengan un 50% de formamida, la hibridación se lleva a cabo a aproximadamente 42ºC.

[0161] Se consideraría que todas las hibridaciones anteriormente indicadas están bajo condiciones de alta rigurosidad.

[0162] A continuación de la hibridación, se procede a lavar el filtro para eliminar toda sonda detectable no específicamente fijada. La rigurosidad que se use para lavar los filtros puede también variarse en dependencia de la naturaleza de los ácidos nucleicos que se hibriden, de la longitud de los ácidos nucleicos que se hibriden, del grado de complementariedad, de la composición de las secuencias de nucleótidos (como p. ej. el contenido de GC frente al contenido de AT) y del tipo de ácido nucleico (como p. ej. RNA frente a DNA). Son ejemplos de lavados en condiciones de rigurosidad progresivamente más alta los siguientes: 2 x SSC, SDS al 0,1% a temperatura ambiente por espacio de 15 minutos (baja rigurosidad); 0,1 x SSC, SDS al 0,5% a temperatura ambiente por espacio de un periodo tiempo de 30 minutos a 1 hora (rigurosidad moderada); 0,1 x SSC, SDS al 0,5% por espacio de 15 a 30 minutos a una temperatura situada entre la temperatura de hibridación y 68ºC (alta rigurosidad); y NaCl 0,15M por espacio de 15 minutos a 72ºC (muy alta rigurosidad). Puede llevarse a cabo un lavado final en condiciones de baja rigurosidad en 0,1 x SSC a temperatura ambiente. Los anteriores ejemplos son meramente ilustrativos de un conjunto de condiciones que pueden usarse para lavar los filtros. Un experto en la materia sabría que hay numerosas recetas para lavados en distintas condiciones de rigurosidad. Se dan a continuación algunos otros ejemplos.

[0163] Los ácidos nucleicos que se han hibridado con la sonda son identificados por autorradiografía o bien mediante otras técnicas convencionales.

[0164] El procedimiento anteriormente descrito puede ser modificado para identificar ácidos nucleicos que tengan decrecientes niveles de homología con la secuencia de la sonda. Por ejemplo, para obtener ácidos nucleicos de decreciente homología con la sonda detectable pueden usarse condiciones menos rigurosas. Por ejemplo, la temperatura de hibridación puede ser reducida en incrementos de 5ºC desde 68ºC hasta 42ºC en un tampón de

hibridación que tenga una concentración de Na+ aproximadamente 1M. A continuación de la hibridación, el filtro puede ser lavado con 2 x SSC, SDS al 0,5% a la temperatura de hibridación. Se considera que estas condiciones son condiciones "moderadas" por encima de 50ºC y condiciones "bajas" por debajo de 50ºC. Un ejemplo específico de condiciones de hibridación "moderadas" es el de cuando la hibridación anteriormente descrita es llevada a cabo a 55ºC. Un ejemplo específico de condiciones de hibridación de "baja rigurosidad" es el de cuando la hibridación anteriormente descrita es llevada a cabo a 45ºC.

[0165] Como alternativa, la hibridación puede realizarse en tampones tales como 6 x SSC que contengan formamida a una temperatura de 42ºC. En este caso, la concentración de formamida en el tampón de hibridación puede ser reducida en incrementos de un 5% desde un 50% hasta un 0% para identificar clones que tengan decrecientes niveles de homología con la sonda. A continuación de la hibridación, el filtro puede ser lavado con 6 x SSC, SDS al 0,5% a 50ºC. Se considera que estas condiciones son condiciones "moderadas" a una concentración de formamida de más de un 25% y condiciones "bajas" a una concentración de formamida de menos de un 25%. Un ejemplo específico de condiciones de hibridación "moderadas" es el de cuando la hibridación anteriormente descrita es llevada a cabo a una concentración de formamida de un 30%. Un ejemplo específico de condiciones de hibridación de "baja rigurosidad" es el de cuando la hibridación anteriormente descrita es llevada a cabo a una concentración de formamida de un 10%.

[0166] Por ejemplo, los métodos precedentes pueden usarse para aislar ácidos nucleicos que tengan una secuencia con una homología de al menos aproximadamente un 90%, al menos un 95%, al menos 90%, al menos un 85%, al menos un 80%, al menos un 70%, al menos un 65%, al menos un 60%, al menos un 55% o al menos un 50% con una secuencia de ácido nucleico seleccionada de entre los miembros del grupo que consta de una de las secuencias del grupo de secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o fragmentos que comprendan al menos aproximadamente 75, 100, 150, 200, 300, 400 o 500 bases consecutivas de la misma, y las secuencias complementarias de la misma. La homología puede medirse usando el algoritmo de alineación. Por ejemplo, los polinucleótidos homólogos pueden tener una secuencia codificadora que sea una variante alélica que se dé de manera natural de una de las secuencias codificadoras que aquí se describen. Tales variantes alélicas pueden tener una sustitución, deleción o adición de uno o varios nucleótidos en comparación con los ácidos nucleicos de las secuencias de ácido nucleico de la ID SEC Nº: 1 o las secuencias complementarias de las mismas.

[0167] Adicionalmente, los procedimientos anteriormente descritos pueden usarse para aislar ácidos nucleicos que codifiquen polipéptidos que tengan una homología de al menos aproximadamente un 99%, un 95%, al menos un 90%, al menos un 85%, al menos un 80%, al menos un 75%, al menos un 70%, al menos un 65%, al menos un 60%, al menos un 55% o al menos un 50% con un polipéptido que tenga la secuencia de una de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma, o fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de la misma según determinación efectuada usando un algoritmo de alineación de secuencias (p. ej. tal como el algoritmo FASTA versión 3.0t78 con los parámetros por defecto).

[0168] En otro aspecto la presente invención se refiere a un polipéptido aislado, sintético o recombinante seleccionado de entre los miembros del grupo que consta de:

(a) una secuencia de aminoácidos como la expuesta en la ID SEC Nº: 2 o un fragmento enzimáticamente activo de la ID SEC Nº: 2 que tenga actividad de amilasa de la ID SEC Nº: 2 de una longitud de al menos 75, 100 o 150 residuos de aminoácidos;
(b) una secuencia de aminoácidos codificada por un ácido nucleico de la reivindicación 1 (a) a 1 (d) que tenga actividad de amilasa de la ID SEC Nº: 2;
(c) una secuencia de aminoácidos (I) que tenga al menos una identidad secuencial de un 95% a todo lo largo de la ID SEC Nº: 2; o (II) que tenga un fragmento enzimáticamente activo de (I) una longitud de al menos 150 residuos de aminoácidos y tenga una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o un fragmento enzimáticamente activo de (I) una longitud de al menos 75 residuos de aminoácidos y que tenga una identidad secuencial de al menos un 98% con la ID SEC Nº: 2, en donde el polipéptido de (I) o (II) tiene una actividad de amilasa de la ID SEC Nº: 2;
en donde la identidad secuencial se determina usando un algoritmo de comparación de secuencias que consta del algoritmo FASTA, versión 3.0t7, con los parámetros por defecto.

[0169] Otro aspecto de la invención es un polipéptido aislado o purificado que comprende la secuencia de una de las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de la misma. Como se ha expuesto anteriormente, tales polipéptidos pueden ser obtenidos insertando un ácido nucleico que codifique al polipéptido en un vector de forma tal que la secuencia codificadora quede operativamente enlazada a una secuencia capaz de activar la expresión del polipéptido codificado en una adecuada célula huésped. Por ejemplo, el vector de expresión puede comprender un promotor, un sitio de unión de ribosomas para la iniciación de la traducción y un terminador de transcripción. El vector puede también incluir secuencias apropiadas para amplificar la expresión.

[0170] Los promotores que son adecuados para expresar el polipéptido o fragmento de mismo en bacterias incluyen a los promotores *E. coli lac* o *trp*, al promotor *lacI*, al promotor *lacZ*, al promotor *T3*, al promotor *T7*, al promotor *gpt*, al promotor *lambda* P_R, al promotor *lambda* P_L, promotores de operones que codifican enzimas glicolíticas tales como 3-

fosfoglicerato quinasa (PGK), y al promotor de fosfatasa ácida. Los promotores fungales incluyen al promotor de factor V. Los promotores eucarióticos incluyen al promotor temprano inmediato del CMV, al promotor de timidina quinasa del HSV, a promotores de shock térmico, al promotor temprano y tardío de SV40, a LTRs (LTRs = repeticiones terminales largas) de retrovirus, y al promotor de metalotioneína-I de ratón. Pueden también usarse otros promotores de los que se sepa que controlan la expresión de genes en células procarióticas o eucarióticas o sus virus.

[0171] Los vectores de expresión de mamífero pueden también comprender un origen de replicación, cualesquiera necesarios sitios de unión de ribosomas, un sitio de poliadenilación, sitios dadores y aceptores de corte y empalme, secuencias de terminación transcripcional y secuencias no transcritas que flanquean al extremo 5'. En algunas realizaciones pueden usarse secuencias de DNA derivadas de sitios de poliadenlización y corte y empalme de SV40 para proporcionar los requeridos elementos genéticos no transcritos.

[0172] Los vectores para expresar el polipéptido o fragmento del mismo en células eucarióticas pueden también contener potenciadores para incrementar los niveles de expresión. Los potenciadores son elementos de DNA que actúan en cis, tienen habitualmente una longitud de aproximadamente 10 a aproximadamente 300 bp (bp = pares de bases) y actúan en un promotor para incrementar su transcripción. Los ejemplos incluyen al potenciador de SV40 en el lado tardío del origen de replicación de los bp 100 a 270, el potenciador promotor temprano de citomegalovirus, el potenciador de polioma en el lado tardío del origen de replicación, y los potenciadores de adenovirus.

[0173] Adicionalmente, los vectores de expresión típicamente contienen uno o varios genes marcadores seleccionables para permitir la selección de células huésped que contengan el vector. Tales marcadores seleccionables incluyen a genes que codifican dihidrofolato reductosa o genes que confieren resistencia a la neomicina para cultivo de células eucarióticas, genes que confieren resistencia a la tetraciclina o a la ampicilina en E. coli, y al gen TRP1 de S. cerevisiae.

[0174] En algunas realizaciones, el ácido nucleico que codifica uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas, o fragmentos que comprenden al menos 75, 100 o 150 aminoácidos consecutivos de las mismas es ensamblado en fase apropiada con una secuencia conductora capaz de dirigir la secreción del polipéptido traducido o fragmento del mismo. Opcionalmente, el ácido nucleico puede codificar un polipéptido de fusión en el cual uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas, o fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas es fusionado con péptidos o polipéptidos heterólogos tales como péptidos N-terminales de identificación que imparten características deseadas, tales como una incrementada estabilidad o una purificación simplificada.

[0175] La apropiada secuencia de DNA puede ser insertada en el vector mediante los de una variedad de procedimientos. En general, la secuencia de DNA es ligada a la deseada posición en el vector a continuación de la digestión del inserto y del vector con apropiadas endonucleasas de restricción. Como alternativa pueden ligarse extremos romos tanto del inserto como del vector. Se da a conocer una variedad de técnicas de clonación en Ausubel et al. Current Protocols in Molecular Biology, John Wiley & Sons, Inc. 1997 y en Sambrook et al., Molecular Cloning: A Laboratory Manual 2ª Ed., Cold Spring Harbor Laboratory Press (1989). Se considera que están al alcance de los expertos en la materia tales procedimientos y otros.

[0176] El vector puede por ejemplo estar en forma de un plásmido, de una partícula viral o de un fago. Otros vectores incluyen a miembros del grupo que consta de secuencias de DNA cromosomal, no cromosomal y sintético, derivados de SV40; plásmidos bacterianos, DNA de fago, baculovirus, plásmidos de levadura, vectores derivados de combinaciones de plásmidos y DNA de fago, DNA viral tal como vacunas, adenovirus, virus de la viruela aviar y pseudorrabia. Los de una variedad de vectores de clonación y expresión destinados a ser usados con huéspedes procarióticos y eucarióticos están descritos por Sambrook et al., Molecular Cloning: A Laboratory Manual, 2ª Ed., Cold Spring Harbor, N.Y.,(1989).

[0177] Los vectores bacterianos particulares que pueden ser usados incluyen a los plásmidos que están disponibles comercialmente y comprenden elementos genéticos del perfectamente conocido vector de clonación pBR322 (ATCC 37017), prKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suecia), GEM1 (Promega Biotec, Madison, WI, EE.UU.), pQE70, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia), pKK232-8 y pCM7. Particulares vectores eucarióticos incluyen a los miembros del grupo que consta de pSV2CAT, pOG44, pXT1, pSG (Stratagene) pPSVK3, pBPV, pMSG y pSVL (Pharmacia). Sin embargo, puede usarse cualquier otro vector siempre que sea replicable y viable en la célula huésped.

[0178] La célula huésped puede ser cualquiera de las células huésped con las que están familiarizados los expertos en la materia, incluyendo a células procarióticas, células eucarióticas, células de mamífero, células de insecto o células vegetales. Como ejemplos representativos de huéspedes apropiados, pueden mencionarse los siguientes: células bacterianas tales como E. coli, Streptomyces, Bacillus subtilis, Salmonella typhimurium y varias especies dentro de los géneros Pseudomonas, Streptomyces y Staphylococcus, células fungales, tales como levadura, células de insecto tales como Drosophila S2 y Spodoptera Sf9, células animales tales como CHO, COS o melanoma de Bowes, y adenovirus. Está dentro de las capacidades de los expertos en la materia la selección de un huésped apropiado.

[0179] El vector puede ser introducido en las células huésped usando cualquiera de las de una variedad de técnicas entre las que se incluyen las de transformación, transfección, transducción, infección viral, pistolas génicas o transferencia génica mediada por Ti. Los métodos particulares incluyen a los miembros del grupo que consta de transfección por fosfato cálcico, transfección mediada por DEAE-Dextrano, lipofección o electroporación (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

[0180] Allí donde ello sea apropiado, las células huésped manipuladas pueden ser cultivadas en medios nutrientes convencionales modificados según sea apropiado para activar promotores, seleccionar transformantes o amplificar los genes de la invención. A continuación de la transformación de una adecuada cepa huésped y del cultivo de la cepa huésped hasta alcanzar una apropiada densidad de células, el promotor seleccionado puede ser inducido mediante medios apropiados (como p. ej. desplazamiento de temperatura o inducción química) y las células pueden ser cultivadas por espacio de un adicional periodo de tiempo para permitirles producir el deseado polipéptido o fragmento del mismo.

[0181] Las células son típicamente recolectadas por centrifugación y rotas por medios físicos o químicos, y el extracto crudo resultante es conservado para su adicional purificación. Las células microbianas que se empleen para la expresión de proteínas pueden ser rotas por cualquier método conveniente, entre los que se incluyen los de ciclación por congelación-descongelación, sonicación, ruptura mecánica o el uso de agentes lisantes de células. Tales métodos son perfectamente conocidos para los expertos en la materia. El polipéptido o fragmento del mismo expresado puede ser recuperado y purificado a partir de los cultivos de células recombinantes por métodos entre los que se incluyen los de precipitación en sulfato amónico o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrofóbica, cromatografía de afinidad, cromatografía en hidroxiapatita y cromatografía en lectina. Según sea necesario, pueden usarse pasos de replegamiento de proteínas para completar la configuración de polipéptido. Si se desea, puede emplearse para los pasos de purificación final cromatografía de líquidos de alta resolución (HPLC).

[0182] Pueden también emplearse para expresar proteína recombinante varios sistemas de cultivo de células de mamífero. Los ejemplos de sistemas de expresión de mamífero incluyen a las líneas COS-7 de fibroblastos de riñón de mono (descritas por Gluzman, Cell, 23:175, 1981) y otras líneas celulares capaces de expresar proteínas desde un vector compatible, tales como las líneas celulares C127, 3T3, CHO, HeLa y BHK.

[0183] Los constructos en las células huésped pueden usarse de manera convencional para producir el producto génico codificado por la secuencia recombinante. En dependencia del huésped que se emplee en un procedimiento de producción recombinante, los polipéptidos producidos por células huésped que contengan el vector pueden ser glicosilados o pueden ser no glicosilados. Los polipéptidos de la invención pueden incluir o pueden también no incluir un residuo de aminoácido de metionina inicial.

[0184] Como alternativa, los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas, o fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas pueden ser producidos sintéticamente mediante sintetizadores de péptidos convencionales. En otras realizaciones pueden emplearse fragmentos o partes de los polipéptidos para producir el correspondiente polipéptido de longitud completa mediante síntesis de péptidos; y por consiguiente los fragmentos pueden emplearse como intermedios para producir los polipéptidos de longitud completa.

[0185] Pueden también emplearse sistemas de traducción sin células para producir uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas, o fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas usando mRNAs transcritos desde un constructo de DNA que comprenda un promotor ligado operativamente a un ácido nucleico que codifique el polipéptido o fragmento del mismo. En algunas realizaciones, el constructo de DNA puede ser linealizado antes de llevar a cabo una reacción de transcripción *in vitro*. El mRNA transcrito es luego incubado con un apropiado extracto de traducción exento de células, tal como un extracto de reticulocitos de conejo, para producir el deseado polipéptido o fragmento del mismo.

[0186] La invención también se refiere a variantes de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y a secuencias sustancialmente idénticas a las mismas, o a fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas. El vocablo “variante” incluye a derivados o análogos de estos polipéptidos. En particular, las variantes pueden diferenciarse en la secuencia de aminoácidos de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y de secuencias sustancialmente idénticas a las mismas, en una o varias sustituciones, adiciones, deleciones, fusiones y truncaciones, que pueden estar presentes en cualquier combinación.

[0187] Las variantes pueden darse de manera natural o bien pueden crearse *in vitro*. En particular, tales variantes pueden ser creadas usando técnicas de manipulación genética tales como la mutagénesis dirigida por sitio, la mutagénesis química aleatoria, los procedimientos de deleción por Exonucleasa III y técnicas de clonación estándar. Como alternativa, tales variantes, fragmentos, análogos o derivados pueden crearse usando procedimientos de síntesis o modificación química.

[0188] Los expertos en la materia están también familiarizados con otros métodos de hacer variantes. Éstos incluyen a procedimientos en los cuales secuencias de ácido nucleico obtenidas de aislados naturales son modificadas para

generar ácidos nucleicos que codifiquen polipéptidos que tengan características que acrecienten su valor en aplicaciones industriales o de laboratorio. En tales procedimientos se generan y caracterizan las de un gran número de secuencias variantes que tengan una o varias diferencias nucleotídicas con respecto a la secuencia obtenida del aislado natural. Típicamente, estas diferencias nucleotídicas redundan en cambios de aminoácidos con respecto a los polipéptidos codificados por los ácidos nucleicos de los aislados naturales.

[0189] Por ejemplo pueden crearse variantes usando PCR propensa a error. En la PCR propensa a error, la PCR se lleva a cabo bajo condiciones en las que es baja la fidelidad de copia de la DNA polimerasa, de forma tal que se obtiene una alta tasa de mutaciones puntuales a todo lo largo del producto de PCR. La PCR propensa a error está descrita en Leung, D.W. et al., Technique, 1:11-15, 1989) y en Caldwell, R. C. & Joyce G.F., PCR Methods Applic., 2:28-33, 1992. Dicho brevemente, en tales procedimientos los ácidos nucleicos a mutagenizar son mezclados con cebadores de PCR, tampón de reacción, MgCl2, MnCl2, Taq polimerasa y una concentración de dNTPs apropiada para alcanzar una alta tasa de mutación puntual a todo lo largo del producto de PCR. Por ejemplo, la reacción puede llevarse a cabo usando 20 fmoles de ácido nucleico a mutagenizar, 30 pmoles de cada cebador de PCR, un tampón de reacción que comprenda KCl 50mM, Tris HCl 10mM (pH 8,3) y un 0,01% de gelatina, MgCl2 7mM, MnCl2 0,5mM, 5 unidades de Taq polimerasa, DGTP 0,2mM, dATP 0,2mM, dCTP 1mM y dTTP 1mM. La PCR puede llevarse a cabo con 30 ciclos de 94ºC por espacio de 1 min., 45ºC por espacio de 1 min. y 72ºC por espacio de 1 min. Sin embargo, se apreciará que estos parámetros pueden variarse según sea apropiado. Los ácidos nucleicos mutagenizados son clonados en un apropiado vector, y se evalúan las actividades de los polipéptidos codificados por los ácidos nucleicos mutagenizados.

[0190] Pueden también crearse variantes usando mutagénesis dirigida por oligonucleótidos para generar mutaciones específicas de sitio en todo DNA clonado de interés. La mutagénesis por oligonucleótidos está descrita en Reidhaar-Olson, J.F. & Sauer, R.T., et al., Science, 241:53-57, 1988. Dicho brevemente, en tales procedimientos los de una pluralidad de oligonucleótidos de doble hebra que llevan una o varias mutaciones a introducir en el DNA clonado son sintetizados e insertados en el DNA clonado a mutagenizar. Se recuperan los clones que contienen el DNA mutagenizado y se valoran las actividades de los polipéptidos que codifican.

[0191] Otro método para generar variantes es la PCR de ensamblaje. La PCR de ensamblaje supone el ensamblaje de un producto de PCR a partir de una mezcla de pequeños fragmentos de DNA. Se producen en paralelo en el mismo vial las de un gran número de distintas reacciones PCR, cebando los productos de una reacción a los productos de otra reacción. La PCR de ensamblaje está descrita en la Patente U.S. Nº 5.965.408, presentada el 9 de julio de 1996 y titulada “Method of DNA Reassembly by Interrupting Synthesis” (“Método de Reensamblaje de DNA Mediante la Interrupción de la Síntesis”).

[0192] Aun otro método para generar variantes es la mutagénesis por PCR sexual. En la mutagénesis por PCR sexual se produce recombinación homóloga forzada entre moléculas de DNA de distinta pero altamente afín secuencia de DNA *in vitro*, como resultado de la fragmentación aleatoria de la molécula de DNA basada en la homología secuencial, seguida por fijación del sobrecruzamiento mediante extensión con cebador en una reacción PCR. La mutagénesis por PCR sexual está descrita en Stemmer, W. P., PNAS, EE.UU., 91:10747-10751, 1994. Dicho brevemente, en tales procedimientos los de una pluralidad de ácidos nucleicos a recombinar son digeridos con DNAsa para generar fragmentos que tienen un tamaño medio de 50-200 nucleótidos. Los fragmentos del deseado tamaño medio son purificados y puestos nuevamente en suspensión en una mezcla de PCR. La PCR se lleva a cabo bajo condiciones que facilitan la recombinación entre los fragmentos de ácido nucleico. Por ejemplo, la PCR puede ser llevada a cabo poniendo de nuevo en suspensión los fragmentos purificados a una concentración de 10-30 ng/l en una solución 0,2mM de cada dNTP, MgCl2 2,2mM, KCl 50mM, Tris HCl 10mM, pH 9,0, y un 0,1% de Triton X-100. Se añaden 2,5 unidades de Taq polimerasa por cada 100:1 de mezcla de reacción, y la PCR se lleva a cabo usando el régimen siguiente: 94ºC por espacio de 60 segundos, 94ºC por espacio de 30 segundos, 50-55ºC por espacio de 30 segundos, 72ºC por espacio de 30 segundos (30-45 veces) y 72ºC por espacio de 5 minutos. Sin embargo, se apreciará que estos parámetros pueden variarse según sea apropiado. En algunas realizaciones pueden incluirse oligonucleótidos en las reacciones PCR. En otras realizaciones puede usarse en un primer conjunto de reacciones PCR el fragmento de Klenow de la DNA polimerasa I, y puede usarse Taq polimerasa en un posterior conjunto de reacciones PCR. Se aíslan las secuencias recombinantes y se valoran las actividades de los polipéptidos que codifican.

[0193] Las variantes pueden también variarse mediante mutagénesis *in vivo*. En algunas realizaciones se generan mutaciones aleatorias en una secuencia de interés propagando la secuencia de interés en una cepa bacteriana, tal como una cepa de E. coli, que lleve mutaciones en una o varias de las rutas de reparación de DNA. Tales cepas “mutadoras” tienen una tasa de mutación aleatoria más alta que la de un progenitor de tipo salvaje. La propagación del DNA en una de estas cepas generará finalmente mutaciones aleatorias dentro del DNA. Cepas mutadoras adecuadas para ser usadas para mutagénesis *in vivo* están descritas en la Publicación Nº WO 91/16427 al amparo del PCT, publicada el 31 de octubre de 1991 y titulada “Methods for Phenotype Creation from Multiple Gene Populations” (“Métodos para la Creación de Fenotipos a partir de Múltiples Poblaciones Génicas”).

[0194] Pueden también generarse variantes usando mutagénesis en casete. En la mutagénesis en casete una pequeña región de una molécula de DNA de doble hebra es sustituida por un “casete” oligonucleotídico sintético que se diferencia de la secuencia nativa. El oligonucleótido a menudo contiene secuencia nativa completa y/o parcialmente aleatorizada.

[0195] Puede también usarse para generar variantes la mutagénesis de conjunto recursivo. La mutagénesis de conjunto recursivo es un algoritmo para la manipulación de proteínas (mutagénesis de proteínas) desarrollado para producir diversas poblaciones de mutantes fenotípicamente afines cuyos miembros se diferencian en la secuencia de aminoácidos. Este método usa un mecanismo de realimentación para controlar sucesivos ciclos de mutagénesis combinatoria en casete. La mutagénesis de conjunto recursivo está descrita en Arkin, A.P. y Youvan, D.C., PNAS, EE.UU., 89:7811-7815, 1992.

[0196] En algunas realizaciones se crean variantes usando mutagénesis de conjunto exponencial. La mutagénesis de conjunto exponencial es un proceso para generar bibliotecas combinatorias con un alto porcentaje de mutantes singulares y funcionales, en donde pequeños grupos de residuos son aleatorizados en paralelo para identificar, en cada posición alterada, aminoácidos que conduzcan a proteínas funcionales. La mutagénesis de conjunto exponencial está descrita en Delegrave, S. y Youvan, D.C., Biotechnology Research, 11:1548-1552, 1993. La mutagénesis aleatoria y la mutagénesis dirigida por sitio están descritas en Arnold, F.H., Current Opinion in Biotechnology, 4:450-455, 1993.

[0197] En algunas realizaciones, las variantes se crean usando procedimientos de reordenamiento en donde partes de los de una pluralidad de ácidos nucleicos que codifican distintos polipéptidos son fusionadas entre sí para crear secuencias de ácido nucleico quimérico que codifican polipéptidos quiméricos como se describe en la Patente U.S. Nº 5.965.408, presentada el 9 de julio de 1996 y titulada “Method of DNA Reassembly by Interrupting Synthesis” (“Método de Reensamblaje de DNA Mediante la Interrupción de la Síntesis”) y en la Patente U.S. Nº 5.939.250, presentada el 22 de mayo de 1996 y titulada “Production of Enzymes Having Desired Activities by Mutagenesis” (“Producción de Enzimas que Tienen Actividades Deseadas por Mutagénesis”).

[0198] Las variantes de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 pueden ser variantes en las cuales uno o varios de los residuos de aminoácidos de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 están sustituidos por un residuo de aminoácido conservado o no conservado (preferiblemente un residuo de aminoácido conservado), y tal residuo de aminoácido sustituido puede ser o puede no ser uno codificado por el código genético.

