

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 545 639**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 9/88 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **02.04.2004** **E 04758802 (5)**

97 Fecha y número de publicación de la concesión europea: **03.06.2015** **EP 1613733**

54 Título: **Pectato liasas, ácidos nucleicos que las codifican y métodos para su preparación y uso**

30 Prioridad:

04.04.2003 US 460842 P
03.07.2003 US 484798 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
14.09.2015

73 Titular/es:

BASF ENZYMES LLC (100.0%)
3550 John Hopkins Court
San Diego, CA 92121, US

72 Inventor/es:

KEROVUO, JANNE;
SOLBAK, ARNE;
GRAY, KEVIN;
MCCANN, RYAN;
PUROHIT, SHALAKA;
GERENDASH, JOEL;
JANSSEN, GISELLE y
DAHOD, SAMUN

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 545 639 T3

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

DESCRIPCIÓN

Pectato liasas, ácidos nucleicos que las codifican y métodos para su preparación y uso

5 Campo técnico

La presente invención se refiere a biología molecular y celular, bioquímica y biotecnología. La invención se refiere a polipéptidos que tienen una actividad de pectato liasa, por ejemplo, una pectinasa, polinucleótidos que codifican los polipéptidos, y métodos para la preparación y uso de estos polinucleótidos y polipéptidos. Los polipéptidos de la invención se pueden usar como pectato liasas para catalizar la beta-eliminación o hidrólisis de pectina y/o ácido poligalacturónico, tal como ácido alfa-D-galacturónico unido en la posición 1,4. Se pueden usar en una diversidad de aplicaciones industriales, por ejemplo, para tratar paredes de células vegetales, tales como las de algodón u otras fibras naturales. En otra aplicación industrial a modo de ejemplo, dos polipéptidos de la invención se pueden usar como limpiadores textiles.

15 Antecedentes

La fibra de algodón consta de una pared celular primaria y una pared celular secundaria. La pared celular secundaria es prácticamente celulosa pura, mientras que la pared celular primaria es una estructura reticular compleja de pectina, proteína, ceras, pigmentos, hemicelulosa y celulosa. En la limpieza textil de materiales de celulosa (por ejemplo, tejido de algodón de punto o tejido), se necesitan condiciones alcalinas (hasta un 10 % de NaOH) y temperaturas elevadas (hasta 100 °C) para una eliminación eficaz de los componentes de la pared celular primaria. Este tratamiento químico riguroso da como resultado pérdidas de materias primas y carga ambiental sustancial. Existen varias enzimas diferentes que tienen la capacidad de degradar la pectina; estas son las pectinasas, pectina metilesterasas, pectina liasas y pectato liasas.

"Apresto" es el nombre dado a la sustancia o mezcla de sustancias que se aplica al hilo de urdimbre antes de tejer. El apresto forma un revestimiento alrededor de la superficie del hilo antes de tejer. Este revestimiento proporciona lubricación y evita la rotura del hilo de urdimbre durante la operación de tejido. Algunas sustancias químicas habituales que se usan para preparar aprestos son Ácido Poliacrílico (PA), Alcohol Polivinílico (PVA), Almidón y Almidón Modificado. A las fibras celulósicas que incluyen algodón, rayón y mezcla de éstos con fibras sintéticas tales como poliéster, por lo general se les da apresto con aprestos a base de almidón. El proceso de retirar el apresto elimina el apresto antes del teñido, estampado y/o acabado. Los aprestos de almidón se pueden eliminar mediante lavado con ácido caliente, que hidrolizará el almidón. Sin embargo, la hidrólisis ácida da como resultado la pérdida de materia prima ya que la celulosa también es propensa a la hidrólisis ácida. Los aprestos de almidón también se pueden eliminar mediante el uso de peróxido de hidrógeno para degradar el almidón por oxidación. La retirada del apresto también puede ser un proceso enzimático. Durante muchos años se han usado amilasas en la industria textil para la eliminación de aprestos de almidón. Algunas condiciones (por ejemplo, pH y temperatura) para la retirada enzimática de apresto son dictadas por las condiciones de funcionamiento de la enzima. La mayoría de las amilasas usadas en la solicitud son relativamente termoestables, sin embargo, son enzimas óptimas neutras o ácidas.

La "limpieza" es un proceso en el que el tejido de algodón al que se le retira el apresto se procesa para solubilizar y extraer material no celulósico indeseado encontrado de forma natural en el algodón y también para retirar impurezas aplicadas tales como lubricantes de maquinaria. La limpieza usa agentes altamente alcalinos para retirar el material no celulósico, que tiene un grave impacto ambiental. Además, los agentes químicos degradan parcialmente la celulosa en la fibra de algodón que causa una pérdida de resistencia de la fibra y materias primas y como tal no es un proceso óptimo. La etapa final en el proceso de tratamiento del tejido de algodón es el blanqueado en el que los pigmentos naturales y la materia presente en la fibra se blanquean. Una enzima pectinolítica alcalina termoestable que podría tener como objetivo de forma específica el material no celulósico podría reducir o eliminar el uso de agentes químicos rigurosos que disminuyen la carga en el entorno a la vez que mantienen la integridad y resistencia de la fibra de algodón.

Algunas divulgaciones mencionadas anteriormente en el examen incluyen el documento de patente WO 00/26464 que analiza pectato liasas y su termoestabilidad, así como las entradas AF279364 y Q9F7L3 en bases de datos que desvelan secuencias que tienen alguna coincidencia con las de la SEC ID N°: 77 y la SEC ID N°: 78, pero que están fuera del alcance de la presente invención.

Sumario

La invención proporciona ácidos nucleicos sintéticos o recombinantes, aislados que comprenden una secuencia de ácidos nucleicos que tiene una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más, o completa (100 %) con la SEC ID N°: 77, sobre una región de al menos aproximadamente 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1550, 1600, 1650, 1700, 1750, 1800, 1850, 1900, 1950, 2000, 2050, 2100, 2200, 2250, 2300, 2350, 2400, 2450, 2500, o más restos, que codifica al menos un polipéptido que tiene una actividad de pectato, y las identidades de las secuencias se determinan mediante análisis con un algoritmo de comparación de secuencias o mediante inspección visual.

Algunos ácidos nucleicos a modo de ejemplo de la invención también incluyen ácidos nucleicos aislados o recombinantes que codifican un polipéptido que tiene una secuencia como se establece en la SEC ID N°: 78, y ciertas subsecuencias de la misma y variantes de la misma. El polipéptido tiene una actividad de pectato liasa.

En un aspecto, la invención también proporciona dichos ácidos nucleicos que codifican pectato liasa con una novedad común por que se derivan de cultivos mixtos.

En un aspecto, la invención también proporciona dichos ácidos nucleicos que codifican pectato liasa con una novedad común por que se derivan de fuentes ambientales, por ejemplo, fuentes ambientales mixtas.

En un aspecto, el algoritmo de comparación de secuencias es un algoritmo BLAST en la versión 2.2.2 en el que una configuración de filtrado se ajusta para blastall -p blastp -d "nr pataa" -F F, y todas las otras opciones se ajustan por defecto.

La actividad de pectato liasa puede comprender catálisis de beta-eliminación (trans-eliminación) o hidrólisis de pectina o ácido poligalacturónico (pectato). La actividad de pectato liasa puede comprender la descomposición o disolución de paredes celulares vegetales. La actividad de pectato liasa puede comprender beta-eliminación (trans-eliminación) o hidrólisis de ácido alfa-D-galacturónico unido en la posición 1,4. La actividad de pectato liasa puede comprender catálisis de beta-eliminación (trans-eliminación) o hidrólisis de ácido galacturónico esterificado con metilo. La actividad de pectato liasa puede ser de acción exo o de acción endo. En un aspecto, la actividad de pectato liasa tiene actuación endo y actúa en sitios aleatorios dentro de una cadena del polímero para dar una mezcla de oligómeros. En un aspecto, la actividad de pectato liasa tiene actuación exo y actúa desde un extremo de una cadena de polímero y produce en monómeros o dímeros. La actividad de pectato liasa puede catalizar la escisión aleatoria de enlaces alfa-1,4-glicosídicos en el ácido péctico (ácido poligalacturónico) mediante trans-eliminación por hidrólisis. La actividad de pectato liasa puede comprender una actividad igual o similar a la de la pectato liasa (EC 4.2.2.2), poli(1,4-alfa-D-galacturónido) liasa, poligalacturonato liasa (EC 4.2.2.2), pectina liasa (EC 4.2.2.10), poligalacturonasa (EC 3.2.1.15), exo-poligalacturonasa (EC 3.2.1.67), exo-poligalacturonato liasa (EC 4.2.2.9) o exo-poli-alfa-galacturonosidasa (EC 3.2.1.82). La actividad de pectato liasa puede comprender beta-eliminación (trans-eliminación) o hidrólisis de galactano en galactosa o galactooligómeros. La actividad de pectato liasa puede comprender beta-eliminación (trans-eliminación) o hidrólisis de una fibra vegetal. La fibra vegetal puede comprender fibra de algodón, fibra de cáñamo o fibra de lino.

El ácido nucleico aislado o recombinante puede codificar un polipéptido que tiene una actividad de pectato liasa que es termoestable. El polipéptido puede mantener una actividad de pectato liasa en condiciones que comprenden un intervalo de temperatura entre aproximadamente 37 °C y aproximadamente 95 °C; entre aproximadamente 55 °C y aproximadamente 85 °C, entre aproximadamente 70 °C y aproximadamente 95 °C, o, entre aproximadamente 90 °C y aproximadamente 95 °C.

El ácido nucleico aislado o recombinante puede codificar un polipéptido que tiene una actividad de pectato liasa que es termotolerante. El polipéptido puede mantener una actividad de pectato liasa después de la exposición a una temperatura en el intervalo de superior a 37 °C a aproximadamente 95 °C o en cualquier parte en el intervalo de superior a 55 °C a aproximadamente 85 °C. En un aspecto, el polipéptido mantiene una actividad de pectato liasa después de exposición a una temperatura en el intervalo de superior a 90 °C a aproximadamente 95 °C a pH 4,5.

La invención proporciona ácidos nucleicos aislados o recombinantes que comprenden una secuencia que se hibrida en condiciones rigurosas a un ácido nucleico de acuerdo con la invención. Las condiciones rigurosas incluyen una etapa de lavado que comprende un lavado en 0,2X SSC a una temperatura de aproximadamente 65 °C durante aproximadamente 15 minutos.

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

La invención proporciona una sonda de ácido nucleico para identificar un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa, en la que la sonda comprende un ácido nucleico que comprende una secuencia de aproximadamente 60 a 100 o de aproximadamente 60 a 150 bases consecutivas de una secuencia de ácidos nucleicos de la invención: SEC ID N°: 77, SEC ID N°: 131 o SEC ID N°: 133.

Un par de la secuencia cebadora de amplificación para amplificar un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa puede ser capaz de amplificar un ácido nucleico que comprende una secuencia

de la invención, o fragmentos o subsecuencias de la misma. Uno o cada miembro del par de la secuencia cebadora de amplificación puede comprender un oligonucleótido que comprende al menos de aproximadamente 10 a 50 bases consecutivas de la secuencia.

- 5 Algunos métodos para la amplificación de un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa pueden comprender la amplificación de un ácido nucleico molde con tal parte la secuencia cebadora de amplificación.

10 La invención proporciona casetes de expresión que comprenden un ácido nucleico de la invención o una subsecuencia del mismo. En un aspecto, el casete de expresión puede comprender el ácido nucleico que se une de forma operativa a un promotor. El promotor puede ser un promotor viral, bacteriano, mamífero o vegetal. En un aspecto, el promotor vegetal puede ser un promotor de patata, arroz, maíz, trigo, tabaco o cebada. El promotor puede ser un promotor constitutivo. El promotor constitutivo puede comprender CaMV35S. En otro aspecto, el promotor puede ser un promotor inducible. En un aspecto, el promotor puede ser un promotor específico de tejido o un promotor regulado de forma ambiental o un promotor regulado por el desarrollo. Por lo tanto, el promotor puede ser, por ejemplo, un promotor específico de semilla, uno específico de hoja, uno específico de raíz, uno específico de tallo o un promotor inducido por abscisión. En un aspecto, el casete de expresión puede comprender adicionalmente un vector de expresión vegetal o de virus vegetal.

20 La invención proporciona vehículos de clonación que comprenden en un casete de expresión (por ejemplo, un vector) de la invención o un ácido nucleico de la invención. El vehículo de clonación puede ser un vector viral, un plásmido, un fago, un fagémido, un cósmido, un fósido, un bacteriófago o un cromosoma artificial. El vector viral puede comprender un de vector adenovirus, un vector retroviral o un vector viral adeno-asociado. El vehículo de clonación puede comprender un cromosoma artificial bacteriano (BAC), un plásmido, un vector derivado de bacteriófago P1 (PAC), un cromosoma artificial de levadura (YAC), o un cromosoma artificial de mamífero (MAC).

25 La invención proporciona célula transformada que comprende un ácido nucleico de la invención o un casete de expresión (por ejemplo, un vector) de la invención, o un vehículo de clonación de la invención. En un aspecto, la célula transformada puede ser una célula bacteriana, una célula de mamífero, una célula fúngica, una célula de levadura, una célula de insecto o una célula vegetal. En un aspecto, la célula vegetal puede ser una célula de patata, trigo, arroz, maíz, tabaco o cebada.

30 La invención proporciona animales no humanos transgénicos que comprenden un ácido nucleico de la invención o un casete de expresión (por ejemplo, un vector) de la invención. En un aspecto, el animal es un ratón.

35 La invención proporciona plantas transgénica se comprenden un ácido nucleico de la invención o un casete de expresión (por ejemplo, un vector) de la invención. La planta transgénica puede ser una planta de maíz, una planta de patata, una planta de tomate, una planta de trigo, una planta de colza oleaginosa, una planta de colza, una planta de soja, una planta de arroz, una planta de cebada o una planta de tabaco.

40 La invención proporciona semillas transgénica se comprenden un ácido nucleico de la invención o un casete de expresión (por ejemplo, un vector) de la invención. La semilla transgénica puede ser una semilla de maíz, una semilla de trigo, una semilla de colza oleaginosa, una semilla de colza, una semilla de soja, una semilla de palma, una semilla de girasol, una semilla de sésamo, una semilla de cacahuete o una semilla de planta de tabaco.

45 Un oligonucleótido antisentido puede comprender una secuencia de ácidos nucleicos complementaria, capaz de hibridarse en condiciones rigurosas con un ácido nucleico de la invención. Algunos métodos para inhibir la traducción de un mensaje de pectato liasa en una célula pueden comprender la administración a la célula o la expresión en la célula de un oligonucleótido antisentido que comprende una secuencia de ácidos nucleicos complementaria, no capaz de hibridarse en condiciones rigurosas con un ácido nucleico de la invención.

50 La invención proporciona un polipéptido aislado o recombinante que comprende una secuencia de aminoácidos que tiene una identidad de secuencias de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más, o completa (100 %) con la SEC ID N°: 78 sobre una región de al menos 350 o más restos, o sobre la longitud completa del polipéptido, y las identidades de las secuencias se determinan mediante análisis con un algoritmo de comparación de secuencias o mediante una inspección visual. Un péptido de la divulgación puede ser, por ejemplo, un fragmento inmunogénico, un motivo (por ejemplo, un sitio de unión), una secuencia señal, una secuencia prepro o un sitio activo. Algunas secuencias de polipéptidos o péptidos a modo de ejemplo de la invención incluyen secuencias codificada por un ácido nucleico de la invención. Algunas secuencias de polipéptidos o péptidos a modo de ejemplo incluyen polipéptidos o péptidos que se unen de forma específica mediante un anticuerpo de la divulgación. El polipéptido aislado o recombinante de la invención (con o sin secuencia señal) tiene actividad de pectato liasa.

60 En un aspecto, el polipéptido aislado o recombinante puede comprender el polipéptido de la invención que carece de una secuencia señal. En un aspecto, el polipéptido aislado o recombinante puede comprender el polipéptido de la

invención que comprende una secuencia señal heteróloga, tal como una señal de secuencia de pectato liasa o no de pectato liasa heteróloga.

5 En un aspecto, la invención proporciona proteínas quiméricas que comprenden un primer dominio que comprende una secuencia señal de la invención (por ejemplo, como se establece en la Tabla 2, que sigue a continuación) y al menos un segundo dominio. La proteína puede ser una proteína de fusión. The el segundo dominio puede comprender una enzima. La enzima puede ser una pectato liasa.

10 En un aspecto, la actividad de pectato liasa comprende una actividad específica a aproximadamente 37 °C en el intervalo de aproximadamente 100 a aproximadamente 1000 unidades por miligramo de proteína. En otro aspecto, la actividad de pectato liasa comprende una actividad específica de aproximadamente 500 a aproximadamente 750 unidades por miligramo de proteína. Como alternativa, la actividad de pectato liasa comprende una actividad específica a 37 °C en el intervalo de aproximadamente 500 a aproximadamente 1200 unidades por miligramo de proteína. En un aspecto, la actividad de pectato liasa comprende una actividad específica a 37 °C en el intervalo de aproximadamente 750 a aproximadamente 1000 unidades por miligramo de proteína. En otro aspecto, la termotolerancia comprende la retención de al menos la mitad de la actividad específica de la pectato liasa a 37 °C después de ser calentada a la temperatura elevada. Como alternativa, la termotolerancia puede comprender la retención de la actividad específica a 37 °C en el intervalo de aproximadamente 500 a aproximadamente 1200 unidades por miligramo de proteína después de ser calentada a la temperatura elevada.

20 La invención proporciona el polipéptido aislado o recombinante de la invención, en la que el polipéptido comprende al menos un sitio de glicosilación. En un aspecto, la glicosilación puede ser una glicosilación unida a N. En un aspecto, el polipéptido se puede glicosilar después de su expresión en una *P. pastoris* o una *S. pombe*.

25 El polipéptido puede mantener una actividad de pectato liasa en condiciones que comprenden aproximadamente pH 6,5, pH 6, pH 5,5, pH 5, pH 4,5 o pH 4. Como alternativa, el polipéptido puede mantener una actividad de pectato liasa en condiciones que comprenden aproximadamente pH 7, pH 7,5, pH 8,0, pH 8,5, pH 9, pH 9,5, pH 10, pH 10,5 o pH 11.

30 En un aspecto, el polipéptido aislado o recombinante puede comprender el polipéptido de la invención que carece de una secuencia señal y/o un dominio prepro. En un aspecto, el polipéptido aislado o recombinante puede comprender el polipéptido de la invención que comprende una secuencia señal heteróloga y/o dominio prepro, tal como una secuencia señal de pectato liasa heteróloga.

35 Una secuencia señal puede comprender un péptido que comprende/ que consiste en una secuencia como se expone en los restos 1 a 12, 1 a 13, 1 a 14, 1 a 15, 1 a 16, 1 a 17, 1 a 18, 1 a 19, 1 a 20, 1 a 21, 1 a 22, 1 a 23, 1 a 24, 1 a 25, 1 a 26, 1 a 27, 1 a 28, 1 a 28, 1 a 30, 1 a 31, 1 a 32, 1 a 33, 1 a 34, 1 a 35, 1 a 36, 1 a 37, 1 a 38, 1 a 39, 1 a 40, 1 a 41, 1 a 42, 1 a 43, 1 a 44 de un polipéptido de la invención. Algunas proteínas quiméricas pueden comprender un primer dominio que comprende tal secuencia señal y al menos un segundo dominio. La proteína puede ser una proteína de fusión. El segundo dominio puede comprender una enzima. La enzima puede ser una pectato liasa, por ejemplo, una enzima de la invención.

45 La invención proporciona polipéptidos quiméricos que comprenden al menos un primer dominio que comprende péptido señal (SP), un dominio prepro, un dominio catalítico (CD), o an a un sitio activo de una pectato liasa de la invención y al menos un segundo dominio que comprende un polipéptido o péptido heterólogo, en el que el polipéptido o péptido heterólogo no está asociado de forma natural con el péptido señal (SP), dominio prepro o dominio catalítico (CD). En un aspecto, el polipéptido o péptido heterólogo no es una pectato liasa. El polipéptido o péptido heterólogo pueden ser amino terminal con respecto a, carboxi terminal con respecto a o estar en ambos extremos del péptido señal (SP), dominio prepro o dominio catalítico (CD).

50 La invención proporciona preparaciones de proteína que comprenden un polipéptido de la invención, en la que la preparación de proteína comprende un líquido, un sólido o un gel.

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

60 La invención proporciona polipéptidos inmovilizados que tienen una actividad de pectato liasa, en la que el polipéptido comprende un polipéptido de la invención, un polipéptido codificado por un ácido nucleico de la invención, o un polipéptido que comprende un polipéptido de la invención y un segundo dominio. En un aspecto, el polipéptido puede estar inmovilizado en una célula, un metal, una resina, un polímero, una cerámica, un vidrio, un microelectrodo, una partícula de grafito, una perla, un gel, una placa, una matriz o un tubo capilar.

65 Algunas matrices pueden comprender un ácido nucleico inmovilizado de la invención o un anticuerpo de la divulgación.

Algunos anticuerpos aislados o recombinantes se pueden unir de forma específica a un polipéptido de la invención o a un polipéptido codificado por un ácido nucleico de la invención. El anticuerpo puede ser un anticuerpo monoclonal o un anticuerpo policlonal. Algunos hibridomas pueden comprender tal anticuerpo.

5 La invención proporciona suplementos alimenticios para un animal que comprenden un polipéptido de la invención, por ejemplo, un polipéptido codificado por el ácido nucleico de la invención. En un aspecto, el polipéptido en el suplemento alimenticio puede estar glicosilado. La invención proporciona matrices de administración de enzimas comestibles que comprenden un polipéptido de la invención, por ejemplo, un polipéptido codificado por el ácido nucleico de la invención. En un aspecto, la matriz de administración comprende un gránulo. En un aspecto, el polipéptido puede estar glicosilado. En un aspecto, la actividad de pectato liasa es termotolerante. En otro aspecto, la actividad de pectato liasa es termoestable.

Un método para aislar o identificar un polipéptido que tiene una actividad de pectato liasa comprende las etapas de: (a) proporcionar un anticuerpo de la divulgación; (b) proporcionar una muestra que comprende polipéptidos; y (c) poner en contacto la muestra de la etapa (b) con el anticuerpo de la etapa (a) en condiciones en las que el anticuerpo se puede unir de forma específica con el polipéptido, aislando o identificando de este modo un polipéptido que tiene una actividad de pectato liasa.

Algunos métodos para preparar un anticuerpo anti-pectato liasa comprenden la administración a un animal no humano un ácido nucleico de la invención o un polipéptido de la invención o subsecuencias del mismo en una cantidad suficiente para generar una respuesta inmune humoral, preparando de este modo un anticuerpo anti-pectato liasa. Algunos métodos para preparar una anti-pectato liasa inmune comprenden la administración a un animal no humano de un ácido nucleico de la invención o un polipéptido de la invención o subsecuencias del mismo en una cantidad suficiente para generar una respuesta inmune.

La invención proporciona métodos para producir un polipéptido recombinante que comprenden las etapas de: (a) proporcionar un ácido nucleico de la invención unido de forma operativa a un promotor; y (b) expresar el ácido nucleico de la etapa (a) en condiciones que permiten la expresión del polipéptido, produciendo de este modo un polipéptido recombinante. En un aspecto, el método puede comprender adicionalmente la transformación de una célula hospedadora con el ácido nucleico de la etapa (a) seguido de la expresión del ácido nucleico de la etapa (a), produciendo de este modo un polipéptido recombinante en una célula transformada.

Algunos métodos para identificar un polipéptido que tiene una actividad de pectato liasa comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención; o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar un sustrato de pectato liasa; y (c) poner en contacto el polipéptido o un fragmento o variante del mismo de la etapa (a) con el sustrato de la etapa (b) y detectar una disminución en la cantidad de sustrato o un aumento en la cantidad de un producto de reacción, en los que una disminución en la cantidad del sustrato o un aumento en la cantidad del producto de reacción detecta un polipéptido que tiene una actividad de pectato liasa.

Algunos métodos para identificar un sustrato de pectato liasa comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención; o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar un sustrato de ensayo; y (c) poner en contacto el polipéptido de la etapa (a) con el sustrato de ensayo de la etapa (b) y detectar una disminución en la cantidad de sustrato o un aumento en la cantidad de producto de reacción, en los que una disminución en la cantidad del sustrato o un aumento en la cantidad de un producto de reacción identifica el sustrato de ensayo como un sustrato de pectato liasa.

Algunos métodos para determinar si un compuesto de ensayo se une de forma específica a un polipéptido comprenden las siguientes etapas: (a) expresar un ácido nucleico o un vector que comprende el ácido nucleico en condiciones que permiten la traducción del ácido nucleico a un polipéptido, en los que el ácido nucleico comprende un ácido nucleico de la invención, o, proporcionar un polipéptido de la invención; (b) proporcionar un compuesto de ensayo; (c) poner en contacto el polipéptido con el compuesto de ensayo; y (d) determinar si el compuesto de ensayo de la etapa (b) se une de forma específica al polipéptido.

Algunos métodos para identificar un modulador de una actividad de pectato liasa comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar un compuesto de ensayo; (c) poner en contacto el polipéptido de la etapa (a) con el compuesto de ensayo de la etapa (b) y medir una actividad en la pectato liasa, en los que un cambio en la actividad de pectato liasa medido en presencia del compuesto de ensayo en comparación con la actividad en ausencia del compuesto de ensayo proporciona una determinación de que el compuesto de ensayo modula la actividad de pectato liasa. La actividad de pectato liasa se puede medir proporcionando un sustrato de pectato liasa y detectando una disminución en la cantidad del sustrato o un aumento en la cantidad de un producto de reacción, o, un aumento en la cantidad del sustrato o una disminución en la cantidad de un producto de reacción. Una disminución en la cantidad del sustrato o un aumento en la cantidad del producto de reacción con el compuesto de ensayo en comparación con la cantidad de sustrato o producto de reacción sin el compuesto de ensayo identifican el compuesto de ensayo como un activador de la actividad de pectato liasa. Un aumento en la cantidad del sustrato o una disminución en la

cantidad del producto de reacción con el compuesto de ensayo en comparación con la cantidad de sustrato o producto de reacción sin el compuesto de ensayo identifican el compuesto de ensayo como un inhibidor de la actividad de pectato liasa.

5 Algunos sistemas informáticos pueden comprender un procesador y un dispositivo de almacenamiento de datos en los que dicho dispositivo de almacenamiento de datos ha almacenado en el mismo una secuencia de polipéptidos o una secuencia de ácidos nucleicos de la invención (por ejemplo, un polipéptido codificado por un ácido nucleico de la invención). El sistema informático puede comprender adicionalmente un algoritmo de comparación de secuencias y un dispositivo de almacenamiento de datos que tiene al menos una secuencia de referencia almacenada en el mismo. El algoritmo de comparación de secuencias puede comprender un programa informático que indica polimorfismos. El sistema informático puede comprender adicionalmente un identificador que identifica una o más características en dicha secuencia. Algunos medios de lectura en soporte informático pueden tener almacenada en el mismo una secuencia de polipéptidos o una secuencia de ácidos nucleicos de la invención. Algunos métodos para identificar una característica en una secuencia comprende las etapas de: (a) lectura de la secuencia usando un programa informático que identifica una o más características en una secuencia, en el que la secuencia comprende una secuencia de polipéptidos o secuencia de ácidos nucleicos de la invención; y (b) identificación de una o más características en la secuencia con el programa informático. Algunos métodos para comparar una primera secuencia con una segunda secuencia comprende las etapas de: (a) lectura de la primera secuencia y la segunda secuencia a través del uso de un programa informático que compara secuencias, en el que la primera secuencia comprende una secuencia de polipéptidos o secuencia de ácidos nucleicos de la invención; y (b) determinación de diferencias entre la primera secuencia y la segunda secuencia con el programa informático. La etapa de determinación de diferencias entre la primera secuencia y la segunda secuencia puede comprender adicionalmente la etapa de identificación de polimorfismos. El método puede comprender adicionalmente un identificador que identifica una o más características en una secuencia. El método puede comprender la lectura de la primera secuencia usando un programa informático en identificación de una o más características en la secuencia.

Algunos métodos para aislar o recuperar un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa a partir de una muestra ambiental comprende las etapas de: (a) proporcionar una secuencia cebadora de amplificación para amplificar un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa, en la que el par cebador es capaz de amplificar un ácido nucleico de la invención; (b) aislar un ácido nucleico a partir de la muestra ambiental o tratar la muestra ambiental de modo que el ácido nucleico en la muestra es accesible a la hibridación con el par cebador de amplificación; y, (c) combinar el ácido nucleico de la etapa (b) con el par cebador de amplificación de la etapa (a) y amplificar ácido nucleico a partir de la muestra ambiental, aislando o recuperando de este modo un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa a partir de una muestra ambiental. Uno o cada miembro de la secuencia cebadora de amplificación puede comprender un oligonucleótido que comprende al menos de aproximadamente 10 a 50 bases consecutivas de una secuencia de la invención.

Algunos métodos para aislar o recuperar un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa a partir de una muestra ambiental comprende las etapas de: (a) proporcionar una sonda de polinucleótidos que comprende un ácido nucleico de la invención o una subsecuencia del mismo; (b) aislar un ácido nucleico a partir de la muestra ambiental o tratar la muestra ambiental de modo que el ácido nucleico en la muestra es accesible para hibridación con una sonda de polinucleótidos de la etapa (a); (c) combina el ácido nucleico aislado o la muestra ambiental tratada de la etapa (b) con la sonda de polinucleótidos de la etapa (a); y (d) aislar un ácido nucleico se hibrida en forma específica con la sonda de polinucleótidos de la etapa (a), aislando o recuperando de este modo un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa a partir de una muestra ambiental. La muestra ambiental puede comprender una muestra de agua, una muestra de líquido, una muestra de suelo, una muestra de aire o una muestra biológica. En un aspecto, la muestra biológica se puede derivar una célula bacteriana, una célula de protozoo, una célula insecto, una célula de levadura, una célula vegetal, una célula fúngica o una célula de mamífero.

Algunos métodos para generar una variante de un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa comprenden las etapas de: (a) proporcionar un ácido nucleico molde que comprende un ácido nucleico de la invención; y (b) modificar, suprimir o añadir uno o más nucleótidos en la secuencia molde, o una combinación de los mismos, para generar una variante del ácido nucleico molde. El método puede comprender adicionalmente la expresión de la variante de ácido nucleico para generar una variante de polipéptido pectato liasa. Las modificaciones, adiciones o supresiones se pueden introducir mediante un método que comprende PCR propensa a error, redistribución, mutagénesis dirigida a oligonucleótidos, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casete, mutagénesis de conjunto recursiva, mutagénesis de conjunto exponencial, mutagénesis específica de sitio, reensamblaje genético, mutagénesis por saturación de sitio genético (GSSM™), reensamblaje de ligamiento genético (SLR) o una combinación de los mismos. En otro aspecto, las modificaciones, adiciones o supresiones se introducen mediante un método que comprende recombinación, recombinación de secuencias recursiva, mutagénesis de ADN modificada con fosfotioato, mutagénesis de molde que contiene uracilo, mutagénesis de dúplex con huecos, mutagénesis de reparación de falta de coincidencia de puntos, mutagénesis de cepa hospedadora con déficit de reparación, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de restricción-selección, mutagénesis de restricción-

purificación, síntesis de genes artificiales, mutagénesis de conjunto, creación de multímeros de ácidos nucleicos quiméricos Y una combinación de los mismos.

El método se puede repetir de forma repetitiva hasta que se produce una pectato liasa que tiene una actividad alterada o diferente o una estabilidad alterada o diferente de la de un polipéptido codificado por el ácido nucleico molde. La variante de polipéptido pectato liasa puede ser termotolerante, y mantiene una cierta actividad después de su exposición a una temperatura elevada. La variante de polipéptido pectato liasa puede presentar un aumento de glicosilación en comparación con la pectato liasa codificada por un ácido nucleico molde. Como alternativa, la variante de polipéptido pectato liasa tiene una actividad de pectato liasa a una temperatura elevada, en la que la pectato liasa codificada por el ácido nucleico molde no es activa a la temperatura elevada. El método se puede repetir de forma repetitiva hasta que se produce una secuencia de codificación de pectato liasa que tiene un uso de codón alterado con respecto a la del ácido nucleico molde. El método se puede repetir de forma repetitiva hasta que se produce un gen de pectato liasa que tiene un nivel de expresión de mensajes mayor o menor o estabilidad que la del ácido nucleico molde.

Algunos métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa, para aumentar su expresión en una célula hospedadora, comprenden las siguientes etapas: (a) proporcionar un ácido nucleico de la invención que codifica un polipéptido que tiene una actividad de pectato liasa; y, (b) identificar un codón no preferente o uno menos preferente en el ácido nucleico de la etapa (a) y sustituirlo con un codón usado preferente o usado de forma neutra que codifica el mismo aminoácido que el del codón sustituido, en los que un codón preferente es un codón que se sobrerrepresenta en secuencias de codificación en genes en la célula hospedadora y un codón no preferente o menos preferente es un codón que se subrepresenta en secuencias de codificación en genes en la célula hospedadora, modificando de este modo al ácido nucleico para aumentar su expresión en una célula hospedadora.

Algunos métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa comprenden las siguientes etapas: (a) proporcionar un ácido nucleico de la invención; y, (b) identificar un codón en el ácido nucleico de la etapa (a) y sustituirlo con un codón diferente que codifica el mismo aminoácido que el del codón sustituido, modificando de este modo los codones en un ácido nucleico que codifican una pectato liasa.

Algunos métodos para modificar codones en un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa para aumentar su expresión en una célula hospedadora, comprenden las siguientes etapas: (a) proporcionar un ácido nucleico de la invención que codifica un polipéptido pectato liasa; y, (b) identificar un codón no preferente o uno menos preferente en el ácido nucleico de la etapa (a) y sustituirlo con un codón usado preferente o usado de forma neutra que codifica el mismo aminoácido que el del codón sustituido, en los que un codón preferente es un codón que se sobrerrepresenta en secuencias de codificación en genes en la célula hospedadora y un codón no preferente o menos preferente es un codón que se subrepresenta en secuencias de codificación en genes en la célula hospedadora, modificando de este modo al ácido nucleico para aumentar su expresión en una célula hospedadora.

Algunos métodos para modificar un codón en un ácido nucleico que codifica un polipéptido que tiene una actividad de pectato liasa para disminuir su expresión en una célula hospedadora comprenden las siguientes etapas: (a) proporcionar un ácido nucleico de la invención; y (b) identificar al menos un codón preferente en el ácido nucleico de la etapa (a) y sustituirlo con un codón no preferente o menos preferente que codifica el mismo aminoácido que el del codón sustituido, en los que un codón preferente es un codón que se sobrerrepresenta en secuencias de codificación en genes en una célula hospedadora y un codón no preferente o menos preferente es un codón que se subrepresenta en secuencias de codificación en genes en la célula hospedadora, modificando de este modo al ácido nucleico para disminuir su expresión en una célula hospedadora. La célula hospedadora puede ser una célula bacteriana, una célula fúngica, una célula de insecto, una célula de levadura, una célula vegetal o una célula de mamífero.

Algunos métodos para producir una biblioteca de ácidos nucleicos que codifican una pluralidad de sitios activos de pectato liasa modificados o sitios de unión al sustrato, en los que los sitios activos modificados o sitios de unión a sustrato se derivan de un primer ácido nucleico que comprende una secuencia que codifica un primer sitio activo o un primer sitio de unión a sustrato, comprenden las siguientes etapas: (a) proporcionar un primer ácido nucleico que codifica un primer sitio activo o primer sitio de unión a sustrato, en los que la primera secuencia de ácidos nucleicos comprende una secuencia que se hibrida en condiciones rigurosas con un ácido nucleico de la invención, y el ácido nucleico codifica un sitio activo de pectato liasa o un sitio de unión a sustrato y pectato liasa; (b) proporcionar un conjunto de oligonucleótidos mutagénicos que codifican variantes de aminoácidos de origen natural en una pluralidad de codones dirigidos en el primer ácido nucleico; y, (c) usar el conjunto de oligonucleótidos mutagénicos para generar un conjunto de variantes de ácidos nucleicos que codifican sitio activo o que codifican sitio de unión a sustrato que codifica una serie de variaciones de aminoácidos en cada codón del aminoácido que se sometió a mutagénesis, reduciendo de este modo una biblioteca de ácidos nucleicos que codifican una pluralidad de sitios activos de pectato liasa o sitios de unión a sustrato. El método puede comprender la mutagénesis del primer ácido nucleico de la etapa (a) mediante un método que comprende un sistema de evolución dirigida optimizado,

mutagénesis por saturación de sitio genético (GSSM™), reensamblaje de ligamiento genético (SLR), PCR propensa a error, redistribución, mutagénesis dirigida a oligonucleótidos, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casete, mutagénesis de conjunto recursiva, mutagénesis de conjunto exponencial, mutagénesis específica de sitio, reensamblaje genético, mutagénesis por saturación de sitio genético (GSSM™), reensamblaje de ligamiento genético (SLR) y una combinación de los mismos. O, el método puede comprender la mutagénesis del primer ácido nucleico de la etapa (a) o variantes mediante un método que comprende recombinación, recombinación de secuencias recursiva, mutagénesis de ADN modificada con fosfotioato, mutagénesis de molde que contiene uracilo, mutagénesis de dúplex con huecos, mutagénesis de reparación de falta de coincidencia de puntos, mutagénesis de cepa hospedadora con déficit de reparación, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de restricción-selección, mutagénesis de restricción-purificación, síntesis de genes artificiales, mutagénesis de conjunto, creación de multímeros de ácidos nucleicos quiméricos y una combinación de los mismos.

Algunos métodos para preparar una molécula pequeña pueden comprender las siguientes etapas: (a) proporcionar una pluralidad de enzimas biosintéticas capaces de sintetizar o modificar una molécula pequeña, en los que una de las enzimas comprende una enzima pectato liasa codificada por un ácido nucleico de la invención; (b) proporcionar un sustrato para al menos una de las enzimas de la etapa (a); y (c) hacer reaccionar el sustrato de la etapa (b) con las enzimas en condiciones que facilitan una pluralidad de reacciones biocatalíticas para generar una molécula pequeña mediante una serie de reacciones biocatalíticas. Algunos métodos para modificar una molécula pequeña comprenden las siguientes etapas: (a) proporcionar una enzima pectato liasa, en los que la enzima comprende un polipéptido de la invención, o, un polipéptido codificado por un ácido nucleico de la invención, o una subsecuencia del mismo; (b) proporcionar una molécula pequeña; y (c) hacer reaccionar la enzima de la etapa (a) con la molécula pequeña de la etapa (b) en condiciones que facilitan una reacción enzimática catalizada por la enzima pectato liasa, modificando de este modo una molécula pequeña mediante una reacción enzimática de pectato liasa. El método puede comprender una pluralidad de sustratos de molécula pequeña para la enzima de la etapa (a), generando de este modo una biblioteca de moléculas pequeñas modificadas producida mediante al menos una reacción enzimática catalizada por la enzima pectato liasa. El método puede comprender una pluralidad de enzimas adicionales en condiciones que facilitan una pluralidad de reacciones biocatalíticas mediante las enzimas para formar una biblioteca de moléculas pequeñas modificadas producida mediante la pluralidad de reacciones enzimáticas. El método puede comprender adicionalmente la etapa de someter a ensayo la biblioteca para determinar si una molécula pequeña modificada en particular que presenta una actividad deseada está presente dentro de la biblioteca. La etapa de someter a ensayo la biblioteca puede comprender adicionalmente las etapas de eliminar de forma sistemática todas excepto una de las reacciones biocatalíticas usadas para producir una parte de la pluralidad de las moléculas pequeñas modificadas dentro de la biblioteca sometiendo al ensayo la parte de la molécula pequeña modificada para la presencia o ausencia de la molécula pequeña modificada en particular con una actividad deseada, e identificar al menos una reacción biocatalítica específica que produce la molécula pequeña modificada en particular de actividad deseada.

Algunos métodos para determinar un fragmento funcional de una enzima pectato liasa pueden comprender las etapas de: (a) proporcionar una enzima pectato liasa, en los que la enzima comprende un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención, o una subsecuencia del mismo; y (b) suprimir una pluralidad de restos de aminoácidos de la secuencia de la etapa (a) y someter a ensayo la subsecuencia restante para una actividad de pectato liasa, determinando de este modo un fragmento funcional de una enzima pectato liasa. La actividad de pectato liasa se puede medir proporcionando un sustrato de pectato liasa y detectando una disminución en la cantidad del sustrato o un aumento en la cantidad de un producto de reacción.

Algunos métodos para ingeniería de células completas de serotipos nuevos modificados mediante el uso de análisis de flujo metabólico en tiempo real comprenden las siguientes etapas: (a) preparar una célula modificada mediante la modificación de la composición genética de una célula, en los que la composición genética se modifica mediante la adición a las células de un ácido nucleico de la invención; (b) cultivar la célula modificada para generar una pluralidad de células modificadas; (c) medir al menos un parámetro metabólico de la célula controlando el cultivo celular de la etapa (b) en tiempo real; y, (d) analizar los datos de la etapa (c) para determinar si el parámetro medido diferencia de una medida comparable en una célula sin modificar en condiciones similares, identificando de este modo un fenotipo modificado por ingeniería en la célula usando análisis de flujo metabólico en tiempo real. La composición genética de la célula se puede modificar mediante un método que comprende la supresión de una secuencia o la modificación de una secuencia en la célula, o, genosupresión de la expresión de un gen. El método puede comprender adicionalmente seleccionar una célula que comprende un fenotipo recién modificado por ingeniería.

El método puede comprender el cultivo de la célula seleccionada, generando de este modo una nueva cepa celular que comprende un fenotipo recién modificado por ingeniería.

Algunos métodos para aumentar la termotolerancia o la termoestabilidad de un polipéptido pectato liasa, pueden comprender la glicosilación de un polipéptido pectato liasa, en los que el polipéptido comprende al menos treinta aminoácidos contiguos de un polipéptido de la invención; o un polipéptido codificado por secuencia de ácidos nucleicos de la invención, aumentando de este modo la termotolerancia o la termoestabilidad del polipéptido pectato

liasa. La actividad específica de pectato liasa puede ser termoestable o termotolerante a una temperatura en el intervalo de superior a aproximadamente 37 °C a aproximadamente 95 °C.

Algunos métodos para sobreexpresar una polipéptido pectato liasa recombinante en una célula pueden comprender la expresión de un vector que comprende un ácido nucleico que comprende un ácido nucleico de la invención o secuencia de ácidos nucleicos de la invención, en los que las identidades de las secuencias se determinan por análisis con un algoritmo de comparación de secuencias o mediante inspección visual, en los que la sobreexpresión se realiza mediante el uso de un promotor de actividad elevada, un vector dicistrónico o mediante la amplificación genética del vector.

Algunos métodos para preparar una planta transgénica comprenden las siguientes etapas: (a) introducir una secuencia de ácidos nucleicos heteróloga en la célula, en los que la secuencia de ácidos nucleicos heteróloga comprende la secuencia de ácidos nucleicos de la invención, produciendo de este modo una célula vegetal transformada; y (b) producir una planta transgénica a partir de la célula transformada. La etapa (a) puede comprender adicionalmente la introducción de la secuencia de ácidos nucleicos heteróloga mediante electroporación o microinyección de protoplastos de célula vegetal. La etapa (a) puede comprender adicionalmente la introducción de la secuencia de ácidos nucleicos heteróloga directamente en tejido vegetal mediante bombardeo de partículas de ADN. Como alternativa, la etapa (a) puede comprender adicionalmente la introducción de la secuencia de ácidos nucleicos heteróloga en el ADN de célula vegetal usando un hospedador de *Agrobacterium tumefaciens*. En un aspecto, la célula vegetal puede ser una célula de patata, maíz, arroz, trigo, tabaco, o cebada.

Algunos métodos para expresar una secuencia de ácidos nucleicos heteróloga en una célula vegetal comprenden las siguientes etapas: (a) transformar la célula vegetal con una secuencia de ácidos nucleicos heteróloga unida de forma operativa a un promotor, en los que la secuencia de ácidos nucleicos heteróloga comprende un ácido nucleico de la invención; (b) cultivar la planta en condiciones en las que la secuencia de ácidos nucleicos heteróloga se expresa en la célula vegetal.

Algunas secuencias de señal aisladas o recombinantes comprenden o consisten en péptidos señal (SP) como se establece en la Tabla 2. Algunas secuencias de señal aisladas o recombinantes consisten en una secuencia como se establece en los restos de 1 a 15, de 1 a 16, de 1 a 17, de 1 a 18, de 1 a 19, de 1 a 20, de 1 a 21, de 1 a 22, de 1 a 23, de 1 a 24, de 1 a 25, de 1 a 26, de 1 a 27, de 1 a 28, de 1 a 29, de 1 a 30, de 1 a 31, de 1 a 32, de 1 a 33, de 1 a 34, de 1 a 35, de 1 a 36, de 1 a 37, de 1 a 38, de 1 a 39, de 1 a 40, de 1 a 41, de 1 a 42, de 1 a 43, y/o de 1 a 44, de la SEC ID N°: 2, SEC ID N°: 4, SEC ID N°: 6, SEC ID N°: 8, SEC ID N°: 10, SEC ID N°: 12, SEC ID N°: 14, SEC ID N°: 16, SEC ID N°: 18, SEC ID N°: 20, SEC ID N°: 22, SEC ID N°: 24, SEC ID N°: 26, SEC ID N°: 28, SEC ID N°: 30, SEC ID N°: 32, SEC ID N°: 34, SEC ID N°: 36, SEC ID N°: 38, SEC ID N°: 40, SEC ID N°: 42, SEC ID N°: 44, SEC ID N°: 46, SEC ID N°: 48, SEC ID N°: 50, SEC ID N°: 52, SEC ID N°: 54, SEC ID N°: 56, SEC ID N°: 58, SEC ID N°: 60, SEC ID N°: 62, SEC ID N°: 64, SEC ID N°: 66, SEC ID N°: 68, SEC ID N°: 70, SEC ID N°: 72, SEC ID N°: 74, SEC ID N°: 76, SEC ID N°: 78, SEC ID N°: 80, SEC ID N°: 82, SEC ID N°: 84, SEC ID N°: 86, SEC ID N°: 88, SEC ID N°: 90, SEC ID N°: 92, SEC ID N°: 94, SEC ID N°: 96, SEC ID N°: 98, SEC ID N°: 100, SEC ID N°: 102, SEC ID N°: 104, SEC ID N°: 106, SEC ID N°: 108, SEC ID N°: 110, SEC ID N°: 112, SEC ID N°: 114, SEC ID N°: 116, SEC ID N°: 118, SEC ID N°: 120, SEC ID N°: 122, SEC ID N°: 124, SEC ID N°: 126, SEC ID N°: 128 o SEC ID N°: 130, SEC ID N°: 132, SEC ID N°: 134.

La invención proporciona péptidos aislados o recombinantes que consisten en un dominio de pectina metil esterasa (PED) o un dominio catalítico (CD) como se establece en la Tabla 2.

La invención proporciona polipéptidos quiméricos que comprenden al menos un primer dominio que comprende péptido señal (SP), un dominio de pectina metil esterasa (PED) o un dominio catalítico (CD) como se establece en la Tabla 2 y al menos un segundo dominio que comprende un polipéptido o péptido heterólogo, en los que el polipéptido o péptido heterólogo no está asociado de forma natural con el péptido señal (SP), dominio de pectina metil esterasa (PED) o dominio catalítico (CD). En un aspecto, el polipéptido o péptido heterólogo no es una pectato liasa. El polipéptido o péptido heterólogo puede ser amino terminal con respecto a, carboxi terminal con respecto a o estar en ambos extremos del péptido señal (SP), dominio de pectina metil esterasa (PED) o dominio catalítico (CD).

La invención proporciona ácidos nucleicos aislados o recombinantes que codifican un polipéptido quimérico, en los que el polipéptido quimérico comprende al menos un primer dominio que comprende péptido señal (SP), un dominio de pectina metil esterasa (PED) o un dominio catalítico (CD) como se establece en la Tabla 2 y al menos un segundo dominio que comprende un polipéptido o péptido heterólogo, en los que el polipéptido o péptido heterólogo no está asociado de forma natural con el péptido señal (SP), dominio de pectina metil esterasa (PED) o dominio catalítico (CD).

Algunos métodos para aumentar la termotolerancia o la termoestabilidad de una pectato liasa pueden comprender la glicosilación de una pectato liasa, en los que el polipéptido comprende al menos treinta aminoácidos contiguos de un polipéptido de la invención, aumentando de este modo la termotolerancia o la termoestabilidad de la pectato liasa.

Algunos métodos para la sobreexpresión de una pectato liasa recombinante en una célula comprenden la expresión de un vector que comprende un ácido nucleico de la invención, en los que la sobreexpresión se ve influida por el uso de un promotor de actividad elevada, un vector dicistrónico o mediante la amplificación genética del vector.

Algunos métodos para preparar una planta transgénica comprenden las siguientes etapas: (a) introducir una secuencia de ácidos nucleicos heteróloga en la célula, en la que la secuencia de ácidos nucleicos heteróloga comprende un ácido nucleico de la invención, produciendo de este modo una célula vegetal transformada; (b) producir una planta transgénica a partir de la célula transformada. En un aspecto, la etapa (a) comprende adicionalmente la introducción de la secuencia de ácidos nucleicos heteróloga mediante electroporación o microinyección de protoplastos de célula vegetal. La etapa (a) puede comprender la introducción de la secuencia de ácidos nucleicos heteróloga directamente ha tejido vegetal mediante bombardeo de partículas de ADN o mediante el uso de un hospedador de *Agrobacterium tumefaciens*.

Algunos métodos para expresar una secuencia de ácidos nucleicos heteróloga en una célula vegetal comprenden las siguientes etapas: (a) transformar la célula vegetal con una secuencia de ácidos nucleicos heteróloga unido de forma operativa a un promotor, en los que la secuencia de ácidos nucleicos heteróloga comprende una secuencia de la invención; (b) cultivar la planta en condiciones en las que la secuencia de ácidos nucleicos heteróloga se expresa en la célula vegetal.

La invención proporciona métodos para hidrolizar, descomponer o alterar una composición que comprende pectina o pectato (ácido poligalacturónico) que comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar una composición que comprende una pectina o un pectato; y (c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) en condiciones en las que el polipéptido hidroliza, descompone o altera la composición que comprende pectina o pectato. En un aspecto, la composición comprende una pared de célula vegetal o una pared de célula bacteriana. La planta puede ser una planta de algodón, una planta de cáñamo o una planta de lino.

La invención proporciona métodos para licuar o retirar una pectina o pectato (ácido poligalacturónico) de una composición que comprende las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar una composición que comprende una pectina o pectato (ácido poligalacturónico); y (c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) en condiciones en las que el polipéptido retira o licúa la pectina o pectato (ácido poligalacturónico).

La invención proporciona composiciones de detergente que comprenden un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención, en las que el polipéptido tiene una actividad de pectato liasa. En un aspecto, la pectato liasa es una pectato liasa que no tiene superficie activa o una pectato liasa de superficie activa. La pectato liasa se puede formular en una composición líquida no acuosa, un sólido fundido, una forma granular, una forma en partículas, un comprimido formado por compresión, una forma de gel, una pasta o una forma de suspensión.

La invención proporciona métodos para lavar un objeto que comprende las siguientes etapas: (a) proporcionar una composición que comprende un polipéptido de la invención que tiene una actividad de pectato liasa; (b) proporcionar un objeto; y (c) poner en contacto el polipéptido de la etapa (a) y el objeto de la etapa (b) en condiciones en las que la composición puede lavar el objeto.

La invención proporciona textiles o tejidos que comprenden un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención. La invención proporciona métodos para la limpieza de fibra, hilo, textil o tejido que comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar una fibra, un hilo, un textil o un tejido; y (c) poner en contacto el polipéptido de la etapa (a) y el textil o tejido de la etapa (b) en condiciones en las que la pectato liasa puede limpiar la fibra, hilo, textil tejido. En un aspecto, la pectato liasa es una pectato liasa alcalina activa y termoestable. Los tratamientos para retirar el apresto y limpieza se pueden combinar en un solo baño. El método puede comprender adicionalmente la adición de una amilasa alcalina y termoestable en la puesta en contacto de la etapa (c). Los tratamientos para retirar el apresto o limpieza pueden comprender condiciones entre aproximadamente pH 8,5 y pH 10,0 y temperaturas de aproximadamente 40 °C. El método puede comprender adicionalmente la adición de una etapa de blanqueado. Los tratamientos para retirar el apresto, limpieza y blanqueado se pueden realizar de forma simultánea o de forma secuencial en un recipiente de un solo baño. El tratamiento de blanqueado puede comprender peróxido de hidrógeno o al menos un compuesto peroxi que puede generar peróxido de hidrógeno cuando se disuelve en agua, o combinaciones de los mismos, y al menos un activador del blanqueado. La fibra, hilo, textil o tejido pueden comprender un material celulósico. El material celulósico puede comprender una fibra en bruto, un hilo, un textil tejido o de punto, un algodón, un hilo de lino, una fibra de lino, un ramio, un rayón, un cáñamo, un yute o una mezcla de fibras naturales o sintéticas.

También se desvelan comidas o alimentos que comprenden un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención. La invención proporciona métodos para mejorar la extracción de un material vegetal rico en aceite que comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b)

proporcionaron material vegetal rico en aceite; y (c) poner en contacto el polipéptido de la etapa (a) y el material vegetal rico en aceite. En un aspecto, el material vegetal rico en aceite comprende una semilla rica en aceite. El aceite puede ser un aceite de soja, un aceite de oliva, un aceite de colza (canola) o un aceite de girasol.

La invención proporciona métodos para preparar un zumo, jarabe, puré o extracto de frutas o vegetales que comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar una composición o un líquido que comprende un material de fruta o vegetal; y (c) poner en contacto el polipéptido de la etapa (a) y la composición, preparando de este modo el zumo, jarabe, puré o extracto de frutas o vegetales.

La invención proporciona papeles o productos de papel o pastas de papel que comprenden una pectato liasa de la invención, o un polipéptido codificado por un ácido nucleico de la invención. La invención proporciona métodos para el tratamiento de un papel o una pasta de papel o de madera que comprenden las siguientes etapas: (a) proporcionar un polipéptido de la invención que tiene una actividad de pectato liasa, o un polipéptido codificado por un ácido nucleico de la invención; (b) proporcionar una composición que comprende un papel o una pasta de papel o de madera; y (c) poner en contacto el polipéptido de la etapa (a) y la composición de la etapa (b) en condiciones en las que la pectato liasa puede tratar el papel o pasta de papel o de madera.

La invención proporciona composiciones que comprenden un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención. La composición puede actuar como una ayuda digestiva.

También se desvelan productos para el cuidado oral que comprenden un polipéptido de la invención, o un polipéptido codificado por un ácido nucleico de la invención. El producto para el cuidado oral puede comprender una pasta dentífrica, una crema dental, un gel o un polvo para dientes, una sustancia odontica, un enjuague bucal, una formulación de aclarado para antes o después del cepillado, un chicle, una pastilla para chupar o un caramelo.

La invención proporciona ácidos nucleicos aislados o recombinantes que tienen una secuencia que comprende una modificación de la secuencia de la SEC ID N°: 131, en la que la modificación de la SEC ID N°: 131 comprende uno más de los siguientes cambios: los nucleótidos en los restos 352 a 354 son CAT o CAC, los nucleótidos en los restos 544 a 546 son GTG, GTT, GTC, o GTA, los nucleótidos en los restos 568 a 570 son TTG, TTA, CTT, CTC, CTA, o CTG, los nucleótidos en los restos 589 a 591 son GGT, GGC, GGA, o GGG, los nucleótidos en los restos 622 a 624 son AAG o AAA, los nucleótidos en los restos 655 a 657 son ATG, los nucleótidos en los restos 667 a 669 son GAG o GAA, los nucleótidos en los restos 763 a 765 son CGG, CGT, CGC, CGA, AGA, AGG, los nucleótidos en los restos 787 a 789 son AAG o AAA, los nucleótidos en los restos 823 a 825 son TAT o TAC, los nucleótidos en los restos 925 a 927 son TGG, o los nucleótidos en los restos 934 a 936 son GTT, GTG, GTC, o GTA. En un aspecto, el ácido nucleico codifica un polipéptido que tiene una actividad de pectato liasa, que puede ser termotolerante o termoestable.

La divulgación proporciona ácidos nucleicos aislados o recombinantes que tienen una secuencia que comprende la modificación de la secuencia de un ácido nucleico de la divulgación (por ejemplo, SEC ID N°: 1, SEC ID N°: 3, SEC ID N°: 5, SEC ID N°: 7, SEC ID N°: 9, SEC ID N°: 11, SEC ID N°: 13, SEC ID N°: 15, SEC ID N°: 17, SEC ID N°: 19, SEC ID N°: 21, SEC ID N°: 23, SEC ID N°: 25, SEC ID N°: 27, SEC ID N°: 29, SEC ID N°: 31, SEC ID N°: 33, SEC ID N°: 35, SEC ID N°: 37, SEC ID N°: 39, SEC ID N°: 41, SEC ID N°: 43, SEC ID N°: 45, SEC ID N°: 47, SEC ID N°: 49, SEC ID N°: 51, SEC ID N°: 53, SEC ID N°: 55, SEC ID N°: 57, SEC ID N°: 59, SEC ID N°: 61, SEC ID N°: 63, SEC ID N°: 65, SEC ID N°: 67, SEC ID N°: 69, SEC ID N°: 71, SEC ID N°: 73, SEC ID N°: 75, SEC ID N°: 77, SEC ID N°: 79, SEC ID N°: 81, SEC ID N°: 83, SEC ID N°: 85, SEC ID N°: 87, SEC ID N°: 89, SEC ID N°: 91, SEC ID N°: 93, SEC ID N°: 95, SEC ID N°: 97, SEC ID N°: 99, SEC ID N°: 101, SEC ID N°: 103 SEC ID N°: 105 SEC ID N°: 107 SEC ID N°: 109, SEC ID N°: 111, SEC ID N°: 113, SEC ID N°: 115, SEC ID N°: 117, SEC ID N°: 119, SEC ID N°: 121, SEC ID N°: 123, SEC ID N°: 125, SEC ID N°: 127, SEC ID N°: 129), en los que la modificación de la secuencia comprende uno o más de los siguientes cambios: los nucleótidos en la posición equivalente de los restos 352 a 354 de la SEC ID N°: 131 se cambian a CAT o CAC, los nucleótidos en la posición equivalente de los restos 544 a 546 de la SEC ID N°: 131 se cambian a GTG, GTT, GTC, o GTA, los nucleótidos en la posición equivalente de los restos 568 a 570 de la SEC ID N°: 131 se cambian a TTG, TTA, CTT, CTC, CTA, o CTG, los nucleótidos en la posición equivalente de los restos 589 a 591 de la SEC ID N°: 131 se cambian a GGT, GGC, GGA, o GGG, los nucleótidos en la posición equivalente de los restos 622 a 624 de la SEC ID N°: 131 se cambian a AAG o AAA, los nucleótidos en la posición equivalente de los restos 667 a 669 de la SEC ID N°: 131 son GAG o GAA, los nucleótidos en la posición equivalente de los restos 763 a 765 de la SEC ID N°: 131 se cambian a CGG, CGT, CGC, CGA, AGA, AGG, los nucleótidos en la posición equivalente de los restos 787 a 789 de la SEC ID N°: 131 se cambian a AAG o AAA, los nucleótidos en la posición equivalente de los restos 823 a 825 de la SEC ID N°: 131 se cambian a TAT o TAC, los nucleótidos en la posición equivalente de los restos 925 a 927 de la SEC ID N°: 131 se cambian a TGG, o los nucleótidos en la posición equivalente de los restos 934 a 936 de la SEC ID N°: 131 se cambian a GTT, GTG, GTC, o GTA. En un aspecto, el ácido nucleico codifica un polipéptido que tiene una actividad de pectato liasa, que puede ser termotolerante o termoestable.

La invención proporciona polipéptidos aislados o recombinantes que tienen una secuencia que comprende una modificación de la secuencia de la SEC ID N°: 132, en los que la modificación de la SEC ID N°: 132 comprende una o más de las siguientes mutaciones: la alanina en la posición del aminoácido 118 es histidina, la alanina en la posición del aminoácido 182 es valina, la treonina en la posición del aminoácido 190 es leucina, la alanina en la posición del aminoácido 197 es glicina, la serina en la posición del aminoácido 208 es lisina, la treonina en la posición del aminoácido 219 es metionina, la treonina en la posición del aminoácido 223 es ácido glutámico, la serina en la posición del aminoácido 255 es arginina, la serina en la posición del aminoácido 263 es lisina, la asparagina en la posición del aminoácido 275 es tirosina, la tirosina en la posición del aminoácido 309 es triptófano, o, la serina en la posición del aminoácido 312 es valina. En un aspecto, el polipéptido tiene una actividad de pectato liasa, que puede ser termotolerante o termoestable.

La divulgación proporciona polipéptidos aislados o recombinantes que tiene una secuencia que comprende una modificación de la secuencia de un polipéptido de la divulgación (por ejemplo, SEC ID N°: 2, SEC ID N°: 4, SEC ID N°: 6, SEC ID N°: 8, SEC ID N°: 10, SEC ID N°: 12, SEC ID N°: 14, SEC ID N°: 16, SEC ID N°: 18, SEC ID N°: 20, SEC ID N°: 22, SEC ID N°: 24, SEC ID N°: 26, SEC ID N°: 28, SEC ID N°: 30, SEC ID N°: 32, SEC ID N°: 34, SEC ID N°: 36, SEC ID N°: 38, SEC ID N°: 40, SEC ID N°: 42, SEC ID N°: 44, SEC ID N°: 46, SEC ID N°: 48, SEC ID N°: 50, SEC ID N°: 52, SEC ID N°: 54, SEC ID N°: 56, SEC ID N°: 58, SEC ID N°: 60, SEC ID N°: 62, SEC ID N°: 64, SEC ID N°: 66, SEC ID N°: 68, SEC ID N°: 70, SEC ID N°: 72, SEC ID N°: 74, SEC ID N°: 76, SEC ID N°: 78, SEC ID N°: 80, SEC ID N°: 82, SEC ID N°: 84, SEC ID N°: 86, SEC ID N°: 88, SEC ID N°: 90, SEC ID N°: 92, SEC ID N°: 94, SEC ID N°: 96, SEC ID N°: 98, SEC ID N°: 100, SEC ID N°: 102, SEC ID N°: 104, SEC ID N°: 106, SEC ID N°: 108, SEC ID N°: 110, SEC ID N°: 112, SEC ID N°: 114, SEC ID N°: 116, SEC ID N°: 118, SEC ID N°: 120, SEC ID N°: 122, SEC ID N°: 124, SEC ID N°: 126, SEC ID N°: 128, SEC ID N°: 130), la modificación de la secuencia comprende uno o más de los siguientes cambios: el aminoácido en la posición equivalente de la alanina en el resto 118 de la SEC ID N°: 132 se cambia por una histidina, el aminoácido en la posición equivalente de la alanina en el resto 182 de la SEC ID N°: 132 se cambia por una valina, el aminoácido en la posición equivalente de la treonina en el resto 190 de la SEC ID N°: 132 se cambia por una leucina, el aminoácido en la posición equivalente de la alanina en el resto 197 de la SEC ID N°: 132 se cambia por una glicina, el aminoácido en la posición equivalente de la serina en el resto 208 de la SEC ID N°: 132 se cambia por una lisina, el aminoácido en la posición equivalente de la treonina en el resto 219 de la SEC ID N°: 132 se cambia por una metionina, el aminoácido en la posición equivalente de la treonina en el resto 223 de la SEC ID N°: 132 se cambia por una ácido glutámico, el aminoácido en la posición equivalente de la serina en el resto 255 de la SEC ID N°: 132 se cambia por una arginina, el aminoácido en la posición equivalente de la serina en el resto 263 de la SEC ID N°: 132 se cambia por una lisina, el aminoácido en la posición equivalente de la asparagina en el resto 275 de la SEC ID N°: 132 se cambia por una tirosina, el aminoácido en la posición equivalente de la tirosina en el resto 309 de la SEC ID N°: 132 se cambia por un triptófano, o, el aminoácido en la posición equivalente de la serina en el resto 312 de la SEC ID N°: 132 se cambia por una valina. En un aspecto, el polipéptido tiene una actividad de pectato liasa, que puede ser termotolerante o termoestable.

Algunos métodos para generar un ácido nucleico que codifica una pectato-liasa modificada comprenden la preparación de una o más modificaciones de la secuencia en un ácido nucleico que codifica pectato-liasa, en los que los cambios en el ácido nucleico que codifica una pectato-liasa son equivalentes a uno o más de los siguientes: cambio de nucleótidos en la posición equivalente de los restos 352 a 354 de la SEC ID N°: 131 por CAT o CAC, cambio de nucleótidos en la posición equivalente de los restos 544 a 546 de la SEC ID N°: 131 por GTG, GTT, GTC, o GTA, cambio de nucleótidos en la posición equivalente de los restos 568 a 570 de la SEC ID N°: 131 por TTG, TTA, CTT, CTC, CTA, o CTG, cambio de nucleótidos en la posición equivalente de los restos 589 a 591 de la SEC ID N°: 131 por GGT, GGC, GGA, o GGG, cambio de nucleótidos en la posición equivalente de los restos 622 a 624 de la SEC ID N°: 131 por AAG o AAA, cambio de nucleótidos en la posición equivalente de los restos 655 a 657 de la SEC ID N°: 131 por ATG, cambio de nucleótidos en la posición equivalente de los restos 667 a 669 de la SEC ID N°: 131 por GAG o GAA, los nucleótidos en la posición equivalente de los restos 763 a 765 de la SEC ID N°: 131 por CGG, CGT, CGC, CGA, AGA, AGG, cambio de nucleótidos en la posición equivalente de los restos 787 a 789 de la SEC ID N°: 131 por AAG o AAA, cambio de nucleótidos en la posición equivalente de los restos 823 a 825 de la SEC ID N°: 131 por TAT o TAC, cambio de nucleótidos en la posición equivalente de los restos 925 a 927 de la SEC ID N°: 131 por TGG, o cambio de nucleótidos en la posición equivalente de los restos 934 a 936 de la SEC ID N°: 131 por GTT, GTG, GTC, o GTA. En un aspecto, la actividad de pectato liasa modificada tiene una actividad termotolerante o termoestable. El ácido nucleico que codifica una pectato-liasa puede comprender un ácido nucleico que tiene una secuencia como se establece en las SEC ID N°: 1, SEC ID N°: 3, SEC ID N°: 5, SEC ID N°: 7, SEC ID N°: 9, SEC ID N°: 11, SEC ID N°: 13, SEC ID N°: 15, SEC ID N°: 17, SEC ID N°: 19, SEC ID N°: 21, SEC ID N°: 23, SEC ID N°: 25, SEC ID N°: 27, SEC ID N°: 29, SEC ID N°: 31, SEC ID N°: 33, SEC ID N°: 35, SEC ID N°: 37, SEC ID N°: 39, SEC ID N°: 41, SEC ID N°: 43, SEC ID N°: 45, SEC ID N°: 47, SEC ID N°: 49, SEC ID N°: 51, SEC ID N°: 53, SEC ID N°: 55, SEC ID N°: 57, SEC ID N°: 59, SEC ID N°: 61, SEC ID N°: 63, SEC ID N°: 65, SEC ID N°: 67, SEC ID N°: 69, SEC ID N°: 71, SEC ID N°: 73, SEC ID N°: 75, SEC ID N°: 77, SEC ID N°: 79, SEC ID N°: 81, SEC ID N°: 83, SEC ID N°: 85, SEC ID N°: 87, SEC ID N°: 89, SEC ID N°: 91, SEC ID N°: 93, SEC ID N°: 95, SEC ID N°: 97, SEC ID N°: 99, SEC ID N°: 101, SEC ID N°: 103 SEC ID N°: 105, SEC ID N°: 107, SEC ID N°: 109, SEC ID N°: 111, SEC ID N°: 113, SEC ID N°: 115, SEC ID N°: 117, SEC ID N°: 119, SEC ID N°: 121, SEC ID N°: 123, SEC ID N°: 125, SEC ID N°: 127, SEC ID N°: 129, SEC ID N°: 131 o SEC ID N°: 133.

Algunos métodos para generar una pectato liasa modificada comprenden preparar una o más modificaciones de secuencia en una pectato liasa, en los que los cambios en la pectato liasa son equivalentes a uno o más de los siguientes cambio: el aminoácido en la posición equivalente de la alanina en el resto 118 de la SEC ID N°: 132 se cambia por una histidina, el aminoácido en la posición equivalente de la alanina en el resto 182 de la SEC ID N°: 132 se cambia por una valina, el aminoácido en la posición equivalente de la treonina en el resto 190 de la SEC ID N°: 132 se cambia por una leucina, el aminoácido en la posición equivalente de la alanina en el resto 197 de la SEC ID N°: 132 se cambia por una glicina, el aminoácido en la posición equivalente de la serina en el resto 208 de la SEC ID N°: 132 se cambia por una lisina, el aminoácido en la posición equivalente de la treonina en el resto 219 de la SEC ID N°: 132 se cambia por una metionina, el aminoácido en la posición equivalente de la treonina en el resto 223 de la SEC ID N°: 132 se cambia por una ácido glutámico, el aminoácido en la posición equivalente de la serina en el resto 255 de la SEC ID N°: 132 se cambia por una arginina, el aminoácido en la posición equivalente de la serina en el resto 263 de la SEC ID N°: 132 se cambia por una lisina, el aminoácido en la posición equivalente de la asparagina en el resto 275 de la SEC ID N°: 132 se cambia por una tirosina, el aminoácido en la posición equivalente de la tirosina en el resto 309 de la SEC ID N°: 132 se cambia por un triptófano, o, el aminoácido en la posición equivalente de la serina en el resto 312 de la SEC ID N°: 132 se cambia por una valina. La pectato liasa puede comprender una secuencia de la divulgación (por ejemplo, SEC ID N°: 2, SEC ID N°: 4, SEC ID N°: 6, SEC ID N°: 8, SEC ID N°: 10, SEC ID N°: 12, SEC ID N°: 14, SEC ID N°: 16, SEC ID N°: 18, SEC ID N°: 20, SEC ID N°: 22, SEC ID N°: 24, SEC ID N°: 26, SEC ID N°: 28, SEC ID N°: 30, SEC ID N°: 32, SEC ID N°: 34, SEC ID N°: 36, SEC ID N°: 38, SEC ID N°: 40, SEC ID N°: 42, SEC ID N°: 44, SEC ID N°: 46, SEC ID N°: 48, SEC ID N°: 50, SEC ID N°: 52, SEC ID N°: 54, SEC ID N°: 56, SEC ID N°: 58, SEC ID N°: 60, SEC ID N°: 62, SEC ID N°: 64, SEC ID N°: 66, SEC ID N°: 68, SEC ID N°: 70, SEC ID N°: 72, SEC ID N°: 74, SEC ID N°: 76, SEC ID N°: 78, SEC ID N°: 80, SEC ID N°: 82, SEC ID N°: 84, SEC ID N°: 86, SEC ID N°: 88, SEC ID N°: 90, SEC ID N°: 92, SEC ID N°: 94, SEC ID N°: 96, SEC ID N°: 98, SEC ID N°: 100, SEC ID N°: 102, SEC ID N°: 104, SEC ID N°: 106, SEC ID N°: 108, SEC ID N°: 110, SEC ID N°: 112, SEC ID N°: 114, SEC ID N°: 116, SEC ID N°: 118, SEC ID N°: 120, SEC ID N°: 122, SEC ID N°: 124, SEC ID N°: 126, SEC ID N°: 128 o SEC ID N°: 130). La actividad de pectato liasa modificada puede tener una actividad termotolerante o termoestable.

La invención proporciona formulaciones que comprenden al menos una enzima de la invención que comprenden dosificaciones en el intervalo entre aproximadamente 1 gramo por ton y 100 o más gramos por ton (por ton de material tratado, por ejemplo, por ton de tejido, tela o similar), entre aproximadamente 10 gramos por ton y 90 gramos por ton, entre aproximadamente 20 gramos por ton y 80 gram por ton, entre aproximadamente 30 gramos por ton y 70 gramos por ton, entre aproximadamente 40 gramos por ton y 50 gramos por ton. Por ejemplo, algunas formulaciones a modo de ejemplo comprenden de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, etc. a 100, 200, 300, 400, 500, etc. o más gramos por ton.

Como alternativa, la invención proporciona formulaciones que comprenden al menos una enzima de la invención que comprende dosificaciones en el intervalo de entre aproximadamente 1 µg por gramo y 100 o más µg por gramo (por gramo de materia tratado, por ejemplo, por gramo de tejido, tela o similares), entre aproximadamente 10 µg por gramo y 90 µg por gramo, entre aproximadamente 20 µg por gramo y 80 µg por gramo, entre aproximadamente 30 µg por gramo y 70 µg por gramo, entre aproximadamente 40 µg por gramo y 50 µg por gramo. Por ejemplo, algunas formulaciones a modo de ejemplo comprenden de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, etc. a 100, 200, 300, 400, 500, etc. o más µg por gramo (por ejemplo, por gramo de tejido, tela o similares).

Como alternativa, la invención proporciona formulaciones que comprenden al menos una enzima de la invención que comprende dosificaciones en el intervalo de entre aproximadamente 0,5 mg por 0,45 kg (1 libra) y 50 o más mg por 0,45 kg (1 libra) (por 0,45 kg de material tratado, por ejemplo, por 0,45 kg de tejido, tela o similares), entre aproximadamente 1 mg por 0,45 kg y 45 mg por 0,45 kg, entre aproximadamente 5 mg por 0,45 kg y 40 mg por 0,45 kg, entre aproximadamente 10 mg por 0,45 kg y 35 mg por 0,45 kg, entre aproximadamente 15 mg por 0,45 kg y 30 mg por 0,45 kg. Por ejemplo, algunas formulaciones a modo de ejemplo comprenden de aproximadamente 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, etc. a 50 o más mg por 0,45 kg (por ejemplo, por 0,45 kg de tejido, tela o similares).

La invención proporciona formulaciones que comprenden al menos una enzima de la invención que comprenden dosificaciones que comprenden una potencia enzimática entre aproximadamente 100 y 40.000 unidades/ml, de 200 a 30.000 unidades/ml, de 500 a 30.000 unidades/ml, de 1000 a 20.000 unidades/ml, de 1000 a 25.000 unidades/ml, de 1000 a 15.000 unidades/ml, de 1000 a 10000 unidades/ml, de 1000 a 5000 unidades/ml, entre aproximadamente 2000 y 20000 unidades/ml, entre aproximadamente 2000 y 15000 unidades/ml, entre aproximadamente 2000 y 10000 unidades/ml, o entre aproximadamente 2000 y 4000 unidades/ml, o, entre aproximadamente 200 y 25.000 unidades/ml, de 200 a 20.000 unidades/ml, de 200 a 15000 unidades/ml, de 200 a 10.000 unidades/ml, entre aproximadamente 400 y 8000 unidades/ml, entre aproximadamente 600 y 6000 unidades/ml, entre aproximadamente 800 y 4000 unidades/ml, o entre aproximadamente 1000 y 2000 unidades/ml, o, en las que la dosificación comprende una potencia enzimática de aproximadamente 1000 u/ml o, en las que en la dosificación comprende una potencia enzimática de aproximadamente 3000 unidades/ml.

La formulación puede comprender una enzima liofilizada (por ejemplo, una enzima de la invención), o, la formulación puede comprender una formulación basada en agua que comprende una enzima de la invención. La formulación puede comprender una enzima liofilizada resuspendida en agua, o puede comprender adicionalmente un glicerol, sacarosa, cloruro sódico, dextrina, propilenglicol, sorbitol, sulfato sódico o TRIS, o un equivalente. En un aspecto, una formulación de la invención comprende adicionalmente un tampón, por ejemplo, un tampón que comprende un pH 7, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 7, glicerol al 35 %, 300 ppm de proxel; pH 7, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 7, cloruro sódico al 10 %, glicerol al 25 %, 300 ppm de proxel; pH 5,5, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 5,5, glicerol al 35 %, 300 ppm de proxel; pH 5,5, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; o, tampón acetato 20 mM, pH 5,5, glicerol al 35 %; MOPS 20 mM, pH 7 o MOPS 25 mM, NaCl 50 mM, pH 7,5; pH 5,0, TRIS 40 mM; pH 7,0, TRIS 40 mM; pH 8,0, TRIS 40 mM; pH 7,5, glicerol al 50 %; pH 7,5, NaCl al 20 %; pH 7,5, propilenglicol al 30 %; pH 7,5, sulfato sódico 100 mM; pH 5,5, glicerol al 35 %; o, cualquier combinación de los mismos, o equivalentes de los mismos.