[0199] Son sustituciones conservadoras aquéllas que sustituyen un aminoácido dado en un polipéptido por otro aminoácido de iguales características. Las siguientes sustituciones se ven típicamente como sustituciones conservadoras: sustituciones de un aminoácido alifático tal como Alanina, Valina, Leucina e Isoleucina por otro aminoácido alifático. Sustitución de la Serina por una Treonina o viceversa; sustitución de un residuo ácido tal como ácido aspártico o ácido glutámico por otro residuo ácido; sustitución de un residuo que lleve un grupo amida, tal como Asparagina y Glutamina, por otro residuo que lleve un grupo amida; intercambio de un residuo básico tal como Lisina y Arginina por otro residuo básico; y sustitución de un residuo aromático tal como Fenilalanina o Tirosina por otro residuo aromático.

[0200] Son otras variantes aquéllas en las cuales uno o varios de los residuos de aminoácidos de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 incluye(n) un grupo sustituyente.

[0201] Son aun otras variantes aquéllas en las cuales el polipéptido es asociado con otro compuesto, tal como un compuesto para incrementar la vida media del polipéptido (como por ejemplo polietilenglicol).

[0202] Son adicionales variantes aquéllas en las cuales aminoácidos adicionales son fusionados con el polipéptido, tal como una secuencia conductora, una secuencia secretora, una secuencia proproteína o una secuencia que facilite la purificación, el enriquecimiento o la estabilización del polipéptido.

[0203] En algunas realizaciones, los fragmentos, derivados y análogos conservan la misma función o actividad biológica como los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y de secuencias sustancialmente idénticas a las mismas. En otras realizaciones, el fragmento, derivado o análogo incluye una proteína, de tal manera que el fragmento, derivado o análogo puede ser activado por escisión de la parte de proproteína para producir un polipéptido activo.

[0204] Otro aspecto de la invención es polipéptidos o fragmentos de los mismos que tengan una homología de un 95% o de más de un 95% con uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y con secuencias sustancialmente idénticas a las mismas, o con fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas. La homología puede determinarse usando cualesquiera de los programas que se han descrito anteriormente, que alinea los polipéptidos o fragmentos que se comparan y determina el grado de identidad de los aminoácidos o de similitud entre ellos. Se apreciará que la “homología” de los aminoácidos incluye a las sustituciones de aminoácidos conservadoras tales como las que se han descrito anteriormente.

[0205] Los polipéptidos o fragmentos que tengan homología con uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y con secuencias sustancialmente idénticas a las mismas, o con un fragmento que comprenda el menos 75, 100 o 150 aminoácidos consecutivos de las mismas pueden ser obtenidos aislando los ácidos nucleicos que los codifican usando las técnicas que se han descrito anteriormente.

[0206] Como alternativa, los polipéptidos o fragmentos homólogos pueden ser obtenidos mediante procedimientos de purificación o enriquecimiento bioquímico. La secuencia de polipéptidos o fragmentos potencialmente homólogos puede determinarse por digestión proteolítica, electroforesis en gel y/o microsecuenciación. La secuencia del probable polipéptido o fragmento homólogo puede ser comparada con uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y con secuencias sustancialmente idénticas a la misma, o con un fragmento que comprenda al menos 75, 100 o 150 aminoácidos consecutivos de la misma usando cualquiera de los programas que se han descrito anteriormente.

[0207] Otro aspecto de la invención es un ensayo para identificar fragmentos o variantes de las secuencias de aminoácidos de la ID SEC Nº: 2 y de secuencias sustancialmente idénticas a las mismas que conservan la función enzimática de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas. Por ejemplo los fragmentos o variantes de dichos polipéptidos pueden usarse para catalizar reacciones bioquímicas que indiquen que el fragmento o la variante conserva la actividad enzimática de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2.

[0208] El ensayo para determinar si los fragmentos o variantes conservan la actividad enzimática de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y de secuencias sustancialmente idénticas a las mismas incluye los pasos de: poner al fragmento o variante de polipéptido en contacto con una molécula de sustrato bajo condiciones que le permitan al fragmento o variante de polipéptido funcionar, y detectar ya sea una disminución del nivel de sustrato o bien un incremento del nivel del específico producto de reacción de la reacción entre el polipéptido y el sustrato.

[0209] Pueden usarse en una variedad de aplicaciones los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias prácticamente idénticas a las mismas o los fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas. Por ejemplo, los polipéptidos o fragmentos de los mismos pueden usarse para catalizar reacciones bioquímicas. Según un aspecto de la invención, se aporta un proceso para utilizar los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias prácticamente idénticas a las mismas o los polinucleótidos que codifican tales polipéptidos para hidrolizar enlaces glicosídicos. En tales procedimientos, una sustancia que contenga un enlace glicosídico (como p. ej. un almidón) es puesta en contacto con uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 o con secuencias sustancialmente idénticas a la misma bajo condiciones que faciliten la hidrólisis del enlace glicosídico.

[0210] Los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a las mismas o los fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas pueden también usarse en la licuefacción y sacarificación de almidón. Usando los polipéptidos o fragmentos de los mismos de esta invención, la licuefacción puede realizarse a un pH más bajo que con enzimas anteriores. En una realización, la licuefacción se lleva a cabo a un pH de 4,5. Adicionalmente, los polipéptidos o fragmentos de los mismos de esta invención son menos dependientes del calcio que las enzimas anteriormente usadas en estos procesos. En la licuefacción se usan amilasas para hidrolizar el almidón. En una realización preferida, los polipéptidos o fragmentos de los mismos de esta invención son termoestables a 90-95ºC.

[0211] Los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a las mismas o los fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas pueden también usarse para generar anticuerpos que se unan específicamente a los polipéptidos o fragmentos. Los anticuerpos resultantes pueden usarse en procedimientos de cromatografía de inmunoafinidad para aislar o purificar el polipéptido o para determinar si el polipéptido está presente una muestra biológica. En tales procedimientos, una preparación proteica, tal como un extracto, o una muestra biológica es puesta en contacto con un anticuerpo capaz de fijarse específicamente a uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y a secuencias sustancialmente idénticas a la misma, o a fragmentos que comprendan al menos 75-100 o 150 aminoácidos consecutivos de la misma.

[0212] En los procedimientos de inmunoafinidad, el anticuerpo es unido a un soporte sólido, tal como una perla u otra matriz de columna. Se pone a la preparación proteica en contacto con el anticuerpo bajo condiciones en las cuales el anticuerpo se fije específicamente a uno de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y a secuencias sustancialmente idénticas al mismo, o a fragmentos del mismo. Tras un lavado para eliminar las proteínas no específicamente fijadas, se eluyen los polipéptidos específicamente unidos.

[0213] La capacidad de las proteínas de una muestra biológica para unirse al anticuerpo puede determinarse usando cualquiera de los de una variedad de procedimientos con los que están familiarizados los expertos en la materia. Por ejemplo, la unión puede determinarse etiquetando el anticuerpo con una etiqueta detectable tal como un agente fluorescente, una etiqueta enzimática o un radioisótopo. Como alternativa, la unión del anticuerpo a la muestra puede detectarse usando un anticuerpo secundario que tenga en el mismo una etiqueta detectable de este tipo. Se incluyen entre los ensayos particulares los ensayos ELISA, los ensayos en sándwich, los radioinmunoensayos y los Western Blots.

[0214] Anticuerpos policlonales generados contra los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas o fragmentos que comprendan al menos 75, 100 o 150

aminoácidos consecutivos de las mismas pueden obtenerse mediante inyección directa de los polipéptidos a un animal o bien administrando los polipéptidos a un animal, como por ejemplo un no humano. El anticuerpo así obtenido se unirá entonces al propio polipéptido. De esta manera, incluso puede usarse una secuencia que codifique tan sólo un fragmento del polipéptido para generar anticuerpos que puedan unirse al polipéptido nativo entero. Tales anticuerpos pueden ser entonces usados para aislar al polipéptido de células que expresen ese polipéptido.

[0215] Para la preparación de anticuerpos monoclonales puede usarse cualquier técnica que proporcione anticuerpos producidos mediante cultivos continuos de líneas celulares. Los ejemplos incluyen la técnica del hibridoma (Kohler y Milstein, Nature, 256:495-497, 1975), la técnica del trioma, la técnica del hibridoma de las células B humanas (Kozbor et al., Immunology Today 4:72, 1983) y la técnica del EBV-hibridoma (Cole et al., 1985, en Monoclonal Antibodies and Cancer Therapy, Alan H. Liss, Inc., pp. 77-96).

[0216] Las técnicas que se han descrito para la producción de anticuerpos de cadena única (Patente U.S. Nº 4.946.778) pueden ser adaptadas para producir anticuerpos de cadena única para los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y para secuencias sustancialmente idénticas a las mismas, o para fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas. Como alternativa pueden usarse ratones transgénicos para expresar anticuerpos humanizados para estos polipéptidos o fragmentos de los mismos.

[0217] Los anticuerpos generados contra los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a las mismas o los fragmentos que comprendan al menos 75, 100 o 150 aminoácidos consecutivos de las mismas pueden usarse para el cribaje para la localización de similares polipéptidos de otros organismos y muestras. En tales técnicas los polipéptidos del organismo son puestos en contacto con el anticuerpo y se detectan aquéllos polipéptidos que se unen específicamente al anticuerpo. Puede usarse para detectar la unión de anticuerpos cualquiera de los procedimientos que se han descrito anteriormente. Un ensayo de cribaje de este tipo está descrito en “Methods for Measuring Cellulase Activities”, Methods in Enzymology, Vol 160, pp. 87-116.

[0218] En el sentido en el que se la utiliza en la presente, la expresión “secuencia de ácido nucleico como la expuesta en la ID SEC Nº: 1” engloba a las secuencias nucleotídicas de las secuencias de ácido nucleico de la ID SEC Nº: 1 y a las secuencias sustancialmente idénticas a las mismas, así como a las secuencias homólogas a las secuencias de ácido nucleico de la ID SEC Nº: 1 y a fragmentos de las mismas y a secuencias complementarias de todas las secuencias precedentes. Los fragmentos incluyen a las partes de la ID SEC Nº: 1 que comprendan al menos 75, 100, 150, 200, 300, 400 o 500 nucleótidos consecutivos de las secuencias de ácido nucleico de la ID SEC Nº: 1 y a las secuencias sustancialmente idénticas a las mismas. Las expresiones “secuencias y fragmentos homólogos a las secuencias de ácido nucleico de la ID SEC Nº: 1” y “secuencias sustancialmente idénticas a las mismas” hacen referencia a una secuencia que tenga una homología de al menos un 99%, un 98%, un 97%, un 96%, un 95% o un 90% con estas secuencias. La homología puede determinarse usando cualesquiera de los parámetros y programas de ordenador que aquí se describen, incluyendo el programa FASTA versión 3.0t78 con los parámetros por defecto. Las secuencias homólogas también incluyen a las secuencias de RNA en las cuales uridinas sustituyen a las timinas en las secuencias de ácido nucleico que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1. Las secuencias homólogas pueden obtenerse usando cualesquiera de los procedimientos que aquí se describen, o bien pueden resultar de la corrección de un error de secuenciación. Se apreciará que las secuencias de ácido nucleico que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y a las secuencias sustancialmente idénticas a las mismas pueden representarse en el tradicional formato de caracteres individuales (véase la contraportada interior de Stryer, Lubert. Biochemistry, 3ª Ed., W. H Freeman & Co., Nueva York) o en cualquier otro formato que registre la identidad de los nucleótidos en una secuencia.

[0219] En el sentido en el que se la utiliza en la presente, la expresión “una secuencia polipeptídica como la que se expone en la ID SEC Nº: 2” engloba a la secuencia polipeptídica de las secuencias de aminoácidos de la ID SEC Nº: 2 y a las secuencias sustancialmente idénticas a la misma, que son codificadas por una secuencia como la que se expone en la ID SEC Nº: 2, por secuencias polipeptídicas homólogas a los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y por las secuencias sustancialmente idénticas a las mismas, o por fragmentos de cualesquiera de las secuencias precedentes. La expresión “secuencias polipeptídicas homólogas” se refiere a una secuencia polipeptídica que tenga una homología de al menos un 99%, un 98%, un 97%, un 96% o un 95% con una de las secuencias polipeptídicas de las secuencias de aminoácidos de la ID SEC Nº: 2. La homología puede determinarse usando cualesquiera de los parámetros y programas de ordenador que aquí se describen, incluyendo el programa FASTA versión 3.0t78 con los parámetros por defecto o con cualesquiera parámetros modificados. Las secuencias homólogas pueden ser obtenidas usando cualesquiera de los procedimientos que aquí se describen, o bien pueden resultar de la corrección de un error de secuenciación. Los fragmentos polipeptídicos comprenden al menos 75, 100 o 150 aminoácidos consecutivos de los polipéptidos de las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas. Se apreciará que los códigos de polipéptido que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y en las secuencias sustancialmente idénticas a las mismas pueden representarse en el tradicional formato de caracteres individuales o formato de tres letras (véase la contraportada interior de Stryer, Lubert. Biochemistry, 3ª Ed., W. H Freeman & Co., Nueva York) o en cualquier otro formato que relate la identidad de los polipéptidos en una secuencia.

[0220] Los expertos en la materia apreciarán que una secuencia de ácido nucleico como la que se expone en la ID SEC Nº: 1 y una secuencia polipeptídica como la que se expone en la ID SEC Nº: 2 puede ser almacenada, registrada y manipulada en cualquier soporte que pueda ser leído por un ordenador y al cual pueda acceder un ordenador. En el sentido en el que se las utiliza en la presente, las palabras "registrada" y "almacenada" se refieren a un proceso para almacenar información en un soporte informático. Un experto en la materia puede fácilmente adoptar cualquiera de los métodos que son actualmente conocidos para registrar información en un soporte legible por ordenador para generar manufacturas que comprendan una o varias de las secuencias de ácido nucleico que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas, una o varias de las secuencias polipeptídicas que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas. Es útil en el contexto de la invención un soporte legible por ordenador que tenga registradas en el mismo al menos 2, 5, 10, 15 o 20 secuencias de ácido nucleico como las que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas.

[0221] Otro ejemplo es un soporte legible por ordenador que tenga registradas en el mismo una o varias de las secuencias de ácido nucleico que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas. Otro ejemplo es un medio legible por ordenador que tenga registradas en el mismo una o varias de las secuencias polipeptídicas que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas. Otro ejemplo es un medio legible por ordenador que tenga registradas en el mismo al menos 2, 5, 10, 15 o 20 de las secuencias que se han expuesto anteriormente.

[0222] Los soportes legibles por ordenador incluyen a los miembros del grupo que consta de soportes legibles magnéticamente, soportes legibles ópticamente, soportes legibles electrónicamente y soportes magnético-ópticos. Por ejemplo, los soportes legibles por ordenador pueden ser un disco duro, un disco flexible, una cinta magnética, una CD-ROM, un Disco Digital Versátil (DVD), una Memoria de Acceso Aleatorio (RAM) o una Memoria de Sólo Lectura (ROM), así como otros tipos de otros soportes que son conocidos para los expertos en la materia.

[0223] Los ejemplos incluyen a los miembros del grupo que consta de sistemas (como p. ej. sistemas basados en Internet), y en particular sistemas informáticos que almacenen y manipulen la información secuencial que aquí se describe. Está ilustrado en forma de diagrama de bloques en la Figura 1 un ejemplo de un sistema informático 100. En el sentido en el que se la utiliza en la presente, la expresión "un sistema informático" se refiere a los componentes de hardware (hardware = soporte físico), a los componentes de software (software = soporte lógico informático) y a los componentes de almacenamiento de datos que se usan para analizar una secuencia nucleotídica de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2. El sistema informático 100 típicamente incluye un procesador para procesar y manipular los datos secuenciales y para acceder a los mismos. El procesador 105 puede ser cualquier tipo perfectamente conocido de unidad central de proceso, tal como por ejemplo la unidad Pentium III de la Intel Corporation, o un procesador similar de Sun, Motorola, Compaq, AMD o International Business Machines.

[0224] Típicamente el sistema informático 100 es un sistema para usos generales que comprende el procesador 105 y uno o varios componentes de almacenamiento interno de datos 110 para almacenar datos, y uno o varios dispositivos de recuperación de datos para recuperar los datos almacenados en los componentes de almacenamiento de datos. Un experto en la materia puede apreciar fácilmente que son adecuados cualesquiera de los sistemas informáticos que están disponibles en la actualidad.

[0225] En una realización particular, el sistema informático 100 incluye un procesador 105 conectado a un bus que está conectado a una memoria principal 115 (implementada preferiblemente como RAM) y a uno o varios dispositivos de almacenamiento interno de datos 110, tales como una unidad de disco duro y/u otros soportes legibles por ordenador en los que se registran los datos. En algunas realizaciones, el sistema informático 100 incluye adicionalmente uno o varios dispositivos de recuperación de datos 118 para leer los datos almacenados en los dispositivos de almacenamiento interno de datos 110.

[0226] El dispositivo de recuperación de datos 118 puede representar, por ejemplo, una unidad de disco flexible, una unidad de disco compacto, una unidad de cinta magnética o un módem capaz de establecer conexión con un sistema de almacenamiento de datos remoto (p. ej. a través de Internet), etc. En algunas realizaciones, el dispositivo de almacenamiento interno de datos 110 es un soporte amovible legible por ordenador tal como un disco flexible, un disco compacto, una cinta magnética, etc. que contenga lógica de control y/o datos registrados en el mismo. El sistema informático 100 puede ventajosamente incluir o ser programado por un software apropiado para leer la lógica de control y/o los datos a partir del componente de almacenamiento de datos una vez insertado en el dispositivo de recuperación de datos.

[0227] El sistema informático 100 incluye un monitor o visualizador 120 que se usa para visualizar la salida para un usuario del ordenador. Hay que observar también que el sistema informático 100 puede ser enlazado a otros sistemas informáticos 125a-c en una red o red de área ancha para así proporcionar un acceso centralizado al sistema informático 100.

[0228] El software para acceder a y procesar las secuencias nucleotídicas de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a las mismas, o una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a las mismas (tales como herramientas de búsqueda, herramientas de comparación y herramientas de modelización, etc.) puede residir en la memoria principal 115 durante la ejecución.

[0229] El sistema informático 100 puede además comprender un algoritmo de comparación de secuencias para comparar a una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y a las secuencias sustancialmente idénticas a la misma, o a una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y a las secuencias sustancialmente idénticas a la misma, almacenadas en un soporte legible por ordenador, con una o varias secuencias nucleotídicas o polipeptídicas de referencia almacenadas en un soporte legible por ordenador. La expresión “algoritmo de comparación de secuencias” se refiere a uno o varios programas que se implementan (local o remotamente) en el sistema informático 100 para comparar una secuencia nucleotídica con otras secuencias nucleotídicas y/o compuestos almacenados en unos medios de almacenamiento de datos. Por ejemplo, el algoritmo de comparación de secuencias puede comparar las secuencias nucleotídicas de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas, o una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma, almacenadas en un soporte legible por ordenador, con secuencias de referencia almacenadas en un soporte legible por ordenador para identificar homologías o motivos estructurales. Se contemplan particularmente para ser usados en este aspecto de la invención varios programas de comparación de secuencias que se identifican en otros sitios en esta memoria descriptiva. Las homologías de las secuencias de proteína y/o las secuencias de ácido nucleico pueden evaluarse usando cualesquiera de los de la variedad de programas y algoritmos de comparación de secuencias que son conocidos en la técnica. Tales algoritmos y programas incluyen, aunque en absoluto sin carácter limitativo, a los miembros del grupo de los llamados TBLASTN, BLASTP, FASTA, TFASTA y CLUSTALW (Pearson and Lipman, Proc. Natl. Acad. Sci. EE.UU. 85(8):2444-2448, 1988; Altschul et al., J. Mol. Biol. 215(3):403-410, 1990; Thompson et al., Nucleic Acids Res. 22(2):4673-4680, 1994; Higgins et al., Methods Enzymol. 266:383-402, 1996; Altschul et al., J. Mol. Biol. 215(3):403-410, 1990; Altschul et al., Nature genetics 3:266-272, 1993).

[0230] La homología o la identidad se mide a menudo usando software de análisis de secuencias (como p. ej. el Paquete de Software de Análisis de Secuencias del Genetics Computer Group, del University of Wisconsin Biotechnology Center, de 1710 University Avenue, Madison, WI 53705). Tal software casa las secuencias similares asignando grados de homología a las distintas deleciones, sustituciones y otras modificaciones. En el contexto de dos o más ácidos nucleicos o secuencias polipeptídicas, los vocablos “homología” e “identidad” se refieren a dos o más secuencias o subsecuencias que son iguales o tienen un porcentaje especificado de residuos de aminoácidos o nucleótidos que son iguales al ser comparados y alineados para una máxima correspondencia en una región designada o ventana de comparación según medición efectuada usando cualquier cantidad de algoritmos de comparación de secuencias o bien mediante alineación manual e inspección visual.

[0231] Para la comparación de secuencias, típicamente una secuencia hace de secuencia de referencia, con la cual se comparan las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, se introducen en un ordenador las secuencias de ensayo y de referencia, de ser necesario se designan coordenadas de subsecuencia, y se designan los parámetros del programa del algoritmo de comparación de secuencias. Pueden usarse parámetros por defecto del programa, o bien pueden designarse parámetros alternativos. El algoritmo de comparación de secuencias calcula entonces el porcentaje de identidades secuenciales para las secuencias de ensayo con respecto a la secuencia de referencia, sobre la base de los parámetros del programa.

[0232] En el sentido en el que la expresión se usa en la presente, una “ventana de comparación” hace referencia a un segmento de cualquiera de las series de posiciones contiguas que se selecciona de entre los miembros del grupo que consta de un número de 20 a 600, habitualmente de poco más o menos 50 a poco más o menos 200, más habitualmente de poco más o menos 100 a poco más o menos 150, en el cual una secuencia puede ser comparada con una secuencia de referencia de igual número de posiciones contiguas tras haber sido las dos secuencias óptimamente alineadas. Los métodos de alineación de secuencias para comparación son perfectamente conocidos en la técnica. Puede llevarse a cabo una óptima alineación de secuencias para comparación p. ej. por medio del algoritmo de homología local de Smith & Waterman, Adv. Appl. Math. 2:482, 1981, por medio del algoritmo de alineación de homología de Needleman & Wunsch, J. Mol. Biol. 48:443, 1970, por medio del método de búsqueda de la similitud de Person & Lipman, Proc. Natl. Acad. Sci. USA 85:2444, 1988, por medio de implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA y TFASTA en el Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), o bien mediante alineación manual e inspección visual. Otros algoritmos para determinar la homología o la identidad incluyen, por ejemplo, además de un programa BLAST (Basic Local Alignment Search Tool del National Center for Biological Information), a los llamados ALIGN, AMS (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, BLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, algoritmo de Smith-Waterman, DARWIN, algoritmo de 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 Comparison), LALIGN (Local Sequence Alignment), LCP (Local Content Program), MACAW (Multiple Alignment Construction & 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 pueden también usarse para cribar bases de datos genómicos para identificar secuencias polinucleotídicas que tengan secuencias sustancialmente idénticas. Están disponibles las de una serie de bases de datos genómicos, y así por ejemplo, una parte considerable del genoma humana está disponible como parte del Proyecto de Secuenciación del Genoma Humano (J. Roach, http://weber.u.Washington.edu/~roach/human_genome_progress.2.html) (Gibbs, 1995). Han sido ya secuenciados al menos veintiún otros genomas, incluyendo, por ejemplo, a los de M. genitalium (Fraser et al., 1995), M. jannaschii (Bult et al., 1996), H. influenzae (Fleischmann et al., 1995), E. coli (Blattner et al., 1997) y levadura (S. cerevisiae) (Mewes et al., 1997) y D. melanogaster (Adams et al., 2000). Se han hecho también importantes progresos en la secuenciación de los genomas de organismos modelo, tales como el ratón, C. elegans y Arabadopsis sp. Varias bases de datos que contienen información genómica con anotaciones con alguna información funcional son mantenidas por distintas organizaciones y son accesibles a través de Internet, por ejemplo en http://wwwtigr.org/tdb; http://www.genetics.wisc.edu; http://genome-www.stanford.edu/~ball; http://hiv-web-lanl.gov; http://www.ncbi.nim.nih.gov; http://www.ebi.ac.uk; http://Pasteur.fr/other/biology y http://www.genome.wi.mit.edu.

[0233] Un ejemplo de un algoritmo útil es el de los algoritmos BLAST y BLAST 2.0, que están descritos en Altschul et al., Nuc. Acids Res. 25:3389-3402, 1977 y en Altschul et al., J. Mol. Biol. 215:403-410, 1990, respectivamente. Software para llevar a cabo análisis BLAST está a disposición del público a través del Centro Nacional de Información Biotecnológica (http://www.ncbi.nim.nih.gov/). Este algoritmo supone identificar primeramente pares de secuencias con alta puntuación (HSPs) identificando palabras cortas de longitud W en la secuencia de consulta que casan o satisfacen alguna puntuación umbral valorada como positiva T al ser alineadas con una palabra de la misma longitud de una secuencia de la base de datos. A T se le denomina el umbral de puntuación de palabras vecinas (Altschul et al., supra). Estas iniciales coincidencias de palabras vecinas hacen de semillas para iniciar búsquedas para encontrar HSPs que las contengan. Las coincidencias de palabras se extienden en ambas direcciones a lo largo de cada secuencia en la medida en que pueda incrementarse la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan usando, para secuencias nucleotídicas, los parámetros M (puntuación de recompensa para un par de residuos coincidentes; siempre > 0). Para secuencias de aminoácidos se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de las coincidencias de palabras en cada dirección se detiene cuando: la puntuación de alineación acumulativa cae en la cantidad X desde su máximo valor alcanzado; la puntuación acumulativa llega a cero o más abajo, debido a la acumulación de una o varias alineaciones de residuos con puntuación negativa; o cuando se llega al final de cualquier secuencia. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y velocidad de la alineación. El programa BLASTN (para secuencias nucleotídicas) usa por defecto una longitud de palabra (W) de 11, una esperanza (E) de 10, M = 5, N = -4 y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa como valores por defecto una longitud de palabra de 3 y una esperanza (E) de 10, y la matriz de puntuación BLOSUM62 (véase Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915, 1989) alineaciones (B) de 50, una esperanza (E) de 10, M = 5, N = -4 y una comparación de ambas hebras.

[0234] El algoritmo BLAST también lleva a cabo un análisis estadístico de la similitud entre dos secuencias (véase, p. ej., Karlin & Altschul, Proc. Natl. Acad. Sci. USA 90:5873, 1993). Una medida de la similitud que proporciona el algoritmo BLAST es la menor probabilidad de suma (P(N)), que proporciona una indicación de la probabilidad con la cual se produciría una coincidencia entre dos secuencias nucleotídicas o de aminoácidos. Por ejemplo, se considera que un ácido nucleico es similar a una secuencia de referencia si la menor probabilidad de suma en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es de menos de aproximadamente 0,2, más preferiblemente de menos de aproximadamente 0,01, y con la máxima preferencia de menos de aproximadamente 0,001.

[0235] En una realización, las homologías de secuencias de proteínas de ácidos nucleicos se evalúan usando la herramienta llamada Basic Local Alignment Search Tool (“BLAST”). En particular se usan cinco programas BLAST específicos para llevar a cabo las tareas siguientes:

(1) Los programas BLASTP y BLAST3 comparan una secuencia de aminoácidos de consulta contra una base de datos de secuencias de proteínas;
(2) el programa BLASTN compara una secuencia nucleotídica de consulta contra una base de datos de secuencias nucleotídicas;
(3) el programa BLASTX compara los productos de traducción conceptual de seis marcos de una secuencia nucleotídica de consulta (ambas hebras) contra una base de datos de secuencias de proteínas;
(4) el programa TBLASTN compara una secuencia proteica de consulta contra una base de datos de secuencias nucleotídicas traducidas en todos los seis marcos de lectura (ambas hebras); y
(5) el programa TBLASTX compara las traducciones de seis marcos de una secuencia nucleotídica de consulta contra las traducciones de seis marcos de una base de datos de secuencias nucleotídicas.

[0236] Los programas BLAST identifican secuencias homólogas identificando segmentos similares, a los que se les denomina en la presente “pares de segmentos con alta puntuación”, entre una secuencia de aminoácidos o de ácido nucleico de consulta y una secuencia de ensayo que se obtiene preferiblemente de una base de datos de secuencias de

proteínas o ácidos nucleicos. Los pares de segmentos con alta puntuación se identifican (es decir, se alinean) preferiblemente por medio de una matriz de puntuación, siendo muchas de éstas conocidas en la técnica. Preferiblemente, la matriz de puntuación que se usa es la matriz BLOSUM62 (Gonnet et al., Science 256:1443-1445, 1992; Henikoff y Henikoff, Proteins 17:49-61, 1993). Menos preferiblemente pueden también usarse las matrices PAM o PAM250 (véase, p. ej. Schwartz y Dayhoff, redactores en jefe, 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation). Los programas BLAST son accesibles a través de la U.S. National Library of Medicine, p. ej. en el sitio www.ncbi.nlm.nih.gov.

[0237] Los parámetros que se usan con los algoritmos anteriormente indicados pueden ser adaptados en dependencia de la longitud de las secuencias y del grado de homología que se estudie. En algunas realizaciones, los parámetros pueden ser los parámetros por defecto que usan los algoritmos en ausencia de instrucciones del usuario.

[0238] La Figura 2 es un diagrama de flujo que ilustra una realización de un proceso 200 para comparar una nueva secuencia nucleotídica o proteica con una base de datos de secuencias a fin de determinar los niveles de homología entre la nueva secuencia y las secuencias de la base de datos. La base de datos de secuencias puede ser una base de datos privada almacenada en el sistema informático 100, o bien una base de datos pública tal como la GENBANK, que está disponible a través de Internet.

[0239] El proceso 200 comienza en el estado de inicio 201 y pasa luego a un estado 202 en el que la nueva secuencia a comparar es almacenada en una memoria en un sistema informático 100. Como se ha expuesto anteriormente, la memoria podría ser cualquier tipo de memoria, incluyendo una memoria RAM o un dispositivo de almacenamiento interno.

[0240] El proceso 200 pasa entonces a un estado 204 en el que se abre una base de datos de secuencias a efectos de análisis y comparación. El proceso 200 pasa entonces a un estado 206 en el que la primera secuencia almacenada en la base de datos es leída y almacenada en una memoria del ordenador. Se lleva luego a cabo una comparación en un estado 210 para determinar si la primera secuencia es la misma como la segunda secuencia. Es importante observar que este paso no queda limitado a llevar a cabo una comparación exacta entre la nueva secuencia y la primera secuencia de la base de datos. Son conocidos para los expertos en la materia métodos perfectamente conocidos para comparar dos secuencias nucleotídicas o proteicas, aunque no sean idénticas. Por ejemplo, pueden introducirse lagunas en una secuencia a fin de incrementar el nivel de homología entre las dos secuencias sometidas a ensayo. Los parámetros que controlan si durante la comparación se introducen en una secuencia lagunas u otras características son normalmente introducidos por el usuario del sistema informático.

[0241] Una vez llevada a cabo una comparación de las dos secuencias en el estado 210, se hace una determinación en un estado de decisión 210 sobre si las dos secuencias son iguales. Naturalmente, el vocablo “iguales” no se limita a secuencias que sean absolutamente idénticas. Las secuencias que estén dentro de los parámetros de homología introducidos por el usuario serán marcadas como “iguales” en el proceso 200.

[0242] Si se hace una determinación de que las dos secuencias son iguales, el proceso 200 pasa a un estado 214 en el que es visualizado para el usuario el nombre de la secuencia de la base de datos. Este estado le notifica al usuario que la secuencia con el nombre visualizado satisface las exigencias en materia de homología que fueron introducidas. Una vez visualizado para el usuario el nombre de la secuencia almacenada, el proceso 200 pasa a un estado de decisión 218 en el que se hace una determinación de 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 de fin 220. No obstante, si de hecho existen más secuencias en la base de datos, el proceso 200 pasa entonces a un estado 224 en el que un puntero es llevado a la siguiente secuencia de la base de datos para que la misma pueda ser comparada con la nueva secuencia. De esta manera, la nueva secuencia es alineada y comparada con cada secuencia de la base de datos.

[0243] Hay que observar que si en el estado de decisión 212 se hizo una determinación de que las secuencias no eran homólogas, el proceso 200 pasaría entonces inmediatamente al paso de decisión 218 a fin de determinar si están disponibles otras secuencias en la base de datos para comparación.

[0244] En consecuencia, un ejemplo que puede usarse en el contexto de la invención es un sistema informático que comprende un procesador, un dispositivo de almacenamiento de datos que tiene almacenadas en el mismo una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma, un dispositivo de almacenamiento de datos que tiene almacenadas recuperablemente en el mismo secuencias nucleotídicas de referencia o secuencias polipeptídicas de referencia a comparar con una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y con secuencias sustancialmente idénticas a la misma, o con una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y con secuencias sustancialmente idénticas a la misma, y un comparador de secuencias para llevar a cabo la comparación. El comparador de secuencias puede indicar un nivel de homología entre las secuencias comparadas o identificar motivos estructurales en el anteriormente descrito código de ácido nucleico de las secuencias de ácido nucleico de la ID SEC Nº: 1, y en secuencias sustancialmente idénticas al mismo, o en una secuencia polipeptídica como la que se expone en

las secuencias de aminoácidos de la ID SEC Nº: 2 y en secuencias sustancialmente idénticas a la misma, o bien puede identificar motivos estructurales en secuencias que se comparen con estos códigos de ácido nucleico y códigos polipeptídicos. En algunos ejemplos, el dispositivo de almacenamiento de datos puede tener almacenadas en el mismo las secuencias de al menos 2, 5, 10, 15, 20, 25, 30 o 40 o más de las secuencias de ácido nucleico que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a las mismas, o las secuencias polipeptídicas que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a las mismas.

[0245] Es útil en el contexto de la invención un método para determinar el nivel de homología entre una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o una secuencia polipeptídica como se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma, y una secuencia nucleotídica de referencia. El método incluye los pasos de leer el código de ácido nucleico o el código polipeptídico y la secuencia nucleotídica o polipeptídica de referencia mediante el uso de un programa de ordenador que determina los niveles de homología, y determinar la homología entre el código de ácido nucleico o el código polipeptídico y la secuencia nucleotídica o polipeptídica de referencia con el programa de ordenador. El programa de ordenador puede ser cualquiera de los de una serie de programas de ordenador para determinar niveles de homología, incluyendo los aquí específicamente enumerados (como p. ej. el programa BLAST2N con los parámetros por defecto o con cualesquiera parámetros modificados). El método puede implementarse usando los sistemas informáticos que se han descrito anteriormente. El método puede también ejecutarse leyendo al menos 2, 5, 14, 15, 20, 25, 30 o 40 o más de las anteriormente descritas secuencias de ácido nucleico como las que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 o de las secuencias polipeptídicas como las que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 mediante el uso del programa de ordenador, y determinando la homología entre los códigos de ácido nucleico o códigos polipeptídicos y las secuencias nucleotídicas o secuencias polipeptídicas de referencia.

[0246] 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 pasa luego a un estado 254 en el que se almacena en una memoria una primera secuencia a comparar. La segunda secuencia a comparar es luego almacenada en una memoria en un estado 256. El proceso 250 pasa entonces a un estado 260 en el que se lee el primer carácter de la primera secuencia, y luego a un estado 262 en el que se lee el primer carácter de la segunda secuencia. Debe entenderse que si la secuencia es una secuencia nucleotídica, el carácter sería entonces normalmente una letra A, T, C, G o U. Si la secuencia es una secuencia proteica, entonces está preferiblemente en el código de aminoácidos de letra única, con lo cual pueden compararse fácilmente las secuencias primera y segunda.

[0247] Entonces se hace una determinación en un estado de decisión 264 sobre si los dos caracteres son iguales. Si son iguales, el proceso 250 pasa entonces a un estado 268 en el que se leen los siguientes caracteres de las secuencias primera y segunda. Se hace entonces una determinación sobre si los siguientes caracteres son iguales. Si lo son, el proceso 250 continúa entonces con este bucle hasta que dos caracteres no sean iguales. Si se hace una determinación de que los siguientes dos caracteres no son iguales, el proceso 250 pasa a un estado de decisión 274 para determinar si hay más caracteres de cualquier secuencia a leer.

[0248] Si no hay más caracteres para leer, el proceso 250 pasa entonces a un estado 276 en el que es visualizado para el usuario el nivel de homología entre las secuencias primera y segunda. El nivel de homología se determina calculando la proporción de caracteres entre las secuencias que eran iguales del número total de secuencias en la primera secuencia. Así, si cada carácter de una primera secuencia de 100 nucleótidos quedó alineado con cada carácter de una segunda secuencia, el nivel de homología sería del 100%.

[0249] Como alternativa, el programa de ordenador puede ser un programa de ordenador que compare las secuencias nucleotídicas de una secuencia de ácido nucleico como la que se expone en la invención con una o varias secuencias nucleotídicas de referencia a fin de determinar si el código de ácido nucleico de las secuencias de ácido nucleico de la ID SEC Nº: 1 y de las secuencias sustancialmente idénticas a la misma se diferencia de una secuencia de ácido nucleico de referencia en una o varias posiciones. Opcionalmente, un programa de este tipo registra la longitud y la entidad de los nucleótidos insertados, delecionados o sustituidos con respecto a la secuencia ya sea del polinucleótido de referencia o bien de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico en la ID SEC Nº: 1, y con respecto a las secuencias sustancialmente idénticas a la misma. En una realización el programa de ordenador puede ser un programa que determine si una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a la misma contiene(n) un único polimorfismo nucleotídico (SNP) con respecto a una secuencia nucleotídica de referencia.

[0250] En consecuencia, otro ejemplo que es útil en el contexto de la invención es un método para determinar si una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a la misma se diferencia(n) en uno o varios nucleótidos de una secuencia nucleotídica de referencia, comprendiendo dicho método los pasos de leer el código del ácido nucleico y la secuencia nucleotídica de referencia mediante el uso de un programa de ordenador que identifique las diferencias entre secuencias de ácido nucleico, e identificar las diferencias entre el código del ácido nucleico y la secuencia nucleotídica de referencia con el programa de ordenador. En algunos ejemplos, el programa de ordenador es un programa que

identifica polimorfismos nucleotídicos individuales. El método puede ser implementado por medio de los sistemas informáticos que se han descrito anteriormente y del método que se ilustra en la Figura 3. El método puede también ejecutarse leyendo al menos 2, 5, 10, 15, 20, 25, 30 o 40 o más de las secuencias de ácido nucleico como las que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y de las secuencias sustancialmente idénticas a las mismas y de las secuencias nucleotídicas de referencia mediante el uso del programa de ordenador, e identificando las diferencias entre los códigos de ácido nucleico y las secuencias nucleotídicas de referencia con el programa de ordenador.

[0251] En otros ejemplos, el sistema informático puede adicionalmente comprender un identificador para identificar características dentro de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 o dentro de una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y dentro de secuencias sustancialmente idénticas a la misma.

[0252] El vocablo "identificador" se refiere a uno o varios programas que identifican ciertas características dentro de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y dentro de secuencias sustancialmente idénticas a la misma, o dentro de una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y de secuencias sustancialmente idénticas a la misma. El identificador puede comprender un programa que identifique un marco de lectura abierto en una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y en las secuencias sustancialmente idénticas a la misma.

[0253] La Figura 4 es un diagrama de flujo que ilustra una realización 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 pasa luego a un estado 304 en el que se almacena en una memoria 115 en el sistema informático 100 una primera secuencia de la que deben verificarse las características. El proceso 300 pasa luego a un estado 306 en el que se abre una base de datos de características secuenciales. Una base de datos de este tipo incluiría una lista de los atributos de cada característica junto con el nombre de la característica. Por ejemplo, un nombre de característica podría ser "Codón de Iniciación" y el atributo sería "ATG". Otro ejemplo sería el nombre de característica "Caja TAATAA" y el atributo de la características sería "TAATAA". Un ejemplo de una base de datos de este tipo es producido por el Grupo Informático de Genética de la Universidad de Wisconsis (www.gcg.com). Como alternativa, las características pueden ser motivos polipeptídicos estructurales tales como hélices alfa, láminas beta o motivos polipeptídicos funcionales tales como sitios activos enzimáticos, motivos hélice-giro-hélice u otros motivos conocidos para los expertos en la materia.

[0254] Una vez abierta la base de datos de características en el estado 306, el proceso 300 pasa a un estado 308 en el que se lee la primera característica en la base de datos. Luego se hace en un estado 310 una comparación del atributo de la primera característica con la primera secuencia. Entonces se hace en un estado de decisión 316 una determinación sobre si el atributo de la característica fue encontrado en la primera secuencia. Si el atributo fue encontrado, el proceso 300 pasa entonces a un estado 318 en el que es visualizado para el usuario el nombre de la característica hallada.

[0255] El proceso 300 pasa entonces a un estado de decisión 320 en el que se hace una determinación sobre si existen más características en la base de datos. Si no existen más características, el proceso 300 termina entonces en un estado de fin 324. Sin embargo, si de hecho existen más características en la base de datos, el proceso 300 lee entonces la siguiente característica secuencial en un estado 326 y va de regreso al estado 310, en el que el atributo de la siguiente característica es comparado contra la primera secuencia.

[0256] Hay que señalar que si el atributo de característica no es encontrado 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.

[0257] En consecuencia, otro ejemplo que es útil en el contexto de la invención es un método que es para identificar una característica dentro de una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y dentro de las secuencias sustancialmente idénticas a la misma, o dentro de una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y de las secuencias sustancialmente idénticas a la misma, y comprende el paso de leer el (los) código(s) de ácido nucleico o el (los) código(s) polipeptídico(s) mediante el uso de un programa de ordenador que identifica características en los mismos e identifica características dentro del (de los) código(s) de ácido nucleico con el programa de ordenador. En un ejemplo, el programa de ordenador comprende un programa de ordenador que identifica marcos de lectura abiertos. El método puede ejecutarse leyendo una única secuencia o bien al menos 2, 5, 10, 15, 20, 25, 30 o 40 de las secuencias de ácido nucleico como las que se exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a las mismas, o las secuencias polipeptídicas como las que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a las mismas, mediante el uso del programa de ordenador, e identificando características dentro de los códigos de ácido nucleico o de los códigos polipeptídicos con el programa de ordenador.

[0258] Una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y secuencias sustancialmente idénticas a la misma, o bien una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y secuencias sustancialmente idénticas a la misma pueden almacenarse y manipularse en una variedad de programas de procesadores de datos en una variedad de formatos. Por ejemplo, una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y las secuencias sustancialmente idénticas a la misma, o bien una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y las secuencias sustancialmente idénticas a la misma pueden almacenarse como texto en un archivo de procesamiento de palabras tal como el WORD de Microsoft o el WORDPERFECT o bien como un archivo ASCII en los de una variedad de programas de bases de datos con los que están familiarizados los expertos en la materia, tales como los programa DB2, SYBASE u ORACLE. Por añadidura, pueden usarse muchos programas de ordenador y bases de datos como algoritmos de comparación de secuencias, como identificadores o bien como fuentes de secuencias nucleotídicas o secuencias polipeptídicas de referencia a comparar con una secuencia de ácido nucleico como la que se expone en las secuencias de ácido nucleico de la ID SEC Nº: 1 y con las secuencias sustancialmente idénticas a la misma, o bien con una secuencia polipeptídica como la que se expone en las secuencias de aminoácidos de la ID SEC Nº: 2 y con las secuencias sustancialmente idénticas a la misma. La lista que se da a continuación no pretende limitar la invención, sino servir de guía sobre los programas y bases de datos que son útiles con las secuencias de ácido nucleico como las que exponen en las secuencias de ácido nucleico de la ID SEC Nº: 1 y con las secuencias sustancialmente idénticas a las mismas, o con las secuencias polipeptídicas como las que se exponen en las secuencias de aminoácidos de la ID SEC Nº: 2 y con las secuencias sustancialmente idénticas a las mismas.

[0259] Los programas y bases de datos que pueden usarse incluyen, aunque sin carácter limitativo, a los miembros del grupo que consta de los siguientes: MacPattern (EMBL), DiscoveryBase (Molecular Applications Group), GeneMine (Molecular Applications Group), Look (Molecular Applications Group), MacLook (Molecular Applications Group), BLAST y BLAST2 (NCBI), BLASTN y BLASTX (Altschul et al., J. Mol. Biol. 215: 403, 1990), FASTA (Pearson and Lipman, Proc. Natl. Acad. Sci. USA, 85: 2444, 1988), FASTDB (Brutlag et al. Comp. App. Biosci. 6:237-245, 1990) Catalyst (Molecular Simulations Inc.), Catalyst/SHAPE (Molecular Simulations Inc.), Cerius2.DBAaccess (Molecular Simulations Inc.), HypoGen (Molecular Simulations Inc.), Insight II, (Molecular Simulations Inc.), Discover (Molecular Simulations Inc.), CHARMm (Molecular Simulations Inc.), Felix (Molecular Simulations Inc.), DelPhi, (Molecular Simulations Inc.), QueanteMM, (Molecular Simulations Inc.), Homology (Molecular Simulations Inc.), Modeler (Molecular Simulations Inc.), ISIS (Molecular Simulations Inc.), Quanta/Protein Design (Molecular Simulations Inc.), WebLab (Molecular Simulations Inc.), WebLab Diversity Explorer (Molecular Simulations Inc.), Gene Explorer (Molecular Simulations Inc.), SeqFold (Molecular Simulations Inc.), la base de datos Chemicals Directoy disponible en MDL, la base de datos Drug Data Report de MDL, la base de datos Comprehensive Medicinal Chemistry, la base de datos World Drug Index de Derwent, la base de datos BioByteMasterFile, la base de datos Genbank y la base de datos Genseqn. Dada la presente publicación, le resultarán obvios a un experto en la materia muchos otros programas y bases de datos.

[0260] Los motivos que pueden ser detectados usando los programas anteriormente mencionados incluyen a los miembros del grupo que consta de secuencias que codifican cremalleras de leucina, motivos hélice-giro-hélice, sitios de glicosilación, sitios de ubiquitinación, hélices alfa y láminas beta, secuencias señal que codifican péptidos señal que dirigen la secreción de las proteínas codificadas, secuencias implicadas en la regulación de la transcripción tales como cajas homeóticas (“homeoboxes”), tramos ácidos, sitios activos enzimáticos, sitios de unión al sustrato y sitios de corte enzimático.

[0261] La presente invención explota las singulares propiedades catalíticas de las enzimas. Mientras que el uso de biocatalizadores (es decir, enzimas crudas o purificadas o células vivientes o no vivientes) en las transformaciones químicas requiere normalmente la identificación de un determinado biocatalizador que reaccione con un específico compuesto de partida, la presente invención usa biocatalizadores y condiciones de reacción seleccionados que son específicos de grupos funcionales que están presentes en muchos compuestos de partida, tales como pequeñas moléculas. Cada biocatalizador es específico de un grupo funcional o de varios grupos funcionales afines, y puede reaccionar con muchos compuestos de partida que contengan este grupo funcional.

[0262] Las reacciones biocatalíticas producen una población de derivados a partir de un único compuesto de partida. Estos derivados pueden ser sometidos a otro ciclo de reacciones biocatalíticas para producir una segunda población de compuestos derivados. Con cada iteración de derivatización biocatalítica pueden producirse miles de variaciones de la pequeña molécula o del compuesto original.

[0263] Las enzimas reaccionan en sitios específicos de un compuesto de partida sin afectar al resto de la molécula, el cual es un proceso que es muy difícil de lograr usando métodos químicos tradicionales. Este alto grado de especificidad biocatalítica proporciona los medios para identificar un único compuesto activo dentro de la biblioteca. La biblioteca está caracterizada por la serie de reacciones biocatalíticas usada para producirla, a la cual se la denomina “historial biosintético”. Cribando la biblioteca con respecto a las actividades biológicas y efectuando un seguimiento del historial biosintético se identifica la secuencia de reacción específica que produce el compuesto activo. Se repite la secuencia de reacción y se determina la estructura del compuesto sintetizado. A diferencia de otros métodos de síntesis y cribaje, este modo de identificación no requiere tecnologías de inmovilización, y los compuestos pueden ser sintetizados y sometidos a ensayo libremente en solución usando virtualmente cualquier tipo de ensayo de cribaje. Es importante

observar que el alto grado de especificidad de las reacciones enzimáticas sobre grupos funcionales permite el "seguimiento" de las reacciones enzimáticas específicas que constituyen la biblioteca producida biocatalíticamente.

[0264] Muchos de los pasos procesales son llevados a cabo usando automatización robótica, que permite la ejecución de muchos miles de reacciones biocatalíticas y ensayos de cribaje por día, así como asegurar un alto nivel de precisión y reproducibilidad. Como resultado de ello, puede producirse en cuestión de semanas una biblioteca de compuestos de derivados que llevaría años producir usando los métodos químicos corrientes.

[0265] Es útil en el contexto de la invención un método que es para modificar pequeñas moléculas y comprende el paso de poner a un polipéptido codificado por un nucleótido aquí descrito o fragmentos enzimáticamente activos del mismo en contacto con una pequeña molécula para así producir una pequeña molécula modificada. Se somete a ensayo a una biblioteca de pequeños moléculas modificadas para determinar si está presente dentro de la biblioteca una pequeña molécula modificada que presente una actividad deseada. Se identifica una específica reacción biocatalítica que produce la pequeña molécula modificada de actividad deseada eliminando sistemáticamente cada una de las reacciones biocatalíticas usadas para producir una parte de la biblioteca, y sometiendo luego a ensayo a las pequeñas moléculas producidas en la parte de la biblioteca para determinar la presencia o ausencia de la pequeña molécula modificada con la actividad deseada. Se repiten opcionalmente las específicas reacciones biocatalíticas que producen la pequeña molécula modificada de actividad deseada. Las reacciones biocatalíticas son llevadas a cabo con un grupo de biocatalizadores que reaccionan con distintas mitades estructurales que se encuentran dentro de la estructura de una pequeña molécula, siendo cada biocatalizador específico de una mitad estructural o de un grupo de mitades estructurales afines; y cada biocatalizador reacciona con muchas distintas pequeñas moléculas que contienen la distinta mitad estructural.

[0266] Se describe a continuación más ampliamente la invención haciendo referencia a los ejemplos siguientes; si bien debe entenderse que la invención no queda limitada a tales ejemplos.

Ejemplos

Ejemplo 1

Identificación y Caracterización de α-Amilasas Termoestables

[0267] El presente ejemplo muestra la identificación de nuevas amilasas ácidas. El programa de cribaje fue realizado bajo condiciones de pH neutro y bajo. Se apuntó para el descubrimiento a bibliotecas de DNA generadas a partir de muestras de bajo pH. Este esfuerzo proporcionó el descubrimiento de cientos de clones que tienen la capacidad de degradar el almidón. Análisis bioinformáticos y de la secuencia de DNA clasificaron a muchos de estos genes como amilasas anteriormente no identificadas.

Estudios Bioquímicos

[0268] El análisis bioquímico de los clones genómicos de amilasa puso de manifiesto que muchos tenían unos óptimos valores pH de menos de pH 6. Lisados de estos clones genómicos fueron sometidos a ensayo para determinar la tolerancia térmica mediante incubación a 70ºC, 80ºC, 90ºC o 100ºC por espacio de 10 minutos y medición de la actividad residual a pH 4,5. Se eligieron para su adicional análisis los clones que conservaban una actividad de > 50% tras tratamiento térmico a 80ºC. Estos clones fueron incubados a 90ºC por espacio de 10 minutos a pH 6,0 y 4,5 y fueron sometidos a ensayo para determinar la actividad residual a pH 4,5. Los de una serie de clones conservaron más de un 40% de su actividad a continuación de este tratamiento. A efectos comparativos, la actividad residual de una amilasa desarrollada, ID SEC Nº: 2, era equivalente a la mejor de las enzimas de segunda generación; y la actividad específica de la ID SEC Nº: 2 era mayor.

[0269] La actividad térmica de los clones con actividad residual tras tratamiento térmico a 90ºC a pH 4,5 fue medida a temperatura ambiente, a 70ºC y a 90ºC a pH 4,5. La Tabla 1 muestra que las tasas de hidrólisis de la ID SEC Nº: 3 (amilasa de *B. stearothermophilus*) y de la ID SEC Nº: 4 (amilasa de *B. licheniformis*) disminuyen a las temperaturas más altas, mientras que la tasa para ID SEC Nº: 5 continúa aumentando al ser la temperatura aumentada hasta 70ºC y tan sólo se reduce en aproximadamente un 50% a 90ºC.

Evaluación de los Candidatos

[0270] Sobre la base de la actividad residual a pH 4,5 tras tratamiento térmico a 90ºC, la actividad específica y la tasa de hidrólisis de almidón a 90ºC en comparación con la amilasa de *B. licheniformis*, ID SEC Nº: 5, es comparada con la amilasa desarrollada de la ID SEC Nº: 2 en un ensayo de licuefacción de almidón.

Tabla 1

	Temperatura ambiente	70ºC	90ºC
ID SEC Nº: 3	1,25	1,43	0,33
ID SEC Nº: 4[2]	3,3	1,9	0,39
ID SEC Nº: 5	1,9	47	19

Tabla 1. Tasas de hidrólisis de almidón etiquetado con colorante (unidades relativas de fluorescencia/seg.) de tres clones genómicos a pH 4,5 y 3 temperaturas distintas. [1]Amilasa de *B. stearothermophilus*, [2]amilasa de *B. licheniformis*

Ejemplo 2

Ensayo de la Actividad de Amilasa: Ensayo de los Extremos Reductores por el Método del Ácido Bicinconínico (BCA)

[0271]

1. Preparar 2 soluciones de sustrato como se indica a continuación:

a) solución pH 8 de almidón (de patata) soluble al 2% disolviendo 2 g de almidón de patata en 100 ml de fosfato sódico 100mM pH 8.

b) solución pH 10 de almidón (de patata) soluble al 2% disolviendo 2 g de almidón de patata en 100 ml de carbonato sódico 100mM.

Calentar ambas soluciones en un baño de agua hirviendo mientras se efectúa mezcla por espacio de 30-40 minutos hasta disolverse el almidón.

2. Preparar Solución A a partir de 64 mg/ml de monohidrato de carbonato sódico, 24 mg/ml de bicarbonato sódico y 1,95 mg/ml de BCA (sal disódica de 4,4'-dicarboxi-2,2'-biquinolina (Sigma Chemical, Nº de cat. D-8284)). Añadida anteriormente a dH2O (dH2O = agua destilada).

3. Preparar solución B combinando 1,24 mg/ml de pentahidrato de sulfato cúprico y 1,26 mg/ml de L-serina. Añadir la mezcla a dH2O.

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

5. Preparar una solución estándar de maltosa 10mM en dH2O, donde la maltosa 10mM se combina en almidón soluble al 2% al pH deseado hasta una concentración final 0, 100, 200, 300, 400 y 600 µM. La curva estándar será generada para cada conjunto de puntos en el tiempo. Puesto que la curva se determina añadiendo 10 µl de los patrones al reactivo de trabajo, sale a 0, 1, 2, 3, 4, 6 nmoles de maltosa.