La invención proporciona procesos de biolimpieza que comprenden las siguientes etapas: (a) proporcionar una pectato liasa de la invención; (b) proporcionar un material que comprende pectina o ácido poligalacturónico; (c) poner en contacto la pectato liasa de (a) con el material de (b) en condiciones alcalinas, por ejemplo, a pH superior a 7,5, o, condiciones que comprenden entre aproximadamente pH 8 y pH 9 o superior, por ejemplo, pH 8,5, en tampón de bicarbonato buffer o equivalente. En un aspecto, el método también comprende un agente humectante no iónico, por ejemplo, a aproximadamente 1 g/l. En un aspecto, la proporción de pectato liasa se encuentra en un baño enzimático entre aproximadamente 10:1 y 50:1 l de pectato liasa:kg de material. En un aspecto, la dosis de pectato liasa está entre aproximadamente 0,1 y 0,2 ml de un extracto concentrado por kg de material, o equivalente. Como alternativa, la dosis de pectato liasa está entre aproximadamente 0,1 ml y 1 ml de un extracto concentrado por kg de material, o equivalente. En un aspecto, el intervalo de temperatura está entre aproximadamente 50 °C y 70 °C. En un aspecto, el tiempo de tratamiento es de aproximadamente 20 min. En un aspecto de los procesos de biolimpieza de la invención, el material comprende un tejido o una tela. En un aspecto, la dosis de pectato liasa es de aproximadamente 0,137 ml de un extracto concentrado por kg de material, o equivalente. En un aspecto, la etapa de poner en contacto comprende adicionalmente el uso de un agente quelante, en la que el agente quelante se excluye del baño enzimático y se añade después de aproximadamente 20 minutos de tratamiento enzimático y se mantiene durante aproximadamente 10 minutos antes del baño de descarga.

Algunos aspectos adicionales de la invención se exponen en las reivindicaciones adjuntas.

Los detalles de una o más realizaciones de la invención se exponen en las figuras adjuntas en la descripción que sigue a continuación. Otras características, objetos, y ventajas de la invención serán evidentes a partir de la descripción y las figuras, y a partir de las reivindicaciones.

Descripción de las figuras

La patente o archivo de solicitud contiene al menos una figura realizada en color. Las copias de esta patente o publicación de solicitud de patente con figura o figuras en color las proporcionará la Oficina previa solicitud y pago de la tasa necesaria.

La Figura 1 es un diagrama de bloque de un sistema informático.

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

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

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

La Figura 5 es un resumen gráfico de la especificidad relativa del sustrato pH, actividad enzimática, descripción por caracterización y caracterización de sustrato de algunas pectato liasas a modo de ejemplo de la invención.

La Figura 6 es un resumen de polipéptido pectato liasas de la invención, caracterizadas como "mutantes positivos", como se analiza con detalle a continuación.

La Figura 7 es una tabla que resume temperaturas de fusión a modo de ejemplo y actividades específicas (SA) de enzimas a modo de ejemplo de la invención a diversas temperaturas.

La Figura 8 resume datos de ensayos de actividad de enzimas termotolerantes a modo de ejemplo de la invención, como se describe en el Ejemplo 4, que sigue a continuación.

Los símbolos de referencia similares en las diversas figuras indican elementos similares.

Descripción detallada

La invención proporciona polipéptidos que tienen una actividad de pectato liasa, polinucleótidos que codifican los polipéptidos, y métodos para preparar y usar estos polinucleótidos y polipéptidos. En un aspecto, las pectato liasas

de la invención se usan para catalizar la beta-eliminación (trans-eliminación) y/o hidrólisis de pectina y/o ácido poligalacturónico (pectato) u otros componentes de la pared vegetal, por ejemplo, homogalacturonano o ramnogalacturonano, incluyendo ácido alfa-D-galacturónico unido en la posición 1,4. Las pectato liasas de la invención también se pueden usar para la hidrólisis de paredes de célula vegetal, por ejemplo, en el tratamiento de fibras naturales que comprenden pectina, por ejemplo, fibras de algodón.

El uso de las pectato liasas de la invención para hidrolizar pectina de la pared celular primaria puede eliminar la necesidad de agentes cáusticos y temperaturas elevadas en la limpieza de fibra de algodón. El uso de las pectato liasas de la invención también puede reducir de forma significativa la cantidad de agua usada para aclarar fibras tratadas, por ejemplo, tejido de algodón de punto o tejido, después de la limpieza química. El uso de las pectato liasas de la invención también puede reducir pérdidas de materia prima en la limpieza química. En un aspecto, una pectato liasa de la invención, por ejemplo, una pectato liasa alcalina y/o termoestable, se usa para biolimpieza. Por lo tanto, la invención proporciona procesos en los que se procesan tejidos de algodón sin apresto para solubiliza y extraer material no celulósico indeseado en tejidos y otros materiales celulósicos usando una enzima de la invención. Los procesos de la invención se pueden usar para solubilizar y/o extraer materiales encontrados de forma natural en el algodón y/o para retirar impurezas aplicadas, tales como lubricantes de maquinaria.

Las Figuras 5 y 7 son resúmenes gráficos, entre otros, de la especificidad relativa del sustrato, valor relativo de la especificidad del sustrato, temperatura de la actividad de caracterización, pH de la actividad de caracterización, actividad enzimática, descripción de la caracterización y caracterización del sustrato de las pectato liasas de la divulgación.

Las preparaciones de pectato liasa de la invención (incluyendo aquéllas para el tratamiento o procesamiento de comidas o alimentos, tratamiento de fibras y textiles, tratamientos de residuos, tratamientos de vegetales, y similares) pueden comprender adicionalmente una o más enzimas, por ejemplo, proteasas, celulasas (endo-beta-1,4-glucanasas), beta-glucanasas (endo-beta-1,3(4)-glucanasas), lipasas, cutinasas, peroxidasas, lacasas, amilasas, pectato liasas, pectinasas, reductasas, oxidasas, fenoloxidasas, ligninasas, pululaninas, arabinanasas, hemicelulasas, mananasas, xiloglucanasas, xilanasas, pectin acetil esterasas, ramnogalacturonano acetil esterasas, poligalacturonasas, ramnogalacturonasas, galactanasas, pectin liasas, pectin metilesterasas, celobiohidrolasas, transglutaminasas; o de las de las mismas.

Definiciones

La expresión "pectato liasa" incluye todos los polipéptidos que tienen una actividad de pectato liasa, o pectinasa, incluyendo la beta-eliminación (trans-eliminación) y/o hidrólisis de pectina y/o ácido poligalacturónico (pectato) u otros componentes de la pared vegetal, por ejemplo, homogalacturonano o ramnogalacturonano, incluyendo ácido alfa-D-galacturónico unido en la posición 1,4. En un aspecto, la actividad de pectato liasa incluye catálisis de la escisión de uniones glicosídicas de sustancias pécticas, por ejemplo, catálisis de la beta-eliminación (trans-eliminación) y/o hidrólisis de paredes de células vegetales (por ejemplo, la descomposición o disolución de paredes celulares que comprenden pectina, por ejemplo, paredes de células vegetales). En un aspecto, la actividad de pectato liasa incluye la catálisis de la beta-eliminación (trans-eliminación) y/o hidrólisis de ácido galacturónico esterificado con metilo, incluyendo ácido poligalacturónico esterificado con metilo parcial o completamente. En un aspecto, la actividad de pectato liasa es principalmente de actuación endo, por ejemplo, cortando el polímero (por ejemplo, ácido poligalacturónico) en sitios aleatorios dentro de una cadena para dar una mezcla de oligómeros, o la actividad de pectato liasa puede ser de actuación exo, atacando desde un extremo del polímero y produciendo monómeros o dímeros, o, una combinación de los mismos. En un aspecto, la actividad de pectato liasa comprende la catálisis de la escisión aleatoria de enlaces alfa-1,4-glicosídicos en el ácido péctico (ácido poligalacturónico) mediante trans-eliminación. En un aspecto, la actividad de la pectato liasa incluye polipéptidos que tienen una actividad igual o similar a la de la pectato liasa (EC 4.2.2.2), poli(1,4-alfa-D-galacturónido) liasa, poligalacturonato liasa (EC 4.2.2.2), pectina liasa (EC 4.2.2.10), poligalacturonasa (EC 3.2.1.15), exo-poligalacturonasa (EC 3.2.1.67), exo-poligalacturonato liasa (EC 4.2.2.9) y/o exo-poli-alfa-galacturonosidasa (EC 3.2.1.82).

Un polipéptido se puede someter a ensayo de forma rutinaria para la actividad de pectato liasa (por ejemplo, sometido a ensayo para observar si la proteína está dentro del alcance de la invención) mediante cualquier método, por ejemplo, un ensayo de PGA para pectato liasas. En este ensayo, la actividad de pectato liasa se mide a la temperatura y pH deseados usando ácido poligalacturónico al 0,2 % (Sigma, P3850) en tampón de Tris HCl 25 mM - Glicina NaOH 25 mM. Una unidad de actividad enzimática se define como la cantidad de proteína que produce 1 μ mol de oligogalacturónidos insaturados por minuto equivalente a 1 μ mol de digalacturónido insaturado, usando el valor del coeficiente de extinción molecular de 4600 $M^{-1}cm^{-1}$ a 235 nm para el dímero. La proteína se puede determinar para proteína purificada homogénea mediante la medida de la absorbancia a 280 nm, usando el valor del coeficiente de extinción específico para cada proteína basándose en la secuencia.

El término "anticuerpo" incluye un péptido o polipéptido derivados de, modelados después de o básicamente codificado por un gen de inmunoglobulina o genes de inmunoglobulina, o fragmentos de los mismos, capaces de unirse de forma estética y cara a un antígeno o epítipo, véase, por ejemplo Fundamental Immunology, Tercera Edición, W.E. Paul, ed., Raven Press, N.Y. (1993); Wilson (1994) J. Immunol. Methods 175: 267-273; Yarmush

(1992) J. Biochem. Biophys. Methods 25: 85-97. El término anticuerpo incluye partes de unión a antígeno, es decir, "sitios de unión a antígeno", (por ejemplo, fragmentos, subsecuencias, regiones que determinan la complementariedad (CDR)) que mantienen la capacidad de unirse al antígeno, incluyendo (i) un fragmento de Fab, un fragmento monovalente que consiste en los dominios VL, VH, CL y CH1; (ii) un fragmento de F(ab')₂, un fragmento bivalente que comprende dos fragmentos de Fab unidos por un puente disulfuro en la región bisagra; (iii) un fragmento de Fd que consiste en los dominios VH y CH1; (iv) un fragmento de Fv que consiste en los dominios VL y VH de una sola rama de un anticuerpo, (v) un fragmento de dAb (Ward *et al.*, (1989) Nature 341: 544-546), que consiste en un dominio VH; y (vi) una región aislada que determina la complementariedad (CDR). En el término "anticuerpo" también se incluyen por referencia anticuerpos de una sola cadena.

Los términos "matriz" o "micromatriz" o "biochip" o "chip" como se usan en el presente documento se refieren a una pluralidad de elementos diana, comprendiendo cada elemento diana una cantidad definida de uno o más polipéptidos (incluyendo anticuerpos) o ácidos nucleicos inmovilizados en una zona definida de una superficie del sustrato, como se analiza con más detalle a continuación.

Como se usa en el presente documento, los términos "ordenador", "programa informático" y "procesador" se usan en sus contextos generales más amplios e incorporan todos estos dispositivos, como se describe con detalle a continuación. Una "secuencia de codificación de" o una "secuencia que codifica" un polipéptido o proteína en particular, es una secuencia de ácidos nucleicos que se transcribe y traduce en un polipéptido o proteína cuando se coloca bajo el control de las secuencias reguladoras apropiadas.

La expresión "casete de expresión", como se usa en el presente documento, se refiere a una secuencia de nucleótidos que es capaz de realizar la expresión de un gen estructural (es decir, una secuencia que codifica proteínas, tal como una pectato liasa de la invención) en un hospedador compatible con tales secuencias. Algunos casetes de expresión incluyen al menos un promotor unido de forma operativa con la secuencia que codifica polipéptidos; y, opcionalmente, con otras secuencias, por ejemplo, señales de terminación de la transcripción. También se pueden usar algunos factores adicionales necesarios o útiles en la realización de la expresión, por ejemplo, potenciador es. Por lo tanto, los casetes de expresión también incluyen plásmidos, vectores de expresión, virus recombinantes, cualquier forma de vector de "ADN desnudo" recombinante, y similares.

"Unido de forma operativa", como se usa en el presente documento, se refiere a una relación funcional entre dos o más segmentos de ácido nucleico (por ejemplo, ADN). Por lo general, se refiere a la relación funcional de la secuencia reguladora de la transcripción con respecto a una secuencia transcrita. Por ejemplo, un promotor se une de forma operativa a una secuencia de codificación, tal como un ácido nucleico de la invención, si estimula o modula la transcripción de la secuencia de codificación en célula hospedadora una apropiada u otro sistema de expresión. Por lo general, las secuencias reguladoras de la transcripción promotoras que se unen de forma operativa a una secuencia transcrita son físicamente contiguas a la secuencia transcrita, es decir, son de actuación *cis*. Sin embargo, no es necesario que algunas secuencias reguladoras de la transcripción, tales como potenciadores, estén físicamente contiguas o estén localizadas en las proximidades de las secuencias de codificación cuya transcripción potencian.

Un "vector" comprende un ácido nucleico que puede infectar, transfectar, transducir de forma transitoria o permanente una célula. Se observará que un vector puede ser un ácido nucleico desnudo, o un ácido nucleico que forma complejos con proteínas o lípidos. El vector comprende opcionalmente ácidos nucleicos virales o bacterianos y/o proteínas, y/o membranas (por ejemplo, una membrana celular, una envoltura lipídica viral, etc.). Algunos vectores incluyen, pero no se limitan a replicones (por ejemplo, replicones de ARN, bacteriófagos) a los que se pueden unir fragmentos de ADN y se llegan a replicar. Por lo tanto, algunos vectores incluyen, pero no se limitan a ARN, ADN o ARN circular o lineal de autorreplicación autónoma (por ejemplo, plásmidos, virus, y similares, véase, por ejemplo, la Patente de Estados Unidos N° 5.217.879), e incluyen los plásmidos tanto de expresión como de no expresión. Cuando se describe un microorganismo recombinante o cultivo celular como hospedador de un "vector de expresión", éste incluye ADN extra-cromosómico circular o lineal como ADN que se ha incorporado en el cromosoma o cromosomas hospedadores. Cuando una célula hospedadora está manteniendo un vector, el vector se puede replicar de forma estable por las células durante la mitosis como una estructura autónoma, o se incorpora dentro del genoma del hospedador.

Como se usa en el presente documento, el término "promotor" incluye todas las secuencias capaces de dirigir la transcripción de una secuencia de codificación en una célula, por ejemplo, una célula vegetal. Por lo tanto, los promotores usados en los constructos de la invención incluyen elementos de control de la transcripción de actuación *cis* y secuencias reguladoras que están implicadas en la regulación o la modulación del momento y/o tasa de transcripción de un gen. Por ejemplo, un promotor puede ser un elemento de control de la transcripción de actuación *cis*, incluyendo un potenciador, un promotor, un terminador de la transcripción, un origen de replicación, una secuencia de integración cromosómica, regiones sin traducir en las posiciones 5' y 3', o una secuencia intrónica, que están implicadas en la regulación de la transcripción. Estas secuencias de actuación *cis* por lo general interactúan con proteínas podrás biomoléculas para realizar (activar/desactivar, regular, modular, etc.) la transcripción. Los promotores "constitutivos" son los que dirigen la expresión de forma continua en la mayoría de las condiciones ambientales y estados de desarrollo o diferenciación celular. Algunos promotores "inducibles" o "regulables" dirigen

la expresión del ácido nucleico de la invención bajo la influencia de condiciones ambientales o condiciones de desarrollo. Algunos ejemplos de condiciones ambientales que pueden influir en la transcripción mediante promotores inducibles incluyen condiciones anaerobias, temperatura elevada, sequía, o la presencia de luz.

Los promotores "específicos de tejido" son elementos para el control de la transcripción que solamente son activos en células o tejidos u órganos en particular, por ejemplo, en plantas o animales. La regulación específica de tejido se puede conseguir mediante ciertos factores intrínsecos que aseguran que se expresen los genes que codifican proteínas específicas para un tejido dado. Se sabe que tales factores existen en mamíferos y plantas de modo que permiten que se desarrollen tejidos específicos.

El término "planta" incluye plantas completas, partes de plantas (por ejemplo, hojas, tallos, flores, raíces, etc.), protoplastos de plantas, semillas y células vegetales y progenie de los mismos. La clase de plantas que se puede usar en el método de la invención es generalmente tan amplia como la clase de plantas superiores susceptibles de técnicas de transformación, incluyendo angiospermas (plantas monocotiledóneas y dicotiledóneas), así como gimnospermas. El término incluye plantas de una diversidad de niveles de ploidía, incluyendo poliploide, diploide, haploide y estados hemigigoto. Como se usa en el presente documento, la expresión "planta transgénica" incluye plantas o células de plantas en las que se ha insertado una secuencia de ácidos nucleicos heteróloga, por ejemplo, los ácidos nucleicos y diversos constructos recombinantes (por ejemplo, casetes de expresión) de la invención.

Los "plásmidos" pueden estar disponibles en el mercado, a disposición del público sin restricciones, o se pueden construir a partir de plásmidos disponibles de acuerdo con procedimientos publicados. En la técnica se conocen algunos plásmidos equivalentes a los que se describen en el presente documento y serán evidentes para los expertos habituales en la materia.

El término "gen" incluye secuencia de ácidos nucleicos que comprende un segmento de ADN implicado en la producción de un producto de transcripción (por ejemplo, un mensaje), que a su vez se traduce para producir una cadena polipeptídica, o regula la transcripción, la reproducción o la estabilidad genética. Los genes pueden incluir regiones que preceden y siguen a la región de codificación, tales como el director y transporte, promotores y potenciadores, así como, en su caso, secuencias intermedias (intrones) entre segmentos de codificación individuales (exones).

Las expresiones "ácido nucleico" o "secuencia de ácidos nucleicos" incluyen oligonucleótido, nucleótido, polinucleótido, o un fragmento de cualquiera de los mismos, a ADN o ARN (por ejemplo, ARNm, ARNr, ARNt) de origen genómico o sintético que pueden ser monocatenario o bicatenario y pueden representar una hebra sentido o antisentido, para ácido nucleico peptídico (PNA), o para cualquier material similar al ADN o similar al ARN, de origen natural o sintético, que incluye, por ejemplo, ARNi, ribonucleoproteínas (por ejemplo, iRNP). Las expresiones incluyendo ácidos nucleicos, es decir, oligonucleótidos, que contienen análogos conocidos de nucleótidos naturales. Las expresiones también incluyen estructuras similares a los ácidos nucleicos con estructuras principales sintéticas, véase por ejemplo, Mata (1997) Toxicol. Appl. Pharmacol. 144: 189-197; Strauss-Soukup (1997) Biochemistry 36: 8692-8698; Samstag (1996) Antisense Nucleic Acid Drug Dev 6: 153-156.

"Aminoácido" o "secuencia de aminoácido" incluyen un oligopéptido, péptido, polipéptido, o secuencia de proteínas, o un fragmento, parte, o subunidad de cualquiera de los mismos, y la moléculas de origen natural o sintético. Los términos "polipéptido" y "proteína" incluyen aminoácidos unidos entre sí mediante enlaces peptídicos o enlaces peptídicos modificados, es decir, isoésteres peptídicos, y pueden contener aminoácidos modificados distintos de los aminoácidos codificados de 20 genes. El término "polipéptido" también incluye péptidos y fragmentos de polipéptido, motivos y similares. El término también incluye polipéptidos glicosilados. Los péptidos y polipéptidos de la invención también incluyen todas las formas "miméticas" y "peptidomiméticas", como se describe con más detalle a continuación.

El término "aislado" incluye un material retirado de su entorno original, por ejemplo, el entorno natural si es de origen natural. Por ejemplo, un polinucleótido o polipéptido de origen natural presente en un animal vivo no está aislado, pero el mismo polinucleótido o polipéptido, separado de algunos o todos los materiales coexistentes en el sistema natural, se aísla. Tales polinucleótidos podrían ser parte de un vector y/o tales polinucleótidos o polipéptidos podrían ser parte de una composición, y además ser aislados en un vector o composición tal que no forma parte de su entorno natural. Como se usa en el presente documento, un material o composición aislado también puede ser una composición "purificada", es decir, no necesita una pureza absoluta; en su lugar, se pretende como una definición relativa. Algunos ácidos nucleicos individuales obtenidos a partir de la biblioteca se pueden purificar de forma convencional hasta homogeneidad electroforética. En aspectos alternativos, la invención proporciona partidos nucleicos que se han purificado a partir de ADN genómico o a partir de otras secuencias en una biblioteca u otro entorno con al menos uno, dos, tres, cuatro, cinco o más órdenes de magnitud.

Como se usa en el presente documento, el término "recombinante" puede incluir ácidos nucleicos adyacentes a un ácido nucleico de la "estructura principal" que no es adyacente en su entorno natural. En un aspecto, algunos ácidos nucleicos representan un 5 % , es del número de insertos de ácido nucleico en una población de "moléculas de la estructura principal" de ácido nucleico. Las "moléculas de la estructura principal" de acuerdo con la invención

incluyen ácidos nucleicos tales como vectores de expresión, ácidos nucleicos de autorreplicación, virus, ácidos nucleicos de integración, y otros vectores o ácidos nucleicos usados para mantener o manipular un inserto de ácido nucleico de interés. En un aspecto, los ácidos nucleicos enriquecidos representan un 10 %, 15 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 %, 98 % o más del número de insertos de ácido nucleico en la población de moléculas de la estructura principal recombinantes. Polipéptidos o proteínas "recombinantes" se refiere a polipéptidos o proteínas producidos mediante técnicas de ADN recombinante; por ejemplo, producidos a partir de células transformadas mediante un constructor de ADN exógeno que codifica el polipéptido o proteína deseados. Los polipéptidos o proteínas "sintéticos" son los preparados mediante síntesis química, como se describe con más detalle a continuación.

Una secuencia promotora se puede "unir de forma operativa a" una secuencia de codificación cuando la ARN polimerasa que inicia la transcripción en el promotor transcribe la secuencia de codificación en el ARNm, como se analiza con más detalle a continuación.

"Oligonucleótido" incluye cualquiera de un polidesoxinucleótido monocatenario o dos hebras de polidesoxinucleótido complementarias que se pueden sintetizar de forma química. Tales oligonucleótidos sintéticos no tienen fosfato en la posición 5' y por lo tanto no se ligarán con otro oligonucleótido sin añadir un fosfato con un ATP en presencia de una quinasa. Un oligonucleótido sintético se puede ligar con un fragmento que no se ha desfosforilado.

La expresión "básicamente idéntico" en el contexto de dos ácidos nucleicos o polipéptidos, puede hacer referencia a dos o más secuencias que tienen, por ejemplo, al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más identidad de (secuencia) de restos de nucleótidos o aminoácidos, cuando se compara y se alinea para la correspondencia máxima, como se mide usando cualquier algoritmo de comparación de secuencias conocido, como se analiza con detalle a continuación, o mediante inspección visual. En aspectos alternativos, la invención proporciona secuencias de ácidos nucleicos y polipéptidos que tienen una identidad sustancial con una secuencia de la invención a modo de ejemplo sobre una región de al menos aproximadamente 10, 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000 o más restos, o una región que varía entre aproximadamente 50 restos con respecto a la longitud completa del ácido nucleico o polipéptido. Las secuencias de ácidos nucleicos de la invención pueden ser básicamente idénticas con respecto a la longitud total de una región que codifica polipéptido.

Una secuencia de aminoácidos "básicamente idéntica" también puede incluir una secuencia que difiere de una secuencia de referencia en una o más sustituciones, supresiones, o inserciones de aminoácidos conservativas o no conservativas, en particular cuando tal sustitución se produce en un sitio que no es el sitio activo de la molécula, y con la condición de que el polipéptido mantenga básicamente sus propiedades funcionales. Por ejemplo, una sustitución de aminoácido conservativa, sustituye un aminoácido por otro de la misma clase (por ejemplo, sustitución de un aminoácido hidrófobo, tal como isoleucina, valina, leucina, o metionina, por otro, o sustitución de un aminoácido polar por otro, tal como sustitución de arginina por lisina, ácido glutámico por ácido aspártico o glutamina por asparagina). Uno o más aminoácidos se pueden suprimir, por ejemplo, de una pectato liasa, dando como resultado la modificación de la estructura del polipéptido, sin alterar de forma significativa su actividad biológica. Por ejemplo, se pueden eliminar los aminoácidos amino o carboxilo terminales que no son necesarios para la actividad de pectato liasa.

"Hibridación" incluye el proceso mediante el que una hebra de ácido nucleico que une con una hebra complementaria a través de emparejamiento de bases. Las reacciones de hibridación pueden ser sensibles y selectivas de modo que se puede identificar una secuencia de interés en particular incluso en muestras en las que está presente a concentraciones bajas. Las condiciones rigurosas se pueden definir, por ejemplo, las concentraciones de sal o formamida en las soluciones de hibridación previa y de hibridación, o mediante la temperatura de hibridación, y son bien conocidas en la técnica. Por ejemplo, la rigurosidad se puede aumentar mediante reduciendo la concentración de sal, aumentando la concentración de formamida, o elevando la temperatura de hibridación, alterando el tiempo de hibridación, como se describe con detalle a continuación. En aspectos alternativos, los ácidos nucleicos de la invención se definen por su capacidad para hibridarse en diversas condiciones de rigurosidad (por ejemplo, elevada, media, y baja), como se establece en el presente documento.

"Variante" incluye polinucleótidos o polipéptidos de la invención modificados en uno o más pares de bases, codones, intrones, exones, o restos de aminoácidos (respectivamente) que aún mantienen la actividad biológica de una pectato liasa de la invención. Las variantes se pueden producir tiene cualquier número de medios incluidos en los métodos tales como, por ejemplo, PCR propensa a error, redistribución, mutagénesis dirigida a oligonucleótidos, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casete, mutagénesis de conjunto recursiva, mutagénesis de conjunto exponencial, mutagénesis específica de sitio, reensamblaje genético, GSSM™ y cualquier combinación de los mismos. En el presente documento se incluyen técnicas para producir variante de pectato liasa que tiene actividad a un pH o temperatura, por ejemplo, que es diferente de una pectato liasa de tipo silvestre.

La expresión "mutagénesis por saturación" o "GSSM™" incluye un método que usaron cebadores de oligonucleótidos degenerados para introducir mutaciones puntuales en un polinucleótido, como se describe con detalle a continuación.

La expresión "sistema de evolución dirigida optimizado" o "evolución dirigida optimizada" incluye un método para el reensamblaje de fragmentos de secuencias de ácidos nucleicos relacionados, por ejemplo, genes relacionados, y se explica con detalle a continuación.

La expresión "reensamblaje de ligamiento genético" o "SLR" incluye un método de ligamiento de fragmentos de oligonucleótidos de una manera no estocástica, y se explica con detalle a continuación.

Generación y Manipulación de Ácidos Nucleicos

La invención proporciona ácidos nucleicos aislados y recombinantes, por ejemplo, polinucleótidos que tienen una identidad de secuencias con la SEC ID N°: 77, SEC ID N°: 131 o SEC ID N°: 133; ácidos nucleicos que codifican polipéptidos de la invención que tienen una secuencia como se establece en la SEC ID N°: 78, SEC ID N°: 132 o SEC ID N°: 134.

Los ácidos nucleicos de la invención también pueden comprender casetes de expresión, tales como vectores de expresión, en los que en un aspecto, codifican a un polipéptido de la invención. Se desvelan métodos para descubrir nuevas secuencias de pectato liasa usando los ácidos nucleicos de la invención, métodos para inhibir la expresión de genes que pectato liasa, transcritos y polipéptidos usando los ácidos nucleicos de la invención, y métodos para modificar los ácidos nucleicos de la invención, por ejemplo, mediante reensamblaje de ligamiento genético, sistema de evolución dirigida optimizado y/o mutagénesis por saturación de sitio genético (GSSM™).

Los ácidos nucleicos de la invención se pueden preparar, aislar y/o manipular, por ejemplo, mediante clonación y expresión de bibliotecas de ADNc, amplificación de mensajes o ADN genómico mediante PCR, y similares. Los genes homólogos se puede modificar mediante manipulación de un ácido nucleico molde, como se describe en el presente documento. La presente invención se puede poner en práctica en conjunto con cualquier método o protocolo o dispositivo conocido en la técnica, que se describen bien en la bibliografía científica y de patente.

Técnicas Generales

Los ácidos nucleicos usados para poner en práctica la presente intervención, bien ARN, iARN (es decir, ARNi), ácido nucleico antisentido, ADNc, ADN genómico, vectores, virus o híbridos de los mismos, se pueden aislar a partir de una diversidad de fuentes, se pueden modificar por ingeniería genética, amplificar, y/o expresar/generar de forma recombinante. Los polipéptidos recombinantes generados a partir de estos ácidos nucleicos se pueden aislar o clonar de forma individual y someter a ensayo para una actividad deseada. Se puede usar cualquier sistema de expresión recombinante, incluyendo sistemas de expresión bacteriana, mamífero, levadura, insecto o célula vegetal.

Como alternativa, estos ácidos nucleicos se pueden sintetizar *in vitro* mediante técnicas de síntesis química bien conocidas, como se describe, por ejemplo, en Adams (1983) J. Am. Chem. Soc. 105: 661; Belousov (1997) Nucleic Acids Res. 25: 3440-3444; Frenkel (1995) Free Radic. Biol. Med. 19: 373-380; Blommers (1994) Biochemistry 33: 7886-7896; Narang (1979) Meth. Enzymol. 68: 90; Brown (1979) Meth. Enzymol. 68: 109; Beaucage (1981) Tetra. Lett. 22: 1859; documento de Patente de Estados Unidos N° 4.458.066.

Las técnicas para la manipulación de ácidos nucleicos, tales como, por ejemplo, subclonación, sondas de marcado (por ejemplo, marcado de cebador aleatorio usando Klenow polimerasa, traducción de la cabeza, amplificación), secuenciación, hibridación y similares se describen bien en la bibliografía científica y de patente, véase, por ejemplo, Sambrook, ed., MOLECULAR CLONING: A LABORATORY MANUAL (2ª ED.), Vols. 1-3, Cold Spring Harbor Laboratory, (1989); CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Ausubel, ed. John Wiley & Sons, Inc., New York (1997); LABORATORY TECHNIQUES IN BIOCHEMISTRY AND MOLECULAR BIOLOGY: HYBRIDIZATION WITH NUCLEIC ACID PROBES, Parte I. Theory and Nucleic Acid Preparation, Tijssen, ed. Elsevier, N.Y. (1993).

Otro medio útil para obtener y manipular ácidos nucleicos usados en la práctica de los métodos de la invención es clonar a partir de muestras genómicas, y, si se desea, identifica sistemáticamente volver a clonar insertos aislados o amplificados de, por ejemplo, clones genómicos o clones de ADNc. Algunas fuentes de ácidos nucleicos usados en los métodos de la invención incluyen bibliotecas genómicas o de ADNc contenidas, por ejemplo, en cromosomas artificiales de mamífero (MAC), véase, por ejemplo, los documentos de Patente de Estados Unidos N°s Unidos 5.721.118; 6.025.155; cromosomas artificiales humanos, véase, por ejemplo, Rosenfeld (1997) Nat. Genet. 15: 333-335; cromosomas artificiales de levadura (YAC); cromosomas artificiales bacterianos (BAC); cromosomas artificiales P1, véase, por ejemplo, Woon (1998) Genomics 50: 306-316; vectores derivados de P1 (PAC), véase, por ejemplo, Kern (1997) Biotechniques 23: 120-124; cósmidos, virus recombinantes, fagos o plásmidos.

En un aspecto, un ácido nucleico que codifica un polipéptido de la invención se ensambla en una fase apropiada con una secuencia directora capaz de dirigir la secreción del polipéptido traducido o fragmentos del mismo.

La invención proporciona proteínas de fusión y ácidos nucleicos que las codifican. Un polipéptido de la invención se puede fusionar con un péptido o polipéptido heterólogo, tales como péptidos de identificación N-terminal que transmiten características deseadas, tales como aumento de la estabilidad o purificación simplificada. Los péptidos y polipéptidos de la invención también se pueden sintetizar y expresar como proteínas de fusión con uno o más dominios adicionales unidos a los mismos para, por ejemplo, producir un péptido más inmunogénico, para aislar más fácilmente un péptido sintetizado de forma recombinante, para identificar y aislar anticuerpos y linfocitos B que expresan anticuerpo, y similares. Los dominios que facilitan detección y purificación incluyen, por ejemplo, péptidos quelantes de metales tales como tramos de polihistidina y módulos de histidina-triptófano que permiten la purificación sobre metales inmovilizados, dominios de proteína A que permiten la purificación sobre inmunoglobulina inmovilizada, y el dominio usado en el sistema de purificación por extensión/afinidad FLAGS (Immunex Corp, Seattle WA). La inclusión de una secuencia conectora de escisión tal como Factor Xa o enteroquinasa (Invitrogen, San Diego CA) entre un dominio de purificación y el péptido o polipéptido que comprende motivo para facilitar la purificación. Por ejemplo, un vector de expresión puede incluir una secuencia de ácidos nucleicos que codifica epítomos unidos a seis restos de histidina seguido de una tiorredoxina y un sitio de escisión de enteroquinasa (véase por ejemplo, Williams (1995) *Biochemistry* 34: 1787-1797; Dobeli (1998) *Protein Expr. Purif.* 12: 404-414). Los restos de histidina facilitan la detección y purificación mientras que el sitio de escisión de enteroquinasa proporciona un medio para purificar el epítomo del resto de la proteína de fusión. La tecnología perteneciente a vectores que codifican proteínas de fusión y la aplicación de proteínas de fusión se describen bien en la bibliografía científica y de patente, véase por ejemplo, Kroll (1993) *DNA Cell. Biol.*, 12: 441-53.

Secuencias de control de transcripción y traducción

La invención proporciona secuencias de ácidos nucleicos (por ejemplo, ADN) de la invención unidas de forma operativa a secuencia o secuencias de control de expresión (por ejemplo, transcripción o traducción), por ejemplo, promotores o potenciadores, para dirigir o modular la síntesis/expresión del ARN. La secuencia de control de expresión puede ser un vector de expresión. Algunos promotores bacterianos a modo de ejemplo incluyen *lacI*, *lacZ*, T3, T7, *gpt*, *lambda PR*, *PL* y *trp*. Algunos promotores eucariotas a modo de ejemplo incluyen promotor inmediato temprano de CMV, timidina quinasa de VHS, SV40 temprano y tardío, LTR de retrovirus, y metalotioneína I de ratón.

Algunos promotores adecuados para la expresión de un polipéptido en bacterias incluyen los promotores de *lac* o *trp* *E. coli*, el promotor *lacI*, el promotor *lacZ*, el promotor T3, el promotor T7, el promotor *gpt*, el promotor *lambda PR*, el promotor *lambda PL*, los promotores de operones que codifican enzimas glicolíticas tales como 3-fosfoglicerato quinasa (PGK), y el promotor de fosfatasa ácida. Algunos promotores eucariotas incluyen el promotor temprano inmediato de CMV, el promotor de timidina quinasa de VHS, promotores de choque térmico, el promotor SV40 temprano y tardío, LTR de retrovirus, y el promotor de metalotioneína I de ratón. También se pueden usar otros promotores conocidos porque controlan la expresión de genes en células procariotas o eucariotas o sus virus.

Promotores Vegetales Específicos de Tejido

La invención proporciona casetes de expresión que se pueden expresar de una manera específica de tejido, por ejemplo, que pueden expresar una pectato liasa de la invención de una manera específica de tejido. La invención también proporciona plantas o semillas que expresan una pectato liasa de la invención de una manera específica de tejido. La especificidad de tejido puede ser específica de semilla, específica de tallo, específica de hoja, específica de raíz, específica de fruta y similares.

En un aspecto, un promotor constitutivo tal como el promotor CaMV 35S se puede usar para expresión en partes específicas de la planta o semilla o por toda la planta. Por ejemplo, para sobreexpresión, se puede usar un fragmento de promotor vegetal que dirigirá la expresión de un ácido nucleico en algunos o todos los tejidos de una planta, por ejemplo, una planta regenerada. En el presente documento tales promotores se denominan promotores "constitutivos" y son activos en la mayoría de las condiciones ambientales y estados de desarrollo o diferenciación celular. Algunos ejemplos de promotores constitutivos incluyen la región de inicio de la transcripción del virus del mosaico de la coliflor (CaMV) 35S, el promotor 1' o 2' derivados de T-ADN de *Agrobacterium tumefaciens*, cualquier otra región de inicio de la transcripción de diversos genes vegetales conocidos por los expertos. Tales genes incluyen, por ejemplo, *ACT11* de *Arabidopsis* (Huang (1996) *Plant Mol. Biol.* 33: 125-139); *Cat3* de *Arabidopsis* (GenBank Nº U43147, Zhong (1996) *Mol. Gen. Genet.* 251: 196-203); el gen que codifica la proteína desaturasa vehículo de estearoil-acilo de *Brassica napus* (Genbank Nº X74782, Solocombe (1994) *Plant Physiol.* 104: 1167-1176); *Gpc1* de maíz (GenBank Nº X15596; Martinez (1989) *J. Mol. Biol.* 208: 551-565); el *Gpc2* de maíz (GenBank Nº U45855, Manjunath (1997) *Plant Mol. Biol.* 33: 97-112); promotores vegetales que se describen en los documentos de Patente de Estados Unidos Nºs 4.962.028; 5.633.440.

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

Como alternativa, el promotor vegetal puede dirigir la expresión del ácido nucleico que expresa pectato liasa en un tejido, órgano o tipos celulares específicos (es decir promotor específico de tejidos) o de otro modo puede estar bajo un control ambiental o de desarrollo más preciso o bajo el control de un promotor inducible. Algunos ejemplos de

condiciones ambientales que pueden influir en la transcripción incluyen condiciones anaerobias, temperatura elevada, la presencia de luz, o pulverización es con agentes químicos/humanos. Por ejemplo, la invención incorpora el promotor inducible de la sequía del maíz (Busk (1997) mencionado anteriormente); el promotor del frío, sequía, y alto contenido salino inducible de la patata (Kirch (1997) Plant Mol. Biol. 33: 897 909).

Los promotores específicos de tejido pueden estimular la transcripción solamente dentro de un cierto marco de tiempo de la etapa de desarrollo dentro de ese tejido. Véase, por ejemplo, Blazquez (1998) Plant Cell 10: 791-800, que caracteriza el promotor del gen LEAFY de *Arabidopsis*. Véase también Cardon (1997) Plant J 12: 367-77, que describe el factor de transcripción SPL3, que reconoce un motivo de secuencia conservada en la región del promotor del gen AP1 de identidad del meristemo floral de *A. thaliana*; y Mandel (1995) Plant Molecular Biology, Vol. 29, pp 995-1004, que describe el promotor del meristemo eIF4. Se pueden usar promotores específicos de tejido que son activos durante todo el ciclo de vida de un tejido particular. En un aspecto, los ácidos nucleicos de la invención se unen de forma operativa a un promotor activo principalmente solo en células de fibra de algodón. En un aspecto, los ácidos nucleicos de la invención según en de forma operativa a un promotor activo principalmente durante las etapas de elongación de la célula de fibra de algodón, por ejemplo, como se describe en Rinehart (1996) mencionado anteriormente. Los ácidos nucleicos se pueden unir de forma operativa al promotor del gen Fb12A para que se expresen de forma preferente en células de fibra de algodón (En el mismo lugar). Véase también, John (1997) Proc. Natl. Acad. Sci. USA 89: 5769-5773; John, *et al.*, documentos de Patente de Estados Unidos N^{os} 5.608.148 y 5.602.321, que describen promotores específicos de fibra de algodón y métodos para la construcción de plantas de algodón transgénicas. También se pueden usar promotores específicos de raíz para expresar los ácidos nucleicos de la invención. Algunos ejemplos de promotores específicos de raíz incluyen el promotor del gen de la alcohol deshidrogenasa (DeLisle (1990) Int. Rev. Cytol. 123: 39-60). Otros promotores que se pueden usar para expresar los ácidos nucleicos de la invención incluyen, por ejemplo, promotores específicos de óvulo, específicos de embrión, específicos de endospermo, específicos de integumento, específicos de revestimiento de semillas, o algunas combinaciones de los mismos; un promotor específico de hoja (véase, por ejemplo, Busk (1997) Plant J. 11: 1285 1295, que describe un promotor específico de hoja en maíz); el promotor ORF13 de *Agrobacterium rhizogenes* (que presenta una actividad elevada en raíces, véase, por ejemplo, Hansen (1997) mencionado anteriormente); un promotor específico del polen de maíz (véase, por ejemplo, Guerrero (1990) Mol. Gen. Genet. 224: 161 168); se puede usar un promotor del tomate activo durante la maduración del fruto, senescencia y abscisión de las hojas y, en menor medida, de las flores (véase, por ejemplo, Blume (1997) Plant J. 12: 731 746); un promotor específico de pistilo del gen SK2 de la patata (véase, por ejemplo, Ficker (1997) Plant Mol. Biol. 35: 425 431); el gen Blec4 del guisante, que es activo en el tejido epidérmico de ápices de raíz vegetativos y florales de alfalfa transgénica convirtiéndole en una herramienta útil para dirigir la expresión de genes extraños a la capa epidérmica de de lotes fibras con crecimiento activo; el gen BEL1 específico de óvulo (véase, por ejemplo, Reiser (1995) Cell 83: 735-742, GenBank N^o U39944); y/o, el promotor en Klee, documento de Patente de Estados Unidos N^o 5.589.583, que describe una región promotora de planta que es capaz de transmitir niveles elevados de transcripción en el tejido meristemático y/o células que se dividen rápidamente.

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

Los ácidos nucleicos de la invención también se pueden unir de forma operativa a promotores de plantas que son inducibles después de su exposición a reactivos químicos que se pueden aplicar a la planta, tales como herbicidas o antibióticos. Por ejemplo, se puede usar el promotor In2-2 del maíz, activado por protectores del herbicida de bencenosulfonamida (De Veylder (1997) Plant Cell Physiol. 38: 568-577); la aplicación de diferentes protectores de herbicidas induce distintos patrones de expresión genética, incluyendo la expresión en la raíz, hidatodos, y en el meristemo apical de la raíz. La secuencia de codificación puede estar bajo el control de, por ejemplo, un promotor inducible por tetraciclina, por ejemplo, como se describe con plantas de tabaco transgénicas que contienen el gen de la arginina descarboxilasa de *Avena sativa* L. (avena) (Masgrau (1997) Plant J. 11: 465-473); o, un elemento sensible al ácido salicílico (Stange (1997) Plant J. 11: 1315-1324). Usando promotores inducidos de forma química (por ejemplo, hormona o pesticida), es decir, promotor sensible a un agente químico que se puede aplicar a la planta transgénica en el campo, la expresión de un polipéptido de la invención se puede inducir en una etapa del desarrollo de la planta en particular. Por lo tanto, la invención también proporciona plantas transgénicas que contienen un gen inducible para polipéptidos de la invención cuyo gama de hospedador es se limita a especies vegetales diana, tales como maíz, arroz, cebada, trigo, patata u otros cultivos, inducible en cualquier etapa del desarrollo del cultivo.

Un experto reconocerá que un promotor vegetal específico de tejido de dirigir la expresión de secuencias unidas de forma operativa en tejidos distintos del tejido diana. Por lo tanto, un promotor específico de tejido es uno que dirige la expresión preferentemente en el tejido diana o tipo de célula, pero también puede conducir a una cierta expresión en otros tejidos.

Los ácidos nucleicos de la invención también se pueden unir de forma operativa a promotores vegetales que son inducibles después de su exposición a reactivos químicos. Estos reactivos incluyen, por ejemplo, herbicidas, auxinas sintéticas, o antibióticos que se pueden aplicar, por ejemplo, pulverizar, en plantas transgénicos. La expresión inducible de los ácidos nucleicos que producen pectato liasa de la invención permitirá al cultivador seleccionar plantas con la expresión y/o actividad de pectato liasa óptima. El desarrollo de partes de plantas se puede controlar de este modo. Por lo tanto, la invención proporciona los medios para facilitar la cosecha de plantas y partes de plantas. Por ejemplo, en diversas realizaciones, se usa el promotor In2-2 del maíz, activado por protectores de herbicidas de bencenosulfonamida (De Veylder (1997) *Plant Cell Physiol.* 38: 568-577); la aplicación de diferentes protectores de herbicidas induce distintos patrones de expresión genética, incluyen la expresión en la raíz, hidatodos, y en el meristemo apical de la raíz. Las secuencias de codificación de la invención también están bajo el control de un promotor inducible por tetraciclina, por ejemplo, como se describe con plantas de tabaco transgénica si contienen el gen de la arginina descarboxilasa de la *Avena sativa* L. (avena) (Masgrau (1997) *Plant J.* 11: 465-473); o, un elemento sensible al ácido salicílico (Stange (1997) *Plant J.* 11: 1315-1324).

Si se desea una expresión apropiada del polipéptido, se debería incluir una región de poliadenilación en el extremo en la posición 3' de la región de codificación. La región de poliadenilación se puede derivar del gen natural, a partir de una diversidad de otros genes vegetales, o de genes en el T-ADN de *Agrobacterium*.

Vectores de expresión y vehículos de clonación

La invención proporciona vectores de expresión y vehículos de clonación que comprenden ácidos nucleicos de la invención, por ejemplo, secuencias que codifican las pectato liasas de la invención. Los vectores de expresión y vehículos de clonación de la invención pueden comprender partículas virales, baculovirus, fago, plásmidos, fagémidos, cósmidos, fósmodos, cromosomas artificiales bacterianos, ADN viral (por ejemplo, vaccinia, adenovirus, virus de la viruela aviar, pseudorrabia y derivados de SV40), cromosomas artificiales basados en P1, plásmidos de levadura, cromosomas artificiales de levadura, y cualquier otro vector específico de hospedadores específicos de interés (tales como *Bacillus*, *Aspergillus* y levadura). Algunos vectores de invención pueden incluir secuencias de ADN cromosómico, no cromosómico y sintético. Los expertos de la materia conocen grandes números de vectores adecuados, y están disponibles en el mercado. Algunos vectores a modo de ejemplo incluyen: bacterianos: vectores pQE (Qiagen), plásmidos pBluescript, vectores pNH, (vectores lambda-ZAP (Stratagene); ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eucariotas: pXT1, pSG5 (Stratagene), pSVK3, pBPV, pMSG, pSVLSV40 (Pharmacia). Sin embargo, se puede usar cualquier otro plásmido u otro vector siempre y cuando se puedan replicar y sean viables en el hospedador. Con la presente invención se pueden usar vectores con número de copias bajo o número de copias elevado.

El vector de expresión puede comprender un promotor, un sitio de unión a ribosoma para inicio de la traducción y un terminado de la transcripción. El vector también puede incluir secuencias apropiadas para la amplificación de la expresión. Algunos vectores de expresión de mamíferos pueden comprender un origen de replicación, si fuera necesario sitios de unión ribosomas, un sitio de poliadenilación, sitios donantes y aceptores de corte y empalme, secuencias de terminación de la transcripción, y secuencias no transcritas de flanco en la posición 5'. En algunos aspectos, las secuencias de ADN derivadas de los sitios de corte y empalme y poliadenilación de SV40 se pueden usar para proporcionar los elementos genéticos no transcritos necesarios.

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

Los vectores para expresión del polipéptido o fragmento del mismo en células eucariotas también pueden contener potenciadores para aumentar los niveles de expresión. Algunos potenciales son elementos de ADN de actuación *cis*, normalmente de aproximadamente 10 pb a aproximadamente 300 pb de longitud que actúan sobre un promotor para aumentar su transcripción. Los ejemplos incluyen el potenciador SV40 en el lado tardío del origen de replicación de 100 pb a 270 pb, el potenciador del promotor temprano del citomegalovirus, el potenciador de polioma en el lado tardío del origen de replicación, y los potenciales de adenovirus.

Una secuencia de ácidos nucleicos se puede insertar en un vector mediante una diversidad de procedimientos. En general, la secuencia se liga con la posición deseada en el vector después de la digestión del inserto y el vector con endonucleasas de restricción apropiadas. Como alternativa, se pueden ligar extremos romos tanto en el inserto como en el vector. En la técnica se conoce una diversidad de técnicas de clonación, por ejemplo, como se describe

en Ausubel y Sambrook. Se considera que tales procedimientos y otros están dentro del alcance de los expertos en la materia.

El Héctor puede estar en forma de un plásmido, una partícula viral, o un fago. Otros vectores incluyen secuencias de ADN cromosómico, no cromosómico y sintético, derivados de SV40; plásmidos bacterianos, ADN de fagos, baculovirus, plásmidos de levadura, vectores derivados de combinaciones de plásmidos y ADN de fagos, ADN viral tal como vaccinia, adenovirus, virus de la viruela aviar, y pseudorrabia. Una diversidad de vectores de clonación y expresión para uso con hospedadores procariotas y eucariotas se describen, por ejemplo, en Sambrook.

Algunos vectores bacterianos en particular que se pueden usar incluyen los plásmidos disponibles en el mercado que comprenden elementos genéticos del vector de clonación bien conocido pBR322 (ATCC 37017), pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Suecia), GEM1 (Promega Biotec, Madison, WI, USA) pQE70, pQE60, pQE-9 (Qiagen), pD10, psiX174 pBluescript II KS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene), ptrc99a, pKK223-3, pKK233-3, DR540, pRIT5 (Pharmacia), pKK232-8 y pCM7. Algunos vectores eucariotas en particular incluyen pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, y pSVL (Pharmacia). Sin embargo, se puede usar cualquier otro vector siempre y cuando se pueda replicar y sea viable en la célula hospedadora.

Los ácidos nucleicos de la invención se pueden expresar en casetes de expresión, vectores o virus y se pueden expresar de forma transitoria o estable en células vegetales y semillas. Un sistema de expresión transitorio a modo de ejemplo la sistemas de expresión episomales, por ejemplo, ARN viral del virus del mosaico de la coliflor (CaMV) generado en el núcleo por transcripción de un minicromosoma episomal que contiene ADN superenrollado, véase, por ejemplo, Covey (1990) Proc. Natl. Acad. Sci. USA 87: 1633-1637. Como alternativa, se pueden insertar secuencias de codificación, es decir, todos o subfragmentos de secuencias de la invención en un genoma de célula hospedadora vegetal que se convierte en parte integral del ADN cromosómico del hospedador. Los transcritos sentido o antisentido se pueden expresar de esta manera. Un vector que comprende las secuencias (por ejemplo, promotores o regiones de codificación) de ácidos nucleicos de la invención puede comprender un gen marcador que confiere un fenotipo seleccionable en una célula vegetal o una semilla. Por ejemplo, el marcador puede codificar resistencia a biocidas, en particular resistencia a antibióticos, tal como resistencia a la kanamicina, G418, bleomicina, higromicina, o resistencia a herbicidas, tal como resistencia a clorosulfurón o Basta.

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

En un aspecto, el vector de expresión puede tener dos sistemas de replicación para permitir que se mantenga en dos organismos, por ejemplo en células de mamífero o de insecto para expresión y en un hospedador procariota para clonación y amplificación. Además, para lectores de expresión de integración, el vector de expresión puede contener al menos una secuencia homóloga con el genoma de la célula hospedadora. Puede contener dos secuencias homólogas que flanquean el constructo de expresión. El director de integración se puede dirigir a un locus específico en la célula hospedadora por selección de la secuencia homóloga apropiada para inclusión en el vector. Los constructos para vectores de integración se conocen bien en la técnica.

Algunos vectores de expresión de la invención también pueden incluir un gen marcador seleccionable para permitir la selección de cepas bacterianas que se han transformado, por ejemplo, genes que hacen que la bacteria sea resistente a fármacos tales como ampicilina, cloramfenicol, eritromicina, kanamicina, neomicina y tetraciclina. Los marcadores seleccionables también pueden incluir genes biosintéticos, tales como los que se encuentran en las rutas biosintéticas de histidina, triptófano y leucina.

Células hospedadoras y células transformadas

La invención también proporciona una célula transformada que comprende secuencias de ácidos nucleicos de la invención, por ejemplo, una secuencia que codifica un pectato liasa de la invención, o un vector de la invención. La célula hospedadora puede ser cualquiera de las células hospedadoras familiares para los expertos en la materia, que incluyen células procariotas, células eucariotas, tales como células bacterianas, célula fúngicas, células de levadura, células de mamífero, células de insecto, o células vegetales. Algunas células bacterianas a modo de ejemplo incluyen *E. coli*, *Streptomyces*, *Bacillus subtilis*, *Bacillus cereus*, *Salmonella typhimurium* y diversas especies dentro de los géneros *Bacillus*, *Streptomyces*, y *Staphylococcus*. Algunas células de insecto a modo de ejemplo incluyen S2 de *Drosophila* y Sf9 de *Spodoptera*. Algunas células de levadura a modo de ejemplo incluyen *Pichia pastoris*, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*. Algunas células animales a modo de ejemplo

incluyen CHO, COS o melanoma de Bowes o cualquier línea celular de ratón o humana. La selección de un hospedador apropiado está dentro de las capacidades de los expertos en la materia. Algunas técnicas para transformar una gran diversidad de especies de plantas superiores se conocen bien y se describen en la bibliografía técnica y científica. Véase, por ejemplo, Weising (1988) Ann. Rev. Genet. 22: 421-477, documento de Patente de Estados Unidos N° 5.750,870.

El vector se puede introducir en las células hospedadoras usando cualquiera de una diversidad de técnicas, que incluyen transformación, transfección, transducción, infección viral, pistolas genéticas, o transferencia de genes mediada por Ti. Algunos métodos en particular incluyen transfección con fosfato cálcico, transfección mediada con DEAE-Dextrano, lipofección, o electroporación (Davis, L., Dibner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

En un aspecto, los ácidos nucleicos o vectores de la invención se introducen en las células a la identificación sistemática, de este modo, los ácidos nucleicos entran en las células de una manera adecuada para expresión posterior del ácido nucleico. El método de introducción va dictado en gran medida por el tipo de célula a la que se dirige. Algunos métodos a modo de ejemplo incluyen precipitación con CaPO_4 , fusión de liposomas, lipofección (por ejemplo, LIPOFECTIN™), electroporación, infección viral, etc. Los ácidos nucleicos candidatos se pueden integrar de forma estable en el genoma de la célula hospedadora (por ejemplo, con introducción retroviral) o pueden existir de forma transitoria o de forma estable en el citoplasma (es decir a través del uso de plásmidos tradicionales, usando secuencias reguladoras convencionales, marcadores de selección, etc.). Dado que muchas identificaciones sistemáticas farmacéuticamente importantes requieren dianas celulares humanas o de mamífero modelo, son preferentes los vectores retrovirales capaces de transfectar tales dianas.

Cuando sea apropiado, las células hospedadoras modificadas por ingeniería se pueden cultivar en medios nutrientes convencionales modificados cuando sea apropiado para activación de promotores, selección de transformantes o amplificación de los genes de la invención. Después de la transformación de una cepa hospedadora adecuada y crecimiento de la cepa hospedadora hasta una densidad celular apropiada, el promotor seleccionado se puede inducir mediante medios apropiados (por ejemplo, desplazamiento de temperatura o inducción química) y las células se pueden cultivar durante un período adicional para permitir que produzcan el polipéptido o fragmento del mismo deseados.

Algunas células se pueden cosechar por centrifugación, interrumpir mediante medios físicos o químicos, y el extracto en bruto resultante se mantiene para purificación adicional. Algunas células victorianas usadas para expresión de proteínas se pueden interrumpir mediante cualquier método conveniente, que incluye ciclo de congelación-descongelación, sonicación, interrupción mecánica, o uso de agentes de lisado celular. Tales métodos son bien conocidos por los expertos en la materia. El polipéptido expresado o fragmento del mismo se puede recuperar y purificar a partir de cultivos de células recombinantes mediante métodos que incluyen precipitación con sulfato de amonio o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía con fosfocelulosa, cromatografía por interacción hidrófoba, cromatografía por afinidad, cromatografía con hidroxilapatito y cromatografía con lectina. Se pueden usar etapas de replegamiento de proteínas, cuando sea necesario, para completar la configuración del polipéptido. Si se desea, se puede usar cromatografía líquida de alto rendimiento (HPLC) para las etapas de purificación final.

También se pueden usar diversos sistemas de cultivo de células de mamífero para expresar proteína recombinante. Algunos ejemplos de sistemas de expresión de mamífero incluyen las líneas COS-7 de fibroblastos de riñón de mono y otras líneas celulares capaces de expresar proteínas a partir de un vector compatible, tal como las líneas celulares C127, 3T3, CHO, HeLa y BHK.

Los constructos en las células hospedadoras se pueden usar de manera convencional para producir el producto genético codificado por la secuencia recombinante. Dependiendo del hospedador usado en un procedimiento de producción recombinante, los polipéptidos producidos por las células hospedadoras que contienen el vector pueden estar glicosilados o pueden estar no glicosilados. Los polipéptidos de la invención también pueden incluir o no un resto inicial del aminoácido metionina.

También se pueden usar sistemas de traducción libres de células para producir un polipéptido de la invención. Algunos sistemas de traducción libres de células pueden usar ARNm transcritos a partir de un constructor de ADN que comprende promotor unido de forma operativa a un ácido nucleico que codifica el polipéptido o fragmentos del mismo. En algunos aspectos, el constructo de ADN se puede linealizar antes de la realización de una reacción de transcripción *in vitro*. A continuación, el ARNm transcritos se incuban con un extracto de traducción libre de células apropiado, tal como un extracto de reticulocitos de conejo, para producir el polipéptido deseado o fragmento del mismo.

Los vectores de expresión pueden contener uno o más genes marcadores seleccionables para proporcionar un rasgo fenotípico para selección de células hospedadoras transformadas tales como dihidrofolato reductasa o resistencia a neomicina para cultivo celular eucariota, o tal como resistencia a tetraciclina o ampicilina en *E. coli*.

Amplificación de Ácidos Nucleicos

En la práctica de la invención, los ácidos nucleicos de la invención que codifican las pectato liasas de la invención, o ácidos nucleicos modificados de la invención, se pueden reproducir mediante amplificación. La amplificación también se puede usar para clonar o modificar los ácidos de de la invención. Por lo tanto, la divulgación proporciona secuencias cebadoras de amplificación para amplificar ácidos nucleicos de la invención. Un experto en la materia puede diseñar secuencias cebadoras de amplificación para cualquier parte o la longitud completa de estas secuencias. Un ácido nucleico se puede amplificar mediante un par cebador cómo se establecen mediante aproximadamente los primeros (la posición 5') 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, o 25 restos de un ácido nucleico de la divulgación, y aproximadamente los primeros (la posición 5') 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, o 25 restos de la hebra complementaria (por ejemplo, de la SEC ID N°: 1, SEC ID N°: 3, SEC ID N°: 5, SEC ID N°: 7, SEC ID N°: 9, SEC ID N°: 11, SEC ID N°: 13, SEC ID N°: 15, SEC ID N°: 17, SEC ID N°: 19, SEC ID N°: 21, SEC ID N°: 23, SEC ID N°: 25, SEC ID N°: 27, SEC ID N°: 29, SEC ID N°: 31, SEC ID N°: 33, SEC ID N°: 35, SEC ID N°: 37, SEC ID N°: 39, SEC ID N°: 41, SEC ID N°: 43, SEC ID N°: 45, SEC ID N°: 47, SEC ID N°: 49, SEC ID N°: 51, SEC ID N°: 53, SEC ID N°: 55, SEC ID N°: 57, SEC ID N°: 59, SEC ID N°: 61, SEC ID N°: 63, SEC ID N°: 65, SEC ID N°: 67, SEC ID N°: 69, SEC ID N°: 71, SEC ID N°: 73, SEC ID N°: 75, SEC ID N°: 77, SEC ID N°: 79, SEC ID N°: 81, SEC ID N°: 83, SEC ID N°: 85, SEC ID N°: 87, SEC ID N°: 89, SEC ID N°: 91, SEC ID N°: 93, SEC ID N°: 95, SEC ID N°: 97, SEC ID N°: 99, SEC ID N°: 101, SEC ID N°: 103, SEC ID N°: 105, SEC ID N°: 107, SEC ID N°: 109, SEC ID N°: 111, SEC ID N°: 113, SEC ID N°: 115, SEC ID N°: 117, SEC ID N°: 119, SEC ID N°: 121, SEC ID N°: 123, SEC ID N°: 125, SEC ID N°: 127, SEC ID N°: 129, SEC ID N°: 131, SEC ID N°: 133).

También se pueden usar reacciones de amplificación para cuantificar la cantidad de ácido nucleico en una muestra (tal como la cantidad de mensaje en una muestra de célula), marcar el ácido nucleico (por ejemplo, para aplicarlo a una matriz o una transferencia), detectar el ácido nucleico, o cuantificar la cantidad de un ácido nucleico específico en una muestra. Se puede amplificar el mensaje aislado a partir de una célula o una biblioteca de ADNc.

El experto en la materia puede seleccionar y diseñar cebadores de amplificación de oligonucleótidos adecuados. En la técnica también se conocen algunos métodos de amplificación, e incluyen, por ejemplo, reacción en cadena de la polimerasa, PCR (véase, por ejemplo, PCR PROTOCOLS, A GUIDE TO METHODS AND APPLICATIONS, ed. Innis, Academic Press, N.Y. (1990) y PCR STRATEGIES (1995), ed. Innis, Academic Press, Inc., N.Y., reacción en cadena de ligasa (LCR) (véase, por ejemplo, Wu (1989) Genomics 4: 560; Landegren (1988) Science 241: 1077; Barringer (1990) Gene 89:117); amplificación de la transcripción (véase, por ejemplo, Kwok (1989) Proc. Natl. Acad. Sci. USA 86: 1173); y, replicación de secuencias autosostenida (véase, por ejemplo, Guatelli (1990) Proc. Natl. Acad. Sci. USA 87: 1874); amplificación de la Q Beta replicasa (véase, por ejemplo, Smith (1997) J. Clin. Microbiol. 35: 1477-1491), ensayo de amplificación automatizada de la Q-beta replicasa (véase, por ejemplo, Burg (1996) Mol. Cell. Probes 10: 257-271) otras técnicas mediadas por la ARN polimerasa (por ejemplo, NASBA, Cangene, Mississauga, Ontario); véase también Berger (1987) Methods Enzymol. 152: 307-316; Sambrook; Ausubel; Patentes de Estados Unidos N° 4.683.195 y 4.683.202; Sooknanan (1995) Biotechnology 13: 563-564.

Determinación del grado de identidad de secuencias

La invención proporciona ácidos nucleicos que comprenden secuencias que tienen una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más, o completa (100 %) con la SEC ID N°: 77 y ácidos nucleicos que codifican la SEC ID N°: 78, sobre una región de al menos aproximadamente 950, 1000, 1050, 1100, 1150 o más, restos. La invención proporciona polipéptidos que comprenden secuencias que tienen una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o más, o completa (100 %) con un polipéptido a modo de ejemplo de la invención. El alcance de la identidad de la secuencia (homología) se puede determinar usando cualquier programa informático y parámetros asociados, que incluyen los que se describen en el presente documento, tal como BLAST 2.2.2, o FASTA versión 3.0t78, con los parámetros por defecto.

Algunas secuencias homólogas también incluyen secuencias de ARN en las que las uridinas sustituyen a las timinas en las secuencias de ácidos nucleicos. Las secuencias homólogas se pueden obtener usando cualquiera de los procedimientos que se describen en el presente documento o pueden resultar de la corrección de un error de secuenciación. Se observará que las secuencias de ácidos nucleicos como se establece en el presente documento se pueden representar en el formato de caracteres individuales tradicional (véase, por ejemplo, Stryer, Lubert. Biochemistry, 3ª Ed., W. H Freeman & Co., New York) o bien cualquier otro formato que registre la identidad de los nucleótidos en una secuencia.

En el presente documento se usan diversos programas de comparación de secuencias identificados. Las identidades (homologías) de las secuencias de proteínas y/o ácidos nucleicos se pueden evaluar usando cualquiera de la diversidad de algoritmos de comparación de secuencias y programas conocidos en la técnica. Tales algoritmos y programas incluyen, pero no se limitan a, TBLASTN, BLASTP, FASTA, TFASTA, y CLUSTALW (Pearson y Lipman, Proc. Natl. Acad. Sci. USA 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).

La homología o identidad se pueden medir usando software de análisis de secuencia (por ejemplo, el Paquete de Software de Análisis de Secuencias del Grupo de Informática Genética, Centro de Biotecnología de la Universidad de Wisconsin, 1710 University Avenue, Madison, WI 53705). Tal software empareja secuencias similares mediante la asignación de grados de homología a diversas supresiones, instituciones y otras modificaciones. Los términos "homología" e "identidad" en el contexto de dos o más secuencias de ácidos nucleicos o polipéptidos, se refieren a dos o más secuencias o subsecuencias que son las mismas o tienen un porcentaje de restos de aminoácidos polinucleótidos especificado que son los mismos cuando se comparan y alinean para una correspondencia máxima sobre una ventana de comparación por región designada como se mide usando cualquier número de algoritmos de comparación de secuencias mediante alineación manual e inspección visual. Para comparación de secuencias, una secuencia puede actuar como una secuencia de referencia, por ejemplo, una secuencia de la invención, con la que se comparan las secuencias de ensayo. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de ensayo y de referencia se introducen en un ordenador, se designan las coordenadas de la subsecuencia, si fuera necesario, y se designan los parámetros del programa de algoritmo de secuencias. Se pueden usar parámetros del programa por defecto, o parámetros alternativos. El algoritmo de comparación de secuencias calcula a continuación el porcentaje de identidades de secuencias para las secuencias de ensayo con respecto a la secuencia de referencia, basándose en los parámetros del programa.

Una "ventana de comparación", como se usa en el presente documento, incluye referencia a un segmento de uno cualquiera de los números de restos contiguos. Por ejemplo, los restos contiguos que varían en cualquier parte de 20 a la longitud completa de una secuencia de polipéptidos o ácidos nucleicos a modo de ejemplo de la invención se comparan con una secuencia de referencia del mismo número de posiciones contiguas después de haber alineado las dos secuencias de forma óptima. Si la secuencia de referencia tiene el requisito de identidad de secuencia con una secuencia de polipéptidos o ácidos nucleicos a modo de ejemplo de la invención, por ejemplo, una identidad de secuencia de un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o superior con una secuencia de la invención, esa secuencia está dentro del alcance de la invención. En realizaciones alternativas, las subsecuencias que varían de aproximadamente 20 a 600, de aproximadamente 50 a 200, y de aproximadamente 100 a 150 se comparan con una secuencia de referencia del mismo número de posiciones contiguas después de haber alineado las dos secuencias de forma óptima. En la técnica se conocen bien algunos métodos para alineación de secuencias para comparación. La alineación de secuencias óptima para comparación se puede realizar, por ejemplo, mediante el algoritmo de homología local de Smith y Waterman, Adv. Appl. Math. 2: 482, 1981, mediante el algoritmo de alineación de homología de Needleman y Wunsch, J. Mol. Biol. 48: 443, 1970, mediante la búsqueda del método de similitud de Person y Lipman, Proc. Nat'l. Acad. Sci. USA 85: 2444, 1988, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el Paquete de Software de Genética de Wisconsin, Grupo de Informática Genética, 575 Science Dr., Madison, WI), o mediante alineación manual e inspección visual. Otros algoritmos para determinar la homología o identidad incluyen, por ejemplo, además de un programa BLAST (Herramienta de Búsqueda de Alineación Local Básica en el Centro Nacional para Información Biológica), ALIGN, AMAS (Análisis de Múltiples Secuencias Alineadas), AMPS (Alineación de Múltiples Secuencias de Proteínas), ASSET (Herramienta de Evaluación Estadística de Segmentos Alineados), BANDS, BESTSCOR, BIOSCAN (Nodo de Análisis Comparativo de Secuencias Biológicas), BLIMPS (Buscador BLocks IMProved), FASTA, Intervals & Points, BMB, CLUSTAL V, CLUSTAL W, CONSENSUS, LCONSENSUS, WCONSENSUS, Smith-Waterman algorithm, algoritmo DARWIN Las Vegas, FNAT (Herramienta de Alineación Forzada de Nucleótidos), Framealign, Framesearch, DYNAMIC, FILTER, FSAP (Paquete de Análisis de Secuencias de Fristensky), GAP (Programa de Alineación Global), GENAL, GIBBS, GenQuest, ISSC (Comparación Sensible de Secuencias), LALIGN (Alineación de Secuencia Local), LCP (Programa de Contenido Local), MACAW (Mesa de Laboratorio de Construcción de Alineaciones múltiples y Análisis), MAP (Programa de Alineación Múltiple), MBLKP, MBLKN, PIMA (Alineación Multi-secuencia Inducida por Patrón), SAGA (Alineación de Secuencias mediante Algoritmo Genético) y WHAT-IF. Tales programas de alineación también se pueden usar para identificar sistemáticamente bases de datos del genoma para identificar secuencias de polinucleótidos que tienen secuencias básicamente idénticas. Una serie de bases de datos del genoma están disponibles, por ejemplo, una parte básica del genoma humano está disponible como parte del Proyecto de Secuenciación del Genoma Humano (Gibbs, 1995). Se han secuenciado varios genomas, por ejemplo, *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). También se han realizado progresos significativos en la secuenciación de los genomas de organismos modelo, tales como ratón, *C. elegans*, y *Arabidopsis* sp. Las bases de datos que contienen información genómica anotada con alguna información funcional se mantienen mediante diferente organización, y se puede acceder a ellas a través de internet.

Los algoritmos BLAST, BLAST 2.0 y BLAST 2.2.2 también se usan para poner en práctica la invención. Éstos se describen, por ejemplo, en Altschul (1977) Nuc. Acids Res. 25: 3389-3402; Altschul (1990) J. Mol. Biol. 215: 403-410. El software para realizar análisis BLAST está disponible al público a través del Centro Nacional para Información de Biotecnología. Este algoritmo implica primero la identificación de pares de secuencias con alta puntuación (HSP) mediante la identificación de palabras cortas de longitud W en la secuencia de consulta, que se empareja o satisface alguna puntuación de umbral con valor positivo T cuando se alinea con una palabra de la misma longitud en una secuencia de bases de datos. T se refiere al umbral de puntuación de la palabra vecina (Altschul (1990) mencionado anteriormente). Estos aciertos en la palabra vecina inicial actúan como semillas para

iniciar búsquedas para encontrar HSP de mayor longitud que las contienen. Estos aciertos de palabras se extienden en ambas direcciones a lo largo de cada secuencia tan lejos como se pueda aumentar la puntuación de alineación acumulativa. Las puntuaciones acumulativas se calculan usando secuencias de nucleótidos, los parámetros M (puntuación de recompensa para un par de restos coincidentes; siempre >0). Para secuencias de aminoácidos, se usa una matriz de puntuación para calcular la puntuación acumulativa. La extensión de los aciertos de palabra en cada dirección se detienen cuando: la puntuación de la alineación acumulativa disminuye en la cantidad X desde su valor máximo conseguido; la puntuación acumulativa va a cero o inferior, debido a la acumulación de una o más alineaciones de restos de puntuación negativa; se alcanza o el extremo de cualquier secuencia. Los parámetros del algoritmo BLAST, W, T, y X, determinan la sensibilidad y la velocidad de la alineación. El programa BLASTN (para secuencias de nucleótidos) usa como defectos una longitud de palabra (W) de 11, una expectativa (E) de 10, M = 5, N = -4 y una comparación de ambas hebras. Para secuencias de aminoácidos, el programa BLASTP usa como defectos una longitud de palabra de 3, y expectativas (E) de 10, y la matriz de puntuación de BLOSUM62 (véase Henikoff y Henikoff (1989) Proc. Natl. Acad. Sci. USA 89: 10915) alineaciones (B) de 50, expectativa (E) de 10, M = 5, N = -4, y una comparación de ambas hebras. El algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, por ejemplo, Karlin y Altschul (1993) Proc. Natl. Acad. Sci. USA 90: 5873). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad es una más pequeña (P(N)), que proporciona una indicación de la probabilidad mediante la que se podría producir un emparejamiento entre dos secuencias de nucleótidos o aminoácidos por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencias si la probabilidad es una más pequeña en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es inferior a aproximadamente 0,2, más preferentemente inferior a aproximadamente 0,01, y lo más preferentemente inferior aproximadamente 0,001. En un aspecto, las homologías es de las secuencias de proteínas y ácidos nucleicos se evalúan usando la Herramienta de Búsqueda de Alineación Local Básica ("BLAST"). Por ejemplo, se pueden usar cinco programas BLAST específicos para realizar la siguiente tarea: (1) BLASTP y BLAST3 comparan una secuencia de consulta de aminoácidos frente a una base de datos de secuencias de proteínas; (2) BLASTN compara una secuencia de consulta de nucleótidos frente a una base de datos de secuencias de nucleótidos; (3) BLASTX compara los productos de traducción conceptual de seis marcos de una secuencia de nucleótidos de consulta (ambas hebras) frente a una base de datos de secuencias de proteínas; (4) TBLASTN compara una secuencia de proteínas de consulta frente a una base de datos de secuencias de nucleótidos traducida en los seis marcos de lectura (ambas hebras); y, (5) TBLASTX compara las traducciones de los seis marcos de una secuencia de consulta de nucleótidos frente a las traducciones de los seis marcos de una base de datos de secuencias de nucleótidos. Los programas BLAST identifican secuencias homólogas mediante la identificación de segmentos similares, que en el presente documento se denominan " pares de segmentos de puntuación elevada", entre una secuencia de consulta de aminoácidos o ácidos nucleicos y una secuencia de ensayo que se obtiene preferentemente a partir de una base de datos de secuencias de proteínas o de ácidos nucleicos. Los pares de segmentos de puntuación elevada se identifican preferentemente (es decir, se alinean) por medio de una matriz de puntuación, de las cuales se conocen muchas en la técnica. Preferentemente la matriz de puntuación usada es la matriz BLOSUM62 (Gonnet *et al.*, Science 256: 1443-1445, 1992; Henikoff y Henikoff, Proteins 17: 49-61, 1993). Menos preferentemente, también se pueden usar las matrices PAM o PAM250 (véase, por ejemplo, Schwartz y Dayhoff, eds., 1978, Matrices for Detecting Distance Relationships: Atlas of Protein Sequence and Structure, Washington: National Biomedical Research Foundation).

Para determinar si un ácido nucleico tiene la identidad de secuencias necesaria para estar dentro del alcance de la invención, se usa al programa NCBI BLAST 2.2.2, opciones por defecto a blastp. Existen aproximadamente 38 opciones de ajuste en el programa BLAST 2.2.2. En este ejemplo, se usan todos los valores por defecto excepto para el ajuste de filtro por defecto (es decir, todos los parámetros se ajustan por defecto excepto el filtro que se ajusta en OFF); en su lugar se usa un ajuste "-F F", que desactiva el filtro. El uso del filtro por defecto a menudo da como resultado violaciones de Karlin-Altschul debido a la longitud corta de la secuencia.