6. Introducir partes alícuotas de 1 ml de solución de sustrato en tubos de microcentrífuga, y equilibrar a la temperatura deseada (5 min.) en bloque de calor o en baño de agua calentado. Añadir 60 µl de solución enzimática al interior de la tapa del tubo.

7. Mientras se equilibra la solución, mezclar 5 ml de ambas soluciones A y B. Poner partes alícuotas de 100 µl en una placa de PCR de 96 pocillos. Poner la placa sobre hielo.

8. Tras 5 minutos de equilibrado de la temperatura, cerrar la tapa de los tubos, invertirlos y someterlos 3 veces a agitación vorticial. Poner inmediatamente partes alícuotas de 10 µl en la placa en el t = 0 (punto cero en el tiempo). Dejar la mezcla enzimática en el bloque de calor y añadir partes alícuotas de 10 µl en cada deseado punto en el tiempo (p. ej. a los 0, 5, 10, 15, 20 y 30 minutos).

9. Asegurarse de que se dejen vacíos 12 pocillos (solamente con la adición de partes alícuotas de reactivo de trabajo) para la adición de 10 µl de patrones, para la curva estándar.

10. Cuando se hayan efectuado las recolecciones y se hayan añadido los patrones en todos los puntos en el tiempo, cubrir la placa y calentarla a 80ºC por espacio de 35 min. Enfriar la placa sobre hielo por espacio de 10 min. Añadir 100 µl de H2O a todos los pocillos. Mezclar y poner partes alícuotas de 100 µl en una placa de 96 pocillos de fondo plano y leer la absorbancia a 560 nm.

11. Poner a cero cada uno de los puntos en el tiempo de la muestra contra su propio t = 0 (restar el valor medio de A560 en el t = 0 de los otros valores medios de A560). Convertir la $A560_{(experimental)}$ en µmoles (dividir la $A560_{(experimental)}$ por la pendiente de la curva estándar (A560/ µmol)). Generar una pendiente de los puntos en el tiempo y los µmoles (en µmoles/min.), y multiplicar por 100 (puesto que el valor de µmoles tan sólo da cuenta de los 10 µl usados en el ensayo, y no de la cantidad hecha en la cantidad de 1 ml de rxn). Para obtener la actividad específica, dividir la pendiente (en µmoles/min.) por los mg de proteína. Todos los puntos deberán hacerse como mínimo por duplicado, siendo lo mejor tres. Se expone en la Figura 5 un ejemplo de curva estándar.

Tabla 2 Datos de la muestra:

Clon	Dilución	Minutos	A560-1	A560-2	A560 media	A560 puesta a cero	(Pendiente A560exp/std) µmoles
ENZ	50	0	0,1711	0,1736	0,17235	0	0,0000
		5	0,2104	0,2165	0,21345	0,0411	0,0005
		10	0,2492	0,2481	0,24865	0,0763	0,0009
		15	0,2984	0,2882	0,2933	0,12095	0,0014
		20	0,3355	0,3409	0,3382	0,16585	0,0020
		30	0,3942	0,3805	0,38735	0,215	0,0026
		40	0,4501	0,4412	0,44565	0,2733	0,0033

Ejemplo 3

Cribaje para la Determinación de la Actividad de α-Amilasa

[0272] El número de placas cribadas por cada placa de cultivo debería ser de aproximadamente 10.000 pfu's (pfu's = unidades formadoras de placa). Para cada biblioteca de DNA deberían cribarse al menos 50.000 placas por cada biblioteca aislada y 200.000 placas por cada biblioteca no aislada, en dependencia del título de pfu para el lisado amplificado de λ Zap Express.

Determinación del Título de la Biblioteca Lambda

[0273]
1) µl de material de biblioteca amplificado con Lambda Zap Express añadidos a 600 µl de células MRF' de *E. coli* (OD_{600} = 1,0). Para diluir el material de MRF', se usa $MgSO_4$ 10mM.
2) Incubar a 37ºC por espacio de 15 minutos.
3) Transferir la suspensión a 5-6 ml de agar superior NZY a 50ºC y mezclar lentamente
4) Verter inmediatamente la solución de agar en una gran placa de cultivo (de 150 mm) con medio NZY.
5) Dejar que el agar superior se solidifique por completo (en aproximadamente 30 minutos) y luego invertir la placa de cultivo.
6) Incubar la placa de cultivo a 39ºC por espacio de 8-12 horas.
7) Se aproxima el número de placas. Se determina el título de fago para obtener 10.000 pfu/placa de cultivo. Diluir una parte alícuota de fago de biblioteca con tampón SM de ser necesario.

Cribaje del Sustrato

[0274]
1) Se añade Lambda Zap Express (50.000 pfu) de la biblioteca amplificada a 600 µl de células MRF' de E. coli (OD600 = 1,0). Para bibliotecas no ambientales, preparar 4 tubos (50.000 pfu por tubo).
2) Incubar a 37ºC por espacio de 15 minutos.
3) Mientras se está incubando la suspensión de fago/células, se añade 1,0 ml de sustrato de almidón rojo (al 1,2% en peso/volumen) a 6,0 ml de agar superior NZY a 50ºC y se efectúa mezcla a fondo. La solución se mantiene a 50ºC hasta que se la necesite.
4) Transferir 1/5 (10.000 pfu) de la suspensión celular a la solución de sustrato/agar superior y mezclar lentamente.
5) Inmediatamente se vierte la solución en una gran placa de cultivo (de 150 mm) con medio NZY.
6) Dejar que el agar superior se solidifique por completo (en aproximadamente 30 minutos), y luego invertir la placa de cultivo.
7) Repetir los procedimientos 4-6 4 veces para el resto de la suspensión celular (1/5 de la suspensión cada vez).
8) Incubar las placas de cultivo a 39ºC por espacio de 8-12 horas
9) Se observa la placa de cultivo para detectar las zonas de inhibición (halos) en torno a las placas.
10) Las placas con halos se extraen del agar y se transfieren a un microtubo estéril. Una punta de pipeta de 200 µl con gran orificio funciona bien para retirar (extraer) el tapón de agar que contiene la placa deseada.
11) Se ponen de nuevo en suspensión los fagos en 500 µl de tampón SM. Se añaden 20 µl de cloroformo para inhibir todo adicional crecimiento celular.
12) Se incuba suspensión de fago puro a temperatura ambiente por espacio de 4 horas o durante la noche antes del paso siguiente.

Aislamiento de los Clones Puros

[0275]
1) Se añaden 10 µl de suspensión de fago puesto de nuevo en suspensión a 500 µl de células MRF' de *E. coli* (OD600 = 1,0).
2) Incubar a 37ºC por espacio de 15 minutos.

3) Mientras está incubándose la suspensión de fago/células, se añade 1 ml de sustrato de almidón rojo (al 1,2% en peso/volumen) a 6,0 ml de agar superior NZY a 50ºC y se mezcla a fondo. Mantener la solución a 50ºC hasta que sea necesaria.
4) Se transfiere la suspensión celular a solución de agar superior/sustrato y se mezcla lentamente.
5) Se vierte inmediatamente la solución en una gran placa de cultivo (de 150 mm) con medio NZY.
6) Dejar que el agar superior se solidifique por completo (en aproximadamente 30 minutos), y luego invertir la placa de cultivo.
7) Se incuba la placa de cultivo a 37ºC por espacio de 8-12 horas.
8) Se observa la placa de cultivo para detectar una zona de inhibición (halo) en torno a una placa individual (clon puro). Si no puede aislarse una placa individual, ajustar el título y poner de nuevo en placa de cultivo la suspensión de fago.
9) La placa individual con halo se extrae del agar y se transfiere a un microtubo estéril.
Una punta de pipeta de 200 µl de gran orificio funciona bien para retirar (extraer) el tapón de agar que contiene la placa deseada. Para amplificar el título, pasar 5 placas activas individuales al interior de un microtubo.
10) Se ponen nuevamente en suspensión los fagos en 500 µl de tampón SM. Se añaden 20 µl de cloroformo para inhibir todo adicional crecimiento celular.
11) Se incuba la suspensión de fago puro a temperatura ambiente por espacio de 4 horas o bien durante la noche antes del paso siguiente. La suspensión de fago puro se almacena a -80ºC añadiendo DMSO a la suspensión de fago (al 7% volumétrico).

Escisión del Clon Puro

[0276]
1) Se añaden 100 µl de suspensión de fago puro a 200 µl de células MRF' de *E. coli* (OD600 = 1,0). Se añade a esto 1,0 µl de fago ayudante ExAssist (> 1 x 106 pfu/ml; Stratagene). Se usa tubo Falcon 2059 para la escisión.
2) Se incuba la suspensión a 37ºC por espacio de 15 minutos.
3) Se añaden a la suspensión celular 3,0 ml de medio 2 x YT.
4) Incubar a 30ºC por espacio de al menos 6 horas o bien durante la noche mientras se efectúa sacudimiento.
5) Se transfiere el tubo a 70ºC por espacio de 20 minutos. La suspensión de fagémido puede tenerse almacenada a 4ºC por espacio de 1 a 2 meses.
6) Se transfieren 100 µl de suspensión de fagémido a un microtubo que contiene 200 µl de células Exp 555 de E. coli (OD600 = 1,0).
7) Se incuba la suspensión a 37ºC por espacio de 15 minutos.
8) Se añaden a la suspensión 300 µl de SOB.
9) Se incuba la suspensión a 37ºC por espacio de un periodo de tiempo de 30 a 45 minutos.
10) Se transfieren 100 µl de suspensión a una pequeña placa de cultivo (de 90 mm) con medio LB que contiene canamicina (medio LB con 50 µg/ml de canamicina) para las bibliotecas de DNA Zapp Express o ampicilina (medio LB con 100 µg/ml de canamicina) para las bibliotecas de DNA Zap II.
11) El resto de la suspensión se transfiere a otra pequeña placa de cultivo con medio LB.
12) Usar perlas de vidrio estéril para distribuir uniformemente la suspensión en la placa de cultivo.
13) Se incuban las placas de cultivo a 30ºC por espacio de un periodo de tiempo de 12 a 24 horas.
14) Se observa la placa de cultivo para detectar las colonias.
15) Inocular una colonia individual en medio líquido LB que contenga un adecuado antibiótico e incubar a 30ºC por espacio de un periodo de tiempo de 12 a 24 horas.
16) Puede prepararse solución concentrada de glicerol añadiendo un 80% de glicerol a cultivo líquido (al 15% volumétrico) y almacenando a -80ºC.

Verificación de la Actividad

[0277]
1) Se transfieren a un microtubo 50 µl de cultivo líquido. Añadir al interior del mismo tubo 500 µl de Azur de Amilopectina pH 7 al 8%. Preparar 2 tubos para cada clon.
2) La actividad se verifica a 50ºC por espacio de 3 horas y durante la noche. Usar tampón a pH 7 como control.
3) Enfriar la muestra de ensayo en baño de agua helada por espacio de 5 minutos.
4) Añadir 750 µl de etanol y mezclar a fondo.
5) Centrifugar a 13000 rpm (16000 g's) por espacio de 5 minutos.
6) Medir la OD (OD = densidad óptica) del supernatante a 595 nm.

Análisis RFLP (Análisis del Polimorfismo de los Fragmentos de Restricción)

[0278]
1) Se transfiere a un microtubo estéril 1,0 ml de cultivo líquido.
2) Se centrifuga a 13200 rpm (16000 g's) por espacio de 1 minuto.
3) Desechar el supernatante. Añadir otro 1,0 ml de cultivo líquido al mismo microtubo estéril.
4) Centrifugar a 13200 rpm (16000 g's) por espacio de 1 minuto.
5) Desechar el supernatante.
6) Seguir el protocolo del QIAprep spin mini kit para el aislamiento del plásmido.

7) Verificar la concentración de DNA usando un biofotómetro
8) Usar Sac I y Kpn I para la primera doble digestión. Incubar a 37ºC por espacio de 1 hora.
9) Usar Pst I y Xho I para la segunda doble digestión. Incubar a 37ºC por espacio de 1 hora.
10) Añadir colorante Loading a la muestra digerida.
11) Pasar la muestra digerida por un gel de agarosa al 1,0% por espacio de 1-1,5 horas a 120 voltios.
12) Ver el gel con generador de imágenes de gel. Todos los clones con un distinto patrón de digestión se enviarán al análisis de secuencias.

Ejemplo 4

Ensayo para Amilasas

Preparación de Cultivos Huésped

[0279]
1. Iniciar un cultivo durante la noche de células huésped MRF' con XL1-Blue. Usar una colonia individual de una placa de huellas para inocular 10 ml de LB suplementado con 20 µg/ml de tetraciclina. Cultivar durante la noche con sacudimiento a 37ºC por espacio de al menos 16 horas.
2. Usando una técnica aséptica, inocular 100 ml de cultivo de día LB_{Tet} fresco con huésped MRF' con XL1-Blue del cultivo LB_{Tet} de durante la noche.
3. Cultivar en un sacudidor a 37ºC hasta que la OD llegue a ser de 0,75 - 1,0.
4. Pelletizar las células huésped a 1000 x g por espacio de 10 minutos y ponerlas lentamente de nuevo en suspensión en $MgSO_4$ 10mM a una OD de 5.
5. Diluir una pequeña cantidad de células huésped hasta una OD de 1 para su uso en la titulación y en la aplicación de la tecnología de inoculación con el inoculador de puntas múltiples.
6. Las preparaciones huésped pueden usarse por espacio de hasta 1 semana si se tienen almacenadas en hielo o a 4ºC.

Comentarios

[0280]
- Para acortar el tiempo de cultivo para el cultivo de día, usar 1/2 x la habitual concentración de Tet en LB (1/2 x = 10 µg/ml), u omitir del todo el antibiótico.
- No usar NZY cuando se seleccione con tetraciclina. La alta concentración de Mg^{++} en el medio NZY hace que la Tet sea inactiva.

Titulación de las Bibliotecas Lambda

[0281]
7. Se ponen en un soporte tres tubos de microfuga estériles.
8. Se ponen partes alícuotas de 995 µl de las células huésped preparadas en un tubo y 45 µl de las células huésped a OD 1 preparadas en cada uno de los dos tubos restantes.
9. Se añaden 5 µl de biblioteca lambda al tubo que contiene 995 µl de células huésped y se efectúa mezcla vorticial. Esto redundó en un factor de dilución de 200.
10. Se preparan diluciones 1/2.000 y 1/20.000 añadiendo consecutivamente 5 µl de la dilución anterior a los dos tubos restantes que contienen 45 µl de las células huésped preparadas. Se efectúa mezcla vorticial tras haber sido hecha cada dilución.
11. Se deja que el fago se adsorba en el huésped incubando a 37ºC por espacio de 15 minutos.
12. En el entretanto se pipetan 100 µl de las células huésped a OD 1 preparadas en el interior de cada uno de tres tubos Falcon 2059.
13. Se añaden 5 µl de cada dilución a un tubo 2059 aparte que contiene células huésped.
14. Se cultiva cada preparación añadiendo 3 ml de agar superior a cada tubo y vertiendo rápidamente en placas de cultivo de 90 mm con NZY. Se asegura una regular y uniforme distribución antes de que se endurezca el agar superior.
15. Las placas de cultivo se invierten y se incuban durante la noche a 37ºC.
16. Se cuentan las placas y se calcula el título del material de la biblioteca (en unidades formadoras de placa (pfu) por µl).

1.

Cribaje de Microtitulación Lambda para Amilasas

Preparación

[0282] Como se ha descrito anteriormente, preparar una cantidad de cultivo huésped MRF' XL1-Blue suficiente para la cantidad de cribaje prevista.

1. Autoclavear varias botellas compatibles con el dispensador QFiII2. Éstas son botellas Corning de boca ancha de 250 ml que contienen una junta anular en torno al labio.
2. Asegurarse de que estén disponibles para la criba suficientes cantidades de placas de cultivo, agar superior, almidón BODIPY, solución de almidón rojo, etc.
3. Programar el ciclo robótico del Día 2 con un representante de Automation.

Día 1

[0283]

1. Etiquetar las placas de cultivo de 1536 pocillos (negras) con el número de placa de cultivo y de criba de biblioteca. Las etiquetas para tubos Tough-Tags^{MF} (MF = marca de fábrica), cortadas por la mitad a lo ancho, son ideales para etiquetar placas de cultivo de 1536 pocillos.
2. Calcular los volúmenes de células huésped y de librería y de medio NZY necesarias para la criba. Esto se hace fácilmente con una hoja de cálculo Excel.
3. Combinar los volúmenes calculados de células huésped a OD 5 y de biblioteca en una botella Corning estéril de boca ancha de 250 ml (que contiene una junta anular).
4. Dejar que se produzca adsorción a 37ºC por espacio de 15 minutos.
5. Añadir el volumen calculado de medio NZY y mezclar a fondo. A esto se le denomina la suspensión de medio, fago y células.
6. Llevar a cabo una titulación concomitante combinando 50 µl de la suspensión de medio, fago y células con 250 µl de células huésped a OD 1 en un tubo Falcon 2059, y luego cultivando en placa de cultivo con 9 ml de agar superior en una placa de cultivo de 150 mm con NZY. Incubar la placa de cultivo de titulación concomitante a 37ºC durante la noche.
7. Cargar el dispensador con el resto de la suspensión y colocar cada placa de cultivo etiquetada de 1536 pocillos con 4 µl por pocillo. Si el dispensador deja burbujas de aire en algunos pocillos, las mismas pueden ser eliminadas centrifugando las placas de cultivo a 200 x g por espacio de 1 minuto.
8. Añadir 0,5 µl de fago de control positivo a la posición de pocillo AD46 de al menos dos de las placas de cultivo de ensayo. Usar con esta finalidad un fuerte clon lambda amilasa-positivo.
9. Incubar las placas de cultivo de ensayo a 37ºC durante noche en un incubador humidificado (≥ 95%).

Día 2

[0284]

17. Contar las pfu en la placa de titulación concomitante y determinar la densidad de siembra media por pocillo (en pfu por pocillo).
18. Inocular con el inoculador de puntas múltiples al menos 2 placas de cultivo de cada tamiz de biblioteca (conteniendo preferiblemente las 2 controles positivos) de la manera siguiente:
a) Preparar 2 placas de césped huésped para que hagan de superficie sobre la cual efectuar la inoculación con inoculador de puntas múltiples: combinar 250 µl de células huésped a OD 1 con 2 ml de almidón rojo al 2% y cultivar con 9 ml de agar superior en placas de 150 mm con NZY. Mantener cada placa de cultivo tan horizontal como sea posible al solidificarse el agar superior, a fin de producir una uniforme tonalidad de rojo en toda la placa de cultivo.
b) Usando un inoculador de puntas múltiples de 1536 posiciones esterilizado dos veces con llama, replicar al menos 2 de las placas de cribaje en las placas de césped huésped.
c) Poner las placas receptoras inoculadas en una campana de flujo laminar con las tapas quitadas por espacio de aproximadamente 15-30 minutos (para eliminar el exceso de humedad).
d) Colocar de nuevo las tapas e incubar en posición invertida a 37ºC durante la noche.
19. Preparar el 2 x tampón de sustrato de almidón con BODIPY de la forma siguiente:
a) Calcular el volumen total de 2 x solución tampón de sustrato necesario para todas las placas de cribaje a razón de 4 µl por pocillo (incluyendo todo volumen de espacio muerto extra requerido por el dispensador) y medir e introducir esta cantidad de CAPS 100mM pH 10,4 en un recipiente apropiado para el dispensador que se use.
b) Recuperar suficientes tubos de 0,5 mg de almidón con BODIPY para producir el requerido volumen de 2 x tampón de sustrato [calculado en el anterior paso a)] a una concentración final de 20-30 µg/ml.
c) Disolver cada tubo de 0,5 mg en 50 µl de DMSO a temperatura ambiente, con protección frente a la luz y con frecuente agitación vorticial. Esto lleva más de 15 minutos; disolviéndose algunos lotes de producción de almidón con BODIPY mejor que otros.
d) Añadir 50 µl de tampón de CAPS 100mM pH 10,4 a cada tubo y mezclar con agitación vorticial.
e) Reunir los contenidos de todos los tubos y retirar todos los agregados no disueltos centrifugando por espacio de 1 minuto a velocidad máxima en una microfuga.
f) Añadir el supernatante al resto del tampón de CAPS 100mM medido en el anterior paso a).
g) Proteger el 2 x tampón de sustrato de la luz envolviéndolo con folio.
20. Llevar las placas de cultivo y el tampón de sustrato a la sala de automatización y programar el robot con los parámetros siguientes:
a) dispensar 4 µl de tampón de sustrato por pocillo
b) 1ª lectura al haber transcurrido 1 hora post-sustrato, 2ª lectura a las 9 horas y tercera lectura a las 17 horas; con incubación a 37ºC entre lecturas
c) filtro de excitación: 485 nm; filtro de emisión: 535 nm
d) ajustar la ganancia del Spectrafluor a 70.

e) asegurarse de que el incubador que se use protegerá a las placas de ensayo de la luz.

Día 3

[0285]
1. Verificar las placas inoculadas para detectar las zonas de inhibición en el césped bacteriano en todas las posiciones que correspondan a pocillos de la correspondiente placa de ensayo. Verificar también las zonas de inhibición en el almidón rojo en cualesquiera de las posiciones de las puntas. Si se usaron para la inoculación con el inoculador de puntas múltiples placas de cultivo que contuviesen controles positivos, debería poder verse una gran zona de inhibición en el trasfondo rojo. Tener cuidado con los contaminantes que también forman zonas de inhibición en almidón rojo (véase el comentario "Contaminantes que Forman Zonas de Inhibición en el Almidón Rojo" al final del ejemplo 7).
2. Identificar las posibles coincidencias a partir del archivo de datos producido por el ordenador del robot. El programa KANAL producido por Engineering simplifica el análisis de los datos. Como regla empírica, una posible coincidencia está caracterizada como un pocillo que tiene una intensidad de señal incrementada al menos 1,5 veces en comparación con el trasfondo.
3. Para cada posible coincidencia, quitar del pocillo 2 µl e introducirlos en un tubo que contengan 500 µl de tampón SM y 50 µl de $CHCl_3$. Efectuar mezcla vorticial y almacenar a 4ºC. De aquí en adelante se llamará a esta solución la solución concentrada 4e-3. Las placas de cribaje originales deberán tenerse almacenadas a 4ºC y protegidas de la luz, al menos hasta que hayan concluido los destacamientos.

[0286] Éste es el método recomendado para destacar las supuestas coincidencias. Es un ensayo en fase líquida que se basa en la confirmación de la actividad en almidón con BODIPY. Como alternativa, las supuestas coincidencias pueden cultivarse directamente en placas de fase sólida que contengan almidón rojo de forma tal que se examinen 2.000-3.000 pfu por coincidencia para las zonas de inhibición. Sin embargo, la incapacidad de observar zonas de inhibición en el almidón rojo no es necesariamente una indicación de que una supuesta coincidencia fuese un falso positivo. La coincidencia tendría entonces que analizarse usando el formato en el cual fue originalmente identificada (es decir, en fase líquida usando almidón con BODIPY como sustrato). Además, los positivos muy débiles se identifican más fácilmente usando el método que se detalla a continuación.

Día 1

[0287]
1. En un tubo cónico estéril de 50 ml, combinar 0,5 ml de células huésped a OD 5 con 45,5 ml de NZY. A esto se le denominará la suspensión de medio y huésped.
2. Para cada supuesta coincidencia a analizar, introducir una parte alícuota de 1 ml de suspensión de medio y huésped en cada uno de 3 tubos estériles de microfuga.
3. Ajustar la pipeta pipetman de 12 canales a modo multidispensador con un tamaño de parte alícuota de 20 µl y un número de partes alícuotas de 2 x. Montar la pipeta pipetman con un conjunto limpio de puntas estériles.
4. Verter aproximadamente 1 ml de suspensión de medio y huésped en una nueva cubeta de solución estéril y cargar la pipeta pipetman de varios canales.
5. Dispensar 20 µl por pocillo en la última fila (fila P) de una placa negra de 384 pocillos (12 canales x 2 = 24 pocillos). Esta fila se usará más tarde para los controles.
6. Expulsar el líquido que queda en las puntas poniendo a las puntas en contacto contra la superficie de la cubeta y pulsar el botón de RESET de la pipeta pipetman. Acostar la pipeta pipetman de manera que se impida la contaminación de las puntas. No hay necesidad de cambiar las puntas en este punto.
7. Verter el resto del fluido de la cubeta al interior de un recipiente para desechos (tal como un vaso de boca ancha) procurando evitar la contaminación por salpicadura.
8. Para la primera supuesta coincidencia a analizar, tomar 111 µl de la solución concentrada 4e-3 (véase Día 2 en *Cribaje de Microtitulación Lambda para Amilasas*) y añadirlos a lo primero en un conjunto de tres tubos que contengan 1 ml de suspensión de medio y huésped (paso 2). Efectuar mezcla vorticial para mezclar. Esto es la *Dilución A*.
9. Tomar 111 µl de Dilución A y añadirlos al siguiente tubo del conjunto. Mezclar en régimen vorticial. Esto es la *Dilución B*.
10. Tomar 111 µl de Dilución B y añadirlos al último tubo del conjunto. Efectuar mezcla vorticial. Esto es la *Dilución C*. Deberían ahora tenerse tres diluciones de fago, donde las concentraciones de cada una se diferencian en un factor de 10.
11. Verter el contenido de Dilución C (la más diluida de las 3 muestras) en el interior de la cubeta de solución y cargar la pipeta pipetman de varios canales.
12. Dispensar 20 µl por pocillo en la primera fila de la placa de 384 pocillos (12 canales x 2 = 24 pocillos).
13. Expulsar el líquido que queda en las puntas poniendo a las puntas en contacto contra la superficie de la cubeta y pulsar el botón de RESET de la pipeta pipetman. Acostar la pipeta pipetman de una manera que impida la contaminación de las puntas. No hay necesidad de cambiar las puntas en este punto.
14. Vaciar la cubeta como se ha descrito anteriormente.
15. Verter el contenido de Dilución B al interior de la misma cubeta y cargar la pipeta pipetman de varios canales.
16. Dispensar 20 µl por pocillo al interior de la segunda fila de la placa de 384 pocillos.
17. Llevar análogamente a cabo los pasos 13-16 para dispensar Dilución A al interior de la tercera fila de la placa de cultivo.

18. Tras haber sido dispuestas ordenadamente las tres diluciones en las 3 primeras filas de la placa de cultivo, desechar todas las puntas y la cubeta de solución en el contenedor de desechos biológicamente peligrosos.
19. Montar la pipeta pipetman con un conjunto limpio de puntas estériles y abrir una nueva cubeta de solución estéril.
20. Repetir los pasos 8-19 para cada supuesta coincidencia restante, usando las restantes filas de la placa hasta la fila O. Pueden analizarse cinco supuestas coincidencias en una placa de 384 pocillos, reservando la última fila (fila P) para los controles.
21. Añadir 0,5 µl de cada control a un pocillo aparte. Usar al menos 2-3 controles aparte, que cubran preferiblemente una gama de actividades.
22. Incubar las placas de ensayo a 37ºC durante la noche en un incubador humidificado (≥ 95%).