Los valores por efecto usados en este aspecto de la invención a modo de ejemplo incluyen:

"Filtro para complejidad baja: ON
Tamaño de Palabra: 3
Matriz: Blosum62
Costes de Huecos: Existencia: 11
Extensión:1"

Otros ajustes por defecto pueden ser: filtro para baja complejidad OFF, tamaño de palabra de 3 para proteína, matriz BLOSUM62, penalización por existencia de huecos de -11 y una penalización de extensión de huecos de -1. Un ajuste del programa NCBI BLAST 2.2.2 a modo de ejemplo tiene la opción por defecto "-W" en 0. Esto significa que, sino se realiza al ajuste, el tamaño de palabra tiene un defecto de 3 para proteínas y de 11 para nucleótidos.

Sistemas informáticos y productos de programas informáticos

Para determinar e identificar identidades de secuencias, homologías estructurales, motivos y similares hechos mediante simulación computacional, la secuencia de la invención se puede almacenar, registrar, y manipular en cualquier medio que se pueda leer y al que se pueda acceder mediante un ordenador. Por consiguiente,

ordenadores, sistemas informáticos, medios de lectura en ordenador, productos de programas informáticos y similares pueden tener registradas o almacenadas en los mismos las secuencias de ácidos nucleicos y polipéptidos de la invención. Como se usa en el presente documento, los términos "registrado" y "almacenado" se refieren a un proceso para almacenar información en un medio informático. Un experto en la materia puede adoptar fácilmente cualquier método conocido para registrar información en un medio de lectura en ordenador para generar preparaciones que comprenden una o más de las secuencias de ácidos nucleicos y/o polipéptidos de la invención.

Un medio de lectura en ordenador puede tener registrado en el mismo al menos una secuencia de ácidos nucleicos y/o polipéptidos de la invención. Algunos medios de lectura en ordenador incluyen medios de lectura por vía magnética, medios de lectura por vía óptica, medios de lectura por vía electrónica y medios magnéticos/ópticos. Por ejemplo, los medios de lectura por ordenador pueden ser un disco duro, un disquete, una cinta magnética, CD-ROM, Disco Versátil Digital (DVD), Memoria de Acceso Aleatorio (RAM), o Memoria de Solo Lectura (ROM) así como otros tipos u otros medios conocidos por los expertos en la materia.

Un sistema 100 a modo de ejemplo, que almacena y manipulan las secuencias y la información de las secuencias que se describen en el presente documento, se ilustra en forma de diagrama del bloque en la Figura 1. Como se usa en el presente documento, "sistema informático" se refiere a los componentes del hardware, componentes del software, y componentes de almacenamiento de datos usados para analizar una secuencia de nucleótidos o polipéptidos de la invención. El sistema informático 100 puede incluir un procesador para procesar, acceder y manipular a los datos de la secuencia. El procesador 105 puede ser cualquier tipo bien conocido de unidad de procesamiento central, tal como, por ejemplo, el Pentium III de Intel Corporation, o procesador similar de Sun, Motorola, Compaq, AMD o International Business Machines. El sistema informático 100 es un sistema con finalidad general que comprende el procesador 105 y uno o más componentes internos de almacenamiento de datos 110 para almacenamiento de datos, uno o más dispositivos de recuperación de datos para recuperar los datos almacenados en los componentes de almacenamiento de datos. Un experto en la materia puede observar fácilmente que son adecuados uno cualquiera de los sistemas informáticos disponibles en la actualidad.

En un aspecto, el sistema informático 100 incluye un procesador 105 conectado a un puerto que se conecta a una memoria principal 115 (preferentemente implementada como RAM) uno o más dispositivos internos de almacenamiento de datos 110, tales como un disco duro y/o otros medios de lectura en ordenador que tienen datos registrados en los mismos. El sistema informático 100 puede incluir adicionalmente uno o más dispositivo de recuperación de datos 118 para la lectura de los datos almacenados en el dispositivo interno de almacenamiento de datos 110. El dispositivo de recuperación de datos 118 de representar, por ejemplo, una unidad de disquete, una unidad de disco compacto, una unidad de cinta magnética, o un módem capaz de conexión a un sistema remoto de almacenamiento de datos (por ejemplo, a través de internet) etc. En algunas realizaciones, el dispositivo de almacenamiento de datos interno 110 es un medio de lectura en ordenador extraíble tal como un disquete, un disco compacto, una cinta magnética, etc. que contiene lógica de control y/o datos registrados en el mismo. El sistema informático 100 puede incluir de forma ventajosa o se puede programar mediante un software apropiado para la lectura de la lógica de control y/o los datos desde el componente de almacenamiento de datos una vez insertado en el dispositivo de recuperación de datos. El sistema informático 100 incluyen una pantalla 120 que se usa para presentar la salida a un usuario del ordenador. También se debería indicar que el sistema informático 100 se puede unir a otros sistemas informáticos 125a-c en una red o red de área amplia para proporcionar acceso centralizado el sistema informático 100. El software para acceder y procesar las secuencias de nucleótidos o de aminoácidos de la invención puede residir en la memoria principal 115 durante la ejecución. En algunos aspectos, el sistema informático 100 puede comprender adicionalmente un algoritmo de comparación de secuencias para comparar las secuencias de ácidos nucleicos de la invención. El algoritmo y la secuencia o secuencias se pueden almacenar en un medio de lectura en ordenador. Un "algoritmo de comparación de secuencias" se refiere a uno o más programas que se implementan (de forma local o de forma remota) en el sistema informático 100 para comparar una secuencia de nucleótidos con otras secuencias de nucleótidos y/o compuestos almacenados dentro de un medio de almacenamiento de datos. Por ejemplo, el algoritmo de comparación de secuencias puede comparar las secuencias de nucleótidos de la invención almacenadas en un medio de lectura en ordenador con secuencias de referencia almacenadas en un medio de lectura en ordenador para identificar homología es o motivos estructurales.

Los parámetros usados con los algoritmos mencionados anteriormente se pueden adaptar dependiendo de la longitud de la secuencia y el grado de homología estudiados. En algunos aspectos, los parámetros pueden ser los parámetros por defecto usados por los algoritmos paren ausencia de instrucciones del usuario. La Figura 2 es un diagrama de flujo que ilustra un aspecto de un proceso 200 para comparar una nueva secuencia de nucleótidos de proteínas con una base de datos de secuencias para determinar los niveles de homología entre la nueva secuencia y las secuencias en la base de datos. La base de datos de secuencias puede ser una base de datos privada almacenará dentro del sistema informático 100, o una base de datos pública tal como GENBANK que está disponible a través de Internet. El proceso 200 comienza en un estado de partida 201 y a continuación se mueve hasta un estado 202 en el que la nueva secuencia a comparar se almacena en una memoria en sistema informático 100. Como se ha analizado anteriormente, la memoria podría ser cualquier tipo de memoria, incluyendo RAM o un dispositivo de almacenamiento interno. El proceso 200 a continuación se mueve hasta un estado 204 en el que se abre una base de datos de secuencias para análisis y comparación. A continuación, el proceso 200 se mueve hasta un estado 206 en el que la primera secuencia almacenada en la base de datos se lee en una memoria en el

ordenador. A continuación se realiza una comparación en un estado 210 para determinar si la primera secuencia es la misma que la segunda secuencia. Es importante indicar que esta etapa no se limita a la realización de una comparación exacta entre la nueva secuencia y la primera secuencia en la base de datos. Los expertos en la materia conocen algunos métodos bien conocidos para comparar dos secuencias de nucleótidos o de proteínas, incluso si no son idénticas. Por ejemplo, se pueden introducir huecos en una secuencia para aumentar el nivel de homología entre las dos secuencias sometidas a ensayo. Los parámetros que controlan si los huecos de otras características se introducen en una secuencia durante la comparación se introducen de forma normal por el usuario del sistema informático. Una vez que se ha realizado una comparación de las dos secuencias en el estado 210, se realizó una determinación en el estado de decisión 210 si las dos secuencias son las mismas. Por supuesto, el término "mismas" no se limita a secuencias que son absolutamente idénticas. Algunas secuencias que están dentro de los parámetros de homología introducidos por el usuario se marcarán como "mismas" en el proceso 200. Si se realiza una determinación de que dos secuencias son las mismas, el proceso 200 se mueve a un estado 214 en el que se presenta al usuario el nombre de la secuencia de la base de datos. Este estado notifica al usuario que la secuencia con el nombre presentado satisface las restricciones de homología que se introdujeron. Una vez que se presenta al usuario el nombre de la secuencia almacenada, el proceso 200 se mueve hasta un estado de decisión 218 en el que se realiza 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 final 220. Sin embargo, si existen más secuencias en la base de datos, entonces el proceso 200 se mueve hasta un estado 224 en el que un puntero se mueve hasta la siguiente secuencia en la base de datos de modo que se puede comparar con la nueva secuencia. Este modo, la nueva secuencia se alinea y se compara con cada secuencia de la base de datos. Se debería indicar que si se había realizado una determinación en el estado de decisión 212 de que las secuencias verán homólogas, entonces el proceso 200 se movería inmediatamente al estado de decisión 218 para determinar si cualquier otra secuencia estaba disponible en la base de datos para comparación. Un sistema informático puede comprender un procesador, un dispositivo de almacenamiento de datos que tiene almacenados en el mismo secuencias de ácidos nucleicos de la invención y un comparador de secuencias para realizar la comparación. El comparador de secuencias puede indicar un nivel de homología entre las secuencias comparadas o identificar motivos estructurales, o puede identificar motivos estructurales en secuencias que se comparan con estos códigos de ácidos nucleicos y códigos de polipéptidos. La Figura 3 es un diagrama de flujo que ilustra una realización de un proceso 250 en un ordenador para determinar si las secuencias homólogas. El proceso 250 comienza en un estado de inicio 252 y a continuación se mueve hasta un estado 254 en el que se almacena una primera secuencia a comparar con una memoria. La segunda secuencia a compararse almacenada continuación en una memoria en un estado 256. El proceso 250 se mueve a continuación hasta un estado 260 en el que el primer carácter en la primera secuencia se lee y a continuación hasta un estado 262 en el que se lee el primer carácter de la segunda secuencia. Se debería entender que si la secuencia es una secuencia de nucleótidos, entonces el carácter sería normalmente cualquiera de A, T, C, G o U. Si la secuencia es una secuencia de proteínas, entonces puede ser código de aminoácidos de una sola letra de modo que la primera y segunda secuencias se pueden comparar fácilmente. A continuación se realiza una determinación en un estado de decisión 264 sobre si los dos caracteres son los mismos. Si son los mismos, entonces el proceso 250 se mueve hasta un estado 268 en el que se leen los siguientes caracteres en la primera y segunda secuencias. A continuación se realizó la determinación de si los siguientes caracteres son los mismos. Si lo son, entonces el proceso 250 continúa este bucle hasta que dos caracteres no son los mismos. Si se realiza una determinación de que los siguientes dos caracteres no son los mismos, el proceso 250 se mueve hasta un estado de decisión 274 para determinar si existe algún carácter más en cualquier secuencia paralela. Si no hay ningún carácter más para leer, entonces el proceso 250 se mueve hasta un estado 276 en el que el nivel de homología entre la primera y segunda secuencias se presenta al usuario. El nivel de homología se determina calculando la proporción de caracteres entre las secuencias que eran las mismas sobre el número total de secuencias en la primera secuencia. Por lo tanto, si cada carácter en la primera secuencia de 100 nucleótidos se alinea con cada carácter en una segunda secuencia, el nivel de homología sería de un 100 %.

Como alternativa, el programa informático puede comparar una secuencia de referencia con una secuencia de la invención para determinar si las secuencias difieren en una o más posiciones. Del programa puede registrar la longitud de identidad de los restos de nucleótidos o aminoácidos insertados, suprimidos o sustituidos con respecto a la secuencia de cualquiera de la referencia o la invención. El programa informático puede ser un programa que determinan si una secuencia de referencia contiene polimorfismo de un solo nucleótido (SNP) con respecto a una secuencia de la invención, o, si una secuencia de la invención comprende un SNP de una secuencia conocida. Por lo tanto, en algunos aspectos, el programa informático es un programa que identifica los SNP. El método se puede implementar mediante los sistemas informáticos que se han descrito anteriormente y el método ilustrado en la Figura 3. El método se puede realizar mediante la lectura de una secuencia de la invención y las secuencias de referencia a través del uso del programa informático y la identificación de diferencias con el programa informático.

El sistema basado en ordenador puede comprender un identificador para identificar características dentro de un ácido nucleico o polipéptido de la invención. Un "identificador" se refiere a uno o más programas que identifican ciertas características dentro de secuencias de ácidos nucleicos. Por ejemplo, un identificado puede comprender un programa que identifica un marco de lectura abierto (ORF) en secuencias de ácidos nucleicos. La Figura 4 es un diagrama de flujo que ilustra un aspecto de un proceso identificador 300 para detectar la presencia de una característica en una secuencia. El proceso 300 comienza en un estado de partida 302 y a continuación se mueve hasta un estado 304 en el que una primera secuencia que se va a comprobar para características se almacena en la

memoria 115 en el sistema informático 100. El proceso 300 a continuación se mueve hasta un estado 306 en el que se abre una base de datos de características de secuencia. Tal base de datos incluiría una lista de cada uno de los atributos de las características un todo con el nombre de la característica. Por ejemplo, un nombre de característica podría ser "Codón de Inicio" y el atributo sería "ATG". Otro ejemplo sería el nombre de la característica "Caja TAATAA" y el atributo de la característica sería "TAATAA". Un ejemplo de tal base de datos se produce en el Grupo de Informática Genética de la Universidad de Wisconsin. Como alternativa, las características pueden ser motivos de polipéptidos estructurales tales como hélices alfa, láminas beta, o motivos de polipéptidos funcionales tales como sitios activos enzimáticos, motivos de hélice-giro-hélice u otros motivos conocidos por los expertos en la materia. Una vez que la base de datos de características se abre en el estado 306, el proceso 300 se mueve hasta un estado 308 en el que la primera característica se lee desde la base de datos. A continuación se realizó una comparación del atributo de la primera característica con la primera secuencia en un estado 310. A continuación se realiza una determinación en un estado de decisión 316 sobre si el atributo de la característica se encontró en la primera secuencia. Si el atributo se encontró, entonces el proceso 300 se mueve hasta un estado 318 en el que en nombre de la sacristía encontrada se presenta al usuario. El proceso 300 se mueve a continuación hasta un estado de decisión 320 en el que se realizó una determinación de si existen características que se mueven en la base de datos. Si no existen más características, entonces el proceso 300 termina en un estado final 324. Sin embargo, si existen más características en la base de datos, entonces el proceso 300 leen las siguientes características de la secuencia en un estado 326 y hace un bucle de nuevo hasta el estado 310 en el que el atributo de la siguiente característica se compara frente a la primera secuencia. Si el atributo de la característica no se encuentra en la primera secuencia en el estado de decisión 316, el proceso 300 se mueve directamente al estado de decisión 320 para determinar si existe cualquier característica más en la base de datos. Por lo tanto, en un aspecto, la invención proporciona un programa informático identifica marcos de lectura abiertos (ORF).

Una secuencia de polipéptidos o de ácidos nucleicos de la invención se pueden almacenar y manipular en una diversidad de programas procesadores de datos en una diversidad de formatos. Por ejemplo, una secuencia se puede almacenar como texto en un archivo de procesamiento de palabras, tal como MicrosoftWORD o WORDPERFECT o como un archivo ASCII en una diversidad de programas de bases de datos familiares para los expertos en la materia, tales como DB2, SYBASE, u ORACLE. Además, se pueden usar muchos programas informáticos y bases de datos como algoritmos de comparación de secuencias, identificadores, o fuentes de secuencias de nucleótidos o secuencias de polipéptidos de referencia para su comparación con las secuencias de ácidos nucleicos de la invención. Los programas y las bases de datos usados para poner en práctica la invención incluyen, pero no se limitan a: 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 y 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 DBAccess (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.), QuanteMM, (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 de MDL Available Chemicals Directory, la base de datos de MDL Drug Data Report, la base de datos de Comprehensive Medicinal Chemistry, la base de datos del Índice de Fármacos Mundiales de Derwent, la base de datos de BioByteMasterFile, la base de datos de Genbank, y la base de datos de the Genseqn. Otros muchos programas y bases de datos serían evidentes para un experto en la materia dada la presente divulgación.

Algunos motivos que se pueden detectar usando los programas mencionados anteriormente incluyen secuencias que codifican cremalleras de leucina, motivos de hélice-giro-hélice, sitios de glicosilación, sitios de ubiquitinación, hélices alfa, y láminas beta, secuencias señalan que codifican péptidos señal que dirigen la secreción de las proteínas codificadas, secuencias implicadas en la regulación de la transcripción tales como homeocajas, tramos ácidos, sitios activos enzimáticos, sitios de unión a sustrato, sitios de escisión enzimática.

Hibridación de ácidos nucleicos

La invención proporciona ácidos nucleicos aislados o recombinantes que se inhibían en condiciones rigurosas con una secuencia de la invención o un ácido nucleico que codifica un polipéptido de la invención.

Las condiciones rigurosas pueden ser condiciones altamente rigurosas, condiciones de rigurosidad media y/o condiciones de rigurosidad baja, que incluyen las condiciones de rigurosidad elevada y reducida que se describen en el presente documento. Puede ser la rigurosidad de las condiciones de lavado lo que establece las condiciones que determinan si un ácido nucleico está dentro del alcance de la invención, como se analiza a continuación.

Los ácidos nucleicos como se definen por su capacidad para hibridarse en condiciones de rigurosidad pueden tener una longitud de aproximadamente cinco restos y la longitud completa del ácido nucleico de la invención; por ejemplo, pueden tener una longitud de al menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 55, 60, 65, 70, 75, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, o más restos. También se incluyen

ácidos nucleicos más cortos que los de la longitud completa. Estos ácidos nucleicos pueden ser útiles, por ejemplo, como sondas de hibridación, sondas de marcado, sondas de oligonucleótidos para PCR, ARNi, antisentido o secuencias que codifican péptidos de unión a anticuerpo (epítomos), motivos, sitios activos y similares.

5 Los ácidos nucleico se pueden definir por su capacidad para hibridarse en condiciones que comprenden rigurosidad elevada de formamida aproximadamente al 50 % a aproximadamente de 37 °C a 42 °C, o mediante su capacidad para hibridarse en condiciones que comprenden rigurosidad reducida de formamida de aproximadamente un 35 % a un 25 % a aproximadamente de 30 °C a 35 °C.

10 Como alternativa, los ácidos nucleicos se pueden definir por su capacidad para hibridarse en condiciones que comprenden rigurosidad elevada a 42 °C en formamida al 50 %, 5X SSPE, SDS al 0,3 %, y una secuencia repetitiva que bloquea ácido nucleico, tal como ADN de cot-1 o de esperma de salmón (por ejemplo, 200 n/ml de ADN de esperma de salmón cizallado y desnaturalizado), o mediante su capacidad para hibridarse en condiciones de rigurosidad reducida que comprenden formamida al 35 % a una temperatura reducida de 35 °C.

15 Después de la hibridación con el filtro se puede lavar con 6X SSC, SDS al 0,5 % a 50 °C. Se considera que estas condiciones son condiciones "moderadas" de formamida por encima de un 25 % y condiciones "bajas" de formamida por debajo de un 25 %. Un ejemplo específico de condiciones de hibridación "moderadas" es cuando la hibridación mencionada anteriormente se realiza en formamida al 30 %. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación anterior se realiza en formamida al 10 %.

20 El intervalo de temperatura que corresponde a un nivel de rigurosidad en particular se puede estrechar adicionalmente mediante el cálculo de la proporción de purina a pirimidina del ácido nucleico de interés y ajustando la temperatura en consecuencia. Los ácidos nucleicos de la invención también se definen por su capacidad para hibridarse en condiciones de rigurosidad elevada, media, y baja como se establece en Ausubel y Sambrook. En la técnica se conocen bien algunas variaciones de los intervalos y condiciones mencionados anteriormente. A continuación se analizan adicionalmente condiciones de hibridación.

25 El procedimiento anterior se puede modificar para identificar ácidos nucleicos que tienen disminución de los niveles de homología con la secuencia de la sonda. Por ejemplo, para obtener ácidos nucleicos con disminución de la homología con la sonda detectable, se pueden usar condiciones menos rigurosas. Por ejemplo, la temperatura de hibridación puede disminuir en incrementos de 5 °C de 68 °C a 42 °C en un tampón de hibridación que tiene una concentración de Na⁺ de aproximadamente 1 M. Después de la hibridación, el filtro se puede lavar con 2X SSC, SDS al 0,5 % a la temperatura de hibridación. Se considera que estas condiciones son condiciones "moderadas" por encima de 50 °C y condiciones "bajas" por debajo de 50 °C. Un ejemplo específico de condiciones de hibridación de "rigurosidad moderada" es cuando la hibridación mencionada anteriormente se realiza a 55 °C. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación mencionada anteriormente se realiza a 45 °C.

30 Como alternativa, la hibridación se puede realizar en tampones, tales como 6X SSC, que contienen formamida a una temperatura de 42 °C. En este caso, la concentración de formamida en el tampón de hibridación se puede reducir en incrementos de un 5 % desde un 50 % a un 0 % para identificar planes que tienen disminución de los niveles de homología con la sonda. Después de la hibridación, el filtro se puede lavar con 6X SSC, SDS al 0,5 % a 50 °C. Se considera que estas condiciones son condiciones "moderadas" por encima de un 25 % de formamida y condiciones "bajas" por debajo de un 25 % de formamida. Un ejemplo específico de condiciones de hibridación de "rigurosidad moderada" es cuando la hibridación mencionada anteriormente se realiza en formamida al 30 %. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación mencionada anteriormente se realiza en formamida al 10 %.

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

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

65 Sondas de oligonucleótidos y métodos para su uso

La invención también proporciona sondas de ácido nucleico que se pueden usar, por ejemplo, para identificar a filosóficos que codifican un polipéptido con una actividad de pectato liasa o fragmentos del mismo o para identificar genes de pectato liasa. La sonda comprende al menos de 60 a 100 o de 60 a 150 bases consecutivas de una secuencia como se establece en un ácido nucleico de la invención. Las sondas identifican un ácido nucleico mediante unión y/o hibridación. Las sondas se pueden usar en matrices, véase el análisis que sigue a continuación, que incluyen, por ejemplo, matrices capilares. Las sondas de la invención también se pueden usar para aislar otros ácidos nucleicos o polipéptidos.

Las sondas de la invención se pueden usar para determinar si una muestra biológica, tal como una muestra de suelo, contiene un organismo que tiene secuencia de ácidos nucleicos de la invención o un organismo a partir del que se obtuvo el ácido nucleico. En tales procedimientos, se obtiene una muestra biológica que alberga potencialmente al organismo a partir del que se aisló el ácido nucleico y los ácidos nucleicos se obtienen a partir de la muestra. Los ácidos nucleicos se ponen en contacto con la sonda en condiciones que permiten que la sonda se y hibride de forma específica con cualquier secuencia complementaria presente en la muestra. Cuando sea necesario, las condiciones que permiten que la sonda se hibride de forma específica con secuencias complementarias se puede determinar colocando la sonda en contacto con secuencias complementarias a partir de muestras conocidas porque contienen la secuencia complementaria, así como secuencias de control que no contienen la secuencia complementaria. Algunas condiciones de hibridación, tales como la concentración de sal del tampón de hibridación, la concentración de formamida del tampón de hibridación, o la temperatura de hibridación, se pueden variar a identificar condiciones que permitan que la sonda se hibride de forma específica con ácidos nucleicos complementarios (véase el análisis sobre condiciones de hibridación específicas).

Si la muestra contiene el organismo a partir del que se aisló el ácido nucleico, entonces se detecta una hibridación específica de la sonda. La hibridación se puede detectar marcando 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. Muchos métodos para el uso de la sondas marcadas para detectar la presencia de ácidos nucleicos complementarios en una muestra son familiares para los expertos en la materia. Éstos incluyen Transferencias de Southern, Transferencias de Northern, procedimientos de hibridación de colonias, y transferencias puntuales. Algunos protocolos para cada uno de estos procedimientos se proporcionan en Ausubel y Sambrook.

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

Algunas sondas derivadas de secuencias cerca de los extremos en las posiciones 3' o 5' de las secuencias de ácidos nucleicos de la invención también se pueden usar en procedimientos de avance de cromosomas para identificar clones que contienen secuencias genómicas, por ejemplo, adicionales. Tales métodos permiten el aislamiento de genes que codifican proteínas de interés adicionales del organismo hospedador.

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

En las reacciones de hibridación de ácidos nucleicos, las condiciones usadas para conseguir un nivel de rigurosidad en particular pueden variar, dependiendo de la naturaleza de los ácidos nucleicos que se están irrigando. Por ejemplo, la longitud, grado de complementariedad, composición de la secuencia de nucleótidos (por ejemplo, contenido de GC con respecto al de AT), y tipo de ácido nucleico (por ejemplo, ARN con respecto a ADN) de las regiones de hibridación de los ácidos nucleicos se pueden considerar en la selección de las condiciones de hibridación. Una consideración adicional es si uno de los ácidos nucleicos está inmovilizado, por ejemplo, en un filtro. La hibridación se puede realizar en condiciones de rigurosidad baja, rigurosidad moderada o rigurosidad elevada. Como un ejemplo de hibridación de ácidos nucleicos, en primer lugar una membrana de polímero que contiene ácidos nucleicos desnaturalizados inmovilizados se hibrida previamente durante 30 minutos a 45 °C en una solución que consiste en NaCl 0,9 M, NaH₂PO₄ 50 mM, pH 7,0, Na₂EDTA 5,0 mM, SDS al 0,5 %, 10X de solución de

Denhardt, y 0,5 mg/ml de ácido poliriboadenílico. A continuación se puede añadir a la solución aproximadamente 2×10^7 cpm (actividad específica de $4-9 \times 10^8$ cpm/ug) de sonda de oligonucleótido marcado en su extremo con ^{32}P . Después de 12-16 horas de incubación, la membrana se lava durante 30 minutos a temperatura ambiente (TA) en 1X SET (NaCl 150 mM, clorhidrato de Tris 20 mM, pH 7,8 Na_2EDTA , 1 mM) que contiene SDS al 0,5 %, seguido de un lavado de 30 minutos en 1X SET recién preparado a $T_f - 10^\circ\text{C}$ para la sonda de oligonucleótido. A continuación, la membrana se expone a una película auto-radiográfica para la detección de señales de hibridación.

Variando la rigurosidad de las condiciones de hibridación usadas para identificar ácidos nucleicos, tales como ADNc o ADN genómico, que se hibridan con la sonda detectable, se pueden identificar y aislar ácidos nucleicos que tienen diferentes niveles de homología con la sonda. La rigurosidad se puede variar mediante la realización de la hibridación a temperaturas variables por debajo de las temperaturas de fusión de las sondas. La temperatura de fusión, T_f , es la temperatura (bajo una fuerza iónica y pH definidos) a la que un 50 % de la secuencia diana se hibrida con una sonda perfectamente complementaria. Las condiciones muy rigurosas se seleccionan para que sean iguales o a aproximadamente 5°C menores que la T_f para una sonda en particular. La temperatura de fusión de la sonda se puede calcular usando las siguientes fórmulas a modo de ejemplo. Para las sondas entre 14 y 70 nucleótidos de longitud, la temperatura de fusión (T_f) se calcula usando la fórmula: $T_f = 81,5 + 16,6(\log [\text{Na}^+]) + 0,41(\text{fracción de G+C}) - (600/N)$ en la que N es la longitud de la sonda. Si la hibridación se realiza en una solución que contiene formamida, la temperatura de fusión se cree calcular usando la ecuación: $T_m = 81,5 + 16,6(\log [\text{Na}^+]) + 0,41(\text{fracción de G+C}) - (\text{formamida al } 0,63 \%) - (600/N)$ en la que N es la longitud de la sonda. La hibridación previa se puede realizar en 6X SSC, 5X reactivo de Denhardt, SDS al 0,5 %, 100 μg de ADN de esperma de salmón fragmentado y desnaturalizado o 6X SSC, 5X reactivo de Denhardt, SDS al 0,5 %, 100 μg de ADN de esperma de salmón fragmentado y desnaturalizado, formamida al 50 %. Algunas fórmulas para SSC y reactivo de Denhardt y otras soluciones se enumeran, por ejemplo, en Sambrook.

La hibridación se realiza mediante la adición de la sonda detectable a las soluciones de hibridación previa enumeradas anteriormente. Cuando la sonda comprende ADN bicatenario, se desnaturaliza antes de la adición de la solución de hibridación. El filtro se pone en contacto con la solución de hibridación durante un periodo de tiempo suficiente para permitir que la sonda se hibride con los ADNc o ADN genómicos que contienen secuencias complementarias con los mismos u homólogas a los mismos. Para sondas con una longitud superior a 200 nucleótidos, la hibridación se puede realizar a $15-25^\circ\text{C}$ por debajo de la T_f . Para sondas más cortas, tales como sondas de oligonucleótidos, la hibridación se puede realizar a $5-10^\circ\text{C}$ por debajo de la T_f . En un aspecto, las hibridaciones en 6X SSC se realizan a aproximadamente 68°C . En un aspecto, las hibridaciones en soluciones que contienen formamida al 50 % se realizan a aproximadamente 42°C . Se consideraría que todas las hibridaciones mencionadas anteriormente están en condiciones de rigurosidad elevada.

Después de la hibridación, el filtro se lava para retirar cualquier sonda detectable unidad de forma no específica. La rigurosidad usada para lavar los filtros también puede variar dependiendo de la naturaleza de los ácidos nucleicos que se están hibridando, la longitud de los ácidos nucleicos que se están hibridando, el grado de complementariedad, la composición de la secuencia de nucleótidos (por ejemplo, contenido de GC con respecto al de AT), y el tipo de ácido nucleico (por ejemplo, ARN con respecto al ADN). Algunos ejemplos de lavados en condiciones de rigurosidad progresivamente más elevada son los que siguen a continuación: 2X SSC, SDS al 0,1 % a temperatura ambiente durante 15 minutos (rigurosidad baja); 0,1X SSC, SDS al 0,5 % a temperatura ambiente de 30 minutos a 1 hora (rigurosidad moderada); 0,1X SSC, SDS al 0,5 % de 15 a 30 minutos entre la temperatura de hibridación y 68°C (rigurosidad elevada); y NaCl 0,15 M de 15 minutos a 72°C (rigurosidad ni elevada). Se puede realizar un lavado final de rigurosidad baja en 0,1X SSC a temperatura ambiente. Los ejemplos mencionados anteriormente son simplemente ilustrativos de un conjunto de condiciones que se pueden usar para lavar los filtros. Un experto en la materia conocería que existen numerosas fórmulas para diferentes lavados con rigurosidad.

Los ácidos nucleicos que se han hibridado con la sonda se pueden identificar mediante autorradiografía u otras técnicas convencionales. El procedimiento mencionado anteriormente se puede modificar para identificar ácidos nucleicos que tienen una disminución de los niveles de homología con la secuencia de la sonda. Por ejemplo, para obtener ácidos nucleicos de disminución de la homología con la sonda detectable, se pueden usar condiciones menos rigurosas. Por ejemplo, la temperatura de hibridación puede disminuir en incrementos de 5°C de 68°C a 42°C en un tampón de hibridación que tiene una concentración de Na^+ de aproximadamente 1 M. Después de la hibridación, el filtro se puede lavar con 2X SSC, SDS al 0,5 % a la temperatura de hibridación. Se considera que estas condiciones son condiciones "moderadas" por encima de 50°C y condiciones "bajas" por debajo de 50°C . Un ejemplo de condiciones de hibridación "de rigurosidad moderada" es cuando la hibridación mencionada anteriormente se realiza a 55°C . Un ejemplo de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación mencionada anteriormente se realiza a 45°C .

Como alternativa, la hibridación se puede realizar en tampones, tales como 6X SSC, que contiene formamida a una temperatura de 42°C . En este caso, la concentración de formamida en el tampón de hibridación se puede reducir en incrementos de un 5 % desde un 50 % a un 0 % para identificar clones que tienen una disminución de los niveles de homología con la sonda. Después de la hibridación, el filtro se puede lavar con 6X SSC, SDS al 0,5 % a 50°C . Se considera que estas condiciones son condiciones "moderadas" por encima de un 25 % de formamida y condiciones "bajas" por debajo de un 25 % de formamida. Un ejemplo específico de condiciones de hibridación "moderada" es

cuando la hibridación mencionada anteriormente se realiza en un 30 % de formamida. Un ejemplo específico de condiciones de hibridación de "rigurosidad baja" es cuando la hibridación mencionada anteriormente se realiza en un 10 % de formamida.

Estas sondas y métodos descritos se pueden usar para aislar ácidos nucleicos que tienen una secuencia con una homología de al menos aproximadamente un 99 %, 98 %, 97 %, al menos 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 las secuencias de ácidos nucleicos de la invención que comprenden al menos aproximadamente 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, 150, 200, 250, 300, 350, 400, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, o más bases consecutivas de los mismos, y las secuencias complementarias a las mismas. La homología se puede medir usando un algoritmo de alineación, como se analiza en el presente documento. Por ejemplo, los polinucleótidos homólogos pueden tener una secuencia de codificación que es una variante alélica de origen natural de una de las secuencias de codificación que se describen en el presente documento. Tales variantes alélicas pueden tener una sustitución, supresión o adición de uno o más nucleótidos cuando se comparan con un ácido nucleico de la invención.

Además, las sondas y métodos se pueden usar para aislar ácidos nucleicos que codifican polipéptidos que tienen una identidad de secuencia (homología) de al menos aproximadamente un 99 %, al menos 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 de la invención que comprende al menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 75, 100, o 150 aminoácidos consecutivos, como se determina usando un algoritmo de alineación de secuencias (por ejemplo, tal como el algoritmo de la versión 3.0t78 de FASTA con los parámetros por defecto, o un programa BLAST 2.2.2 con ajustes a modo de ejemplo como se establece en el presente documento).

Inhibición de la Expresión de Pectato liasa

La invención proporciona ácidos nucleicos complementarios con (por ejemplo, secuencias antisentido para) las secuencias de ácidos nucleicos de la invención. Las secuencias antisentido son capaces de inhibir el transporte, corte y empalme o transcripción de genes que codifican pectato liasa. La inhibición se puede realizar a través de la dirección del ADN genómico o del ARN mensajero. La transcripción o función del ácido nucleico al que se dirige se puede inhibir, por ejemplo, mediante hibridación y/o escisión. Un conjunto de inhibidores particularmente útil proporcionado incluye oligonucleótidos que son capaces de unirse a genes de pectato liasa o de realizar mensajes, en cualquier caso previniendo o inhibiendo la producción o la función de pectato liasa. La asociación se puede realizar a través de hibridación específica de secuencias. Otra clase de inhibidores útiles incluye oligonucleótidos que causan la inactivación o escisión de mensajes de pectato liasa. El oligonucleótido puede tener una actividad enzimática que causa tal escisión, tal como ribozimas. El oligonucleótido se puede modificar químicamente o conjugar con una enzima o composición capaz de escindir el ácido nucleico complementario. Una combinación de muchos de tales oligonucleótidos diferentes se puede identificar sistemáticamente para aquellos que tienen la actividad deseada. Por lo tanto, la invención proporciona diversas composiciones para inhibición de la expresión de pectato liasa en un ácido nucleico y/o nivel de proteína, por ejemplo, antisentido, ARNi, y ribozimas que comprenden secuencias de pectato liasa de la invención y los anticuerpos anti-pectato liasa que se describen en el presente documento.

La inhibición de la expresión de pectato liasa puede tener una diversidad de aplicaciones industriales. Por ejemplo, la inhibición de la expresión de pectato liasa puede ralentizar o prevenir el deterioro por "putrefacción blanda". El deterioro por "putrefacción blanca" se produce cuando la pectina, un polisacárido estructural principal en la pared de célula vegetal, se degrada de forma enzimática. Esto puede conducir al deterioro, o putrefacción, de frutas y verduras. En un aspecto, el uso de composiciones de la invención que inhiben la expresión y/o actividad de pectato liasas, por ejemplo, anticuerpos, oligonucleótidos antisentido, ribozimas y ARNi, se usan para ralentizar o prevenir el deterioro por "putrefacción blanda". Por lo tanto, algunos métodos y composiciones comprenden la aplicación sobre una planta o producto vegetal (por ejemplo, una fruta, semilla, raíz, hoja, etc.) de anticuerpos, oligonucleótidos antisentido, ribozimas y ARNi de la divulgación para ralentizar o prevenir la descomposición por "putrefacción blanda". Estas composiciones también se pueden expresar por la planta (por ejemplo, una planta transgénica) u otro organismo (por ejemplo, una bacteria u otro microorganismo transformado con un gen de pectato liasa de la invención).

La inhibición de la expresión de pectato también puede prevenir o ralentizar el crecimiento normal del agente patógeno *Erysiphe cichoracearum* del oídio. Esta resistencia al oídio representa una forma de resistencia a enfermedades basada en la pérdida de un gen necesario durante una interacción compatible en lugar de la activación de las rutas de defensa del hospedador conocidas. Véase, por ejemplo, Vogel (2002) Plant Cell 14: 2095-2106. Por lo tanto, algunos métodos y composiciones comprenden la aplicación en una planta o producto vegetal (por ejemplo, una fruta, semilla, raíz, hoja, etc.) de anticuerpos, oligonucleótidos antisentido, ribozimas y ARNi de la divulgación para ralentizar o prevenir el crecimiento del patógeno del oídio.

Oligonucleótidos Antisentido

La divulgación proporciona oligonucleótidos antisentido capaces de unirse a mensajes de pectato liasa que pueden inhibir la actividad política mediante la dirección del ARNm. Algunas estrategias para diseñar oligonucleótidos antisentido se describen bien en la bibliografía científica y patente, y el experto en la materia puede diseñar tales oligonucleótidos de pectato liasa usando los nuevos reactivos de la invención. Por ejemplo, en la técnica se conocen bien protocolos de avance genético / formación de mapas de ARN para identificar sistemáticamente oligonucleótidos antisentido eficaces, véase, por ejemplo, Ho (2000) *Methods Enzymol.* 314: 168-183, que describe un ensayo de formación de mapas de ARN, que se basa en técnicas moleculares convencionales para proporcionar un método fácil y de confianza para la selección de secuencias antisentido potentes. Véase también Smith (2000) *Eur. J. Pharm. Sci.* 11: 191-198.

Los ácidos nucleicos de origen natural se usan como oligonucleótidos antisentido. Dos oligonucleótidos antisentido pueden tener cualquier longitud; por ejemplo, en aspectos alternativos, los oligonucleótidos antisentido tienen entre aproximadamente 5 y 100, de aproximadamente 10 a 80, de aproximadamente 15 a 60, de aproximadamente 18 a 40. La longitud óptima se puede determinar mediante identificación sistemática de rutina. Los oligonucleótidos antisentido pueden estar presentes en cualquier concentración. La concentración óptima se puede determinar mediante identificación sistemática de rutina. Se conoce una amplia diversidad de nucleótidos de origen natural, sintéticos y análogos de ácidos nucleicos que pueden dirigirse a este problema potencial. Por ejemplo, se pueden usar estructuras principales no iónicas que contienen ácidos nucleicos peptídicos (PNA), tales como unidades de N-(2-aminoetil) glicina. También se pueden usar oligonucleótidos antisentido que tienen enlaces de fosforotioato, como se describe en el documento WO 97/03211; documento WO 96/39154; Mata (1997) *Toxicol Appl Pharmacol* 144: 189-197; *Antisense Therapeutics*, ed. Agrawal (Humana Press, Totowa, N.J., 1996). Algunos oligonucleótidos antisentido que tienen análogos de estructura principal de ADN sintético proporcionados por la invención también pueden incluir ácidos nucleicos con fosforo-ditioato, metilfosfonato, fosforamidato, fosfotriéster de alquilo, sulfamato, 3'-tioacetal, metilen(metilimino), 3'-N-carbamato, y carbamato de morfolino, como se ha descrito anteriormente.

Se puede usar metodología de química combinatoria para crear grandes números de oligonucleótidos que se pueden identificar sistemáticamente de forma rápida para oligonucleótidos específicos que tienen afinidades de unión y especificidades apropiadas hacia cualquier diana, tal como las secuencias sentido y antisentido de pectato liasa de la invención (véase, por ejemplo, Gold (1995) *J. of Biol. Chem.* 270: 13581-13584).

Ribozimas Inhibitorias

La divulgación proporciona ribozimas capaces de unirse a mensaje de pectato liasa. Estas ribozimas pueden inhibir la actividad de pectato liasa, por ejemplo, mediante dirección de ARNm. Algunas estrategias para diseñar ribozimas y seleccionar la secuencia antisentido específica de pectato para la dirección se describen bien en la bibliografía científica y de patente, y el experto en la materia puede diseñar tales ribozimas usando los nuevos reactivos de la invención. Algunas ribozimas actúan mediante la unión a un ARN diana a través de la parte de unión del ARN diana de una ribozima que se mantiene en una proximidad cercana a una parte enzimática del ARNcr escinde el ARN diana. Por lo tanto, la ribozima reconoce y se une a un ARN diana a través de emparejamiento de bases complementarias, y una vez unido el sitio correcto, actúa de forma enzimática para escindir e inactivar el ARN diana. La escisión de un ARN diana de tal modo destruirá su capacidad para dirigir la síntesis de una proteína codificada si la escisión se produce en la secuencia de codificación. Después de que una ribozima se ha unido y escindido de su diana de ARN, se puede liberar de ese ARN para unir y escindir nuevas dianas repetidamente.

En algunas circunstancias, la naturaleza enzimática de una ribozima puede ser ventajosa sobre otras tecnologías, tales como tecnología antisentido (cuando una molécula de ácido nucleico se une simplemente a una diana de ácido nucleico para bloquear su transcripción, traducción o asociación con otra molécula) ya que la concentración eficaz de ribozima necesaria para realizar un tratamiento terapéutico puede ser inferior a la de un oligonucleótido antisentido. Esta ventaja potencial refleja la capacidad de la ribozima para actuar de forma enzimática. Por lo tanto, una sola molécula de ribozima es capaz de escribir muchas moléculas de ARN diana. Además, una ribozima es por lo general un inhibidor altamente específico, con la especificidad e inician dependiendo no solamente del mecanismo de unión de emparejamiento de bases, sino también del mecanismo mediante el que la molécula inhibe la expresión del ARN al que se une. Es decir, la inhibición está causada por la escisión de la diana de ARN y por lo tanto la especificidad se define como la proporción de la tasa de escisión del ARN dirigido sobre la tasa de escisión del ARN no dirigido. Este mecanismo de escisiones depende de factores adicionales a los implicados en el emparejamiento de bases. Por lo tanto, la especificidad de acción de una ribozima puede ser superior a la de la unión de oligonucleótido antisentido del mismo sitio de ARN.

La ribozima, por ejemplo, una molécula de ARN de ribozima enzimática, se puede formar en un motivo de cabeza de martillo, un motivo de horquilla, como un motivo del virus de la hepatitis delta, un motivo de intrón del grupo I y/o un ARN de tipo RNasaP en asociación con una secuencia guía de ARN. Algunos ejemplos de motivos de cabeza de martillo se describen, por ejemplo, en Rossi (1992) *Aids Research and Human Retroviruses* 8: 183; motivos de horquilla en Hampel (1989) *Biochemistry* 28: 4929, y Hampel (1990) *Nuc. Acids Res.* 18: 299; el motivo del virus de la hepatitis delta en Perrotta (1992) *Biochemistry* 31: 16; el motivo de RNasaP en Guerrier-Takada (1983) *Cell* 35: 849; y el intrón del grupo I en la Patente de Estados Unidos N° 4.987.071 de Cech. La mención de estos motivos

específicos no pretende ser limitante. Los expertos en la materia reconocerán que una ribozima, por ejemplo, una molécula de ARN enzimático puede tener un sitio de unión a sustrato específico complementario con una o más que en las regiones del ARN del gen diana. Una ribozima de la invención puede tener una secuencia de nucleótidos dentro o rodeando ese sitio de unión a sustrato que transmite una actividad de escisión del ADN a la molécula.

ARN de interferencia (ARNi)

La divulgación proporciona una molécula inhibitoria de ARN, una molécula denominada "ARNi", que comprende una secuencia de pectato liasa de la invención. La molécula de ARNi comprende una molécula de ARN bicatenario (ARNds). El ARNi puede inhibir la expresión de un gen de pectato liasa. En un aspecto, el ARNi tiene una longitud de aproximadamente 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más nucleótidos dúplex. Aún que no está limitado mediante ningún mecanismo de acción en particular, el ARNi puede entrar en una célula provocar la degradación de un ARN monocatenario (ARNss) de secuencia similares o idénticas, que incluyen ARNm endógenos. Cuando una célula se expone a un ARN bicatenario (ARNds), el ARNm del gen homólogo se degrada de forma selectiva mediante un proceso denominado interferencia de ARN (ARNi). Un posible mecanismo básico más allá del ARNi es la ruptura de un emparejamiento de ARN bicatenario (ARNds) de una secuencia de genes específicos en piezas cortas denominadas ARNr interferencia corto, que desencadena la degradación del ARNm que empareja su secuencia. En un aspecto, los ARNi se usan en terapia de silenciamiento genético, véase, por ejemplo, Shuey (2002) Drug Discov. Today 7: 1040-1046. Algunos métodos pueden degradar el ARNr de forma selectiva usando los ARNi. El proceso se puede poner en práctica *in vitro*, *ex vivo* o *in vivo*. Las moléculas de ARNi se pueden usar para generar una mutación de pérdida de función en una célula, un órgano o un animal. En la técnica se conoce bien métodos para preparar y usar moléculas de ARNi para degradar el ARN de forma selectiva, véase, por ejemplo, los documentos de patente de Estados Unidos Nº 6.506.559; 6.511.824; 6.515.109; 6.489.127.

Modificación de Ácidos Nucleicos

Algunos métodos pueden generar variantes de los ácidos nucleicos de la invención, las variantes de ácidos nucleicos generados con estos métodos (por ejemplo, SEC ID Nº: 133) y a continuación se analizan polipéptidos codificados por ellos (por ejemplo, SEC ID Nº: 134). Estos métodos se pueden repetir o usar en diversas combinaciones para generar pectato liasas que tienen una actividad alterada o diferente o una estabilidad alterada o diferente de la de una pectato liasa codificada por el molde de ácido nucleico. Estos métodos también se pueden repetir o usar en diversas combinaciones, por ejemplo, para generar variaciones en expresión genética/mensajera, traducción de mensajes o estabilidad del mensaje. En otro aspecto, la composición genética de una célula se vio alterada, por ejemplo, por la modificación de un gen homólogo *in vitro*, *in vivo* o *ex vivo*, seguido de su reinserción en la célula.

Un ácido nucleico de la invención se puede alterar mediante cualquier medio. Por ejemplo, métodos aleatorios o estocásticos, o métodos no estocásticos, o métodos de "evolución dirigida", véase, por ejemplo, el documento de Patente de Estados Unidos Nº 6.361.974. En la técnica se conocen bien de todos para mutación aleatoria de genes, véase, por ejemplo, el documento de Patente de Estados Unidos Nº 5.830.696. Por ejemplo, se pueden usar algunos agentes mutágenos para mutar de forma aleatoria un gen. Algunos agentes mutágenos incluyen, por ejemplo, luz ultravioleta o radiación gamma, o un agente mutágeno químico, por ejemplo, mitomicina, ácido nitroso, psoralenos fotoactivados, solos o en combinación, para inducir rupturas de ADN susceptibles de reparación por recombinación. Otros agentes mutágenos químicos incluyen, por ejemplo, bisulfito sódico, ácido nitroso, hidroxilamina, hidrazina o ácido fórmico. Otros agentes mutágenos son análogos de precursores de nucleótidos, por ejemplo, nitrosoguanidina, 5-bromouracilo, 2-aminopurina, o acridina. Estos agentes se pueden añadir a una reacción de PCR en lugar del precursor de nucleótidos, mutando la secuencia de este modo. Se pueden usar agentes de intercalado tales como proflavina, acriflavina, quinacrina y similares.

Se puede usar cualquier técnica de biología molecular, por ejemplo, mutagénesis de PCR aleatoria, véase, por ejemplo, Rice (1992) Proc. Natl. Acad. Sci. USA 89: 5467-5471; o, mutagénesis de casete múltiple combinatorio, véase, por ejemplo, Cramer (1995) Biotechniques 18: 194-196. Como alternativa, algunos ácidos nucleicos, por ejemplo, genes, se pueden reaensamblar después de fragmentación aleatoria, o "estocástica", véanse, por ejemplo, los documentos de la patente de Estados Unidos Nº 6.291.242; 6.287.862; 6.287.861; 5.955.358; 5.830.721; 5.824.514; 5.811.238; 5.605.793. En aspectos alternativos, se introducen modificaciones, adiciones o supresiones por PCR propensa a error, redistribución, mutagénesis dirigida a oligonucleótidos, PCR de ensamblaje, mutagénesis de PCR sexual, mutagénesis *in vivo*, mutagénesis de casete, mutagénesis de conjunto recursiva, mutagénesis de conjunto exponencial, mutagénesis específica de sitio, reaensamblaje genético, mutagénesis por saturación de sitio genético (GSSM™), reaensamblaje de ligamiento genético (SLR), recombinación, recombinación de secuencias recursiva, mutagénesis de ADN modificada con fosfotioato, mutagénesis de molde que contiene uracilo, mutagénesis de dúplex con huecos, mutagénesis de reparación de falta de coincidencia de puntos, mutagénesis de cepa hospedadora con déficit de reparación, mutagénesis química, mutagénesis radiogénica, mutagénesis de supresión, mutagénesis de restricción-selección, mutagénesis de restricción-purificación, síntesis de genes artificiales, mutagénesis de conjunto, creación de multímeros de ácidos nucleicos quiméricos, y/o una combinación de estos y otros métodos.

Las siguientes publicaciones describen una diversidad de procedimientos de recombinación y/o métodos recursivos que se pueden incorporar en los métodos de la invención: Stemmer (1999) "Molecular breeding of viruses for targeting and other clinical properties" *Tumo Targeting* 4: 1-4; Ness (1999) *Nature Biotechnology* 17: 893-896; Chang (1999) "Evolution of a cytokine using DNA family shuffling" *Nature Biotechnology* 17: 793-797; Minshull (1999) "Protein evolution by molecular breeding" *Current Opinion in Chemical Biology* 3: 284-290; Christians (1999) "Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling" *Nature Biotechnology* 17: 259-264; Cramer (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" *Nature* 391: 288-291; Cramer (1997) "Molecular evolution of an arsenate detoxification pathway by DNA shuffling", *Nature Biotechnology* 15: 436-438; Zhang (1997) "Directed evolution of an effective fucosidase from a galactosidase by DNA shuffling and screening" *Proc. Natl. Acad. Sci. USA* 94: 4504-4509; Patten *et al.* (1997) "Applications of DNA Shuffling to Pharmaceuticals and Vaccines" *Current Opinion in Biotechnology* 8: 724-733; Cramer *et al.* (1996) "Construction and evolution of antibody-phage libraries by DNA shuffling" *Nature Medicine* 2: 100-103; Gates *et al.* (1996) "Affinity selective isolation of ligands from peptide libraries through display on a lac repressor 'headpiece dimer'" *Journal of Molecular Biology* 255: 373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" In: *The Encyclopedia of Molecular Biology*. VCH Publishers, New York. pp. 447-457; Cramer y Stemmer (1995) "Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wildtype cassettes" *BioTechniques* 18: 194-195; Stemmer *et al.* (1995) "Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides" *Gene*, 164: 49-53; Stemmer (1995) "The Evolution of Molecular Computation" *Science* 270: 1510; Stemmer (1995) "Searching Sequence Space" *Bio/Technology* 13: 549-553; Stemmer (1994) "Rapid evolution of a protein in vitro by DNA shuffling" *Nature* 370: 389-391; y Stemmer (1994) "DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution". *Proc. Natl. Acad. Sci. USA* 91: 10747-10751.

Algunos métodos de mutación para generar diversidad incluyen, por ejemplo, mutagénesis dirigida al sitio (Ling *et al.* (1997) "Approaches to DNA mutagenesis: an overview" *Anal Biochem.* 254 (2): 157-178; Dale *et al.* (1996) "Oligonucleotide-directed random mutagenesis using the phosphorothioate method" *Methods Mol. Biol.* 57: 369-374; Smith (1985) "In vitro mutagenesis" *Ann. Rev. Genet.* 19: 423-462; Botstein y Shortle (1985) "Strategies and applications of in vitro mutagenesis" *Science* 229: 1193-1201; Carter (1986) "Site-directed mutagenesis" *Biochem. J.* 237:1-7; y Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" en *Nucleic Acids & Molecular Biology* (Eckstein, F. y Lilley, D. M. J. eds., Springer Verlag, Berlín)); mutagénesis usando moldes que contienen uracilo (Kunkel (1985) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Proc. Natl. Acad. Sci. USA* 82: 488-492; Kunkel *et al.* (1987) "Rapid and efficient site-specific mutagenesis without phenotypic selection" *Methods in Enzymol.* 154, 367-382; y Bass *et al.* (1988) "Mutant Trp repressors with new DNA-binding specificities" *Science* 242: 240-245); mutagénesis dirigida a oligonucleótidos (Methods in Enzymol. 100: 468-500 (1983); Methods in Enzymol. 154: 329-350 (1987); Zoller y Smith (1982) "Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment" *Nucleic Acids Res.* 10: 6487-6500; Zoller y Smith (1983) "Oligonucleotide-directed mutagenesis of DNA fragments cloned in M13 vectors" *Methods in Enzymol.* 100: 468-500; y Zoller y Smith (1987) "Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template" *Methods in Enzymol.* 154: 329-350); mutagénesis de ADN modificada con fosforotioato (Taylor *et al.* (1985) "The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA" *Nucl. Acids Res.* 13: 8749-8764; Taylor *et al.* (1985) "The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA" *Nucl. Acids Res.* 13: 8765-8787 (1985); Nakamaye (1986) "Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 14: 9679-9698; Sayers *et al.* (1988) "Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis" *Nucl. Acids Res.* 16: 791-802; y Sayers *et al.* (1988) "Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide" *Nucl. Acids Res.* 16: 803-814); mutagénesis usando ADN dúplex con huecos (Kramer *et al.* (1984) "The gapped duplex DNA approach to oligonucleotide-directed mutation construction" *Nucl. Acids Res.* 12: 9441-9456; Kramer y Fritz (1987) *Methods in Enzymol.* "Oligonucleotide-directed construction of mutations via gapped duplex DNA" 154: 350-367; Kramer *et al.* (1988) "Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations" *Nucl. Acids Res.* 16: 7207; y Fritz *et al.* (1988) "Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro" *Nucl. Acids Res.* 16: 6987-6999).

Algunos protocolos adicionales que se pueden usar para poner en práctica la invención incluyen reparación de falta de coincidencia puntual (Kramer (1984) "Point Mismatch Repair" *Cell* 38: 879-887), mutagénesis usando cepas hospedadoras con deficiencia de reparación (Carter *et al.* (1985) "Improved oligonucleotide site-directed mutagenesis using M13 vectors" *Nucl. Acids Res.* 13: 4431-4443; y Carter (1987) "Improved oligonucleotide-directed mutagenesis using M13 vectors" *Methods in Enzymol.* 154: 382-403), mutagénesis de supresión (Eghtedarzadeh (1986) "Use of oligonucleotides to generate large deletions" *Nucl. Acids Res.* 14: 5115), restricción-selección y restricción-selección y restricción-purificación (Wells *et al.* (1986) "Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin" *Phil. Trans. R. Soc. Lond. A* 317: 415-423), mutagénesis mediante síntesis genética total (Nambiar *et al.* (1984) "Total synthesis and cloning of a gene coding for the ribonuclease S protein" *Science* 223: 1299-1301; Sakamar y Khorana (1988) "Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin)" *Nucl. Acids Res.* 14: 6361-6372; Wells *et*

al. (1985) "Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites" *Gene* 34: 315-323; y Grundstrom *et al.* (1985) "Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis" *Nucl. Acids Res.* 13: 3305-3316), reparación de la ruptura de la doble hebra (Mandecki (1986); Arnold (1993) "Protein engineering for unusual environments" *Current Opinion in Biotechnology* 4: 450-455, "Oligonucleotide-directed double-strand break repair in plasmids of *Escherichia coli*: a method for site-specific mutagenesis" *Proc. Natl. Acad. Sci. USA*, 83: 7177-7181). Detalles adicionales de muchos de los métodos mencionados anteriormente se pueden encontrar en el Volumen 154 de *Methods in Enzymology*, que también describe controles útiles para problemas de resolución de problemas con diversos métodos de mutagénesis.

Algunos protocolos adicionales que se pueden usar para poner en práctica la invención se describen, por ejemplo, en el documento de Patente de Estados Unidos Nº 5.605.793 de Stemmer (25 de febrero de 1997), "Methods for In Vitro Recombination"; documento de Patente de Estados Unidos Nº 5.811.238 de Stemmer *et al.* (22 de septiembre de 1998) "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; documento de Patente de Estados Unidos Nº 5.830.721 de Stemmer *et al.* (3 de noviembre de 1998), "DNA Mutagenesis by Random Fragmentation and Reassembly"; documento de Patente de Estados Unidos Nº 5.834.252 de Stemmer, *et al.* (10 de noviembre de 1998) "End-Complementary Polymerase Reaction"; documento de Patente de Estados Unidos Nº 5.837.458 de Minshull, *et al.* (17 de noviembre de 1998), "Methods and Compositions for Cellular and Metabolic Engineering"; documento de patente WO 95/22625, Stemmer y Crameri, "Mutagenesis by Random Fragmentation and Reassembly"; documento de patente WO 96/33207 de Stemmer y Lipschutz "End Complementary Polymerase Chain Reaction"; documento de patente WO 97/20078 de Stemmer y Crameri "Methods for Generating Polynucleotides having Desired Characteristics by Iterative Selection and Recombination"; documento de patente WO 97/35966 de Minshull y Stemmer, "Methods and Compositions for Cellular and Metabolic Engineering"; documento de patente WO 99/41402 de Punnonen *et al.* "Targeting of Genetic Vaccine Vectors"; documento de patente WO 99/41383 de Punnonen *et al.* "Antigen Library Immunization"; documento de patente WO 99/41369 de Punnonen *et al.* "Genetic Vaccine Vector Engineering"; documento de patente WO 99/41368 de Punnonen *et al.* "Optimization of Immunomodulatory Properties of Genetic Vaccines"; documento de patente EP 752008 de Stemmer y Crameri, "DNA Mutagenesis by Random Fragmentation and Reassembly"; documento de patente EP 0932670 de Stemmer "Evolving Cellular DNA Uptake by Recursive Sequence Recombination"; documento de patente WO 99/23107 de Stemmer *et al.*, "Modification of Virus Tropism and Host Range by Viral Genome Shuffling"; documento de patente WO 99/21979 de Apt *et al.*, "Human Papillomavirus Vectors"; documento de patente WO 98/31837 de del Cardayre *et al.* "Evolution of Whole Cells and Organisms by Recursive Sequence Recombination"; documento de patente WO 98/27230 de Patten y Stemmer, "Methods and Compositions for Polypeptide Engineering"; documento de patente WO 98/27230 de Stemmer *et al.*, "Methods for Optimization of Gene Therapy by Recursive Sequence Shuffling and Selection", documento de patente WO 00/00632, "Methods for Generating Highly Diverse Libraries", documento de patente WO 00/09679, "Methods for Obtaining in Vitro Recombined Polynucleotide Sequence Banks and Resulting Sequences", documento de patente WO 98/42832 de Arnold *et al.*, "Recombination of Polynucleotide Sequences Using Random or Defined Primers", documento de patente WO 99/29902 de Arnold *et al.*, "Method for Creating Polynucleotide and Polypeptide Sequences", documento del patente WO 98/41653 de Vind, "An in Vitro Method for Construction of a DNA Library", documento de patente WO 98/41622 de Borchert *et al.*, "Method for Constructing a Library Using DNA Shuffling", y documento de patente WO 98/42727 de Pati y Zarlign, "Sequence Alterations using Homologous Recombination."

Algunos protocolos adicionales que se pueden usar para poner en práctica la invención (que proporcionan detalles con respecto a diversos métodos de generación de diversidad) se describen, por ejemplo, en la solicitud de Patente de Estados Unidos con Nº de serie. (USSN) 09/407,800, "SHUFFLING OF CODON ALTERED GENES" de Patten *et al.* presentada el 28 de septiembre de 1999; "EVOLUTION OF WHOLE CELLS AND ORGANISMS BY RECURSIVE SEQUENCE RECOMBINATION" de del Cardayre *et al.*, documento de patente de Estados Unidos Nº 6.379.964; "OLIGONUCLEOTIDE MEDIATED NUCLEIC ACID RECOMBINATION" de Crameri *et al.*, documentos de patente de Estados Unidos Nºs 6.319.714; 6.368.861; 6.376.246; 6.423.542; 6.426.224 y documento PCT/US00/01203; "USE OF CODON-VARIED OLIGONUCLEOTIDE SYNTHESIS FOR SYNTHETIC SHUFFLING" de Welch *et al.*, documento de patente de Estados Unidos Nº 6.436.675; "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" de Selifonov *et al.*, presentado el 18 de enero de 2000, (PCT/US00/01202) y, por ejemplo "METHODS FOR MAKING CHARACTER STRINGS, POLYNUCLEOTIDES & POLYPEPTIDES HAVING DESIRED CHARACTERISTICS" de Selifonov *et al.*, presentado el 18 de julio de 2000 (documento de Estados Unidos con Nº de Serie 09/618.579); "METHODS OF POPULATING DATA STRUCTURES FOR USE IN EVOLUTIONARY SIMULATIONS" de Selifonov y Stemmer, presentado el 18 de enero de 2000 (documento PCT/US00101138); y "SINGLE-STRANDED NUCLEIC ACID TEMPLATE-MEDIATED RECOMBINATION AND NUCLEIC ACID FRAGMENT ISOLATION" de Affholter, presentado el 6 de septiembre de 2000 (documento de Estados Unidos con Nº de Serie 09/656.549); y documentos de patente de Estados Unidos Nºs 6.177.263; 6.153.410.

Algunos métodos no estocásticos, o de "evolución dirigida", incluyen, por ejemplo, mutagénesis por saturación (GSSM™), reensamblaje de ligamiento genético (SLR), o una combinación de los mismos se usan para modificar los ácidos nucleicos de la invención para generar peptato líasas con propiedades nuevas o alteradas (por ejemplo, actividad en condiciones altamente ácidas o alcalinas, temperaturas elevadas, y similares). Algunos polipéptidos codificados por los ácidos nucleicos modificados se pueden identificar sistemáticamente para una actividad antes de

someter a ensayo la actividad proteolítica u otra actividad. Se puede usar cualquier modalidad protocolo de ensayo, por ejemplo, usando una plataforma de matriz capilar. Véanse, por ejemplo, los documentos de patente de Estados Unidos N^{os} 6.361.974; 6.280.926; 5.939.250.

5 *Mutagénesis por saturación, o, GSSMTM*

En un aspecto, se usan cebadores de codón que contienen una secuencia de N,N,G/T degenerada para introducir mutaciones puntuales en un polinucleótido, por ejemplo, una pectato liasa o un anticuerpo de la invención, con el fin de generar un conjunto de polipéptidos de la progenie en los que se representa una gama total de sustituciones de aminoácidos individuales en cada posición del aminoácido, por ejemplo, un resto de aminoácido en un sitio activo enzimático o unión de ligandos dirigida al sitio a modificar. Estos oligonucleótidos pueden comprender una primera secuencia homóloga contigua, una secuencia de N,N,G/T degenerada, y, ocasionalmente, una segunda secuencia homóloga. Los productos de traducción de la progenie corriente abajo a partir del uso de tales oligonucleótido se incluyen todos los posibles cambios de aminoácidos en cada sitio de aminoácidos a lo largo del polipéptido, porque la degeneración de la secuencia de N,N,G/T incluye codones para los 20 aminoácidos. En un aspecto, uno de tales oligonucleótidos degenerados (que consiste por ejemplo en un casete de N,N,G/T degenerado) se usa para someter cada codón original en un molde de polinucleótido precursor a una gama completa de sustituciones de codón. En otro aspecto, se usan al menos dos casetes degenerados – en el mismo oligonucleótido o no, para someter al menos a dos codones originales en un molde de polinucleótido que puso a una gama completa de sustituciones de codón. Por ejemplo, más de una secuencia de N,N,G/T puede estar contenida en un oligonucleótido para introducir mutaciones de aminoácidos en más de un sitio. Esta pluralidad de secuencias de N,N,G/T puede ser directamente contigua, o puede estar separada por una o más secuencias de nucleótidos adicionales. En otro aspecto, se pueden usar oligonucleótidos que sirven para introducir adiciones y supresiones solos o en combinación con los codones que contiene una secuencia de N,N,G/T, para introducir cualquier combinación de fermentación de adiciones, supresiones, y/o sustituciones de aminoácidos.

En un aspecto, se realiza la mutagénesis simultánea de dos o más posiciones de aminoácidos contiguas usando un oligonucleótido que contiene tripletes de N,N,G/T contiguos, es decir una secuencia (N,N,G/T)_n degenerada. En otro aspecto, se usan casetes degenerados que tienen menos de generación que la secuencia de N,N,G/T. Por ejemplo, en algunos casos puede ser deseable usar (por ejemplo, en un oligonucleótido) una secuencia de triplete degenerado que consiste solamente en un N, en el que dicho N puede estar en la primera, segunda o tercera posición del triplete. Se puede usar cualquier otra base que incluye cualquier combinación y permutación de la misma en las dos posiciones restantes del triplete. Como alternativa, en algunos casos puede ser deseable usar (por ejemplo, en un oligo) una secuencia de triplete de N,N,N degenerado.

En un aspecto, el uso de tripletes degenerados (por ejemplo, tripletes de N,N,G/T) permite una generación sistemática y fácil de una gama total de posibles aminoácidos naturales (para un total de 20 aminoácidos) en todas y cada una de las posiciones de los aminoácidos en un polipéptido (en aspectos alternativos, los métodos también incluyen generación de menos de todas las posibles sustituciones por resto de aminoácido, o codón, posición). Por ejemplo, para un polipéptido de 100 aminoácidos, se pueden generar 2000 especies distintas (es decir 20 aminoácidos posibles por posición X 100 posiciones de aminoácidos). A través del el uso de un oligonucleótido o conjunto de oligonucleótidos que contienen un triplete degenerado N,N,G/T, 32 secuencias individuales pueden codificar los 20 aminoácidos naturales posibles. Por lo tanto, en un recipiente de reacción en el que una secuencia de polinucleótidos precursora se somete a mutagénesis por saturación usando al menos uno de tales oligonucleótidos, se generan 32 polinucleótidos de progenie distinta que codifican 20 polipéptidos distintos. Por el contrario, el uso de un oligonucleótido no degenerado en mutagénesis dirigida al sitio conduce solamente a un producto de polipéptido de la progenie por recipiente de reacción. Los oligonucleótidos no degenerados se pueden usar opcionalmente junto con los cebadores degenerados que se desvelan; por ejemplo, los oligonucleótidos no degenerados se pueden usar oligonucleótidos para generar mutaciones puntuales específicas en un polinucleótido de trabajo. Esto proporciona un medio para generar mutaciones puntuales silenciosas específicas, mutaciones puntuales que conducen a los correspondientes cambios de aminoácidos, y mutaciones puntuales que producen la generación de codones de parada y la correspondiente expresión de fragmentos de polipéptidos.

En un aspecto, cada recipiente de reacción de mutagénesis por saturación contiene polinucleótidos que codifican al menos 20 moléculas de polipéptidos de la progenie (por ejemplo, pectato liasas) de modo que los 20 aminoácidos naturales se representan en una posición específica del aminoácido que corresponde a la posición del codón sometido a mutagénesis en el polinucleótido precursor (otros aspectos usan menos de las 20 combinaciones naturales). Los polipéptidos de la progenie degenerada de 32 veces generados a partir de cada recipiente de reacción de mutagénesis por saturación se pueden someter la amplificación clonal (por ejemplo, se pueden clonar en un hospedador adecuado, por ejemplo, hospedador de *E. coli*, usando, por ejemplo, un vector de expresión) y se pueden someter a identificación sistemática de expresión. Cuando un polipéptido de la progenie individual se identifica mediante identificación sistemática para represente un cambio favorable en las propiedades (cuando se compara con el polipéptido precursor, tal, aumento de la actividad proteolítica en condiciones alcalinos o ácidas), se puede secuenciar para que identifique la sustitución de aminoácido correspondientemente favorable contenida en el mismo.

En un aspecto, después de la mutagénesis todas y cada una de las posiciones del aminoácido en un polipéptido precursor usando mutagénesis por saturación, se desvelan el presente documento, se puede identificar cambios favorables de aminoácidos en una o más posiciones de la misma. Se puede generar una o más moléculas de nueva progenie que contienen una combinación de todas o parte de estas instituciones de aminoácidos favorables. Por ejemplo, si 2 cambios de aminoácidos favorables específicos identifican en cada una de las posiciones de 3 aminoácidos en un polipéptido, las permutaciones incluyen 3 posibilidades en cada posición (sin cambio del aminoácido glicina, y cada uno de dos cambios favorables) y 3 posiciones. Por lo tanto, existen $3 \times 3 \times 3$ o 27 posibilidades totales, incluyendo 7 que se examinaron anteriormente - 6 mutaciones puntuales individuales (es decir 2 en cada una de las tres posiciones) y ningún cambio en posición alguna.

En otro aspecto, se puede usar mutagénesis de saturación de sitio junto con otros medios estocásticos o no estocásticos para variar la secuencia, por ejemplo, reensamblaje de ligamiento genético (véase a continuación), redistribución, quimerización, recombinación y otros procesos de mutagénesis y agentes de mutagénesis. La presente invención proporciona el uso de cualquier proceso o procesos de mutagénesis, incluyendo mutagénesis por saturación, de una manera repetitiva.

Reensamblaje de Ligamiento Sintética (SLR)

Un sistema de modificación genética no estocástico denominado "reensamblaje de ligamiento genético", o simplemente "SLR", un "proceso de evolución dirigida", puede generar polipéptidos, por ejemplo, pectato liasas o anticuerpos, con propiedades nuevas o alteradas. El SLR es un método de ligamiento de fragmentos de oligonucleótidos en conjunto de forma no estocástica. Este método se diferencia de la redistribución de oligonucleótidos estocástica en que los componentes básicos del ácido nucleico no se redistribuyen, concatenan o quimerizan de forma aleatoria, sino que en su lugar se ensamblan de forma no estocástica. Véase, por ejemplo, la solicitud de Patente de Estados Unidos con N° de Serie (USSN) 09/332.835 con el título "Reensamblaje de Ligamiento Sintético en Evolución Dirigida" y presentada el 14 de junio de 1999 ("USSN 09/332.835"). En un aspecto, el SLR comprende las siguientes etapas: (a) proporcionar un polinucleótido molde, en el que el polinucleótido molde comprende secuencia que codifica un gen homólogo; (b) proporcionar una pluralidad de polinucleótidos como componentes básicos, en el que los polinucleótidos como componentes básicos se diseñan para reensamblaje cruzado con el polinucleótido molde en una secuencia determinada previamente, y un polinucleótido como componente básico comprende una secuencia que es una variante del gen homólogo y una secuencia homóloga con el polinucleótido molde que franquea secuencia variable; (c) combinar un polinucleótido como componente básico con un polinucleótido molde de modo que el cruce del polinucleótido como componente básico se reensambla con el polinucleótido molde para generar polinucleótidos que comprenden variaciones de secuencias de genes homólogos.

El SLR no depende de la presencia de niveles elevados de homología entre polinucleótidos a redistribuir. Por lo tanto, este método se puede usar para bibliotecas (o conjuntos) generadas de forma no estocástica de moléculas de la progenie que consisten en aproximadamente 10100 quimeras diferentes. El SLR se puede usar para generar bibliotecas que comprenden aproximadamente 101000 quimeras de la progenie diferentes. Por lo tanto, los métodos no estocásticos pueden producir un conjunto de moléculas de ácidos nucleicos quiméricas finalizadas que tienen un orden de ensamblaje global que se eligen mediante el diseño. Este método incluye las etapas de generar mediante diseño una pluralidad de componentes básicos de ácidos nucleicos específicos que tienen extremos que se pueden ligar mutuamente compatibles útiles, y ensamblar estos componentes básicos del ácido nucleico, de manera que se consigue una orden de montaje global diseñada.

Se considera que los extremos que se pueden ligar mutuamente compatibles de los componentes básicos del ácido nucleico a ensamblar son "útiles" para este tipo de conjunto ordenado si permiten que los componentes básicos se acoplen en órdenes predeterminadas. Por lo tanto, el orden general de ensamblaje en el que los componentes básicos del ácido nucleico se pueden acoplar se especifica mediante el diseño de los extremos que se pueden ligar. Si se va a usar más de una etapa de ensamblaje, entonces el orden general de ensamblaje en el que se pueden acoplar los componentes básicos del ácido nucleico también se especifica mediante el orden secuencial de la etapa o etapas de ensamblaje. En un aspecto, las piezas de componentes básicos hibridados se tratan con una enzima, tal como una ligasa (por ejemplo, T4 ADN ligasa), para conseguir la unión covalente de las piezas de los componentes básicos.

En un aspecto, el diseño de los componentes básicos de los oligonucleótidos se obtiene mediante el análisis de un conjunto de moldes de secuencias de ácidos nucleicos precursores que sirven como base para la producción de un conjunto de la progenie de polinucleótidos quiméricos finalizados. Por lo tanto, estos moldes de oligonucleótidos precursores sirven como una fuente de información de la secuencia que ayuda en el diseño de los componentes básicos del ácido nucleico que se van a someter a mutagénesis, por ejemplo, quimerizar o redistribuir. En un aspecto de este método, las secuencias de una pluralidad de moldes de ácido nucleico precursor se alinean con el fin de seleccionar uno o más puntos de demarcación. Los puntos de demarcación pueden estar situados en una zona de homología, y constante uno o más nucleótidos. Estos puntos de demarcación se comparten preferentemente por al menos dos de los moldes precursores. Por lo tanto, los puntos de demarcación se pueden usar para delinear los límites de los componentes básicos de los oligonucleótidos que se generan con el fin de re

ordenar los polinucleótidos precursores. Los puntos de demarcación identificados y seleccionados en las moléculas precursoras sirven como puntos potenciales de quimerización en el ensamblaje de las moléculas finales de la progenie quiméricas. Un punto de demarcación puede ser una zona de homología (que consiste en al menos una base nucleotídica homóloga) compartida por al menos dos secuencias de polinucleótidos precursores. Como alternativa, un punto de demarcación puede ser una zona de homología que se comparte por al menos la mitad de las secuencias de polinucleótidos precursores, o, puede ser una zona de homología que se comparte por al menos dos tercios de las secuencias de polinucleótidos precursores. Incluso más preferentemente un punto de demarcación útil es una zona de homología que se comparte por al menos tres cuartas partes de las secuencias de polinucleótidos precursores, o, se puede compartir por casi todas las secuencias de polinucleótidos precursores. En un aspecto, un punto de demarcación es una zona de homología que se comparte por todas las secuencias de polinucleótidos precursores.

En un aspecto, un proceso de reensamblaje de ligamiento se realiza de forma exhaustiva con el fin de generar una biblioteca exhaustiva de polinucleótidos quiméricos de la progenie. En otras palabras, todas las combinaciones ordenadas posibles de los componentes básicos del ácido nucleico se representan en el conjunto de moléculas de ácido nucleico quimérico finalizado. Al mismo tiempo, en otro aspecto, el orden de montaje (es decir, el orden de montaje de cada componente básico en la secuencia en las posiciones 5' a 3' de cada ácido nucleico quimérico finalizado) en cada combinación se realiza mediante diseño (o no estocástico) como se ha descrito anteriormente. Debido a la naturaleza no estocástica de esto, la posibilidad de productos secundarios no deseados se reduce en gran medida.

El método reensamblaje de ligamiento se puede realizar de forma sistemática. Por ejemplo, el método se realiza con el fin de generar una biblioteca sistemáticamente compartimentada de moléculas de la progenie, con compartimentos que se pueden seleccionar de forma sistemática, por ejemplo, uno a uno. Esto proporciona que, a través del uso selectivo y con criterio de componentes básicos de ácidos nucleicos específicos, junto con el uso selectivo y con criterio de reacciones de ensamblaje en etapas secuencialmente, se puede conseguir un diseño, en el que algunos conjuntos específicos de productos de la progenie se producen en cada uno de varios recipientes de reacción. Esto permite que se realice un procedimiento de examen sistemático y de identificación sistemática. Por lo tanto, estos métodos permiten el examen sistemático de un número potencialmente muy grande de moléculas de la progenie sistemáticamente en grupos más pequeños. Debido a su capacidad para realizar quimerizaciones de una manera que es altamente flexible pero exhaustiva y sistemática también, en particular cuando hay un bajo nivel de homología entre las moléculas precursoras, estos métodos proporcionan la generación de una biblioteca (o conjunto) formada por un gran número de moléculas de la progenie. Debido a la naturaleza no estocástica del reensamblaje de ligamiento instantánea, las moléculas de la progenie generadas preferentemente comprenden una biblioteca de moléculas de ácido nucleico quimérico finalizado que tienen un orden de ensamblaje global que se eligen mediante el diseño. También se pueden usar métodos de mutagénesis por saturación y de evolución dirigida optimizada para generar diferentes especies moleculares de la progenie. Se observa que esto proporciona una libertad de elección y control con respecto a la selección de puntos de demarcación, el tamaño y número de los componentes básicos del ácido nucleico, y el tamaño y el diseño de los acoplamientos. Se observa, además, que el requisito de homología intermolecular es altamente relajado para la operatividad de este proceso. De hecho, incluso los puntos de demarcación se pueden elegir en zonas con poca o ninguna homología intermolecular. Por ejemplo, a causa de la oscilación del codón, es decir, la degeneración de los codones, se pueden introducir sustituciones de nucleótidos en los componentes básicos del ácido nucleico sin alterar el aminoácido codificado originalmente en el correspondiente molde precursor. Como alternativa, un codón puede ser alterado de modo que se altera la codificación de un ácido amino original. Esto proporciona que tales sustituciones se pueden introducir en el componente básico del ácido nucleico con el fin de aumentar la incidencia de puntos de demarcación homóloga intermoleculares y por lo tanto para permitir un aumento del número de acoplamientos a conseguir entre los componentes básicos, que a su vez permite que se genere un mayor número de moléculas quiméricas de la progenie.

En otro aspecto, la naturaleza sintética de la etapa en la que se generan los componentes básicos permite el diseño e introducción de nucleótidos (por ejemplo, uno o más nucleótidos, que pueden ser, por ejemplo, codones o intrones o secuencias reguladoras) que más tarde se pueden retirar opcionalmente en un proceso *in vitro* (por ejemplo, por mutagénesis) o en un proceso *in vivo* (por ejemplo, mediante el uso de la capacidad de corte y empalme de genes de un organismo hospedador). Se observa que en muchos casos la introducción de estos nucleótidos también puede ser deseable por muchas otras razones además del beneficio potencial de crear un punto de demarcación útil.

En un aspecto, un componente básico de ácido nucleico se usa para introducir un intrón. Por lo tanto, se introducen intrones funcionales en un gen artificial preparado de acuerdo con los métodos que se describen en el presente documento. El intrón o intrones introducidos artificialmente pueden ser funcionales en una célula hospedadora para corte y empalme de genes mucho más en la forma en la que los intrones de origen natural sirven funcionalmente en el corte y empalme de genes.

Sistema Optimizado de Evolución Dirigida

Un sistema de modificación genética no estocástico denominado "sistema de evolución dirigida optimizado" puede generar polipéptidos, por ejemplo, pectato liasas o anticuerpos, con propiedades nuevas o alteradas. La evolución dirigida optimizada se refiere al uso de ciclos repetidos de redistribución reductora, recombinación y selección que permite la evolución molecular dirigida de los ácidos nucleicos a través de recombinación. La evolución dirigida optimizada permite la generación de una gran población de secuencias quiméricas evolucionadas, en la que la población generada se enriquece de forma significativa para las secuencias que tienen un número predeterminado de sucesos de cruce.

Un suceso de cruce es un punto en una secuencia quimérica en el que se produce un cambio en la secuencia a partir de una variante precursora a otra variante precursora. Tal punto se encuentra normalmente en la unión en la que los oligonucleótidos de dos precursores se unen entre sí para formar una sola secuencia. Este método permite el cálculo de las concentraciones correctas de secuencias de oligonucleótidos para que la población quimérica final de las secuencias se vea enriquecida para el número elegido de sucesos de cruce. Esto proporciona un mayor control sobre la elección de variantes quiméricas que tienen un número predeterminado de sucesos de cruce.

Además, este método proporciona un medio conveniente para explorar una enorme cantidad del posible espacio de la variante de proteína en comparación con otros sistemas. Anteriormente, si se generaban, por ejemplo, 10^{13} moléculas quiméricas durante una reacción, sería extremadamente difícil someter a ensayo un gran número de variantes quiméricas para una actividad en particular. Por otra parte, una porción significativa de la población de la progenie tendría un número muy elevado de sucesos de cruce que daría como resultado proteínas que sería menos probable que tuvieron un aumento de los niveles de una actividad en particular. Mediante el uso de estos métodos, la población de moléculas quiméricas se puede enriquecer para aquellas variantes que tienen un número de sucesos de cruce en particular. Por lo tanto, aunque todavía se pueden generar 10^{13} moléculas quiméricas durante una reacción, cada una de las moléculas elegidas para el análisis adicional tiene de forma más probable, por ejemplo, solamente tres sucesos de cruce. Dado que la población de la progenie resultante se puede sesgar para que tenga un número predeterminado de sucesos de cruce, los límites de la variedad funcional entre las moléculas quiméricas se reducen. Esto proporciona un número más manejable de variables cuando se calcula el oligonucleótido a partir de que los polinucleótidos precursores originales podrían ser responsables de influir en un rasgo en particular.

Un método para la creación de una secuencia de polinucleótidos de la progenie quimérica es crear oligonucleótidos que corresponden a fragmentos o partes de cada secuencia precursora. Cada oligonucleótido incluye preferentemente una región única de solapamiento de modo que la mezcla de los oligonucleótidos juntos da lugar a una nueva variante que tiene cada fragmento oligonucleótido ensamblado en el orden correcto. También se puede encontrar información, por ejemplo, en el documento USSN 09/332.835; documento de Patente de Estados Unidos N° 6.361.974.

El número de oligonucleótidos generados para cada variante precursora tiene una relación con el número total de cruces resultantes en la molécula quimérica que se crea en última instancia. Por ejemplo, se podrían proporcionar tres variantes de secuencias de nucleótidos precursoras para someterse a una reacción de ligamiento con el fin de encontrar una variante quimérica que tiene, por ejemplo, mayor actividad a temperatura elevada. Como un ejemplo, se puede generar un conjunto de 50 secuencias de oligonucleótidos que corresponde a cada una de las porciones de cada variante precursora. Por consiguiente, durante el proceso de reensamblaje de ligamiento se podrían producir hasta 50 sucesos de cruce dentro de cada una de las secuencias quiméricas. La probabilidad de que cada uno de los polinucleótidos quiméricos generados contenga oligonucleótidos de cada variante precursora en orden alternativo es muy baja. Si cada fragmento de oligonucleótido está presente en la reacción de ligamiento en la misma cantidad molar es probable que en algunas posiciones los oligonucleótidos del mismo polinucleótido precursor se ligan entre sí y por lo tanto no se producirá como resultado un suceso de cruce. Si la concentración de cada oligonucleótido de cada precursor se mantiene constante durante cualquier etapa de ligamiento en este ejemplo, existe una posibilidad de un 1/3 (suponiendo 3 progenitores) de que un oligonucleótido de la misma variante precursora se ligue dentro de la secuencia quimérica y no se produzca cruce.

Por consiguiente, se puede determinar una función de densidad de probabilidad (PDF) para predecir la población de sucesos de cruce que se pueden producir durante cada etapa de una reacción de ligamiento dado un número determinado de variantes de los precursores, una serie de oligonucleótidos que corresponden a cada variante, y las concentraciones de cada variante durante cada etapa en la reacción de ligamiento. Las estadísticas y las matemáticas detrás de la determinación de la PDF se describen a continuación. Mediante el uso de estos métodos, se puede calcular una función de densidad de probabilidad, y por lo tanto enriquecer la población de la progenie quimérica para un número predeterminado de sucesos de cruce resultantes de una reacción de ligamiento en particular. Por otra parte, se puede determinar previamente un número objetivo de sucesos de cruce, y el sistema programado a continuación para calcular las cantidades de partida de cada oligonucleótido precursor durante cada etapa en la reacción de ligamiento para dar como resultado una función de densidad de probabilidad que se centra en el número predeterminado de sucesos de cruce. Estos métodos se dirigen al uso de ciclos repetidos de redistribución reductora, recombinación y selección que permite la evolución molecular dirigida de un ácido nucleico que codifica un polipéptido a través de recombinación. Este sistema permite la generación de una gran población de secuencias quiméricas evolucionadas, en la que la población generada se enriquece significativamente para las

secuencias que tienen un número predeterminado de sucesos de cruce. Un suceso de cruce es un punto en una secuencia quimérica, en el que se produce un cambio en la secuencia a partir de una variante precursora por otra variante precursora. Tal punto se encuentra normalmente en la unión de oligonucleótidos, en la que dos de los precursores se ligan entre sí para formar una sola secuencia. El método permite el cálculo de las concentraciones correctas de secuencias de oligonucleótidos para que la población quimérica final de las secuencias se enriquezca para el número elegido de sucesos de cruce. Esto proporciona un mayor control sobre la elección de variantes quiméricas que tienen un número predeterminado de sucesos de cruce.

Además, estos métodos proporcionan un medio conveniente para explorar una enorme cantidad de espacio variable posible de la proteína en comparación con otros sistemas. Mediante el uso de los métodos que se describen en el presente documento, la población de moléculas quiméricas se puede enriquecer para las variantes que tienen un número de sucesos de cruce en particular. Por lo tanto, aunque todavía se pueden generar 10^{13} moléculas quiméricas durante una reacción, cada una de las moléculas elegida para el análisis adicional tiene de forma más probable, por ejemplo, solamente tres sucesos de cruce. Dado que la población de la progenie resultante se puede sesgar para que tenga un número predeterminado de sucesos de cruce, los límites de la variedad funcional entre las moléculas quiméricas se reduce. Esto proporciona un número de variables más manejable cuando se calcula el oligonucleótido de los polinucleótidos precursores originales que podría ser responsable de influir en un rasgo en particular.

En un aspecto, el método crea una secuencia de polinucleótidos de la progenie quimérica mediante la creación de oligonucleótidos correspondientes a fragmentos o partes de cada secuencia precursora. Cada oligonucleótido incluye preferentemente una región única de solapamiento de manera que la mezcla de los oligonucleótidos juntos da lugar a una nueva variante que tiene cada fragmento oligonucleótido ensamblado en el orden correcto. Véase también el documento USSN 09/332.835.

Determinación de Sucesos de Cruce

Un sistema y un software pueden recibir una función deseada de densidad de probabilidad de cruce (PDF), el número de genes precursores a reensamblar, y el número de fragmentos en el reensamblaje como entradas. La salida de este programa es una "PDF de fragmento" que se puede usar para determinar una fórmula para producir genes reensamblados, y la PDF de cruce estimada de esos genes. El procesamiento que se describe en el presente documento se realiza preferentemente en MATLAB™ (The Mathworks, Natick, Massachusetts), un lenguaje de programación y en torno del desarrollo para informática técnica.

Procesos Repetitivos

Estos procesos se pueden repetir de forma reiterada. Por ejemplo, un ácido nucleico (o, el ácido nucleico) responsable de un fenotipo de pectato liasa alterado o nuevo se identifica, se vuelve a aislar, se modifica de nuevo, se vuelve a someter a ensayo para su actividad. Este proceso se puede repetir de forma reiterada hasta que se modifica por ingeniería un fenotipo deseado. Por ejemplo, toda una ruta anabólica o catabólica bioquímica se puede modificar por ingeniería en una célula, incluyendo, por ejemplo, actividad de hidrólisis de epóxido.

De forma análoga, si se determina que un oligonucleótido en particular oligonucleotide no han influido en absoluto en el rasgo deseado (por ejemplo, un fenotipo de pectato liasa nuevo), se puede eliminar como una variable mediante la síntesis de oligonucleótidos precursores más grandes que incluyen la secuencia a eliminar. Dado que la incorporación de la secuencia dentro de una secuencia más grande evitar cualquier suceso de cruce, ya no se producirá variación alguna en esta secuencia en los polinucleótidos de la progenie. Esta práctica repetitiva para determinar qué oligonucleótidos están más relacionados con el rasgo deseado y cuales no están relacionados, permite una exploración más eficaz de todas las posibles variantes de proteínas a las que se podría proporcionar un rasgo o actividad en particular.

Redistribución in vivo

La redistribución *in vivo* de moléculas se usa en métodos que proporcionan variantes de polipéptidos de la invención. La redistribución *in vivo* se puede realizar usando la propiedad natural de las células para múltiples recombinantes. Aunque la recombinación *in vivo* ha proporcionado la ruta natural principal para la diversidad molecular, la recombinación genética sigue siendo un proceso relativamente complejo que implica 1) el reconocimiento de homologías; 2) escisión de la hebra, invasión de la hebra, y etapas metabólicas que conducen a la producción de quiasma recombinante; y por último 3) la resolución del quiasma en moléculas de combinadas discretas. La formación del quiasma requiere el reconocimiento de secuencias homólogas.

Un método puede producir un polinucleótido híbrido a partir de al menos un primer polinucleótido (por ejemplo, una pectato liasa de la invención) y un segundo polinucleótido (por ejemplo, una enzima, tal como una pectato liasa de la invención o cualquier otra pectato liasa, o, una marca o un epítipo). Un polinucleótido híbrido se puede producir mediante la introducción de al menos un primer polinucleótido y un segundo polinucleótido que comparten al menos una región de homología de secuencia parcial en una célula hospedadora adecuada. Las regiones de homología de

secuencia parcial estimulan procesos que dan como resultado una reorganización de la secuencia produciendo un polinucleótido híbrido. La expresión "polinucleótido híbrido", como se usa en el presente documento, es partir secuencia de nucleótidos que resulta del método descrito y que contiene la secuencia de al menos dos secuencias de polinucleótidos originales. Tales polinucleótidos híbridos pueden resultar de sucesos de recombinación que estimulan la integración de secuencias entre moléculas de ADN. Además, tales polinucleótidos híbridos pueden resultar de procesos de reorganización reductora intramolecular que usan secuencias repetidas para alterar una secuencia de nucleótidos dentro de una molécula de ADN.

Producción de variantes de secuencias

Existen métodos adicionales para preparar variantes de secuencias de las secuencias de ácidos nucleicos (por ejemplo, pectato liasa) de la invención y métodos adicionales para aislar pectato liasas usando los ácidos nucleicos y polipéptidos de la invención. Algunas variantes de una secuencia que codifica pectato liasa (por ejemplo, un gen, ADNc o mensaje) de la invención se pueden alterar mediante cualquier medio, que incluye, por ejemplo, métodos aleatorios o estocásticos, o, no estocásticos, o métodos de "evolución dirigida", como se ha descrito anteriormente.

Las variantes aisladas pueden ser de origen natural. También se pueden crear variantes *in vitro*. Algunas variantes se pueden crear usando técnicas de ingeniería genética tales como mutagénesis dirigida al sitio, mutagénesis química aleatoria, procedimientos de supresión de Exonucleasa III, y técnicas de clonación convencionales. Como alternativa, tales variantes, fragmentos, análogos, o derivados se pueden crear usando procedimientos de síntesis o de modificación química. Otros métodos para preparar variantes tales o familiares para los expertos en la materia. Estos incluyen procedimientos en los que las secuencias de ácidos nucleicos obtenidas a partir de aislados naturales se modifican para generar ácidos nucleicos que codifican polipéptidos que tienen características que aumentan su valor en aplicaciones industriales o de laboratorio. En tales procedimientos, se genera y se caracteriza un gran número de secuencias de variantes que tienen una o más diferencias en los nucleótidos con respecto a la secuencia obtenida a partir del aislado natural. Estas diferencias de nucleótidos pueden dar como resultado cambios de aminoácidos con respecto a los polipéptidos codificados por los ácidos nucleicos a partir de los aislados naturales.

Por ejemplo, se pueden crear variantes usando PCR propensa a error. En la PCR propensa a error, la PCR se realiza en condiciones en las que la fidelidad de la copia de la ADN polimerasa es baja, de modo que se tiene una tasa elevada de mutaciones puntuales a lo largo de toda la longitud del producto de PCR. La PCR propensa a error se describe, por ejemplo, en Leung, D.W., *et al.*, Technique, 1: 11-15, 1989) y Caldwell, R. C. y Joyce G.F., PCR Methods Applic., 2: 28-33, 1992. En resumen, en tales procedimientos, los ácidos nucleicos a someter a mutagénesis se mezclan con cebadores de PCR, tampón de reacción, $MgCl_2$, $MnCl_2$, Taq polimerasa ay una concentración apropiada de dNTP para conseguir una tasa elevada de mutaciones puntuales a lo largo de toda la longitud del producto de PCR. Por ejemplo, la reacción se puede realizar usando 20 fmoles de ácido nucleico a someter a mutagénesis, 30 pmoles de cada cebador de PCR, un tampón de reacción que consiste en KCl 50 mM, Tris HCl 10 mM (pH 8,3) y gelatina al 0,01 %, $MgCl_2$ 7mM, $MnCl_2$ 0,5mM, 5 unidades de Taq polimerasa, dGTP 0,2 mM, dATP 0,2 mM, dCTP 1 mM, y dTTP 1 mM. La PCR se puede realizar durante 30 ciclos de 94 °C durante 1 min, 45 °C durante 1 min, y 72 °C durante 1 min. Sin embargo, se observará que estos parámetros se pueden variar cuando sea apropiado. Los ácidos nucleicos sometidos a mutagénesis se clonan en un vector apropiado y se evalúan las actividades de los polipéptidos codificados por los ácidos nucleicos sometidos a mutagénesis.

También se pueden crear variantes usando mutagénesis dirigirá a oligonucleótidos para generar mutaciones específicas de sitio en cualquier ADN clonado de interés. La mutagénesis de oligonucleótido se describe, por ejemplo, en Reidhaar-Olson (1988) Science 241: 53-57. En resumen, en tales procedimientos, una pluralidad de oligonucleótidos bicatenarios que portan una o más mutaciones a introducir en el ADN clonado se sintetizan e insertan en el ADN clonado a someter a mutagénesis. Los clones que contienen ADN sometido a mutagénesis se recuperan y se evalúan las actividades de los polipéptidos que codifican.

Otro método para generar variantes es la PCR de ensamblaje. La PCR de ensamblaje implica el ensamblaje de un producto de PCR de una mezcla de fragmentos pequeños de ADN. Un gran número de diferentes reacciones de PCR se producen en paralelo en el mismo vial, con los productos de una reacción que ceba los productos de otra reacción. La PCR de ensamblaje se describe, por ejemplo, en el documento de Patente de Estados Unidos N° 5.965.408.

Además, otro método para generar variantes es la mutagénesis de PCR sexual. En la mutagénesis de PCR sexual, se produce una recombinación homóloga forzada entre moléculas de ADN de secuencias de ADN diferentes pero muy relacionadas *in vitro*, como resultado de la fragmentación aleatoria de la molécula de ADN basándose en homología de secuencias, seguido de fijación del cruce por la extensión del cebador en una reacción de PCR. La mutagénesis de PCR sexual se describe, por ejemplo, en Stemmer (1994) Proc. Natl. Acad. Sci. USA 91: 10747-10751. En resumen, en tales procedimientos, una pluralidad de ácidos nucleicos recombinantes se digieren con DNasa para generar fragmentos que tienen un tamaño medio de 50-200 nucleótidos. Los fragmentos del tamaño medio deseado ser política ni se vuelven a suspender en una mezcla de PCR. La PCR se realiza en condiciones que facilitan la recombinación entre los fragmentos de ácido nucleico. Por ejemplo, la PCR se puede realizar mediante

resuspensión de los fragmentos purificados a una concentración de 10-30 ng/μl en una solución 0,2 mM de cada dNTP, MgCl₂ 2,2 mM, KCl 50 mM, Tris HCl 10 mM, pH 9,0, y Triton X-100 al 0,1 %, 2,5 unidades de Taq polimerasa por 100:1 de mezcla de reacción y se realiza la PCR usando el siguiente régimen: 94 °C durante 60 segundos, 94 °C durante 30 segundos, 50-55 °C durante 30 segundos, 72 °C durante 30 segundos (30-45 veces) y 72 °C durante 5 minutos. Sin embargo, se observará que estos parámetros se deben variar cuando sea apropiado. En algunos aspectos, los oligonucleótidos se pueden incluir en las reacciones de PCR. En otros aspectos como el fragmento de Klenow de la ADN polimerasa I se puede usar en un primer conjunto de reacciones de PCR y la Taq polimerasa se puede usar en un conjunto de reacciones de PCR posteriores. Las secuencias recombinantes se aíslan y las actividades de los polipéptidos que codifican se evalúan.

También se pueden crear variantes mediante mutagénesis *in vivo*. En algunos aspectos, las mutaciones aleatorias en una secuencia de interés se generan mediante la propagación de la secuencia de interés en una cepa bacteriana, tal como una cepa de *E. coli*, que porta mutaciones en una o más de las rutas de reparación de ADN. Tales cepas "mutadoras" tienen una tasa de mutación aleatoria más elevada que la de una precursora de tipo silvestre. La propagación del ADN en una de estas cepas en ocasiones generará mutaciones aleatorias dentro de la de. Las cepas mutadoras adecuadas para usurpar a mutagénesis *in vivo* se describen, por ejemplo, en la Publicación de PCT con N° WO 91/16427.

También se puede generar variantes usando mutagénesis de casete. En la mutagénesis de casete, una pequeña región de una molécula de ADN bicatenario se sustituye con un "casete" de oligonucleótido sintético que difiere de la secuencia narrativa. A menudo, lo oligonucleótido contiene secuencias nativas total y/o parcialmente aleatorias.

También se puede usar mutagénesis de conjunto recursiva para generar variantes. La mutagénesis de conjunto recursiva es un algoritmo para ingeniería de proteínas (mutagénesis de proteínas) desarrollado para producir diversas poblaciones de mutantes relacionados fenotípicamente cuyos miembros se diferencian en la secuencia de aminoácidos. Este método usa un mecanismo de retroalimentación para controlar rondas sucesivas de mutagénesis de casete combinatoria. La mutagénesis de conjunto recursiva se describe, por ejemplo, en Arkin (1992) Proc. Natl. Acad. Sci. USA 89: 7811-7815.

En algunos aspectos, las variantes se crean usando mutagénesis de conjunto exponencial. La mutagénesis de conjunto exponencial es un proceso para generar bibliotecas combinatorias con un porcentaje elevado de mutantes únicos y funcionales, en el que pequeños grupos de restos se clasifican al azar en paralelo para identificar, en cada posición alterada, aminoácidos que conducen a proteínas funcionales. La mutagénesis de conjunto exponencial se describe, por ejemplo, en Delegrave (1993) Biotechnology Res. 11: 1548-1552. La mutagénesis aleatoria y dirigida al sitio se describe, por ejemplo, en Arnold (1993) Current Opinion in Biotechnology 4: 450-455.

En algunos aspectos, las variantes se crean usando procedimientos de redistribución en los que las partes de una pluralidad de los nucleicos que codifican distintos polipéptidos se fusionan entre sí para crear secuencias de ácidos nucleicos quiméricos que codifican polipéptidos quiméricos como se describe, por ejemplo, los documentos de Patente de Estados Unidos N°s 5.965.408; 5.939.250 (véase también el análisis, mencionado anteriormente).

La invención también proporciona variantes de polipéptidos de la invención (por ejemplo, pectato liasas) que comprenden secuencias en las que uno o más de los restos de aminoácido (por ejemplo, de un polipéptido a modo de ejemplo de la invención) se sustituyen con un resto de aminoácido conservado o no conservado (por ejemplo, un resto de aminoácido conservado) y tal resto de aminoácido sustituido puede ser uno codificado o no por el código genético. Las sustituciones conservativas son aquellas que sustituyen un aminoácido largo en un polipéptido por otro aminoácido de características similares. Por lo tanto, los polipéptidos de la invención incluyen aquellos que tienen sustituciones conservativas de secuencias de la invención, por ejemplo, los polipéptidos a modo de ejemplo de la invención, que incluyen, pero no se limitan a, las siguientes sustituciones: sustituciones de un aminoácido alifático tal como Alanina, Valina, Leucina e Isoleucina con otro aminoácido Alifático; sustitución de una Serina con una Treonina o viceversa; sustitución de un resto ácido tal como Ácido Aspártico y Glutámico con otro resto ácido; sustitución de un resto que porta un grupo amida, tal como Asparagina y Glutamina, con otro resto que porta un grupo amida; intercambio de un resto básico tal como Lisina y Arginina con otro resto básico; y sustitución de un resto aromático tal como Fenilalanina, Tirosina con otro resto aromático. Otras variantes son aquellas en las que uno o más de los restos de aminoácidos de los polipéptidos de la invención incluye un grupo sustituye.

Otras variantes son aquellas en las que el polipéptido está asociado con otro compuesto, tal como un compuesto para aumentar la vida media del polipéptido, por ejemplo, polietilenglicol.

Las variantes adicionales en las que los aminoácidos adicionales se fusionan con el polipéptido, tales como una secuencia directora, una secuencia secretora, una secuencia de proproteína o una secuencia que facilita la purificación, enriquecimiento, o estabilización del polipéptido.

En algunos aspectos, las variantes, fragmentos, derivados y análogos de dos polipéptidos de la invención mantienen la misma función o actividad biológica que los polipéptidos a modo de ejemplo, por ejemplo, actividad de pectato liasa, como se describe en el presente documento. En otros aspectos, la variante, fragmento, derivado, o análogo

incluye una proproteína, de modo que la variante, fragmento, derivado, o análogo se puede activar por escisión de la parte de proproteína para producir un polipéptido activo.

Optimización de codones para conseguir niveles elevados de expresión de proteínas en células hospedadoras

Algunos métodos pueden modificar ácidos nucleicos que codifican pectato liasa para modificar el uso del codón, por ejemplo modificación de codones en un ácido nucleico que codifica un pectato liasa para aumentar o disminuir su expresión en una célula hospedadora. El método comprende la identificación de un codón "no preferente" o un codón "menos preferente" en ácido nucleico que codifica pectato liasa y sustitución de uno o más de estos codones no preferentes o menos preferentes con un "codón preferente" que codifica el mismo aminoácido que el del codón sustituido y al menos un codón no preferente o menos preferente en el ácido nucleico se ha sustituido con un codón preferente que codifica al mismo aminoácido. Un codón preferente es un codón que se sobrerrepresenta en secuencias de codificación en genes en la célula hospedadora y un codón no preferente o menos preferente es un codón que se subrepresenta en secuencias de codificación en genes en la célula hospedadora.

Las células hospedadas para expresar los ácidos nucleicos, casetes de expresión y vectores de la invención incluyen bacterias, levadura, hongos, células vegetales, células de insecto y células de mamífero. Se desvelan algunos métodos para la optimización del uso del codón en todas estas células. Algunas células hospedadoras a modo de ejemplo incluyen bacterias gram negativas, tales como *Escherichia coli*; bacterias gram positivas, tales como cualquiera de *Streptomyces* *Lactobacillus gasseri* *Lactococcus lactis* *Lactococcus cremoris*, cualquier *Bacillus*, por ejemplo, *Bacillus subtilis*, *Bacillus cereus*. Algunas células hospedadoras a modo de ejemplo también incluyen organismos eucariotas, por ejemplo, diversas levaduras, tales como *Saccharomyces* sp., que incluyen *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, y *Kluyveromyces lactis*, *Hansenula polymorpha*, *Aspergillus niger*, y células y líneas celulares de mamífero y células y líneas celulares de insecto. Por lo tanto, la divulgación también incluye ácidos nucleicos y polipéptidos optimizados para expresión en estos organismos y especies.

Por ejemplo, los codones de un ácido nucleico que codifica una pectato liasa aislados de una célula bacteriana se modifican de modo que el ácido nucleico se expresa de forma óptima en una célula bacteriana diferente de la bacteria de la que se derivó la pectato liasa, una levadura, un hongo, una célula vegetal, una célula de insecto o una célula de mamífero. Algunos métodos para optimizar codones se conocen bien en la técnica, véase, por ejemplo, el documento de Patente de Estados Unidos Nº 5.795.737; Baca (2000) Int. J. Parasitol. 30: 113-118; Hale (1998) Protein Expr. Purif. 12: 185-188; Narum (2001) Infect. Immun. 69: 7250-7253. Véase también Narum (2001) Infect. Immun. 69: 7250-7253, que describe la optimización de codones en sistemas de ratón; Outchkourov (2002) Protein Expr. Purif. 24: 18-24, que describe la optimización de codones en levadura; Feng (2000) Biochemistry 39: 15399-15409, que describe la optimización de codones en *E. coli*; Humphreys (2000) Protein Expr. Purif. 20: 252-264, que describe la optimización del uso de codones que influye en la secreción en *E. coli*.

Reensamblaje genético sintético

Un método no estocástico denominado reensamblaje genético sintético (por ejemplo, GeneReassembly™, véase, por ejemplo, el documento de Patente de Estados Unidos Nº 6.537.776) puede ser para modificar pectato liasas de la invención o para preparar nuevas pectato liasas dentro del alcance de la invención. GeneReassembly™ difiere de la redistribución estocástica en que los componentes básicos del ácido nucleico no se redistribuyen ni concatenan ni quimerizan de forma aleatoria, pero en su lugar se ensamblan de forma no estocástica.

El método de reensamblaje genético sintético no depende de la presencia de un nivel elevado de homología entre polinucleótidos a redistribuir. La invención se puede usar para generar de forma no estocástica bibliotecas (o conjuntos) de moléculas de la progenie que comprenden aproximadamente 10^{100} quimeras diferentes. Posiblemente, el reensamblaje genético sintético se puede usar incluso para generar bibliotecas que comprenden aproximadamente 10^{1000} quimeras de la progenie diferentes.

Tal método no estocástico puede producir un conjunto de moléculas de ácido nucleico quimérico finalizado que tiene un orden de ensamblaje global que se dirige mediante diseño. Consiste en las etapas de generar mediante diseño una pluralidad de componentes básicos de ácidos nucleicos específicos que tienen extremos que se pueden ligar mutuamente compatibles útiles y ensamblaje de estos componentes básicos de ácido nucleico, de modo que se consigue un orden de ensamblaje global diseñado.

En un aspecto, el reensamblaje genético sintético comprende un método para: 1) preparar una generación de molécula o moléculas de la progenie (incluyendo una molécula que comprende una secuencia de polinucleótidos, por ejemplo, una molécula que comprende una secuencia de codificación de polipéptidos), que se sometía mutagénesis para conseguir al menos una mutación, adición, supresión, y/o formación de quimeras puntuales, de uno o más moldes de generación ancestral o sobra; 2) identifica sistemáticamente la molécula o moléculas de la generación de la progenie, por ejemplo, usando un método de alto rendimiento, para al menos una propiedad de interés (tal como una mejora en una actividad enzimática); 3) opcionalmente obtener y/o catalogar una información estructural y/o y funcionar con respecto a las moléculas de la generación precursora y/o de la progenie; y 4)

opcionalmente repetir cualquiera de las etapas 1) a 3). En un aspecto, se genera (por ejemplo, a partir de un monte de polinucleótido precursor), en lo que se denomina "mutagénesis por saturación del sitio del codón", una generación de polinucleótidos de la progenie, cada uno de los cuales tiene al menos un conjunto de hasta tres mutaciones puntuales contiguas (es decir, diferentes bases que comprenden un nuevo codón), de modo que cada codón (o cada familia de codones degenerados que codifica al mismo aminoácido) se representa en cada posición del codón. Con respecto a, y codificado por, esta generación de polinucleótidos de la progenie, también se genera un conjunto de polipéptidos de la progenie, cada uno de los cuales tiene al menos una mutación puntual de un solo aminoácido. En un aspecto, se genera, en lo que se denomina "mutagénesis por saturación del sitio del aminoácido", uno de tales polipéptidos mutantes para cada una de las 19 sustituciones de alfa-aminoácidos que forman polipéptidos codificados de forma natural en todas y cada una de las posiciones del aminoácido a lo largo del polipéptido. Esto produce, para todas y cada una de las posiciones del aminoácido largo del polipéptido precursor, un total de 20 polipéptidos de la progenie distintos que incluyen el aminoácido original, o potencialmente más de 21 polipéptidos de la progenie distintos si se usan aminoácidos adicionales en lugar de o además de los 20 aminoácidos codificados de forma natural.

Este enfoque también es útil para generar mutantes que contienen, además de y/o en combinación con los 20 alfa-aminoácidos que forman polipéptidos codificados de forma natural, otros aminoácidos y derivados de aminoácidos raros y/o no codificados de forma natural. Este enfoque también es útil para generar mutantes mediante el uso de, además de y/o en combinación con sistemas de reconocimiento de codones naturales o sin alterar de hospedadores adecuados, sistemas de reconocimiento de codones alterados, sometidos a mutagénesis, y/o de diseño (tal como en una célula hospedadora con una o más moléculas de ARNt alterado).

La divulgación se refiere a recombinación y de forma más específica a un método para preparar polinucleótidos que codifican un polipéptido mediante un método de reclasificación de secuencias de polinucleótidos *in vivo* que contienen regiones de homología parcial, ensamblar los polinucleótidos para formar al menos un polinucleótido e identificar sistemáticamente los polinucleótidos para la producción de polipéptido o polipéptidos que tienen una propiedad útil.

La divulgación es útil para analizar y clasificar, con respecto a cualquier otra propiedad molecular (por ejemplo, una actividad enzimática) o combinación de propiedades permitidas por la tecnología actual, los efectos de cualquier cambio educacional conseguido (incluyendo en particular la mutagénesis por saturación). Por lo tanto, se proporciona un método integral para determinar el efecto del cambio de cada aminoácido en un polipéptido precursor con cada una de las al menos 19 sustituciones posibles. Esto permite la caracterización y clasificación de cada aminoácido en un polipéptido precursor de acuerdo con su espectro de efectos potenciales en una propiedad que se puede medir del polipéptido.

Se puede introducir un intrón en una molécula de la progenie quimérica mediante un componente básico de ácido nucleico. A menudo, los intrones tienen secuencias consenso en ambos extremos para hacerlos operativos. Además, para permitir el coste y empalme genético, los intrones pueden servir para un fin adicional a otros ácidos nucleicos para permitir la recombinación homóloga. Para este fin, y potencialmente para otros, en ocasiones puede ser deseable generar componente básico de ácido nucleico grande para introducir un intrón. Si el tamaño es demasiado grande para generar fácilmente mediante síntesis química directa dos oligos monocatenarios, tal componente básico de ácido nucleico especializado también se puede generar mediante síntesis química directa de más de dos oligos monocatenarios o usando una reacción de amplificación basada en polimerasa.

Se considera que los extremos que se pueden ligar mutuamente compatibles de los componentes básicos del ácido nucleico a ensamblar son "útiles" para este tipo de ensamblaje ordenado y permiten que los componentes básicos se acoplen en órdenes predeterminados. Por lo tanto, el orden de ensamblaje global en el que se pueden acoplar los componentes básicos del ácido nucleico se especifica mediante el diseño de los extremos que se pueden ligar y, si se va a usar más de una etapa de ensamblaje, entonces el orden de ensamblaje global en el que se pueden acoplar los componentes básicos del ácido nucleico también se especifica mediante el orden secuencial de la etapa o etapas de ensamblaje. Las piezas básicas hibridadas se pueden tratar con una enzima, tal como una ligasa (por ejemplo, T4 ADN ligasa) para conseguir la unión covalente de las piezas básicas.

El acoplamiento se puede producir de una forma que no hace uso de cada nucleótido en un saliente que participa. Es bastante probable que el acoplamiento sobreviva (por ejemplo, en un hospedador transformado) si el acoplamiento reforzado por el tratamiento con una enzima ligasa para formar lo que se puede denominar una "ligamiento del hueco" o una "ligamiento con huecos". Este tipo de acoplamiento puede contribuir a la generación de producto o productos de fondo no deseados, pero también se puede usar de forma ventajosa para aumentar la diversidad de la biblioteca de la progenie generada por el reensamblaje de ligamiento diseñado. Ciertos salientes son capaces de experimentar autoacoplamiento para formar un acoplamiento palindrómico. Un acoplamiento se refuerza sustancialmente si se refuerza por tratamiento con una enzima ligasa. La falta de fosfatos en la posición 5' en estos salientes se puede usar de forma ventajosa para evitar este tipo de autoligamiento palindrómica. Por consiguiente, se pueden preparar (u ordenar) componentes básicos de ácidos nucleicos por vía química, que carecen de un grupo fosfato en la posición 5'. Como alternativa, se pueden eliminar, por ejemplo por tratamiento con

una enzima fosfatasa, tal como una fosfatasa alcalina intestinal de ternero (CIAP), para prevenir autoligaciones palindrómicas en los procesos de reensamblaje de ligamiento.

Animales no humanos transgénicos

La invención proporciona animales no humanos transgénicos que comprenden un ácido nucleico, un polipéptido (por ejemplo, una pectato liasa), un casete o vector de expresión o una célula transformada o transfectada de la invención. La invención también proporciona métodos para preparar y usar estos animales no humanos transgénicos.

Los animales no humanos transgénicos pueden ser, por ejemplo, cabras, conejos, ovejas, cerdos, vacas, ratas y ratones, que comprenden los ácidos nucleicos de la invención. Estos animales se pueden usar, por ejemplo, como modelos *in vivo* para estudiar la actividad de pectato liasa, o, como modelos para identificar sistemáticamente agentes que cambien la actividad de pectato liasa *in vivo*. Las secuencias de codificación para los polipéptidos para expresar en los animales no humanos transgénicos se pueden diseñar para que sean constitutivas, o, bajo el control de factores reguladores de la transcripción específicos de tejido, específicos de desarrollo o inducibles. Los animales no humanos transgénicos se pueden diseñar y generar usando cualquier método con en la técnica; véanse, por ejemplo, los documentos de Patente de Estados Unidos N^{os} 6.211.428; 6.187.992; 6.156.952; 6.118.044; 6.111.166; 6.107.541; 5.959.171; 5.922.854; 5.892.070; 5.880.327; 5.891.698; 5.639.940; 5.573.933; 5.387.742; 5.087.571, que describen la preparación y uso de células y embriones transformados y ratones, y ratas, conejos, ovejas, cerdos y vacas transgénicos. Véanse también, por ejemplo, Pollock (1999) J. Immunol. Methods 231: 147-157, que describe la producción de proteínas recombinantes en la leche de animales para lácteos transgénicos; Baguisi (1999) Nat. Biotechnol. 17: 456-461, que demuestra la producción de cabras transgénicas. El documento de Patente de Estados Unidos N^o 6.211.428, describe la preparación y uso de mamíferos no humanos transgénicos que expresan en sus cerebros un constructo de ácido nucleico que comprende una secuencia de ADN. El documento de Patente de Estados Unidos N^o 5.387.742, describe la inyección de secuencias de ADN recombinante o sintético clonado en embriones de ratón fertilizados, implantando los embriones inyectados en hembras pseudopreñadas, y crecimiento a término de ratones transgénicos cuyas células expresan proteínas relacionadas con la patología de la enfermedad de Alzheimer. El documento de Patente de Estados Unidos N^o 6.187.992, describe la preparación y uso de un ratón transgénico cuyo genoma comprende una alteración del gen que codifica una proteína precursora amiloide (APP).

También se pueden usar "animales genosuprimidos" para poner en práctica los métodos de la invención. Por ejemplo, en un aspecto, los animales transgénicos o modificados de la invención comprenden un "animal genosuprimido", por ejemplo, un "ratón genosuprimido", modificado por ingeniería para que no expresen un gen endógeno, que se sustituye con un gen que expresa una pectato liasa de la invención, o, una proteína de fusión que comprende una pectato liasa de la invención.

Plantas y Semillas Transgénicas

La invención proporciona plantas y semillas transgénicas que comprenden un ácido nucleico, un polipéptido (por ejemplo, una pectato liasa), un casete de expresión o un vector o una célula transformada o transfectada de la invención. La invención también proporciona productos de plantas, por ejemplo, aceites, semillas, hojas, extractos similares, que comprenden un ácido nucleico y/o un polipéptido (por ejemplo, una pectato liasa) de la invención. La planta transgénica puede ser dicotiledónea (una dicot) o monocotiledónea (una monocot). La invención también proporciona métodos para preparar y usar estas plantas y semillas transgénica es. La planta o célula vegetal transgénica que expresa un polipéptido de la presente invención se puede construir de acuerdo con cualquier método conocido en la técnica. Véase, por ejemplo, el documento de Patente de Estados Unidos N^o 6.309.872.

Los ácidos nucleicos y constructos de expresión de la invención se pueden introducir en una célula vegetal mediante cualquier medio. Por ejemplo, los ácidos nucleicos o constructos de expresión se pueden introducir en el genoma de un hospedador de planta deseado, o, los ácidos nucleicos o constructos de expresión pueden ser episomas. La introducción en el genoma de una planta deseada se puede realizar de modo que la producción de la pectato liasa del hospedador se regula mediante elementos de control de la transcripción o de la traducción endógenos. La invención también proporciona "plantas genosuprimidas" en las que la inserción de la secuencia genética, por ejemplo, mediante recombinación homóloga, ha alterado la expresión del gen endógeno. Algunos medios para generar plantas "genosuprimidas" se conocen bien en la técnica, véanse, por ejemplo, Strepp (1998) Proc Natl. Acad. Sci. USA 95: 4368-4373; Miao (1995) Plant J 7:359-365. Véase el análisis sobre plantas transgénica se sigue a continuación.

Los ácidos nucleicos de la invención se pueden usar para conferir rasgos deseados en cualquier parte esencial de la planta, por ejemplo, en plantas que producen almidón, tales como patata, trigo, arroz, cebada, y similares. Los ácidos nucleicos de la invención se pueden usar para manipular rutas metabólicas de una planta para optimizar o para alterar la expresión del hospedador, de pectato liasa. Pueden cambiar la actividad de pectato liasa en una planta. Como alternativa, una pectato liasa de la invención se puede usar en la producción de una planta transgénica para producir un compuesto no producidos de forma natural por la planta. Esto puede reducir los costos de producción o puede crear un producto.

En un aspecto, la primera etapa en la producción de una planta transgénica implica la preparación de un constructor de expresión para la expresión en una célula vegetal. Estas técnicas se conocen bien en la técnica. Pueden incluir la selección y clonación de un promotor, una secuencia de codificación para facilitar una unión eficaz de ribosomas a ARNm y selección de las secuencias terminadas de genes apropiados. Un promotor constitutivo a modo de ejemplo es CaMV35S, del virus del mosaico de la coliflor, que generalmente da como resultado un grado elevado de expresión en plantas. Otros promotores son más específicos y responden a señales en el entorno interno o externo de la planta. Un promotor inducible por la luz a modo de ejemplo es el promotor del gen cab, que codifica las proteínas de unión a clorofila a/b principal.

En un aspecto, el ácido nucleico se modifica para conseguir una mayor expresión en una célula vegetal. Por ejemplo, es probable que una secuencia de la invención tenga un porcentaje más elevado de pares de nucleótidos A-T en comparación con el observado en una planta, algunas de las cuales prefieren pares de nucleótidos G-C. Por lo tanto, los nucleótidos A-T en la secuencia de codificación se pueden sustituir con nucleótidos G-C sin cambiar de forma significativa la secuencia de aminoácidos para aumentar la producción del producto genético en células vegetales.

Se pueden añadir genes marcadores seleccionables para el constructo genético para identificar células o tejidos vegetales que se han integrado de forma satisfactoria en el transgén. Esto puede ser necesario porque la consecución de la incorporación y la expresión de en células vegetales es un suceso raro, que se produce solamente en un pequeño tanto por ciento de los tejidos o células dirigidas. Los genes marcadores que se pueden seleccionar codifican proteínas que proporcionan resistencia a agentes que normalmente son tóxicos para las plantas, tales como antibióticos o herbicidas. Solamente sobrevivirán las células vegetales que han integrado el gen marcador seleccionable cuando se cultiven en un medio que contiene el antibiótico o herbicida apropiado. Al igual que para otros genes insertados, los genes marcadores también necesitan secuencias promotoras y determinación para un funcionamiento apropiado.

En un aspecto, la preparación de plantas o semillas transgénicas comprende la incorporación de secuencias de la invención y, ocasionalmente, genes marcadores en un constructor de expresión de dianas (por ejemplo, un plásmido), junto con la colocación del promotor y las secuencias terminadoras. Esto puede implicar la transferencia del gen modificado en la planta a través de un método adecuado. Por ejemplo, se puede introducir un constructor directamente en el ADN genómico de la célula vegetal usando técnicas tales como electroporación y microinyección de protoplastos de células vegetales, o los constructos se pueden introducir directamente en tejido vegetal usando métodos balísticos, tales como bombardeo de partículas de arena. Por ejemplo, véase, por ejemplo, Christou (1997) Plant Mol. Biol. 35: 197-203; Pawlowski (1996) Mol. Biotechnol. 6: 17-30; Klein (1987) Nature 327: 70-73; Takumi (1997) Genes Genet. Syst. 72: 63-69, que analizan el uso del bombardeo de partículas de introducir transgenes en trigo; y Adam (1997) mencionado anteriormente, para uso del bombardeo de partículas para introducir YAC en células vegetales. Por ejemplo, Rinehart (1997) mencionado anteriormente, usó bombardeo de partículas para generar plantas de algodón transgénicas. El aparato para hacer de la partícula se describe en el documento de Patente de Estados Unidos Nº 5.015.580; y, el instrumento de aceleración de partículas PDS-2000 de BioRad (Biolistics) disponible en el mercado; véase también, John, documento de Patente de Estados Unidos Nº 5.608.148; y Ellis, documento de Patente de Estados Unidos Nº 5.681.730, que describe la transformación de gimnospermas mediada por partículas.

En un aspecto, los protoplastos se pueden inmovilizar e inyectar con ácidos nucleicos, por ejemplo, un constructor de expresión. Aunque la regeneración de plantas a partir de protoplastos no es tan fácil como con los cereales, la regeneración de plantas es posible en legumbres usando embriogénesis somática a partir de callo derivado de protoplasto. Los tejidos organizados se pueden transformar con ADN desnudo usando técnicas de pistola genética, en la que el ADN se recubre en microproyectiles de tungsteno, con disparos de 1/100ª del tamaño de las células, que cortan el ADN profundo en células y orgánulos. A continuación, el tejido transformado se induce para su regeneración, normalmente mediante embriogénesis somática. Esta técnica ha sido satisfactoria en varias especies especiales que incluyen maíz y arroz.

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

Como alternativa, Los ácidos nucleicos, por ejemplo, un constructor de expresión, se pueden combinar con regiones de flaqueo de T-ADN y se pueden introducir en un vector hospedador de *Agrobacterium tumefaciens*. Las funciones de virulencia del hospedador de *Agrobacterium tumefaciens* dirigirán la inserción del constructo y el marcador adyacente en el ADN de la célula vegetal cuando la célula se infecta con las bacterias. En la bibliografía científica se describen bien algunas técnicas de transformación mediadas por *Agrobacterium tumefaciens*, incluyendo el desarme y uso de vectores binarios. Véanse, por ejemplo, Hors (1984) Science 233: 496-498; Fraley (1983) Proc. Natl. Acad. Sci. USA 80: 4803 (1983); Gene Transfer to Plants, Potrykus, ed. (Springer-Verlag, Berlín 1995). El ADN en una célula de *A. tumefaciens* está contenido en el cromosoma bacteriano así como en otra estructura conocida como un plásmido Ti (inductor tumoral). El plásmido Ti contiene un tramo de ADN denominado

T-ADN (~20 kb de longitud) que se transfiere a la célula vegetal en el proceso de infección en la serie de genes de vir (virulencia) que dirigen el proceso de infección. *A. tumefaciens* solamente puede infectar a una planta a través de heridas: cuando una raíz o tallo de la planta está herido emite ciertas señales químicas, como respuesta a lo cual, los genes vir de *A. tumefaciens* se activan y dirigen una serie de sucesos necesarios para transferencia del T-ADN desde el plásmido Ti al cromosoma de la planta. Entonces, el T-ADN entra en la célula vegetal a través de la herida. Una especulación es que el T-ADN espera hasta que el ADN de la planta se está replicando o transcribiendo, entonces se introduce a sí mismo en el ADN expuesto de la planta. Para usar *A. tumefaciens* como un vector de transgén, la sección que induce el tumor de T-ADN se tiene que retirar, a la vez que se mantienen las regiones del límite del T-ADN y los genes vir. A continuación, el transgén se inserta entre las regiones del límite del T-ADN, en las que se transfiere a la célula vegetal y se integra en los cromosomas de la planta.

La invención proporciona la transformación de plantas monocotiledóneas usando los ácidos nucleicos de la invención, que incluyen cereales importantes, véase Hiei (1997) Plant Mol. Biol. 35: 205-218. Véase también, por ejemplo, Horsch, Science (1984) 233: 496; Fraley (1983) Proc. Natl Acad. Sci USA 80: 4803; Thykjaer (1997) mencionado anteriormente; Park (1996) Plant Mol. Biol. 32: 1135-1148, que analiza la integración del T-ADN en el ADN genómico. Véase también D'Halluin, documento de Patente de Estados Unidos N° 5.712.135, que describe un proceso para la integración establece un ADN que comprende un gen que es funcional en una célula de un cereal, u otra planta monocotiledónea.

En un aspecto, la tercera etapa puede implicar la selección y regeneración de plantas completas capaces de transmitir el gen día no incorporado a la siguiente generación. Tales técnicas de regeneración se basan en la manipulación de ciertas fitohormonas en medio de crecimiento de cultivo tisular, por lo general basándose en un marcador biocida y/o herbicida que se ha introducido junto con las secuencias de nucleótidos deseadas. La regeneración de plantas a partir del protoplastos cultivados se describe en Evans *et al.*, Protoplasts Isolation and Culture, Handbook of Plant Cell Culture, pp. 124-176, MacMillan Publishing Company, New York, 1983; y Binding, Regeneration of Plants, Plant Protoplasts, pp. 21-73, CRC Press, Boca Raton, 1985. La regeneración también se puede obtener a partir de callos, explantes, órganos de plantas, o partes de las mismas. Por lo general, tales técnicas de regeneración se describen en Klee (1987) Ann. Rev. of Plant Phys. 38: 467-486. Para obtener plantas completas a partir de tejidos transgénicos tales como embriones inmaduros, se pueden cultivar en condiciones ambientales controladas en una serie de medios que contienen nutrientes y hormonas, un proceso conocido como cultivo tisular. Una vez que las plantas se generan y producen semillas, comienza la evaluación de la progenie.

Después de incorporar el casete de expresión de forma estable en plantas transgénicas, se puede introducir en otras plantas mediante cruce sexual. Se puede usar cualquiera de las técnicas convencionales de reproducción, dependiendo de las especies a cruzar. Dado que la expresión transgénica de los ácidos nucleicos de la invención conduce a cambios fenotípicos, las plantas que comprende los ácidos nucleicos recombinantes de la invención se pueden cruzar sexualmente con la segunda planta para obtener un producto final. Por lo tanto, la semilla de la invención se puede derivar de un cruce entre dos plantas transgénica es de la invención, o un cruce entre una planta de la invención y otra planta. Los efectos deseados (por ejemplo, expresión de los polipéptidos de la invención para producir una planta en la que el comportamiento de floración está alterado) se pueden aumentar cuando ambas plantas precursoras expresan los polipéptidos (por ejemplo, una pectato liasa) de la invención. Los efectos deseados se pueden pasar a generaciones futuras de plantas mediante medios de propagación convencionales.

Los ácidos nucleicos y polipéptidos de la invención se expresan en se insertan en cualquier planta o semilla. Las plantas más escénicas de la invención pueden ser dicotiledóneas o monocotiledóneas. Algunos ejemplos de plantas transgénicas monocotiledóneas de la invención son hierbas, tales como hierba de la pradera (blue grass, *Poa*), lleva forrajera tal como festuca, *lolium*, tierra templada, tal como *Agrostis*, y cereales, por ejemplo, trigo, arena, centeno, cebada, arroz, sorgo y maíz (maíz). Algunos ejemplos de plantas transgénicas dicotiledóneas de la invención son tabaco, legumbres, tales como faltantes altramuces, patata, remolacha azucarera, guisante, judía y soja, y plantas crucíferas (familia *Brassicaceae*), tales como coliflor, semilla de colza, y el organismo modelo muy relacionado *Arabidopsis thaliana*. Por lo tanto, las plantas y semillas transgénicas de la invención incluyen una amplia variedad de plantas, que incluyen, pero no se limitan a, especies de los géneros *Anacardium*, *Arachis*, *Asparagus*, *Atropa*, *Avena*, *Brassica*, *Citrus*, *Citrullus*, *Capsicum*, *Carthamus*, *Cocos*, *Coffea*, *Cucumis*, *Cucurbita*, *Daucus*, *Elaeis*, *Fragaria*, *Glycine*, *Gossypium*, *Helianthus*, *Heterocallis*, *Hordeum*, *Hyoscyamus*, *Lactuca*, *Linum*, *Lolium*, *Lupinus*, *Lycopersicon*, *Malus*, *Manihot*, *Majorana*, *Medicago*, *Nicotiana*, *Olea*, *Oryza*, *Panieum*, *Pannisetum*, *Persea*, *Phaseolus*, *Pistachia*, *Pisum*, *Pirus*, *Prunus*, *Raphanus*, *Ricinus*, *Secale*, *Senecio*, *Sinapis*, *Solanum*, *Sorghum*, *Theobromus*, *Trigonella*, *Triticum*, *Vicia*, *Vitis*, *Vigna*, y *Zea*.

En realizaciones alternativas, los ácidos nucleicos de la invención se expresan en plantas que contienen células de fibra, que incluyen, por ejemplo, algodón, árbol de algodón de seda (Kapok, *Ceiba pentandra*), Sauce del desierto, arbusto de la creosota, winterfat, balsa, ramio, kenaf, cáñamo, hibisco, yute, sisal, abacá y fibra de lino. En realizaciones alternativas, las plantas más clínicas de la invención pueden ser miembros del género *Gossypium*, que incluye miembros de cualquiera de las especies *Gossypium*, tales como *G. arboreum*; *G. herbaceum*, *G. barbadense*, y *G. hirsutum*.

La invención también proporciona la transgénica sabrosa para la producción de grandes cantidades de los polipéptidos (por ejemplo, una pectato liasa o anticuerpo) de la invención. Por ejemplo, véase Palmgren (1997) Trends Genet. 13: 348; Chong (1997) Transgenic Res. 6: 289-296 (producción de beta-caseína de proteína de leche humana en plantas de patata transgénicas usando un promotor de manopina sintasa bidireccional, inducible con auxina, (mas1',2') con métodos de transformación de disco de hoja mediado por *Agrobacterium tumefaciens*).

Usando procedimientos conocidos, un experto puede identificar sistemáticamente plantas de la invención detectando el aumento o disminución de ARNm del transgén o proteína en plantas transgénicas. En la técnica se conocen bien algunos medios para detectar y cuantificar ARNm o proteínas.

Polipéptidos y péptidos

En un aspecto, la invención proporciona polipéptidos aislados o recombinantes que tienen una identidad de secuencia (por ejemplo, una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o superior, o completa (100 %)) con la SEC ID N°: 78. El polipéptido tiene una actividad de pectato liasa (por ejemplo, pectinasa).

La identidad puede ser sobre la longitud completa del polipéptido, o, la identidad puede ser sobre una región de al menos 300, 350, 400, 450, 500, 550, 600, 650 o más restos. Los polipéptidos también pueden ser más cortos que la longitud completa de los polipéptidos a modo de ejemplo. Los polipéptidos (péptidos, fragmentos) pueden tener un tamaño que varía entre aproximadamente 5 y la longitud completa de un polipéptido, por ejemplo, una enzima, tal como una pectato liasa; algunos tamaños a modo de ejemplo son de aproximadamente 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650 o más restos, por ejemplo, textos contiguos de una pectato liasa a modo de ejemplo de la invención. Tales péptidos pueden ser útiles, por ejemplo, como sondas de marcado, antígenos, tolerágenos, motivos, sitio activo de pectato liasas, dominios de unión a carbohidrato, y similares. Tales polipéptidos también incluyen anticuerpos capaces de unirse a una enzima de la invención.

Los polipéptidos de la invención incluyen pectato liasas en una forma activa o inactiva. Por ejemplo, los polipéptidos de la invención incluyen proproteínas antes de la "maduración" o procesamiento de secuencias prepro, por ejemplo, mediante una enzima de procesamiento de proproteína, tal como una proproteína convertasa para generar una proteína madura "activa". Los polipéptidos de la invención incluyen pectato liasas inactivas por otras razones, por ejemplo, antes de "activación" mediante un suceso de procesamiento posterior a la traducción, por ejemplo, una acción de endo o exo-peptidasa o proteinasa, un suceso de fosforilación, una amidación, una glicosilación o una sulfatación, un suceso de dimerización, y similares. En la técnica se conocen bien algunos métodos para identificar secuencias de dominio "prepro" y secuencias señal, véase, por ejemplo, Van de Ven (1993) Crit. Rev. Oncog. 4 (2): 115-136. Por ejemplo, para identificar una secuencia prepro, la proteína se purifica a partir del espacio extracelular y la secuencia de proteína N-terminal se determina se compara con la forma sin procesar.

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

La invención incluye polipéptidos con o sin una secuencia señal y/o una secuencia prepro. La invención incluye polipéptidos consecuencias señal y/o secuencias prepro heterólogas. La secuencia prepro (que incluye una secuencia de la invención usada como un dominio prepro heterólogo) se puede localizar en el extremo amino terminal o en el extremo carboxi terminal de la proteína. La invención también incluye secuencias señal aisladas o recombinantes, secuencias prepro y dominios catalíticos (por ejemplo, "sitios activos") que comprenden secuencias de la invención.

Los polipéptidos péptidos de la invención se pueden aislar de fuentes naturales, pueden ser polipéptidos sintéticos, o pueden ser polipéptidos generados de forma recombinante. Los polipéptidos y proteínas se pueden expresar de forma recombinante *in vitro* o *in vivo*. Los péptidos y polipéptidos de la invención se pueden preparar y aislar usando cualquier método conocido en la técnica. Los polipéptidos y péptidos de la invención también se pueden sintetizar, total o parcialmente, usando métodos químicos bien conocidos en la técnica. Véase, por ejemplo, Caruthers (1980) Nucleic Acids Res. Symp. Ser. 215-223; Horn (1980) Nucleic Acids Res. Symp. Ser. 225-232; Banga, A.K., Therapeutic Peptides and Proteins, Formulation, Processing and Delivery Systems (1995) Technomic Publishing Co. Lancaster, PA. Por ejemplo, la síntesis de péptidos se puede realizar usando diversas técnicas en fase sólida (véase por ejemplo, Roberge (1995) Science 269: 202; Merrifield (1997) Methods Enzymol. 289: 3-13) y la síntesis automatizada se puede conseguir, por ejemplo, usando el Sintetizador de Péptidos ABI 431A (Perkin Elmer) de acuerdo con las instrucciones proporcionadas por el fabricante.

Los péptidos y polipéptidos de la invención también se pueden glicosilar. La glicosilación se pueda añadir después de la traducción ya sea de forma química o mediante mecanismos de biosíntesis celular, en los que el último incorpora el uso de motivos de glicosilación conocidos, que pueden ser nativos para la secuencia se pueden añadir como un péptido o se pueden añadir en la secuencia de codificación del ácido nucleico. La glicosilación puede ser unida a O o unida a N.

Los péptidos y polipéptidos de la invención, como se ha definido anteriormente, incluyen todas las formas "miméticas" y "peptidomiméticas". Los términos "mimético" y "peptidomimético" se refieren a un compuesto químico sintético que tiene básicamente las mismas características estructurales y/o funcionales de los polipéptidos de la invención. El mimético puede estar formado totalmente de análogos de aminoácidos sintéticos, no naturales, o, es una molécula quimérica de aminoácidos peptídicos parcialmente naturales y análogos de aminoácidos parcialmente naturales. El mimético también por incorporar cualquier cantidad de funciones conservativas de aminoácidos naturales siempre y cuando tales sustituciones tampoco alteren básicamente la estructura y/o actividad del mimético. Al igual que con los polipéptidos de la invención que son variantes conservativas, la experimentación de rutina determinará si un mimético está dentro del alcance de la invención, es decir, que su estructura y/o función no está alterada básicamente. Por lo tanto, en un aspecto, una composición mimética está dentro del alcance de la invención si tiene una actividad de peptidasa.

Las composiciones miméticas de polipéptidos de la invención pueden contener cualquier combinación de componentes estructurales no naturales. En un aspecto alternativo, las composiciones miméticas de la invención incluyen uno o los tres grupos estructurales siguientes: a) grupos de unión a restos distintos de los enlaces de unión amida natural ("enlace peptídico"); b) restos no naturales en lugar de restos de aminoácidos de origen natural; o c) restos que inducen una mimética estructural secundaria, es decir, para inducir por estabilizar una estructura secundaria, por ejemplo, un giro beta, giro gamma, lámina beta, conformación de hélice alfa, y similares. Por ejemplo, un polipéptido de la invención se puede caracterizar como un mimético cuando todos o algunos de sus restos se unen mediante un medio químico distinto de enlaces peptídico naturales. Los restos peptidomiméticos individuales se pueden unir mediante enlaces específicos, otros enlaces químicos o medios de acoplamiento, tales como, por ejemplo, glutaraldehído, ésteres de N-hidroxisuccinimida, maleimidas bifuncionales, N,N'-díciclohexilcarbodiimida (DCC) o N,N'-diisopropilcarbodiimida (DIC). Los grupos de unión que pueden ser una alternativa a los enlaces de unión de amida tradicional ("enlace peptídico") incluyen, por ejemplo, cetometileno (por ejemplo, -C(=O)-CH₂- para -C(=O)-NH-), aminometileno (CH₂-NH), etileno, olefina (CH=CH), éter (CH₂-O), tioéter (CH₂-S), tetrazol (CN₄-), tiazol; retroamida, tioamida, o éster (véase, por ejemplo, Spatola (1983) en Chemistry and Biochemistry of Amino Acids, Peptides and Proteins, Vol. 7, pp 267-357, "Peptide Backbone Modifications", Marcell Dekker, NY).

Un polipéptido de la invención también se puede caracterizar como un mimético al contener todos o algunos restos no naturales en lugar de restos de aminoácidos de origen natural. Los restos no naturales se describen bien en la bibliografía científica y de patente; a continuación se describen unas pocas composiciones no naturales a modo de ejemplo útiles como miméticos de restos de aminoácidos naturales y directrices. Se pueden generar miméticos de aminoácidos aromáticos por sustitución, por ejemplo, con D- o L- fenilalanina; D- o L- fenilglicina; D- o L-2 tienoilalanina; D- o L-1, -2, 3-, o 4- pirenolalanina; D- o L-3 tienoilalanina; D- o L-(2-piridinil)-alanina; D- o L-(3-piridinil)-alanina; D- o L-(2-pirazinil)-alanina; D- o L-(4-isopropil)-fenilglicina; D-(trifluorometil)-fenilglicina; D-(trifluorometil)-fenilalanina; D-p-fluoro-fenilalanina; D- o L-p-bifenilfenilalanina; D- o L-p-metoxi-bifenilfenilalanina; D- o L-2-indol(alquil)alaninas; y, D- o L-alquilalaninas, en los que el alquilo puede ser metilo sustituido o sin sustituir, etilo, propilo, hexilo, butilo, pentilo, isopropilo, iso-butilo, sec-isotilo, iso-pentilo, o un aminoácido no ácido. Los anillos aromáticos de un aminoácido no natural incluyen, por ejemplo, tiazolilo, tiofenilo, pirazolilo, benzoimidazolilo, naftilo, furanilo, pirrolilo, y anillos aromáticos de piridilo.

Se pueden generar miméticos de aminoácidos ácidos mediante sustitución, por ejemplo, con aminoácidos que no son carboxilato a la vez que se mantiene una carga negativa; (fosfono)alanina; treonina sulfatada. Los grupos laterales carboxilo (por ejemplo, aspártico o glutámico) también se prevé modificar en forma selectiva por reacción con carbodiimidas (R'-N-C-N-R') tales como, por ejemplo, 1-ciclohexil-3(2-morfolinil-(4-etil) carbodiimida o 1-etil-3(4-azonia-4,4-dimetolpentil) carbodiimida. El aspártico o el glutámico también se pueden convertir en restos de asparaginilo y glutaminilo por reacción con iones amonio. Se pueden generar miméticos de aminoácidos básicos por sustitución, por ejemplo, (además de lisina y arginina) con los aminoácidos ornitina, citrulina, o ácido (guanidino)-acético, o ácido (guanidino)alquil-acético, en el que alquilo se ha definido anteriormente. El derivado de nitrilo (por ejemplo, que contiene el resto CN en lugar de COOH) se puede sustituir para asparagina o glutamina. Los restos de asparaginilo y glutaminilo se pueden desaminar a los correspondientes restos de aspártico o glutámico. Se pueden generar miméticos de restos de arginina haciendo reaccionar arginilo con, por ejemplo, uno o más reactivos convencionales, que incluyen, por ejemplo, fenilgloxal, 2,3-butanodiona, 1,2-ciclohexanodiona, o ninhidrina, preferentemente en condiciones alcalinas. Se pueden generar miméticos de restos de tirosina haciendo reaccionar tirosilo, por ejemplo, con compuestos de diazonio aromáticos o tetranitrometano. Se puede usar N-acetilimidazol y tetranitrometano para formar especies de O-acetil tirosilo y derivados 3-nitro, respectivamente. Se pueden generar miméticos de restos de cisteína haciendo reaccionar restos de cisteinilo, por ejemplo, con alfa-haloacetatos tales como ácido 2-cloroacético o cloroacetamida y las aminas correspondientes; para dar carboximetilo o derivados de

carboxiamidometilo. También se puede generar miméticos de restos de cisteína haciendo reaccionar restos de cisteinilo, por ejemplo, con bromo-trifluoroacetona ácido, alfa-bromo-beta-(5-imidozoi) propiónico; fosfato de cloroacetilo, N-alquilmaleimidas, disulfuro de 3-nitro-2-piridilo; disulfuro de metil 2-piridilo; p-cloromercuribenzoato; 2-cloromercuri-4-nitrofenol; o, cloro-7-nitrobenzo-oxa-1,3-diazol. Se pueden generar miméticos de lisina (y los restos amino terminales se pueden alterar) haciendo reaccionar lisinilo, por ejemplo, ácido succínico u otros anhídridos de ácido carboxílico. También se pueden generar miméticos de lisina y otros restos que contienen alfa-amino por reacción con imidoésteres, tales como picolinimidato de metilo, fosfato de piridoxal, piridoxal, cloroborohidruro, ácido trinitrobenzenosulfónico, O-metilisourea, 2,4, pentanodiona, y reacciones catalizadas por transamidasa con glioxilato. Se pueden generar miméticos de metionina por reacción, por ejemplo, con sulfóxido de metionina. Algunos miméticos de prolina incluyen, por ejemplo, ácido pipercolico, ácido tiazolidina carboxílico, 3- o 4- hidroxil prolina, deshidroprolina, 3- o 4-metilprolina, o 3,3,-dimetilprolina. Se pueden generar miméticos de restos de histidina haciendo reaccionar histidilo, por ejemplo, con dietilprocarbonato o bromuro de para-bromofenacilo. Otros miméticos incluyen, por ejemplo, los generados por hidroxilación de prolina y lisina; fosforilación de los grupos hidroxilo de restos de serilo o treonilo; metilación de los grupos alfa-amino de lisina, arginina e histidina; acetilación de la N-terminal amina; metilación de restos de amida de cadena principal o sustitución con N-metil aminoácidos; o amidación de grupos carboxilo C-terminales.

Un resto, por ejemplo, un aminoácido de un polipéptido de la invención también se puede sustituir con un aminoácido (o resto peptidomimético) de la quiralidad opuesta. Por lo tanto, cualquier aminoácido de origen natural en la configuración L (que también se puede denominar como R o S, dependiendo de la estructura de la entidad química) se puede sustituir con el aminoácido del mismo tipo de estructura química o un peptidomimético, pero de la quiralidad opuesta, denominado D-aminoácido, pero también denominado como forma R o S.

Algunos métodos pueden modificar los polipéptidos de la invención ya sea mediante procesos naturales, tales como procesamiento posterior a la traducción (por ejemplo, fosforilación, acilación, etc), o mediante técnicas de modificación química, y los polipéptidos modificados resultantes. Se pueden producir modificaciones en cualquier parte en el polipéptido, incluyendo la estructura principal del péptido, las cadenas laterales del aminoácido y los extremos amino o carboxilo. Se observará que el mismo tipo de modificación puede estar presente en el mismo o grados variables en diversos sitios en un polipéptido dado. También, un polipéptido dado puede presentar muchos tipos de modificaciones. Algunas modificaciones incluyen acetilación, acilación, ADP-ribosilación, amidación, unión covalente de flavina, unión covalente de un resto heme, unión covalente de un nucleótido o derivado de nucleótido, unión covalente de un lípido o derivado de lípido, unión covalente de un fosfatidilinositol, ciclación por reticulación, formación de enlaces disulfuro, desmetilación, formación de reticulaciones covalentes, formación de cisteína, formación de piroglutamato, formilación, gamma-carboxilación, glicosilación, formación de anclas de GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, pegilación, procesamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfatación, y adición mediada por ARN de transferencia de aminoácidos a proteínas tal como arginilación. Véase, por ejemplo, Creighton, T.E., *Proteins - Structure and Molecular Properties* 2ª Ed., W.H. Freeman y Company, New York (1993); *Posttranslational Covalent Modification of Proteins*, B.C. Johnson, Ed., Academic Press, New York, pp. 1-12 (1983).

También se pueden usar métodos de síntesis química de péptidos químicos para sintetizar el polipéptido o fragmentos de la invención. Tal método se ha conocido en la técnica desde el comienzo de 1960 (Merrifield, R. B., *J. Am. Chem. Soc.*, 85: 2149-2154, 1963) (Véase también Stewart, J. M. e Young, J. D., *Solid Phase Peptide Synthesis*, 2ª Ed., Pierce Chemical Co., Rockford, Ill., pp. 11-12)) y se han usado recientemente en kits de diseño y síntesis de péptidos de laboratorio disponibles en el mercado (Cambridge Research Biochemicals). Tales kits de laboratorio disponibles en el mercado por lo general han usado las enseñanzas de H. M. Geysen *et al.*, *Proc. Natl. Acad. Sci., USA*, 81: 3998 (1984) y proporcionan la síntesis de péptidos sobre las puntas de una multitud de "varillas" o "sujeciones" todos los cuales se conectan a una sola placa. Cuando se usa tal sistema, una placa de varillas o sujeciones se invierte y se inserta en una segunda placa de pocillos o depósitos correspondientes, que contienen soluciones para que un aminoácido apropiado se una o se ancle a la punta de la sujeción o la varilla. Mediante la repetición de tal etapa del proceso, es decir, mediante la inversión e inserción de las puntas de la varilla y la sujeción en soluciones apropiadas, los aminoácidos se preparan en péptidos deseados. Además, está disponible un número de sistemas de síntesis de péptidos de FMOC. Por ejemplo, el ensamblaje de un polipéptido o fragmento se puede realizar en un soporte sólido usando un sintetizador de péptidos automatizado de Applied Biosystems, Inc. Modelo 431A™. Tal equipo proporciona un acceso fácil a los péptidos de la invención, ya sea por síntesis directa o por síntesis de una serie de fragmentos que se pueden acoplar usando otras técnicas conocidas.

La invención incluye pectato liasas de la invención con y sin señal. El polipéptido que comprende una secuencia señal de la invención puede ser una pectato liasa de la invención u otra pectato liasa u otra enzima u otro polipéptido.

La divulgación incluye pectato liasas inmovilizadas, anticuerpos anti-pectato liasa y fragmentos de los mismos. La divulgación proporciona métodos para inhibir la actividad de pectato liasa, por ejemplo, usando mutantes negativos dominantes o anticuerpos anti-pectato liasa. La invención incluye heterocomplejos, por ejemplo, proteínas de fusión, heterodímeros, etc., que comprenden las pectato liasas de la invención.

Los polipéptidos de la invención pueden tener una actividad de pectato liasa en diversas condiciones, por ejemplo, extremas en pH y/o temperatura, agentes oxidantes, y similares. Algunos métodos pueden conducir a preparaciones alternativas de pectato liasa con diferentes eficacias catalíticas y estabilidades, por ejemplo, hacia temperatura, agentes oxidantes y cambios en las condiciones de lavado. En un aspecto, se pueden producir variantes de pectato liasa usando técnicas de mutagénesis dirigida al sitio y/o mutagénesis aleatoria. En un aspecto, se puede usar evolución dirigida para producir una gran diversidad de variantes de pectato liasa con especificidades y estabilidad alternativas.

Las proteínas de la invención también son útiles como reactivos de investigación para identificar moduladores de pectato liasa, por ejemplo, activadores o inhibidores de actividad de pectato liasa. En resumen, las muestras de ensayo (compuestos, caldos de cultivo, extractos, y similares) se añaden a ensayos de pectato liasa para determinar su capacidad para inhibir la escisión del sustrato. Los inhibidores identificados de este modo se pueden usar en industria y en investigación para reducir o prevenir la proteólisis no deseada. Al igual que con las pectato liasas, los inhibidores se pueden combinar para aumentar el espectro de actividad.

La divulgación también proporciona métodos para descubrir nuevas pectato liasas usando los ácidos nucleicos, polipéptidos y anticuerpos que se describen en el presente documento, se pueden identificar sistemáticamente bibliotecas de fagos lambda para descubrimiento basado en la expresión de pectato liasas, para permitir la detección de clones tóxicos; aumento del acceso al sustrato; reducción de la necesidad de modificación por ingeniería de un hospedador, derivando el potencial de cualquier sesgo resultante de la escisión de masa de la biblioteca; y, crecimiento más rápido a bajas densidades de clones. La identificación sistemática de bibliotecas de fagos lambda se puede realizar en fase líquida o en fase sólida. En un aspecto, la invención proporciona identificación sistemática en fase líquida. Esto proporciona una mayor flexibilidad en condiciones de ensayo; flexibilidad del sustrato adicional; sensibilidad más elevada para clones débiles; y facilidad de automatización sobre identificación sistemática en fase sólida.

Algunos métodos de identificación sistemática pueden usar las proteínas y ácidos nucleicos de la invención y la automatización robótica permite la ejecución de muchos miles de reacciones biocatalíticas y ensayos de identificación sistemática en un breve periodo de tiempo, por ejemplo, al día, así como asegurar un nivel elevado de precisión y reproducibilidad (véase el análisis de matrices, a continuación). Como resultado, se puede producir una biblioteca de compuestos derivados en cuestión de semanas. Para enseñanzas adicionales sobre modificación de moléculas, incluyendo moléculas pequeñas, véase el documento PCT/US94/09174.

La presente invención incluye enzimas pectato liasa que son variantes de carbonilo hidrolasa de origen no natural (por ejemplo, variantes de pectato liasa) que tienen una actividad proteolítica, estabilidad, especificidad de sustrato, perfil de pH y/o rendimiento diferentes en comparación con la carbonil hidrolasa precursora a partir de la que se deriva la secuencia de aminoácidos de la variante. De forma específica, tales variantes de pectato liasa tienen una secuencia de aminoácidos que no se encuentra en la naturaleza, que se deriva por sustitución de una pluralidad de restos de aminoácidos de una pectato liasa precursora con diferentes aminoácidos. La pectato liasa precursora puede ser una pectato liasa de origen natural o una pectato liasa recombinante. Las variantes de pectato liasa útiles incluyen la sustitución de cualquiera de los L-aminoácidos de origen natural en las posiciones designadas de los restos de aminoácidos.