Día 2

[0288]

1. Inocular con un inoculador de puntas múltiples todas las placas de destacar coincidencias en un césped huésped con almidón rojo usando el mismo método que se ha descrito para el Día 2 en *Cribaje de Microtitulación Lambda para Amilasas*, exceptuando el hecho de que se usa un inoculador de puntas mútliples de 384 posiciones.
2. Preparar el 2 x tampón de sustrato de almidón con BODIPY de la forma siguiente:
a) Calcular el volumen total de 2 x solución tampón de sustrato que será necesario para todas las placas usadas para destacar las supuestas coincidencias a razón de 20 µl por pocillo (incluyendo todo volumen de espacio muerto extra que requiera el dispensador) y medir e introducir esta cantidad de CAPS 100mM pH 10,4 en el interior de un recipiente apropiado para el dispensador que se use.
b) Recuperar tubos de 0,5 mg de almidón con BODIPY suficientes para producir el requerido volumen de 2 x tampón de sustrato [calculado en el anterior paso a)] a una concentración final de 20-30 µg/ml.
c) Disolver cada tubo de 0,5 mg en 50 µl de DMSO a temperatura ambiente, protegiendo de la luz y con frecuente agitación vorticial. Esto lleva más de 15 minutos; disolviéndose algunos lotes de producción de almidón con BODIPY mejor que otros.
d) Añadir 50 µl de tampón de CAPS 100mM pH 10,4 a cada tubo y efectuar mezcla vorticial.
e) Reunir los contenidos de todos los tubos y retirar todos los agregados no disueltos centrifugando por espacio de 1 minuto a velocidad máxima en una microfuga.
f) Añadir el supernatante al resto del tampón de CAPS 100mM medido en el anterior paso a).
g) Proteger el 2 x tampón de sustrato de la luz envolviéndolo con folio.
3. Dispensar 20 µl por pocillo en todas las placas usadas para destacar las supuestas coincidencias.
4. Envolver todas las placas en folio de aluminio e incubarlas a temperatura ambiente por espacio de 2-6 horas.
5. Leer cada placa en el Spectrafluor con los ajustes siguientes:
a) lectura de fluorescencia (filtro de excitación: 485 nm; filtro de emisión: 535 nm)
b) definición de la placa: negra de 384 pocillos
c) leer desde arriba
d) ganancia óptima
e) número de destellos: 3
6. En la resultante hoja de cálculo Excel, poner en forma de gráfico cada 3 filas de supuestas coincidencias en un gráfico aparte y verificar la actividad. Asegurarse de que los controles positivos produzcan señales destacándose del trasfondo.
7. Para cada supuesta coincidencia que parezca tener una señal real entre los pocillos, recolectar una muestra de un pocillo positivo de la manera siguiente:
a) Seleccionar un pocillo positivo de una fila que represente la más alta dilución inicial.
b) Transferir 2 µl de ese pocillo al interior de un tubo que contenga 500 µl de SM y 50 µl de $CHCl_3$. Ésta es la llamada solución destacadora.
c) Almacenar a 4ºC.
8. Usando los métodos anteriormente descritos, cultivar aproximadamente 10 µl de cada solución destacadora en placas de cultivo de 150 mm con NZY usando almidón rojo. El objetivo es obtener varias (al menos 20) placas bien separadas de las cuales extraer aislados.

Día 3

[0289]

1. Verificar las placas inoculadas para detectar una aceptable incidencia de inhibiciones en el césped bacteriano que correspondan a pocillos de la correspondiente placa de cultivo de ensayo. Verificar también la presencia de inhibiciones en el almidón rojo en los controles positivos y en cualesquiera supuestas coincidencias que se comprueben. Tener cuidado con los contaminantes que también forman zonas de inhibición en el almidón rojo (véase lo expuesto más adelante).
2. De las placas de fase sólida que contienen diluciones de soluciones concentradas destacadoras, extraer varias placas aisladas y poner a cada una en 500 µl de SM con 50 µl de $CHCl_3$. Esto recibe el nombre de solución concentrada de aislado.
3. Las soluciones concentradas de aislado pueden ser luego sometidas individualmente a ensayo sobre almidón con BODIPY usando los métodos anteriormente descritos. Puede prescindirse de este paso si la placa que fue extraída en el paso 2 produjo una zona de inhibición en el trasfondo de almidón rojo. Las soluciones concentradas de aislado pueden

ser entonces sometidas individualmente a ensayo sobre almidón con BODIPY usando los métodos anteriormente descritos. Sin embargo, puede prescindirse de este paso si la placa que fue extraída en el paso 2 produjo una zona de inhibición en el trasfondo de almidón rojo.

Escisiones

Día 1

[0290]

1. En un tubo Falcon 2059, mezclar 200 µl de huésped MRF' XL1-Blue a OD 1, 100 µl de solución concentrada de aislado lambda y 1 µl de solución concentrada de fago ExAssist.
2. Incubar en sacudidor a 37ºC por espacio de 15 minutos.
3. Añadir 3 ml de medio NZY.
4. Incubar en sacudidor a 30ºC durante la noche.

Día 2

[0291]

1. Calentar el tubo de escisión hasta 70ºC por espacio de 20 minutos.
2. Centrifugar a 1000 x g por espacio de 10 minutos.
3. En un tubo Falcon 2059, combinar 50 µl de supernatante con 200 µl de huésped EXP505 a OD 1.
4. Incubar en sacudidor a 37ºC por espacio de 15 minutos.
5. Añadir 300 µl de medio SOB.
6. Incubar en sacudidor a 37ºC por espacio de 30-45 minutos.
7. Cultivar 50 µl en una gran placa de cultivo LB_{Kan50} usando perlas de vidrio estériles. Si las placas están "secas", puede añadirse más medio SOB para ayudar a desembolsar las células.
8. Incubar la placa de cultivo a 30ºC por espacio de al menos 24 horas.
9. Cultivar un aislado para secuenciación y/o RFLP.

[0292] El cultivo a 30ºC reduce el número de copias de plásmido y se usa para mitigar la aparente toxicidad de algunos clones de amilasa.

a. Contaminantes que Forman Zonas de Inhibición en Almidón Rojo

[0293] Cuando se usa almidón rojo sobre medio sólido para el análisis de fago para determinar la actividad de amilasa, es común ver unidades formadoras de colonias (cfu) contaminantes que forman zonas de inhibición en el almidón rojo. Para las placas inoculadas, es importante distinguir los clones de fago amilasa-positivos de estos contaminantes cuando se alineen con una determinada posición de pocillo. La fuente de los microbios contaminantes es presumiblemente la solución concentrada de almidón rojo al 2%, que no puede ser esterilizada mediante autoclaveado o bien mediante autofiltración tras la preparación. Se piensa que se trata de organismos oportunistas que sobreviven metabolizando el almidón rojo. A fin de reducir estos contaminantes, usar técnica estéril al hacer las soluciones de almidón rojo al 2% y almacenar las soluciones concentradas ya sea a 4ºC o bien en hielo. Reponer las soluciones concentradas cada 1-2 semanas o siempre que se observe una alta incidencia de contaminantes.

Ejemplo 5

Reensamblaje por Ligación de Amilasas

[0294] Fueron amplificados nueve fragmentos (cada uno de aproximadamente 150 bp) de cada uno de los clones progenitores ID SEC Nº: 6, ID SEC Nº: 66 e ID SEC Nº: 67, cubriendo todo el marco de lectura abierto. Se indican en la Tabla 3 los cebadores.

TABLA 3

	ID SEC Nº:	
GAACACTAGTAGGAGGTAACTTATGGCAAAGTATTCCGAGCTCGAAG	11	SpeI
GAACGGTCTCATTCCGCCAGCCAGCAAGGGGATGAGCGG	12	BsaI
GAACCGTCTCAAAACACGGCCCATGCCTACGGC	13	BsmBI
GAACGTCTCACCTCGACTTCCACCCCAACGAGGTCAAG	14	BsmAI
GAACGTCTCAGGCGCTTTGACTACGTGAAGGGC	15	BsmAI
GAACGGTCTCAACAAGATGGATGAGGCCTTTG	16	BsaI
GAACCGTCTCACGATATAATCTGGAACAAGTACCTTGC	17	BsmBI
GAACCGTCTCAGAAGCACGAGCATAGTTTACTACG	18	BsmBI
GAACCGTCTCAAAGGTGGGTTTATGTGCCG	19	BsmBI
GAACGTCTCAGGAATCCAAATGGCGGATATTCCCGC	20	BsmAI
GAACGGTCTCAGTTTATCATATTGATGAGCTCC	21	BsaI
GAACCGTCTCAGAGGTAGTTGGCAGTATATTTG	22	BsmBI
GAACGTCRCACGCCAGGCATCAACGCCGATG	23	BsmAI
GAACGTCTCATTGTAGTAGAGCGGGAAGTC	24	BsmBI
GAACGGTCTCAATCGGTGTCGTGGTTTGCTAC	25	BsaI
GAACCGTCTCACTTCCACCTGCGAGGTGGTC	26	BsmBI
GAACCGTCTCACCTTCCAACCTTGCTCGAGC	27	BsmBI
TCGAGACTGACTCTCACCCAACACCGCAATAGC	28	
GAACACTAGTAGGAGGTAACTTATGGCCAAGTACCTGGAGCTCGAAG AGG	29	SpeI
GAACGGTCTCATTCCCCCGGCGAGCAAGGGC	30	BsaI
GAACCGTCTCAAAACACCGCCCACGCCTACGG	31	BsmBI
GAACGTCTCACCTCGACTTCCACCCCAAC	32	BsmAI
GAACGTCTCAGGCGCTTCGACTACGTCAAGG	33	BsmAI
GAACGGTCTCAACAAGATGGACGCGGCCTTTGAC	34	BsaI
GAACCGTCTCACGATATAATTTGGAACAAGTACCC	35	BsmBI
GAACCGTCTCAGAAGCACCGACATAGTCTAC	36	BSMBI
GAACCGTCTCAAAGGTGGGTCTACGTTCCG	37	BsmBI
GAACGTCTCAGGAATCCATATTGCGGAGATTCCGGC	38	BsmAI
GAACGGTCTCAGTTTATCATGTTCACGAGCTC	39	BsaI
GAACCGTCTCAGAGGTAGTTGGCCGTGTACTTG	40	BsmBI
GAACGTCTCAGCCATGCGTCAACGCCGATG	41	BsmAI
GAACGTCTCATTGTAGTAGAGCGGGAAGTCG	42	BsmAI
GACGGTCTCAATCGGTGTCGTGGTTTGCAACG	43	BsaI
GAACCGTCTCACTTCCACCGGCGAGGTGGTCGTG	44	BsmBI
GAACCGTCTCACCTTCCGGCCTTGCTCGAGCC	45	BsmBI
TCGAGACTGACTCTCAGCCCACCCCGCAGTAGCTC	46	
GAACACTAGTAGGAGGTAACTTATGGCCAAGTACTCCGAGCTGGAAG AGG	47	SpeI
GAACGGTCTCATTCCTCCCGCGAGCAAGGG	48	BsaI
GAACCGTCTCAAAACACCGCCCACGCCTATG	49	BsmBI
GAACGTCTCACCTCGACTTCCACCCGAACGAGC	50	BsmAI
GAACGTCTCAGGCGCTTCGACTACGTCAAGG	51	BsmAI
GAACGGTCTCAACAAGATGGACGAGGCCTTCG	52	BsaI
GAACCGTCTCACGATATAATCTGGAACAAG	53	BsmBI
GAACCGTCTCAGAAGCACTGACATCGTTTACTACG	54	BsmBI
GAACCGTCTCAAAGGTGGGTTTACGTTCCG	55	BsmBI
GAACGTCTCAGGAATCCATATCGCCGAAAT	56	BsmAI
GAACGGTCTCAGTTTATCATGTTTATGAGC	57	BsaI
GAACCGTCTCAGAGGTAGTTGGCCGTGTATTTAC	58	BsmBI
GAACGTCTCACGCCAGGCATCGATGCCGAT	59	BsmAI
GAACGTCTCATTGTAGTAGAGGGCGAAGTCAAAG	60	BsmAI
GAACGGTCTCAATCGGTATCGTGGTTGGCTACAAAC	61	BsaI
GAACCGTCTCACTTCCTCCGGCGAGGTTGTCATG	62	BsmBI
GAACCGTCTCACCTTCCGGCTTTGCTTGAGGC	63	BsmBI
TCGAGACTGACTCTCACCCAACACCGCAGTAGCTCC	64	
CACACAGCAGCAACCAACCTCGAGACTGACTCTCASCC	65	BbvI

[0295] Las condiciones que se usaron para la PCR fueron las siguientes: 3 min. a 94ºC (30 seg. a 94ºC; 30 seg. a 55ºC; 30 seg. a 68ºC) x 30 ciclos, seguidos por 10 min. a 68ºC. Los productos de PCR correspondientes a regiones homólogas de los tres progenitores fueron reunidos (1:1:1), cortados con la apropiada enzima de restricción (véase la Tabla 3) y purificados en gel. Cantidades iguales de los combinados de fragmentos fueron combinadas y ligadas (a 16ºC; durante la noche). Los resultantes productos de ligación de 450 bp fueron purificados en gel y ligados para así obtener genes de amilasa de longitud completa. Los productos de longitud completa resultantes fueron purificados en gel y amplificados por PCR usando una mezcla de cebadores F1 ID SEC Nº: 6, ID SEC Nº: 66, ID SEC Nº: 67 y cebador ID SEC Nº: 65. Las condiciones que se usaron para la PCR fueron las siguientes: 3 min. a 94ºC (30 seg. a 94ºC; 30 seg. a 50ºC; 60 seg. a 68ºC) x 30 ciclos, seguidos por 10 min. a 68ºC. Los productos de PCR resultantes (~ 1,4 kbp) fueron purificados, cortados con *SpeI* y *BbvI*, purificados en gel, ligados en pMYC (vector de Mycogen, cortado con SpeI/XhoI) y transformados en Top10 de *E. coli*. DNA plásmido de un combinado de ~ 21000 colonias fue aislado y transformado en *Pseudomonas*.

Cribaje de la α-Amilasa Reensamblada

[0296] Las *Pseudomonas fluorescens* transformadas (MB214) que contenían pMYC derivado de los clones progenitores ID SEC Nº: 6, ID SEC Nº: 66 e ID SEC Nº: 68 fueron clasificadas en placas de 96 o 384 pocillos por FACS y tratadas con urea 6M. El cribaje primario usando RBB-almidón y/o FITC-almidón como sustratos fue realizado como se describe más ampliamente más adelante. Los clones activos elevados fueron cribados usando RBB-almidón como sustrato usando cultivos inducidos y mediante ensayo de licuefacción. Se realizó acopio y secuenciación de los nuevos clones activos elevados sobre la base de los datos de licuefacción.

[0297] La biblioteca de amilasa reensamblada transformada (MB214 (Pf)) fue recolectada y clasificada en placas de 96 pocillos (o placas de 384 pocillos) a razón de 1 célula/pocillo en 50 µl de LB+Tet. Las placas de cultivo fueron incubadas por espacio de 24 horas a 30ºC. Se hicieron placas replicadas correspondientes a cada pocillo para almacenamiento. Fueron añadidos cuarenta y cinco (45) µl de urea 12M a cada pocillo y las placas de cultivo fueron sacudidas por espacio de 10 minutos. Las placas de cultivo fueron mantenidas a temperatura ambiente por espacio de al menos 1 hora y el lisado se almacenó a 4ºC.

Ensayo Usando RBB-Almidón

[0298] Fueron añadidos a cada pocillo de una nueva placa de 96 pocillos (de fondo en V) 75 µl de sustrato de RBB-almidón (almidón de maíz insoluble en RBB al 1% en tampón de NaAc 50mM, pH = 4,5). Cinco microlitros de lisado enzimático fueron transferidos a cada pocillo con sustrato usando Biomek o Zymark. Las placas de cultivo fueron selladas con cinta selladora de aluminio y sacudidas brevemente en el sacudidor. Las placas de cultivo fueron incubadas a 90ºC por espacio de 30 minutos, efectuándose a continuación enfriamiento a temperatura ambiente por espacio de un periodo de tiempo de aproximadamente 5 a 10 minutos. Fueron añadidos a cada pocillo cien microlitros de etanol al 100%, y las placas de cultivo fueron selladas y sacudidas brevemente en el sacudidor. Las placas de cultivo fueron luego centrifugadas a 4000 rpm por espacio de 20 minutos usando una centrifuga de sobremesa. 100 µl del supernatante fueron transferidos a una nueva placa de 96 pocillos (de fondo plano) mediante Biomek y dieron una lectura OD_{595}. Controles: ID SEC Nº: 6, ID SEC Nº: 66, ID SEC Nº: 68.

Ensayo Usando FITC-Almidón

[0299] Se añadieron 50 µl de sustrato (FITC-almidón al 0,01% en tampón de NaAc 100mM, pH = 4,5) a cada pocillo de una nueva placa de cultivo de 384 pocillos. Se transfirieron 5 µl de lisado enzimático a cada pocillo con sustrato y se incubó la placa de cultivo a temperatura ambiente durante la noche. Se leyó para cada pocillo el cambio de polarización del sustrato, con excitación a 485 nm y emisión a 535 nm. Controles: ID SEC Nº: 6, ID SEC Nº: 66, ID SEC Nº: 68. Preferiblemente se usaron para todos los ensayos placas de 96 pocillos.

Confirmación de los Nuevos Clones Activos

[0300] Cada clon positivo del cribaje fue cultivado e inducido usando un protocolo estándar. Cada clon fue examinado para determinar el crecimiento (es decir, la densidad de células a lo largo del tiempo), la actividad al nivel de cada célula (ensayo del RBB-almidón y ensayo de licuefacción), la expresión (gel de proteína) y la solubilidad de la proteína (mediante análisis microscópico). Los nuevos clones elevados confirmados fueron transferidos para fermentación. Uno de los clones que tienen actividad de amilasa mayor que la del progenitor incluye la enzima que se expone en la ID SEC Nº: 2 (codificada por la ID SEC Nº: 1).

Ejemplo 6

Licuefacción por Amilasas

[0301] Este ejemplo caracteriza los patrones de oligosacáridos de licuefacción resultantes de las enzimas de la invención, ID SEC Nº: 6 e ID SEC Nº: 2, frente a las amilasas comerciales de *Bacillus licheniformis* y *Bacillus*

stearothermophilus. Estos resultados quedaron de manifiesto en el progreso de la sacarificación y en los niveles finales de dextrosa de jarabes generados mediante amilasas comerciales y de la invención.

[0302] Tres enzimas comerciales, la Spezyme AA de Genencor y otras dos, requirieron todas ellas más del doble de la dosificación recomendada para alcanzar el equivalente de dextrosa (DE) perseguido. El equivalente de dextrosa (DE) es el estándar industrial para medir la concentración de azúcares reductores totales, calculados como D-glucosa sobre la base del peso en seco. El almidón granular no hidrolizado tiene un DE de virtualmente cero, mientras que el DE de la D-glucosa es de 100 por definición.

[0303] Esto confirma el efecto de "doble dosificación" para todas las amilasas de Bacillus y da más credibilidad a la propuesta de que la dosificación observada para la ID SEC Nº: 2 en las pruebas es también el doble del valor que se requeriría en condiciones más normales. La dosificación "normal" proyectada de 60-70 unidades/kilo de almidón a pH 4,5 para alcanzar un DE de 19 es consistente con los datos de licuefacción de laboratorio.

[0304] El patrón de oligosacáridos generado por las amilasas de la invención es distinto de los perfiles para las amilasas de Bacillus. La distribución del peso molecular para las amilasas de Bacillus (cromatografía de permeación en gel con detección por difusión de luz y viscosidad) es bimodal con una considerable fracción dentro de la gama de pesos moleculares muy altos (> 300.000) incluso para un DE de 18. La ID SEC Nº: 2 con un DE de 18 presenta una distribución uniforme con nada de más de 20.000. Esto es consistente con la más baja viscosidad para los jarabes de la invención. Los perfiles de los DP (grados de polimerización) según medición efectuada por HPLC también reflejan esta diferencia en el patrón de actuación.

[0305] En este estudio, así como en los estudios anteriores, el nivel de glucosa final tras sacarificación de los jarabes licuados con amilasas de la invención frente a los jarabes correspondientes a Bacillus es el mismo para ambos casos. Sin embargo, hemos adquirido ahora suficientes datos de sacarificación de estudios internos así como de estudios de GPC para confirmar que los residuos no de dextrosa para las amilasas de la invención son distintos de los jarabes de amilasa de Bacillus. A pesar de que los niveles de dextrosa y maltosa son esencialmente iguales para ambas, las amilasas de la invención tienen una más alta fracción de DP3 pero una menor cantidad de las "más altas" en comparación con la enzima de Bacillus. De manera consistente con la ausencia de fragmentos de alto peso molecular tras la licuefacción, los jarabes de la invención post-sacarificación tienen un más bajo contenido de la fracción > DP7.

	Glucosa	**DP2**	**DP3**	**>DP7**
ID SEC Nº: 2	95,25	2,39	1,13	0,91
Comercial	94,16	2,10	0,39	2,91
ID SEC Nº: 6	94,77	2,27	1,48	0,82

[0306] El concentrado de amilasa de la ID SEC Nº: 2 fue preparado a partir de caldos de fermentación mediante tratamiento térmico, lavado celular y extracción alcalina usando microfiltración y ultrafiltración (rendimiento total del 48%). El concentrado de UF fue neutralizado con ácido acético y hecho con un 30% de glicerol a pH 4,5. El nivel de actividad de la formulación en forma de lechada era representativo de un producto comercial (120 U^1/g = 0,5 kg/tonelada de almidón).

Ensayo de la Actividad de Amilasa Estándar

[0307] Una cubeta de 1 ml que contenía 950 µl de MOPS 50mM pH 7,0 que contenía PNP-α-D-hexa-(1→4)-glucopiranósido 5mM se puso en el controlador de temperatura Peltier del espectrofotómetro Beckman DU-7400 precalentado hasta 80ºC. El espectrofotómetro fue puesto a cero a 405 nm, y se añadieron a la cubeta 50 µl de la solución enzimática, se efectuó mezcla a fondo y se supervisó el incremento de la OD_{450nm} a lo largo de un intervalo de tiempo de 1 minuto. La tasa $\Delta OD_{405nm/min}$ se convierte a unidades estándar de µmoles/minuto a partir de la respuesta a una OD_{405nm} de 50 µl de 1 µmol/ml de PNP en 950 ml de MOPS 50mM a un pH de 7,0 y a -80ºC. Una unidad Diversa estándar de alfa-amilasa 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 la Actividad de Glucoamilasa Estándar

[0308] Una cubeta de 1 ml que contenía 950 µl de MOPS 50mM pH 7,0 que contenía pNP-α-D-glucopiranósido 5mM fue puesta en el controlador de temperatura Peltier del espectrofotómetro Beckman DU-7400 precalentado hasta 60ºC. El espectrofotómetro fue puesto a cero a 405 nm y se añadieron a la cubeta 50 µl de la solución enzimática, se efectuó mezcla a fondo y se supervisó el incremento de la OD_{405nm} a lo largo de un intervalo de tiempo de un minuto. La tasa ΔOD_{405nm}/min. se convierte en unidades estándar de pmoles/minuto a partir de la respuesta a una OD_{405nm} de 50 µl de 1 µmol/ml de pNP en 950 ml de MOPS 50mM a pH 7,0 y -60ºC. Una unidad Diversa estándar de glucoamilasa (DGA) es igual a la cantidad de enzima que catalizará la liberación de 1 µmol/ml/minuto de pNP bajo las condiciones definidas del ensayo.

Determinación del Equivalente de Dextrosa

[0309] Para medir el DE se usó el método de la neocuproína. Las muestras seleccionadas fueron medidas tanto mediante el procedimiento de la invención como por un analista de GPC (GPC = cromatografía de permeación en gel) usando el procedimiento de GPC de Fehling.

Ensayo de la Neocuproína

[0310] Una muestra de 100 µl fue añadida a 2,0 ml de solución A de neocuproína (40 g/l de carbonato sódico, 16 g/l de licina, 0,45 g/l de sulfato de cobre). A esto le fueron añadidos 2,0 ml de solución B de neocuproína (1,2 g/l de neocuproína clorhidrato Sigma N-1626). Los tubos fueron sometidos a mezcla y calentamiento en un baño de agua hirviendo por espacio de 12 minutos y fueron enfriados y sus contenidos fueron diluidos hasta un volumen de 10 ml con agua desionizada, y la OD fue leída a 450 nm en el espectrofotómetro. El equivalente de glucosa de la muestra fue extrapolado a partir de la respuesta de un patrón de 0,2 mg/ml de glucosa sometido simultáneamente a ensayo.

[0311] Se diluye la muestra de almidón aproximadamente de 1 a 16 veces con agua desionizada, registrándose la dilución exacta. Diez mililitros de la muestra diluida fueron añadidos a 20 ml de agua desionizada. Se añadieron al almidón diluido diez mililitros de solución A y B de Fehling. La muestra fue hervida por espacio de 3 minutos y enfriada sobre hielo. Fueron añadidos diez mililitros de KI al 30% y 10 ml de H_2SO_4 6N. La solución fue titrada contra tiosulfato sódico 0,1N. El volumen de titrante se registra y se usa para calcular el DE.

Determinación del Almidón Residual

[0312] Muestras post-sacarificación fueron sometidas a ensayo para determinar el almidón residual usando el procedimiento del yodo de Staley.

[0313] Se pesaron en un gran platillo de pesaje veinte gramos de muestra. Se añaden al platillo de pesaje 45 µl de solución de yodo, y se mezcla a fondo la solución de almidón. El color azul oscuro indica la presencia de almidón; un verde azulado claro indica una ligera presencia de almidón; el color verde claro indica la presencia de trazas de almidón; y el color rojo amarillento indica la ausencia de almidón. La solución de yodo se prepara disolviendo 21,25 gramos de yodo y 40,0 gramos de yoduro potásico en un litro de agua.

Perfil de Oligosacáridos

[0314] Los perfiles de carbohidratos de licuefacción y sacarificación fueron medidos por HPLC (columna Bio-Rad Aminex HPX-87C en forma de calcio - 80ºC) usando detección del índice refractivo.

Cromatografía de Permeación en Gel

[0315] La distribución del peso molecular fue determinada por cromatografía en una columna PL Aquagel-OH con detección de masa por índice refractivo (Modelo 2410 de Waters). Se usó un detector Viscotek Modelo T60 para mediciones continuas de la viscosidad y la difusión de luz.

Electroforesis Capilar

[0316] Sistema Beckman Coulter P/ACE MDQ para Glicoproteínas - separación de oligosacáridos derivatizados con APTS (APTS = aminopropiltrietoxisilano) en un capilar de sílice fundida - detección por fluorescencia inducida por láser.

Licuefacción Primaria

[0317] Almidón en línea procedente directamente del proceso de GPC es bombeado al interior de un tanque de alimentación de 60 litros donde pueden ajustarse antes de la licuefacción el pH, los DS (sólidos secos) y el nivel de calcio. Se añade la amilasa a la lechada. La lechada con un 32% de DS es bombeada a razón de 0,7 litros/minuto por una bomba de desplazamiento positivo al chorro - una cámara de mezcla a presión donde la lechada de almidón es calentada instantáneamente hasta una temperatura de más de 100ºC mediante inyección de vapor. El almidón gelatinizado parcialmente licuado es bombeado a través de una red de tubería (aún bajo presión) para así tener el deseado tiempo de permanencia (5 minutos) a temperatura. Se descarga la presión al interior de un tanque de vaporización rápida, y pueden tomarse muestras. Las muestras se tomaron por duplicado.

Licuefacción Secundaria

[0318] El almidón licuado fue recogido en botellas de vidrio de un litro y mantenido en un baño de agua a 95ºC por espacio de 90 minutos.