Variantes de mutagénesis de sustitución de sitio genético (GSSM™)

La invención proporciona variantes de pectato liasa y los ácidos nucleicos que las codifican. En un aspecto, la invención proporciona la SEC ID N°: 134, codificada por la SEC ID N°: 133, respectivamente. La SEC ID N°: 133 es una variante de ácido nucleico generada por mutagénesis por saturación de sitio genético (GSSM™) de la SEC ID N°: 131 (que codifica la SEC ID N°: 132). La SEC ID N°: 131 y la SEC ID N°: 132 son variaciones truncadas del ácido nucleico como se establece en la SEC ID N°: 77, que codifica la SEC ID N°: 78, respectivamente. La siguiente Tabla 1 resume los cambios de aminoácidos resultantes de las variaciones generadas por GSSM™ en sus respectivos ácidos nucleicos de codificación (la SEC ID N°: 78 de longitud completa codificada por la SEC ID N°: 77, y la SEC ID N°: 132 "precursora" truncada codificada por la SEC ID N°: 131:

Tabla 1 posición del aminoácido que incluye mutación en las SEC ID N°s: 131, 132	Posición del nucleótido en el gen de tipo silvestre truncado (SEC ID N°s: 131, 132)	Posición del nucleótido en el gen de tipo silvestre de longitud completa (SEC ID N°s: 77, 78)	Posición del aminoácido en el gen de tipo silvestre de longitud completa (SEC ID N°s: 77, 78)
A118H	352-354	1423-1425	475
A182V	544-546	1615-1617	539
T190L	568-570	1639-1641	547
A197G	589-591	1660-1662	554

Tabla 1 posición del aminoácido que incluye mutación en las SEC ID N ^{os} : 131, 132	Posición del nucleótido en el gen de tipo silvestre truncado (SEC ID N ^{os} : 131, 132)	Posición del nucleótido en el gen de tipo silvestre de longitud completa (SEC ID N ^{os} : 77, 78)	Posición del aminoácido en el gen de tipo silvestre de longitud completa (SEC ID N ^{os} : 77, 78)
S208K	622-624	1693-1695	565
T219M	655-657	1726-1728	576
T223E	667-669	1738-1740	580
S255R	763-765	1834-1836	612
S263K	787-789	1858-1860	620
N275Y	823-825	1894-1896	632
Y309W	925-927	1996-1998	666
S312V	934-936	2005-2007	669

La Figura 6 es una tabla que resume cambios en la secuencia a modo de ejemplo en los polipéptidos pectato liasa de la invención, caracterizados como "mutantes positivos". Los mutantes positivos identificados como A-S son mutantes positivos conminatorios (cada uno tiene varios cambios generados por GSSM™). Los mutantes positivos identificados como AA-LL son mutantes positivos individuales (cada uno con un cambio generado por GSSM™).

La Figura 7 es una tabla que resume temperaturas de fusión y actividades específicas a modo de ejemplo (SA) de enzimas a modo de ejemplo de la invención a diversas temperaturas. La actividad específica (U/mg de enzima pura) se midió a diferentes temperaturas a pH 9,5 en Glicina NaOH 25 mM, tampón Tris HCl 25 mM. Una unidad de actividad enzimática se definió como la cantidad de enzima que producía 1 μ mol de oligogalacturónidos insaturados equivalente a 1 μ mol de digalacturónido insaturado por minuto. Las concentraciones de proteína de las preparaciones de enzimas puras se midieron a A280 usando un coeficiente de extinción molar de 73800 M⁻¹cm⁻¹ (1 A280 equiv. para 0,50 mg/ml). Las temperaturas de fusión se determinaron con un calorímetro de barrido diferencial.

En estas Figuras, el mutante "N" tiene una secuencia como se establece en la SEC ID N^o: 134, codificada por la SEC ID N^o: 133.

Secuencias señal de pectato liasa, dominios de pectina metil esterasa y dominios catalíticos, módulos de unión a carbohidratos y dominios prepro

La divulgación proporciona secuencias señal (por ejemplo, péptidos señal (SP)), dominios prepro y dominios catalíticos (CD). Los SP, dominios prepro y/o CD pueden ser péptidos aislados o recombinantes o pueden formar parte de una proteína de fusión, por ejemplo, como un dominio heterólogo en una proteína quimérica. La divulgación proporciona ácidos nucleicos que codifican estos dominios catalíticos (CD), dominios prepro y secuencias señal (SP, por ejemplo, un péptido que tiene una secuencia que comprende/que consiste en restos terminales amino de un polipéptido de la invención).

La divulgación proporciona secuencias señal de pectato liasa (por ejemplo, péptidos señal (SP)) y ácidos nucleicos que codifican estas secuencias señal, por ejemplo, un péptido que tiene una secuencia que comprende/que consiste en restos terminales amino de un polipéptido de la invención, por ejemplo, péptidos señal (SP) como se establece en la Tabla 2, que sigue a continuación. Una secuencia señal puede comprender un péptido que comprende/que consiste en una secuencia como se establece en los restos de 1 a 15, de 1 a 16, de 1 a 17, de 1 a 18, de 1 a 19, de 1 a 20, de 1 a 21, de 1 a 22, de 1 a 23, de 1 a 24, de 1 a 25, de 1 a 26, de 1 a 27, de 1 a 28, de 1 a 28, 1 a 30, 1 a 31, 1 a 32, 1 a 33, 1 a 34, 1 a 35, 1 a 36, 1 a 37, 1 a 38, 1 a 39, de 1 a 40, de 1 a 41, de 1 a 42, de 1 a 43, de 1 a 44 de un polipéptido de la divulgación, por ejemplo, SEC ID N^o: 2, SEC ID N^o: 4, SEC ID N^o: 6, SEC ID N^o: 8, SEC ID N^o: 10, SEC ID N^o: 12, SEC ID N^o: 14, SEC ID N^o: 16, SEC ID N^o: 18, SEC ID N^o: 20, SEC ID N^o: 22, SEC ID N^o: 24, SEC ID N^o: 26, SEC ID N^o: 28, SEC ID N^o: 30, SEC ID N^o: 32, SEC ID N^o: 34, SEC ID N^o: 36, SEC ID N^o: 38, SEC ID N^o: 40, SEC ID N^o: 42, SEC ID N^o: 44, SEC ID N^o: 46, SEC ID N^o: 48, SEC ID N^o: 50, SEC ID N^o: 52, SEC ID N^o: 54, SEC ID N^o: 56, SEC ID N^o: 58, SEC ID N^o: 60, SEC ID N^o: 62, SEC ID N^o: 64, SEC ID N^o: 66, SEC ID N^o: 68, SEC ID N^o: 70, SEC ID N^o: 72, SEC ID N^o: 74, SEC ID N^o: 76, SEC ID N^o: 78, SEC ID N^o: 80, SEC ID N^o: 82, SEC ID N^o: 84, SEC ID N^o: 86, SEC ID N^o: 88, SEC ID N^o: 90, SEC ID N^o: 92, SEC ID N^o: 94, SEC ID N^o: 96, SEC ID N^o: 98, SEC ID N^o: 100, SEC ID N^o: 102, SEC ID N^o: 104, SEC ID N^o: 106, SEC ID N^o: 108, SEC ID N^o: 110, SEC ID N^o: 112, SEC ID N^o: 114, SEC ID N^o: 116, SEC ID N^o: 118, SEC ID N^o: 120, SEC ID N^o: 122, SEC ID N^o: 124, SEC ID N^o: 126, SEC ID N^o: 128, SEC ID N^o: 130, SEC ID N^o: 132, SEC ID N^o: 134.

La invención también proporciona dominios de pectato liasa pectina metil esterasa (PED) y dominios catalíticos (CD) como se establece en la Tabla 2, que sigue a continuación.

Las secuencias señal de pectato liasa (SP), CD, y/o secuencias prepro pueden ser péptidos aislados, o, secuencias unidas a otra hidrolasa o un polipéptido que no es pectato liasa, por ejemplo, como una proteína de fusión (química). Algunos polipéptidos pueden comprender tales secuencias señal de pectato liasa. Algunos polipéptidos pueden comprender tales secuencias señal de pectato liasa SP, CD, y/o prepro y comprenden secuencias heterólogas para pectato liasas de la invención (por ejemplo, una proteína de fusión que comprende una SP, CD, y/o prepro del polipéptido de la invención y secuencias de otra pectato liasa o una proteína que no es pectato liasa). En un aspecto, la invención proporciona pectato liasas de la invención con secuencias SP, CD, y/o prepro heterólogas, por ejemplo, secuencias con una secuencia señal de levadura. Una pectato liasa de la invención puede comprender un SP heterólogo y/o prepro en un vector, por ejemplo, un vector de la serie pPIC (Invitrogen, Carlsbad, CA).

La Tabla 2 resume secuencias señal (es decir, señales de péptidos en su forma aislada), dominios catalíticos, módulos de unión a carbohidratos y dominios de pectina metil esterasa. Por ejemplo, la Tabla 2 describe: en la fila 1, un péptido señal (SP) en los restos de 1 a 28 de la SEC ID N°: 102 (codificados por la SEC ID N°: 101) y un dominio catalítico (CD) en los restos 78-459 de la SEC ID N°: 102; en la fila 2, un péptido señal (SP) en los restos de 1 a 21 de la SEC ID N°: 2 (codificados por la SEC ID N°: 1), un dominio de pectina metil esterasa (PED) en los restos 28-308, y un dominio catalítico (CD) en los restos 309-638; en la fila 3, etc.

Tabla 2	
SEC ID N°:	Módulos (SP = péptido señal, CD = dominio catalítico, CBM = módulo de unión a carbohidrato, PED = dominio de pectina metil esterasa)
101, 102	SP; 1-28, CD; 78-459
1, 2	SP; 1-21, PED; 28-308, CD; 309-638
103, 104	SP; 1-26, CD; 27-366
105, 106	SP; 1-43, CD; 44-400
107, 108	SP; 1-31, CD; 32-357
109, 110	SP; 1-21, PED; 28-308, CD; 309-637
11, 12	CD; 1-388
111, 112	SP; 1-27, CD; 82-461
113, 114	SP; 1-18, CD; 19-388
115, 116	CD; 1-331
117, 118	SP; 1-24, CD; 25-574
119, 120	CBM; 1-61, CBM; 134-257, CD; 258-615
121, 122	SP; 1-29, CD; 30-348
123, 124	SP; 1-21, CD; 22-390
125, 126	CD; 24-325
127, 128	SP; 1-24, CD; 125-482
129, 130	CD; 38-326
13, 14	SP; 1-22, CD; 23-354
15, 16	SP; 1-33, CD; 34-359
17, 18	CD; 1-348
19, 20	CD; 1-373
21, 22	SP; 1-23, CD; 24-422
23, 24	SP; 1-18, CD; 19-393
25, 26	SP; 1-15, CD; 16-397
27, 28	SP; 1-21, PED; 28-308, CD; 309-638
29, 30	SP; 1-27, CD; 77-459
3, 4	SP; 1-28, CD; 81-476
31, 32	CD; 1-348
33, 34	SP; 1-18, CD; 19-346
35, 36	CD; 1-356
37, 38	SP; 1-35, CD; 36-387
39, 40	SP; 1-32, CD; 33-358
41, 42	SP; 1-21, CD; 22-359
43, 44	CBM; 4-89, CBM; 152-275, CD; 277-633
45, 46	SP; 1-20, CD; 21-328
47, 48	SP; 1-21, CD; 22-358
49, 50	SP; 1-16, CD; 17-340
5, 6	CD; 1-358
51, 52	CD; 1-376
53, 54	SP; 1-31, CBM; 32-124, CBM; 180-303, CD; 304-658
55, 56	CD; 1-374
57, 58	CD; 1-389
59, 60	SP; 1-24, CD; 25-359
61, 62	CD; 90-407
63, 64	SP; 1-16, CD; 17-340
65, 66	SP; 1-28, CD; 29-436

SEC ID N°:	Módulos (SP = péptido señal, CD = dominio catalítico, CBM = módulo de unión a carbohidrato, PED = dominio de pectina metil esterasa)
67, 68	SP; 1-32, CBM; 33-126, CBM; 184-307, CD; 308-664
69, 70	SP; 1-22, CD; 23-344
7, 8	CD; 1-374
71, 72	SP; 1-20, CD; 21-345
73, 74	SP; 1-22, CD; 23-406
75, 76	SP; 1-34, CD; 110-555
77, 78	SP; 1-33, CBM; 34-126, CBM; 199-322, CD; 323-680
79, 80	SP; 1-28, CD; 81-458
81, 82	SP; 1-30, CD; 31-354
83, 84	PED; 268-556, CD; 782-1164
85, 86	CD; 1-383
87, 88	SP; 1-32, CD; 33-375
89, 90	SP; 1-31, CD; 32-459
9, 10	SP; 1-29, CD; 30-371
91, 92	CD; 1-374
93, 94	CD; 1-353
95, 96	SP; 1-31, CD; 32-357
97, 98	PED; 45-333, CD; 336-698
99, 100	SP; 1-35, CD; 36-593

En un aspecto, los SP, CD, y/o las secuencias prepro se identifican después de la identificación de nuevos polipéptidos pectato liasa. Las rutas mediante las que las proteínas se clasifican y se transportan a su propia ubicación celular a menudo se denominan rutas de dirección de proteína. Uno de los elementos más importantes de todos estos sistemas de dirección es una secuencia corta de aminoácidos en el extremo amino de un polipéptido recién sintetizado denominado la secuencia señal. Esta secuencia señal dirige una proteína a su ubicación apropiada en la célula y se retira durante del transporte o cuando la proteína alcanza su destino final. La mayoría de las proteínas lisosomales, de membrana, o secretadas tienen una secuencia señal amino-terminal que las marcas para translocación en el lumen del retículo endoplasmático. La longitud de las secuencias señal puede variar de 13 a 45 o más restos de aminoácidos. Los expertos en la materia conocen diversos métodos de reconocimiento de secuencias señal. Por ejemplo, en un aspecto, se identifican nuevas señales de péptidos señal de pectato liasa mediante un método denominado SignalP. El SignalP usa una red neural combinada que reconoce tanto péptidos señal como sus sitios de escisión. (Nielsen, *et al.*, " Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites". Protein Engineering, vol. 10, no. 1, p. 1-6 (1997).

En algunos aspectos, algunas pectato liasas de la invención no tienen SP y/o secuencias prepro, y/o dominios catalíticos (CD). En un aspecto, la invención proporciona polipéptidos (por ejemplo, pectato liasas) que carecen de todo o parte de un SP, un CD y/o un dominio prepro. En un aspecto, la invención proporciona secuencias de ácidos nucleicos que codifican la secuencia señal (SP), un CD, y/o prepro de una pectato liasa unida de forma operativa a secuencias de ácidos nucleicos de una pectato liasa diferente u, opcionalmente, puede ser deseable una secuencia señal (SP) y/o dominio prepro de una proteína que no es pectato liasa.

La invención también proporciona polipéptidos aislados o recombinantes que comprenden secuencias señal (SP), dominios prepro, dominios de pectina metil esterasa (PED) y dominios catalíticos (CD) de la invención y secuencias heterólogas. Las secuencias heterólogas son secuencias no asociadas de forma natural (por ejemplo, a una pectato liasa) con una SP, dominio prepro, PED, y/o CD. La secuencia a la que la SP, dominios prepro, PED y/o CD no se asocian de forma natural puede estar en el extremo amino terminal, extremo carboxi terminal, y/o en ambos extremos de la SP, del dominio prepro, de PED, y/o de CD de la SP, dominio prepro, PED y/o CD. En un aspecto, la invención proporciona un polipéptido aislado o recombinante que comprende (o que consiste en) un polipéptido que comprende una secuencia señal (SP), dominio prepro, dominio de pectina metil esterasa (PED) y/o dominio catalítico (CD) de la invención con la condición de que no esté asociado con ninguna secuencia con la que se asocia de forma natural (por ejemplo, una secuencia de pectato liasa). De forma análoga en un aspecto, la invención proporciona ácidos nucleicos aislados o recombinantes que codifican estos polipéptidos. Por lo tanto, en un aspecto, el ácido nucleico aislado o recombinante de la invención comprende secuencia de codificación para una secuencia señal (SP), dominio prepro, dominio de pectina metil esterasa (PED) y/o dominio catalítico (CD) de la invención y una secuencia heteróloga (es decir, una secuencia no asociada de forma natural con la secuencia señal (SP), dominio prepro, dominio de pectina metil esterasa (PED) y/o dominio catalítico (CD) de la invención). La secuencia heteróloga puede estar en el extremo terminal en la posición 3', en el extremo terminal en la posición 5', y/o en ambos extremos de la secuencia de codificación de SP, dominio prepro, PED y/o CD.

Glicosilación

Los péptidos y polipéptidos de la invención (por ejemplo, pectato liasas, anticuerpos) también pueden estar glicosilados, por ejemplo, en un aspecto, comprendiendo al menos un sitio de glicosilación, por ejemplo, una glicosilación unida a N o unida a O. En un aspecto, el polipéptido se puede glicosilar después de su expresión en una cepa de *P. pastoris* o de *S. pombe*. La glicosilación se puede añadir después de la traducción por vía química o

mediante cualquier mecanismo de biosíntesis celular, en la que el último incorpora el uso de motivos de glicosilación conocidos, que pueden ser nativos para la secuencia o se pueden añadir como un péptido o se pueden añadir en la secuencia de codificación de ácido nucleico.

5 Bibliotecas de pectato liasas polipéptidos híbridos (quiméricos)

En un aspecto, la invención proporciona pectato liasas y proteínas de fusión híbridas, que incluyen bibliotecas de péptidos, que comprenden secuencias de la invención. Las bibliotecas de péptidos de la invención se pueden usar para aislar moduladores de péptidos (por ejemplo, activadores o inhibidores) de dianas, tales como sustratos de pectato liasa, receptores, enzimas. Las bibliotecas de péptidos de la invención se pueden usar para identificar asociados de unión de dianas adecuados, tales como ligandos, por ejemplo, citoquinas, hormonas y similares. En un aspecto, la invención proporciona proteínas quiméricas que comprenden una secuencia señal (SP), dominio de pectina metil esterasa (PED) y/o dominio catalítico (CD) del polipéptido de la invención y una secuencia heteróloga (véase anteriormente).

En un aspecto, las proteínas de fusión de la invención (por ejemplo, el resto peptídico) se estabilizan de forma convencional (con respecto a péptidos lineales) para permitir una afinidad de unión más elevada para las dianas. La invención proporciona fusiones de pectato liasas de la invención y otros péptidos, que incluyen péptidos conocidos y aleatorios. Se pueden fusionar un modo tal que la estructura de las pectato liasas no se ve alterada de forma significativa y el péptido se estabiliza conformacionalmente de forma metabólica o estructural. Esto permite la creación de una biblioteca de péptidos que se controla fácilmente tanto para su presencia como su cantidad dentro de las células.

Las variantes de secuencias de aminoácidos de la invención se pueden caracterizar mediante una naturaleza de la variación determinada previamente, una característica que las diferencia de una forma de origen natural, por ejemplo, una variación alélica o interespecie de una secuencia de pectato liasa. En un aspecto, las variantes de la invención presentan la misma actividad biológica cualitativa que el análogo de origen natural. Como alternativa, las variantes se pueden seleccionar para que tengan características modificadas. En un aspecto, cuando el sitio o región para la introducción de una variación de la secuencia de aminoácidos se determina previamente, no es necesario que la mutación, *per se*, se determine previamente. Por ejemplo, para optimizar el rendimiento de una mutación en un sitio dado, se puede realizar mutagénesis aleatoria en el codón o región diana y las variantes de pectato liasa expresadas se identifiquen sistemáticamente para la combinación de actividad deseada óptima. Se conocen bien algunas técnicas para preparar mutaciones de sustitución en sitios determinados previamente en el ADN que tiene la secuencia conocida, como se analiza en el presente documento, por ejemplo, mutagénesis del cebador M13 y mutagénesis de PCR. La identificación sistemática de los mutantes se puede realizar usando ensayos de actividades proteolíticas. En aspectos alternativos, las sustituciones de aminoácidos pueden ser restos individuales; las inserciones pueden ser del orden de aproximadamente de 1 a 20 aminoácidos, aunque se pueden realizar inserciones considerablemente mayores. Las supresiones pueden variar de aproximadamente 1 a aproximadamente 20, 30, 40, 50, 60, 70 restos o más. Para obtener un derivado final con las propiedades óptimas, se pueden usar sustituciones, supresiones, inserciones o cualquier combinación de las mismas. Generalmente, estos cambios se realizan en unos pocos aminoácidos para minimizar la alteración de la molécula. Sin embargo, se pueden tolerar cambios mayores en ciertas circunstancias.

La invención proporciona pectato liasas en las que la estructura de la cadena principal de polipéptido, la estructura secundaria o la estructura terciaria, por ejemplo, una estructura alfa-helicoidal o de lámina beta, se ha modificado. En un aspecto, se ha modificado la carga o hidrofobia. En un aspecto, se ha modificado el tamaño de una cadena lateral. Se realizan cambios sustanciales en la función o identidad inmunológica seleccionando sustituciones que son menos conservativas. Por ejemplo, se pueden realizar sustituciones que influyen de forma más significativa: la estructura de la cadena principal de polipéptido en la zona de la alteración, por ejemplo una estructura alfa-helicoidal o una lámina beta; una carta o un sitio hidrófobo de la molécula, que pueden estar en un sitio activo; o una cadena lateral. La invención proporciona sustituciones en polipéptidos de la invención en los que (a) un resto hidrófilo, por ejemplo serilo o treonilo, se sustituye por (o con) un resto hidrófobo, por ejemplo leucilo, isoleucilo, fenilalanilo, valilo o alanilo; (b) una cisteína o prolina se sustituye por (o con) cualquier otro resto; (c) un resto que tiene una cadena lateral electropositiva, por ejemplo lisilo, arginilo, o histidilo, se sustituye por (o con) un resto electronegativo, por ejemplo glutamilo o aspartilo; o (d) un resto que tiene una cadena lateral grande, por ejemplo fenilalanina, se sustituye por (o con) una que no tiene una cadena lateral, por ejemplo glicina. Las variantes pueden presentar la misma actividad biológica cualitativa (es decir, actividad de pectato liasa) aunque se pueden seleccionar variantes para modificar las características de las pectato liasas cuando sea necesario.

En un aspecto, las pectato liasas de la invención comprenden epítomos o marcas de purificación, secuencias señal u otras secuencias de fusión, etc. En un aspecto, las pectato liasas de la invención se pueden fusionar con un péptido aleatorio para formar un polipéptido de fusión. En el presente documento, con "fusionado" o "unido de forma operativa" se hace referencia a que el péptido aleatorio y la pectato liasa se unen en conjunto, de un modo tal que se minimiza la alteración de la estabilidad de la estructura de la pectato liasa, por ejemplo, mantiene la actividad de pectato liasa. El polipéptido de fusión (o polinucleótido de fusión que codifica el polipéptido de fusión) también puede comprender adiciones de componentes, que incluyen múltiples péptidos en múltiples bucles.

En un aspecto, los péptidos y los ácidos nucleicos que los codifican son aleatorios, ya sea totalmente aleatorios o están sesgados en su aleatorización, por ejemplo en la secuencia de nucleótidos/restos generalmente o por posición. "Aleatorio" se refiere a que cada ácido nucleico y péptido consiste básicamente en nucleótidos y aminoácidos aleatorios, respectivamente. En un aspecto, los ácidos nucleicos que dan lugar a los péptidos se pueden sintetizar de forma química, y por lo tanto pueden incorporar cualquier nucleótido en cualquier posición. Por lo tanto, cuando los ácidos nucleicos se expresan para formar péptidos, se puede incorporar cualquier resto de aminoácido en cualquier posición. El proceso de síntesis se puede diseñar para generar ácidos nucleicos aleatorios, para permitir la formación de todas o la mayoría de las posibles combinaciones en la longitud del ácido nucleico, formando de este modo una biblioteca de ácidos nucleicos aleatorios. La biblioteca puede proporcionar una población estructuralmente diversa de forma suficiente de productos de expresión aleatorios para influir en un intervalo probabilísticamente suficiente de respuestas celulares para proporcionar una o más células que presentan una respuesta deseada. Una biblioteca de interacción puede ser lo suficientemente grande como para que al menos uno de sus miembros tenga una estructura que le proporciona afinidad hacia alguna molécula, proteína, u otro factor.

Metodologías de Identificación Sistemática y Dispositivos de Control "En línea"

En la práctica de estos métodos, se puede usar una diversidad de aparatos y metodologías en conjunto con los polipéptidos y ácidos nucleicos de la invención, por ejemplo, para identificar sistemáticamente la actividad de pectato liasa, para identificar sistemáticamente compuestos como moduladores potenciales, por ejemplo, activadores o inhibidores, de una actividad de pectato liasa, para anticuerpos que se unen a un polipéptido de la invención, para ácidos nucleicos que hibridan a un ácido nucleico de la invención, para identificar sistemáticamente células que expresan un polipéptido de la invención y similares.

Matrices Capilares

Algunas matrices capilares, tales como la GIGAMATRIX™, Diversa Corporation, San Diego, CA, se pueden usar en los métodos de la invención. Los ácidos nucleicos o polipéptidos de la invención se pueden inmovilizar en o aplicar a una matriz, incluyendo matrices capilares. Las matrices se pueden usar para identificar sistemáticamente o para controlar bibliotecas de composiciones (por ejemplo, moléculas pequeñas, anticuerpos, ácidos nucleicos, etc.) para su capacidad para unirse al cu para modular la actividad de un ácido nucleico o un polipéptido de la invención. Las matrices capilares proporcionan otro sistema para mantener e identificar sistemáticamente muestras. Por ejemplo, un aparato de identificación sistemática de muestras puede incluir una pluralidad de capilares formados en una matriz de capilares adyacentes, en la que cada capilar comprende al menos una pared que define un lumen capilar que comprende al menos una pared que define un lumen para retener una muestra. El aparato puede incluir adicionalmente material intersticial colocado entre capilares adyacentes en la matriz, y uno o más vicios de referencia formados dentro del material intersticial. Un capilar para identificar sistemáticamente una muestra, en la que el capilar se adapta para que se una en una matriz de capilares, puede incluir una primera pared que define un lumen para retener la muestra, un ataque formado por un material de filtro, para filtrar la energía de excitación proporcionada al lumen para excitar la muestra.

Un polipéptido o ácido nucleico, por ejemplo, un ligando, se puede introducir en un primer componente en al menos una parte de un capilar de una matriz capilar. Cada capilar de la matriz capilar puede comprender al menos una pared que define un lumen para retener el primer componente. Una burbuja de aire se puede introducir en el capilar detrás del primer componente. Un segundo componente se puede introducir en el capilar, en el que el segundo componente se separa del primer componente mediante la burbuja de aire. Una muestra de interés se puede introducir en un primer líquido marcado con una partícula detectable en un capilar de una matriz capilar, en el que cada capilar de la matriz capilar comprende al menos una pared que define un lumen para retener el primer líquido y la partícula detectable, y en el que al menos una prensa reviste con material de unión para unir la partícula detectable a la han menos una parte. El método puede incluir adicionalmente la eliminación del primer líquido del tubo capilar, en el que la partícula detectable unida se mantiene dentro del capilar, y la introducción de un segundo líquido en el tubo capilar.

La matriz capilar puede incluir una pluralidad de capilares individuales que comprenden al menos una pared externa que define un lumen. La pared externa del capilar puede ser una o más paredes fusionadas en conjunto. De forma análoga, la pared de definir un lumen que es cilíndrico, cuadrado, hexagonal o de cualquier otra forma geométrica siempre y cuando las paredes formen un lumen para la retención de un líquido o muestra. Los capilares de la matriz capilar se pueden mantener en conjunto en las proximidades para formar una estructura plana. Los capilares se pueden unir en conjunto, mediante su fusión (por ejemplo, en el que los capilares están hechos de vidrio), pegado, unión, o sujetos lado a lado. La matriz capilar puede estar formada con cualquier número de capilares individuales, por ejemplo, una serie de 100 a 4.000.000 capilares. Una matriz capilar puede formar una placa de microtitulación que tiene aproximadamente 100.000 o más capilares individuales unidos en conjunto.

Matrices, o "Biochips"

Los ácidos nucleicos o polipéptidos de la invención se pueden inmovilizar en o aplicar a una matriz. Las matrices se pueden usar para identificar sistemáticamente o para controlar bibliotecas de composiciones (por ejemplo, moléculas pequeñas, anticuerpos, ácido sulfúrico, etc.) para su capacidad para unirse o para modular la actividad de un ácido nucleico o un polipéptido de la invención. Por ejemplo, en un aspecto de la invención, un parámetro controlado es la expresión del transcrito de un gen de pectato liasa. Uno o más, o, todos los transcritos de una célula se pueden medir mediante hibridación de una muestra que comprende transcritos de la célula, o, ácidos nucleicos representativos o complementarios de transcritos de una célula, mediante hibridación con ácidos nucleicos inmovilizados en una matriz, o "biochip". Mediante el uso de una "matriz" de ácidos nucleicos en un microchip, algunos o todos los transcritos de una célula se pueden cuantificar de forma simultánea. Como alternativa, también se pueden usar matrices que comprenden ácido nucleico genómico para determinar el genotipo de una cepa recién modificada por ingeniería preparada con los métodos que se desvelan en el presente documento. También se pueden usar "matrices de polipéptidos" para cuantificar de forma simultánea una pluralidad de proteínas. La presente invención se puede poner en práctica con cualquier "matriz" conocida, también denominada una "micromatriz" o "matriz de ácido nucleico" o "matriz polipeptídica" o "matriz de anticuerpo" o "biochip", o variación de las mismas. Las matrices son genéricamente una pluralidad de "sitios" o "elementos diana", comprendiendo cada elemento diana una cantidad definida de una o más moléculas biológicas, por ejemplo, oligonucleótidos, inmovilizadas en una zona definida de una superficie del sustrato para unión específica a una molécula de nuestra, por ejemplo, transcritos de ARNm.

En la práctica de los métodos, se puede incorporar cualquier matriz y/o método conocidos para preparar y usar matrices total o parcialmente, o variaciones de los mismos, como se describe, por ejemplo, en los documentos de Patente de Estados Unidos N^{os} 6.277.628; 6.277.489; 6.261.776; 6.258.606; 6.054.270; 6.048.695; 6.045.996; 6.022.963; 6.013.440; 5.965.452; 5.959.098; 5.856.174; 5.830.645; 5.770.456; 5.632.957; 5.556.752; 5.143.854; 5.807.522; 5.800.992; 5.744.305; 5.700.637; 5.556.752; 5.434.049; véase también, por ejemplo, el documento de patente WO 99/51773; documento de patente WO 99/09217; documento de patente WO 97/46313; documento de patente WO 96/17958; véase también, por ejemplo, Johnston (1998) Curr. Biol. 8: R171-R174; Schummer (1997) Biotechniques 23: 1087-1092; Kern (1997) Biotechniques 23: 120-124; Solinas-Toldo (1997) Genes, Chromosomes & Cancer 20: 399-407; Bowtell (1999) Nature Genetics Supp. 21: 25-32. Véanse también las solicitudes de Patente de Estados Unidos publicadas N^{os} 20010018642; 20010019827; 20010016322; 20010014449; 20010014448; 20010012537; 20010008765,

Métodos de identificación sistemática de anticuerpos y basados en anticuerpos

La divulgación proporciona anticuerpos aislados o recombinantes que se unen de forma específica a una pectato liasa de la invención. Estos anticuerpos se pueden usar para aislar, identificar o cuantificar las pectato liasas de la invención o polipéptidos relacionados. Estos anticuerpos pueden usar para aislar otros polipéptidos dentro del alcance la invención u pectato liasas otras relacionadas. Los anticuerpos se pueden diseñar para que se unan a un sitio activo de una pectato liasa. Algunos métodos para la inhibición de pectato liasas pueden usar estos anticuerpos.

La divulgación proporciona fragmentos de las enzimas de la invención, que incluyen fragmentos inmunogénicos de un polipéptido de la invención que pueden comprender adicionalmente adyuvantes, vehículos y similares.

Los anticuerpos se pueden usar en inmunoprecipitación, tinción, columnas de inmutafinidad, y similares. Si se desea, se pueden generar secuencias de ácidos nucleicos que codifican antígenos específicos mediante inmunización seguido de aislamiento del polipéptido o ácido nucleico, amplificación o clonación e inmovilización de polipéptidos en una matriz. Como alternativa, estos métodos se pueden usar para modificar la estructura de un anticuerpo producido por una célula a modificar, por ejemplo, una afinidad de un anticuerpo se puede aumentar o disminuir. Además, la capacidad para preparar o modificar anticuerpos puede ser un fenotipo modificado por ingeniería en una célula por estos métodos.

Los expertos en la materia conocen algunos métodos de inmunización, producción y aislamiento de anticuerpos (policlonales y monoclonales) y se describen en la bibliografía científica y de patente, véase, por ejemplo, Coligan, CURRENT PROTOCOLS IN IMMUNOLOGY, Wiley/Greene, NY (1991); Stites (eds.) BASIC AND CLINICAL IMMUNOLOGY (7^a ed.) Lange Medical Publications, Los Altos, CA ("Stites"); Goding, MONOCLONAL ANTIBODIES: PRINCIPLES AND PRACTICE (2^a ed.) Academic Press, New York, NY (1986); Kohler (1975) Nature 256: 495; Harlow (1988) ANTIBODIES, A LABORATORY MANUAL, Cold Spring Harbor Publications, New York. Los anticuerpos donde se pueden generar *in vitro*, por ejemplo, usando bibliotecas de presentación de fagos que expresan sitio de unión a anticuerpo recombinante, además de los métodos *in vivo* tradicionales usando animales. Véase, por ejemplo, Hoogenboom (1997) Trends Biotechnol. 15: 62-70; Katz (1997) Annu. Rev. Biophys. Biomol. Struct. 26: 27-45.

Se pueden usar polipéptidos o péptidos para generar anticuerpos que se unen de forma específica a los polipéptidos, por ejemplo, las pectato liasas, de la invención. Los anticuerpo resultante se pueden usar en procedimientos de cromatografía por afinidad para aislar o purificar el polipéptido o para determinar si el polipéptido está presente en una muestra biológica. En tales procedimientos, una preparación de proteína, tal como un extracto,

o una muestra biológica se pone en contacto con un anticuerpo capaz de unirse de forma específica uno de los polipéptidos de la invención.

En los procedimientos de inmunoafinidad, el anticuerpo se une a soporte sólido, tal como una perla u otra matriz de columna. La preparación de proteínas se coloca en contacto con el anticuerpo en condiciones en las que el anticuerpo se une de forma específica a uno de los polipéptidos de la invención. Después de un lavado para retirar las proteínas no unidas de forma específica, los polipéptidos unidos de forma específica se eluyen.

La capacidad de las proteínas en una muestra biológica para unirse al anticuerpo se puede determinar usando cualquiera de una diversidad de procedimientos familiares para los expertos en la materia. Por ejemplo, la unión se puede determinar marcando el anticuerpo con una marca detectable tal como un agente fluorescente, una marca enzimática, o un radioisótopo. Como alternativa, la unión del anticuerpo a la muestra se puede detectar usando un anticuerpo secundario que tiene tal marca detectable en el mismo. Algunos ensayos en particular incluyen ensayos de ELISA, ensayos de sándwich, radioinmunoensayos, y Transferencias de Western.

Se pueden obtener anticuerpos policlonales generados frente a los polipéptidos de la invención por inyección directa de los polipéptidos en un animal o mediante la administración de los polipéptidos a un animal no humano. El anticuerpo obtenido de este modo se unirá al polipéptido en sí mismo. De esta manera, se puede usar incluso una secuencia que codifica solamente un fragmento del polipéptido para generar anticuerpos que se pueden unir a todo el polipéptido nativo. Tales anticuerpos se pueden unir a continuación para aislar el polipéptido de células que expresan ese polipéptido.

Para la preparación de anticuerpos monoclonales, se puede usar cualquier técnica que proporciona anticuerpos producidos mediante cultivos continuos de líneas celulares. Los ejemplos incluyen la técnica del hibridoma, la técnica del trioma, la técnica del hibridoma de linfocitos B humanos, y la técnica del hibridoma de EBV (véase, por ejemplo, Cole (1985) en Monoclonal Antibodies and Cancer Therapy, Alan R. Liss, Inc., pp. 77-96).

Algunas técnicas descritas para la producción de anticuerpos de una sola cadena (véase, por ejemplo, el documento de Patente de Estados Unidos Nº 4.946.778) se pueden adaptar para producir anticuerpos de una sola cadena para los polipéptidos de la invención. Como alternativa, se pueden usar ratones transgénicos para expresar anticuerpos humanizados para estos polipéptidos o fragmentos.

Se pueden usar anticuerpos generados frente a los polipéptidos de la invención en la identificación sistemática de polipéptidos similares (por ejemplo, pectato liasas) a partir de otros organismos y muestras. En tales técnicas, los polipéptidos del organismo se ponen en contacto con el anticuerpo y se detectan esos polipéptidos que se unen de forma específica al anticuerpo. Se puede usar cualquiera de los procedimientos que se han descrito anteriormente para detectar la unión al anticuerpo.

Kits

Los kits pueden comprender las composiciones, por ejemplo, ácidos nucleicos, casetes de expresión, vectores, células, semillas o plantas o partes de plantas transgénicas, polipéptidos (por ejemplo, pectato liasas) y/o anticuerpos que se desvelan en el presente documento. Los kits también pueden contener material con instrucciones que enseña las metodologías y usos industriales de la invención, como se describe en el presente documento.

Medida de Parámetros Metabólicos

La evolución de la célula completa, o la modificación por ingeniería de la célula completa, de una célula puede desarrollar una nueva cepa celular que tiene un nuevo fenotipo, por ejemplo, una nueva o modificada actividad de pectato liasa, mediante la modificación de la composición genética de la célula. La composición genética se puede modificar mediante la adición a la célula de un ácido nucleico de la invención. Para detectar un nuevo fenotipo, al menos un parámetro metabólico de una célula modificada se controla en la célula en un marco de "tiempo real" o de tiempo "en línea". Una pluralidad de células, tal como un cultivo celular, se puede controlar en "tiempo real" o "en línea". Una pluralidad de parámetros metabólicos se puede controlar en "tiempo real" o "en línea". Algunos parámetros metabólicos se pueden controlar usando las pectato liasas de la invención.

El análisis de flujo metabólico (MFA) se basa en una región marco conservada de bioquímica conocida. Se construye una matriz metabólica linealmente independiente basándose en la ley de conservación de masa y en la hipótesis del estado pseudoestacionario (PSSH) en los metabolitos intracelulares. En la práctica de estos métodos, se establecen redes metabólicas, que incluyen la:

- Identidad de todos los sustratos, productos y metabolitos intermedios de la ruta
- Identidad de todas las reacciones químicas que interconvierten los metabolitos de la ruta, la estequiometría de las reacciones de la ruta,

- Identidad de todas las enzimas que catalizan las reacciones, la cinética de la reacción enzimática,
- Las interacciones reguladoras entre los componentes de la ruta, por ejemplo interacciones alostéricas, interacciones enzima-enzima etc,
- compartimentalización intracelular de enzimas o cualquier otra organización supramolecular de las enzimas, y,
- 5 • la presencia de cualquier gradiente de concentración de metabolitos, enzimas o moléculas efectoras o barreras de difusión para su movimiento.

Una vez que se construye la red metabólica de una cepa dada, se puede introducir una presentación matemática por concepto de matriz para estimar los flujos metabólicos intracelulares si están disponibles los datos de metaboloma en línea. El fenotipo metabólico se basa en los cambios de toda la red metabólica dentro de una célula. El fenotipo metabólico se basa en el cambio de utilización del uso de la ruta con respecto a condiciones ambientales, regulación genética, estado de desarrollo y el genotipo, etc. Después del cálculo de MFA en línea, el comportamiento dinámico de las células, su fenotipo y otras propiedades se analizan mediante la investigación del uso de la ruta. Por ejemplo, si el suministro de glucosa aumenta y el de oxígeno disminuye durante la fermentación de la levadura, el uso de las vías respiratorias se reducirá y/o detendrá, y el uso de las rutas fermentativas dominará. El control del estado fisiológico de los cultivos celulares será posible después del análisis de la ruta. Los métodos de la invención pueden ayudar a determinar cómo manipular la fermentación mediante la determinación de cómo cambiar el suministro de sustrato, temperatura, uso de inductores, etc., para controlar el estado fisiológico de las células para moverse a lo largo de la dirección deseable. Los resultados del MFA también se pueden comparar con datos de transcriptoma y proteoma para diseñar experimentos y protocolos para ingeniería metabólica o redistribución de genes, etc.

Cualquier fenotipo modificado o nuevo se puede consultar y detectar, incluyendo características nuevas o mejoradas en la célula. Se puede controlar cualquier aspecto del metabolismo o crecimiento.

25 *Control de la expresión de un transcrito de ARNm*

El fenotipo modificado por ingeniería puede comprender el aumento o una disminución de la expresión de un transcrito de ARNm (por ejemplo, un mensaje de pectato liasa) o generación de nuevos transcritos (por ejemplo, pectato liasa) en una célula. Este aumento o disminución de la expresión se puede trazar sometiendo a ensayo la presencia de una pectato liasa de la invención o mediante ensayos de actividad de pectato liasa. También se pueden detectar transcritos de ARNm, o mensajes, y cuantificar mediante cualquier método conocido en la técnica, que incluye, por ejemplo, transferencias de Northern, reacciones de amplificación cuantitativa, hibridación a matrices, y similares. Las reacciones de amplificación cuantitativa incluyen, por ejemplo, PCR cuantitativa, que incluye, por ejemplo, reacción en cadena de la polimerasa por transcripción inversa cuantitativa, o RT-PCR; RT-PCR en tiempo real cuantitativa, o "RT-PCR cinética en tiempo real" (véase, por ejemplo, Kreuzer (2001) Br. J. Haematol. 114: 313-318; Xia (2001) Transplantation 72: 907-914).

El fenotipo modificado por ingeniería se puede generar por genosupresión de la expresión de un gen homólogo. La secuencia que codifica el gen o uno o más elementos de control de la transcripción se pueden genosuprimir, por ejemplo, promotores o potenciadores. Por lo tanto, la expresión de un transcrito se puede suprimir completamente o solamente se puede disminuir.

El fenotipo modificado por ingeniería puede comprender el aumento de la expresión de un gen homólogo. Esto se puede realizar mediante genosupresión de un elemento de control negativo, incluyendo un elemento regulador de la transacción que actúa en cis o trans, o, en la mutagénesis de un elemento de control positivo. Uno o más, o, todos los transcritos de una célula se pueden medir por hibridación de una muestra que comprende transcritos de la célula, o, ácidos nucleicos representativos de o complementarios para transcritos de una célula, por hibridación con ácidos nucleicos inmovilizados en una materia.

50 *Control de la expresión de polipéptidos, péptidos y aminoácidos*

El fenotipo modificado por ingeniería puede comprender el aumento o la disminución de la expresión de un polipéptido (por ejemplo, una pectato liasa) o la generación de nuevos polipéptidos en una célula. Este aumento o disminución de la expresión se puede trazar mediante la determinación de la cantidad de pectato liasa presente o mediante ensayos de actividad de pectato liasa. También se pueden detectar y cuantificar polipéptidos, péptidos y aminoácidos mediante cualquier método conocido en la técnica, que incluye, por ejemplo, resonancia magnética nuclear (RMN), espectrofotometría, radiografía (radiomarcado de proteínas), electroforesis, electroforesis capilar, cromatografía líquida de alto rendimiento (HPLC), cromatografía en capa fina (TLC), cromatografía de hiperdifusión, diversos métodos inmunológicos, por ejemplo inmunoprecipitación, inmunodifusión, inmunoelectroforesis, radioinmunoensayos (RIA), ensayos de inmunoabsorción ligada a enzimas (ELISA), ensayos de inmunofluorescencia, electroforesis en gel (por ejemplo, SDS-PAGE), tinción con anticuerpos, clasificación celular activada por fluorescencia (FACS), espectrometría de masas por pirólisis, Espectrometría de Infrarrojos con Transformada de Fourier, espectrometría de Raman, GC-MS, y LC-Electronebulización y espectrometrías de masas con electronebulización de protección-LC-tándem, y similares. Las nuevas bioactividades también se pueden identificar sistemáticamente usando métodos, o variaciones de los mismos, que se describen en el documento de

Patente de Estados Unidos N° 6.057.103. Además, como se analiza a continuación en detalle, uno o más, o, todos los polipéptidos de una célula se pueden medir usando una matriz de proteína.

Aplicaciones industriales

5

Composiciones de Detergente

Las composiciones de detergente pueden comprender uno o más polipéptidos (por ejemplo, pectato liasas) de la invención. Para los métodos para preparar y usar composiciones detergente, véase, por ejemplo, los documentos de Patente de Estados Unidos N°s 6.413.928; 6.399.561; 6.365.561; 6.380.147. Las composiciones de detergente pueden ser una composición acuosa de una y dos partes, una composición líquida no acuosa, un sólido fundido, una forma granular, una forma de partículas, un comprimido formado por compresión, un gel y/o una pasta y la forma de suspensión. Las pectato liasas de la invención también se pueden usar como un producto aditivo para detergente en una forma sólida o una forma líquida. Se pretende que tales productos aditivos complementen o refuercen el rendimiento de composiciones de detergente convencionales y se pueden añadir en cualquier etapa del proceso de limpieza.

La invención también proporciona métodos capaces de eliminar tierra de alimentos recién recogidos, películas de restos de alimentos y otras composiciones de alimentos secundarias usando estas composiciones de detergente. Las pectato liasas de la invención pueden facilitar la eliminación de manchas de almidón por medio de hidrólisis catalítica o trans-eliminación de pectinas, incluyendo la alteración de paredes celulares de vegetales y bacterianas. Las pectato liasas de la invención se pueden usar en detergentes para lavavajillas en detergentes para lavado de textiles.

El contenido de enzima activa real depende del método de preparación de una composición de detergente y no es crítico, suponiendo que la solución de detergente tenga la actividad enzimática deseada. En un aspecto, la cantidad de pectato liasa presente en la solución final varía de aproximadamente 0,001 mg a 0,5 mg por gramo de la composición de detergente. La enzima en particular elegida para uso en el proceso y productos de la presente invención depende de las condiciones del uso final, incluyendo la forma física del producto, pH de uso, temperatura de un solo, y tipos de suelo a degradar o alterar. La enzima se puede elegir para que proporcione una actividad y estabilidad óptimas para cualquier conjunto de condiciones de utilidad dadas. En un aspecto, las pectato liasas de la presente invención son activas en los intervalos de pH de aproximadamente 4 a aproximadamente 12 y en el intervalo de temperatura de aproximadamente 20 °C a aproximadamente 95 °C. Los detergentes de la invención pueden comprender tensioactivos catiónicos, no iónicos semipolares o zwitteriónicos; o, mezclas de los mismos.

Las pectato liasas de la invención se pueden formular en detergentes en polvo y líquidos que tienen un pH entre 4,0 y 12,0 a niveles de aproximadamente un 0,01 % a aproximadamente un 5 % (preferentemente de un 0,1 % a un 0,5 %) en peso. Estas composiciones de detergente también pueden incluir otras enzimas tales como proteasas, celulasas, lipasas o endoglicosidasas, endo-beta-1,4-glucanasas, beta-glucanasas, endo-beta-1,3(4)-glucanasas, cutinasas, peroxidasas, lacasas, amilasas, glucoamilasas, pectinasas, reductasas, oxidasas, fenoloxidasas, ligninasas, pululaninas, arabinanasas, hemicelulasas, mananasas, xiloglucanasas, xilanasas, pectina acetil esterases, ramnogalacturonano acetil esterases, poligalacturonasas, ramnogalacturonasas, galactanasas, pectina liasas, pectina metilesterases, celobiohidrolasas y/o transglutaminasas. Estas composiciones de detergente también pueden incluir aditivos y estabilizantes.

La adición de las pectato liasas de la invención a composiciones de limpieza convencionales no crea ninguna limitación especial en uso. En otras palabras, cualquier temperatura y pH adecuados para el detergente también son adecuados para las composiciones de la invención siempre y cuando la enzima sea activa en o tolerante para el pH y/o temperatura del uso pretendido. Además, las pectato liasas de la invención se pueden usar en una composición de limpieza sin detergentes, de nuevo solas o en combinación con aditivos y estabilizantes.

La presente invención proporciona composiciones de limpieza que incluyen composiciones de detergente para la limpieza de superficies duras, composiciones de detergente para la limpieza de tejidos, composiciones para lavavajillas, composiciones para limpieza oral, composiciones para limpieza de dentaduras, y composiciones para limpieza de lentes de contacto.

En un aspecto, la invención proporciona un método para lavar un objeto que comprende poner en contacto el objeto con un polipéptido de la invención en condiciones suficientes para lavado. Una pectato liasa de la invención se puede incluir como un aditivo de detergente. La composición de detergente, por ejemplo, se puede formular como una composición de detergente lavado a mano o a máquina que comprende un polipéptido de la invención. Un aditivo para lavandería adecuado para tratamiento previo de tejidos manchados puede comprender un polipéptido de la invención. Una composición suavizante de tejido puede comprender una pectato liasa de la invención. Como alternativa, una pectato liasa de la invención se puede formular como una composición de detergente para uso en operaciones de limpieza de superficies duras del hogar en general. En aspectos alternativos, los aditivos detergentes y las composiciones de detergente pueden comprender una u otras enzimas más tales como una proteasa, una lipasa, una cutinasa, otra pectato liasa, una carbohidrasa, una celulasa, una pectinasa, una

mananasa, una arabinasa, una galactanasa, una xilanasa, una oxidasa, por ejemplo, una lactasa, y/o una peroxidasa (véase también, la mencionada anteriormente). Las propiedades de la enzima o enzimas de la invención se eligen para que sean compatibles con el detergente seccionado (es decir, pH óptimo, compatibilidad con otros ingredientes enzimáticos y no enzimáticos, etc.) y la enzima o enzimas están presentes en cantidades eficaces. En un aspecto, las enzimas de pectato liasa de la invención se usan para retirar materiales de tejidos con malos olores. Diversas composiciones de detergente y métodos para preparar las que se pueden usar en la práctica de la invención se describen, por ejemplo, en los documentos de Patente de Estados Unidos N^{os} 6.333.301; 6.329.333; 6.326.341; 6.297.038; 6.309.871; 6.204.232; 6.197.070; 5.856.164.

Cuando se formulan como composiciones adecuadas para su uso en un método de lavado de ropa a máquina, las pectato liasas de la invención pueden comprender un compuesto tanto tensioactivo como aditivo. Pueden comprender adicionalmente uno o más componentes detergentes, por ejemplo, compuestos poliméricos orgánicos, agentes de blanqueado, enzimas adicionales, supresores de espuma, agentes dispersantes, agentes dispersantes de cal-jabón, agentes de suspensión y agentes anti-redeposición de suciedad e inhibidores de la corrosión. Las composiciones de limpieza de la invención también pueden contener agentes suavizantes, como componentes adicionales de detergente. Tales composiciones que contienen carbohidrasa pueden proporcionar limpieza de tejidos, eliminación de manchas, mantenimiento de la blancura, suavizado, aspecto de color, inhibición de la transferencia del tinte y desinfección cuando se formulan como composiciones de detergente para lavado de ropa.

La densidad de las composiciones de detergente para lavado de ropa de la invención puede variar de aproximadamente 200 a 1500 g/litro, o, de aproximadamente 400 a 1200 g/litro, o, de aproximadamente 500 a 950 g/litro, o, de 600 a 800 g/litro, de composición; esto se puede medir a aproximadamente 20 °C.

La forma "compacta" de las composiciones de detergente para lavado de ropa se refleja de la mejor manera mediante la densidad y, en términos de composición, mediante la cantidad de sal de carga inorgánica. Las sales de carga inorgánica son ingredientes convencionales de las composiciones de detergente en forma de polvo. En las composiciones de detergente convencionales, las sales de carga están presentes en cantidades básicas, por lo general de un 17 % a un 35 % en peso de la composición total. En un aspecto de las composiciones compactas, la sal de carga está presente en cantidades que no superan un 15 % de la composición total, o, que no superan un 10 %, o, no superan un 5 % en peso de la composición. Las sales de carga inorgánica se pueden seleccionar a partir de las sales de metales alcalinos y alcalinotérreos de sulfatos y cloruros, por ejemplo, sulfato sódico.

Algunas composiciones de detergente líquido se puede presentar en una "forma concentrada". En un aspecto, Las composiciones de detergente líquido pueden contener una cantidad menor de agua, en comparación con los detergentes líquidos convencionales. En aspectos alternativos, el contenido de agua del detergente líquido concentrado es inferior a un 40 %, o, inferior a un 30 %, o, inferior a un 20 % en peso de la composición de detergente. Los compuestos de detergente de la invención pueden comprender formulaciones como se describe en el documento WO 97/01629.

Tratamiento de fibras y textiles

La invención proporciona métodos para el tratamiento de fibras, tejidos o cualquier material que comprende pectato o ácido poligalacturónico usando una o más pectato liasas de la invención. Las pectato liasas se pueden usar en cualquier método de tratamiento de fibra o tejido, que se conocen bien en la técnica, véanse, por ejemplo, los documentos de Patente de Estados Unidos N^{os} 6.261.828; 6.077.316; 6.024.766; 6.021.536; 6.017.751; 5.980.581; Publicación de Patente de Estados Unidos N^o 20020142438 A1. Por ejemplo, las pectato liasas de la invención se pueden usar en limpieza de fibra y/o tejido. En un aspecto, la sensación y el aspecto de un tejido mejora con un método de la invención que comprende poner en contacto el tejido con pectato liasa de la invención en una solución. En un aspecto, el tejido se trata con la solución a presión. Por ejemplo, las pectato liasas de la invención se pueden usar en la eliminación de la.

En un aspecto, las pectato liasas de la invención se aplican durante o después del tejido de textiles, o durante la etapa de retirada del apresto, o durante una o más etapas adicionales del procesamiento del tejido. Durante el tejido de textiles, los hilos se exponen a una tensión mecánica considerable. Antes del tejido en telares mecánicos, los hilos de urdimbre a menudo se revisten con almidón o derivados de almidón para encolado para aumentar su resistencia a la tracción y para evitar la ruptura. Después de haber tejido los textiles, un tejido puede evolucionar a una etapa de retirada del apresto. Esto puede ir seguido de una o más etapas adicionales de procesamiento del tejido. El des encolado es el acto de eliminar la "cola" de los textiles. Después del tejido, el revestimiento de encolado se debe eliminar antes del procesamiento adicional del tejido con el fin de garantizar un resultado homogéneo y a prueba de lavado.

Las enzimas de la invención se pueden usar para limpiar tejidos o cualquier material que comprende pectato o ácido poligalacturónico, que incluye tejidos que contienen algodón, como aditivos detergentes, por ejemplo, en composiciones acuosas. Para la fabricación de ropa, el tejido se puede cortar y coser en ropa o prendas de vestir. Estas se pueden acabar antes o después del tratamiento. En particular, para la fabricación de pantalones vaqueros, se han desarrollado diferentes métodos de acabado enzimático. El acabado de prendas vaqueras comienza con una

etapa de retirada del apresto enzimática, durante la que las prendas someten a la acción de enzimas amilolíticas para proporcionar suavidad al tejido y hacer que el algodón sea más accesible a las etapas de acabado enzimático posteriores. La invención proporciona métodos de acabado de prendas vaqueras, retirada enzimática del apresto y proporcionan suavidad a los tejidos mediante el uso de cualquier combinación de enzimas, tales como amilasas, endoglucanasas, una pectato liasa de la invención.

En un aspecto, una amilasa alcalina y termoestable y la pectato liasa se combinan en una retirada del apresto y biolimpieza en un solo baño. Entre las ventajas de la combinación de la retirada del apresto y limpieza en una etapa se encuentran la reducción de costes y el impacto ambiental menor debido a ahorros de energía y uso de agua y una producción de residuos menos. Las condiciones de aplicación para la retirada del apresto y biolimpieza pueden estar entre aproximadamente pH 8,5 y pH 10,0 y temperaturas a aproximadamente 40 °C y superiores. Se pueden usar dosificaciones enzimáticas bajas (por ejemplo, aproximadamente 5 g por una ton de algodón) y tiempos de reacción cortos (por ejemplo, aproximadamente 15 minutos) para obtener una retirada del apresto y limpieza eficaces sin calcio añadido.

Las pectato liasas de la invención se pueden usar en combinación con otras enzimas que degradan carbohidratos, por ejemplo, celulasa, arabinanasa, xiloglucanasa, pectinasa, xilanasa, similares, para la preparación de fibroso para la limpieza de fibras. Algunas proteasas también se pueden usar en combinación. Estas se pueden usar en combinación con detergentes. En un aspecto, las pectato liasas de la invención se pueden usar enterramientos para evitar la aparición de colores grises en textil.

Las pectato liasas de la invención se pueden usar para tratar cualquier material celulósico, incluye fibras (por ejemplo, fibras de algodón, cáñamo, fibra de lino o hilo de lino), tejidos cosidos y sin coser, por ejemplo, tejidos de punto, tejidos, pantalones vaqueros, hilos y toallas, hechos de algodón, mezclas de algodón o agentes celulósicos naturales o hechos por el hombre (por ejemplo, que proceden de fibras de celulosa que contienen xilano como tales como de pasta de madera) o mezclas de los mismos. Algunos ejemplos de mezclas son mezclas de algodón o rayón/viscosa con uno o más materiales de compañía tales como lana, fibras sintéticas (por ejemplo fibras de poliamida, fibras acrílicas, fibras de poliéster, fibras de alcohol de polivinilo, fibras de cloruro de polivinilo, fibras de cloruro de polivinilideno, fibras de poliuretano, fibras de poliurea, fibras de aramida), y fibras que contienen celulosa (por ejemplo rayón/viscosa, ramio, cáñamo, lino/hilo de lino, yute, fibras de acetato de celulosa, lyocell).

Los procesos de tratamiento de textiles de la invención (por ejemplo, limpieza usando las pectato liasas de la invención) se pueden usar en combinación con otros tratamientos textiles, por ejemplo, retirada del apresto y blanqueado. La limpieza es la retirada del material no celulósico de la fibra de algodón, por ejemplo, la cutícula (que consiste principalmente en ceras) y la pared celular primaria (que consiste principalmente en pectina, proteína y xiloglucano). Una eliminación de cera apropiada es necesaria para obtener una alta humectabilidad. Esto es necesario para el teñido. La eliminación de las paredes celulares primarias con los procesos de la invención mejora la eliminación de cera y asegura un tenido más uniforme. El tratamiento de textiles con los procesos de la invención puede mejorar la blancura en el proceso de blanqueado. El producto químico principal usado en la limpieza es el hidróxido sódico en concentraciones elevadas y a temperaturas elevadas. El blanqueado comprende la oxidación del textil. Por lo general, el blanqueado implica el uso de peróxido de hidrógeno como el agente oxidante para obtener ya sea un tejido totalmente blanqueado (blanco) o para garantizar un tono limpio del colorante.

La invención proporciona un proceso de un solo baño para retirar el apresto, limpieza y blanqueado de materiales celulósicos. En un aspecto, la retirada del apresto, limpieza y blanqueados se realizan en un solo baño poniendo en contacto los materiales celulósicos de forma simultánea o secuencialmente en un recipiente (un "solo baño") con un sistema enzimático y un sistema de blanqueado que comprende peróxido de hidrógeno o al menos un compuesto peroxi que puede generar peróxido de hidrógeno cuando se disuelve en agua, o combinaciones de los mismos, y al menos un activador del blanqueado. Algunos materiales celulósicos que incluyen fibras en bruto, hilo, o textiles tejidos o de punto, hechos de algodón, hilo de lino, fibra de lino, ramio, rayón, cáñamo, yute, o mezclas de estas fibras entre sí o con otras fibras naturales o sintéticas, se pueden tratar con los procesos de la invención.

La invención también proporciona pectinasas alcalinas (pectato liasas activas en condiciones alcalinas). Estas tienen amplias aplicaciones en el procesamiento de textiles, desgomado de fibras vegetales (por ejemplo, fibras del líber de la planta), tratamiento de aguas residuales pécticas, fabricación de papel, y fermentaciones de café y té. Véase, por ejemplo, Hoondal (2002) Applied Microbiology and Biotechnology 59: 409-418.

Tratamiento de alimentos y procesamiento de alimentos

Las pectato liasas de la invención tiene numerosas aplicaciones en la industria de procesamiento de alimentos. Por ejemplo, en un aspecto, las pectato liasas de la invención se usan para mejorar la extracción de aceite de materiales de plantas ricas en aceite, por ejemplo, semillas ricas en aceite, por ejemplo, aceite de semilla de soja de semillas de soja, aceite de oliva de aceitunas, aceite de colza de colza y/o aceite de girasol de semillas de girasol.

Las pectato liasas de la invención se pueden usar para separación de componentes de materiales de células vegetales. Por ejemplo, las pectato liasas de la invención se pueden usar en la separación de materiales ricos en

pectina (por ejemplo, paredes celulares), azúcar material de plantas rico en almidón en componentes, por ejemplo, sacarosa de remolacha azucarera o almidón o azúcares de la patata, pulpa o fracciones de cáscara. En un aspecto, las pectato liasas de la invención se pueden usar para separar cultivos ricos en proteínas o ricos en aceite en fracciones valiosas de proteína y aceite y cáscara. El proceso de separación se puede realizar mediante el uso de métodos conocidos en la técnica.

Las pectato liasas de la invención se pueden usar en la preparación de zumos, jarabes, extractos de frutas o vegetales y similares para aumentar el rendimiento. Las pectato liasas de la invención se pueden usar en el tratamiento enzimático (por ejemplo, hidrólisis de pectinas y/o ácido poligalacturónico, tal como ácido alfa-D-galacturónico unido en la posición 1,4) de diversos materiales derivados de la pared de célula vegetal, por ejemplo de la producción de vino o zumo, o restos agrícolas tales como cáscaras de vegetales, cáscaras de judía, culpa de remolacha azucarera, culpa de aceituna, culpa de patata, y similares. Las pectato liasas de la invención se pueden usar para modificar la consistencia y aspecto de frutas o vegetales procesados. Por ejemplo, las pectato liasas de la invención se pueden usar en la producción de zumos transparentes, por ejemplo, de manzanas, peras o bayas; para enturbiar zumos estables, por ejemplo, de manzanas, peras, bayas, cítricos o tomates; y para tratar purés, por ejemplo, de zanahorias y tomates, y para tratar sirope de dátíl (véase, por ejemplo, Sidhu (2002) Food Chemistry 79: 215-220). En estos procesos, las pectato liasas de la invención se pueden usar con otras enzimas (por ejemplo, celulasas, amilasas, etc.) u otras composiciones. Por ejemplo, en un aspecto, se usan enzimas de pectinasa y celulasa para mejorar el rendimiento, estabilidad y calidad del zumo de una fruta, por ejemplo, un higo chumbo. Una pectinasa de la invención puede aumentar el rendimiento, estabilidad y color (ensayo de color a medida como liberación de antocianinas o carotenoides) y claridad de un zumo. En un aspecto, se usa una combinación de pectinasa y celulasa; pectinasa a un 0,50 % en v/v puede producir un rendimiento elevado, un zumo transparente libre de sedimentos y zumo de alta calidad. Véase, por ejemplo, Essa, Hesham A., *et al.*, 2002, Nahrung, 46 (4): 245-250,

En un aspecto, una enzima o preparación enzimática de la invención se usa para la despectinación y reducción de la viscosidad en zumo de vegetales y/o frutas, por ejemplo, en zumos de manzana o pera u otras preparaciones alimentarias de manzana o pera (por ejemplo, salsas). En un aspecto, el zumo de fruta o vegetal se trata con una preparación enzimática de la invención en una cantidad eficaz para degradar el material que contiene pectina contenido en el zumo de fruta o vegetal.

En un aspecto, la enzima o preparación enzimática se usa en el tratamiento de puré de frutas y verduras para mejorar la capacidad de extracción o de degradación del puré. La preparación enzimática se puede usar en el tratamiento de puré de manzanas y peras para producción de zumo, y en el tratamiento del puré de uvas para producción de vino.

Las pectato liasas de la invención se pueden usar para tratar materiales vegetales para facilitar el procesamiento del material vegetal, incluyendo alimentos, para partir a la purificación extracción de los componentes vegetales tales como galactanos, pectinas y/o ácidos poligalacturónicos. Las pectato liasas de la invención se pueden usar para purificar pectinas de cítricos, para aumentar el valor alimentario, disminuye la capacidad de unión a agua, aumentar la capacidad de degradación en plantas de agua residual y/o aumenta la conversión de material vegetal para ensilado, y similares.

Alimentos y comida o aditivos alimentarios para animales

La invención proporciona métodos para el tratamiento de alimentos o comidas y aditivos de comida y alimentarios de animales usando las pectato liasas de la invención, animales que incluyen mamíferos (por ejemplo, seres humanos), pájaros, peces y similares. La invención proporciona alimentos, comidas, y aditivos para animales que comprenden las pectato liasas de la invención. En un aspecto, el tratamiento de alimentos, comidas y aditivos para animales usando las pectato liasas de la invención puede ayudar en la disponibilidad de nutrientes, por ejemplo, almidón, en el alimento o aditivo para animales. Esto puede dar como resultado la liberación de nutrientes y azúcares absorbidos digeribles y de fácil absorción.

Las pectato liasas de la presente invención, en la modificación del alimento o una comida para un animal, pueden procesar el alimento o la comida *in vitro* (modificando los componentes del alimento o comida) o *in vivo*. Las pectato liasas se pueden añadir a composiciones para alimento o comida para animales que contienen cantidades elevadas arabinogalactanos o galactanos, por ejemplo alimento o comida que contiene material vegetal a partir de semilla de soja, semilla de colza, altramuzy y similares. Cuando la pectato liasa se añade al alimento o comida aumenta de forma significativa la descomposición *in vivo* del material de la pared de célula vegetal, por lo que se consigue un mejor uso de los nutrientes de la planta por el animal (por ejemplo, ser humano). En un aspecto, la tasa de crecimiento y/o proporción de conversión del alimento (es decir, el peso del alimento ingerido con respecto al aumento de peso) del animal aumenta. Por ejemplo el galactano que no se puede digerir se degrada con una pectato liasa de la invención, por ejemplo en combinación con beta-galactosidasa, en galactosa o galactooligómeros. Estos productos de digestión enzimática son más digeribles por el animal. Por lo tanto, pueden contribuir a la energía disponible del alimento. Además, mediante la degradación del galactano, la pectato liasa de la invención

puede aumentarla la capacidad de digestión y absorción de los componentes del alimento que no son carbohidrato tales como proteína, grasa y minerales.

En otro aspecto, la pectato liasa de la invención se puede proporcionar mediante la expresión de las enzimas directamente en cultivos de alimentos transgénicos (como por ejemplo, plantas transgénicas, semillas y similares), tales como maíz, semilla de soja, semilla de colza, altramuces y similares. Como se ha analizado anteriormente, la invención proporciona plantas transgénicas, partes de plantas y células vegetales que comprenden secuencias de ácidos nucleicos que codifican un polipéptido de la invención. En un aspecto, el ácido nucleico se expresa de modo que la pectato liasa de la invención se produce en cantidades recuperables. Las pectato liasas se pueden recuperar de cualquier planta o parte de planta. Como alternativa, la planta o parte de planta que contiene el polipéptido recombinante se puede usar como tal para aumentar la calidad de un alimento o comida, por ejemplo, aumentando el valor nutricional, sabor, y propiedades reológicas, o para destruir un factor antinutritivo.

Tratamiento de papel o pasta

Las pectato liasas de la invención se pueden encontrar en el tratamiento de papel o pasta o eliminación de tinta del papel. Por ejemplo, en un aspecto, la invención proporciona un proceso de tratamiento de papel que usarlas pectato liasas de la invención. En un aspecto, las pectato liasas se pueden usar para modificar pectina y/o ácido poligalacturónico, tal como ácido alfa-D-galacturónico unido en la posición 1,4. En otro aspecto, los componentes del papel de papel fotocopiado reciclado durante procesos de eliminación de tinta químicos y enzimáticos. En un aspecto, las pectato liasas de la invención se pueden usar en combinación con celulasas. El papel se puede tratar mediante los siguientes tres procesos: 1) disgregación en presencia de las pectato liasas de la invención, 2) disgregación con un agente químico de eliminación de tinta y las pectato liasas de la invención, y/o 3) disgregación después de empapar con las pectato liasas de la invención. El papel reciclado tratado con las pectato liasas puede tener un brillo más elevado debido a la eliminación de partículas de tóner en comparación con el papel tratado solamente con celulosa. Aunque la invención no se limita mediante ningún mecanismo en particular, el efecto de las pectato liasas de la invención se puede deber a su comportamiento como agentes de superficie activa en suspensión de pasta.

La invención proporciona métodos para el tratamiento de papel y pasta de papel usando una o más pectato liasas de la invención. Las pectato liasas de la invención se pueden usar en cualquier método de tratamiento de papel o pasta, que se conocen bien en la técnica, véase, por ejemplo, el documento de Patente de Estados Unidos Nº 6.241.849; 6.066.233; 5.582.681. Por ejemplo, en un aspecto, la invención proporciona un método para eliminar la tinta y decolorar un papel impreso que contienen colorantes, que comprende la preparación de pasta de un papel impreso para obtener una suspensión de pasta, y retirar una tinta de la suspensión de pasta en presencia de las pectato liasas de la invención (también se pueden añadir otras enzimas). En otro aspecto, la invención proporciona un método para aumentar la liberación de la pasta, por ejemplo, pasta preparada a partir de fibra secundaria, mediante la adición de una mezcla enzimática que comprende las pectato liasas de la invención (también puede incluir otras enzimas, por ejemplo, las enzimas celulasa, amilasa o glucoamilasa) a la pasta y para el tratamiento en condiciones que provocan una reacción para producir una pasta tratada de forma enzimática. La liberación de la pasta tratada de forma enzimática aumenta desde la liberación inicial de la pasta de fibras secundarias sin una pérdida del brillo.

Repulpado: tratamiento de materiales lignocelulósicos

La invención también proporciona un método para el tratamiento de fibras lignocelulósicas, en el que las fibras se tratan con las pectato liasas de la invención, en una cantidad que es eficaz para la mejora de las propiedades de la fibra. Las pectato liasas de la invención también se pueden usar en la producción de materiales lignocelulósicos, tales como pasta, papel y cartón, a partir de papel usado reforzado con almidón y cartón, especialmente cuando el repulpado se produce a pH superior a 7 y en el que pectato liasas pueden facilitar la disgregación del material de desecho a través de la degradación de las paredes celulares. Las pectato liasas de la invención pueden ser útiles en un proceso para producir una pasta de fabricación de papel a partir de papel impreso revestido con almidón. El proceso se puede realizar como se describe, por ejemplo, en el documento WO 95/14807.

Un proceso a modo de ejemplo comprende la disgregación del papel para producir una pasta, mediante el tratamiento con una enzima que degrada pectina de la invención antes, durante o después de la disgregación, y separando las partículas de tinta de la pasta después de disgregación y tratamiento enzimático. Véase también el documento de Patente de Estados Unidos Nº 6.309.871 y otros documentos de patente de Estados Unidos mencionados en el presente documento. Por lo tanto, la invención incluye un método para eliminación enzimática de tinta de la pasta de papel reciclado, en el que las pectato liasas se aplican en una cantidad que es eficaz para la eliminación de tinta de la superficie de la fibra.

Tratamiento de residuos

Las pectato liasas de la invención se pueden usar en una diversidad de otras aplicaciones industriales, por ejemplo, en el tratamiento de residuos. Por ejemplo, en un aspecto, la invención proporciona un proceso de digestión de residuos sólidos usando las pectato liasas de la invención. Los métodos pueden comprender la reducción de la

masa y el volumen de los residuos sólidos básicamente sin tratar. Los residuos sólidos se pueden tratar con un proceso digestivo enzimático en presencia de una solución enzimática (incluyendo las pectato liasas de la invención) a una temperatura controlada. Esto da como resultado una reacción sin fermentación bacteriana apreciable a partir de microorganismos añadidos. El residuo sólido se convierte en un residuo licuado y cualquier residuo sólido residual. El residuo licuado resultante se puede separar de cualquier residuo solidificado residual mencionado. Véase, por ejemplo, el documento de Patente de Estados Unidos N° 5.709.796.

Productos para el cuidado oral

La invención proporciona productos para el cuidado oral que comprenden pectato liasas de la invención. Algunos productos para el cuidado oral a modo de ejemplo incluyen pastas dentífricas, cremas dentales, ángeles o polvos para dientes, agentes odontológicos, enjuagues bucales, formulaciones de aclarado para antes o después del cepillado, gomas de mascar, pastillas para chupar, o caramelos. Véase, por ejemplo, el documento de Patente de Estados Unidos N° 6.264.925.

Fabricación de cerveza y fermentación

La invención proporciona métodos de fabricación de cerveza (por ejemplo, fermentación) que comprenden las pectato liasas de la invención. En un proceso a modo de ejemplo, las materias primas que contienen almidón se desintegran y se procesan para formar una malta. Una pectato liasa de la invención se usa en cualquier punto en el proceso de fermentación. Por ejemplo, las pectato liasas de la invención se pueden usar en el procesamiento de malta de cebada. La materia prima principal de la fabricación de cerveza es la malta de cebada. Se puede tratar de un proceso de tres etapas. En primer lugar, el grano de cebada se puede impregnar para aumentar el contenido de agua, por ejemplo, hasta aproximadamente un 40 %. En segundo lugar, el grano se puede germinar mediante incubación de 15 °C a 25 °C de 3 a 6 días cuando la síntesis enzimática se estimula bajo el control de giberelinas. En un aspecto, las pectato liasas de la invención se añaden en esta etapa (o cualquier otra) del proceso. La acción de las pectato liasas da como resultado un aumento de azúcares de reducción fermentables. Esto se puede expresar como la potencia diastática, DP, que puede aumentar de aproximadamente 80 a 190 en 5 días a 12 °C. Las pectato liasas de la invención se pueden usar en cualquier proceso de producción de cerveza o bebida alcohólica, como se describe, por ejemplo, en los documentos de patente de Estados Unidos N°s 5.762.991; 5.536.650; 5.405.624; 5.021.246; 4.788.066.

Otras aplicaciones industriales

La invención también incluye un método para aumentar el flujo de fluidos de producción a partir de una formación subterránea retirando un fluido dañino, que contiene pectina, viscoso formado durante las operaciones de producción y que se encuentra dentro de la formación subterránea que rodea una perforación del pozo completada que comprende permitir a los fluidos de producción que fluyan desde la perforación del pozo; reducir el flujo de fluidos de producción desde la formación por debajo de caudales esperados; formular un tratamiento enzimático mezclando en conjunto un fluido acuoso y un polipéptido de la invención; bombear el tratamiento enzimático a una ubicación deseada dentro de la perforación del pozo; permitir que el tratamiento enzimático degrade el fluido dañino, que contiene pectina, viscoso, mediante el cual el fluido se puede retirar de la formación subterránea a la superficie del pozo; y en el que el tratamiento enzimático es eficaz para atacar la pectina en paredes celulares.

La invención se describirá adicionalmente con referencia a los siguientes ejemplos; sin embargo, se debe entender que la invención no se limita a tales ejemplos.

Ejemplos

EJEMPLO 1: Ensayos de actividad de pectato liasa

El siguiente ejemplo describe ensayos de actividad de pectato liasa a modo de ejemplo para determinar la actividad catalítica de una pectato liasa. Estos ensayos a modo de ejemplo se pueden usar para determinar si un polipéptido está dentro del alcance de la invención.

Ensayo de viscosidad de la unidad APSU

Unidades APSU: El ensayo de la unidad APSU es una medida de viscosidad que usa el ácido poligalacturónico sustrato sin calcio añadido.