Sacarificación

[0319] El almidón licuado fue enfriado hasta 60ºC, el pH fue ajustado a 4,5, y las muestras fueron tratadas con glucoamilasa. El progreso de la sacarificación fue supervisado a lo largo del tiempo mediante HPLC.

Preparación de Jarabes Licuados para Análisis

[0320] Estas pruebas se hicieron para obtener jarabes licuados a niveles de DE de ~ 12 y 18 para tres amilasas: ID SEC Nº: 6, ID SEC Nº: 2 y *amilasa de B. licheniformis* comercial. Los jarabes fueron sacarificados con tres niveles de glucoamilasa. Los jarabes licuados fueron también analizados por HPLC y cromatografía de permeación en gel.

Sacarificación

[0321] Los jarabes licuados producidos con cada amilasa fueron ajustados a un pH de aproximadamente 2,5 con HCl 6N inmediatamente después de la licuefacción secundaria de 90 minutos para inactivar a toda amilasa residual. Los jarabes fueron luego ajustados a un pH de 4,5, fueron puestos en un baño de agua a 60ºC y fueron sacarificados con tres niveles de glucoamilasa. El grado de sacarificación fue supervisado por HPLC en puntos en el tiempo que van desde las 18 hasta las 88 horas.

[0322] Los jarabes licuados fueron sacarificados con la dosificación estándar - un 0,04% de una glucoamilasa de doble fortaleza - y con dos dosificaciones más bajas (de un 50% y un 25%) para supervisar cualesquiera diferencias en el progreso de la sacarificación. Como indican los gráficos anteriores y las tablas que aparecen a continuación, los niveles de glucosa son más altos en los puntos más tempranos en el tiempo con los jarabes de la ID SEC Nº: 2. Algo de la diferencia se debe a un punto de partida más alto, pero un factor que también contribuye es la diferencia entre los perfiles del peso molecular (al ser más pequeños y más uniformes, los oligosacáridos de los jarabes licuados con la ID SEC Nº: 2 deberían ser y aparentemente son mejores sustratos para glucoamilasa).

Progreso de la Sacarificación - Evolución del % de Dextrosa Referido al Tiempo - Glucoamilasa al 0,04%

[0323]

Amilasa	18 h	24 h	40 h	44 h	88 h
Comercial	70,2	78,4	86,1	86,7	94,2
ID SEC Nº: 2	79	88,6	92,5	92,8	95,3
ID SEC Nº: 6	74,1	85,9	91,9	91,6	94,8

Progreso de la Sacarificación - Evolución del % de Dextrosa Referido al Tiempo - Glucoamilasa al 0,02%

[0324]

Amilasa	18 h	24 h	40 h	44 h	88 h
Amilasa de B. licheniformis	54,5	66,7	76,1	77,2	90,9
ID SEC Nº: 2	60,1	72	84,8	85,3	93,6
ID SEC Nº: 6	57,1	70	84	86,5	92,5

Perfil de Azúcares Post-sacarificación

[0325] En estos estudios y en todos los anteriores estudios de sacarificación, el nivel de glucosa final que se alcanza tras la sacarificación de jarabes licuados con amilasa de la invención y con amilasa de B. licheniformis es esencialmente idéntico. Es también similar el nivel de DP2 (maltosa). Estos grandes fragmentos son malos sustratos para glucoamilasa y tienden a convertirse lentamente, en caso de hacerlo en absoluto, en fragmentos más pequeños y finalmente en glucosa.

	Glucosa	DP2	DP3	>DP7
ID SEC Nº: 2	95,25	2,39	1,13	0,91
Comercial	94,16	2,10	0,39	2,91
ID SEC Nº: 6	94,77	2,27	1,48	0,82

Distribución del Peso Molecular

[0326] La distribución del peso molecular de los jarabes licuados hasta unos DE's de 12 y 18 mediante amilasas de la invención (ID SEC Nº: 6 e ID SEC Nº: 2) y amilasa de Bacillus licheniformis comercial y amilasa de Bacillus stearothermophilus comercial fueron medidos mediante cromatografía de permeación en gel usando detección por índice refractivo, difusión de luz y viscosidad. Ambas amilasas de licheniformis y de stearothermophilus generan una distribución bimodal - estando el pico primario centrado en 2000 y estando un pico secundario en 32.000 con un punto de inflexión que se extiende hasta más allá del nivel de 160.000. El peso molecular más bajo representa aproximadamente un 60% de la masa total de la muestra. Las amilasas de la invención presentan un único pico en 2000 con muy poca cosa por encima de 30.000.

HPLC

[0327] Los jarabes con un DE de 12 y 18 producidos mediante las amilasas comerciales y de la invención fueron analizados por HPLC. Ambas técnicas (cromatogramas/electroferogramas que se muestran en los Apéndices) producen huellas características de cada clase de amilasa; siendo los patrones de oligosacáridos distintos para las amilasas de licheniformis en comparación con las amilasas de stearothermophillus en comparación con las amilasas de la invención. Los jarabes licuados con amilasas de la invención presentan evidencia de mayor ramificación en los oligosacáridos. Los análisis por HPLC tan sólo resuelven los oligosacáridos dentro de la gama de valores del grado de polimerización < DP 25; no siendo visibles en estas técnicas los fragmentos mayores. Se sabe de las amilasas de Bacillus que licuan el almidón de forma tal que la fracción de amilopectina es hidrolizada menos extensivamente que la fracción de amilosa. Estos fragmentos de amilopectina > DP 30 están contenidos en la fracción de alto peso molecular centrada en 32.000, y en consecuencia se ve poca evidencia de ramificación en los análisis por HPLC de los jarabes licuados con amilasas de Bacillus. Los oligosacáridos < DP 15 derivados de las amilasas de la invención contienen fragmentos tanto de amilosa como de amilopectina.

[0328] A pesar de que la invención ha sido descrita con referencia a la realización actualmente preferida, debe entenderse que pueden hacerse varias modificaciones sin por ello salir fuera del espíritu de la invención. En consecuencia, la invención queda limitada tan sólo por las reivindicaciones siguientes.

REIVINDICACIONES

1. Ácido nucleico aislado, sintético o recombinante que comprende una secuencia seleccionada de entre los miembros del grupo que consta de:

(a) la ID SEC Nº: 1,
(b) que codifica un polipéptido que tiene una secuencia como la expuesta en la ID SEC Nº: 2, o un fragmento enzimáticamente activo que tiene actividad de amilasa de la ID SEC Nº: 2 de una longitud de al menos 75, 100 o 150 residuos de aminoácidos;
(c) que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y (I) tiene al menos una identidad secuencial de un 90% en toda la longitud de la ID SEC Nº: 1, o (II) tiene una longitud de al menos 200 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 1, o (III) tiene una longitud de al menos 150 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 97% con la ID SEC Nº: 1, o (IV) tiene una longitud de al menos 75 nucleótidos de la ID SEC Nº: 1 y tiene una identidad secuencial de al menos un 99% con la ID SEC Nº: 1;
(d) que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y (I) tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o (II) tiene un fragmento enzimáticamente activo de (I) y al menos una longitud de 150 residuos de aminoácidos y tiene una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o un fragmento enzimáticamente activo de (I) una longitud de al menos 75 residuos de aminoácidos y tiene una identidad secuencial de al menos un 98% con la ID SEC Nº: 2; o bien
(e) secuencias plenamente complementarias de (a), (b), (c) o (d),
en donde la identidad secuencial se determina usando un algoritmo de comparación de secuencias que consta del algoritmo FASTA, versión 3.0t78, con los parámetros por defecto.

2. Polipéptido aislado, sintético o recombinante seleccionado de entre los miembros del grupo que consta de:

(a) una secuencia de aminoácidos como la expuesta en la ID SEC Nº: 2 o un fragmento enzimáticamente activo de la ID SEC Nº: 2 que tenga actividad de amilasa de la ID SEC Nº: 2 de una longitud de al menos 75, 100 o 150 residuos de aminoácidos;
(b) una secuencia de aminoácidos codificada por un ácido nucleico de la reivindicación 1 (a) a 1 (d) que tenga actividad de amilasa de la ID SEC Nº: 2;
(c) una secuencia de aminoácidos (I) que tenga al menos una identidad secuencial de un 95% a todo lo largo de la ID SEC Nº: 2; o (II) que tenga un fragmento enzimáticamente activo de (I) una longitud de al menos 150 residuos de aminoácidos y tenga una identidad secuencial de al menos un 95% con la ID SEC Nº: 2, o un fragmento enzimáticamente activo de (I) una longitud de al menos 75 residuos de aminoácidos y que tenga una identidad secuencial de al menos un 98% con la ID SEC Nº: 2, en donde el polipéptido de (I) o (II) tiene una actividad de amilasa de la ID SEC Nº: 2;
en donde la identidad secuencial se determina usando un algoritmo de comparación de secuencias que consta del algoritmo FASTA, versión 3.0t7, con los parámetros por defecto

3. Ácido nucleico según la reivindicación 1 que codifica un polipéptido que tiene una secuencia de aminoácidos de (a), (b), (c), (d) o (e) de la reivindicación 1, que comprende una secuencia señal o un péptido o polipéptido heterólogo.

4. Polipéptido según la reivindicación 2 que comprende una secuencia de aminoácidos de (a), (b) o (c) de la reivindicación 2 que comprende una secuencia señal o un péptido o polipéptido heterólogo.

5. Anticuerpo aislado o recombinante que se une específicamente a un polipéptido seleccionado de entre los miembros del grupo que consta de:

(a) un polipéptido como el que se expone en la reivindicación 2; y
(b) un polipéptido codificado por un ácido nucleico como el que se expone en la reivindicación 1.

6. El anticuerpo de la reivindicación 5, en donde los anticuerpos son policlonales o monoclonales.

7. Método que es para producir un polipéptido y comprende los pasos de introducir en una célula huésped un ácido nucleico que codifica un polipéptido, y cultivar la célula huésped bajo condiciones que permitan la expresión del polipéptido, en donde el polipéptido es seleccionado de entre los miembros del grupo que consta de: (a) un polipéptido como el que se expone en la reivindicación 2; y (b) un polipéptido codificado por un ácido nucleico como el que se expone en la reivindicación 1.

8. Método que es para generar un polinucleótido variante y comprende los pasos de:

obtener un ácido nucleico que comprenda a un polinucleótido como el que se expone en la reivindicación 1 o que codifique un polipéptido como el que se expone en la reivindicación 2; y

modificar uno o varios nucleótidos de dicho polinucleótido para convertirlo en otro nucleótido, delecionar uno o varios nucleótidos de dicho polinucleótido, o añadir uno o varios nucleótidos a dicho polinucleótido.

9. El método de la reivindicación 8, en donde las modificaciones son introducidas por un método que se selecciona de entre los miembros del grupo que consta de PCR sujeta a error, reordenamiento, mutagénesis dirigida por oligonucleótidos, PCR de ensamblaje, mutagénesis por PCR sexual, mutagénesis in vivo, mutagénesis en casete, mutagénesis de ensamblaje recursivo, mutagénesis de ensamblaje exponencial, mutagénesis específica de sitio, reensamblaje de genes, mutagénesis a saturación en sitio génico y cualquier combinación de estos métodos.

10. Método que es para hidrolizar un enlace de almidón o un enlace glicosídico en una sustancia que contiene un enlace glicosídico y comprende el paso de poner a una sustancia que contenga el enlace de almidón o enlace glicosídico en contacto con un polipéptido seleccionado de entre los miembros del grupo que consta de

(a) un polipéptido capaz de hidrolizar un enlace de almidón y codificado por un polinucleótido que tiene una secuencia como la que se expone en la reivindicación 1; y
(b) un polipéptido capaz de hidrolizar un enlace de almidón y que tiene una secuencia como la que se expone en la reivindicación 2
bajo condiciones que facilitan la hidrólisis del enlace de almidón.

11. El método de la reivindicación 10, en donde el método comprende el paso de catalizar la hidrólisis de un enlace alfa-1,4-glucosídico, que comprende el paso de poner a una muestra que contiene una composición que comprende un enlace alfa-1,4-glucosídico en contacto con un polipéptido que tiene una secuencia seleccionada de entre los miembros del grupo que consta de:

(a) un polipéptido que es capaz de hidrolizar un enlace alfa-1,4-glucosídico y es codificado por un polinucleótido que tiene una secuencia como la que se expone en la reivindicación 1; y
(b) un polipéptido que es capaz de hidrolizar un enlace alfa-1,4-glucosídico y que tiene una secuencia como se expone en la reivindicación 2
bajo condiciones que faciliten la hidrólisis del enlace alfa-1,4-glucosídico.

12. Ensayo para identificar a un polipéptido enzimáticamente activo, comprendiendo dicho ensayo los pasos de:

poner a un polipéptido que tenga una secuencia como la que se expone en la reivindicación 2 en contacto con una molécula de sustrato bajo condiciones que le permitan al polipéptido funcionar; y
detectar ya sea una disminución de la cantidad del sustrato o bien un incremento de una cantidad de un producto de reacción que resulta de una reacción entre dicho polipéptido y dicho sustrato; en donde una disminución de la cantidad del sustrato o un incremento de la cantidad del producto de reacción es indicativa(o) de un polipéptido funcional.

13. Sonda de ácido nucleico que comprende una secuencia de ácido nucleico como la que se expone en la reivindicación 1 o codifica un polipéptido como el que se expone en la reivindicación 2.

14. La sonda de la reivindicación 13, en donde es oligonucleótido es DNA o RNA o un ácido nucleico peptídico (PNA).

15. Sonda de la reivindicación 13 o 14, en donde la sonda comprende además una etiqueta isotópica detectable, o bien comprende además una etiqueta no isotópica detectable seleccionada de entre los miembros del grupo que consta de una molécula fluorescente, una molécula quimioluminiscente, una enzima, un cofactor, un sustrato enzimático y un hapteno.

16. Preparación proteica que comprende un polipéptido seleccionado de entre los miembros del grupo que consta de:

(a) un polipéptido codificado por un polinucleótido que tiene una secuencia como la que se expone en la reivindicación 1; y
(b) un polipéptido que tiene una secuencia como la que se expone en la reivindicación 2.

17. La preparación proteica de la reivindicación 16, en donde la preparación proteica comprende un sólido o un líquido.

18. Vector de clonación que comprende un polinucleótido que tiene una secuencia seleccionada de entre los miembros del grupo que consta de:

(a) una secuencia de ácido nucleico como la que se expone en la reivindicación 1; y
(b) una secuencia de ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2.

19. Célula huésped que comprende un vector de clonación como el que se reivindica en la reivindicación 18.

20. Vector de expresión que es capaz de replicarse en una célula huésped y comprende un polinucleótido que tiene una secuencia seleccionada de entre los miembros del grupo que consta de

(a) una secuencia de ácido nucleico como la que se expone en la reivindicación 1; y
(b) una secuencia de ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2.

21. Vector como el que se reivindica en las reivindicaciones 18 a 20, en donde el vector es seleccionado de entre los miembros del grupo que consta de vectores virales, vectores plasmídicos, vectores de fago, vectores de fagémido, cósmidos, fósmidos, bacteriófagos, cromosomas artificiales, vectores de adenovirus, vectores retrovirales y vectores virales adeno-asociados.

22. Célula huésped que comprende un vector de expresión como el que se reivindica en la reivindicación 21.

23. Célula huésped como la que se reivindica en la reivindicación 19 o 22, en donde la célula huésped es una célula procariótica, una célula eucariótica, una célula fungal, una célula de levadura, una célula de insecto, una célula de mamífero o una célula vegetal.

24. El método de la reivindicación 10, en donde el método comprende el paso de licuar una composición con contenido de almidón, que comprende el paso de poner al almidón en contacto con un polipéptido

(a) codificado por una secuencia de ácido nucleico como la que se expone en la reivindicación 1 o por un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) como el que se expone en la reivindicación 2.

25. Jarabe licuado que es producido por el método de la reivindicación 24 y tiene una relación Mn/Mw de menos de aproximadamente 2,0 al nivel de un DE de 18 y de menos de aproximadamente 4,0 al nivel de un DE de 12.

26. El método de la reivindicación 10, en donde el método comprende el paso de lavar un objeto, que comprende el paso de poner a dicho objeto en contacto con un polipéptido de la reivindicación 2 bajo condiciones suficientes para dicho lavado.

27. El método de la reivindicación 10, en donde el método comprende el paso de desencolar un textil, que comprende el paso de poner a dicho textil en contacto con un polipéptido de la reivindicación 2 bajo condiciones suficientes para dicho desencolado.

28. El método de la reivindicación 10, en donde el método comprende un tratamiento de fibras lignocelulósicas, en donde las fibras son tratadas con un polipéptido de la reivindicación 2.

29. El método según la reivindicación 28 para el desentintado enzimático de pasta papelera reciclada, en donde el polipéptido de la reivindicación 2 es aplicado en una cantidad que es eficiente para un efectivo desentintado de la superficie de las fibras.

30. Aditivo detergente que comprende un polipéptido de la reivindicación 2.

31. El método de la reivindicación 10, en donde el método comprende el paso de producir un jarabe con alto contenido de maltosa o un jarabe con alto contenido de glucosa o un jarabe mixto, comprendiendo los pasos de:

licuar almidón usando una cantidad eficaz de un polipéptido de la reivindicación 2 para así obtener un hidrolizado de almidón soluble; y sacarificar el hidrolizado de almidón soluble, obteniéndose con ello un jarabe.

32. El método de las reivindicaciones 10 o 24, en donde el almidón es de un material seleccionado de entre los miembros del grupo que consta de arroz, arroz germinado, maíz, cebada, trigo, legumbres, batata, sorgo, centeno, bulgur o una combinación de los mismos.

33. El método de la reivindicación 10, en donde el método comprende el paso de incrementar el flujo de fluidos de producción de una formación subterránea eliminando un fluido perjudicial viscoso con contenido de almidón formado durante las operaciones de producción y que se encuentra dentro de la formación subterránea que rodea a una perforación de pozo completada, comprendiendo dicho paso los pasos de:

dejar que los fluidos de producción fluyan desde la perforación del pozo;
reducir el flujo de fluidos de producción de la formación a caudales inferiores a los previstos;
hacer un tratamiento enzimático mezclando juntamente a un fluido acuoso y un polipéptido de la reivindicación 2;
bombear el fluido de tratamiento enzimático a un punto deseado dentro de la perforación del pozo;
dejar que el tratamiento enzimático degrade al fluido perjudicial viscoso con contenido de almidón, con lo cual el fluido puede sacarse de la formación subterránea a la superficie del pozo; y

en donde el tratamiento enzimático es eficaz para atacar los enlaces alfa-glucosídicos en el fluido con contenido de almidón.

34. El método de la reivindicación 10, en donde el método comprende el paso de licuar una composición que comprende almidón, comprendiendo el método el paso de poner al almidón en contacto con un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2, o
(b) un polipéptido como el que se expone en la reivindicación 2.

35. Almidón licuado que es producido por el método de la reivindicación 34 y comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

36. El método de la reivindicación 10, en donde el método comprende el paso de hacer un aditivo alimentario, un espesante, un ingrediente para carga de bajo nivel calórico o un agente peliculígeno, comprendiendo dicho paso el paso de poner a una composición que comprende almidón en contacto con suficiente polipéptido que tenga actividad de alfa-amilasa para licuar el almidón, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

37. El método de la reivindicación 10, en donde el método comprende el paso de hacer un pienso, comprendiendo dicho paso el paso de poner a una composición de pienso que comprende almidón en contacto con una cantidad de un polipéptido que tenga actividad de alfa-amilasa suficiente para licuar almidón en el pienso, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

38. El método de cualquiera de las reivindicaciones 31, 34-36, en donde el almidón es un material seleccionado de entre los miembros del grupo que consta de arroz, arroz germinado, maíz, cebada, trigo, legumbres, batata, milo, sorgo, centeno y/o bulgur o una combinación de los mismos.

39. El método de cualquiera de las reivindicaciones 26-27, 31, 33-38, que comprende además la adición de una segunda alfa-amilasa o de una beta-amilasa o de una combinación de las mismas.

40. Pienso animal que comprende un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

41. Composición que comprende una composición de maltodextrina uniforme hecha por un método que comprende el paso de licuar una composición que comprende almidón con un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende
(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

42. Aditivo alimentario, espesante, ingrediente para carga de bajo nivel calórico o agente peliculígeno que comprende jarabes licuados, jarabes licuados tratados con carbón o jarabes licuados secados por pulverización que comprenden un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que expone en la reivindicación 2, o
(b) un polipéptido como el que se expone en la reivindicación 2.

43. Composición de pienso que comprende jarabes licuados, jarabes licuados tratados con carbón o jarabes licuados secados por pulverización que comprenden un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

44. La composición de pienso de la reivindicación 43, en donde la composición de pienso comprende un pienso que comprende almidón, o la composición de pienso comprende maíz, trigo, milo, sorgo, centeno y/o bulgur.

45. El método de la reivindicación 10, en donde el método comprende el paso de hacer un pienso, que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2, o

(b) un polipéptido como el que se expone en la reivindicación 2.

46. El método de la reivindicación 8, en donde la variación comprende a una variación en especificidad por sustrato, unión a sustrato, patrón de escisión de sustrato, estabilidad térmica, perfil de pH/actividad, perfil de pH/estabilidad, incrementada estabilidad a un bajo pH de < 6, incrementada estabilidad a un bajo pH de < 5, incrementada estabilidad a un alto pH de > 9, estabilidad frente a la oxidación, dependencia del Ca^{2+} o actividad específica, o bien es menos sensible a las condiciones depletoras de los iones de calcio.

47. El método de la reivindicación 10, en donde el método comprende el paso de hacer un alcohol, que comprende el paso de poner a una composición que comprende almidón en contacto con un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

48. El método de la reivindicación 10, en donde el método comprende a un proceso de molienda de maíz en húmedo que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

49. El proceso de molienda de maíz en húmedo de la reivindicación 48, en donde el proceso adicionalmente comprende el uso de un segundo polipéptido que tiene actividad de alfa-amilasa, o de una beta-amilasa, o de otra enzima.

50. Detergente que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

51. Proceso de horneo que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

52. Bebida que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

53. Masa, producto de panadería o producto horneado que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

54. Premezcla para hornear que comprende una harina que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

55. Aditivo para horneo que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

56. Textil que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

57. Papel, pasta papelera o papel reciclado que comprende un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

58. El método de la reivindicación 10, en donde el método comprende a un proceso de perforación que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

59. El método de la reivindicación 10, en donde el método comprende un proceso de molienda en seco que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa para la licuefacción de grano entero molido, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

60. El método de la reivindicación 10, en donde el método comprende un proceso de fermentación que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

61. El método de la reivindicación 10, en donde el método comprende un método de procesamiento de textiles que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

62. El método de la reivindicación 10, en donde el método comprende a un método de procesamiento de papel, pasta papelera o papel reciclado que comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o
(b) un polipéptido como el que se expone en la reivindicación 2.

63. El método de la reivindicación 10, en donde el método comprende a un método que es para hacer una bebida y comprende el uso de un polipéptido que tiene actividad de alfa-amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

64. El polipéptido de la reivindicación 2, en donde la secuencia señal tiene una longitud de 13 a 36 residuos de aminoácidos.

65. Método que es para hacer un aditivo alimentario, un espesante, un ingrediente para carga de bajo nivel calórico o un agente peliculígeno y comprende el uso de un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

66. Método que es para hacer un etanol combustible y comprende el uso de un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

67. Proceso de campo petrolífero que comprende el uso de un polipéptido que tiene actividad de amilasa, en donde el polipéptido comprende

(a) una secuencia de aminoácidos codificada por un ácido nucleico como el que se expone en la reivindicación 1, o un ácido nucleico que codifica un polipéptido como el que se expone en la reivindicación 2; o

(b) un polipéptido como el que se expone en la reivindicación 2.

68. Ácido nucleico aislado, sintético o recombinante según la reivindicación 1 que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y tiene una identidad secuencial de al menos un 97% con la ID SEC Nº: 2.

69. Ácido nucleico aislado, sintético o recombinante según la reivindicación 1 que codifica un polipéptido que tiene actividad de amilasa de la ID SEC Nº: 2 y tiene una identidad secuencial de al menos un 98% con la ID SEC Nº: 2.

70. Polipéptido aislado, sintético o recombinante según la reivindicación 2, que comprende una secuencia de aminoácidos que tiene una identidad secuencial de al menos un 97% dentro de toda la longitud de la ID SEC Nº: 2 y tiene actividad de amilasa de la ID SEC Nº: 2.

71. Polipéptido aislado, sintético o recombinante según la reivindicación 2, que comprende una secuencia de aminoácidos que tiene una identidad secuencial de al menos un 98% dentro de toda la longitud de la ID SEC Nº: 2 y tiene actividad de amilasa de la ID SEC Nº: 2.

72. Polipéptido aislado, sintético o recombinante según la reivindicación 2, que carece de secuencia señal.

73. Polipéptido aislado, sintético o recombinante según la reivindicación 2 que carece de secuencia señal y adicionalmente comprende una secuencia peptídica o polipeptídica heteróloga, en donde la secuencia peptídica o polipeptídica heteróloga preferiblemente comprende a una secuencia conductora, una secuencia secretora o una secuencia proproteína.

74. Polipéptido aislado, sintético o recombinante según la reivindicación 2, 72 o 73, que comprende una modificación que comprende a los miembros del grupo que consta de glicosilación, acetilación, acilación, ribosilación del ADP, amidación, unión covalente de flavina, unión covalente de una mitad heme, unión covalente de un nucleótido o un derivado nucleotídico, unión covalente de un lípido o un derivado lipídico, unión covalente de un fosfatidilinositol, ciclización reticulante, formación de enlaces disulfuro, desmetilación, formación de uniones intermoleculares covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicoxilación, un anclaje GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, pegilación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación o sulfatación.

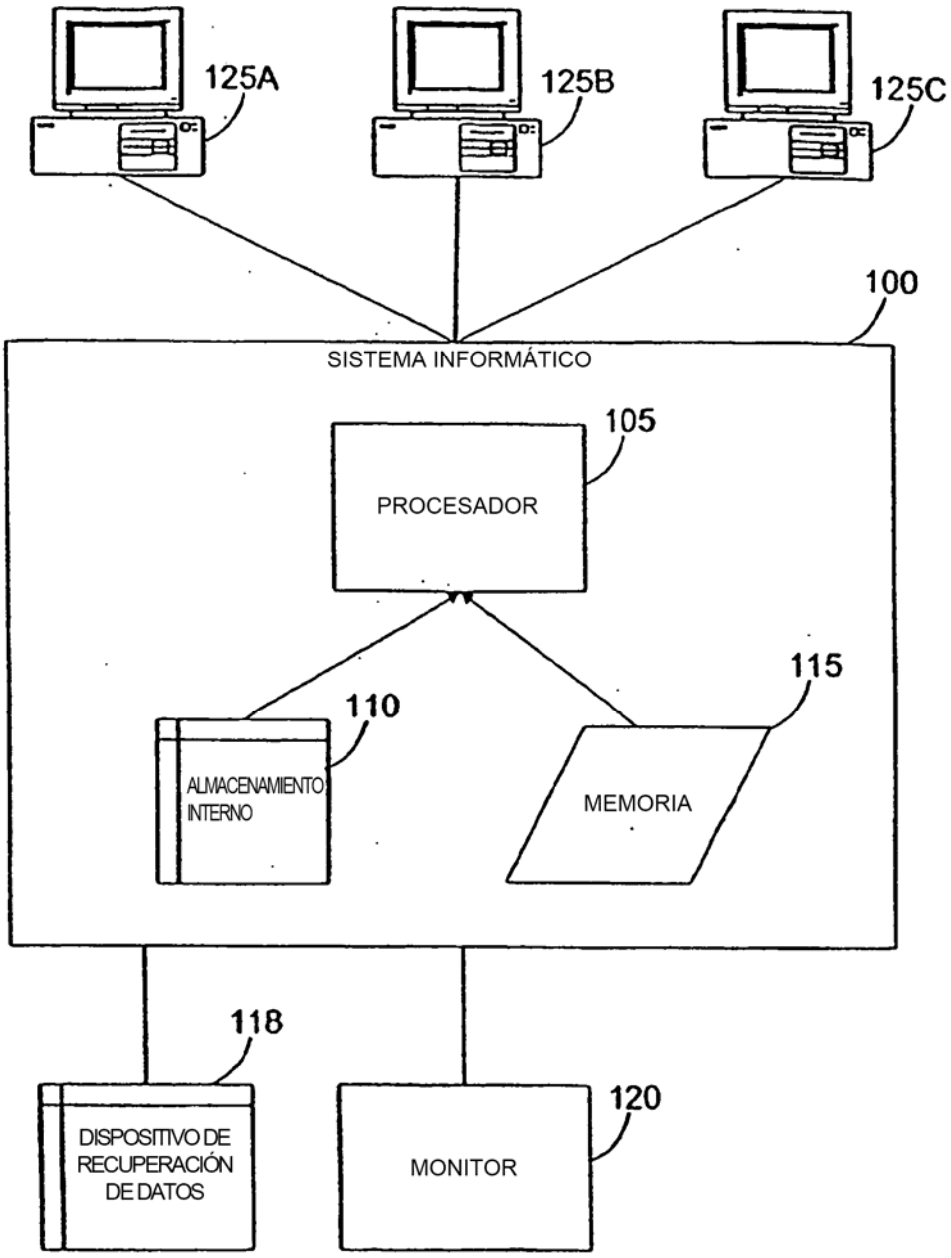

125A
125B
125C
100
SISTEMA INFORMÁTICO
105
PROCESADOR
115
110
ALMACENAMIENTO INTERNO
MEMORIA
118
DISPOSITIVO DE RECUPERACIÓN DE DATOS
120
MONITOR

FIGURA 1

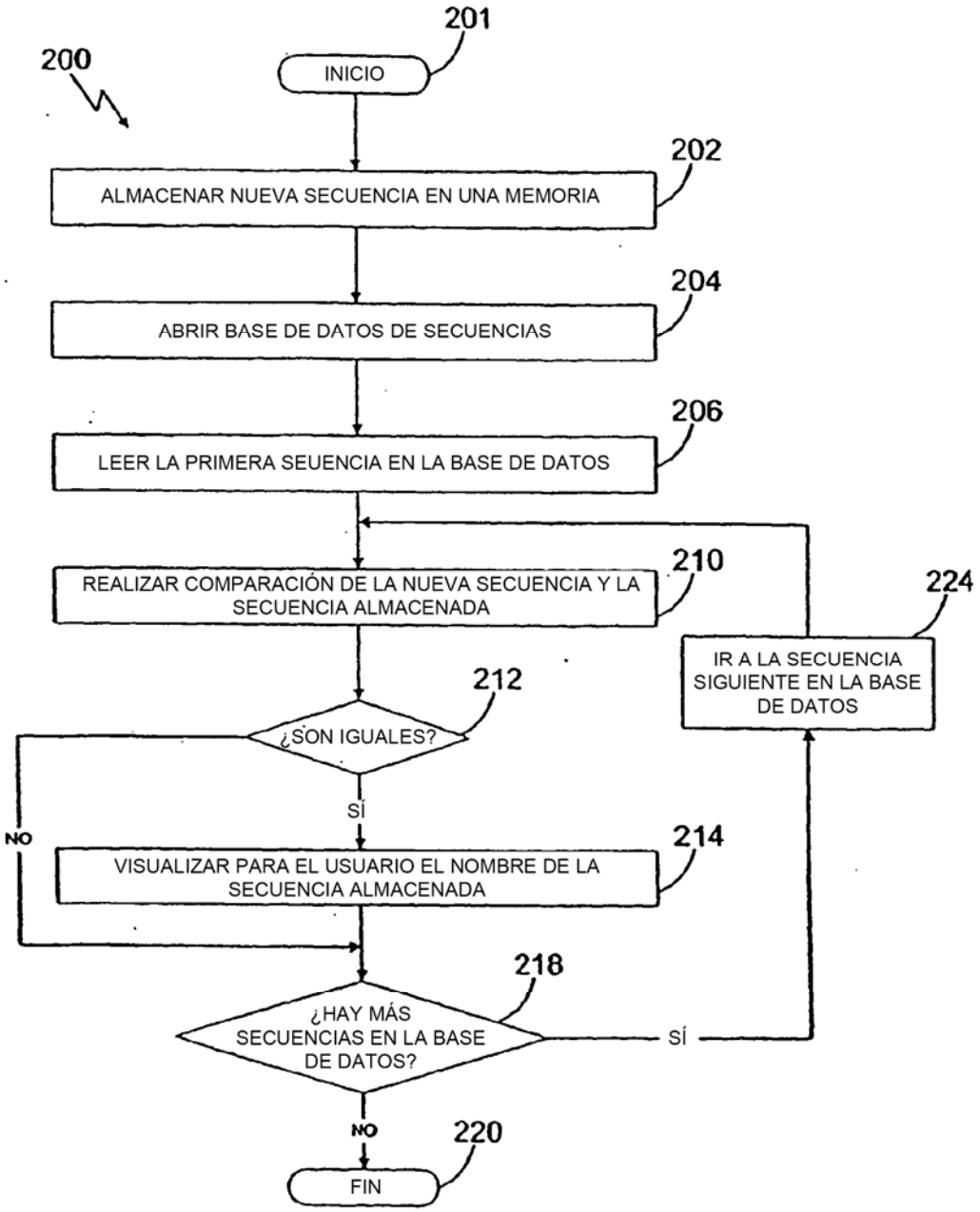

200
201
INICIO
202
ALMACENAR NUEVA SECUENCIA EN UNA MEMORIA
204
ABRIR BASE DE DATOS DE SECUENCIAS
206
LEER LA PRIMERA SEUENCIA EN LA BASE DE DATOS
210
REALIZAR COMPARACIÓN DE LA NUEVA SECUENCIA Y LA SECUENCIA ALMACENADA
212
¿SON IGUALES?
SÍ
NO
214
VISUALIZAR PARA EL USUARIO EL NOMBRE DE LA SECUENCIA ALMACENADA
218
¿HAY MÁS SECUENCIAS EN LA BASE DE DATOS?
SÍ
NO
220
FIN
224
IR A LA SECUENCIA SIGUIENTE EN LA BASE DE DATOS

FIGURA 2

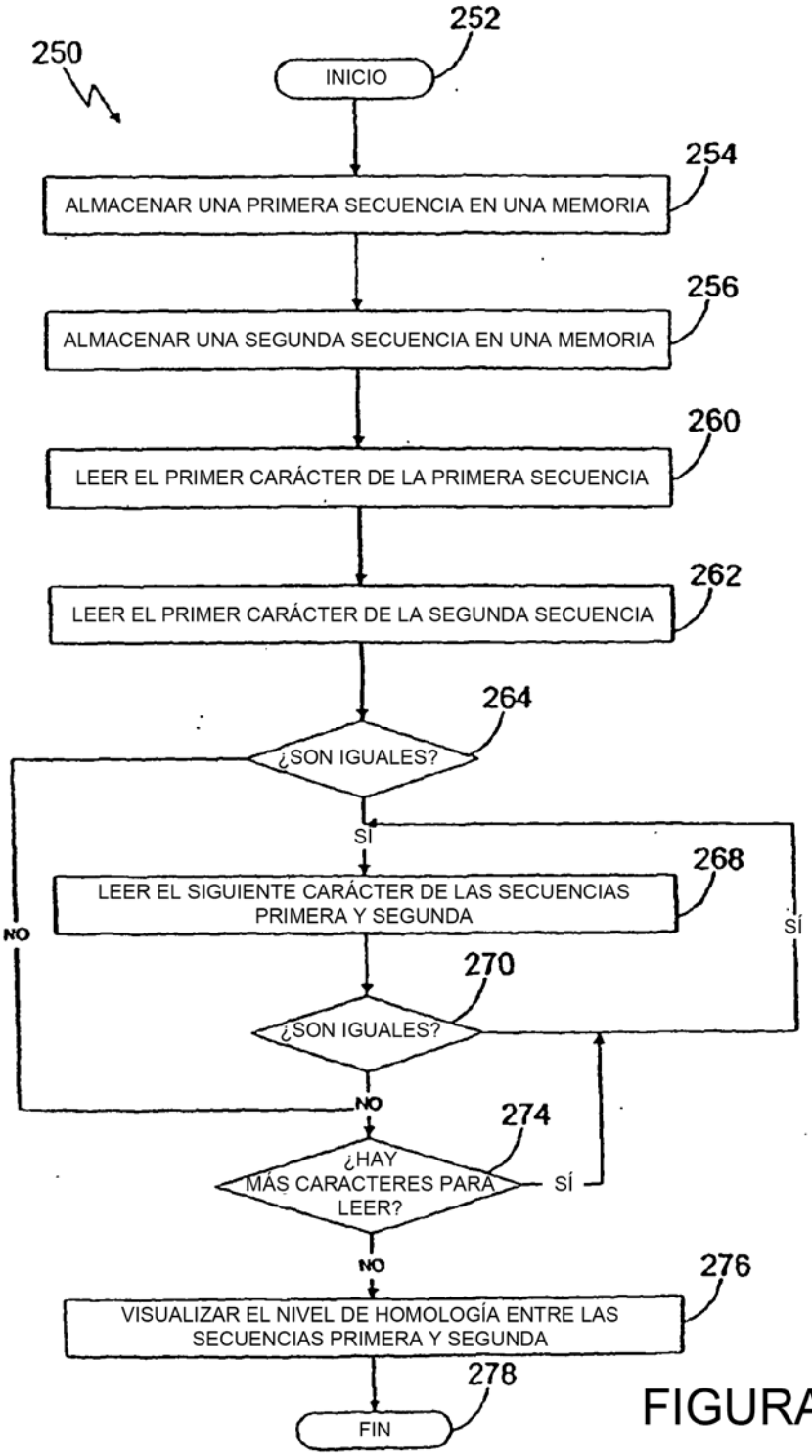

250
252
INICIO
254
ALMACENAR UNA PRIMERA SECUENCIA EN UNA MEMORIA
256
ALMACENAR UNA SEGUNDA SECUENCIA EN UNA MEMORIA
260
LEER EL PRIMER CARÁCTER DE LA PRIMERA SECUENCIA
262
LEER EL PRIMER CARÁCTER DE LA SEGUNDA SECUENCIA
264
¿SON IGUALES?
SÍ
NO
268
LEER EL SIGUIENTE CARÁCTER DE LAS SECUENCIAS PRIMERA Y SEGUNDA
SÍ
270
¿SON IGUALES?
NO
274
¿HAY MÁS CARACTERES PARA LEER?
SÍ
NO
276
VISUALIZAR EL NIVEL DE HOMOLOGÍA ENTRE LAS SECUENCIAS PRIMERA Y SEGUNDA
278
FIN

FIGURA 3

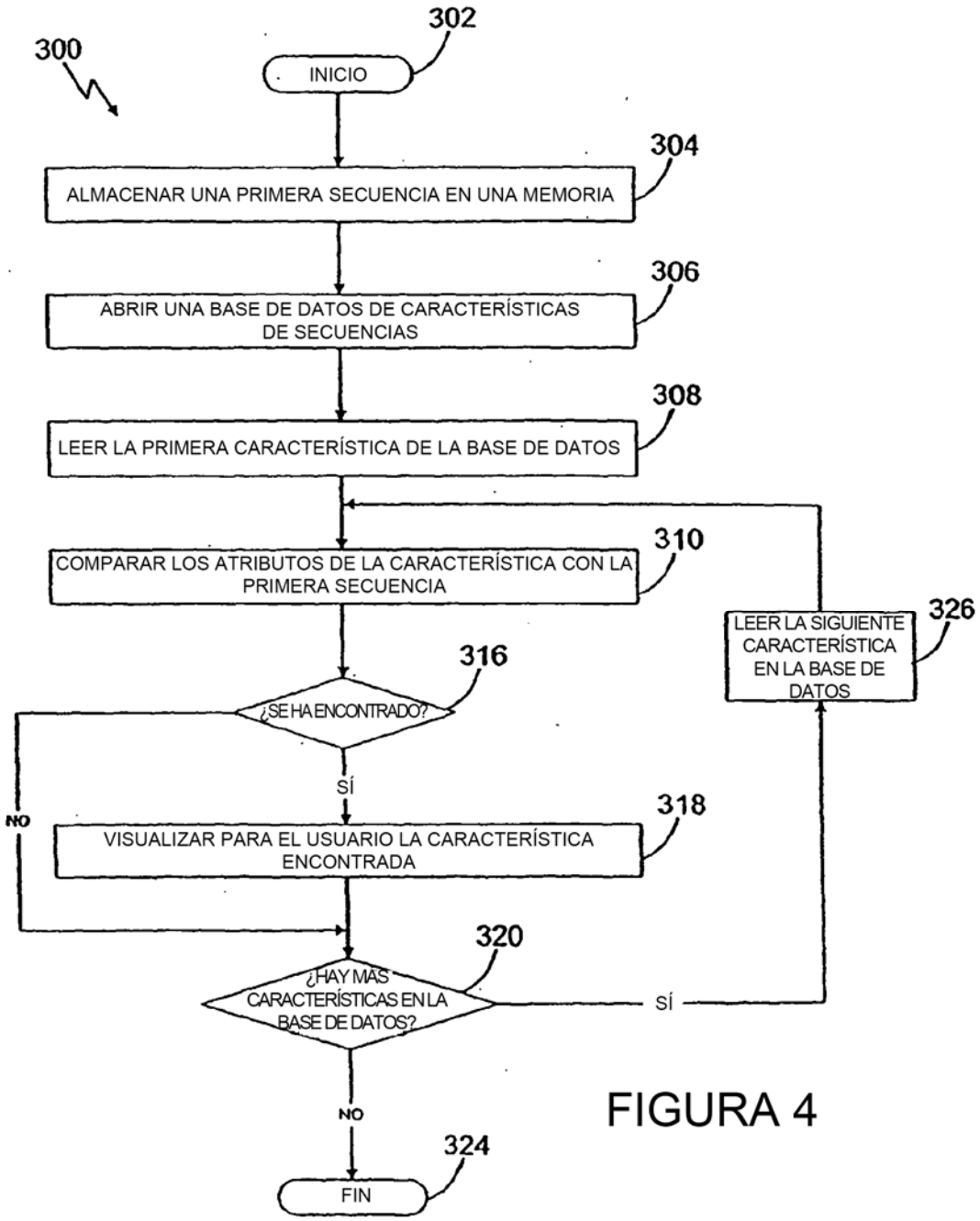
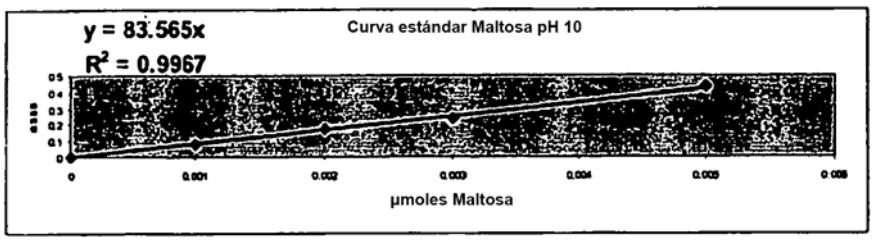

300
302
INICIO
304
ALMACENAR UNA PRIMERA SECUENCIA EN UNA MEMORIA
306
ABRIR UNA BASE DE DATOS DE CARACTERÍSTICAS DE SECUENCIAS
308
LEER LA PRIMERA CARACTERÍSTICA DE LA BASE DE DATOS
310
COMPARAR LOS ATRIBUTOS DE LA CARACTERÍSTICA CON LA PRIMERA SECUENCIA
316
¿SE HA ENCONTRADO?
SÍ
NO
318
VISUALIZAR PARA EL USUARIO LA CARACTERÍSTICA ENCONTRADA
320
¿HAY MÁS CARACTERÍSTICAS EN LA BASE DE DATOS?
SÍ
NO
324
FIN
326
LEER LA SIGUIENTE CARACTERÍSTICA EN LA BASE DE DATOS

FIGURA 4

Figura 5: Ejemplo de Curva Estándar del ensayo del Ejemplo 2.



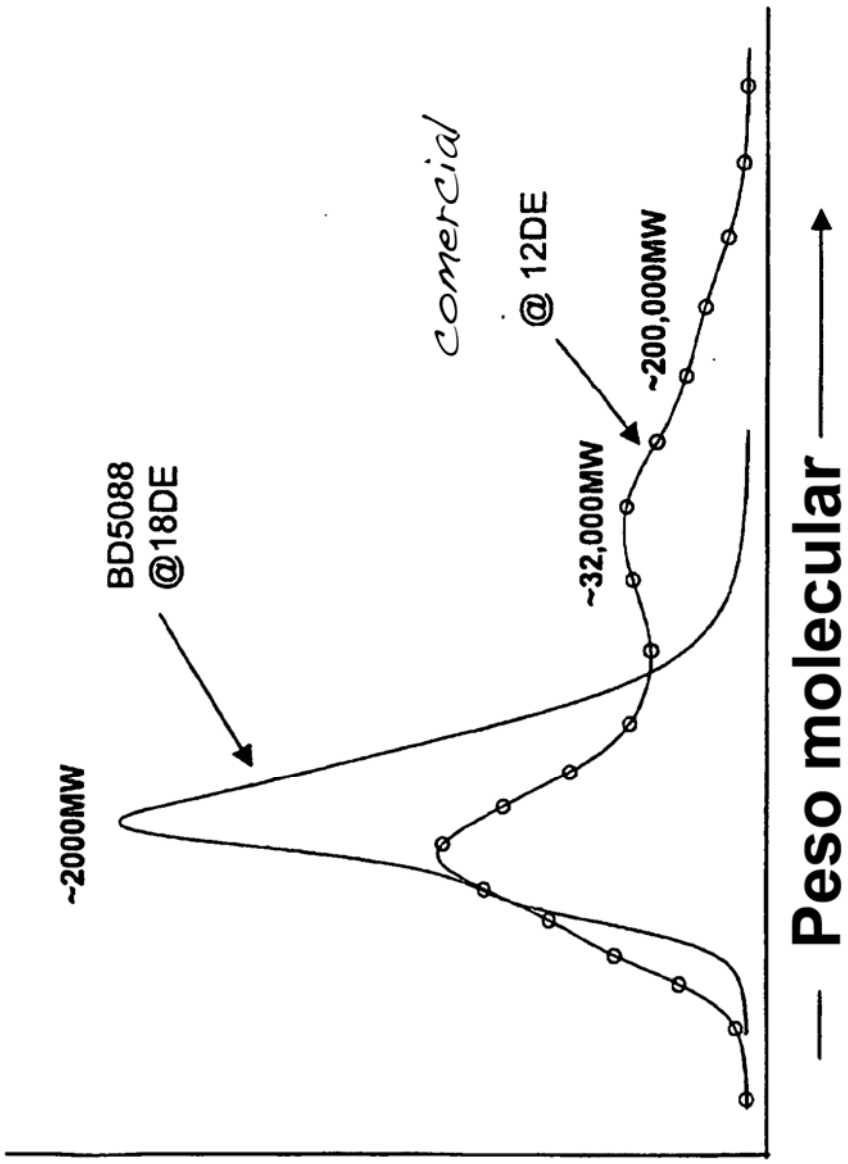

~2000MW
BD5088
@18DE
comercial
@ 12DE
~32,000MW
~200,000MW
— Peso molecular —

FIGURA 6a

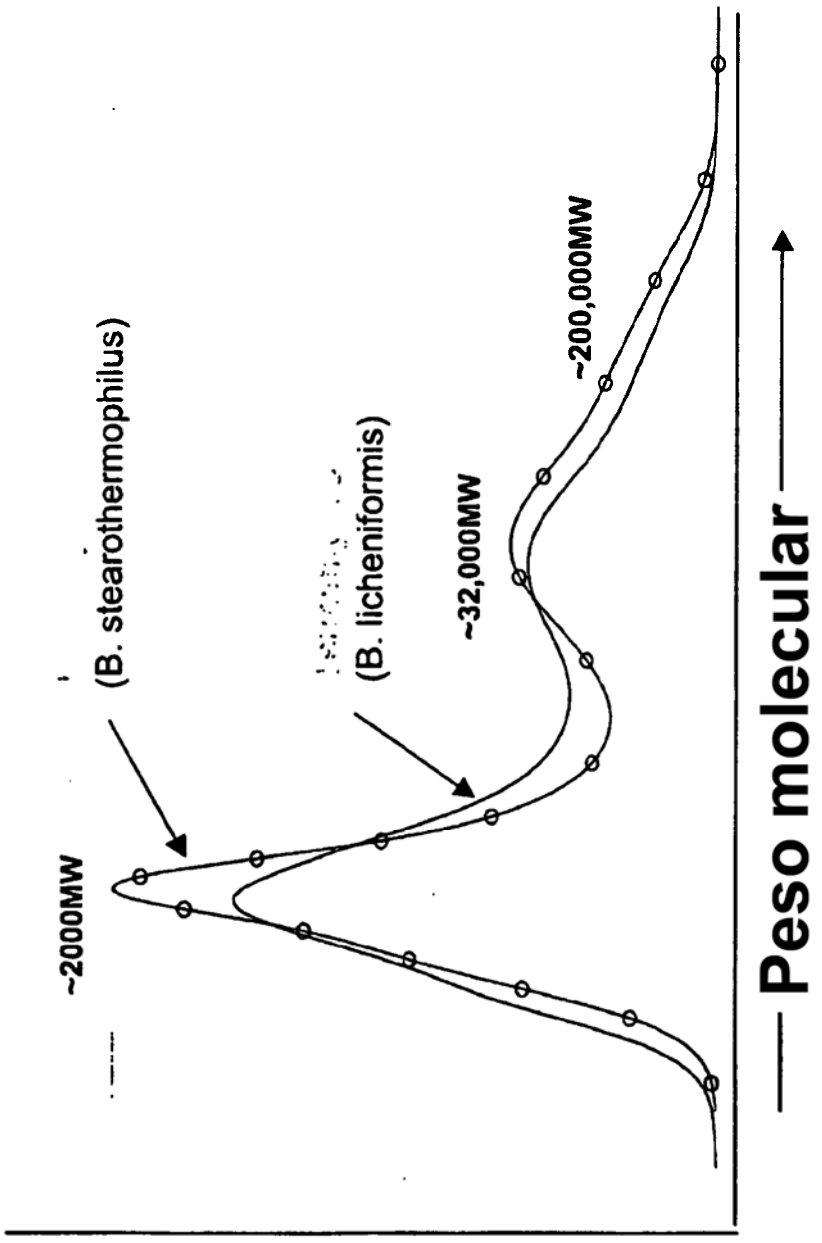

~2000MW
(B. stearothermophilus)
(B. licheniformis)
~32,000MW
~200,000MW
Peso molecular

FIGURA 6b

Figura 7:

ID SEC Nº: 1

atggccaagtactccgagctggaaaagggcggggtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaat
acggcagaagataccggagtggtacgatgccggaatctccgcaatatggattcccccggcgagcaagggcatgggcggcgcctattcgatg
ggctacgacccctacgacttctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaa
catgataaacaccgcccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggtgacctggagtggaaccccctt
cgtgaacgactatacctggaccgacttctcaaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccatgc
gggcgattccggaacatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctac
gcggcatatctcaggagcatcggcatcgatgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcaaggactggctgaactggt
ggggaggctgggcggttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcg
ccctctactacaagatggatgaggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgac
ccgttcaaggccgtaacctttgtagcaaaccacgacaccgatataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggcc
agccgacaatattctaccgcgactacgaggagtggctcaacaaggataagctcaagaacctcatctggatacatgagaacctcgccggagga
agcaccgacatagtctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctag
gctcgagcaaggccggaaggtgggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagac
aagtacgtctactcaagcggctgggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctactgcg
gggtgggctga

ID SEC Nº: 2

Met Ala Lys Tyr Ser Glu Leu Glu Lys Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser
Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp
Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp
Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn
Met Ile Asn Thr Ala His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly
Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys
Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly
Tyr Pro Asp Ile Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala
Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val
Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala
Val Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu
Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg
Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala
Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn
Lys Asp Lys Leu Lys Asn Leu Ile Trp Ile His Glu Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr
Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile
Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu
Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala
Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

ID SEC Nº: 3

atgttcctgctcgcgttttgctcactgcctcgctgttctgcccaacaggacagcccgccaaggctgccgcaccgtttaacggcaccatgatgca
gtattttgaatggtacttgccggatgatggcacgttatggaccaaagtggccaatgaagccaacaacttatccagccttggcatcaccgctctttg
gctgccgcccgcttacaaaggaacaagccgcagcgacgtagggtacggagtatacgacttgtatgacctcggcgaattcaatcaaaaaggga
ccgtccgcacaaaatacggaacaaaagctcaatatcttcaagccattcaagccgcccacgccgctggaatgcaagtgtacgccgatgtcgtgtt
cgaccataaaggcggcgctgacggcacggaatgggtggacgccgtcgaagtcaatccgtccgaccgcaaccaagaaatctcgggcacctat
caaatccaagcatggacgaaatttgattttcccgggcggggcaacacctactccagctttaagtggcgctggtaccattttgacggcgttgattgg
gacgaaagccgaaaattgagccgcatttacaaattccgcggcatcggcaaagcgtgggattgggaagtagacacggaaaacggaaactatg
actacttaatgtatgccgaccttgatatggatcatcccgaagtcgtgaccgagctgaaaaactgggggaaatggtatgtcaacacaacgaacatt
gatgggttccggcttgatgccgtcaagcatattaagttcagttttttcctgattggttgtcgtatgtgcgttctcagactggcaagccgctatttaccg
tcggggaatattggagctatgacatcaacaagttgcacaattacattacgaaaacagacggaacgatgtctttgtttgatgccccgttacacaaca

aattttataccgcttccaaatcaggggcgcatttgatatgcgcacgttaatgaccaatactctcatgaaagatcaaccgacattggccgtcaccttcgttgataatcatgacaccgaacccggccaagcgctgcagtcatgggtcgacccatggttcaaaccgttggcttacgcctttattctaactcggcaggaaggataccgtgcgtctttatggtgactattatggcattccacaatataacattccttcgctgaaaagcaaaatcgatccgctcctcatcgcgcgcagggattatgcttacggaacgcaacatgattatcttgatcactccgacatcatcgggtggacaagggaaggggtcactgaaaaaccaggatccgggctggccgcactgatcaccgatgggccgggaggaagcaaatggatgtactgttggcaaacaacacgctggaaaagtgttctatga