La sal sódica del ácido poligalacturónico al 5 % sustrato (Sigma P-1879) se solubiliza en tampón de glicina 0,1 M a pH 10. Los 4 ml de sustrato se incuban previamente durante 5 min a 40 °C. La enzima se añade (en un volumen de 250 µl) y se mezcla durante 10 segundos en una mezcladora a velocidad máxima, a continuación se incuba durante 20 min a 40 °C. Para una determinación doble de una curva estándar de una dilución de concentración de enzima en el intervalo de 5 APSU/ml hasta superior a 100 APSU/ml con un mínimo de 4 concentraciones entre 10 y 60 APSU por ml. La viscosidad se puede medir usando un MIVI600™ (Sofraser, Villemandeur, Francia). La discordia se puede

medir como mV después de 10 segundos. El programa PRISM™ Prism de GRAFAPAD, que usa un ajuste no lineal con una desintegración exponencial de una fase con una meseta, se puede usar para los cálculos. La meseta más el lapso es el mV obtenido sin enzima. Véase, por ejemplo, la Patente de Estados Unidos N° 6.368.843.

5 Ensayo de beta-eliminación

Un ensayo de liasa a modo de ejemplo (a 235 nm) para la determinación de la actividad de beta-eliminación mide de aumentos en la absorbancia a 235 nm. La sal sódica del ácido poligalacturónico al 0,1 % sustrato (Sigma P-1879) se solubiliza en tampón de glicina 0,1 M a pH 10. Para el cálculo de la tasa catalítica, un aumento de 5,2 de absorbancia a 235 unidades por min corresponde a la formación de 1 μmol de producto insaturado (véase, por ejemplo, Nasuna (1966) J. Biol. Chem. 241: 5298-5306; Bartling (1995) Microbiology 141: 873-881). La condición en el estado estacionario se mide usando una cubeta de 0,5 ml con una trayectoria de la luz de 1 cm en un espectrofotómetro de matriz de diodo HP en un soporte de cubeta con temperatura controlada con la medida continua de la absorbancia a 235 nm. Para el estado estacionario, se puede usar un aumento lineal durante al menos 200 segundos para el cálculo de la tasa. Se usa para convertir pmol por min de producto. Véase, por ejemplo, la Patente de Estados Unidos N° 6.368.843.

Ensayo de Agar

La actividad de pectato liasa se puede medir mediante un ensayo de agar. Una solución de ensayo se aplica a agujeros de 4 mm perforados en placas de agar (por ejemplo, LB agar), que contienen poligalacturonato sódico al 0,7 % en p/v (Sigma P 1879). A continuación, las placas se incuban durante 6 h a una temperatura en particular (por ejemplo, 75 °C.). A continuación, las placas empapan en cualquiera de (i) CaCl_2 1 M durante 0,5 h o (ii) Br de alquil trimetilamonio mixto al 1 % (MTAB, Sigma M-7635) durante 1 h. Ambos procedimientos provocan la precipitación del poligalacturonato dentro del agar. La actividad de pectato liasa se puede detectar mediante la aparición de zonas transparentes dentro de un fondo de poligalacturonato precipitado. La sensibilidad del ensayo se calibra usando dilución de una preparación estándar de pectato liasa.

Análisis del Punto Final -- Trans-eliminación a 235 nm para Pectato Liasas (Método de Calcio Elevado: Calcio 1 mM en la Mezcla de Incubación Final). En este método, el sustrato y la enzima se incuban durante 20 min a 37 °C seguido de la medida a 235 nm de la formación de uniones dobles. Por último, la tasa de la degradación se calcula basándose en el coeficiente de extinción molar en términos de Unidades Trans.

Procedimiento: Mezcla de 0,5 ml de dilución enzimática con 0,5 ml de solución de sustrato. Sustrato: Ácido poligalacturónico de Sigma P-1879, lote 77H3784, Tampón 2x de Glicina 0,1 M a pH 10+, 2,0 mmol de CaCl_2 , Reactivo de parada: H_3PO_4 0,02 M, Temperatura de incubación 37 °C, Tiempo de reacción 20 min. Coeficiente de extinción de la trans-eliminación 0,0052 $\mu\text{mol cm}^{-1}$. Enzima diluida en agua libre de iones de 0,5 a 5 APSU por ml. Valor principal por duplicado de 0,5 ml. El sustrato al 2 % en p/v en tampón 2x se mezcla con 0,5 ml de enzima diluida. Ambos se incuban previamente durante 5 min en un baño de agua a 37 °C. Incubar durante 20 min. Detener usando 5 ml de reactivo de parada y mezclar. Mezclar la enzima en blanco usada como blanco y el reactivo de parada primero y a continuación añadir el sustrato, todo en el mismo volumen.

Enzima	0,5 ml
Sustrato	0,5 ml
Parada	5 ml
Volumen Total	6 ml

Medir la absorbancia a 235 nm en una cubeta de 1 cm. Calcular la formación de la trans-eliminación por min usando el coeficiente de extinción de 0,0052 $\mu\text{mol cm}^{-1}$. Véase, por ejemplo, la Patente de Estados Unidos N° 6.368.843.

EJEMPLO 2: Ensayo de Aplicación de Bio-Limpieza en Algodón

El siguiente ejemplo describe un Ensayo de Aplicación de Limpieza en Algodón a modo de ejemplo usando las enzimas pectato liasa de la invención. El uso de las pectato liasas de la invención para hidrolizar la pectina de la pared celular primaria ("biolimpieza") puede eliminar la necesidad de agentes cáusticos y temperaturas elevadas en la limpieza de fibra de algodón.

Preparación de los Materiales:

- Requiere tampón de Bicarbonato Sódico 50 mM a pH óptimo
- Dilución a 1:10 del tensioactivo Calloway 1663
- Tampón fosfato 50 mM a pH 6

- Mezclar 43,3 ml de Na-P monobásico 1,0 M, 6,6 ml de Na-P dibásico 1,0 M, ajustar el volumen a 1 l con agua D.I. Ajustar el pH a 6
- Rojo de Rutenio (R-2751 SIGMA)
- Añadir 0,5 g de Rojo de Rutenio a 1 l de Tampón Fosfato produciendo una concentración final de 0,05 %.
- 5 • NaOAc (5 g/l a pH 5)
- Tejido de algodón 400R (Testfabrics Inc.) al que se le quita el apresto antes de la limpieza

Procedimiento de Limpieza:

- 10 1. Colocar 1,0 g de tejido de algodón sin apresto (en cada vaso de precipitados Labomat).
2. Cada experimento debería usar un algodón sin tratar, como blanco (sin enzimas añadidas).
3. Añadir 50 ml de Tampón de Bicarbonato Sódico 50 mM a pH 8,5-9 a cada vaso de precipitados y 2,5 ml de dilución a 1:10 de Tensioactivo Calloway 1663.
4. Ajustar las tapas usando una llave Allen e instalar los vasos de precipitados en el Labomat, asegurándose de
- 15 que los vasos de precipitados se distribuyen de manera uniforme en la rejilla giratoria. Conectar el vaso de precipitados 1 con el cable detector de temperatura al conector en la parte media de la rejilla.
5. Elevar la temperatura hasta la temperatura deseada y mantener durante 10 minutos.
6. Añadir 50-200 ul de enzima (por ejemplo, una pectato liasa de la invención) a una concentración diluida previamente a 0,1 ug/ul a través del tabique de separación en el vaso de precipitados usando una jeringa. La
- 20 concentración enzimática total usada para limpiar 1 gramo de tejido de algodón puede estar entre 5-20 ug.
7. Realizar la reacción en el Labomat a temperatura durante 15 minutos.
8. Aclarar el tejido de algodón dos veces by poniendo el tejido de algodón en la mano y apretando el algodón seco, colocar el algodón de nuevo en el vaso de precipitados y rellenar el vaso de precipitados con agua D.I. y repetir esta etapa de nuevo, acabar escurriendo el exceso de agua del algodón.
- 25 9. Empapar el tejido de algodón en NaOAc (5 g/l de pH 5) durante 2 minutos.
10. Repetir el ciclo de aclarado 2X en la etapa 8.
11. Colocar el tejido de algodón en bandejas de pesada y permitir que el tejido se seque durante una noche en las biocampanas de flujo laminar.

30 *Procedimiento de Teñido:*

1. Colocar el tejido de algodón tratado en los vasos de precipitados Labomat.
2. Añadir 100 ml de Rojo de Rutenio al 0,05 %, tampón de Na-P a pH 6 a cada vaso de precipitados.
3. Ajustar las tapas usando una llave Allen e instalar los vasos de precipitados en el Labomat asegurándose de
- 35 que los vasos de precipitados se distribuyen de manera uniforme en la rejilla giratoria. Conectar el vaso de precipitados 1 con el cable detector de temperatura al conector en la parte media de la rejilla.
4. Elevar la temperatura hasta 50 °C y mantener durante 30 minutos.
5. Aclarar el tejido dos veces poniendo el algodón en la mano y apretando el algodón seco, colocar el tejido de nuevo en el vaso de precipitados y rellenar el vaso de precipitados con agua D.I. y repetir esta etapa de nuevo, acabando con la retirada del exceso de agua del tejido.
- 40 6. Colocar el tejido teñido en el vaso de precipitados y añadir 100 ml de agua D.I..
7. Ajustar las tapas usando una llave Allen e instalar los vasos de precipitados en el Labomat asegurándose de que los vasos de precipitados se distribuyen de manera uniforme en la rejilla giratoria.
8. Elevar la temperatura a 100 °C y mantener durante 10 minutos; enfriar los vasos de precipitados hasta 60 °C.
- 45 9. Repetir el ciclo de aclarado 2X en la etapa 5.
10. Colocar el tejido teñido en bandejas de pesada y permitir que el tejido se seque durante una noche en las biocampanas de flujo laminar.

Cuantificación de la Limpieza Enzimática:

- 50 1. Calibrar el GretagMacbeth Color Eye 7000A seleccionando el Icono de Color Eye en el escritorio del ordenador.
2. Colocar la lente de color negro sobre el orificio y pulsar enter cuando el programa solicite la calibración del umbral negativo.
3. Colocar el filtro de color blanco sobre el edificio y pulsar enter cuando el programa solicite la calibración del
- 55 equilibrio de color blanco.
4. Colocar el tejido teñido seco sobre el orificio y pulsar F4 para leer la blancura del tejido.
5. Registrar el número L*, dar la vuelta al tejido para leer el otro lado y registrar el número L*. calcular el número medio L* para cada muestra.
- 60 6. Hacer el gráfico del L delta para cada muestra de algodón limpiada restando el número L* de las muestras con el L* del tejido sin tratar.

EJEMPLO 3: Un proceso de un solo baño para retirar el apresto y limpieza

El siguiente ejemplo describe un proceso de un solo baño a modo de ejemplo para retirar el apresto y limpieza. La invención proporciona secciones para retirar el apresto, limpieza y blanqueado de materiales celulósicos poniendo en contacto los materiales celulósicos de forma simultánea o secuencial en un proceso de un solo baño con sistema enzimático que comprende una pectato liasa de la invención. El proceso de un solo baño puede comprender adicionalmente sistema de blanqueado que comprende peróxido de hidrógeno o al menos un compuesto peroxi que genera peróxido de hidrógeno cuando se disuelve en agua, o combinaciones de los mismos, y al menos un activador del blanqueado.

Algunos materiales celulósicos incluyen fibras en bruto, hilo, o textiles tejidos o de punto, hechos de algodón, hilo de lino, fibra de lino, ramio, rayón, cáñamo, yute, o mezclas de estas fibras entre sí o con otras fibras naturales o sintéticas, se pueden tratar mediante este proceso de un solo baño de la invención. En un aspecto, un peso de tejido se carga en un recipiente, que se rellena posteriormente con una solución tampón (por ejemplo, tampón de fosfato Na 20 mM, pH 9,2) que comprende una pectato liasa de la invención (por ejemplo, 3000 APSU/kg por fibra de pectato liasa), agente humectante (por ejemplo, 0,5 g/l), H₂O₂ (por ejemplo, 1,7 g/l) y la gente estabilizante (por ejemplo, 0,75 g/l). El tejido se puede tratar, por ejemplo, a 55 °C durante aproximadamente 15 min, tras lo que la temperatura se elevó de 5 °C/min a 70 °C durante 1 h. A continuación, el tejido se lava minuciosamente con agua para retirar los agentes químicos residuales y se seca a temperatura ambiente durante una noche.

EJEMPLO 4: Ensayo para detectar enzimas termotolerantes

El siguiente ejemplo describe un ensayo a modo de ejemplo para detectar enzimas termotolerantes que se puede usar para determinar si una enzima está dentro del alcance de la invención. Este ejemplo describe un ensayo de identificación sistemática basada en la absorbancia ("descubrimiento") para detectar enzimas termotolerantes, que, en un aspecto, se pueden caracterizar como "mutantes positivos" para un gen de pectato liasa "precursora". Este protocolo a modo de ejemplo se puede usar para variantes, o mutantes, generados por mediante cualquiera de los métodos de GSSM™ o conminatorios.

Materiales y preparaciones

- a. Ácido Poligalacturónico (PGA), y [Sigma P-3889] al 2 %
- b. Placas de Fondo Plano de 96 Pocillos (resistentes a) UV [Thomson Instrument 931801B]
- c. Placas de 96 Pocillos COSTAR
- d. Sellados de Papel de Aluminio para PCR Adhesivos [Marsh AB-0626]
- e. B-PER [PIERCE 78248]
- f. LBamp100 o LBcarb100
- g. TRIS a pH 8,0 (250 mM, 10X)
- h. Glicina (250 mM, 10X)
- i. Sustrato de Ácido Poligalacturónico al 0,2 % para detección de la actividad enzimática: 100 ml de cada uno de los siguientes: 10X Tris, 10X Glicina, y PGA al 2 %, más 700 ml de agua sagrada.
- j. Placas: alícuotas de 200 µl de medio, LBamp100 o LBcarb100, en los concilios tanto de las placas COSTAR como de las de fondo plano de 96 pocillos resistentes a UV.

Selección de colonias y replicación de placas

Colonias de clones GSSM™ o mutantes combinatorias se seleccionaron con un seleccionador de colonias Autogen (Framingham, MA) y la células se inocularon en medio Lamp100. Un total de 168 clones GSSM™ se identificaron sistemáticamente por sitio de residuo o se identificaron sistemáticamente 13.000 clones de la biblioteca combinatoria 2328. Los clones mutados se seleccionaron en las filas A, B, C, E, F, G, y H de la placa de 96 pocillos. Los clones de tipo silvestre (wt) (SEC ID N°: 132, codificados por la SEC ID N°: 131) se seleccionaron en la fila D como un control. Después de finalizar la selección de las colonias, las placas se incubaron durante una noche (aproximadamente 18 horas) a 37 °C, agitando la 150 RPM. A partir de este momento se hará referencia a estas placas como las placas madre.

Se prepararon copias de cada placa madre en placas resistentes a UV (ahora denominadas "placas de ensayo") usando en replicados de placas automatizado. Una vez que la replicación fue completa, las placas de ensayo se colocaron en una incubadora humidificada a 30 °C durante una noche.

Ensayo Primario

Las densidades celulares en cada pocillo de todas las placas se determinaron a DO₆₀₀ usando un sistema SPECTRAMAX™ (Molecular Devices Corporation, Sunnyvale CA). A continuación, todas las placas de ensayo se sellaron con Sellos de Papel de Aluminio para PCR y a continuación se centrifugaron 2200 rpm en una centrifugadora Eppendorf durante 10 minutos. Usando el sistema PowerWasher, el sobrenadante se aspiró a continuación fuera de las placas de ensayo dejando solamente las células. A continuación se añadieron 20 µl de B-PER™ (Pierce Biotechnology, Rockford, IL) a cada pocillo y las placas de ensayo se volvieron a sellar. A continuación, las placas se colocaron en un agitador de placas durante 10 minutos con el fin de garantizar la lisis

celular apropiada. A continuación, las placas se colocaron en una incubadora calentada previamente a 50 °C durante 50 minutos para el ensayo GSSM™ o a 70 °C durante 25 minutos para el ensayo de mutante positivo combinatorio. A continuación, las placas de ensayo se retiraron de la incubadora después del tiempo de estimulación por calor apropiado y se enfriaron rápidamente a temperatura ambiente. El SPECTRAMAX™ se usó para leer la cinética a la longitud de onda de 235 nm durante un periodo de 2 minutos. Cualquier supuesto acierto que se realizaba mejor que el de tipo silvestre se liberaba para un ensayo secundario. La Figura 8 ilustra un residuo con múltiples aciertos positivos. En la Figura 8, la Fila D contiene la actividad residual del tipo silvestre (wt), SEC ID N°: 132, y las filas A, B, C, E, F, G, H son los clones GSSM™ del sitio de mutación 182.

10 Ensayo Secundario

Se identificaron todos los pocillos que presentaban un aumento de la tasa enzimática en comparación con el rendimiento de tipo silvestre y los clones de la respectiva placa madre se liberaron. Usando técnicas a sépticas, se usó un palillo estéril en el pocillo de un supuesto acierto de la placa madre. Las células que se adhieren al palillo se transfirieron a una nueva placa seleccionando un nuevo pocillo con 200 µl de LBamp100. Además, del mismo modo, la fila D se llenó con WT de cada placa liderada. Las placas madres secundarias se colocaron en la incubadora humidificada a 30 °C durante una noche. Las placas madres secundarias se manipularon con un alfiler en placas de 96 pocillos resistentes UV. Las placas de ensayo secundarias colocaron a continuación en una incubadora humidificada a 30 °C durante una noche. Las densidades celulares en cada pocillo de todas las placas se determinaron a DO₆₀₀ usando los sistemas SpectraMax. A continuación, todas las placas de ensayos se sellaron con Sellos de Papel de Aluminio para PCR y a continuación se centrifugaron a 2200 rpm en una centrifugadora Eppendorf durante 10 minutos. Las etapas restantes del ensayo secundario fueron las mismas que se indican para el ensayo final. Cualquier acierto confirmado que se realiza mejor que el del tipo silvestre se separó y se sometió a ensayo de nuevo en el ensayo terciario.

25 Ensayo Terciario

Se tomaron alícuotas de 5 µl de cultivo de pocillos que confirmaban un aumento de la actividad de termotolerancia a partir del clon de tipo silvestre en una placa de Petri LBcarb100 pequeña para preparar placas sembradas en estrías. También se usaron 5 µl de de los bolsillos de control de "tipo silvestre" (wt) (SEC ID N°: 132, codificada por la SEC ID N°: 131) para preparar una placa sembrada en estrías. Las placas sembradas en estrías se incubaron a 37 °C durante una noche. Se raspó una pequeña sección de una colonia individual y se inocularon 5 ml de LBcarb100 a las células. Se permitió que el cultivo creciera durante una noche a 37 °C a 200 RPM. El acierto confirmado se diluyó a continuación hasta DO₆₀₀ = 0,2. Se tomaron alícuotas de 200 µl de un clon confirmado en un pocillo, rellenando una fila entera en la placa resistente a UV de 96 pocillos. Lo mismo se realizó para un control de wt. A continuación, todas las placas se sellaron y se centrifugaron a 2000 rpm en la centrifugadora Eppendorf durante 10 minutos. Las etapas restantes del ensayo terciario fueron las mismas que las que se indican para el ensayo primario. Al final, todos los aciertos supuestos que se realizaban mejor que el tipo silvestre se enviaron para secuenciación. También se prepararon reservas de glicerol.

40 EJEMPLO 5: Procesos y formulaciones para enzimas de la invención

El siguiente ejemplo describe procesos a modo de ejemplo (por ejemplo, un proceso de biolimpieza) y formulaciones de la invención. Las composiciones y procesos de la invención se sometieron a ensayo usando la pectato liasa a modo de ejemplo que tiene una secuencia como se establece en la SEC ID N°: 134, codificada, por ejemplo, por la SEC ID N°: 133 ("SEC ID N°: 134").

Definición de unidad:

La actividad de pectato liasa (SEC ID N°: 134) se midió de forma rutinaria usando ácido poligalacturónico al 0,2 % (p/v) (Sigma, P3850) en Tris HCl 25 mM – tampón de Glicina NaOH 25 mM. Una unidad de actividad enzimática se definió como la cantidad de proteína que producía 1 µmol de oligogalacturónidos insaturados por minuto equivalente a 1 µmol de digalacturónido insaturado, usando el valor del coeficiente de extinción molecular de 4600 M⁻¹cm⁻¹ a 235 nm para el dímero.

El aparato SpectraMax se usó en placas de UV de 96 pocillos.

Potencia de la formulación:

Las enzimas de la invención se pueden formular en cualquier dosificación para ajustarse a una necesidad en particular; en la técnica se conocen ensayos para determinar la dosificación óptima de cualquier formulación en particular, y varios se describen en el presente documento. En aspectos alternativos, las formulaciones de la invención pueden tener una potencia baja entre aproximadamente 2000 y 4000 u/ml (en las que u = unidad). Esto es comparable con los otros productos que hay en el mercado. En un aspecto, la formulación mínima es de aproximadamente 1000 u/ml, y, en otro aspecto, la formulación máxima es de aproximadamente 10.000 u/ml, por

ejemplo, la formulación puede comprender una enzima de la invención en una cantidad entre aproximadamente 1000 u/ml y 10.000 u/ml.

- 5 Los estudios de solubilidad con producto liofilizado (SEC ID N°: 134) resuspendido en agua indicaban que la solubilidad de la enzima puede ser tan elevada como 25000 u/ml a 4 °C. Por lo tanto, en un aspecto, la invención proporciona formulaciones que tienen un nivel tan elevado como aproximadamente 25000 u/ml, o superior. En un aspecto la invención proporciona formulaciones que comprenden una enzima de la invención en una cantidad entre aproximadamente 100 u/ml and 25000 u/ml, 30000 u/ml, 35000 u/ml o 40000 u/ml, o superior.

10 Diseño de la formulación:

- 15 La invención proporciona formulaciones que comprenden al menos una enzima de la invención, y, en aspectos alternativos, comprende adicionalmente cualquier aditivo o aditivos. Las formulaciones de la invención se pueden basar en aditivos conocidos en otras formulaciones enzimáticas, por ejemplo, análogas. Por ejemplo, las formulaciones de la invención pueden comprender los aditivos y/o condiciones que se exponen en las Tablas 3, 4, 5 y 6, que siguen a continuación, o cualquier variación de las mismas. Por ejemplo, las formulaciones de la invención pueden comprender glicerol, sacarosa, cloruro sódico, dextrina, propilenglicol, sorbitol, sulfato sódico o TRIS, o un equivalente.

- 20 En un aspecto, una formulación de la invención puede ser una formulación basada en agua, o, una formulación basada en aceite.

- 25 Se realizaron dos fases de estudios de estabilidad de la formulación; estos estudios usaron la enzima a modo de ejemplo de la SEC ID N°: 134.

Estudio de estabilidad acelerada a 37 °C

Nota: se trata de formulaciones basadas en tampón.

- 30
- Identificar sistemáticamente diversos aditivos
 - Someter a ensayo diferentes valores de pH
 - Formulaciones a aproximadamente 2000 u/ml.
 - La SEC ID N°: 134 fue la enzima a modo de ejemplo sometida a ensayo

35 Tabla 3 (SN = número de muestra)

SN	pH	Glicerol	Sacarosa	Cloruro sódico	Dextrina	Propilen glicol	sorbitol	Sulfato sódico	Conc. Eficaz de TRIS
1	5								40 mM
2	6								40 mM
3	7								40 mM
4	8								40 mM
5	7,5	35 %							20 mM
6	7,5	50 %							20 mM
7	7,5		35 %						20 mM
8	7,5			20 %					20 mM
9	7,5				10 %				20 mM
10	7,5					30 %			20 mM
11	7,5							100 mM	20 mM
12	7,5						35 %		20 mM
13	5,5	35 %							40 mM
14	5,5	50 %							40 mM

Formulaciones con el mayor rendimiento (usando la SEC ID N°: 134 como una enzima a modo de ejemplo de la invención) basadas en el aspecto físico y en la retención de actividad superior a un 80 %:

Tabla 4	
Formulación	Aditivo
1	pH 5,0, TRIS 40 mM
3	pH 7,0, TRIS 40 mM
4	pH 8,0, TRIS 40 mM
6	pH 7,5, glicerol al 50 %
8	pH 7,5, NaCl al 20 %
10	pH 7,5, propilenglicol al 30 %
11	pH 7,5, sulfato sódico 100 mM
13	pH 5,5, glicerol al 35 %

En aspectos alternativos, las formulaciones de la invención pueden tener aproximadamente 10.000 u/ml, o estar en una cantidad entre aproximadamente 100 u/ml, 200 u/ml, 300 u/ml, 400 u/ml o 500 u/ml y 10.000 u/ml, 15.000 u/ml, 20.000 u/ml, 25000 u/ml, 30000 u/ml, 35000 u/ml o 40000 u/ml, o superior. En aspectos alternativos, las formulaciones de la invención pueden ser de aproximadamente 500 a 30.000 unidades/ml, de 1000 a 25.000 unidades/ml, o, entre aproximadamente 1000 y 20.000 unidades/ml, de 1000 a 15.000 unidades/ml, de 1000 a 10000 unidades/ml, de 1000 a 5000 unidades/ml, entre aproximadamente 2000 y 20000 unidades/ml, entre aproximadamente 2000 y 15000 unidades/ml, entre aproximadamente 2000 y 10000 unidades/ml, o entre aproximadamente 2000 y 4000 unidades/ml. En aspectos alternativos, las formulaciones de la invención pueden comprender una formulación basada en agua, por ejemplo, cuando es factible ningún tampón; se puede usar cualquier sistema tampón basado en agua.

Tabla 5:

ESTUDIO EN FASE II DE ESTABILIDAD DE LA FORMULACIÓN DE PECTATO LIASA					
Nº Formulación	pH	Tampón	ADITIVOS		
1	pH 7,0	ND			benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
2	pH 7,0	ND			300 ppm de proxel
3	pH 7,0	ND	cloruro sódico al 15 %		benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
4	pH 7,0	ND	cloruro sódico al 15 %		300 ppm de proxel
5	pH 7,0	ND		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
6	pH 7,0	ND		glicerol al 35 %	300 ppm de proxel
7	pH 7,0	ND	cloruro sódico al 10 %	glicerol al 25 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
8	pH 7,0	ND	cloruro sódico al 10 %	glicerol al 25 %	300 ppm de proxel
9	pH 5,5	ND			benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
10	pH 5,5	ND		300 ppm de proxel	
11	pH 5,5	ND	cloruro sódico al 15 %		benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
12	pH 5,5	ND	cloruro sódico al 15 %		300 ppm de proxel
13	pH 5,5	ND		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
14	pH 5,5	ND		glicerol al 35 %	300 ppm de proxel
15	pH 5,5	ND	cloruro sódico al 10 %	glicerol al 25 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
16	pH 5,5	ND	cloruro sódico al 10 %	glicerol al 25 %	300 ppm de proxel
CONTROLES					
17	pH 7,0	TRIS		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
18	pH 5,5	TRIS		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
19	pH 7,0	Acetato		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
20	pH 5,5	Acetato		glicerol al 35 %	benzoato sódico al 0,1 %, sorbato potásico al 0,1 %

Se pueden usar tapones adicionales en una formulación de la invención: MOPS 20 mM, pH 7 o MOPS 25 mM, NaCl 50 mM, pH 7,5.

5

Formulaciones con el mayor rendimiento (usando la SEC ID N°: 134 como una enzima a modo de ejemplo de la invención) basadas en el aspecto físico y en la retención de actividad superior a un 80 %:

Tabla 6	
Formulación N°	Detalles
5	pH 7, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
6	pH 7, glicerol al 35 %, 300 ppm de proxel
7	pH 7, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
8	pH 7, cloruro sódico al 10 %, glicerol al 25 %, 300 ppm de proxel
13	pH 5,5, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
14	pH 5,5, glicerol al 35 %, 300 ppm de proxel
15	pH 5,5, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %
20*	tampón acetato 20 mM, pH 5,5, glicerol al 35 %

Por ejemplo, la invención proporciona formulaciones que comprenden al menos una enzima de la invención y que comprenden un tampón (formulación) de: pH 7, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 7, glicerol al 35 %, 300 ppm de proxel; pH 7, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 7, cloruro sódico al 10 %, glicerol al 25 %, 300 ppm de proxel; pH 5,5, glicerol al 35 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; pH 5,5, glicerol al 35 %, 300 ppm de proxel; pH 5,5, cloruro sódico al 10 %, glicerol al 25 %, benzoato sódico al 0,1 %, sorbato potásico al 0,1 %; o, tampón acetato 20 mM, pH 5,5, glicerol al 35 %; MOPS 20 mM, pH 7 o MOPS 25 mM, NaCl 50 mM, pH 7,5; pH 5,0, TRIS 40 mM; pH 7,0, TRIS 40 mM; pH 8,0, TRIS 40 mM; pH 7,5, glicerol al 50 %; pH 7,5, NaCl al 20 %; pH 7,5, propilenglicol al 30 %; pH 7,5, sulfato sódico 100 mM; pH 5,5, glicerol al 35 %; o, cualquier combinación de los mismos, o, con equivalentes de los mismos.

Aplicación de Biolimpieza a modo de Ejemplo

- En un aspecto, el pH es pH 8,5 (tampón de bicarbonato)
- Agente humectante no iónico (1 g/l) [por ejemplo: Apollowet NFW]
- Proporción de agua madre en el baño en el baño enzimático: de 10:1 a 50:1 (l de agua madre:kg de tejido)
- Dosis de enzima: 0,137 ml del extracto concentrado por kg de tejido
- Intervalo de temperatura: entre aproximadamente 50 °C y 70 °C
- Tiempo de tratamiento de aproximadamente 20 min
- Los agentes quelantes se deberían excluir del baño enzimático, y solamente se deberían añadir después de 20 minutos de tratamiento enzimático y mantener durante 10 minutos antes del baño de descarga.

Por lo tanto, en la invención proporciona un proceso de biolimpieza que usa al menos una enzima de la invención que comprende al menos, varias o todas las siguientes etapas/limitaciones: el pH es pH 8,5, en tampón de bicarbonato, que comprende una gente humectante no iónico (a, por ejemplo, 1 g/l), en la que la proporción de agua madre en el baño enzimático es de aproximadamente 10:1 a 50:1 (l de agua madre:kg de tejido), en la que la dosis enzimática está entre aproximadamente 0,1 y 0,2 ml, por ejemplo, a aproximadamente 0,137 ml del extracto concentrado por kg de tejido, a un intervalo de temperatura entre aproximadamente 50 °C y 70 °C; con un tiempo de tratamiento de aproximadamente 20 min; y, en un aspecto, comprende agentes quelantes, que se deberían excluir del baño enzimático y solamente se debían añadir después de 20 minutos de tratamiento enzimático y mantener durante 10 minutos antes del baño de descarga.

La enzima de la SEC ID N°: 134 funcionó bien en el intervalo de 5 a 25 gramos de enzima pura por ton de tejido tratado.

LISTADO DE SECUENCIAS

<110> Diversa Corporation
Kerovuo, Janne
Solbak, Arne
Gray, Kevin
McCann, Ryan
Purohit, Shalaka

Gerendash, Joel
Janssen, Giselle
Dahod, Samun

5 <120> PECTATO LIASAS, ÁCIDOS NUCLEICOS QUE LAS CODIFICAN Y MÉTODOS PARA SU PREPARACIÓN Y USO

<130> 564462009640

10 <140> Pendiente de asignación
<141> Simultáneamente con la presente

<150> 60/460,842

<151> 04-04-2003

15 <150> 60/484,798

<151> 07-03-2003

<160> 134

20 <170> FastSEQ para Windows Versión 4.0

<210> 1

<211> 1917

25 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

30 <400> 1

```

gtgtctctct ttagaaaact cgcactgctg gttctgtgog gtctactgct ttctgtcgga 60
gcagaaaccc gagcgtcgaa gcgaattgtc gtggcgcgtg atggatcggg tgacgtcagg 120
acaattcaac aagcgttgga ccagggtccc aaagacaata cacaccgggt cttgattcag 180
atcaaaccgg gtgtgtatca ggaacagggt cgtgtcgccg ccggcaaacg ctttatcaact 240
tttcgcggcg acgatgcgag caagaccgtc atcacctatc gattgagcgc actgcaagcg 300
ggaaataccc ggctggcatt caccaccttc gttaatgcag acgactttcg cgccgagaac 360
ctgacgtttg aaaactcctt cggcacccgt tcacaagogg ttgctttgtt tgcgatgcg 420
gaccgcgcga cgtttgaaaa ctgccggttc ctccggttggc aggacacttt gtttgtgaac 480
ggcagccgcc acttcttcaa agactgctac gtcgaaggcc acgtcgattt cttttcggc 540
acggcctccg ccgtgtttga gaactgcacc attcacagca aaggcgaagg ttatgtgacc 600
gcacactatc gcaccagcga tgagatggat accggttttg tctttcatcg ttgtcgtttg 660
accggacgag acacgggccg cggagtttat ctccgaaggc cttggcgacc ttacgcgcgc 720
gtcgtcttta tcgattgctg gctggacgca cacatcagac ctgaaggctg ggataattgg 780
agagatcctg aacgagagaa gaccgcgtgg tttgccgagt acaagtcaaa agggcccggg 840
gctaataccc tagctcgtgt cgcgtggctc aggcagagca cgacagaaca aggcaccgag 900
ttttcgcggg aacgcttttt cagccgcgct gttcgcgggc tctctgggca ggccaaccag 960
gcagtcggaa cgatcgcgtg ggacgatgcg cagaaaaaac cgaacgagtg gtatgcgagc 1020
gccgaggcgt tgcgcattgc cgacaacggt gttctttatc aacgtgactc cggcggttgg 1080
cccaagaaca tcgacatggg gaagccgctc gaogaaaagg gtcgagccgg tcttctgcgc 1140
gtgcgtaaga agaacgattc cacgatcgac aatggcgcga cttacacgca actctcgttt 1200
ctggcgcggg tttacacggc gcaaaaagcag gagcggcatc gcgagtcggt tctgaaggga 1260
ctcgattacc tgttgaaggc gcagtatcca aacggaggct ggccgcagtt ctatcccaac 1320
ctcaacggct attacaacaa catcactttc aacgacaacg ccatgatcgg cgtgatgaaa 1380
ctgctgcgcg acgtagcgac agcgaacccg gcgtatgcgt tcgtcgacga agcacgacgg 1440
acgagtgcgg cgaaggcggg cgaaaaagga atcgagtga tactgaagac gcaggtggtt 1500
gtgaatggcc ggcgaccgt gtggtgtgcy caacatgacg aagtcacgct cgcgcctgcc 1560
ccggcgagga cgtttgaatt agtttcgctg agtggtggtg aaagcgttga gatcgtgcgc 1620
tttttgatgt cgatcaagaa cccgtgcgcg gcggttgcg aggcgatcga gtcggcgggt 1680
gcgtggttcg agcaatcgca agtgaaagat cccgccggca aacctgcgtg ggcgcgattt 1740
tatgagatcg gcaactaatc tccgatcttc gccgggcgtg acggcgtcgt taagtatgat 1800
gtgaaacaga tcgatgagga acgacgaaag aattacgcat ggtacgttga cgacgcagcg 1860
aaactactga aaaccgacta tcctgagtggt aaagaaaaga acgccaaga tcaatga 1917

```

<210> 2
 <211> 638
 <212> PRT
 5 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 10 <222> (1)...(21)
 <221> DOMINIO
 <222> (28)...(308)
 <223> Dominio de pectin metil esterasa
 <221> DOMINIO
 15 <222> (309)...(638)
 <223> Dominio catalítico

<400> 2

	Met	Ser	Leu	Phe	Arg	Lys	Leu	Ala	Leu	Leu	Val	Leu	Cys	Gly	Leu	Leu
	1				5					10					15	
	Leu	Ser	Val	Gly	Ala	Glu	Thr	Arg	Ala	Ser	Lys	Arg	Ile	Val	Val	Ala
				20					25					30		
	Ala	Asp	Gly	Ser	Gly	Asp	Val	Arg	Thr	Ile	Gln	Gln	Ala	Val	Asp	Gln
			35					40					45			
	Val	Pro	Lys	Asp	Asn	Thr	His	Pro	Val	Leu	Ile	Gln	Ile	Lys	Pro	Gly
		50				55						60				
	Val	Tyr	Gln	Glu	Gln	Val	Arg	Val	Ala	Ala	Gly	Lys	Arg	Phe	Ile	Thr
	65				70						75				80	
	Phe	Arg	Gly	Asp	Asp	Ala	Ser	Lys	Thr	Val	Ile	Thr	Tyr	Arg	Leu	Ser
				85						90					95	
	Ala	Leu	Gln	Ala	Gly	Asn	Thr	Arg	Leu	Ala	Phe	Thr	Thr	Phe	Val	Asn
			100						105					110		
	Ala	Asp	Asp	Phe	Arg	Ala	Glu	Asn	Leu	Thr	Phe	Glu	Asn	Ser	Phe	Gly
		115					120					125				
	Thr	Gly	Ser	Gln	Ala	Val	Ala	Leu	Phe	Val	Asp	Ala	Asp	Arg	Ala	Thr
		130				135						140				
	Phe	Glu	Asn	Cys	Arg	Phe	Leu	Gly	Trp	Gln	Asp	Thr	Leu	Phe	Val	Asn
	145				150						155				160	
	Gly	Ser	Arg	His	Phe	Phe	Lys	Asp	Cys	Tyr	Val	Glu	Gly	His	Val	Asp
				165						170					175	
	Phe	Ile	Phe	Gly	Thr	Ala	Ser	Ala	Val	Phe	Glu	Asn	Cys	Thr	Ile	His
			180					185					190			
20	Ser	Lys	Gly	Glu	Gly	Tyr	Val	Thr	Ala	His	Tyr	Arg	Thr	Ser	Asp	Glu

Met	Asp	Thr	Gly	Phe	Val	Phe	His	Arg	Cys	Arg	Leu	Thr	Gly	Arg	Asp
210						215					220				
Thr	Gly	Arg	Gly	Val	Tyr	Leu	Gly	Arg	Pro	Trp	Arg	Pro	Tyr	Ala	Arg
225					230					235					240
Val	Val	Phe	Ile	Asp	Cys	Trp	Leu	Asp	Ala	His	Ile	Arg	Pro	Glu	Gly
				245					250					255	
Trp	Asp	Asn	Trp	Arg	Asp	Pro	Glu	Arg	Glu	Lys	Thr	Ala	Trp	Phe	Ala
			260					265					270		
Glu	Tyr	Lys	Ser	Lys	Gly	Pro	Gly	Ala	Asn	Pro	Val	Ala	Arg	Val	Ala
		275					280					285			
Trp	Ser	Arg	Gln	Leu	Thr	Thr	Glu	Gln	Ala	Ala	Glu	Phe	Ser	Arg	Glu
	290					295					300				
Arg	Phe	Phe	Ser	Arg	Ala	Val	Arg	Gly	Leu	Ser	Gly	Gln	Ala	Asn	Gln
305					310					315					320
Ala	Val	Gly	Thr	Ile	Ala	Trp	Asp	Asp	Ala	Gln	Lys	Lys	Pro	Asn	Glu
				325					330					335	
Trp	Tyr	Ala	Ser	Ala	Glu	Ala	Leu	Arg	Ile	Ala	Asp	Asn	Val	Val	Leu
			340					345					350		
Tyr	Gln	Arg	Asp	Ser	Gly	Gly	Trp	Pro	Lys	Asn	Ile	Asp	Met	Gly	Lys
		355					360					365			
Pro	Leu	Asp	Glu	Lys	Gly	Arg	Ala	Gly	Leu	Leu	Arg	Val	Arg	Lys	Lys
	370					375					380				
Asn	Asp	Ser	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Tyr	Thr	Gln	Leu	Ser	Phe
385					390					395					400
Leu	Ala	Arg	Val	Tyr	Thr	Ala	Gln	Lys	Gln	Glu	Arg	His	Arg	Glu	Ser
				405					410					415	
Phe	Leu	Lys	Gly	Leu	Asp	Tyr	Leu	Leu	Lys	Ala	Gln	Tyr	Pro	Asn	Gly
			420					425					430		
Gly	Trp	Pro	Gln	Phe	Tyr	Pro	Asn	Leu	Asn	Gly	Tyr	Tyr	Lys	His	Ile
		435					440					445			
Thr	Phe	Asn	Asp	Asn	Ala	Met	Ile	Gly	Val	Met	Lys	Leu	Leu	Arg	Asp
	450					455					460				
Val	Ala	Thr	Ala	Lys	Pro	Ala	Tyr	Ala	Phe	Val	Asp	Glu	Ala	Arg	Arg
465					470					475					480
Thr	Ser	Ala	Ala	Lys	Ala	Val	Glu	Lys	Gly	Ile	Glu	Cys	Ile	Leu	Lys
				485					490					495	
Thr	Gln	Val	Val	Val	Asn	Gly	Arg	Arg	Thr	Val	Trp	Cys	Ala	Gln	His
			500					505					510		
Asp	Glu	Val	Thr	Leu	Ala	Pro	Ala	Pro	Ala	Arg	Thr	Phe	Glu	Leu	Val
		515					520					525			
Ser	Leu	Ser	Gly	Gly	Glu	Ser	Val	Glu	Ile	Val	Arg	Phe	Leu	Met	Ser
	530					535					540				
Ile	Lys	Asn													

<210> 3
<211> 1416
<212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 3

```

atgtcgtcac gacgcgagtt cattagagat ctgttgactg ggggcgcact gatcgccgtc      60
gcgcgcgctc tgtctgcgtt tgcagcggag gagaatccgt gggaaaacggt gatgccttcg      120
atcgtgaaac gcatcaagcg acctcgtttc ccgatgcgca cgtttgatct caccggagttt      180
ggagcgaaag gtgatggacg aacagattgc acgttggctt tccgtcgcgc gatcgatcga      240
tgcacgaacg ccggtggtgg gagagtagtt gtcccaccgg gttcgtatct cactggcgcc      300
attcatttga agagcaacgt cgaccttcac atctcagaag gtactacggt caagttcagc      360
cagaacccga aagactacct gcccggtgtt ttctcgcgtt gggaaggcgt cgaggtgttc      420
aactactcgc cttttatcta cgcttcgaa caaacgaaca ttgcgatcac tggcaagggc      480
acgctcaacg gtcaaagcga caacgaacac tggtgccctt ggaaacggacg tgcgcgtac      540
ggctggaag aagggatgag caatcagcgt cccgatcgaa atgcgctgtt tgcgatggcc      600
gaaaaaggtg tcccggttca ggagcgcatt tttggtgagg gccattactt aaggccgcag      660
ttcattcaac cttatcgttg tgagaacgtg ctgatcgaag gtgtcactat tcgaaactcg      720
ccgatgtggg aaattcatcc ggtgctctgc cggaatgtca tcgtccaaaa tgtgatcatc      780
aacagtcatt gtccaaacaa cgacgggtgt aatcctgagt cgtgcacgga tgtgttgatt      840
aaggattgtg acttcgacac tggtgacgat tgtatcgaga tcaagtcagg ccgaaatgca      900
gatggggcgg gactgaaggc tctactgaa aacattatcg tgactgggtg tcgcatgaaa      960
gatggtcacg gcgggattac ggtgggcagc gagatttcgg gtggggtgcg aaatcttttc     1020
gcatccaact gccggtctga cagtccgaac ctggaccatg cattgcgggt taagaataac     1080
gctatgcgtg gcgggctgtt ggagaatctg cacttcogaa atatcgacgt cgggcaagtg     1140
gcgcacgcgg tgatcacgat cgatttcaat tatgaggaag gcgcgaaggg atcgttcacg     1200
ccagtcgttc gtgattacac cgtcgtatgg cttcgcagca cgaagagtaa gtacgcgctc     1260
gatgtgcagg gcttggcgac ggcgcggatc gtgaatctgc gtctaaccac ctgcatcttc     1320
gacaatgtcg ctgaaggaaa tgttgtgaag aacgtaaagg atgcaactat cgagaatgtc     1380
aaaatcaatg gaaaaagcgt tgatgcagtg ccgtag                                1416

```

<210> 4

5 <211> 471

<212> PRT

<213> Desconocido

<220>

10 <223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(28)

<221> DOMINIO

<222> (81)...(476)

15 <223> Dominio catalítico

<400> 4

```

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

```

Leu	Thr	Gly	Ala	Ile	His	Leu	Lys	Ser	Asn	Val	Asp	Leu	His	Ile	Ser
			100					105					110		
Glu	Gly	Thr	Thr	Val	Lys	Phe	Ser	Gln	Asn	Pro	Lys	Asp	Tyr	Leu	Pro
			115				120					125			
Val	Val	Phe	Ser	Arg	Trp	Glu	Gly	Val	Glu	Val	Phe	Asn	Tyr	Ser	Pro
			130			135					140				
Phe	Ile	Tyr	Ala	Phe	Glu	Gln	Thr	Asn	Ile	Ala	Ile	Thr	Gly	Lys	Gly
145					150					155					160
Thr	Leu	Asn	Gly	Gln	Ser	Asp	Asn	Glu	His	Trp	Trp	Pro	Trp	Asn	Gly
			165						170					175	
Arg	Ala	Ala	Tyr	Gly	Trp	Lys	Glu	Gly	Met	Ser	Asn	Gln	Arg	Pro	Asp
			180					185					190		
Arg	Asn	Ala	Leu	Phe	Ala	Met	Ala	Glu	Lys	Gly	Val	Pro	Val	Gln	Glu
			195				200					205			
Arg	Ile	Phe	Gly	Glu	Gly	His	Tyr	Leu	Arg	Pro	Gln	Phe	Ile	Gln	Pro
210					215						220				
Tyr	Arg	Cys	Glu	Asn	Val	Leu	Ile	Glu	Gly	Val	Thr	Ile	Arg	Asn	Ser
225				230						235					240
Pro	Met	Trp	Glu	Ile	His	Pro	Val	Leu	Cys	Arg	Asn	Val	Ile	Val	Gln
			245						250					255	
Asn	Val	Ile	Ile	Asn	Ser	His	Gly	Pro	Asn	Asn	Asp	Gly	Cys	Asn	Pro
			260					265					270		
Glu	Ser	Cys	Thr	Asp	Val	Leu	Ile	Lys	Asp	Cys	Asp	Phe	Asp	Thr	Gly
			275				280					285			
Asp	Asp	Cys	Ile	Ala	Ile	Lys	Ser	Gly	Arg	Asn	Ala	Asp	Gly	Arg	Arg
290					295						300				
Leu	Lys	Ala	Pro	Thr	Glu	Asn	Ile	Ile	Val	Thr	Gly	Cys	Arg	Met	Lys
305					310					315					320
Asp	Gly	His	Gly	Gly	Ile	Thr	Val	Gly	Ser	Glu	Ile	Ser	Gly	Gly	Val
			325					330						335	
Arg	Asn	Leu	Phe	Ala	Ser	Asn	Cys	Arg	Leu	Asp	Ser	Pro	Asn	Leu	Asp
			340					345					350		
His	Ala	Leu	Arg	Val	Lys	Asn	Asn	Ala	Met	Arg	Gly	Gly	Leu	Leu	Glu
			355				360					365			
Asn	Leu	His	Phe	Arg	Asn	Ile	Asp	Val	Gly	Gln	Val	Ala	His	Ala	Val
370					375						380				
Ile	Thr	Ile	Asp	Phe	Asn	Tyr	Glu	Glu	Gly	Ala	Lys	Gly	Ser	Phe	Thr
385					390					395					400
Pro	Val	Val	Arg	Asp	Tyr	Thr	Val	Asp	Gly	Leu	Arg	Ser	Thr	Lys	Ser
			405						410					415	
Lys	Tyr	Ala	Leu	Asp	Val	Gln	Gly	Leu	Ala	Thr	Ala	Pro	Ile	Val	Asn
			420					425					430		
Leu	Arg	Leu	Thr	Asn	Cys	Ile	Phe	Asp	Asn	Val	Ala	Glu	Gly	Asn	Val
			435				440					445			
Val	Lys	Asn	Val	Lys	Asp	Ala	Thr	Ile	Glu	Asn	Val	Lys	Ile	Asn	Gly
			450			455					460				
Lys	Ser	Val	Asp	Ala	Val	Pro									
465						470									

<210> 5
 <211> 1077
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 5

ttggacgaca agtgggctga gcggacatcg cccgatttca atctcgtctc gtggaatgaa

60

```

attctaaagc agccgaaact ttggtacgcg gtcgacgaag cgacgcggat cgcaaatacag 120
gtgatccttt atcaacgcga caacggtggt tggcgaaga atatacagat ggccgccatg 180
ctcatgcagg cagaacgcga aaaacttagt cgcgagaaga gcgagaccga cagacaatc 240
gacaacggcg cgacgacaac ccagctcgcg tatctggcga aggtcatcac ggccaagaat 300
atcgaaagcc atcgctcgcg gtttttcaaa ggccctcgatt ttcttttcgc catgcagtac 360
gggaatggcg gcttcccgcga attttttcct ctgctgacg attattcgcg cgagattacg 420
ttcaacgaca acgcgatgat aaatgtgctt cggttgctcc gcgacatagc cgatcgaaag 480
aacgattatg tgtttgtcga tgaagagcgg cgagcgaagg ccgagcaggc tgtaaggcgt 540
gcgatcccggt tgatcctcag cacgcaggtc gtcgtcgatg gaaagaaaac cgtctgggct 600
tgcgagtagt atgaagccg attgaagccg gccgcggcgc gaaagttcga gccggcatca 660
ttgaccggcg gcgagagcgt tggcatcgct cgggttttga tgctagaaaa accaacaccc 720
gagatcatta acgcgatcga atccgccatc gcttgggtaca aggcgaacaa catctcgga 780
cttaggtggg agaggcgaaa cggcgagaac attgtgatca aagacaagaa cgcgcgcgcc 840
gtctggggcg gcttttatca gatcgaaacg atgaggccga tcttcgccgg tcgcatgctg 900
gtcatcagat acgatgtgat gcagatcgag tcggaacgctc gaaacggata tgcattgtac 960
gtatccgaac cgaatgagtt gttgaatgaa gattatccga agtggaggac aaggagtgcg 1020
aagcgtgccg agatctttca acgtccgcct cttggttcga gatttcggac cgtgtag 1077

```

<210> 6

<211> 358

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10 <221> DOMINIO

<222> (1)...(358)

<223> Dominio catalítico

<400> 6

15

```

Met Asp Asp Lys Trp Ala Glu Arg Thr Ser Pro Asp Phe Asn Leu Val
1      5      10      15
Ser Trp Asn Glu Ile Leu Lys Gln Pro Lys Leu Trp Tyr Ala Val Asp
20     25     30
Glu Ala Thr Arg Ile Ala Asn Gln Val Ile Leu Tyr Gln Arg Asp Asn
35     40     45
Gly Gly Trp Pro Lys Asn Ile Asp Met Ala Ala Met Leu Met Gln Ala
50     55     60
Glu Arg Glu Lys Leu Ser Arg Glu Lys Ser Glu Thr Asp Thr Thr Ile
65     70     75     80
Asp Asn Gly Ala Thr Thr Thr Gln Leu Ala Tyr Leu Ala Lys Val Ile
85     90     95
Thr Ala Lys Asn Ile Glu Ser His Arg Val Ala Phe Phe Lys Gly Leu
100    105    110
Asp Phe Leu Phe Ala Met Gln Tyr Gly Asn Gly Gly Phe Pro Gln Phe
115    120    125
Phe Pro Leu Arg Asp Asp Tyr Ser Arg Glu Ile Thr Phe Asn Asp Asn
130    135    140
Ala Met Ile Asn Val Leu Arg Leu Leu Arg Asp Ile Ala Asp Arg Lys
145    150    155    160
Asn Asp Tyr Val Phe Val Asp Glu Glu Arg Arg Ala Lys Ala Glu Gln
165    170    175
Ala Val Arg Arg Ala Ile Pro Leu Ile Leu Ser Thr Gln Val Val Val
180    185    190
Asp Gly Lys Lys Thr Val Trp Ala Ala Gln Tyr Asp Glu Lys Thr Leu
195    200    205
Lys Pro Ala Ala Ala Arg Lys Phe Glu Pro Ala Ser Leu Thr Ala Gly
210    215    220

```

Glu Ser Val Gly Ile Val Arg Phe Leu Met Leu Glu Lys Pro Thr Pro
 225 230 235 240
 Glu Ile Ile Asn Ala Ile Glu Ser Ala Ile Ala Trp Tyr Lys Ala Asn
 245 250 255
 Asn Ile Ser Gly Leu Arg Trp Glu Arg Arg Asn Gly Glu Asn Ile Val
 260 265 270
 Ile Lys Asp Lys Asn Ala Pro Pro Val Trp Ala Arg Phe Tyr Gln Ile
 275 280 285
 Glu Thr Met Arg Pro Ile Phe Ala Gly Arg Asp Ala Val Ile Arg Tyr
 290 295 300
 Asp Val Met Gln Ile Glu Ser Glu Arg Arg Asn Gly Tyr Ala Trp Tyr
 305 310 315 320
 Val Ser Glu Pro Asn Glu Leu Leu Asn Glu Asp Tyr Pro Lys Trp Arg
 325 330 335
 Thr Arg Ser Ala Lys Arg Ala Gln Ile Phe Gln Arg Pro Pro Leu Gly
 340 345 350
 Ser Arg Phe Arg Thr Val
 355

<210> 7
 <211> 1125
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 7

gtgcatgcgg ggcgcaaaca cgtgagccga tggcgcgaaag agttcctgcg cgacttcgcc 60
 gcgcggctct cccgaaccat tccgtcctcg ccggcgcgaga gcgctgcggt cagcgggggt 120
 ccggcgcgga tccgtggtgg agcggagctc ctgcggcgaga agccggagtg gtatgcctcg 180
 cgagaggcga ggacgatcgc cgacagcgtc atccagtacc aggcggcgga cggcggtggt 240
 cccaagaaca ccgacctcgg gactccgccc acggctgaat cagcgccggg cgcggcgggc 300
 gacgtgacgt cgagcaccat cgacaacaac ggcaacgaga tgccgatgca gttccttgcg 360
 ctggtggcgg acgcgaccgg cgaggctcgc tatcgcgctg cgttccttcg cggcttcgac 420
 tacctgctcg ccgcgcagta tcccaacggc ggctggccgc agttctttcc gctccgcccgc 480
 gggattaca cccacatcac cttcaacgac aacgcgatgg tcaacgtgct gacggtgctg 540
 cgcgatgccg cggccggtca ggcgccatac gccttcgtgg acgagccccg ccgcgcgaag 600
 gccgcgcccg ccgtgtcccg ggggatcgac gtcacctga agaccgaagt gaaacagaac 660
 ggcaagctga cggcggtgtg cgcgcagcac gacgagaaga ccctcgcgcc ggcggtggcg 720
 cgcgcttacg agccgccatc gctctccggc agcgaaaccg tcggcatcgt ccgcttcctg 780
 atggagatcg agaagccgtc accggagatc gtcgccgga tcgaaggggc cgtcgccctg 840
 ctgaagtcgg tggcgattcc ggggctgcgc tacgagtcct tcaccggcgc ggacggacag 900
 agggaccgcc gcgtcgttcc ggatccatcg gccggactcc tgtgggcgcg gttctacgag 960
 ctccgcacca accggccgat cttcctcggc cgcgactccg tggttcgcgc cgcgctcagt 1020
 gacatcgaac gcgagcgccg cgcgggctac gcctactacg gtacgtggcc ggcgagcctg 1080
 attgctgcgg actaccgcgc ctggcggtcg acgctccggc gctga 1125

<210> 8
 <211> 374
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> DOMINIO
 <222> (1)...(374)
 <223> Dominio catalítico

<400> 8

Met His Ala Gly Ala Lys His Val Ser Arg Trp Arg Glu Glu Phe Leu
1 5 10 15
Arg Asp Phe Ala Ala Arg Leu Ser Arg Thr Ile Pro Ser Ser Pro Ala
20 25 30
Gln Ser Ala Ala Val Ser Gly Val Pro Ala Ala Ile Arg Trp Gly Ala
35 40 45
Asp Val Leu Arg Gln Lys Pro Glu Trp Tyr Ala Ser Arg Glu Ala Arg
50 55 60
Thr Ile Ala Asp Ser Val Ile Gln Tyr Gln Ala Ala Asp Gly Gly Trp
65 70 75 80
Pro Lys Asn Thr Asp Leu Gly Thr Pro Pro Thr Ala Glu Ser Arg Ala
85 90 95
Gly Ala Ala Ala Asp Val Thr Ser Ser Thr Ile Asp Asn Asn Gly Thr
100 105 110
Thr Met Pro Met Gln Phe Leu Ala Leu Val Ala Asp Ala Thr Gly Glu
115 120 125
Ala Arg Tyr Arg Ala Ser Phe Leu Arg Gly Phe Asp Tyr Leu Leu Ala
130 135 140
Ala Gln Tyr Pro Asn Gly Gly Trp Pro Gln Phe Phe Pro Leu Arg Arg
145 150 155 160
Gly Tyr Tyr Thr His Ile Thr Phe Asn Asp Asn Ala Met Val Asn Val
165 170 175
Leu Thr Val Leu Arg Asp Ala Ala Ala Gly Gln Ala Pro Tyr Ala Phe
180 185 190
Val Asp Glu Pro Arg Arg Ala Lys Ala Arg Ala Ala Val Ser Arg Gly
195 200 205
Ile Asp Val Ile Leu Lys Thr Gln Val Lys Gln Asn Gly Lys Leu Thr
210 215 220
Ala Trp Cys Ala Gln His Asp Glu Lys Thr Leu Ala Pro Ala Trp Ala
225 230 235 240
Arg Ala Tyr Glu Pro Pro Ser Leu Ser Gly Ser Glu Thr Val Gly Ile
245 250 255
Val Arg Phe Leu Met Glu Ile Glu Lys Pro Ser Pro Glu Ile Val Ala
260 265 270
Ala Ile Glu Gly Ala Val Ala Trp Leu Lys Ser Val Ala Ile Pro Gly
275 280 285
Leu Arg Tyr Glu Ser Phe Thr Gly Ala Asp Gly Gln Arg Asp Arg Arg
290 295 300
Val Val Pro Asp Pro Ser Ala Gly Leu Leu Trp Ala Arg Phe Tyr Glu
305 310 315 320
Leu Gly Thr Asn Arg Pro Ile Phe Leu Gly Arg Asp Ser Val Val Arg
325 330 335
Ala Ala Leu Ser Asp Ile Glu Arg Glu Arg Arg Ala Gly Tyr Ala Tyr
340 345 350
Tyr Gly Thr Trp Pro Ala Ser Leu Ile Ala Ala Asp Tyr Pro Arg Trp
355 360 365
Arg Ser Thr Leu Arg Arg
370

<210> 9
<211> 1116
5 <212> ADN
<213> Desconocido

<220>
10 <223> Obtenido de una muestra ambiental

<400> 9

```

ttgatcggtg gcatgaaaac gattctctca aatctgaacg cggcgctgct ttcattgcgcc 60
ctgctctttg cggcagccac acaggggaacc aagccgcccg aagtgcggtg gaatgagtgc 120
ctaaacccaaa aacctgcctg gtacggcagc ccggaagcgg tgcgcattgc tgacaacctg 180
ttgctttacc aacgcgacca cggcggctgg cacaagaata tcgaaatggc tgcggtcttg 240
accgaacagc aacaggcaga gttgaaagcg caaaaggcaa ccgacgattc gacgattgat 300
aacggcgcgga cctataccca ggtgatttat ctggcgcgcg tcttcaatgc gacgaagcag 360
gagcgattca aaaccgcgtt tctcaaagga ttcgattatc tgctcaaggc tcagtatgcg 420
aacggcggtt ggccgcagta ttaccgcgtt ttgcagggtt attacaaca catcacgttc 480
aacgatgacg cgatggtcgg cgtgcttgat cttctgcgcg atgttgcgcg cggcgattcc 540
ggttatccgt tcgtggacag cgaccggcgc gcccgcgcca gccaggccgt gcaaaaagga 600
attgaagtga tcttgaaatg ccagatcgtg gtccgcccga aaaagaccgc ctggtgcgcg 660
caacacgatg aagtgcattt cgcggcgcg cgggcacgca cctacgagaa aatttcgctg 720
agcggcagcg aatcggtttg cctgatccgc ttctgatgg gcattgaaca accggacgcg 780
cgtgtagttg aggcgattga gtccgccgtt gcctggctca agcaagccaa gctgaccggc 840
atcaaagtgg ttcagaaggc ggatgcttcg aagcccaatg gcttcgaccg ggtcgtcgtt 900
gaagatgcac aagccgggcc attgtggcg cgcttttacg agatcgggtac gggccgcccg 960
atcttttccg gacgtgacgg catcgtcaaa tacagcttg cggaaatcga acacgaacgg 1020
cgcacgggct acggctggtg cacgaatgcg cccgcgaaat tgctggaaca agattatccg 1080
gcctggcaaa tcaaacgcgg gggcaagaaa aagtaa 1116

```

<210> 10
 <211> 371
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 10 <221> SEÑAL
 <222> (1)...(29)
 <221> DOMINIO
 <222> (30)...(371)
 <223> Dominio catalítico

15 <400> 10

```

Met Ile Gly Ser Met Lys Thr Ile Leu Ser Asn Leu Asn Ala Ala Leu
 1           5           10           15
Leu Ser Cys Ala Leu Leu Phe Ala Ala Ala Thr Gln Gly Thr Lys Pro
          20          25          30
Pro Glu Val Arg Trp Asn Glu Cys Leu Asn Gln Lys Pro Ala Trp Tyr
          35          40          45
Gly Ser Pro Glu Ala Val Arg Ile Ala Asp Asn Leu Leu Tyr Gln
          50          55          60
Arg Asp His Gly Gly Trp His Lys Asn Ile Glu Met Ala Ala Val Leu
          65          70          75          80
Thr Glu Gln Gln Gln Ala Glu Leu Lys Ala Gln Lys Ala Thr Asp Asp
          85          90          95
Ser Thr Ile Asp Asn Gly Ala Thr Tyr Thr Gln Val Ile Tyr Leu Ala
          100         105         110
Arg Val Phe Asn Ala Thr Lys Gln Glu Arg Phe Lys Thr Ala Phe Leu
          115         120         125
Lys Gly Phe Asp Tyr Leu Leu Lys Ala Gln Tyr Ala Asn Gly Gly Trp
          130         135         140
Pro Gln Tyr Tyr Pro Arg Leu Gln Gly Tyr Tyr Lys His Ile Thr Phe
          145         150         155         160
Asn Asp Asp Ala Met Val Gly Val Leu Asp Leu Leu Arg Asp Val Ala
          165         170         175

```

Arg Gly Asp Ser Gly Tyr Arg Phe Val Asp Ser Asp Arg Arg Ala Arg
180 185 190
Ala Ser Gln Ala Val Gln Lys Gly Ile Glu Cys Ile Leu Lys Cys Gln
195 200 205
Ile Val Val Ala Gly Lys Lys Thr Ala Trp Cys Ala Gln His Asp Glu
210 215 220
Val Thr Phe Ala Pro Ala Pro Ala Arg Thr Tyr Glu Lys Ile Ser Leu
225 230 235 240
Ser Gly Ser Glu Ser Val Gly Leu Ile Arg Phe Leu Met Gly Ile Glu
245 250 255
Gln Pro Asp Ala Arg Val Val Glu Ala Ile Glu Ser Ala Val Ala Trp
260 265 270
Leu Lys Gln Ala Lys Leu Thr Gly Ile Lys Val Val Gln Lys Ala Asp
275 280 285
Ala Ser Lys Pro Asn Gly Phe Asp Arg Val Val Val Glu Asp Ala Gln
290 295 300
Ala Gly Pro Leu Trp Ala Arg Phe Tyr Glu Ile Gly Thr Gly Arg Pro
305 310 315 320
Ile Phe Ser Gly Arg Asp Gly Ile Val Lys Tyr Ser Leu Ala Glu Ile
325 330 335
Glu His Glu Arg Arg Thr Gly Tyr Gly Trp Tyr Thr Asn Ala Pro Ala
340 345 350
Lys Leu Leu Glu Gln Asp Tyr Pro Ala Trp Gln Ile Lys Arg Gly Gly
355 360 365
Lys Lys Lys
370

<210> 11
<211> 1167
5 <212> ADN
<213> Desconocido

<220>
10 <223> Obtenido de una muestra ambiental

<400> 11

atgtcgttgg	gaccaggtgc	taatccgaaa	gctcgcgttc	cctgggtccaa	acaactatcg	60
ggtgttgagg	caaaagtgtt	cgatcgcgag	cggttcttca	gcctcgtctg	ggaacgaacc	120
tctaagaaga	atgaccagca	agtcggcgcc	atcgcgtgga	aagatgcaca	cggaaaggca	180
gatgagtgg	atgcgagcgt	tgaggcactt	cgtatagccg	ataacgtcgt	tttctatcaa	240
cgtgactcag	gtggctggcc	caagaatata	gagatggcga	agacgttgag	cgatcgtgag	300
aaggctgcga	ttctccgcga	gaagaaaaag	aatgactcaa	caatcgacaa	tggcgcgact	360
cacactcagt	tatcttttct	ggcgcgcgtc	tatacagcac	aacagcagga	gcgacatcgc	420
gagtcatttt	taaaaggact	ggattactta	ctgaaggcgc	agtattcaaa	tgggtgctgg	480
ccacagtctt	atccaaactt	gaatggctac	tacaaacgga	tcacgtacaa	cgatggcgcg	540
atgatcgggtg	tgatgaagct	tctgcgtgat	gttgccggcag	cgaaacctga	atacgcgttt	600
gtcgatgaaa	ctcggcgtgc	gaaggctgcg	aacgcggtgg	aaaaaggcat	cgtgtgcatt	660
ttgaaaacgc	aggtggttgt	tgatgggcgt	cgcactgttt	ggtgtgcaca	acacgacgaa	720
gtgacgtttg	cgcccgcgcc	tgaagaaag	tttgagttag	cttcgttgag	cggcgggtgag	780
agcgtcgata	ttgttcgatt	tctaattgtcg	ataaaggatc	catcgcgtaa	cgtggttgaa	840
tcgattgaat	cggcagttaa	atggttttag	cagtcggagc	taaaaggcgt	taagtgggtc	900
aagaaaaccg	acgtactca	acctaattgg	ttcgattgtg	tcgttggtta	agatccggag	960
agctctgttt	gggcgcgctt	ttacgagatt	ggcacgaacc	gcccgatctt	tgccgggcgt	1020
gatggagtgc	ctaagtatga	cgtcgcgcag	atcgaacacg	agcgacgaac	gggttacgaa	1080
tggtacgttg	atgagcgagc	aaaactgctg	aaaaagatt	atccggcggtg	gaagaaacga	1140
catgtcgtca	cgacgcgagt	tcattag				1167

15 <210> 12
<211> 388
<212> PRT
<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

<221> DOMINIO

<222> (1)...(388)

5 <223> Dominio catalítico

<400> 12

```

Met Ser Leu Gly Pro Gly Ala Asn Pro Lys Ala Arg Val Pro Trp Ser
 1      5      10      15
Lys Gln Leu Ser Gly Val Glu Ala Lys Leu Phe Asp Arg Glu Arg Phe
      20      25      30
Phe Ser Leu Ala Ala Glu Arg Thr Ser Lys Lys Asn Asp Gln Gln Val
      35      40      45
Gly Ala Ile Ala Trp Lys Asp Ala His Gly Lys Ala Asp Glu Trp Tyr
      50      55      60
Ala Ser Val Glu Ala Leu Arg Ile Ala Asp Asn Val Val Phe Tyr Gln
      65      70      75      80
Arg Asp Ser Gly Gly Trp Pro Lys Asn Ile Glu Met Ala Lys Thr Leu
      85      90      95
Ser Asp Arg Glu Lys Ala Ala Ile Leu Arg Glu Lys Lys Lys Asn Asp
      100      105      110
Ser Thr Ile Asp Asn Gly Ala Thr His Thr Gln Leu Ser Phe Leu Ala
      115      120      125
Arg Val Tyr Thr Ala Gln Gln Gln Glu Arg His Arg Glu Ser Phe Leu
      130      135      140
Lys Gly Leu Asp Tyr Leu Leu Lys Ala Gln Tyr Ser Asn Gly Gly Trp
      145      150      155      160
Pro Gln Phe Tyr Pro Asn Leu Asn Gly Tyr Tyr Lys Arg Ile Thr Tyr
      165      170      175
Asn Asp Gly Ala Met Ile Gly Val Met Lys Leu Leu Arg Asp Val Ala
      180      185      190
Ala Ala Lys Pro Glu Tyr Ala Phe Val Asp Glu Thr Arg Arg Ala Lys
      195      200      205
Ala Ala Asn Ala Val Glu Lys Gly Ile Val Cys Ile Leu Lys Thr Gln
      210      215      220
Val Val Val Asp Gly Arg Arg Thr Val Trp Cys Ala Gln His Asp Glu
      225      230      235      240
Val Thr Phe Ala Pro Ala Pro Ala Arg Lys Phe Glu Leu Ala Ser Leu
      245      250      255
Ser Gly Gly Glu Ser Val Asp Ile Val Arg Phe Leu Met Ser Ile Lys
      260      265      270
Asp Pro Ser Arg Asn Val Val Glu Ser Ile Glu Ser Ala Val Lys Trp
      275      280      285
Phe Glu Gln Ser Glu Leu Lys Gly Val Lys Trp Val Lys Lys Thr Asp
      290      295      300
Ala Thr Gln Pro Asn Gly Phe Asp Cys Val Val Val Lys Asp Pro Glu
      305      310      315      320
Ser Ser Val Trp Ala Arg Phe Tyr Glu Ile Gly Thr Asn Arg Pro Ile
      325      330      335
Phe Ala Gly Arg Asp Gly Val Pro Lys Tyr Asp Val Ala Gln Ile Glu
      340      345      350
His Glu Arg Arg Thr Gly Tyr Glu Trp Tyr Val Asp Glu Ala Ala Lys
      355      360      365
Leu Leu Lys Lys Asp Tyr Pro Ala Trp Lys Lys Arg His Val Val Thr
      370      375      380
Thr Arg Val His
      385

```

10

<210> 13

<211> 1065

<212> ADN

15 <213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

5 <400> 13

```

atgaaaacga tcagccttat ttgcctcgca atctctgctg ggattctgga ttcggttgcg      60
gcggcacgct ggaacgaatt cgcccagaag gcggatgatt ggtatcgagg tgacgaaggc      120
aggcgcgttg cttcgaatat tctttctcac caatcactgc aaggaagctg gcccaagaat      180
accgatacca ccgcgagatt cttcaatgga gatctagcga agattcaggg cacgttcgac      240
aacggtgcga cgacggacga gttgcgtttc ctggcccgcg cgtttgtcgc cagaaaagaa      300
aaaaactacg agtcagcggt ccgaaaaggc ttcgaacaca ttctcgcggc gcaatacgcg      360
aacggcggat ggccgcaata ttccgcccgc cccaaaagtt accaccgaca cattaccttc      420
aacgataatt cgatgggtgc gctgatgatt ttcttctgcg aggtcacgac ttcgaatctc      480
tactcgttcg tcgaagcgcc gctgcgaaca caagcccgcg aaagtttcga tcgcggtgtg      540
cgggtgcattc ttaagtgccg gatcgctcgt aacgggcaca agaccgcgtg gtgcgcgcaa      600
catgatgaaa cggatttcag ccccgatcc gcgcgtagtt acgaactgcc ttcgctgagc      660
ggttctgaat cagtcggcat tgtgcgcttg ctgatgagcc tcgatcagcc gagccgcgga      720
gtgatcgatg ccatcaccaa cgccgtagcg tggttcgaat cggcgaagct gcccgggatc      780
aaaaccggtc aagagaccga tccgaattcg cccaaaggct ggaatcgctg cgtcgtaaaa      840
gatgaaagtg cccgaccgat gtgggcgctg ttctacgaca tcaacaccaa caaaccgttc      900
ttttgtgatc gcgatggtgt gccaaagccg agtcttgccg agatcggtta tgaacggcgg      960
aacggttatg cgtggctcgg atactggcct gaagacttgc tcgcaagaga gtatccagcg     1020
tggaagatga agtggctgaa gcccaaagag cgcccagcat tttga      1065

```

<210> 14

10 <211> 354

<212> PRT

<213> Desconocido

<220>

15 <223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(22)

<221> DOMINIO

<222> (23)...(354)

20 <223> Dominio catalítico

<400> 14

```

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

```

Ala	Thr	Lys	Glu	Lys	Asn	Tyr	Glu	Ser	Ala	Phe	Arg	Lys	Gly	Phe	Glu
			100					105					110		
His	Ile	Leu	Ala	Ala	Gln	Tyr	Ala	Asn	Gly	Gly	Trp	Pro	Gln	Tyr	Ser
		115					120					125			
Pro	Pro	Pro	Lys	Ser	Tyr	His	Arg	His	Ile	Thr	Phe	Asn	Asp	Asn	Ser
		130				135					140				
Met	Val	Arg	Leu	Met	Ile	Phe	Leu	Arg	Glu	Val	Thr	Thr	Ser	Asn	Leu
145					150					155					160
Tyr	Ser	Phe	Val	Glu	Ala	Pro	Leu	Arg	Thr	Gln	Ala	Arg	Glu	Ser	Phe
				165					170					175	
Asp	Arg	Gly	Val	Arg	Cys	Ile	Leu	Lys	Cys	Gln	Ile	Val	Val	Asn	Gly
			180				185					190			
His	Lys	Thr	Ala	Trp	Cys	Ala	Gln	His	Asp	Glu	Thr	Asp	Phe	Ser	Pro
		195					200					205			
Arg	Ser	Ala	Arg	Ser	Tyr	Glu	Leu	Pro	Ser	Leu	Ser	Gly	Ser	Glu	Ser
		210				215					220				
Val	Gly	Ile	Val	Arg	Leu	Leu	Met	Ser	Leu	Asp	Gln	Pro	Ser	Arg	Gly
225					230					235					240
Val	Ile	Asp	Ala	Ile	Thr	Asn	Ala	Val	Ala	Trp	Phe	Glu	Ser	Ala	Lys
				245					250					255	
Leu	Pro	Gly	Ile	Lys	Thr	Val	Gln	Glu	Thr	Asp	Pro	Asn	Ser	Pro	Lys
			260				265						270		
Gly	Trp	Asn	Arg	Val	Val	Val	Lys	Asp	Glu	Ser	Ala	Arg	Pro	Met	Trp
		275					280					285			
Ala	Arg	Phe	Tyr	Asp	Ile	Asn	Thr	Asn	Lys	Pro	Phe	Phe	Cys	Asp	Arg
		290				295					300				
Asp	Gly	Val	Pro	Lys	Pro	Ser	Leu	Ala	Glu	Ile	Gly	Tyr	Glu	Arg	Arg
305					310					315					320
Asn	Gly	Tyr	Ala	Trp	Leu	Gly	Tyr	Trp	Pro	Glu	Asp	Leu	Leu	Ala	Arg
				325					330					335	
Glu	Tyr	Pro	Ala	Trp	Lys	Met	Lys	Trp	Leu	Lys	Pro	Lys	Glu	Arg	Pro
			340					345					350		
Ala	Phe														

<210> 15
 <211> 1575
 5 <212> ADN
 <213> Bacteria

<400> 15

atgagacgac	cagtcgcact	ccggctccac	gcggcactgg	ccaccctggc	cctggcggcc	60
gcgaccggcg	tggtgctctc	gatccccag	gcatcgggcg	cggccggcgg	cgccaccggc	120
tacgccggcc	agaacggcgg	caccaccggc	ggtgccggcg	gccagaccgt	acggggccacc	180
acgggcaccg	ccatccacgc	ggcctgtgc	ggacggggcca	gcagcagcac	cccgatcacg	240
atcgaggctg	agggaaacgat	caaccaacgc	aacaccggcca	aggtgtccgg	ccccagctgc	300
aacaccggcg	ccggagtgat	cgagctgaag	cagatcagca	acgtcacgct	cgtcgggggtc	360
ggctccggcg	ccgtcttcga	ccaactcggc	atccacatcc	gcgagtcag	caacatcatc	420
atccagaacg	tgacgggtccg	gaacgtcaag	aagtcgggct	cgccgctgtc	caacggcggc	480
gacgccatcg	gcatggagag	cgacgtccgc	aacgtctggg	tcgaccactc	caccctggag	540
gcctcggggc	gcgagtcoga	gggtacgac	ggcctcttcg	acatgaagga	caacaccggg	600
tacgtgaccc	tgctgtacag	catcctggcg	aaatccgggc	gcggcgccct	cgtaggggtcc	660
agcgagaccg	aactctcgaa	cagcttcac	acgtaccacc	acaacctgta	cgagaacatc	720
gactcgcgcg	cgccctgtct	gcggcgccgg	accgcccaca	tgtacaacaa	ccactacctg	780
cggatccagc	agtcgggcac	caactccgct	gcgggagccc	acgccaaggt	ggacaacaac	840
tacttcgagc	actccaagga	cgctctcggc	accttctaca	cggacggcgc	cggtgactgg	900
caggtcagcg	gcaacgtcta	cgacaacgtg	acctgggtccg	cccggggcac	cgacaacaac	960
ccggcggggc	cggaccgcga	gtccaacacc	accgtctcca	tcccctacgc	cttcagcctc	1020
gaccocggcca	cctgcgtgcc	ggacgtcgtg	agccgaacgg	cgggtgccgg	caagggactt	1080

```

caggtgtcga acggcagctg ctccccgcag acacccacgc ccacgccgac gggcacgccg 1140
accacacccg cgccgacgac tcccaccccg agcccgacgc cctccacgcc cggaccgacc 1200
cagccccggc ggacgaacct cagcatcggg gccgggtccg acggttcgag caaggccgac 1260
ggcaccagct acggcaacgt ccgggacggg gacctcggca cccactggtc tccggccggt 1320
tcgaccggct ccgtgtcgat caagtggggc agcgccacca cggctctccg catcgtcatc 1380
cgcgagggcg cgggcgcgac gggcggtcatc ggctcctggc tcgtcctgaa cggcgacacc 1440
ggcgccgtgc tgacctccgg cagcggggcg gggacgatct ccgtcccccg gacggccctg 1500
aagaagatca ccttcgagat cacgggcgcg agcggcacgc cacggatcgc cgagttcgag 1560
acgtacgccg gctag 1575

```

<210> 16
 <211> 524
 <212> PRT
 <213> Bacteria

<220>
 <221> SEÑAL
 <222> (1)...(33)
 <221> DOMINIO
 <222> (34)...(359)
 <223> Dominio catalítico

<400> 16

```

Met Arg Arg Pro Val Ala Leu Arg Leu His Ala Ala Leu Ala Thr Leu
 1          5          10          15
Ala Leu Ala Ala Ala Thr Gly Val Val Leu Ser Ile Pro Gln Ala Ser
          20          25          30
Ala Ala Ala Gly Gly Ala Thr Gly Tyr Ala Gly Gln Asn Gly Gly Thr
 35          40          45
Thr Gly Gly Ala Gly Gly Gln Thr Val Arg Ala Thr Thr Gly Thr Ala
 50          55          60
Ile His Ala Ala Leu Cys Gly Arg Ala Ser Ser Thr Pro Ile Thr
 65          70          75          80
Ile Glu Val Glu Gly Thr Ile Asn His Ala Asn Thr Ala Lys Val Ser
          85          90          95
Gly Pro Ser Cys Asn Thr Ala Ala Gly Val Ile Glu Leu Lys Gln Ile
          100          105          110
Ser Asn Val Thr Leu Val Gly Val Gly Ser Gly Ala Val Phe Asp Gln
          115          120          125
Leu Gly Ile His Ile Arg Glu Ser Ser Asn Ile Ile Ile Gln Asn Val
          130          135          140
Thr Val Arg Asn Val Lys Lys Ser Gly Ser Pro Leu Ser Asn Gly Gly
          145          150          155          160
Asp Ala Ile Gly Met Glu Ser Asp Val Arg Asn Val Trp Val Asp His
          165          170          175
Ser Thr Leu Glu Ala Ser Gly Gly Glu Ser Glu Gly Tyr Asp Gly Leu
          180          185          190
Phe Asp Met Lys Asp Asn Thr Arg Tyr Val Thr Leu Ser Tyr Ser Ile
          195          200          205
Leu Arg Lys Ser Gly Arg Gly Gly Leu Val Gly Ser Ser Glu Thr Glu
          210          215          220
Leu Ser Asn Ser Phe Ile Thr Tyr His His Asn Leu Tyr Glu Asn Ile
          225          230          235          240
Asp Ser Arg Ala Pro Leu Leu Arg Gly Gly Thr Ala His Met Tyr Asn
          245          250          255
Asn His Tyr Leu Arg Ile Asn Glu Ser Gly Ile Asn Ser Arg Ala Gly
          260          265          270

```

Ala His Ala Lys Val Asp Asn Asn Tyr Phe Glu Asp Ser Lys Asp Val
 275 280 285
 Leu Gly Thr Phe Tyr Thr Asp Ala Ala Gly Tyr Trp Gln Val Ser Gly
 290 295 300
 Asn Val Tyr Asp Asn Val Thr Trp Ser Ala Arg Gly Thr Asp Asn Asn
 305 310 315 320
 Pro Ala Gly Pro Asp Pro Gln Ser Asn Thr Thr Val Ser Ile Pro Tyr
 325 330 335
 Ala Phe Ser Leu Asp Pro Ala Thr Cys Val Pro Asp Val Val Ser Arg
 340 345 350
 Thr Ala Gly Ala Gly Lys Gly Leu Gln Val Ser Asn Gly Ser Cys Ser
 355 360 365
 Pro Gln Thr Pro Thr Pro Thr Pro Thr Gly Thr Pro Thr Thr Pro Ala
 370 375 380
 Pro Thr Thr Pro Thr Pro Ser Pro Thr Pro Ser Thr Pro Gly Pro Thr
 385 390 395 400
 Gln Pro Gly Gly Thr Asn Leu Ser Ile Gly Ala Gly Ser Asp Gly Ser
 405 410 415
 Ser Lys Ala Asp Gly Thr Ser Tyr Gly Asn Val Arg Asp Gly Asp Leu
 420 425 430
 Gly Thr His Trp Ser Pro Ala Gly Ser Thr Gly Ser Val Ser Ile Lys
 435 440 445
 Trp Gly Ser Ala Thr Thr Val Ser Arg Ile Val Ile Arg Glu Ala Ala
 450 455 460
 Gly Ala Thr Gly Val Ile Gly Ser Trp Leu Val Leu Asn Gly Asp Thr
 465 470 475 480
 Gly Ala Val Leu Thr Ser Gly Ser Gly Ala Gly Thr Ile Ser Val Pro
 485 490 495
 Arg Thr Ala Leu Lys Lys Ile Thr Phe Glu Ile Thr Gly Ala Ser Gly
 500 505 510
 Thr Pro Arg Ile Ala Glu Phe Glu Thr Tyr Ala Gly
 515 520

<210> 17
 <211> 1047
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 17
 ttgccgcgtg cgcccgggtg tgagtcgtca tcgccagcgc agacgtcatt ggttgccggtc 60
 tccctgggatc agatcctccg tcagcctgcg gcctgggtacg gcggtgcgga ggcgttgcca 120
 gtcgctgaga acgtgctctt gtatcagcgc gcggcaggag ggtggccgaa gaacatcaac 180
 atggcggcgc cgatgaccgc cgctgaccgt gcgaaagtca cggacgagcg cgcgcagaaac 240
 gacgccacga tcgacaacac gtcaacgacg acgcagatcc gttttcttgc gctcgttctt 300
 cgcggcaccg ccgacgcacg attcaaggac gcggcgctga agggcatcga cttcctgctg 360
 gctgcgcaat acgcgaatgg aggttggcct cagtattttc cctgcgcga cgactactcg 420
 cggcgcatca cgttcaatga cgacgcgatg gtgaatgtga tgacgtgct ggcgcgagact 480
 tcgcagggcc agacgcggtt cgagttcgtc gacgcctcgc ggcgcggccg ggcggcgag 540
 tctgtctcac gcggcgctga cgtcatgctg cgcacgcaga ttcgagtcaa cggcgtgctg 600
 accggctggt gccagcagca cgacgagcgg aactttcagc cgggtgaaggc gcgcgcgtac 660
 gaacatccgt cgattgccag caaggaaacc gcgagcatcg caagattcct gatggggatt 720
 gaacggccgt gcggcgagat cgtgtccgcg cgttgatggc cagtcgcgtg gttcgagcg 780
 gcgcagattt caggtgtgcg gacggagcgc cggcccgacg gatcgaatcc gggcggcgac 840
 gtcgtggcgg tgcaggactc cgcgcgcgcg ccaatctggg ccgcttcta cgagattggc 900
 accaaccggc cgatgttttc gggtcgcgac gggtcatca agtacagcct cagcgagatc 960
 gagatcgagc ggcgcgctgg atacagctgg tacggcgact acgccgccag actgctcaga 1020
 gacgactatc cgaagtggaa gaaatga 1047

<210> 18

<211> 348
<212> PRT
<213> Desconocido

5 <220>
<223> Obtenido de una muestra ambiental
<221> DOMINIO
<222> (1)...(348)
<223> Dominio catalítico

10 <400> 18

Met	Pro	Arg	Ala	Pro	Gly	Gly	Glu	Ser	Ser	Ser	Pro	Ala	Gln	Thr	Ser
1				5					10					15	
Ser	Val	Ala	Val	Ser	Trp	Asp	Gln	Ile	Leu	Arg	Gln	Pro	Ala	Ala	Trp
			20					25					30		
Tyr	Gly	Gly	Ala	Glu	Ala	Leu	Arg	Val	Ala	Glu	Asn	Val	Leu	Leu	Tyr
		35					40					45			
Gln	Arg	Ala	Ala	Gly	Gly	Trp	Pro	Lys	Asn	Ile	Asn	Met	Ala	Ala	Pro
		50				55					60				
Met	Thr	Ala	Ala	Asp	Arg	Ala	Lys	Val	Thr	Asp	Glu	Arg	Ala	Gln	Asn
65					70					75					80
Asp	Ala	Thr	Ile	Asp	Asn	Thr	Ser	Thr	Thr	Thr	Gln	Ile	Arg	Phe	Leu
			85					90						95	
Ala	Leu	Val	Leu	Arg	Gly	Thr	Ala	Asp	Ala	Arg	Phe	Lys	Asp	Ala	Ala
			100					105					110		
Leu	Lys	Gly	Ile	Asp	Phe	Leu	Leu	Ala	Ala	Gln	Tyr	Ala	Asn	Gly	Gly
		115					120					125			
Trp	Pro	Gln	Tyr	Phe	Pro	Leu	Arg	Asp	Asp	Tyr	Ser	Arg	Arg	Ile	Thr
		130				135					140				
Phe	Asn	Asp	Asp	Ala	Met	Val	Asn	Val	Met	Thr	Leu	Leu	Arg	Glu	Thr
145					150					155					160
Ser	Gln	Gly	Gln	Thr	Pro	Phe	Glu	Phe	Val	Asp	Ala	Ser	Arg	Arg	Gly
			165					170						175	
Arg	Ala	Ala	Gln	Ser	Val	Ser	Arg	Gly	Val	Asp	Val	Met	Leu	Arg	Thr
			180					185					190		
Gln	Ile	Arg	Val	Asn	Gly	Val	Leu	Thr	Gly	Trp	Cys	Gln	Gln	His	Asp
		195					200					205			
Glu	Arg	Asn	Phe	Gln	Pro	Val	Lys	Ala	Arg	Ala	Tyr	Glu	His	Pro	Ser
		210				215					220				
Ile	Ala	Ser	Lys	Glu	Thr	Ala	Ser	Ile	Ala	Arg	Phe	Leu	Met	Gly	Ile
225					230					235					240
Glu	Arg	Pro	Ser	Pro	Glu	Ile	Val	Ser	Ala	Val	Asp	Gly	Ala	Val	Ala
			245						250					255	
Trp	Leu	Arg	Ala	Ala	Gln	Ile	Ser	Gly	Val	Arg	Thr	Glu	Arg	Arg	Pro
			260					265					270		
Asp	Gly	Ser	Asn	Pro	Gly	Gly	Asp	Val	Val	Ala	Val	Gln	Asp	Ser	Ala
		275					280					285			
Ala	Pro	Pro	Ile	Trp	Ala	Arg	Phe	Tyr	Glu	Ile	Gly	Thr	Asn	Arg	Pro
		290				295					300				
Met	Phe	Ser	Gly	Arg	Asp	Gly	Val	Ile	Lys	Tyr	Ser	Leu	Ser	Glu	Ile
305					310					315					320
Glu	Ile	Glu	Arg	Arg	Ala	Gly	Tyr	Ser	Trp	Tyr	Gly	Asp	Tyr	Ala	Ala
			325						330					335	
Arg	Leu	Leu	Arg	Asp	Asp	Tyr	Pro	Lys	Trp	Lys	Lys				
			340					345							

<210> 19
<211> 1122
<212> ADN
<213> Desconocido

20 <220>
<223> Obtenido de una muestra ambiental

<400> 19

```

gtgaacaggt ggcgcgaaga cttettgccc gacttgcggg cccgcatgct ccggtgcatg      60
gttccccggc cgcagatcca ctggggcggc ggtgtcatcc ggcaggaacc ggaatggtac      120
ggctcggccg aggcgcgtgc gatcgccgac agcgtttctt aataccagtc gaccgctggc      180
ggctggccca agaacaccga cttgacggtc tcgccaccgt ccgcgaatt ccttgcggat      240
gcggatggtc tcacgaacac gatcgacaac gacgccacca cgttgcggat gcgatttctc      300
gctctggtgg cgcacgcgac cggcggcatc aagtacgcgc ccgcgttcga acgcggtctg      360
gactacctgc tcgcgcgtca gtatcccaat ggcggctggc ctcagtattt tcccctgcgt      420
gacggctatt actcgcacat cacctacaac gacaatgcga tgggtcaacgt cctcaccgtt      480
ctgcgcgatg cggcgcgggg ccggcccccct tactcgttcg tcgacagggc ccggcgcgcc      540
agagcagaaa cggccatcgc tcgcggcatc gacatcatcg tcgcactca ggtgagacgg      600
gccggcgtgc tgacgcgatg gtgcgcccag caccagcaaa agacgctcga gccggcgtgg      660
gcgcgcaact acgaaccgcc gacactctcc gggcagcaaa gcgtcggcat cgtgcgcttt      720
ctcatgggaa tcgaaaagcc cagcgcgagg atcgtcgcgg cgggtgcaagg cgccgctgac      780
tgggtgagag ccgtcgcgat cagcgggttg cgtctcgagg aattcaccga ccgcgatggc      840
aggcgcgaca ggcgcgtcgt cgcgatccg gcagcgcgc tcctgtgggc gcgttctac      900
gagcttggca cggaccgtcc cgtcttcacc ggccgcgaca aggtgatccg gtactcgtc      960
agcgaatcgc agcagagcg ccggaacggg tatgcctact atggcacatg gccggccacg      1020
ctcctcagcg aggagtaccc ccgttggcgc gcgaaacacc tggctcgacg gagcgtcagg      1080
caggtagagg agggaatcgc gatacgcgtc cctaaccct ga      1122

```

5

<210> 20
 <211> 373
 <212> PRT
 <213> Desconocido

10

<220>
 <223> Obtenido de una muestra ambiental
 <221> DOMINIO
 <222> (1)...(373)

15

<223> Dominio catalítico

<400> 20

```

Met Asn Arg Trp Arg Glu Asp Phe Leu Arg Asp Phe Ala Ala Arg Met
 1          5          10          15
Leu Arg Cys Met Val Pro Arg Pro Gln Ile His Trp Gly Gly Gly Val
      20      25      30
Ile Arg Gln Glu Pro Glu Trp Tyr Gly Ser Ala Glu Ala Arg Ala Ile
      35      40      45
Ala Asp Ser Val Leu Gln Tyr Gln Ser Thr Ala Gly Gly Trp Pro Lys
      50      55      60
Asn Thr Asp Leu Thr Val Ser Pro Pro Ser Ala Glu Phe Leu Ala Asp
      65      70      75      80
Ala Asp Gly Leu Thr Asn Thr Ile Asp Asn Asp Ala Thr Thr Leu Pro
      85      90      95
Met Arg Phe Leu Ala Leu Val Ala His Ala Thr Gly Gly Ile Lys Tyr
      100      105      110
Arg Ala Ala Phe Glu Arg Gly Leu Asp Tyr Leu Leu Ala Ala Gln Tyr
      115      120      125

```

Pro	Asn	Gly	Gly	Trp	Pro	Gln	Tyr	Phe	Pro	Leu	Arg	Asp	Gly	Tyr	Tyr
130					135					140					
Ser	His	Ile	Thr	Tyr	Asn	Asp	Asn	Ala	Met	Val	Asn	Val	Leu	Thr	Val
145					150					155					160
Leu	Arg	Asp	Ala	Ala	Ala	Gly	Arg	Pro	Pro	Tyr	Ser	Phe	Val	Asp	Arg
				165					170					175	
Ala	Arg	Arg	Ala	Arg	Ala	Glu	Thr	Ala	Ile	Ala	Arg	Gly	Ile	Asp	Ile
			180					185					190		
Ile	Val	Arg	Thr	Gln	Val	Arg	Arg	Ala	Gly	Val	Leu	Thr	Ala	Trp	Cys
	195					200					205				
Ala	Gln	His	Asp	Glu	Lys	Thr	Leu	Glu	Pro	Ala	Trp	Ala	Arg	Asn	Tyr
210					215					220					
Glu	Pro	Pro	Thr	Leu	Ser	Gly	His	Glu	Ser	Val	Gly	Ile	Val	Arg	Phe
225				230						235					240
Leu	Met	Gly	Ile	Glu	Lys	Pro	Thr	Pro	Arg	Ile	Val	Ala	Ala	Val	Gln
				245					250					255	
Gly	Ala	Ala	Asp	Trp	Leu	Arg	Ala	Val	Ala	Ile	Ser	Gly	Leu	Arg	Leu
			260					265					270		
Glu	Glu	Phe	Thr	Asp	Ala	Asp	Gly	Arg	Arg	Asp	Arg	Arg	Val	Val	Ala
	275						280				285				
Asp	Pro	Ala	Ala	Pro	Leu	Leu	Trp	Ala	Arg	Phe	Tyr	Glu	Leu	Gly	Thr
290					295					300					
Asp	Arg	Pro	Val	Phe	Thr	Gly	Arg	Asp	Lys	Val	Ile	Arg	Tyr	Ser	Leu
305				310						315					320
Ser	Glu	Ile	Glu	His	Glu	Arg	Arg	Asn	Gly	Tyr	Ala	Tyr	Tyr	Gly	Thr
				325					330					335	
Trp	Pro	Ala	Thr	Leu	Leu	Ser	Glu	Glu	Tyr	Pro	Arg	Trp	Arg	Ala	Lys
			340				345						350		
His	Leu	Ala	Arg	Arg	Ser	Val	Arg	Gln	Val	Glu	Glu	Gly	Ile	Ala	Ile
	355						360					365			
Arg	Val	Pro	Asn	Pro											
370															

<210> 21

<211> 1269

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 21

atgcgtaaat	cgaactgggc	cgtcacaacg	gccatcctgc	tgcgcgtgag	cgccgcacccg	60
ctggcgccaa	agcccatcgg	acagatcacc	ctcgccgtgc	cgctcagccc	ggcgcgccctg	120
accgaaacgc	cgccctgagca	gggggcgcaa	tggcaggcct	atctcgccac	caccgaggca	180
cagcttaagg	cagacaaggc	ggcgctggct	gccgagcgcg	ccggtctggc	cgaaatcccc	240
gccaagccga	agaccggcag	cgccaacacc	atgcgcgtcg	acaagccgct	ggaatggtac	300
gcgtcgctccg	aggcgcgctt	ggtcgccgat	aatatcgta	gctatcagac	tccggcgaggc	360
ggctggggca	aaaatcaggc	ccgcaacgaa	cccacgcggt	tgaaagggtca	ggcctacact	420
atcgatgacg	ccgatcccac	cggttcgggc	aaatggaact	tcgtcggcac	catcgacaac	480
gacgccacca	tcgtggaaat	tcgctttctc	gcccgcgtag	cgccggcgccg	cacggggcccg	540
gaaggcgacg	tctatcgcg	ctccgccacg	cgccggcatca	cctacttgct	ggcgggcgacg	600
taccccaatg	gcggctggcc	gcaggtctgg	ccgcttcagg	gcggctatca	cgacgccatc	660
accctcaatg	acggcgcgat	gatccatgtg	ctcgaactgt	ttgacgacat	cgccagcgga	720
caggggcgact	tcgccttctc	gcctgagccg	ctgcgcgaca	aggctcaggc	cgcacaggca	780
aagggtcaga	aggtgcttct	cgatcttcag	cttaagcgca	acggcgaaacg	caccctgtgg	840
gcgcagcagt	acgatccgat	taccctcttg	cccagcgcg	cgcgtaacta	cgagccgctcg	900
tcgatcagca	ccggtgaaag	cgccggtgtg	ctgatctacc	tcatgtccct	gcccaccccc	960
tcgcctgaag	tgccgcgacg	catcgaaaaa	ggcgtggccc	tgctgatcaa	acttcagatc	1020

```

aacggcatgg catgggaaaa ggacggcatg cgcaaacgtc tggtcgccaa ggctgacgcc 1080
tcgcgcgtgt ggtcgcgcta tcacgactcg gaaacgctgc tgcccatott cggtgaccgc 1140
gacatgcgca tcttcgacga cgtcaacgac atcagcgacg aacgcagccg cggctatgcc 1200
tggtatggca caagcccggc acggggccatc gccgaatacg aaaaatggaa acagggcaac 1260
ggcaaatga 1269

```

<210> 22
 <211> 422
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 10 <221> SEÑAL
 <222> (1)...(23)
 <221> DOMINIO
 <222> (24)...(422)
 <223> Dominio catalítico

15 <400> 22

```

Met Arg Lys Ser Asn Trp Ala Val Thr Thr Ala Ile Leu Leu Ala Leu
 1      5      10      15
Ser Ala Ala Pro Leu Ala Ala Lys Pro Ile Gly Gln Ile Thr Leu Ala
 20      25      30
Val Pro Leu Ser Pro Ala Arg Leu Thr Glu Thr Pro Pro Glu Gln Arg
 35      40      45
Ala Gln Trp Gln Ala Tyr Leu Ala Thr Thr Glu Ala Gln Leu Lys Ala
 50      55      60
Asp Lys Ala Ala Leu Ala Ala Glu Arg Ala Gly Leu Ala Glu Ile Pro
 65      70      75      80
Ala Lys Pro Lys Thr Gly Ser Ala Asn Thr Met Pro Leu Asp Lys Pro
 85      90      95
Leu Glu Trp Tyr Ala Ser Ser Glu Ala Arg Leu Val Ala Asp Asn Ile
 100     105     110
Val Ser Tyr Gln Thr Pro Ala Gly Gly Trp Gly Lys Asn Gln Ala Arg
 115     120     125
Asn Glu Pro Thr Arg Leu Lys Gly Gln Ala Tyr Thr Ile Asp Asp Ala
 130     135     140
Asp Pro Thr Gly Ser Gly Lys Trp Asn Phe Val Gly Thr Ile Asp Asn
 145     150     155     160
Asp Ala Thr Ile Val Glu Ile Arg Phe Leu Ala Arg Val Ala Ala Ala
 165     170     175
Ala Thr Gly Pro Glu Gly Asp Val Tyr Arg Ala Ser Ala Thr Arg Gly
 180     185     190
Ile Thr Tyr Leu Leu Ala Ala Gln Tyr Pro Asn Gly Gly Trp Pro Gln
 195     200     205
Val Trp Pro Leu Gln Gly Gly Tyr His Asp Ala Ile Thr Leu Asn Asp
 210     215     220
Gly Ala Met Ile His Val Leu Glu Leu Phe Asp Asp Ile Ala Ser Gly
 225     230     235     240
Gln Gly Asp Phe Ala Phe Leu Pro Glu Pro Leu Arg Asp Lys Val Glu
 245     250     255
Ala Ala Gln Ala Lys Gly Gln Lys Val Leu Leu Asp Leu Gln Leu Lys
 260     265     270
Arg Asn Gly Glu Arg Thr Leu Trp Ala Gln Gln Tyr Asp Pro Ile Thr
 275     280     285
Leu Leu Pro Ser Ala Ala Arg Asn Tyr Glu Pro Ser Ser Ile Ser Thr

```



```

      290      295      300
Gly Glu Ser Ala Gly Val Leu Ile Tyr Leu Met Ser Leu Pro Asn Pro
305      310      315      320
Ser Pro Glu Val Arg Asp Ala Ile Glu Lys Gly Val Ala Leu Leu Ile
      325      330      335
Lys Leu Gln Ile Asn Gly Met Ala Trp Glu Lys Asp Gly Met Arg Lys
      340      345      350
Arg Leu Val Ala Lys Ala Asp Ala Ser Pro Leu Trp Ser Arg Tyr His
      355      360      365
Asp Ser Glu Thr Leu Leu Pro Ile Phe Gly Asp Arg Asp Met Arg Ile
      370      375      380
Phe Asp Asp Val Asn Asp Ile Ser Asp Glu Arg Ser Arg Gly Tyr Ala
385      390      395      400
Trp Tyr Gly Thr Ser Pro Ala Arg Ala Ile Ala Glu Tyr Glu Lys Trp
      405      410      415
Lys Gln Gly Asn Gly Lys
      420

```

<210> 23
 <211> 1182
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 23

```

atgaaccgtg gcgtgattgt tttgctggcg gccgctccag ctgcggcgca tggcgcagtg      60
ctggggtata tgacgcctgc gcagccgttg accgaggcgc gcattgccgc gctgccggcg      120
tcggagcagg gcgcctggcg gggctacctc gcccgcctcc gcgcagccat ggacgccgac      180
aaggccgccc tggccgccga gcgcgcgcgc ctcgccaccg taccgccggc gccgccgcat      240
ggcgttggtg atggcgggat ggcgcgcaac cgtccgacgg cttggtatgg gacgccggaa      300
gcgcggcaca tcgcggacaa tatcgtcagc ttccagacgc cgtccggcgg ctggggcaag      360
aacgtggacc gcacgggacc tgtgcgccag cgcggacagc attacgtttc cttcgatggc      420
aaggagtcct ggaacttcac cggcacgata gacaacaacg ccacaacgag cgagctgaaa      480
ttcctggcgc gcgtgcaggc gcaaatgccc ggccggcgcg gcgacgaata ccggaaggcc      540
gccctgcgcg gcacagcta cctgttgaac tcacaatatc ccaacggcgg cttcccgcag      600
gtctatccgc tgcaaggcgg ctaccacgac gccatcacct tcaacgacga tgccttcgcc      660
aacgtgctgc aagtgtgtgt ggaagtggcg aaccgcaggg gcgactatgc cttcgtcccc      720
gaaaccgtgg caaccgatgc ccgcgcggcc gcggacaagg cgtccaagt cctgctggcg      780
agccagatca tcgtcggcgg cgtacgcacc gcctggtgcc agcagcacga tgcgatcacg      840
ctggcgcccc tcggcgcccc caatttcgaa ccggccgcgc tgaccagcac ggaaagcgcg      900
cgctgtctga tgetgttgat gctgctgccc gatccgagcc cggagctgag agcgtcaatc      960
catgcgggga tggcctggct gcagaaagcg gcgctgccgg gggatgtctg gtcgcgctac     1020
tatgacctga acacgatgag gccgatcttt ggggatcgtg accgcagtat ccacgatgat     1080
gtgaaggaat tgagcgagga gaggcaaaaa ggctatggct ggttcagtaa cggaccagcc     1140
agagctaaac aggcctttga ggccctggacg cgcaaacctt ga                        1182

```

<210> 24
 <211> 393
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(18)
 <221> DOMINIO
 <222> (19)...(393)
 <223> Dominio catalítico

<400> 24

```

Met Asn Arg Gly Val Ile Val Leu Leu Ala Ala Ala Pro Ala Ala Ala
1      5      10      15
His Gly Ala Val Leu Gly Tyr Met Thr Pro Ala Gln Pro Leu Thr Glu
20      25      30
Ala Arg Ile Ala Ala Leu Pro Ala Ser Glu Gln Gly Ala Trp Arg Gly
35      40      45
Tyr Leu Ala Arg Ser Arg Ala Ala Met Asp Ala Asp Lys Ala Ala Leu
50      55      60
Ala Ala Glu Arg Ala Ala Leu Ala Thr Val Pro Pro Ala Pro Pro His
65      70      75      80
Gly Gly Gly Asp Gly Gly Met Ala Arg Asn Arg Pro Thr Ala Trp Tyr
85      90      95
Gly Thr Pro Glu Ala Arg His Ile Ala Asp Asn Ile Val Ser Phe Gln
100     105
Thr Pro Ser Gly Gly Trp Gly Lys Asn Val Asp Arg Thr Gly Pro Val
115     120     125
Arg Gln Arg Gly Gln His Tyr Val Ser Phe Asp Gly Lys Glu Ser Trp
130     135     140
Asn Phe Ile Gly Thr Ile Asp Asn Asn Ala Thr Thr Ser Glu Leu Lys
145     150     155     160
Phe Leu Ala Arg Val Gln Ala Gln Met Pro Gly Ala Ala Gly Asp Glu
165     170     175
Tyr Arg Lys Ala Ala Leu Arg Gly Ile Ser Tyr Leu Leu Asn Ser Gln
180     185     190
Tyr Pro Asn Gly Gly Phe Pro Gln Val Tyr Pro Leu Gln Gly Gly Tyr
195     200     205
His Asp Ala Ile Thr Phe Asn Asp Asp Ala Phe Ala Asn Val Leu Gln
210     215     220
Val Leu Leu Glu Val Ala Asn Arg Arg Gly Asp Tyr Ala Phe Val Pro
225     230     235     240
Glu Thr Val Ala Thr Asp Ala Arg Ala Ala Ala Asp Lys Ala Leu Gln
245     250     255
Val Leu Leu Ala Ser Gln Ile Ile Val Gly Gly Val Arg Thr Ala Trp
260     265     270
Cys Gln Gln His Asp Ala Ile Thr Leu Ala Pro Val Gly Ala Arg Asn
275     280     285
Phe Glu Pro Ala Ala Leu Thr Ser Thr Glu Ser Ala Arg Leu Leu Met
290     295     300
Leu Leu Met Leu Leu Pro Asp Pro Ser Pro Glu Leu Arg Ala Ser Ile
305     310     315     320
His Ala Gly Met Ala Trp Leu Gln Lys Ala Ala Leu Pro Gly Asp Val
325     330     335
Trp Ser Arg Tyr Tyr Asp Leu Asn Thr Met Arg Pro Ile Phe Gly Asp
340     345     350
Arg Asp Arg Ser Ile His Asp Asp Val Lys Glu Leu Ser Glu Glu Arg
355     360     365
Gln Lys Gly Tyr Gly Trp Phe Ser Asn Gly Pro Ala Arg Ala Lys Gln
370     375     380
Ala Phe Glu Ala Trp Thr Arg Lys Pro
385     390

```

<210> 25

5 <211> 1194

<212> ADN

<213> Desconocido

<220>

10 <223> Obtenido de una muestra ambiental

<400> 25

```

ttggtcgctg ccctattaag ctgcggcagc gccaatctct atgcagaatc aaccgcaaaa 60
tcggttacgc aatcagcagc cacaaatcaa ttgcaaaatg aaaaaagcag ttgggacagc 120
tattacgcgc catccaaaaa aatacatcag gcagaccagg attttctcgc cgctgaatta 180
aaaaaactcg gtcagaaaaa accaacattg cccgcacaca ccaaagattt tggttttgat 240
gttaagcagg taaatgcaga ttggttttaa agtgacgaag gcaaacgtgt gatggagatt 300
attctctcct tccaaacccc gtccggcggg tgggtcaaag gtaccgacat ggccaaggcg 360
gtgcgacaac ctgggcaagc ctttggcggt gaaaaagggt atatcccaac atttgataat 420
ggcgctacca gcaactcaatt gatgttgctc gcgcaagcac accaagccac cggcgatcac 480
cgcttttagc acgcattttg gcgcggcttg caattaattt tgactgcgca ataccogaat 540
ggtggctggc cacaaaactt tccactaacc ggtagctacc acgattacat cacctacaac 600
gacaatctta cgcgcgacct gatggttagt ctgcacaaaa cagcgcaggc aaaaaatgat 660
tttgcatctg tgaccaaagc gcagcaaata gcagcgctcag ctaggcctgc gcgtgcactt 720
gattgcgtat tgaaatcaca agttgtcgtc aatggcacac gcacactctg gggcgcacag 780
cacgatgta aaacactgca accaaccaaa gcgcgcgcgt ttgaaatggt gtcactcact 840
accactgaaa gcgcagccat gctcagtttt ctgatggata tcaaaaatcc cagcgcggat 900
attattcaat ccatacatgc agccatagcc tggatatgagc aaaataaaat cgtcggaaaa 960
acctggacac gtggtgatgc ggaattaaaa gataataaaa attcgcagcc actctgggcg 1020
cgtttttatg agataggcac taataagcct atatttgggg atcgcgatga cactgtgtat 1080
tacgatttgg caaaagtgtc taaagagcgt cgcgaagggt atgcgtggtg ctccactgac 1140
ccgaataaga cgctaaaaaa atatgctgaa tgggtctaaaa aatatcccaa ataa 1194

```

<210> 26
 <211> 397
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 10 <221> SEÑAL
 <222> (1)...(15)
 <221> DOMINIO
 <222> (16)...(397)
 <223> Dominio catalítico

15 <400> 26

```

Met Val Ala Ala Leu Leu Ser Cys Gly Ser Ala Asn Leu Tyr Ala Glu
 1          5          10          15
Ser Thr Ala Lys Ser Val Thr Gln Ser Ala Ala Thr Asn Gln Leu Gln
          20          25          30
Asn Glu Lys Ser Ser Trp Asp Ser Tyr Tyr Ala Ala Ser Lys Lys Ile
          35          40          45
His Gln Ala Asp Gln Asp Phe Leu Ala Ala Glu Leu Lys Lys Leu Gly
 50          55          60
Gln Lys Lys Pro Thr Leu Pro Ala His Thr Lys Asp Phe Gly Phe Asp
65          70          75          80
Val Lys Gln Val Asn Ala Asp Trp Phe Lys Ser Asp Glu Gly Lys Arg
          85          90          95
Val Met Glu Ile Ile Leu Ser Phe Gln Thr Pro Ser Gly Gly Trp Ser
100          105          110
Lys Arg Thr Asp Met Ala Lys Ala Val Arg Gln Pro Gly Gln Ala Phe
115          120          125

```

Gly	Val	Glu	Lys	Gly	Tyr	Ile	Pro	Thr	Phe	Asp	Asn	Gly	Ala	Thr	Ser
130						135					140				
Thr	Gln	Leu	Met	Leu	Leu	Ala	Gln	Ala	His	Gln	Ala	Thr	Gly	Asp	His
145						150				155					160
Arg	Phe	Ser	Asp	Ala	Phe	Gly	Arg	Gly	Leu	Gln	Leu	Ile	Leu	Thr	Ala
				165					170					175	
Gln	Tyr	Pro	Asn	Gly	Gly	Trp	Pro	Gln	Asn	Phe	Pro	Leu	Thr	Gly	Ser
			180					185					190		
Tyr	His	Asp	Tyr	Ile	Thr	Tyr	Asn	Asp	Asn	Leu	Thr	Arg	Asp	Leu	Met
		195					200					205			
Val	Val	Leu	His	Lys	Thr	Ala	Gln	Ala	Lys	Asn	Asp	Phe	Ala	Phe	Val
210						215					220				
Thr	Lys	Ala	Gln	Gln	Ile	Ala	Ala	Ser	Ala	Ser	Leu	Ala	Arg	Ala	Leu
225						230					235				240
Asp	Cys	Val	Leu	Lys	Ser	Gln	Val	Val	Val	Asn	Gly	Thr	Arg	Thr	Leu
				245					250					255	
Trp	Gly	Ala	Gln	His	Asp	Val	Lys	Thr	Leu	Gln	Pro	Thr	Lys	Ala	Arg
			260					265					270		
Ala	Phe	Glu	Met	Val	Ser	Leu	Thr	Thr	Thr	Glu	Ser	Ala	Ala	Met	Leu
		275					280					285			
Ser	Phe	Leu	Met	Asp	Ile	Lys	Asn	Pro	Ser	Ala	Asp	Ile	Ile	Gln	Ser
290						295					300				
Ile	His	Ala	Ala	Ile	Ala	Trp	Tyr	Glu	Gln	Asn	Lys	Ile	Val	Gly	Lys
305						310				315					320
Thr	Trp	Thr	Arg	Gly	Asp	Ala	Glu	Leu	Lys	Asp	Asn	Lys	Asn	Ser	Gln
				325					330					335	
Pro	Leu	Trp	Ala	Arg	Phe	Tyr	Glu	Ile	Gly	Thr	Asn	Lys	Pro	Ile	Phe
			340				345						350		
Gly	Asp	Arg	Asp	Asp	Thr	Val	Tyr	Tyr	Asp	Leu	Ala	Lys	Val	Ser	Lys
		355					360					365			
Glu	Arg	Arg	Glu	Gly	Tyr	Ala	Trp	Tyr	Ser	Thr	Asp	Pro	Asn	Lys	Thr
370						375					380				
Leu	Lys	Lys	Tyr	Ala	Glu	Trp	Ser	Lys	Lys	Tyr	Pro	Lys			
385					390					395					

<210> 27
 <211> 1917
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 27

gtgtctctct	ttagaaaact	cgcactgccg	gttctgtgcg	gtctactgct	ttctgtcgga	60
gcagaaaccc	gagcgtcgaa	gcgcattgtc	gtggccgctg	atggatcggg	tgacgtcagg	120
acgattcaac	aagcgggtgga	ccaggttccc	aaagacaata	cacacccggg	cttgattcag	180
atcaagccgg	gtgtgtatca	ggaacaagtg	cgtgtcgccg	ccggcaaacg	ctttatcact	240
cttcgcgggc	acgacggag	caagaccgtc	atcacctatc	gattgagcgc	actacaagcg	300
ggaaataccc	ggttggcatt	caccacetta	attaatgcag	acgactttcg	cgccgagaac	360
ctgacgtttg	aaaactcctt	cggcaccggg	tcacaagcgg	ttgctttgtt	tgctgatgcg	420
aaccgcgcga	cgtttgaaaa	ctgccgggtc	ctcgggtggc	aggacacttt	gtttgtgaac	480
ggcagccgcc	acttcttcaa	agactgctac	gtcgaaggcc	atgtcgattt	cattttcggc	540
acggcctccg	cagtgtttga	gaactgcacc	attcacagca	aaggcgaagg	ttatgtgacc	600
gcgcactatc	gcaccagcga	tgagatggat	accggttttg	tctttcatcg	ttgtcgtttg	660
accggacgag	acacggggccg	cggagtttat	ctcgggaaggc	cgtggcgacc	ttacgcgcgc	720
gtcgtcttta	tcgattgctg	gctggacgca	cacatcacag	ctgaaggctg	ggataattgg	780
agagatcctg	aacgagagaa	gaccgcgtgg	tttgccgagt	acaagtcaaa	agggcccggt	840
gctaataccg	tagctcgtgt	cgcgtgggtc	aggcagttga	cgacagaaca	agccgcggag	900

```

ttttcgcggg aacgcttttt cagcgcgcgt gttcgcgggc tctctgggca ggccaaccag 960
gcagtcggaa cgatcgcggt ggacgatgcg cagaaaaaac cgaacgagtg gtatgcgagc 1020
gccgaggggt tgcgcacgct cgacaacgtt gttctttatc aacgtgactc cggcggtgtg 1080
cccaagaaca tcgacatggg gaagccgctc gacgacaagg gtcgagccgg tcttctgcgc 1140
gtgctgaaga agaacgattc caccatcgat aacggcgcgga cttacacgca actctcgttt 1200
ctagcgcggg tttacacggc gcaaaagcag gagcggcgac gcgagtcggt tctgaaggga 1260
ctcgattacc tgttgaaggc gcagtatcca aacggagggt ggccgcagtt ctatcccaat 1320
ctcaacgggt attacaaaca catcactttc aacgacaacg cgatgatcgg cgtgatgaaa 1380
ctgctgcggc acgtagcggc agcgaaaccg gcgtatgcgt ttgtcgacga agcacgacga 1440
acgagtgagg cgaaggcggg cgaaaaagga atcgagtgcg tactgaagac gcaggtgggt 1500
gtgaatggcc ggcgacgggt gtggtgtgct caacatgacg aagtcacgct cgcgcctgcc 1560
cggcgaggga cgtttgaatt agtttcgctg agtgggtgtg aaagcggttg gatcggtgcg 1620
tttttgatgt cgatcaagaa cccgtcgccg gcggtgtgct aggcgatcga gtcggcggtt 1680
gcgtggttcg agcaatcgca agtgaaagat cccgcgggca aacctgcgtg ggcgcgattt 1740
tatgagatcg gcaactaatc tccgatcttc gccgggcgtg acggcgctgt taagtatgat 1800
gtgaaacaga tcgatgagga acgacgaaag aattacgcat ggtacgttga cgcgcgagcg 1860
aaactactga agaccgacta tctgagtggt aaagaaaaga acgccaagaa tcaatga 1917

```

<210> 28
 <211> 638
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(21)
 <221> DOMINIO
 <222> (28)...(308)
 <223> Dominio de pectin metil esterasa
 <221> DOMINIO
 <222> (309)...(638)
 <223> Dominio catalítico

<400> 28

```

Met Ser Leu Phe Arg Lys Leu Ala Leu Pro Val Leu Cys Gly Leu Leu
1          5          10          15
Leu Ser Val Gly Ala Glu Thr Arg Ala Ser Lys Arg Ile Val Val Ala
20          25          30
Ala Asp Gly Ser Gly Asp Val Arg Thr Ile Gln Gln Ala Val Asp Gln
35          40          45
Val Pro Lys Asp Asn Thr His Pro Val Leu Ile Gln Ile Lys Pro Gly
50          55          60
Val Tyr Gln Glu Gln Val Arg Val Ala Ala Gly Lys Arg Phe Ile Thr
65          70          75          80
Leu Arg Gly Asp Asp Ala Ser Lys Thr Val Ile Thr Tyr Arg Leu Ser
85          90          95
Ala Leu Gln Ala Gly Asn Thr Arg Leu Ala Phe Thr Thr Leu Ile Asn
100         105         110
Ala Asp Asp Phe Arg Ala Glu Asn Leu Thr Phe Glu Asn Ser Phe Gly
115         120         125
Thr Gly Ser Gln Ala Val Ala Leu Phe Val Asp Ala Asn Arg Ala Thr
130         135         140
Phe Glu Asn Cys Arg Phe Leu Gly Trp Gln Asp Thr Leu Phe Val Asn
145         150         155         160
Gly Ser Arg His Phe Phe Lys Asp Cys Tyr Val Glu Gly His Val Asp

```

				165					170					175	
Phe	Ile	Phe	Gly	Thr	Ala	Ser	Ala	Val	Phe	Glu	Asn	Cys	Thr	Ile	His
			180					185					190		
Ser	Lys	Gly	Glu	Gly	Tyr	Val	Thr	Ala	His	Tyr	Arg	Thr	Ser	Asp	Glu
		195					200					205			
Met	Asp	Thr	Gly	Phe	Val	Phe	His	Arg	Cys	Arg	Leu	Thr	Gly	Arg	Asp
	210					215					220				
Thr	Gly	Arg	Gly	Val	Tyr	Leu	Gly	Arg	Pro	Trp	Arg	Pro	Tyr	Ala	Arg
225					230					235					240
Val	Val	Phe	Ile	Asp	Cys	Trp	Leu	Asp	Ala	His	Ile	Arg	Pro	Glu	Gly
			245						250					255	
Trp	Asp	Asn	Trp	Arg	Asp	Pro	Glu	Arg	Glu	Lys	Thr	Ala	Trp	Phe	Ala
		260					265						270		
Glu	Tyr	Lys	Ser	Lys	Gly	Pro	Gly	Ala	Asn	Pro	Val	Ala	Arg	Val	Ala
	275						280					285			
Trp	Ser	Arg	Gln	Leu	Thr	Thr	Glu	Gln	Ala	Ala	Glu	Phe	Ser	Arg	Glu
	290					295					300				
Arg	Phe	Phe	Ser	Arg	Ala	Val	Arg	Gly	Leu	Ser	Gly	Gln	Ala	Asn	Gln
305					310					315					320
Ala	Val	Gly	Thr	Ile	Ala	Trp	Asp	Asp	Ala	Gln	Lys	Lys	Pro	Asn	Glu
			325						330					335	
Trp	Tyr	Ala	Ser	Ala	Glu	Ala	Leu	Arg	Ile	Ala	Asp	Asn	Val	Val	Leu
		340					345						350		
Tyr	Gln	Arg	Asp	Ser	Gly	Gly	Trp	Pro	Lys	Asn	Ile	Asp	Met	Gly	Lys
	355						360					365			
Pro	Leu	Asp	Asp	Lys	Gly	Arg	Ala	Gly	Leu	Leu	Arg	Val	Arg	Lys	Lys
	370					375					380				
Asn	Asp	Ser	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Tyr	Thr	Gln	Leu	Ser	Phe
385					390					395					400
Leu	Ala	Arg	Val	Tyr	Thr	Ala	Gln	Lys	Gln	Glu	Arg	His	Arg	Glu	Ser
			405						410					415	
Phe	Leu	Lys	Gly	Leu	Asp	Tyr	Leu	Leu	Lys	Ala	Gln	Tyr	Pro	Asn	Gly
			420					425					430		
Gly	Trp	Pro	Gln	Phe	Tyr	Pro	Asn	Leu	Asn	Gly	Tyr	Tyr	Lys	His	Ile
	435						440					445			
Thr	Phe	Asn	Asp	Asn	Ala	Met	Ile	Gly	Val	Met	Lys	Leu	Leu	Arg	Asp
	450					455					460				
Val	Ala	Ala	Ala	Lys	Pro	Ala	Tyr	Ala	Phe	Val	Asp	Glu	Ala	Arg	Arg
465					470					475					480
Thr	Ser	Ala	Ala	Lys	Ala	Val	Glu	Lys	Gly	Ile	Glu	Cys	Ile	Leu	Lys
			485						490					495	
Thr	Gln	Val	Val	Val	Asn	Gly	Arg	Arg	Thr	Val	Trp	Cys	Ala	Gln	His
		500						505					510		
Asp	Glu	Val	Thr	Leu	Ala	Pro	Ala	Pro	Ala						

<210> 29
<211> 1398

<212> ADN
<213> Desconocido

<220>

5 <223> Obtenido de una muestra ambiental

<400> 29

```

atgattaacg gtcgagattt cataaaagac ctcatcatca cctccgcggg agtcgcgggt 60
ctcccgcgaac tggcggttcgg acaaaacgat ccctggaaaa ctcaataccc gcagatcctc 120
gcgcgggatca aaccgcgcgaa atttcggaag cgcgatttcg tcatcacgaa gttcggcgcg 180
aaggcgggaa ccgatagcac gcaagcgatc gctaaagccc tcgacgcttg cgcgaaagcc 240
ggcgggcgga ggcgtcgtct acccgccggc gaatttctca ccggtgcgat ccatctcaag 300
tcgaacacca atctctacgt ctcaaaaggc gcgactctga agttttcgac cgaccccgaa 360
aaatctctgc cgatcggttc cacgcgggtg gaagggatgg agttgatgca tctctcgccg 420
ttcatctacg cgtacgagca gacgaacatc gcgatcaccg gcgagggcac gtcgacggc 480
caaggcaaat cgttcttttg gaagtggcac ggcaaccgcg gatacggcgg caaccccgaa 540
gtgatcagtc agcaaaaagc gcggggcgga ctttacgaga tgatggacaa gaacgtaccc 600
gtcgcggagc ggcgtgttcgg tatcgggcac tatctccggc cgcagttcat ccagccgtac 660
aaatgtaaga acgtcttgat cgaaggcggt acgatcatcg actcgccgat gtgggaagtt 720
catccggtgc tttgcgagaa tgtcacggtc cgcgaatctc acatctcgtc gcacgggtccg 780
aacaacgacg gctgcgatcc cgagtcgtgc aaagacgtcc tgatcgacaa ctgcttcttc 840
gacacccgtg acgactgcat cgcgatcaag tcgggtcgca ataacgacg tcgtcgtctg 900
aacacaccga ccgagacat catcgtccgc aactgcacga tgaaagacg tcacgggtgt 960
atcacggtcg gaagcgagat ctccggcggc gtgcgaaact tgttcgcaca cgattgcaag 1020
atggacagtg cggatctgtg gaccgcgctc cgggtaaaga acaacgcata gcggggcggc 1080
atgctggaga atttctatct ccgcaacatc aecgtcgggc aagtcgcgcg tgctgtggtc 1140
gagatcgatt tcaactatga agaaggcgcg aagggatcgt acacaccggt catgcgcaac 1200
tacgtggtcg aggatctgac gtgcaccagc gggaaccggc ccgtcgatct gcaaggatta 1260
gacaacgcgc caatttacga tgtgtcgtg cgtaacacga ccttcggcgc gatgaagaac 1320
aagagcgtcg tgaagaatgt ccgaggactg aagatcgaaa acgttacgt cagcggcacg 1380
cgcgtggaga gtttatga
1398

```

10 <210> 30
<211> 465
<212> PRT
<213> Desconocido

15 <220>
<223> Obtained from an environmental sampl
<221> SEÑAL
<222> (1)...(27)
<221> DOMINIO
20 <222> (77)...(459)
<223> Dominio catalítico

<400> 30

```

Met Ile Asn Arg Arg Asp Phe Ile Lys Asp Leu Ile Ile Thr Ser Ala
 1           5           10           15
Gly Val Ala Val Leu Pro Gln Leu Ala Phe Gly Gln Asn Asp Pro Trp
          20           25           30
Lys Thr Gln Tyr Pro Gln Ile Leu Ala Arg Ile Lys Pro Pro Lys Phe
          35           40           45
Pro Lys Arg Asp Phe Val Ile Thr Lys Phe Gly Ala Lys Ala Gly Thr
          50           55           60

```

25

Asp	Ser	Thr	Gln	Ala	Ile	Ala	Lys	Ala	Leu	Asp	Ala	Cys	Ala	Lys	Ala
65					70					75					80
Gly	Gly	Gly	Arg	Val	Val	Val	Pro	Ala	Gly	Glu	Phe	Leu	Thr	Gly	Ala
				85					90					95	
Ile	His	Leu	Lys	Ser	Asn	Thr	Asn	Leu	Tyr	Val	Ser	Lys	Gly	Ala	Thr
			100					105					110		
Leu	Lys	Phe	Ser	Thr	Asp	Pro	Glu	Lys	Tyr	Leu	Pro	Ile	Val	His	Thr
		115					120					125			
Arg	Trp	Glu	Gly	Met	Glu	Leu	Met	His	Leu	Ser	Pro	Phe	Ile	Tyr	Ala
	130					135					140				
Tyr	Glu	Gln	Thr	Asn	Ile	Ala	Ile	Thr	Gly	Glu	Gly	Thr	Leu	Asp	Gly
145				150						155					160
Gln	Gly	Lys	Ser	Phe	Phe	Trp	Lys	Trp	His	Gly	Asn	Pro	Arg	Tyr	Gly
				165					170					175	
Gly	Asn	Pro	Glu	Val	Ile	Ser	Gln	Gln	Lys	Ala	Arg	Ala	Arg	Leu	Tyr
			180					185					190		
Glu	Met	Met	Asp	Lys	Asn	Val	Pro	Val	Ala	Glu	Arg	Val	Phe	Gly	Ile
		195					200					205			
Gly	His	Tyr	Leu	Arg	Pro	Gln	Phe	Ile	Gln	Pro	Tyr	Lys	Cys	Lys	Asn
	210					215					220				
Val	Leu	Ile	Glu	Gly	Val	Thr	Ile	Ile	Asp	Ser	Pro	Met	Trp	Glu	Val
225					230					235					240
His	Pro	Val	Leu	Cys	Glu	Asn	Val	Thr	Val	Arg	Asn	Leu	His	Ile	Ser
				245					250					255	
Ser	His	Gly	Pro	Asn	Asn	Asp	Gly	Cys	Asp	Pro	Glu	Ser	Cys	Lys	Asp
			260					265					270		
Val	Leu	Ile	Asp	Asn	Cys	Phe	Phe	Asp	Thr	Gly	Asp	Asp	Cys	Ile	Ala
		275					280					285			
Ile	Lys	Ser	Gly	Arg	Asn	Asn	Asp	Gly	Arg	Arg	Leu	Asn	Thr	Pro	Thr
	290					295					300				
Glu	Asn	Ile	Ile	Val	Arg	Asn	Cys	Thr	Met	Lys	Asp	Gly	His	Gly	Gly
305				310						315					320
Ile	Thr	Val	Gly	Ser	Glu	Ile	Ser	Gly	Gly	Val	Arg	Asn	Leu	Phe	Ala
				325					330					335	
His	Asp	Cys	Lys	Met	Asp	Ser	Ala	Asp	Leu	Trp	Thr	Ala	Leu	Arg	Val
			340					345					350		
Lys	Asn	Asn	Ala	Ser	Arg	Gly	Gly	Met	Leu	Glu	Asn	Phe	Tyr	Phe	Arg
		355					360					365			
Asn	Ile	Thr	Val	Gly	Gln	Val	Ala	Arg	Ala	Val	Val	Glu	Ile	Asp	Phe
	370					375					380				
Asn	Tyr	Glu	Glu	Gly	Ala	Lys	Gly	Ser	Tyr	Thr	Pro	Val	Met	Arg	Asn
385				390						395					400
Tyr	Val	Val	Glu	Asp	Leu	Thr	Cys	Thr	Ser	Gly	Asn	Arg	Pro	Val	Asp
				405					410					415	
Leu	Gln	Gly	Leu	Asp	Asn	Ala	Pro	Ile	Tyr	Asp	Val	Ser	Leu	Arg	Asn
			420					425					430		
Thr	Thr	Phe	Gly	Ala	Met	Lys	Asn	Lys	Ser	Val	Val	Lys	Asn	Val	Arg
		435					440					445			
Gly	Leu	Lys	Ile	Glu	Asn	Val	Thr	Val	Ser	Gly	Thr	Arg	Val	Glu	Ser
	450					455					460				
Leu															
465															

<210> 31

<211> 1401

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 31


```

atgatcaatc tttatggcgt ctttgacatc cggacctttg gggcccaacc ggacggagaa 60
acgccttcca ctgcggcgat tacggcggcc atcgaaactt gtgcgcgggc cgggggagga 120
gtggtctaca tcccgccggg acggttctct accggtcccc tccgcctcaa aagccacgtc 180
cggctccatc tcgaggccgg agcgcacttg ctctttagtc aggaccgggc cgattatcct 240
gttctggaga cgaggtggga ggggaaggag gtcttgacct atgcacacca gatctacggc 300
gaggacctcg aaggggtcgc gattaccggg cgggggacca tcgacggccg gggcgagact 360
tggtggcgac tcttcgcgcg caaagccttc acccatcccc gaccccgccct catcgccctt 420
acccgctgca aggacatcct gatagaagga gtaacctcg tcaattcacc ggcctggacc 480
atcaatcctg tgatgtgcga gcgggtgacc atcgataagg tgactatcat caaccgccc 540
gactcgccca acaccgacgg gatcgacccc gattcctccc ggaacgtcta tatcactaac 600
tgctacattg acgtaggcga tgactgcata gccatcaaag cgggccgaga ggactccctt 660
tatcgagcgc cttgtgaaaa cattgtcatc gccaaactgcc tcatgcgcca cggtcacggc 720
ggggtggtca tcggcagcga gaccagcggg ggtattcgca aggtagtcat taccaactgc 780
atcttcgagg acaccgacgg gggcattaga cttaagtccc ggcgcggacg cggcggttc 840
gtcgaggacc tccgggcgac gaattattatc atggaaaagg tgctctgtcc ctctgtctc 900
aacatgtact atgataccgg gggaggcgtg atcgacgagc gcgcgcataga cttagaacc 960
catccggtaa gcgaggctac accctccttc cgcgcctct ccttcagtca cattactgcc 1020
cgggaagtgc agcccgccgc gcccttcctc tacggcctgc ccgaacagcc tctggaggac 1080
gtcttatttg acgatatctg gatagagctg gccgcgacg ctctcctgc ccgtccggcc 1140
atgatgcggg ccgtcccgcc catgagccaa ggtggtgtgc tctgctacgg tcgcgcggcg 1200
atctccttcc ggcacatgca cctccgcggg caccgcggtc cggccttcca gatcgaacgc 1260
gcggaggcgg tgcagttgat gggctgctcg accgacggca gtgaagaccc ccagcttgtc 1320
ttgggtcaag cggaggaggt caccatccgt gactgcacct ttaccgcca gcaggacccc 1380
gcaaaagaaa ggcaaaatta a 1401

```

<210> 32

<211> 466

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10 <221> DOMINIO

<222> (1)...(348)

<223> Dominio catalítico

<400> 32

15

```

Met Ile Asn Leu Tyr Gly Val Phe Asp Ile Arg Thr Phe Gly Ala Gln
 1           5           10           15
Pro Asp Gly Glu Thr Pro Ser Thr Ala Ala Ile Thr Ala Ala Ile Glu
           20           25           30
Thr Cys Ala Ala Ala Gly Gly Gly Val Val Tyr Ile Pro Ala Gly Arg
           35           40           45
Phe Leu Thr Gly Pro Leu Arg Leu Lys Ser His Val Arg Leu His Leu
           50           55           60
Glu Ala Gly Ala His Leu Leu Phe Ser Gln Asp Pro Ala Asp Tyr Pro
65           70           75           80
Val Leu Glu Thr Arg Trp Glu Gly Lys Glu Val Leu Thr Tyr Ala His
           85           90           95
Gln Ile Tyr Gly Glu Asp Leu Glu Gly Val Ala Ile Thr Gly Arg Gly
           100          105          110
Thr Ile Asp Gly Arg Gly Glu Thr Trp Trp Arg Leu Phe Arg Ala Lys
           115          120          125
Ala Phe Thr His Pro Arg Pro Arg Leu Ile Ala Phe Thr Arg Cys Lys
           130          135          140

```

Asp Ile Leu Ile Glu Gly Val Thr Leu Val Asn Ser Pro Ala Trp Thr
145 150 155 160
Ile Asn Pro Val Met Cys Glu Arg Val Thr Ile Asp Lys Val Thr Ile
165 170 175
Ile Asn Pro Pro Asp Ser Pro Asn Thr Asp Gly Ile Asp Pro Asp Ser
180 185 190
Ser Arg Asn Val Tyr Ile Thr Asn Cys Tyr Ile Asp Val Gly Asp Asp
195 200 205
Cys Ile Ala Ile Lys Ala Gly Arg Glu Asp Ser Leu Tyr Arg Thr Pro
210 215 220
Cys Glu Asn Ile Val Ile Ala Asn Cys Leu Met Arg His Gly His Gly
225 230 235 240
Gly Val Val Ile Gly Ser Glu Thr Ser Gly Gly Ile Arg Lys Val Val
245 250 255
Ile Thr Asn Cys Ile Phe Glu Asp Thr Asp Arg Gly Ile Arg Leu Lys
260 265 270
Ser Arg Arg Gly Arg Gly Gly Phe Val Glu Asp Leu Arg Ala Thr Asn
275 280 285
Ile Ile Met Glu Lys Val Leu Cys Pro Phe Val Leu Asn Met Tyr Tyr
290 295 300
Asp Thr Gly Gly Gly Val Ile Asp Glu Arg Ala His Asp Leu Glu Pro
305 310 315 320
His Pro Val Ser Glu Ala Thr Pro Ser Phe Arg Arg Leu Ser Phe Ser
325 330 335
His Ile Thr Ala Arg Glu Val Gln Ala Ala Ala Phe Leu Tyr Gly
340 345 350
Leu Pro Glu Gln Pro Leu Glu Asp Val Leu Phe Asp Asp Ile Trp Ile
355 360 365
Glu Leu Ala Ala Asp Ala Ser Pro Ala Arg Pro Ala Met Met Arg Ala
370 375 380
Val Pro Pro Met Ser Gln Gly Gly Val Leu Cys Tyr Gly Ala Arg Arg
385 390 395 400
Ile Ser Phe Arg His Met His Leu Arg Gly His Arg Gly Pro Ala Phe
405 410 415
Gln Ile Glu Arg Ala Glu Ala Val Gln Leu Met Gly Cys Ser Thr Asp
420 425 430
Gly Ser Glu Asp Pro Gln Leu Val Leu Gly Gln Ala Glu Glu Val Thr
435 440 445
Ile Arg Asp Cys Thr Phe Thr Ala Gln Gln Asp Pro Ala Lys Glu Arg
450 455 460
Gln Asn
465

<210> 33
<211> 1041
5 <212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

10 <400> 33

atgaaacttc gatgtctgat gctcaccctg cttcttttgcg gcagcgcctt cgccgcccag 60
cggattacgg ccgacaagat caacaacaag cccgactcct ggcttaccag cgacgaaggc 120
atcaagctga tcgacaacat catcacctgg cagaaccccg aggggtggctg ggccaagtac 180
tacgacgcga ccaatccgca caaacaaggc gaagtctacg gcgactggga cggcgctcggc 240
accatcgaca acggetacac ctacaccgag ctgaatctoc tggcgcacgt ctacaccctc 300
accaagcgcc cggagatcct cgattcgttc aacaagggcc tggagtttct gctcaaagcc 360
caatacccca gcggcggtcg gccgcaacgg tttccggtgc ccaacaacta cggcaagtgc 420

```

atcacgctca acgacaacgc gatggtgaac gtgatgcagt tcctgcagaa cgtcgcaaag 480
ggcaaggaag acttcgcttt cgtcgacgag cagcgtcgcg ccaaagcgaa ggaggcgttt 540
gaccgcggga tcgactgcct tctgaagctc cagattaccg tgaacggcaa gcttaccgcc 600
tgggcccagc agtatgaccc gaagacactc gccgcggcgc ccgcccgggc gtacgagctc 660
ccgggcctca gcggtgcga aagcgcgcc gtcatgcgt tgttcatgtc tttggagaac 720
cccagtcocg aagttcagcg cgcgctccac gcggcgggcg cttggtacga ggcgtcgaag 780
atcacccgca agaagctggt gcgcgagaac aacgacgtga cactggccga cgaccccaac 840
ggcgagccgc tttgggcgcg cttctaogac atcgaaacca accgcccgtt ctattgcggt 900
cgcgacggcg tgaagaagtg gtcgctggac gagatcgagc ccgaacgccg caagggctac 960
gcttgggtoc gcccctgggc gacgagcgtc ctggagcagt atcgcaagtg ggcggcgaag 1020
caccaccccg tgaacagttg a 1041

```

<210> 34

<211> 346

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10 <221> SEÑAL

<222> (1)...(18)

<221> DOMINIO

<222> (19)...(346)

<223> Dominio catalítico

15

<400> 34

```

Met Lys Leu Arg Cys Leu Met Leu Thr Leu Leu Leu Cys Gly Ser Ala
 1      5      10      15
Phe Ala Ala Asp Arg Ile Thr Ala Asp Lys Ile Asn Asn Lys Pro Asp
 20      25      30
Ser Trp Leu Thr Ser Asp Glu Gly Ile Lys Leu Ile Asp Asn Ile Ile
 35      40      45
Thr Trp Gln Asn Pro Glu Gly Gly Trp Ala Lys Tyr Tyr Asp Ala Thr
 50      55      60
Asn Pro His Lys Gln Gly Glu Val Tyr Gly Asp Trp Asp Gly Val Gly
 65      70      75      80
Thr Ile Asp Asn Gly Tyr Thr Tyr Thr Glu Leu Asn Leu Leu Ala His
 85      90      95
Val Tyr Thr Leu Thr Lys Arg Pro Glu Ile Leu Asp Ser Phe Asn Lys
100      105      110
Gly Leu Glu Phe Leu Leu Lys Ala Gln Tyr Pro Ser Gly Gly Trp Pro
115      120      125
Gln Arg Phe Pro Val Pro Asn Asn Tyr Gly Lys Cys Ile Thr Leu Asn
130      135      140
Asp Asn Ala Met Val Asn Val Met Gln Phe Leu Gln Asn Val Ala Lys
145      150      155      160
Gly Lys Glu Asp Phe Ala Phe Val Asp Glu Gln Arg Arg Ala Lys Ala
165      170      175
Lys Glu Ala Phe Asp Arg Gly Ile Asp Cys Leu Leu Lys Leu Gln Ile
180      185      190
Thr Val Asn Gly Lys Leu Thr Ala Trp Ala Gln Gln Tyr Asp Pro Lys
195      200      205
Thr Leu Ala Ala Ala Pro Ala Arg Ala Tyr Glu Leu Pro Gly Leu Ser
210      215      220
Gly Cys Glu Ser Ala Pro Val Met Arg Leu Phe Met Ser Leu Glu Asn
225      230      235      240
Pro Ser Pro Glu Val Gln Arg Ala Val His Ala Ala Ala Ala Trp Tyr

```

				245					250					255					
Glu	Ala	Ser	Lys	Ile	Thr	Gly	Lys	Lys	Leu	Val	Arg	Glu	Asn	Asn	Asp				
			260					265					270						
Val	Thr	Leu	Ala	Asp	Asp	Pro	Asn	Gly	Glu	Pro	Leu	Trp	Ala	Arg	Phe				
		275					280					285							
Tyr	Asp	Ile	Glu	Thr	Asn	Arg	Pro	Phe	Tyr	Cys	Gly	Arg	Asp	Gly	Val				
	290				295					300									
Lys	Lys	Trp	Ser	Leu	Asp	Glu	Ile	Glu	Pro	Glu	Arg	Arg	Lys	Gly	Tyr				
305					310					315					320				
Ala	Trp	Val	Arg	Pro	Trp	Ala	Thr	Ser	Val	Leu	Glu	Gln	Tyr	Arg	Lys				
				325					330					335					
Trp	Ala	Ala	Lys	His	Pro	Pro	Val	Asn	Ser										
			340					345											

<210> 35
 <211> 1071
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 35

atgccaaaaa	attccgacga	cgcgtggcgg	gaaaagactc	cgcccgattg	gagtccttgc	60
acatggagcg	acgtattcaa	acagaagcct	ctctggtacc	aaaccgacga	ggcggctcga	120
gtcgcggacc	aactcctcat	ctatcaaaaa	gagaacggcg	ggtttgagaa	gaatgtcgac	180
atggcggttg	tgtcgacgca	gaaggaaaaa	gaagagctca	ccgcaaagcg	gtcagacgtc	240
tccgaaacga	cgatcgacaa	ccggaccacg	tatcctcagg	tcgcgtatct	cggtcgagta	300
atcacccgaa	gccttcttaa	accttcgcgc	ccggcgaatc	ttccgaaata	caaagacgcc	360
ttcaacaaag	gtcttgatta	cctgcttgcc	tcccagtatg	agaacggagg	atttccgcaa	420
ttctatccgt	tgaaaaaagg	ctattacaca	cacatcacct	tcaacgacga	cgcgatgatc	480
ggcgctcctga	aggtgcttcg	cgacatcgca	aataagaaag	aggattacgt	gttcgtggat	540
gaagcgcgaa	gacttcgcgc	cgagcaagcg	gtcgccaaag	cgctgcctct	tattctgaag	600
cttcaggttg	tcgtcgacgg	aaagaaaacc	gtctgggctg	cgcagtatga	cgagactacg	660
ctggcgccctg	cagcggctcg	caagtttgag	cccggtgctg	tgaccgctgg	tgagagcgctc	720
ggcatcgctcc	gatacctgat	gcaggaaaaa	ccgacgccgg	agatcaccca	tgcgatcgag	780
tctgcgatcg	attggtatcg	aaagaacaag	atcgacggaa	tacgttgagg	gcgcatcaaa	840
ggcgagaaca	cggttgatga	agacaaatcg	gctcccccta	tatgggcacg	gttctatcag	900
atcgaaacga	tgcgtccgat	cttcacggga	cgtgattcgg	ttatcaagta	tgacgtgacg	960
caggtcgaag	ccgagcgtcg	gaatggttac	gcctggtagc	tcaccgcacc	gaatgaattg	1020
gtgaacgagg	attatttgaa	gtggaagggg	aaaagcgccg	gagccaagta	g	1071

15 <210> 36
 <211> 356
 <212> PRT
 <213> Desconocido

20 <220>
 <223> Obtenido de una muestra ambiental
 <221> DOMINIO
 <222> (1)...(356)
 <223> Dominio catalítico

25 <400> 36

Met	Pro	Lys	Asn	Ser	Asp	Asp	Ala	Trp	Arg	Glu	Lys	Thr	Pro	Pro	Asp
1				5					10					15	
Trp	Ser	Leu	Val	Thr	Trp	Ser	Asp	Val	Phe	Lys	Gln	Lys	Pro	Leu	Trp
			20					25					30		
Tyr	Gln	Thr	Asp	Glu	Ala	Ala	Arg	Val	Ala	Asp	Gln	Leu	Leu	Ile	Tyr
		35					40					45			
Gln	Lys	Glu	Asn	Gly	Gly	Phe	Glu	Lys	Asn	Val	Asp	Met	Ala	Leu	Met
	50					55					60				
Leu	Thr	Gln	Lys	Glu	Lys	Glu	Glu	Leu	Thr	Ala	Lys	Arg	Ser	Asp	Val
65					70					75					80
Ser	Glu	Thr	Thr	Ile	Asp	Asn	Arg	Thr	Thr	Tyr	Pro	Gln	Val	Ala	Tyr
				85					90					95	
Leu	Gly	Arg	Val	Ile	Thr	Ala	Ser	Leu	Leu	Lys	Pro	Ser	Pro	Pro	Ala
			100					105					110		
Asn	Leu	Pro	Lys	Tyr	Lys	Asp	Ala	Phe	Asn	Lys	Gly	Leu	Asp	Tyr	Leu
		115					120					125			
Leu	Ala	Ser	Gln	Tyr	Glu	Asn	Gly	Gly	Phe	Pro	Gln	Phe	Tyr	Pro	Leu
	130					135					140				
Lys	Lys	Gly	Tyr	Tyr	Thr	His	Ile	Thr	Phe	Asn	Asp	Asp	Ala	Met	Ile
145					150					155					160
Gly	Val	Leu	Lys	Val	Leu	Arg	Asp	Ile	Ala	Asn	Lys	Lys	Glu	Asp	Tyr
				165					170					175	
Val	Phe	Val	Asp	Glu	Ala	Arg	Arg	Leu	Arg	Ala	Glu	Gln	Ala	Val	Ala
			180					185					190		
Lys	Ala	Leu	Pro	Leu	Ile	Leu	Lys	Leu	Gln	Val	Val	Val	Asp	Gly	Lys
		195					200					205			
Lys	Thr	Val	Trp	Ala	Ala	Gln	Tyr	Asp	Glu	Thr	Thr	Leu	Ala	Pro	Ala
	210					215					220				
Ala	Ala	Arg	Lys	Phe	Glu	Pro	Val	Ser	Leu	Thr	Ala	Gly	Glu	Ser	Val
225					230					235					240
Gly	Ile	Val	Arg	Tyr	Leu	Met	Gln	Glu	Lys	Pro	Thr	Pro	Glu	Ile	Thr
				245					250					255	
Asp	Ala	Ile	Glu	Ser	Ala	Ile	Asp	Trp	Tyr	Arg	Lys	Asn	Lys	Ile	Asp
			260					265					270		
Gly	Ile	Arg	Trp	Glu	Arg	Ile	Lys	Gly	Glu	Asn	Thr	Val	Val	Lys	Asp
		275					280					285			
Lys	Ser	Ala	Pro	Pro	Ile	Trp	Ala	Arg	Phe	Tyr	Gln	Ile	Glu	Thr	Met
	290					295					300				
Arg	Pro	Ile	Phe	Ile	Gly	Arg	Asp	Ser	Val	Ile	Lys	Tyr	Asp	Val	Thr
305					310					315					320
Gln	Val	Glu	Ala	Glu	Arg	Arg	Asn	Gly	Tyr	Ala	Trp	Tyr	Val	Thr	Ala
				325					330					335	
Pro	Asn	Glu	Leu	Val	Asn	Glu	Asp	Tyr	Leu	Lys	Trp	Lys	Gly	Lys	Ser
			340					345					350		
Ala	Gly	Ala	Lys												
			355												

<210> 37
 <211> 1860
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 37

atgttcaacta	ctactggctc	tcattgccc	cggaattccg	cggttttttc	ccttactgcg	60
atagcagccg	ctgttgcggt	gatggcaggc	acttcagcat	ttgcagctgc	gacgggtggc	120
ttctctacca	ctgatgggtg	caacgtatcg	ggcgcccggt	cgtttactgc	atcgacttac	180
cagcaaatca	acaccattat	tgccaacgca	aaactggatg	atgcaggtaa	aaaagtcact	240
gggggtgctt	acccgcttat	cattacctac	accggtaatg	aagactcgct	gattaaccag	300
atgatcaaag	accacacggg	gaattcatcg	ggcaactgcc	ctaaccgcg	ttggagcgaa	360
gcctatcgct	acgtggaaat	taaagagttt	accaagggtg	ttaccattca	aggcgcgaa	420
ggttcttcag	caaacttcgg	cattgtgatt	aataaatctg	acaatgtgat	tgtgcgtaat	480
atgaaaatcg	gtgcgcttgc	tggtgogagt	aacgatgcgg	atatgattcg	tatcgacacc	540
ggcggttaacg	tgtggattga	tcacaacgaa	ttgtttgcgg	taaataatga	atgtaaagggt	600
tcaccggatg	gtgacctgac	atgttgaaagt	gcgattgata	ttaaaaaagc	atcgcaaaat	660
attacgggtgt	cctacaacat	tatccgcgat	agtaaaaaag	tagggctcga	tggttcgagt	720
agcagtgata	ttgcaggtgg	ccgtaagatt	acgttccatc	acaatattta	tgcgaatgtt	780
gggtgcacgtt	taccgttgca	acgcggtggt	tggacacaca	tgtataacaa	totttacgac	840
ggagttacca	gctcgggtat	taacgttcgt	caaggtggct	acgcgcta	cgagaacaac	900
tggttccaaa	atgctgtcaa	cccggttacc	tgcggttttg	acagtagtaa	ctgcggttac	960
tgggatctgc	gcaacaacaa	cgtgcgcaac	cctggtgatt	tctccacct	caacattacc	1020
tggaccagcg	gtggcaccat	cgacgccacc	aactggacta	ccactcaacc	tttcccgatt	1080
agcattccct	acagctactc	gcctgttagc	ccgcagtggt	tcaaagacaa	gttggcaaat	1140
tatgctggtg	tcggtaaaaa	caatgcgcaa	ttaacggcgt	ctgcgtgcag	cgaaaatact	1200
tcacggtag	caccttcac	agtgccagca	tcacggcg	caccttcaag	ccgttcaccc	1260
agcagtgag	cgccatccag	cacaccaact	acatcaagct	cgagttcagt	tgccgcaacc	1320
ggttcaattt	cgctcggtgc	aacggcaacc	aacaacagca	ttgtgttgag	ttggtcacc	1380
aacaatgtga	cgctcggttc	gcaagaagtg	tatcgcgata	ccgacgctga	tccatcgggg	1440
cgtgtgcgta	tcgcatccct	ggctgcttca	gcgcgtatgt	ataccgatag	cacagcggca	1500
tcggggccaaa	cctattacta	ctggattaaa	aataccactt	ctggtgttgt	caccaattcc	1560
aatgctgcat	cagcgcgtat	tggtagcacg	gcgtccagtt	ctggtgcac	aagcagctca	1620
agttcaagcg	gcggcgcgcc	cgtattaggt	ggtactgggt	attatccaag	cggttctcc	1680
aagtgcgctg	atttgggcgg	gacttgttca	gtgtcatcgg	gcgatggctg	ggttgcgttt	1740
ggtcgcaaa	gcaagtgggt	taccaagaaa	gtatcggtag	gtagttcaat	cgctgtacc	1800
gttgcggcat	ttggttcgga	tccacagggc	aaccctaaca	agtgttctta	caaacgttaa	1860

<210> 38

5 <211> 619

<212> PRT

<213> Desconocido

<220>

10 <223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(35)