ID SEC N°: 4

atgaaacaacaaaaacggctttacgcccgattgctgacgctgttatttgcgctcatcttcttgctgcctcattctgcagcagcggcggcaaatcttaatgggacgctgatgcagtattttgaatggtacatgcccaatgacggccaacattggaagcgcttgcaaaacgactcggcatatttggctgaacacggtattactgccgtctggattcccccggcatataagggaacgagccaagcggatgtgggctacggtgcttacgacctttatgatttaggggagtttcatcaaaaagggacggttcggacaaagtacggcacaaaaggagagctgcaatctgcgatcaaaagtcttcattcccgcgacattaacgtttacggggatgtggtcatcaaccacaaaggcggcgctgatgcgaccgaagatgtaaccgcggttgaagtcgatcccgctgaccgcaaccgcgtaatttcaggagaacaccgaattaaagcctggacacattttcattttccggggcgcggcagcacatacagcgattttaaatggcattggtaccatttttgacggaaccgattgggacgagtcccgaaagctgaaccgcatctataagtttcaaggaaaggcttgggattgggaagtttccaatgaaaacggcaactatgattatttgatgtatgccgacatcgattatgaccatcctgatgtcgcagcagaaattaagagatggggcacttggtatgccaatgaactgcaatttggacggtttccgtcttgatgctgtcaaacacattaaatttttcttttttgcgggattgggttaatcatgtcagggaaaaaacggggaaggaaatgtttacggtagctgaatattggcagaatgacttgggcgcgctggaaaactatttgaacaaaacaaatttttaatcattcagtgtttgacgtgccgcttcattatcagttccatgctgcatcgacacagggaggcggctatgatatgaggaaattgctgaacggtacggtcgtttccaagcatccgttgaaagcggttacatttgtcgataaccatgatacacagccggggcaatcgcttgagtcgactgtccaaacatggtttaagccgcttgcttacgctttcattctcacaaggaatctggataccctcaggttttctacggggatatgtacgggacgaaaggagactcccagcgcgaaattcctgccttgaaacacaaaattgaaccgatcttaaaagcgagaaaacagtatgcgtacggagcacagcatgattatttcgaccaccatgacattgtcggctggacaagggaaggcgacagctcggttgcaaattcaggtttggcggcattaataacagacggacccggtggggcaaagcgaatgtatgtcggccggcaaaacgccggtgagacatggcatgacattaccggaaaccgttcggagccggttgtcatcaattcggaaggctggggagagtttcacgtaaacggcgggtcggtttcaatttatgttcaaagatag

ID SEC N°: 5

gtggtgcacatgaagttgaagtaccttgccttagttttgttggctgtggcttcgataggcctactctcgactccagtgggtgctgccaagtactccgaactcgaagagggcggtgttataatgcaggccttctactgggatgttcccggaggggaatctggtgggacaccataagacagaaaatcccggagtggtacgacgctggaatctcggcgatatggattcctccagctagcaaagggatgggcggtggttattccatgggctacgatccctacgatttctttgacctcggcgagtactatcagaagggaacagttgagacgcgcttcggctcaaaggaggaactggtgaacatgataaacaccgcacactcctatggcataaaggtgatagcggacatagtcataaaccaccgcgccggtggagaccttgagtggaacccctttgtaaacaactatacttggacagacttctccaaggtcgcctccggtaaatacacggccaactaccttgacttccacccaaacgaggtcaagtgctgcgatgagggtacatttggtgactttccggacatcgcccacgagaagagctgggatcagtactggctctgggcaagcaatgagagctacgccgcatatctccggagcataggagtcgatgcatggcgtttcgactacgtcaaaggttacggagcgtgggttgttaatgactggctcagctggtggggaggctgggccgttggagagtactgggacacgaacgttgatgcactccttaactgggcatacgacagcggtgccaaggtctttgacttcccgctctactacaagatggacgaagccttttgacaacaccaacatccccgctttggtttacgccctccagaacggaggaacagtcgtttcccgcgatcccttcaaggcagtaactttcgttgccaaccacgatacagatataatctggaacaagtatccggcttatgcgttcatccttacctatgagggacagcctgttatattttaccgcgactacgaggagtggctcaacaaggataagcttaacaaccttatctggatacacgagcaccttgccggaggaagtaccaagatcctctactacgataacgatgagctaatattcatgagggagggctacgggagcaagccgggcctcataacctacataaacctcggaaacgactgggccgagcgctgggtgaacgtcggctcaaagtttgccggctacacaatccatgaatacacaggcaatctcggtggctgggttgacaggtgggttcagtacgatggatgggttaaactgacggcacctcctcatgatccagccaacggatattacggctactcagtctggagctacgcaggcgtcggatga

ID SEC N°: 6

atgaagaagtttgtcgccctgttcataaccatgtttttcgtagtgagcatggcagtcgttgcacagccagctagcgccgcaaagtattccgagctcgaagaaggcggcgttataatgcaggccttctactgggacgtcccaggtggaggaatctggtgggacaccatcaggagcaagatacccggagtggtacgaggcgggaatatccgccatttggattccgccagccagcaaggggatgagcggcggttactcgatgggctacgatccctacgatttcttgacctcggcgagtacaaccagaagggaaccatcgaaacgcgctttggctctaaacaggagctcatcaatatgataaacacggcccatgcctacggcataaaggtcatagcggacatcgtcataaaccaccgcgcaggcggagacctcgagtggaacccgttcgttggggactacacctggacggacttctcaaaggtggcctcgggcaaatatactgccaactacctcgacttccaccccaacgaggtcaagtgctgtgacgagggcacatttggaggcttcccagacatagcccacgagaagagctgggaccagcactggctctgggcgagcgatgagagctacgccgcctacctaaggagcatcggcgttgatgcctggcgctttgactacgtgaaggctacggagcgtgggtcgtcaaggactggctcaactggtggggcggctgggccgttggc

gagtactgggacaccaacgttgatgcactcctcaactgggcctactcgagcggcgccaaggtcttcgacttcccgctctactacaagatggatg
aggcctttgacaacaaaaacattccagcgctcgtctctgcccttcagaacggccagactgttgtctcccgcgacccgttcaaggccgtaaccttt
gtagcaaaccacgacaccgatataatctggaacaagtaccttgcttatgctttcatcctcacctacgaaggccagcccgtcatattctaccgcgac
tacgaggagtggctcaacaaggacaggttgaacaacctcatatggatacacgaccacctcgcaggtggaagcacgagcatagtctactacga
cagcgacgagatgatcttcgtgaggaacggctatggaagcaagcctggccttataacttacatcaacctcggctcgagcaaggttggaaggtg
ggtttatgtgccgaagttcgcgggcgcgtgcatccacgagtatactggtaacctcggaggctgggtagacaagtacgtctactcaagcggctg
ggtctatctcgaagctccagcttacgaccctgccaacgggcagtatggctactccgtgtggagctattgcggtgttgggtga

ID SEC Nº: 7

Met Phe Leu Leu Ala Phe Leu Leu Thr Ala Ser Leu Phe Cys Pro Thr Gly Gln Pro Ala Lys Ala Ala Ala Pro Phe Asn Gly Thr Met Met Gln Tyr Phe Glu Trp Tyr Leu Pro Asp Asp Gly Thr Leu Trp Thr Lys Val Ala Asn Glu Ala Asn Asn Leu Ser Ser Leu Gly Ile Thr Ala Leu Trp Leu Pro Pro Ala Tyr Lys Gly Thr Ser Arg Ser Asp Val Gly Tyr Gly Val Tyr Asp Leu Tyr Asp Leu Gly Glu Phe Asn Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Ala Gln Tyr Leu Gln Ala Ile Gln Ala Ala His Ala Ala Gly Met Gln Val Tyr Ala Asp Val Val Phe Asp His Lys Gly Gly Ala Asp Gly Thr Glu Trp Val Asp Ala Val Glu Val Asn Pro Ser Asp Arg Asn Gln Glu Ile Ser Gly Thr Tyr Gln Ile Gln Ala Trp Thr Lys Phe Asp Phe Pro Gly Arg Gly Asn Thr Tyr Ser Ser Phe Lys Trp Arg Trp Tyr His Phe Asp Gly Val Asp Trp Asp Glu Ser Arg Lys Leu Ser Arg Ile Tyr Lys Phe Arg Gly Ile Gly Lys Ala Trp Asp Trp Glu Val Asp Thr Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Leu Asp Met Asp His Pro Glu Val Val Thr Glu Leu Lys Asn Trp Gly Lys Trp Tyr Val Asn Thr Thr Asn Ile Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Phe Pro Asp Trp Leu Ser Tyr Val Arg Ser Gln Thr Gly Lys Pro Leu Phe Thr Val Gly Glu Tyr Trp Ser Tyr Asp Ile Asn Lys Leu His Asn Tyr Ile Thr Lys Thr Asp Gly Thr Met Ser Leu Phe Asp Ala Pro Leu His Asn Lys Phe Tyr Thr Ala Ser Lys Ser Gly Gly Ala Phe Asp Met Arg Thr Leu Met Thr Asn Thr Leu Met Lys Asp Gln Pro Thr Leu Ala Val Thr Phe Val Asp Asn His Asp Thr Glu Pro Gly Gln Ala Leu Gln Ser Trp Val Asp Pro Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Gln Glu Gly Tyr Pro Cys Val Phe Tyr Gly Asp Tyr Tyr Gly Ile Pro Gln Tyr Asn Ile Pro Ser Leu Lys Ser Lys Ile Asp Pro Leu Leu Ile Ala Arg Arg Asp Tyr Ala Tyr Gly Thr Gln His Asp Tyr Leu Asp His Ser Asp Ile Ile Gly Trp Thr Arg Glu Gly Val Thr Glu Lys Pro Gly Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ser Lys Trp Met Tyr Cys Trp Gln Thr Thr Arg Trp Lys Ser Val Leu

ID SEC Nº: 8

Met Lys Gln Gln Lys Arg Leu Tyr Ala Arg Leu Leu Thr Leu Leu Phe Ala Leu Ile Phe Leu Leu Pro His Ser Ala Ala Ala Ala Ala Asn Leu Asn Gly Thr Leu Met Gln Tyr Phe Glu Trp Tyr Met Pro Asn Asp Gly Gln His Trp Lys Arg Leu Gln Asn Asp Ser Ala Tyr Leu Ala Glu His Gly Ile Thr Ala Val Trp Ile Pro Pro Ala Tyr Lys Gly Thr Ser Gln Ala Asp Val Gly Tyr Gly Ala Tyr Asp Leu Tyr Asp Leu Gly Glu Phe His Gln Lys Gly Thr Val Arg Thr Lys Tyr Gly Thr Lys Gly Glu Leu Gln Ser Ala Ile Lys Ser Leu His Ser Arg Asp Ile Asn Val Tyr Gly Asp Val Val Ile Asn His Lys Gly Gly Ala Asp Ala Thr Glu Asp Val Thr Ala Val Glu Val Asp Pro Ala Asp Arg Asn Arg Val Ile Ser Gly Glu His Arg Ile Lys Ala Trp Thr His Phe His Phe Pro Gly Arg Gly Ser Thr Tyr Ser Asp Phe Lys Trp His Trp Tyr His Phe Asp Gly Thr Asp Trp Asp Glu Ser Arg Lys Leu Asn Arg Ile Tyr Lys Phe Gln Gly Lys Ala Trp Asp Trp Glu Val Ser Asn Glu Asn Gly Asn Tyr Asp Tyr Leu Met Tyr Ala Asp Ile Asp Tyr Asp His Pro Asp Val Ala Ala Glu Ile Lys Arg Trp Gly Thr Trp Tyr Ala Asn Glu Leu Gln Leu Asp Gly Phe Arg Leu Asp Ala Val Lys His Ile Lys Phe Ser Phe Leu Arg Asp Trp Val Asn His Val Arg Glu Lys Thr Gly Lys Glu Met Phe Thr Val Ala Glu Tyr Trp Gln Asn Asp Leu Gly Ala Leu Glu Asn Tyr Leu Asn Lys Thr Asn Phe Asn His Ser Val Phe Asp Val Pro Leu His Tyr Gln Phe His Ala Ala Ser Thr Gln Gly Gly Gly Tyr Asp Met Arg Lys Leu Leu Asn Gly Thr Val Val Ser Lys His Pro Leu Lys Ala Val Thr Phe Val Asp Asn His Asp Thr Gln Pro Gly Gln Ser Leu Glu Ser Thr Val Gln Thr Trp Phe Lys Pro Leu Ala Tyr Ala Phe Ile Leu Thr Arg Glu Ser Gly Tyr Pro Gln Val Phe Tyr Gly Asp Met Tyr Gly Thr Lys Gly Asp Ser Gln Arg Glu Ile Pro Ala Leu Lys His Lys Ile Glu Pro Ile Leu Lys Ala Arg Lys Gln Tyr Ala Tyr Gly Ala Gln His Asp Tyr Phe Asp His His Asp

Ile Val Gly Trp Thr Arg Glu Gly Asp Ser Ser Val Ala Asn Ser Gly Leu Ala Ala Leu Ile Thr Asp Gly Pro Gly Gly Ala Lys Arg Met Tyr Val Gly Arg Gln Asn Ala Gly Glu Thr Trp His Asp Ile Thr Gly Asn Arg Ser Glu Pro Val Val Ile Asn Ser Glu Gly Trp Gly Glu Phe His Val Asn Gly Gly Ser Val Ser Ile Tyr Val Gln Arg

ID SEC Nº: 9

Val Val His Met Lys Leu Lys Tyr Leu Ala Leu Val Leu Leu Ala Val Ala Ser Ile Gly Leu Leu Ser Thr Pro Val Gly Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Glu Glu Leu Val Asn Met Ile Asn Thr Ala His Ser Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asn Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Asp Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Asn Asp Trp Leu Ser Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Asp Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Thr Asn Ile Pro Ala Leu Val Tyr Ala Leu Gln Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu Asn Asn Leu Ile Trp Ile His Glu His Leu Ala Gly Gly Ser Thr Lys Ile Leu Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Met Arg Glu Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Asn Asp Trp Ala Glu Arg Trp Val Asn Val Gly Ser Lys Phe Ala Gly Tyr Thr Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Arg Trp Val Gln Tyr Asp Gly Trp Val Lys Leu Thr Ala Pro Pro His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Ala Gly Val Gly

ID SEC Nº: 10

Met Lys Lys Phe Val Ala Leu Phe Ile Thr Met Phe Phe Val Val Ser Met Ala Val Val Ala Gln Pro Ala Ser Ala Ala Lys Tyr Ser Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Gly Gly Gly Ile Trp Trp Asp Thr Ile Arg Ser Lys Ile Pro Glu Trp Tyr Glu Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asn Gln Lys Gly Thr Ile Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Gly Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln His Trp Leu Trp Ala Ser Asp Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Ser Ala Leu Gln Asn Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Leu Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Val Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Asn Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Ser Ile Val Tyr Tyr Asp Ser Asp Glu Met Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Val Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Tyr Val Tyr Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala Tyr Asp Pro Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

ID SEC Nº: 66

atggctctggaagagggcgggcttataatgcaggcattctactgggacgtccccatgggaggaatctggtgggacacgatagcccagaagat
acccgactgggcaagcgccgggatttcggcgatatggattccccccgcgagcaagggtatgagcggcggctattcgatgggctacgacccct
acgattattttgacctcggtgagtactaccagaagggaacggtggaaacaagattcggctcaaagcaggagctcataaacatgataaacaccg
cccacgcctatggcatgaaggtaatagccgatatagtcatcaaccaccgcgccggcggcgatctggagtggaaccccttcgtgaacgactata
cctggaccgacttctcgaaggtcgcgtcgggtaaatacacggccaactacctcgacttccacccgaacgagctccacgcgggcgattccgga
acatttggaggctatcccgacatatgccacgacaagagctgggaccagtactggctctgggccagccaggagagctacgcggcctatctcag
gagcatcggcatcgacgcctggcgcttcgactacgtcaagggctatgctccctgggtcgtcagggactggctgaactggtggggaggctggg
cagttggagagtactgggacaccaacgtcgacgctgttctcaactgggcatactcgagcggtgccaaggtctttgacttcgccctctactacaag
atggacgaggccttcgataacaacaacattcccgccctggtggacgccctcagatacggccagacagtggtcagccgcgacccgttcaaggc
tgtgacgtttgtagccaaccacgataccgacataatctggaacaagtatccagcctacgcgttcatcctcacctacgagggccagccgacaatat
tctaccgcgactacgaggagtggctcaacaaggacaagctcaagaacctcatctggatacatgacaacctcgccggagggagcactgacatc
gtttactacgacaacgacgagctgatattcgtgagaaacggctacggaagcaagccgggactgataacatacatcaacctcggctcaagcaaa
gccggaaggtgggtttacgttccgaagttcgcaggctcgtgcatacacgagtacaccggcaacctcggcggctgggtggacaagtgggtgga
ctcaagcggctgggtttacctcgaggctcctgcccacgacccggccaacggccagtacggctactccgtttggagctattgcggtgttgggtga

ID SEC Nº: 67

Met Ala Leu Glu Glu Gly Gly Leu Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Met Gly Gly Ile Trp
Trp Asp Thr Ile Ala Gln Lys Ile Pro Asp Trp Ala Ser Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala
Ser Lys Gly Met Ser Gly Gly Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Tyr Phe Asp Leu Gly Glu Tyr
Tyr Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Ile Asn Met Ile Asn Thr Ala
His Ala Tyr Gly Met Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp
Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn
Tyr Leu Asp Phe His Pro Asn Glu Leu His Ala Gly Asp Ser Gly Thr Phe Gly Gly Tyr Pro Asp Ile
Cys His Asp Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Gln Glu Ser Tyr Ala Ala Tyr Leu Arg
Ser Ile Gly Ile Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Ala Pro Trp Val Val Arg Asp Trp
Leu Asn Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Val Leu Asn Trp
Ala Tyr Ser Ser Gly Ala Lys Val Phe Asp Phe Ala Leu Tyr Tyr Lys Met Asp Glu Ala Phe Asp Asn
Asn Asn Ile Pro Ala Leu Val Asp Ala Leu Arg Tyr Gly Gln Thr Val Val Ser Arg Asp Pro Phe Lys
Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile
Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Lys Leu
Lys Asn Leu Ile Trp Ile His Asp Asn Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp
Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Ser Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser
Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ser Cys Ile His Glu Tyr Thr Gly Asn Leu
Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro
Ala Asn Gly Gln Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

ID SEC Nº: 68

atgaagcctgcgaaactcctcgtctttgtgctcgtagtctctatcctcgcggggctctacgcccagcccgcgggggcggccaagtacctggagc
tcgaagagggcggcgtcataatgcaggcgttctactgggacgtgccttcaggaggaatatggtgggacacaatacggcagaagatacccgga
gtggtacgatgccggaatctccgcaatatggattcccccggcgagcaagggcatgggcggcgcctattcgatgggctacgacccctacgactt
ctttgacctcggtgagtacgaccagaagggaacggtagagacgcgctttggctccaagcaggagctcgtgaacatgataaacaccgcccacg
cctacggcatcaaggtcatcgcagacatagtaatcaaccaccgcgccggaggagaccttgagtggaacccccttcgtcaatgactacacctgga
cggacttctcgaaggtcgcttccggcaagtacacggccaactacctcgacttccaccccaacgaggtcaagtgctgcgacgagggcacctttg
gagggttcccggacatagcccacgagaagagctgggaccagtactggctctgggcgagcaacgagagctacgccgcctacctcaggagca
tcggcgttgacgcatggcgcttcgactacgtcaagggctacggagcgtgggtcgtcaaggactggctggactggtggggaggctgggccgt
cggggagtactgggacacaaacgttgatgcactgctcaactgggcctactcgagcgatgcaaaagtcttcgacttcccgctctactacaagatg
gacgcggcctttgacaacaagaacattcccgcactcgtcgaggccctcaagaacgggggcacagtcgtcagccgcgacccgtttaaggccgt
aaccttcgttgcaaaccacgacacggacataatttggaacaagtacccggcctacgccttcatcctcacctacgagggccagccgacgatattc
taccgcgactacgaggagtggctcaacaaggacaggctcaagaacctcatctggatacacgaccacctcgccggtggaagcaccgacatag

tctactacgataacgatgaactcatcttcgtcaggaacggctacggggacaagccggggcttataacctacatcaacctaggctcgagcaagg
ccgggaggtgggtctacgttccgaagttcgcgggagcgtgcatccacgagtacaccggcaacctcggcggctgggtggacaagtgggtgga
ctcaagcgggtgggtgtacctcgaggcccctgcccacgacccggccaacggctattacggctactccgtctggagctactgcggggtgggct
ga

ID SEC Nº: 69

Met Lys Pro Ala Lys Leu Leu Val Phe Val Leu Val Val Ser Ile Leu Ala Gly Leu Tyr Ala Gln Pro Ala Gly Ala Ala Lys Tyr Leu Glu Leu Glu Glu Gly Gly Val Ile Met Gln Ala Phe Tyr Trp Asp Val Pro Ser Gly Gly Ile Trp Trp Asp Thr Ile Arg Gln Lys Ile Pro Glu Trp Tyr Asp Ala Gly Ile Ser Ala Ile Trp Ile Pro Pro Ala Ser Lys Gly Met Gly Gly Ala Tyr Ser Met Gly Tyr Asp Pro Tyr Asp Phe Phe Asp Leu Gly Glu Tyr Asp Gln Lys Gly Thr Val Glu Thr Arg Phe Gly Ser Lys Gln Glu Leu Val Asn Met Ile Asn Thr Ala His Ala Tyr Gly Ile Lys Val Ile Ala Asp Ile Val Ile Asn His Arg Ala Gly Gly Asp Leu Glu Trp Asn Pro Phe Val Asn Asp Tyr Thr Trp Thr Asp Phe Ser Lys Val Ala Ser Gly Lys Tyr Thr Ala Asn Tyr Leu Asp Phe His Pro Asn Glu Val Lys Cys Cys Asp Glu Gly Thr Phe Gly Gly Phe Pro Asp Ile Ala His Glu Lys Ser Trp Asp Gln Tyr Trp Leu Trp Ala Ser Asn Glu Ser Tyr Ala Ala Tyr Leu Arg Ser Ile Gly Val Asp Ala Trp Arg Phe Asp Tyr Val Lys Gly Tyr Gly Ala Trp Val Val Lys Asp Trp Leu Asp Trp Trp Gly Gly Trp Ala Val Gly Glu Tyr Trp Asp Thr Asn Val Asp Ala Leu Leu Asn Trp Ala Tyr Ser Ser Asp Ala Lys Val Phe Asp Phe Pro Leu Tyr Tyr Lys Met Asp Ala Ala Phe Asp Asn Lys Asn Ile Pro Ala Leu Val Glu Ala Leu Lys Asn Gly Gly Thr Val Val Ser Arg Asp Pro Phe Lys Ala Val Thr Phe Val Ala Asn His Asp Thr Asp Ile Ile Trp Asn Lys Tyr Pro Ala Tyr Ala Phe Ile Leu Thr Tyr Glu Gly Gln Pro Thr Ile Phe Tyr Arg Asp Tyr Glu Glu Trp Leu Asn Lys Asp Arg Leu Lys Asn Leu Ile Trp Ile His Asp His Leu Ala Gly Gly Ser Thr Asp Ile Val Tyr Tyr Asp Asn Asp Glu Leu Ile Phe Val Arg Asn Gly Tyr Gly Asp Lys Pro Gly Leu Ile Thr Tyr Ile Asn Leu Gly Ser Ser Lys Ala Gly Arg Trp Val Tyr Val Pro Lys Phe Ala Gly Ala Cys Ile His Glu Tyr Thr Gly Asn Leu Gly Gly Trp Val Asp Lys Trp Val Asp Ser Ser Gly Trp Val Tyr Leu Glu Ala Pro Ala His Asp Pro Ala Asn Gly Tyr Tyr Gly Tyr Ser Val Trp Ser Tyr Cys Gly Val Gly

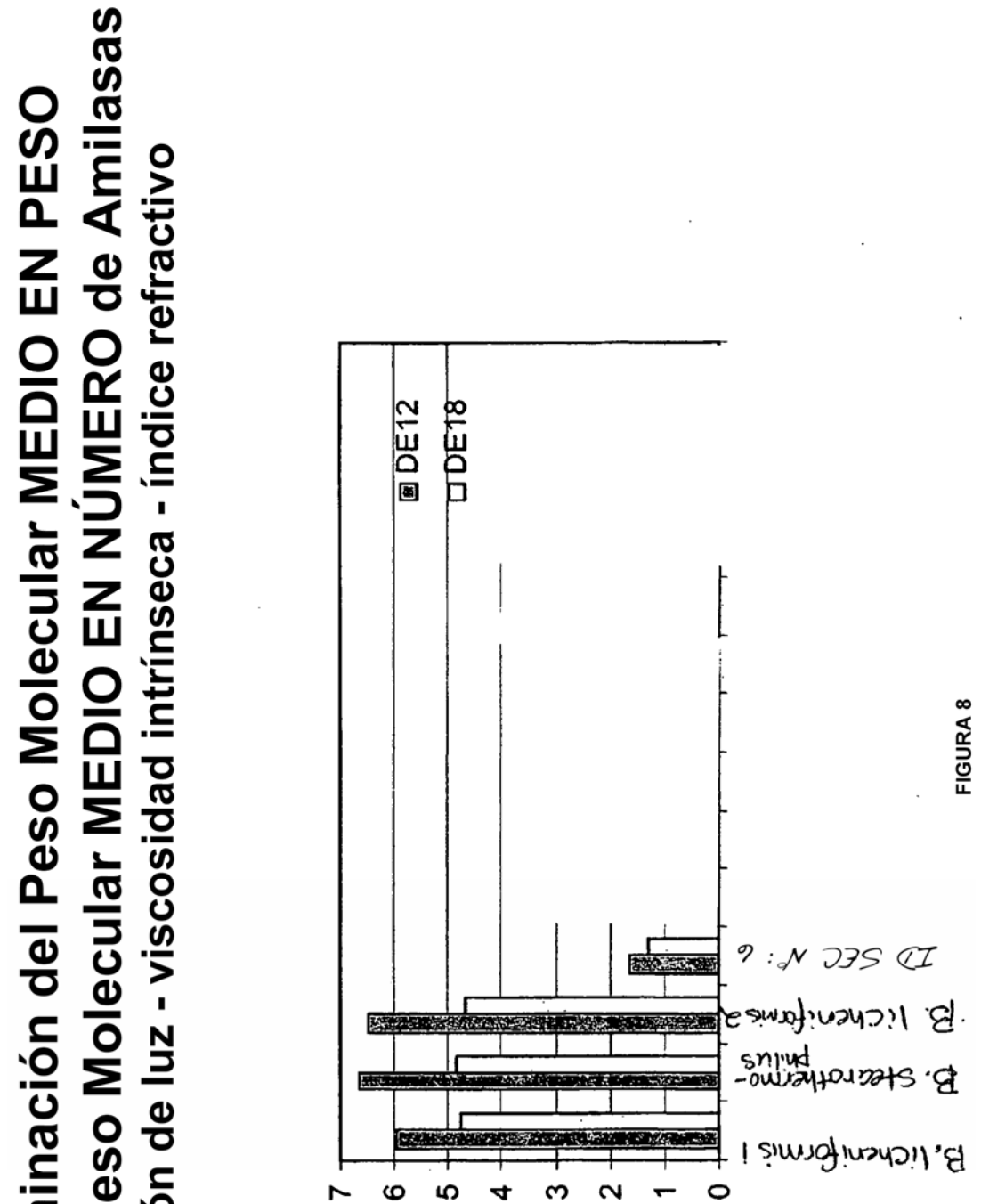

Determinación del Peso Molecular MEDIO EN PESO
frente al Peso Molecular MEDIO EN NÚMERO de Amilasas
difusión de luz - viscosidad intrínseca - índice refractivo
DE12
DE18
0
1
2
3
4
5
6
7
B. licheniformis 1
B. stearothermo-philus
B. licheniformis 2
ID SEC Nº: 6

FIGURA 8