<221> DOMINIO

<222> (36)...(387)

15 <223> Dominio catalítico

<400> 38

Met	Phe	Thr	Thr	Thr	Gly	Ser	His	Cys	Ala	Arg	Asn	Ser	Ala	Arg	Phe
1				5					10					15	
Ser	Leu	Thr	Ala	Ile	Ala	Ala	Ala	Val	Ala	Leu	Met	Ala	Gly	Thr	Ser
			20					25					30		
Ala	Phe	Ala	Ala	Ala	Thr	Gly	Gly	Phe	Ser	Thr	Thr	Asp	Gly	Gly	Asn
		35				40						45			
Val	Ser	Gly	Ala	Arg	Ser	Phe	Thr	Ala	Ser	Thr	Tyr	Gln	Gln	Ile	Asn
	50				55						60				
Thr	Ile	Ile	Ala	Asn	Ala	Lys	Leu	Asp	Asp	Ala	Gly	Lys	Lys	Val	Thr
65				70						75					80
Gly	Gly	Ala	Tyr	Pro	Leu	Ile	Ile	Thr	Tyr	Thr	Gly	Asn	Glu	Asp	Ser
				85					90					95	
Leu	Ile	Asn	Gln	Met	Ile	Lys	Asp	His	Thr	Val	Asn	Ser	Ser	Gly	Asn
			100					105						110	
Cys	Pro	Asn	Pro	Arg	Trp	Ser	Glu	Ala	Tyr	Arg	Tyr	Val	Glu	Ile	Lys
	115					120						125			
Glu	Phe	Thr	Lys	Gly	Ile	Thr	Ile	Gln	Gly	Ala	Asn	Gly	Ser	Ser	Ala
	130					135					140				

Asn	Phe	Gly	Ile	Val	Ile	Asn	Lys	Ser	Asp	Asn	Val	Ile	Val	Arg	Asn	145	150	155	160
Met	Lys	Ile	Gly	Ala	Leu	Ala	Gly	Ala	Ser	Asn	Asp	Ala	Asp	Met	Ile	165	170	175	
Arg	Ile	Asp	Thr	Gly	Val	Asn	Val	Trp	Ile	Asp	His	Asn	Glu	Leu	Phe	180	185	190	
Ala	Val	Asn	Asn	Glu	Cys	Lys	Gly	Ser	Pro	Asp	Gly	Asp	Leu	Thr	Phe	195	200	205	
Glu	Ser	Ala	Ile	Asp	Ile	Lys	Lys	Ala	Ser	Gln	Asn	Ile	Thr	Val	Ser	210	215	220	
Tyr	Asn	Ile	Ile	Arg	Asp	Ser	Lys	Lys	Val	Gly	Leu	Asp	Gly	Ser	Ser	225	230	235	240
Ser	Ser	Asp	Ile	Ala	Gly	Gly	Arg	Lys	Ile	Thr	Phe	His	His	Asn	Ile	245	250	255	
Tyr	Arg	Asn	Val	Gly	Ala	Arg	Leu	Pro	Leu	Gln	Arg	Gly	Gly	Trp	Thr	260	265	270	
His	Met	Tyr	Asn	Asn	Leu	Tyr	Asp	Gly	Val	Thr	Ser	Ser	Gly	Ile	Asn	275	280	285	
Val	Arg	Gln	Gly	Gly	Tyr	Ala	Leu	Ile	Glu	Asn	Asn	Trp	Phe	Gln	Asn	290	295	300	
Ala	Val	Asn	Pro	Val	Thr	Cys	Arg	Phe	Asp	Ser	Ser	Asn	Cys	Gly	Tyr	305	310	315	320
Trp	Asp	Leu	Arg	Asn	Asn	Asn	Val	Arg	Asn	Pro	Gly	Asp	Phe	Ser	Thr	325	330	335	
Tyr	Asn	Ile	Thr	Trp	Thr	Ser	Gly	Gly	Thr	Ile	Asp	Ala	Thr	Asn	Trp	340	345	350	
Thr	Thr	Thr	Gln	Pro	Phe	Pro	Ile	Ser	Ile	Pro	Tyr	Ser	Tyr	Ser	Pro	355	360	365	
Val	Ser	Pro	Gln	Cys	Val	Lys	Asp	Lys	Leu	Ala	Asn	Tyr	Ala	Gly	Val	370	375	380	
Gly	Lys	Asn	Asn	Ala	Gln	Leu	Thr	Ala	Ser	Ala	Cys	Ser	Gly	Asn	Thr	385	390	395	400
Ser	Ser	Val	Ala	Pro	Ser	Ser	Val	Pro	Ala	Ser	Ser	Ala	Ala	Pro	Ser	405	410	415	
Ser	Arg	Ser	Ser	Ser	Ser	Ala	Ala	Pro	Ser	Ser	Thr	Pro	Thr	Thr	Ser	420	425	430	
Ser	Ser	Ser	Ser	Val	Ala	Ala	Thr	Gly	Ser	Ile	Ser	Leu	Gly	Ala	Thr	435	440	445	
Ala	Thr	Asn	Asn	Ser	Ile	Val	Leu	Ser	Trp	Ser	Pro	Asn	Asn	Val	Thr	450	455	460	
Leu	Gly	Ser	Gln	Glu	Val	Tyr	Arg	Asp	Thr	Asp	Ala	Asp	Pro	Ser	Gly	465	470	475	480
Arg	Val	Arg	Ile	Ala	Ser	Leu	Ala	Ala	Ser	Ala	Arg	Met	Tyr	Thr	Asp	485	490	495	
Ser	Thr	Ala	Ala	Ser	Gly	Gln	Thr	Tyr	Tyr	Tyr	Trp	Ile	Lys	Asn	Thr	500	505	510	
Thr	Ser	Gly	Val	Val	Thr	Asn	Ser	Asn	Ala	Ala	Ser	Ala	Arg	Ile	Gly	515	520	525	
Ser	Thr	Ala	Ser	Ser	Ser	Val	Ala	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Gly	530	535	540	
Gly	Ala	Pro	Val	Leu	Gly	Gly	Thr	Gly	Asp	Tyr	Pro	Ser	Gly	Phe	Ser	545	550	555	560
Lys	Cys	Ala	Asp	Leu	Gly	Gly	Thr	Cys	Ser	Val	Ser	Ser	Gly	Asp	Gly	565	570	575	
Trp	Val	Ala	Phe	Gly	Arg	Lys	Gly	Lys	Trp	Val	Thr	Lys	Lys	Val	Ser	580	585	590	
Val	Gly	Ser	Ile	Ala	Cys	Thr	Val	Ala	Ala	Phe	Gly	Ser	Asp	Pro		595	600	605	
Gln	Gly	Asn	Pro	Asn	Lys	Cys	Ser	Tyr	Lys	Arg						610	615		

<210> 39

5 <211> 1077

<212> ADN
<213> Desconocido

<220>

5 <223> Obtenido de una muestra ambiental

<400> 39

```

atggcgccga tccctccgacc caacctcctt tgcacttacg cgctctgcat gggcttgctc 60
gccgtggtga gctgcgcggc ggggcccgtg tcagcgcagc agccggcgcc atggagcacg 120
gccatcgtgg agcaggagga gagcgcgttc gcctcccgt cgatgcgcag cgtcgcgcac 180
aacgtcgtgc gccatcagtc ggccgaaggc ggctggccta agaaccacaa tctggcggcg 240
ccgccatcgg ggccggcgcc ggagggcgtc gccaatacga tcgacaatga tgcgacgacg 300
ctgccgatgg agtttcttggc gctgtgtatc cagcccggcg gcgtccgata caagccggcc 360
ttcgagcgcg ggctggatta tctgcttgcg gtcagtagc cgaacggcgg ctggccgcag 420
ttctatccgc tgcgcggggg ctattacgat cactgtacgt tcaacgacga cgccatgatc 480
cgggtgatga ttctgctcgg cgcagtggcg cgcggcgggg cgccctatga atttgtcgac 540
gccggcgccg gcgcgcgcgc tgcagccgcg gtcgagcggg gcctggcgct catcctgcgc 600
acgcagatcc ggcaggcgcg ggcgctgacg gtctggtgcg cgcagtatga cagcgccacc 660
ttgcagcccg cctgggcgcg cgcctatgag ccgcccgtccc tgtccggcgc ggaagtggtg 720
gggatcgtgc gctatctcat gtcgatcgac catccctcgc ccgaagtcgt cgcgcgcgtc 780
gacggcgctg tggcatgget gcgcgcggcc gccattgccg gcgtgcgcgt ggagaatttc 840
acggacgccg acggccgccc tgaccgcgcg gccgtggccg acgcgggcgc gccgccgatc 900
tgggcgcggt tctacgagtt cggcgccaac cggccgatct tctggggcg tgattccgtt 960
tttcaactaca cgttcggaga aatcgagcgc gagcggcgcg caggctacaa ttattacgga 1020
tactgggcgc gctccgtgct ggaagactat ccggcctggc gcgcgcgcgt gcgatga 1077

```

10

<210> 40
<211> 358
<212> PRT
<213> Desconocido

15

<220>
<223> Obtenido de una muestra ambiental
<221> SEÑAL
<222> (1)...(32)
20 <221> DOMINIO
<222> (33)...(358)
<223> Dominio catalítico

25

<400> 40

```

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

```

[illegible]

<210> 41
<211> 1080
<212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 41

atgaaaaaatt	taaaaatacag	tttagtttca	tttgtaactac	tcattactat	gaatgttttt	60
acgcaagaaa	aaaaagtaac	ttggaaaagc	atcacagaaa	ataacgatga	aaattgggtt	120
gtaagcgaag	aagccaaaaa	aatagccgaa	aatgttttgt	tatatcaacg	cgatattggt	180
ggttggccaa	aaaacactga	aattcaaaat	gaactttcag	aaaaagaaaa	actaacatta	240
aaagaatttaa	aatcggatcc	aaaaggatgt	accatcgaca	atggtgcaac	gtgtcaggaa	300
ttacttttct	tatccaaaat	atataaatcc	aatccagatg	agcgatataa	aatgggtttc	360
ttaaaagggtg	tgatttacct	gattacagct	caatacaaaa	atggtgggtg	gccacaatat	420
taccctttga	gagaaggata	ttacactcat	attacttaca	acgataatgc	aatggtggaat	480
gttttaaagt	tgttgaagata	agttaaagat	aaatctgatt	actactcaat	tcaagcacc	540
gatgaaattt	ccaaaatggc	tgaagtatca	tttaataaag	gagtcgattg	catattaaaa	600
acacagtaca	aacaaaatgg	aatattaacc	gcttggtgtg	cacaacatga	cagggaaaca	660
ttgaaacctg	ctaaagcaag	agcttatgaa	ttgccttcgt	taagcggaaa	agaatcagcc	720
aaaatttgtt	tgttattaat	gtcaatcgaa	aatccatcta	aagaagtaat	tactgccgta	780
aattcagcag	ttaattgggt	tgaaaaaaca	aaaatcaacg	gaattaaaat	tgaaccatt	840
tccaccggga	aaaaggatga	aaaagataga	attgttggtg	aaagtcctga	tgctccgccg	900
ctttggggcaa	gatttatgga	attaagtac	aacaaaccat	ttttttgtga	tcgtgacgga	960
aagaaaaaat	acagcatgtc	agaaattagt	caagagcgta	gaaccggcta	tgcatggtac	1020
accaacgaac	caaaagaagt	tttaaaaaaa	tacgatgatt	ggaagtcata	attaaactaa	1080

<210> 42
 <211> 359
 <212> PRT
 <213> Desconocido

5

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(21)
 <221> DOMINIO
 <222> (22)...(359)
 <223> Dominio catalítico

10

<400> 42

15

Met	Lys	Asn	Leu	Lys	Tyr	Ser	Leu	Val	Ser	Phe	Val	Leu	Leu	Ile	Thr	1	5	10	15
Met	Asn	Val	Phe	Thr	Gln	Glu	Lys	Lys	Val	Thr	Trp	Lys	Ser	Ile	Thr	20	25	30	
Glu	Asn	Asn	Asp	Glu	Asn	Trp	Phe	Val	Ser	Glu	Glu	Ala	Lys	Lys	Ile	35	40	45	
Ala	Glu	Asn	Val	Leu	Leu	Tyr	Gln	Arg	Asp	Ile	Gly	Gly	Trp	Pro	Lys	50	55	60	
Asn	Thr	Glu	Ile	Gln	Asn	Glu	Leu	Ser	Glu	Lys	Glu	Lys	Leu	Thr	Leu	65	70	75	80
Lys	Glu	Leu	Lys	Ser	Asp	Pro	Lys	Gly	Cys	Thr	Ile	Asp	Asn	Gly	Ala	85	90	95	
Thr	Cys	Gln	Glu	Leu	Leu	Phe	Leu	Ser	Lys	Ile	Tyr	Lys	Ser	Asn	Pro	100	105	110	
Asp	Glu	Arg	Tyr	Lys	Met	Ala	Phe	Leu	Lys	Gly	Val	Ile	Tyr	Leu	Ile	115	120	125	
Thr	Ala	Gln	Tyr	Lys	Asn	Gly	Gly	Trp	Pro	Gln	Tyr	Tyr	Pro	Leu	Arg	130	135	140	
Glu	Gly	Tyr	Tyr	Thr	His	Ile	Thr	Tyr	Asn	Asp	Asn	Ala	Met	Val	Asn	145	150	155	160
Val	Leu	Lys	Leu	Leu	Lys	Glu	Val	Lys	Asp	Lys	Ser	Asp	Tyr	Tyr	Ser	165	170	175	
Ile	Gln	Ala	Pro	Asp	Glu	Ile	Ser	Lys	Met	Ala	Glu	Val	Ser	Phe	Asn	180	185	190	
Lys	Gly	Val	Asp	Cys	Ile	Leu	Lys	Thr	Gln	Tyr	Lys	Gln	Asn	Gly	Ile	195	200	205	
Leu	Thr	Ala	Trp	Cys	Ala	Gln	His	Asp	Arg	Glu	Thr	Leu	Lys	Pro	Ala	210	215	220	
Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Pro	Ser	Leu	Ser	Gly	Lys	Glu	Ser	Ala	225	230	235	240
Lys	Ile	Val	Leu	Leu	Leu	Met	Ser	Ile	Glu	Asn	Pro	Ser	Lys	Glu	Val	245	250	255	
Ile	Thr	Ala	Val	Asn	Ser	Ala	Val	Asn	Trp	Phe	Glu	Lys	Thr	Lys	Ile	260	265	270	
Asn	Gly	Ile	Lys	Ile	Glu	Thr	Ile	Ser	Thr	Gly	Lys	Lys	Asp	Glu	Lys	275	280	285	
Asp	Arg	Ile	Val	Val	Glu	Ser	Pro	Asp	Ala	Pro	Pro	Leu	Trp	Ala	Arg	290	295	300	
Phe	Met	Glu	Leu	Ser	Asp	Asn	Lys	Pro	Phe	Phe	Cys	Asp	Arg	Asp	Gly	305	310	315	320
Lys	Lys	Lys	Tyr	Ser	Met	Ser	Glu	Ile	Ser	Gln	Glu	Arg	Arg	Thr	Gly	325	330	335	
Tyr	Ala	Trp	Tyr	Thr	Asn	Glu	Pro	Lys	Glu	Val	Leu	Lys	Lys	Tyr	Asp	340	345	350	
Asp	Trp	Lys	Ser	Ser	Leu	Asn										355			

<210> 43
 <211> 1902

20

<212> ADN
<213> Desconocido

<220>

5 <223> Obtenido de una muestra ambiental

<400> 43

```
gtggatccaa agaattgggg cagcggattt accggcgaaa tcaaagtaac taacaacaca      60
agccaaacag tcaatagctg gtctgtgtca tggcaagagg caggagccag tgtaactaat      120
tcttgaatg caaccttggg agggacgaat ccttataaccg caaccgggtt aggatggaac      180
tcaaccctgg cgcccgaggc ctctgccagt tttggttttc aagcaaacgg cactgcgggg      240
gcaccaaagg taaatggcag tttgtgtggt gcgactgcac catctgcagc gaccagcaaa      300
tccagtgcga gtgttgccag ttcaaagatt gcaagttcaa ttcaatcaag tgcaactagc      360
agtccaatat cgtccagttc tgctgcacct tcaagcacgc caaaatccag tagctctgct      420
ccaacggctg catcattcac tattcaagaa gagcaagccg gtttttgccg tgtagacggt      480
attgcaacgg aaagtaccaa caccggattc accggcaacg gctacacca ttccaataat      540
gtacaagggt ctgccattgt gtgggaggta aatgcaacta ccagtgcacg ccatacaatt      600
actttccgct tcgctaattg tggcactgcg aatcgcaatg gctcgctagt cattaacggc      660
ggcagcaatg gtaattacac ggtgcaatta ccacgcaccg cgagctgggc tgactggcaa      720
acagtaagtc tggaaattga tttggtacaa ggcaataaca atttgcaact caccgcattg      780
actgcagatg gcctcgcaaa tatcgacttc atcaaaattg aaggagcatc aaccaaacgc      840
ggaacctgtg cagggtgcgt cagcagtagc agtgttgcc ttcgggtaaa atccagtgtc      900
agcgcggcaa gcagttctgt accaacgaac accggcgcca tgctaacttt ggatggcaac      960
cctgccgcaa gctggcttaa caaatcgctg acaaagtggg gcgcacgcg cgctgacatt     1020
gttgctcttt atcaacagtc caacggcggc tggccaaaaa atctggatta caattcagtg     1080
agcgtggtta atggcggcag tgcaagcggc accatcgata atggtgcaac tattactgaa     1140
atggtttatc tcgctgaggt ttacaaaacc ggaaacaata ccaagtaccg cgatgcagtt     1200
cgccgtgcag caaactttat cgtgagttcg caatatagca ctggcgcggt gccgcaattt     1260
tatccgctca aaggtggcta tgcagaccac gccaccttta atgataacgc catggcttac     1320
gcattaactg tattggattt cgctgcaaac aagcgcgcgc cttttgatac ggatgtcttt     1380
aatgacacag accgogcaaa atttaaaaca gcggtaacca aaggtgttga ttacatttta     1440
aaagcgcaat ggaaacaaaa tggaaaatta acagcctggt gcgcacaaca tggcgcgact     1500
gactatcaac ctaaaaaagc acgcgcttat gaattggaat cactgagtg tagcgagtc     1560
gttggtgtaa ttgcattttt aatgacgcag ccgcagacag cacaatatca aacggcggtt     1620
aaagcaggcc tcaactggtt caatagcccg agcacctatt tggaagggtta cacctacgat     1680
tcattccaaag cgtccactaa tcccatagtg cagaaagcgg gaagtagaat gtggtatcgc     1740
ttttacgatt taaataccaa ccgtgggttt ttacgcgacc gggacggcag caaattctat     1800
gacattacca aaatgtctga agaacgtcgc acgggttata gttggggtgg cgcttatggt     1860
gagagcatca tcgcctttgg caaaaaagtg ggctatctat aa                        1902
```

10

<210> 44
<211> 633
<212> PRT
<213> Desconocido

15

<220>
<223> Obtenido de una muestra ambiental

<221> UNIÓN

<222> (4)...(89)

20

<223> Módulo de unión a carbohidratos

<221> UNIÓN

<222> (152)...(275)

<223> Módulo de unión a carbohidratos

<221> DOMINIO

25

<222> (277)...(633)

<223> Dominio catalítico

<400> 44

Met	Asp	Pro	Lys	Asn	Trp	Gly	Ser	Gly	Phe	Thr	Gly	Glu	Ile	Lys	Val
1				5					10					15	
Thr	Asn	Asn	Thr	Ser	Gln	Thr	Val	Asn	Ser	Trp	Ser	Val	Ser	Trp	Gln
			20					25					30		
Glu	Ala	Gly	Ala	Ser	Val	Thr	Asn	Ser	Trp	Asn	Ala	Thr	Leu	Gly	Gly
		35					40					45			
Thr	Asn	Pro	Tyr	Thr	Ala	Thr	Gly	Leu	Gly	Trp	Asn	Ser	Thr	Leu	Ala
	50					55					60				
Pro	Gly	Ala	Ser	Ala	Ser	Phe	Gly	Phe	Gln	Ala	Asn	Gly	Thr	Ala	Gly
65					70					75					80
Ala	Pro	Lys	Val	Asn	Gly	Ser	Leu	Cys	Gly	Ala	Thr	Ala	Ser	Ser	Ala
				85					90					95	
Ala	Thr	Ser	Lys	Ser	Ser	Ala	Ser	Val	Ala	Ser	Ser	Lys	Ile	Ala	Ser
			100					105					110		
Ser	Ile	Gln	Ser	Ser	Ala	Thr	Ser	Ser	Ser	Lys	Ser	Ser	Ser	Ser	Ala
	115						120					125			
Ala	Pro	Ser	Ser	Thr	Pro	Lys	Ser	Ser	Ser	Ser	Ala	Pro	Thr	Ala	Ala
	130					135					140				
Ser	Phe	Thr	Ile	Gln	Glu	Gln	Ala	Gly	Phe	Cys	Arg	Val	Asp	Gly	
145				150					155					160	
Ile	Ala	Thr	Glu	Ser	Thr	Asn	Thr	Gly	Phe	Thr	Gly	Asn	Gly	Tyr	Thr
			165					170						175	
Asn	Ser	Asn	Asn	Val	Gln	Gly	Ala	Ala	Ile	Val	Trp	Ala	Val	Asn	Ala
			180					185					190		
Thr	Thr	Ser	Ala	Arg	His	Thr	Ile	Thr	Phe	Arg	Phe	Ala	Asn	Gly	Gly
	195					200						205			
Thr	Ala	Asn	Arg	Asn	Gly	Ser	Leu	Val	Ile	Asn	Gly	Gly	Ser	Asn	Gly
	210				215						220				
Asn	Tyr	Thr	Val	Gln	Leu	Pro	Arg	Thr	Ala	Ser	Trp	Ala	Asp	Trp	Gln
225				230					235					240	
Thr	Val	Ser	Leu	Glu	Ile	Asp	Leu	Val	Gln	Gly	Asn	Asn	Asn	Leu	Gln
			245						250					255	
Leu	Thr	Ala	Leu	Thr	Ala	Asp	Gly	Leu	Ala	Asn	Ile	Asp	Phe	Ile	Lys
			260				265						270		
Ile	Glu	Gly	Ala	Ser	Thr	Lys	Ala	Gly	Thr	Cys	Ala	Gly	Ala	Val	Ser
	275					280						285			
Ser	Ser	Ser	Val	Ala	Ser	Ser	Val	Lys	Ser	Ser	Ala	Ser	Ala	Ala	Ser
	290					295					300				
Ser	Ser	Val	Pro	Thr	Asn	Thr	Gly	Ala	Met	Leu	Thr	Leu	Asp	Gly	Asn
305				310					315					320	
Pro	Ala	Ala	Ser	Trp	Leu	Asn	Lys	Ser	Arg	Thr	Lys	Trp	Ser	Ala	Ser
			325						330					335	
Arg	Ala	Asp	Ile	Val	Ala	Ser	Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro
			340				345						350		
Lys	Asn	Leu	Asp	Tyr	Asn	Ser	Val	Ser	Ala	Gly	Asn	Gly	Gly	Ser	Ala
	355					360						365			
Ser	Gly	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu
	370					375					380				

Ala Glu Val Tyr Lys Thr Gly Asn Asn Thr Lys Tyr Arg Asp Ala Val
 385 390 395 400
 Arg Arg Ala Ala Asn Phe Ile Val Ser Ser Gln Tyr Ser Thr Gly Ala
 405 410 415
 Leu Pro Gln Phe Tyr Pro Leu Lys Gly Gly Tyr Ala Asp His Ala Thr
 420 425 430
 Phe Asn Asp Asn Gly Met Ala Tyr Ala Leu Thr Val Leu Asp Phe Ala
 435 440 445
 Ala Asn Lys Arg Ala Pro Phe Asp Thr Asp Val Phe Asn Asp Thr Asp
 450 455 460
 Arg Ala Lys Phe Lys Thr Ala Val Thr Lys Gly Val Asp Tyr Ile Leu
 465 470 475 480
 Lys Ala Gln Trp Lys Gln Asn Gly Lys Leu Thr Ala Trp Cys Ala Gln
 485 490 495
 His Gly Ala Thr Asp Tyr Gln Pro Lys Lys Ala Arg Ala Tyr Glu Leu
 500 505 510
 Glu Ser Leu Ser Gly Ser Glu Ser Val Gly Val Ile Ala Phe Leu Met
 515 520 525
 Thr Gln Pro Gln Thr Ala Gln Ile Gln Thr Ala Val Lys Ala Gly Leu
 530 535 540
 Asn Trp Phe Asn Ser Pro Ser Thr Tyr Leu Glu Gly Tyr Thr Tyr Asp
 545 550 555 560
 Ser Ser Lys Ala Ser Thr Asn Pro Ile Val Gln Lys Ala Gly Ser Arg
 565 570 575
 Met Trp Tyr Arg Phe Tyr Asp Leu Asn Thr Asn Arg Gly Phe Phe Ser
 580 585 590
 Asp Arg Asp Gly Ser Lys Phe Tyr Asp Ile Thr Lys Met Ser Glu Glu
 595 600 605
 Arg Arg Thr Gly Tyr Ser Trp Gly Gly Ala Tyr Gly Glu Ser Ile Ile
 610 615 620
 Ala Phe Gly Lys Lys Val Gly Tyr Leu
 625 630

<210> 45
 <211> 987
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 45

atgactagac ggcgcttcat cgcggttata tgtttcttcg cggcgtctg cgcgcacgcg 60
 cagtcacacg tgcgtggaa ggacgtgctc gagcagtcg agggctggta ttccacgacc 120
 gccgcgcacg tcgtcgccga caccgtgctg ctgtatcaac gtccatccgg tggatggccg 180
 aaggacatcg acatgacggc gcgcgcggcg gaccgcactc ctccgcgcg tccagacgcg 240
 accatcgaca acggcgccac gaccacgcag atccgcctgc tcgtcgtgc ggcctcgggc 300
 gcaccggcgg ctgcgcgcc caactacacg gcgcgcgcg ttcgcgggat cgattacctg 360
 ctcgaggcgc agtatcccaa cggcggctgg ccgcagttct tccccctgcg caaggactat 420
 tcgcgctacg tcacgttcaa cgacgacgcg atgatgaacg tgatgttcct gctggacgag 480
 gtctcggcgg gagatgcgcc gttcacgttc gtggacgaac aacgcgcgca ccgcgcgcgc 540
 gctgcgcgtg ccaagggggt ctccgtcatc ctgaagtgc aggtccggat cgacgggacg 600
 ctgaccgcct ggtgcgcgca acacgacgag atcaccctgg caccgcgtcc ggcgcgaccc 660
 ttcgagcacg cgtcgctcag cggcaacgag tctgtcgcca tcgtgcgctt cctgatgacc 720
 cgtccgcgga cgcagcgat cgtcgccgcg gtcgatgcgg cggtcgcctg gtcagacgc 780
 gtccgcctcc ctgacggacg gtgggccgga ttctacgagt tcggtaccaa tcgtccgac 840
 ttctcggggc gagacagtgt cgtgcgtac aaactcgagg agatcgaaca ggaacgtcag 900
 gagggctacg cgtggtacgg cactggccg aggcgcctg ttgagaagat gtacctgca 960
 tggaaagtcg ggcttcggg caagtag 987

<210> 46
 <211> 328

<212> PRT
<213> Desconocido

<220>

5 <223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(20)

<221> DOMINIO

<222> (21)...(328)

10 <223> Dominio catalítico

<400> 46

```

Met Thr Arg Arg Ala Phe Ile Ala Val Ile Cys Phe Phe Ala Ala Val
1      5      10      15
Cys Ala His Ala Gln Ser Thr Val Arg Trp Lys Asp Val Leu Glu Gln
20      25      30
Ser Glu Gly Trp Tyr Ser Thr Thr Ala Ala His Val Val Ala Asp Thr
35      40      45
Val Leu Leu Tyr Gln Arg Pro Ser Gly Gly Trp Pro Lys Asp Ile Asp
50      55      60
Met Thr Ala Pro Pro Ala Asp Arg Thr Pro Pro Ala Arg Pro Asp Ala
65      70      75      80
Thr Ile Asp Asn Gly Ala Thr Thr Thr Gln Ile Arg Leu Leu Ala Arg
85      90      95
Ala Ala Ser Gly Ala Pro Ala Ala Ala Ala His Thr Tyr Thr Ala Ala
100     105     110
Ala Leu Arg Gly Ile Asp Tyr Leu Leu Glu Ala Gln Tyr Pro Asn Gly
115     120     125
Gly Trp Pro Gln Phe Phe Pro Leu Arg Lys Asp Tyr Ser Arg Tyr Val
130     135     140
Thr Phe Asn Asp Asp Ala Met Met Asn Val Met Phe Leu Leu Asp Glu
145     150     155     160
Val Ser Ala Gly Asp Ala Pro Phe Thr Phe Val Asp Glu Gln Arg Arg
165     170     175
Asp Arg Ala Arg Ala Ala Val Ala Lys Gly Val Ser Val Ile Leu Lys
180     185     190
Ser Gln Val Arg Ile Asp Gly Thr Leu Thr Ala Trp Cys Ala Gln His
195     200     205
Asp Glu Ile Thr Leu Ala Pro Arg Pro Ala Arg Thr Phe Glu His Ala
210     215     220
Ser Leu Ser Gly Asn Glu Ser Val Ala Ile Val Arg Phe Leu Met Thr
225     230     235     240
Arg Pro Pro Thr Pro Ala Ile Val Ala Ala Val Asp Ala Ala Val Ala
245     250     255
Trp Leu Arg Arg Val Arg Leu Pro Asp Gly Arg Trp Ala Arg Phe Tyr
260     265     270
Glu Phe Gly Thr Asn Arg Pro Ile Phe Ser Gly Arg Asp Ser Val Val
275     280     285
Arg Tyr Lys Leu Glu Glu Ile Glu Gln Glu Arg Gln Glu Gly Tyr Ala
290     295     300
Trp Tyr Gly Thr Trp Pro Arg Thr Leu Val Glu Lys Met Tyr Pro Ala
305     310     315     320
Trp Lys Ser Arg Leu Pro Gly Lys
325

```

15 <210> 47
<211> 1077
<212> ADN
<213> Desconocido

20 <220>
<223> Obtenido de una muestra ambiental

<400> 47

```

atgaaaaatt ttaaaaaatat tgtaggagcg ttacttatat ctgtaacggt ttgtgtgcac 60
gggcaggtaa acaaaaaaatc ctggcgggct attacacagt ctaacgacga tgcattggtt 120
gcatctgatg gagctgcaca gattgcagat aatgtattac tctatcagcg caatgttggc 180
ggatggccta aaaatattga aatgcaggaa ccgcttagtg aggccgacaa aaaaaagctg 240
atagatctta agtctacggc caaagaaagt actacagata atggggctac gtgtcaggaa 300
atggtattcc tctctaagat atataaaca aagcccgaa agaatataa agaggctttt 360
ttaaaaggac ttaattattt gcttgaagca cagtataaaa atggtggatg gccacagttc 420
taccctttaa aaaaagggtt ttataccac attacctata atgacgattc tatggtaaac 480
attcttatga tcttaaagaa tattaaggaa gatgccaact attacagtat tacgccaagc 540
gataaagttt taaagcaggc atcgacagct tttgacagag gcattgactg cattctaaaa 600
acacagtaca agcaaaaagg tgtgcttaca agctgggtgtg cccagcacga tgaggttaca 660
ttagaacctg caaatgcaag ggcttttgag ttggcatcac taagtggtaa agaactctgt 720
aaaataacgt tgttgctaag gtctgtaaaa aatccgtcta aagaggttgt tgctgtgta 780
gatgctgctg tggcgtggtt tgaaaaaaca aaaattgaag gcattaaagt agaagaagta 840
accggagctg atggcaaaaa ggatagggta gtagtacaaa gggctgatgc cgaaccattg 900
tgggcgcgtt ttatggaact ggataccaac aggccatttt tttgcgacag ggacggtata 960
aaaaaatatt cgcttgctga gataggtcat gaacgccgta acggatatgg ctggtacacc 1020
aacgaaccaa aagaagtttt aaagaaatac accaaatgga aaaacagtct taaatag 1077

```

<210> 48
 <211> 358
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(21)
 <221> DOMINIO
 <222> (22)...(358)
 <223> Dominio catalítico

<400> 48

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

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

<210> 49
<211> 1023
5 <212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

10

<400> 49
atgttaagtt tcatcgcggt atcagtggtt cataattact gcacagggca gacagcgctcc 60
accataaatt cagtgggcga aaagatgctt cagtaccagt tgtcaaatgg cgcttgcccc 120
aaacagttgg tagacaaaag tgtcggtgat tacagtcttc cattaacgaa agagcgcta 180
cagcagatca agaaaacaga tattgatcat gctacgctcg acaacagtcg gacaaccggg 240
gaaataactg aattgatcaa ggcttttaag gacactaaaa ataaggcata ttgactgct 300
gtagaaaagg ggattgcata tattttatcg gctcaatatg agaatggcgg atttccacaa 360
tactacccaa ataaattata ctatagagct gagataacat acaacgatga tgcgatgatc 420
aatgcattac tagtgcttta caaagtagcc aataagcgag aggggtttga ggctatcaat 480
cccatatttg tgtcaaaagc gcaaaaagca gttgaaaagg gtataacctg tctcctaaaa 540
acacaggtca tacaagacgg aaaaaggagt atttgggctg cgcaatacga tcagaacact 600
ttacaacctg ctacaggcaag aaagtttgaa ccagcttcat tgagcacaag tgaatctggt 660
tccatcggtc gctttctcat gctacagcct gcaaccactg aaattaagca agcgatcgaa 720
catagcaatc aatgggttcga acagcatgat attgaagggt accgtttcga ccgcatacaa 780
gatagggtga ctggaaaata tcaacggcaa cttgtcgcag atcggacttc cagatttggt 840
gcgcgatttt ataattctga agacaaccgc ccattgtttg gagatcgga caatacaatc 900
aaatacaact ttgaggaggt ttgagaggag cgtagaaatg gctatgcttg gttcggcaac 960
tggccggaag agctgatcca aaaggactat ccaaaatgga aaaaacaata caaaattaaa 1020

taa

1023

<210> 50

<211> 340
 <212> PRT
 <213> Desconocido

5 <220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(16)
 <221> DOMINIO
 10 <222> (17)...(340)
 <223> Dominio catalítico

<400> 50

Met	Leu	Ser	Phe	Ile	Ala	Val	Ser	Val	Phe	His	Asn	Tyr	Cys	Thr	Gly	1	5	10	15
Gln	Thr	Ala	Ser	Thr	Lys	Asn	Ser	Val	Ala	Glu	Lys	Met	Leu	Gln	Tyr	20	25	30	
Gln	Leu	Ser	Asn	Gly	Ala	Trp	Pro	Lys	Gln	Leu	Val	Asp	Lys	Ser	Val	35	40	45	
Val	Asp	Tyr	Ser	Leu	Pro	Leu	Thr	Lys	Glu	Arg	Leu	Gln	Gln	Ile	Lys	50	55	60	
Lys	Thr	Asp	Ile	Asp	His	Ala	Thr	Leu	Asp	Asn	Ser	Ala	Thr	Thr	Arg	65	70	75	80
Glu	Ile	Thr	Glu	Leu	Ile	Lys	Ala	Phe	Lys	Asp	Thr	Lys	Asn	Lys	Ala	85	90	95	
Tyr	Leu	Thr	Ala	Val	Glu	Lys	Gly	Ile	Ala	Tyr	Ile	Leu	Ser	Ala	Gln	100	105	110	
Tyr	Glu	Asn	Gly	Gly	Phe	Pro	Gln	Tyr	Tyr	Pro	Asn	Lys	Leu	Tyr	Tyr	115	120	125	
Arg	Ala	Glu	Ile	Thr	Tyr	Asn	Asp	Asp	Ala	Met	Ile	Asn	Ala	Leu	Leu	130	135	140	
Val	Leu	Tyr	Lys	Val	Ala	Asn	Lys	Arg	Glu	Gly	Phe	Glu	Ala	Ile	Asn	145	150	155	160
Pro	Ile	Phe	Val	Ser	Lys	Ala	Gln	Lys	Ala	Val	Glu	Lys	Gly	Ile	Thr	165	170	175	
Cys	Ile	Leu	Lys	Thr	Gln	Val	Ile	Gln	Asp	Gly	Lys	Arg	Ser	Ile	Trp	180	185	190	
Ala	Ala	Gln	Tyr	Asp	Gln	Asn	Thr	Leu	Gln	Pro	Ala	Gln	Ala	Arg	Lys	195	200	205	
Phe	Glu	Pro	Ala	Ser	Leu	Ser	Thr	Ser	Glu	Ser	Val	Ser	Ile	Val	Arg	210	215	220	
Phe	Leu	Met	Leu	Gln	Pro	Ala	Thr	Thr	Glu	Ile	Lys	Gln	Ala	Ile	Glu	225	230	235	240
His	Ala	Ile	Gln	Trp	Phe	Glu	Gln	His	Asp	Ile	Glu	Gly	Tyr	Arg	Phe	245	250	255	
Asp	Arg	Ile	Gln	Asp	Arg	Val	Thr	Gly	Lys	Tyr	Gln	Arg	Gln	Leu	Val	260	265	270	
Ala	Asp	Arg	Thr	Ser	Thr	Ile	Trp	Ala	Arg	Phe	Tyr	Asn	Leu	Glu	Asp	275	280	285	
Asn	Arg	Pro	Leu	Phe	Gly	Asp	Arg	Asp	Asn	Thr	Ile	Lys	Tyr	Asn	Phe	290	295	300	
Glu	Glu	Val	Ser	Glu	Glu	Arg	Arg	Asn	Gly	Tyr	Ala	Trp	Phe	Gly	Asn	305	310	315	320
Trp	Pro	Glu	Lys	Leu	Ile	Gln	Lys	Asp	Tyr	Pro	Lys	Trp	Lys	Lys	Gln	325	330	335	
Tyr	Lys	Ile	Lys													340			

15
 <210> 51
 <211> 1131
 <212> ADN
 20 <213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

<400> 51

5

```
gtgacgtggg atcagatcct tcgtcagcct gccgcctggt acggcgggtcc ggaagcgcga      60
cggatcgcgga atctggtcct gctgtaccag cgcgcgacgg ggggctggcc caagaacatc      120
gacatggcgc ggctggttgc tccggacgat cgcacgacgc tcgcggcgga acgggcccctc      180
accgactcga cgatcgacaa tggatcgacg acgacgcagt tgcggtttct cgcgatggtg      240
cagcacgccc agcaggcacc cgtgcgcgac gccatcacgc acggcctgga ctatctgctg      300
aacgcgcaat actcgaacgg cggatggccg cagtactttc cgctccgaga cgactactcg      360
cgtcacatca cgttcaacga cgacgcgatg atcaatgtaa tgacggtgct acgcgatgtc      420
gcagaagctc gcatgccctt cgaagggatc gacgcgggtcc gtcgggaccg ggcgcgtgtc      480
gccatcacgc gtggcatcga cgtgattctc gggacgcaaa tccgcgtcgg ggaccgtctg      540
acgggctggt gccagcagca tgacgagcgc tccctcgccc ccaccaaggc tcgcgcctac      600
gagcagccat cgatcgccag caaggaaacg gtaaccatca cgcgcttcct catgaccctc      660
gatcgcccca gtcagcagat catcgcgggc atcgaggcgg ctgtcgagtg gttgcgcgtg      720
gcgaccctgt cgggtgtgcg agttgagcgt cggcgggacc cggcgagtcc gaccggatat      780
gacgtcgtcg ccgcgcggga tgcgcgcgca cctccgacct gggcacgggt ctacgagatc      840
ggcacgaacc gcccaatggt ttccggcgcg gacggcgtga tcagattccg gctcgcggac      900
atcgagattg agcgcgcgac cggctacagc tggatgggcg actatgccgc gaggttgctg      960
aacgaggagt atccggcggt ggcgaggcta cgcggggcga gctttcagaa cgcgcgagctg     1020
cacaaggagt ccggtgaagt cgtacacacg gcgatcgtgc acgatcttgc ctcccttgat     1080
gtcgaagaca aagaccagcc gcagccgaaa gtgcttttgc ctgggcggta g      1131
```

<210> 52

<211> 376

10 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

15 <221> DOMINIO

<222> (1)...(376)

<223> Dominio catalítico

<400> 52

20

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

		100		105		110
Phe	Pro	Leu	Arg	Asp	Asp	Tyr
		115		120		125
Ala	Met	Ile	Asn	Val	Met	Thr
		130		135		140
Met	Pro	Phe	Glu	Gly	Ile	Asp
		145		150		155
Ala	Ile	Thr	Arg	Gly	Ile	Asp
		165		170		175
Gly	Asp	Arg	Leu	Thr	Gly	Trp
		180		185		190
Ala	Pro	Thr	Lys	Ala	Arg	Ala
		195		200		205
Glu	Thr	Val	Thr	Ile	Thr	Arg
		210		215		220
Gln	Gln	Ile	Ile	Ala	Ala	Ile
		225		230		235
Ala	Thr	Leu	Ser	Gly	Val	Arg
		245		250		255
Pro	Thr	Gly	Tyr	Asp	Val	Val
		260		265		270
Thr	Trp	Ala	Arg	Phe	Tyr	Glu
		275		280		285
Gly	Arg	Asp	Gly	Val	Ile	Arg
		290		295		300
Arg	Arg	Thr	Gly	Tyr	Ser	Trp
		305		310		315
Asn	Glu	Glu	Tyr	Pro	Ala	Trp
		325		330		335
Asn	Ala	Glu	Leu	His	Lys	Glu
		340		345		350
Val	His	Asp	Leu	Ala	Phe	Leu
		355		360		365
Pro	Lys	Val	Leu	Phe	Ala	Gly
		370		375		

<210> 53
 <211> 1977
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 53

atgaataact	caacaaaaaa	aatgattcgg	ccactcaagg	catctttttgc	cttggggcgct	60
ctcgcaactgg	caatcgcatc	accctcatgg	gcggtttgct	cttacagcgt	aaccaataat	120
tggggctctg	gctttaccgg	agaaattaaa	gtaaccaacg	atacaacatc	gactgtaaat	180
aattgggtctg	tgtcttgcca	ggaatcaggg	gtgaccgtca	ctaacgcgatg	gaatgcaaca	240
ctgagcggat	caaatcctta	taccgcaaca	tcaactcgggt	ggaacggaac	tctcgtcca	300
aaagcttcag	caagttttgg	ttttcaagca	aatggaacag	cgggcgacc	gaaagtaa	360
ggaaccttgt	gtggtaccag	cacatcatca	acaggtacat	cctcagttgc	accttcaccc	420
gtagcgagta	gcgttgctgt	atcaagcagt	aaatcatcaa	gctctgttgc	aaccatcagt	480
agctctaaat	ccagcagcag	tgtgccgaca	gtttcatcat	tcaactattca	ggaagagcaa	540
gccggtttct	gccgtgtaga	tggcattgca	actgaaagta	ctaacactgg	ctatacaggt	600
aattggctaca	ccaacaccac	taatgcgcaa	ggcgctgcaa	ttgaatgggc	aattaatgct	660
cccaacagca	gccgctacac	cctcaccttc	cgttatgcca	atgctgtgtac	cgctaatacgc	720
aatgggttcgt	tattaattaa	cgacggaagc	aatggtaact	acacagtgca	attgccaaagt	780
accggcgcat	gggcaacctg	gcaaaccgtc	agtgttgaag	tggatttggt	gcaaggcaat	840

```

aatattttga aactcgcttc gcttactgct gatggccttg cgaatataga ttcattaaaa 900
attgaaggcg cacaagccaa agctggtgta tgcagcacta cggtaagtag cagctcttcg 960
tcaattaaat caagttccag ttcattcatcg tccagctcaa ctgcagcagt aaaaacatta 1020
acactggatg gtaaccctgc tgcaaacctgg ttttaataaat ccagaaccaa gtggaatgtc 1080
agcagagctg acatcgtact ttcgtatcag caatcaaagt gtggctggcc aaaaaatttg 1140
gactacaact cggtaggctc aggtaaatgg gttagcgaca gcggcactat tgataatgg 1200
gcaaccataa ccgaaatgg gtacctcgct gaagtgtata aaaaatggcg gaataccaaa 1260
taccgcgacg ccgtgcgcag agcagcgaat tttattgtga gttcacaata cagcactgg 1320
gctttaccgc agttttatcc gctgaaagg gttagcgag atcacgctac cttaaatgat 1380
aatggtatgg cttacgcgtt gactgttctg gatttcgcgg taaataaacg cgcgccattt 1440
gataacgata ttttctctga ctctgaccgc agcaaattta aaactgctgt taccaaaggc 1500
gtcgattaca tattaaaagc gcaatggaaa cagaatggaa aattaaccgt atggtgcgca 1560
caacacggtg ctaatgatta tcaaccgaaa aaagcgcggtg cttacgagtt agaattcattg 1620
agtggtagtg aatctgtcgg tgtactcgct ttcttaaatga ctcaaccaca aaccacgcaa 1680
attgaagcag ctgtgcgtgc aggtgtggcc tggtttaata gcccaagcac ctacttgaat 1740
aattacactt acgattcttc caaagcttcg accaatccaa tcgtgccaaa atccggaagc 1800
aaaaatgtgg atcgccttta tgacctgaat accaaccgcg gtttcttcag tgatcgtgac 1860
ggcagcaagt tctacgacat caccctaatg tcagaagagc gtcgcactgg ttacagttgg 1920
ggtggtgact acggcagctc gattatcagc ttgcacacaa aagtgggata tctctaa 1977

```

<210> 54
 <211> 658
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(31)
 <221> UNIÓN
 <222> (32)...(124)
 <223> Módulo de unión a carbohidratos

<221> UNIÓN
 <222> (180)...(303)
 <223> Módulo de unión a carbohidratos
 <221> DOMINIO
 <222> (304)...(658)
 <223> Dominio catalítico

<400> 54

```

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

```

115				120				125							
Ser	Ser	Thr	Gly	Thr	Ser	Ser	Val	Ala	Pro	Ser	Ser	Val	Ala	Ser	Ser
130						135					140				
Val	Ala	Val	Ser	Ser	Ser	Lys	Ser	Ser	Ser	Ser	Ser	Val	Ala	Thr	Ile
145				150						155					160
Ser	Ser	Lys	Ser	Ser	Ser	Val	Pro	Thr	Val	Ser	Ser	Phe	Thr	Ile	
			165					170						175	
Gln	Glu	Glu	Gln	Ala	Gly	Phe	Cys	Arg	Val	Asp	Gly	Ile	Ala	Thr	Glu
			180					185						190	
Ser	Thr	Asn	Thr	Gly	Tyr	Thr	Gly	Asn	Gly	Tyr	Thr	Asn	Thr	Thr	Asn
		195					200					205			
Ala	Gln	Gly	Ala	Ala	Ile	Glu	Trp	Ala	Ile	Asn	Ala	Pro	Asn	Ser	Ser
210						215					220				
Arg	Tyr	Thr	Leu	Thr	Phe	Arg	Tyr	Ala	Asn	Ala	Gly	Thr	Ala	Asn	Arg
225				230						235					240
Asn	Gly	Ser	Leu	Leu	Ile	Asn	Asp	Gly	Ser	Asn	Gly	Asn	Tyr	Thr	Val
			245						250					255	
Gln	Leu	Pro	Ser	Thr	Gly	Ala	Trp	Ala	Thr	Trp	Gln	Thr	Val	Ser	Val
			260					265						270	
Glu	Val	Asp	Leu	Val	Gln	Gly	Asn	Asn	Ile	Leu	Lys	Leu	Ala	Ser	Leu
		275					280					285			
Thr	Ala	Asp	Gly	Leu	Ala	Asn	Ile	Asp	Ser	Leu	Lys	Ile	Glu	Gly	Ala
290						295					300				
Gln	Ala	Lys	Ala	Gly	Val	Cys	Ser	Thr	Thr	Val	Ser	Ser	Ser	Ser	Ser
305					310					315					320
Ser	Ile	Lys	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Thr	Ala
			325						330					335	
Val	Lys	Thr	Leu	Thr	Leu	Asp	Gly	Asn	Pro	Ala	Ala	Asn	Trp	Phe	Asn
			340					345						350	
Lys	Ser	Arg	Thr	Lys	Trp	Asn	Val	Ser	Arg	Ala	Asp	Ile	Val	Leu	Ser
		355					360					365			
Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro	Lys	Asn	Leu	Asp	Tyr	Asn	Ser
370						375					380				
Val	Gly	Ser	Gly	Asn	Gly	Gly	Ser	Asp	Ser	Gly	Thr	Ile	Asp	Asn	Gly
385					390					395					400
Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu	Ala	Glu	Val	Tyr	Lys	Asn	Gly
			405						410					415	
Gly	Asn	Thr	Lys	Tyr	Arg	Asp	Ala	Val	Arg	Arg	Ala	Ala	Asn	Phe	Ile
			420					425						430	
Val	Ser	Ser	Gln	Tyr	Ser	Thr	Gly	Ala	Leu	Pro	Gln	Phe	Tyr	Pro	Leu
			435				440					445			
Lys	Gly	Gly	Tyr	Ala	Asp	His	Ala	Thr</							

Pro Ile Val Pro Lys Ser Gly Ser Lys Met Trp Tyr Arg Phe Tyr Asp
595 600 605
Leu Asn Thr Asn Arg Gly Phe Phe Ser Asp Arg Asp Gly Ser Lys Phe
610 615 620
Tyr Asp Ile Thr Gln Met Ser Glu Glu Arg Arg Thr Gly Tyr Ser Trp
625 630 635 640
Gly Gly Asp Tyr Gly Ser Ser Ile Ile Ser Phe Ala Gln Lys Val Gly
645 650 655
Tyr Leu

<210> 55
<211> 1125
5 <212> ADN
<213> Desconocido

<220>
10 <223> Obtenido de una muestra ambiental
<400> 55

```

gtggctcctag gtaataacgg cggcagcttg agttgcgtcc aatatattgt gattgtgaaa 60
ggaccgggtg gacctcgacc gccgggtgaaa ccggccgtcc aggcgcccgt tagggttacc 120
tggagcgcac gcctagtcca gcggcccgaa tggtagcgga gtgacgaagc gatccgcac 180
gcggacaacg tcctcctcta ccagcgcaac accggcgggg ggccgaagga catagatatg 240
gccgagccca tcccggaaca caggaaagtcc ttttccctca ccgagaagga gccgaccgat 300
gactcgacca tcgacaacgg tgccaccgtg acccagctca agtatctcgc ccgcgtctac 360
aaggcgacca ggctggaacg gttcaaggag ggcttccctca aaggctctcg ctacctcttg 420
gccgcccagt acccgaaagg cggctggccc cagtattatc ctaacttgag gggtactac 480
gccaacatca cttataacga caatgccatg gtgaacgtgc tcaccctcct ccagagcatc 540
gcaaaaaagg ccccgagta cgacttcgtc gaccgggogc gccgggagaa ggccgcccgg 600
gccgtggcga aagggatcga ctgcatcctc aagaccaga tccgtgtcaa tggaaaactt 660
accgcctggt gcgcccagca tgaccccaag acgctggcgc ccgcgcgggc ccgttcgtat 720
gagcttgagt ccatcagcgg tttcgagagc gtcgggatcg tccggttctt aatgagcctc 780
gagaatccga gcccgaaagg catcgaggcg gtagaggccg ccgtgaaatg gttcgaggag 840
gtcaagctta ccgggatcaa ggtggtcgag aaaccggacc cgtcccttcc gggcggttac 900
gaccgcgtgg tggtcgaaga cccaacgcg ccgcccctct gggcccgggt ctacgagatc 960
ggcaaccaac gtcccttctt ctgcggccgc gatggtatca aaaaatacag cctggcggag 1020
atcgaacacg aacgcgggt cggttactcc tggtagacca atgcccggc ctacctcatc 1080
gagaaggagt atccgctctg gcgggcaaaa caccctacca agtaa 1125

```

15 <210> 56
<211> 374
<212> PRT
<213> Desconocido

20 <220>
<223> Obtenido de una muestra ambiental
<221> DOMINIO
<222> (1)...(374)
<223> Dominio catalítico

25 <400> 56

Met	Val	Leu	Gly	Asn	Asn	Gly	Gly	Ser	Leu	Ser	Cys	Val	Gln	Tyr	Ile
1				5				10						15	
Val	Ile	Val	Lys	Gly	Pro	Gly	Gly	Pro	Arg	Pro	Pro	Val	Lys	Pro	Ala
			20					25					30		
Val	Gln	Ala	Pro	Val	Arg	Val	Thr	Trp	Ser	Ala	Cys	Leu	Val	Gln	Arg
		35					40					45			
Pro	Glu	Trp	Tyr	Gly	Ser	Asp	Glu	Ala	Ile	Arg	Ile	Ala	Asp	Asn	Val
	50					55					60				
Leu	Leu	Tyr	Gln	Arg	Asn	Thr	Gly	Gly	Trp	Pro	Lys	Asp	Ile	Asp	Met
65					70					75					80
Ala	Glu	Pro	Ile	Pro	Glu	His	Arg	Lys	Ser	Phe	Phe	Leu	Thr	Glu	Lys
				85					90					95	
Glu	Arg	Thr	Asp	Asp	Ser	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Val	Thr	Gln
			100					105					110		
Leu	Lys	Tyr	Leu	Ala	Arg	Val	Tyr	Lys	Ala	Thr	Arg	Leu	Glu	Arg	Phe
		115					120					125			
Lys	Glu	Gly	Phe	Leu	Lys	Gly	Leu	Asp	Tyr	Leu	Leu	Ala	Ala	Gln	Tyr
	130					135					140				
Pro	Asn	Gly	Gly	Trp	Pro	Gln	Tyr	Tyr	Pro	Asn	Leu	Arg	Gly	Tyr	Tyr
145					150					155					160
Ala	Asn	Ile	Thr	Tyr	Asn	Asp	Asn	Ala	Met	Val	Asn	Val	Leu	Thr	Leu
				165					170					175	
Leu	Gln	Ser	Ile	Ala	Lys	Lys	Ala	Pro	Glu	Tyr	Asp	Phe	Val	Asp	Pro
		180						185					190		
Ala	Arg	Arg	Glu	Lys	Ala	Ala	Arg	Ala	Val	Ala	Lys	Gly	Ile	Asp	Cys
		195					200					205			
Ile	Leu	Lys	Thr	Gln	Ile	Arg	Val	Asn	Gly	Lys	Leu	Thr	Ala	Trp	Cys
	210					215					220				
Ala	Gln	His	Asp	Pro	Lys	Thr	Leu	Ala	Pro	Ala	Pro	Ala	Arg	Ser	Tyr
225					230					235					240
Glu	Leu	Glu	Ser	Ile	Ser	Gly	Phe	Glu	Ser	Val	Gly	Ile	Val	Arg	Phe
				245					250					255	
Leu	Met	Ser	Leu	Glu	Asn	Pro	Ser	Pro	Lys	Val	Ile	Glu	Ala	Val	Glu
		260						265					270		
Ala	Ala	Val	Lys	Trp	Phe	Glu	Glu	Val	Lys	Leu	Thr	Gly	Ile	Lys	Val
		275					280					285			
Val	Glu	Lys	Pro	Asp	Pro	Ser	Leu	Pro	Gly	Gly	Tyr	Asp	Arg	Val	Val
	290					295					300				
Val	Glu	Asp	Pro	Asn	Ala	Pro	Pro	Ile	Trp	Ala	Arg	Phe	Tyr	Glu	Ile
305					310					315					320
Gly	Thr	Asn	Arg	Pro	Phe	Phe	Cys	Gly	Arg	Asp	Gly	Ile	Lys	Lys	Tyr
				325					330					335	
Ser	Leu	Ala	Glu	Ile	Glu	His	Glu	Arg	Arg	Val	Gly	Tyr	Ser	Trp	Tyr
		340					345						350		
Thr	Asn	Ala	Pro	Ala	Tyr	Leu	Ile	Glu	Lys	Glu	Tyr	Pro	Leu	Trp	Arg
		355					360					365			
Ala	Lys	His	Pro	Thr	Lys										
		370													

<210> 57
 <211> 1170
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 57

atggacaaac	gcgtcaaag	gattcatcag	ctttcaaaag	aagaagcaaa	gcagttcgag	60
cccgaaaatt	tcttcaaagg	caaagacggc	tggaatccga	aaaaggcgga	tgaccgctgg	120
ctcgaaaaaa	caaaacctga	ctggcagctc	gttacgtgga	acgacgcgtt	acgccaggcg	180
ccgctctggt	atcaaaccga	tgaagcggcg	cgcattgccc	accagggtgat	tttgtaccag	240
aaagacaacg	gcggctggga	aaaaaatctc	gatatgacgg	cgatgctcac	gcaagccgaa	300
cgcgaaaagc	tcgccaaga	aaaatcgaac	acgtcggaaa	cgacgatoga	caaccgcacg	360
acctacacgc	aagtcgcttt	tctcgccaaa	gtcattacgg	gcagcttgca	gaaaacgact	420

ccgccgacca	atttcccgaa	acataaggaa	gcttttttca	agggcttgga	ttacctgctc	480
gcgtcgcagt	acgaatcggg	cggctttccg	cagttttatc	cgctcaaaaa	aggttattac	540
acgcacatca	cgttcaacga	cgatgcgatg	attggcggtt	tgaagggttt	gcgcgaaatc	600
gccaaaaaga	aggaagacta	tctttttgtt	gacgaagaac	gcgcctgaa	agcggaaaaa	660
tgggtcgaaa	aagcgtgcc	gctgattctg	aaattgcagg	ttgaagtcgg	cggcaaaaaa	720
acggtttggg	cggcgcagta	tgacgaaaac	actttttaac	ccgcagcggc	gcgaaagtgt	780
gaaccgggtt	ctttaacggc	gggcgaatcg	gtcggcatcg	tccggttttt	aatgtacgat	840
tcaaagcccg	accaggcgac	gattgacgg	attgaatctg	ccattcagtg	gtatcgcgcg	900
aacaaaatcg	aaggcattcg	atgggtgcgc	gaaaacggcg	aaaaccgcgt	cgtcaaggac	960
aaaaacgcgc	cgcgattttg	ggcgcggttt	tacgaaatcg	aaacgatgaa	gccgattttc	1020
atcgggogcg	acgccatcat	tcgttacgac	gtgtctgaaa	tcgaagccga	gcgcgcgaac	1080
ggctacgcgt	ggtaogtctc	ggagccgaac	gagctgcttg	aaaaagatta	cccgaatgg	1140
ctggaaaaaa	ttaaaaaatc	agtaaagtaa				1170

<210> 58

<211> 389

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10 <221> DOMINIO

<222> (1)...(389)

<223> Dominio catalítico

<400> 58

Met	Asp	Lys	Arg	Val	Lys	Trp	Ile	His	Gln	Leu	Ser	Lys	Glu	Glu	Ala
1				5					10					15	
Lys	Gln	Phe	Glu	Pro	Glu	Asn	Phe	Leu	Lys	Gly	Lys	Asp	Gly	Trp	Asn
		20					25						30		
Pro	Lys	Lys	Ala	Asp	Asp	Arg	Trp	Leu	Glu	Lys	Thr	Lys	Pro	Asp	Trp
		35					40					45			
Gln	Leu	Val	Thr	Trp	Asn	Asp	Ala	Leu	Arg	Gln	Ala	Pro	Leu	Trp	Tyr
	50					55					60				
Gln	Thr	Asp	Glu	Ala	Ala	Arg	Ile	Ala	Asp	Gln	Val	Ile	Leu	Tyr	Gln
65					70					75					80
Lys	Asp	Asn	Gly	Gly	Trp	Glu	Lys	Asn	Leu	Asp	Met	Thr	Ala	Met	Leu
			85						90					95	
Thr	Gln	Ala	Glu	Arg	Glu	Lys	Leu	Ala	Lys	Glu	Lys	Ser	Asn	Thr	Ser
		100						105					110		
Glu	Thr	Thr	Ile	Asp	Asn	Arg	Thr	Thr	Tyr	Thr	Gln	Val	Ala	Phe	Leu
	115						120					125			
Ala	Lys	Val	Ile	Thr	Gly	Ser	Leu	Gln	Lys	Thr	Thr	Pro	Pro	Thr	Asn
	130					135					140				
Phe	Pro	Lys	His	Lys	Glu	Ala	Phe	Phe	Lys	Gly	Leu	Asp	Tyr	Leu	Leu
145					150					155					160
Ala	Ser	Gln	Tyr	Glu	Ser	Gly	Gly	Phe	Pro	Gln	Phe	Tyr	Pro	Leu	Lys
			165						170					175	
Lys	Gly	Tyr	Tyr	Thr	His	Ile	Thr	Phe	Asn	Asp	Asp	Ala	Met	Ile	Gly
			180					185					190		
Val	Leu	Lys	Val	Leu	Arg	Glu	Ile	Ala	Lys	Lys	Lys	Glu	Asp	Tyr	Leu
	195						200					205			
Phe	Val	Asp	Glu	Glu	Arg	Arg	Leu	Lys	Ala	Glu	Lys	Ser	Val	Glu	Lys
	210					215					220				
Ala	Leu	Pro	Leu	Ile	Leu	Lys	Leu	Gln	Val	Glu	Val	Gly	Gly	Lys	Lys
225					230					235					240
Thr	Val	Trp	Ala	Ala	Gln	Tyr	Asp	Glu	Asn	Thr	Phe	Lys	Pro	Ala	Ala
			245						250					255	
Ala	Arg	Lys	Phe	Glu	Pro	Val	Ser	Leu	Thr	Ala	Gly	Glu	Ser	Val	Gly
		260						265					270		
Ile	Val	Arg	Phe	Leu	Met	Tyr	Asp	Ser	Lys	Pro	Asp	Gln	Ala	Thr	Ile
	275						280					285			
Asp	Ala	Ile	Glu	Ser	Ala	Ile	Gln	Trp	Tyr	Arg	Ala	Asn	Lys	Ile	Glu
	290				295						300				
Gly	Ile	Arg	Trp	Val	Arg	Glu	Asn	Gly	Glu	Asn	Arg	Val	Val	Lys	Asp
305					310					315					320
Lys	Asn	Ala	Pro	Pro	Ile	Trp	Ala	Arg	Phe	Tyr	Glu	Ile	Glu	Thr	Met
			325						330					335	
Lys	Pro	Ile	Phe	Ile	Gly	Arg	Asp	Ala	Ile	Ile	Arg	Tyr	Asp	Val	Ser
		340						345					350		
Glu	Ile	Glu	Ala	Glu	Arg	Arg	Asn	Gly	Tyr	Ala	Trp	Tyr	Val	Ser	Glu
	355						360					365			
Pro	Asn	Glu	Leu	Leu	Glu	Lys	Asp	Tyr	Pro	Lys	Trp	Leu	Glu	Lys	Ile
	370					375					380				
Lys	Lys	Ser	Val	Lys											
385															

<210> 59
 <211> 1080
 <212> ADN
 <213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

<400> 59

```

atgagaatcc ggtcctcttc aatcgcggtc ggccctgattt gcagtctggc gctaagggtg      60
cctgcgcaag cgcaggtcac cgtgcgctgg ggggacgtcc tgaaccagcc cgccgcctgg      120
tatggcaccg atgaagcccg tcgaattgcc gaccacgtgc tcgagcatca acgagcggaa      180
ggcggatggc caaagaacac ggacatgacc gcagcgcccg atccggcggt gctcacagcc      240
gcgcgagtga agccagactc gacgatcgat aacggcgcgga cggtcactga aatgcgcgtc      300
ctcgcgcgcg tctaccgttc atcacccgat ccccgttatc gcgatgcgct gctcaagggt      360
ctcgactatc tgttggcagc gcagtatgcc aacggcggtt ggccgcagtt ctaccgcgtc      420
cggcaggact attcgcgcta tatcacgttc aacgacaacg cgatgatcaa tgcgtgacg      480
ctgctctcag acgtcgctgc cggaaatggc gactggcggt ttgctgatgc cagccggcgc      540
gagaaaagcc ggacggctgt agagaaggcc gtagaagtca tcctgcgcgc gcaggtgaga      600
gttgacggcc ggctgaccgc gtggtgcgcc caacacgacg aggtgacact cggagccgcgc      660
aaggcccgcg cctacgaaca tccgtcgctg agcggacagg agacggtggg gatcatccgg      720
tttctcatga cccgcgataa accggatcag agagtcgtcg atgcaatcga ggcgtcagtg      780
gcatggctga aggcggtgca gctcaaagga cttcgcgctc accagcgccg cgatccctcg      840
ctgccggagg ggcgtgacgt ggtgaccgtc gctgaccgtt cggcgccgcc gctctgggcg      900
cgcttctacg aaatcgggac caatcgcccg atcttctctg gacgcgacgg cgtgatccga      960
tactcgctgg cagagatcga gcacgaacgc cggatagggt acgcctggct cgggaacctg      1020
ccgcggaagc tgctcgatac cgaataccca tcctggcgac ggaactcaaca aaggccgtga      1080

```

<210> 60
 <211> 359
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> SEÑAL
 <222> (1)...(24)
 <221> DOMINIO
 <222> (25)...(359)
 <223> Dominio catalítico

<400> 60

```

Met Arg Ile Arg Ser Ser Ser Ile Ala Phe Gly Leu Ile Cys Ser Leu
 1      5      10      15
Ala Leu Arg Val Pro Ala Gln Ala Gln Val Thr Val Arg Trp Ala Asp
 20      25      30
Val Leu Asn Gln Pro Ala Ala Trp Tyr Gly Thr Asp Glu Ala Arg Arg
 35      40      45
Ile Ala Asp His Val Leu Glu His Gln Arg Ala Glu Gly Gly Trp Pro
 50      55      60
Lys Asn Thr Asp Met Thr Ala Ala Pro Asp Pro Ala Val Leu Thr Ala
 65      70      75      80
Ala Arg Val Lys Pro Asp Ser Thr Ile Asp Asn Gly Ala Thr Val Thr
 85      90      95
Glu Met Arg Val Leu Ala Arg Val Tyr Arg Ser Ser Pro Asp Pro Arg
100      105      110
Tyr Arg Asp Ala Leu Leu Lys Gly Leu Asp Tyr Leu Leu Ala Ala Gln
115      120      125
Tyr Ala Asn Gly Gly Trp Pro Gln Phe Tyr Pro Leu Arg Gln Asp Tyr
130      135      140
Ser Arg Tyr Ile Thr Phe Asn Asp Asn Ala Met Ile Asn Val Val Thr
145      150      155      160
Leu Leu Ser Asp Val Ala Ala Gly Asn Gly Asp Trp Ala Phe Ala Asp
165      170      175
Ala Ser Arg Arg Glu Lys Ser Arg Thr Ala Val Glu Lys Ala Val Glu
180      185      190
Val Ile Leu Arg Ala Gln Val Arg Val Asp Gly Arg Leu Thr Ala Trp
195      200      205
Cys Ala Gln His Asp Glu Val Thr Leu Glu Pro Arg Lys Ala Arg Ala
210      215      220
Tyr Glu His Pro Ser Leu Ser Gly Gln Glu Thr Val Gly Ile Ile Arg
225      230      235      240
Phe Leu Met Thr Arg Asp Lys Pro Asp Gln Arg Val Val Asp Ala Ile
245      250      255
Glu Ala Ser Val Ala Trp Leu Lys Ala Val Gln Leu Lys Gly Leu Arg
260      265      270
Val Asp Gln Arg Arg Asp Pro Ser Leu Pro Glu Gly Arg Asp Val Val
275      280      285
Thr Val Ala Asp Pro Ser Ala Pro Pro Leu Trp Ala Arg Phe Tyr Glu
290      295      300
Ile Gly Thr Asn Arg Pro Ile Phe Ser Gly Arg Asp Gly Val Ile Arg
305      310      315      320
Tyr Ser Leu Ala Glu Ile Glu His Glu Arg Arg Ile Gly Tyr Ala Trp
325      330      335
Leu Gly Thr Trp Pro Ala Lys Leu Leu Asp Thr Glu Tyr Pro Ser Trp
340      345      350
Arg Arg Thr Gln Gln Arg Pro
355

```

<210> 61

<211> 1224

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 61

```

gtggaattac cagtaaccgg cgcattgggca acctggcaaa ccgcaactgt tgaaattgat      60
ttggtgcaag gtaacaacct gttaaaactt tctgcgatca cggctgatgg tttggcaaat      120
atcgattcgt tgaaaaattga cggcgcacaa accaaagccg gcgtgtgcag cactgtggca      180
agcagcagct cttcatccgt tgcttcatcg attaaatcaa gctccagttc atcctcttcc      240
agttcaacga cgacggtaaa aacattaaca ctggatggca accccgcagc aaactggttt      300
aacaatcca gaaccaaatg gaataccagc agagccgatg ttgtactttc ctatcaacaa      360
tccaacggcg gctggccaaa aaatctcgat tacaattcag taagcgcagg taatggcggc      420
agcgatagcg gcaccatcga taacgggtgca accattactg aaatggttta tctcgcggaa      480
gtttacaaaa atggcaacaa caccaagtat cgcgatgcgg tgcgcagagc cgcaaatttt      540
attgtcagct cgcaatacag cactgggtgca ttaccacaat tttatccatt gaaaggcggc      600
tatgcagacc acgccacctt taacgataac ggcatggcat atgcattaac ggtattggat      660
tttgcagtca acaaacgcgc cccatttgat actgatgttt tctccgattc tgatcgcgcg      720
aaattcaaaa ccgctgttgc caaagggtgtg gattacattt tgaaagcgca gtggaaacaa      780
aacggaaaat taaccgtgtg gtgtgcacaa catggtgcta ccgattatca accgaaaaaa      840
gcgcgcgcct atgaattgga atcactgagt ggcagcgaat ctgttggtgt actcgttttc      900
ttgatgacct aaccgcaaac cgcacaaatt gaagccgctg taaaagccgg tgtagcctgg      960
ttcaatagcc ccaacacgta tttgaacaat tacaettacg actcttcaaa agcgtcaact      1020
aatccaatag ttgccaagtc tggaagcaaa atgtggtatc gcttttacga tttaaatacc      1080
aatcgtggct tcttcagtga tcgcgatggc agcaaatctc atgacatcac ccagatgtca      1140
gaagagcgtc gcactggata tagctggggt ggtgattacg gcacgtcgat tatttccttc      1200
gcgcaaaaag tgggatatct gtaa                                     1224

```

<210> 62
 <211> 407.
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental
 <221> DOMINIO
 <222> (0)...(407)
 <223> Dominio catalítico

<400> 62

Met	Glu	Leu	Pro	Val	Thr	Gly	Ala	Trp	Ala	Thr	Trp	Gln	Thr	Ala	Thr
1				5					10					15	
Val	Glu	Ile	Asp	Leu	Val	Gln	Gly	Asn	Asn	Leu	Leu	Lys	Leu	Ser	Ala
			20					25					30		
Ile	Thr	Ala	Asp	Gly	Leu	Ala	Asn	Ile	Asp	Ser	Leu	Lys	Ile	Asp	Gly
		35				40						45			
Ala	Gln	Thr	Lys	Ala	Gly	Val	Cys	Ser	Thr	Val	Ala	Ser	Ser	Ser	Ser
	50				55					60					
Ser	Ser	Val	Ala	Ser	Ser	Ile	Lys	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
65					70					75					80
Ser	Ser	Thr	Thr	Thr	Val	Lys	Thr	Leu	Thr	Leu	Asp	Gly	Asn	Pro	Ala
			85					90						95	
Ala	Asn	Trp	Phe	Asn	Lys	Ser	Arg	Thr	Lys	Trp	Asn	Thr	Ser	Arg	Ala
			100					105					110		
Asp	Val	Val	Leu	Ser	Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro	Lys	Asn
	115					120						125			
Leu	Asp	Tyr	Asn	Ser	Val	Ser	Ala	Gly	Asn	Gly	Gly	Ser	Asp	Ser	Gly
	130				135					140					
Thr	Ile	Asp	Asn	Gly	Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu	Ala	Glu
145				150						155					160
Val	Tyr	Lys	Asn	Gly	Asn	Asn	Thr	Lys	Tyr	Arg	Asp	Ala	Val	Arg	Arg
			165					170					175		
Ala	Ala	Asn	Phe	Ile	Val	Ser	Ser	Gln	Tyr	Ser	Thr	Gly	Ala	Leu	Pro

180								185					190				
Gln	Phe	Tyr	Pro	Leu	Lys	Gly	Gly	Tyr	Ala	Asp	His	Ala	Thr	Phe	Asn		
		195					200					205					
Asp	Asn	Gly	Met	Ala	Tyr	Ala	Leu	Thr	Val	Leu	Asp	Phe	Ala	Val	Asn		
		210				215					220						
Lys	Arg	Ala	Pro	Phe	Asp	Thr	Asp	Val	Phe	Ser	Asp	Ser	Asp	Arg	Ala		
225					230				235						240		
Lys	Phe	Lys	Thr	Ala	Val	Ala	Lys	Gly	Val	Asp	Tyr	Ile	Leu	Lys	Ala		
				245					250					255			
Gln	Trp	Lys	Gln	Asn	Gly	Lys	Leu	Thr	Val	Trp	Cys	Ala	Gln	His	Gly		
			260						265				270				
Ala	Thr	Asp	Tyr	Gln	Pro	Lys	Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Glu	Ser		
		275					280					285					
Leu	Ser	Gly	Ser	Glu	Ser	Val	Gly	Val	Leu	Ala	Phe	Leu	Met	Thr	Gln		
		290				295					300						
Pro	Gln	Thr	Ala	Gln	Ile	Glu	Ala	Ala	Val	Lys	Ala	Gly	Val	Ala	Trp		
305					310					315					320		
Phe	Asn	Ser	Pro	Asn	Thr	Tyr	Leu	Asn	Asn	Tyr	Thr	Tyr	Asp	Ser	Ser		
				325					330					335			
Lys	Ala	Ser	Thr	Asn	Pro	Ile	Val	Ala	Lys	Ser	Gly	Ser	Lys	Met	Trp		
			340						345				350				
Tyr	Arg	Phe	Tyr	Asp	Leu	Asn	Thr	Asn	Arg	Gly	Phe	Phe	Ser	Asp	Arg		
		355					360					365					
Asp	Gly	Ser	Lys	Phe	Tyr	Asp	Ile	Thr	Gln	Met	Ser	Glu	Glu	Arg	Arg		
		370				375					380						
Thr	Gly	Tyr	Ser	Trp	Gly	Gly	Asp	Tyr	Gly	Thr	Ser	Ile	Ile	Ser	Phe		
385					390					395					400		
Ala	Gln	Lys	Val	Gly	Tyr	Leu											
				405													

<210> 63
<211> 1023
<212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 63						
atgttaagtt	tcatcgcggt	atcagtggtt	cataattact	gtacagggca	gactgcgtcc	60
acaaaaaatt	cagtggccga	aaagatgctt	cagtaccagt	tgtcaaattg	cgcctggccc	120
aaacagtttg	tagacaaaag	tgtcgttgat	tacagtcttc	cattaacgaa	agagctccta	180
cagcagatca	agaaaacaga	tattgatcat	gctacgctcg	acaacagtgc	gacaacccgg	240
gaaataactg	aattgatcaa	ggcttttaag	gacactaaaa	ataaggcata	tttgactgct	300
gcagaaaagg	ggattgcata	tattttatcg	gctcaatatg	agaattggcgg	atttccacaa	360
tactacccaa	ataaattata	ctatagagct	gagataacat	acaacgatga	tgcgatgatc	420
aatgcattac	tagtgcttta	caaagtagcc	aataagcgag	agggggttga	ggctatcaat	480
cccatatttg	tgtcaaaaagc	gcaaaaagca	gttgaaaagg	gtataacctg	tatcctaata	540
acacaggtca	tacaagacgg	aaaaaggagt	atttgggctg	cgcaatacga	tcagaacact	600
ttacaacctg	ctcaggcaag	aaagtttgaa	ccagcttcat	tgagcacaa	tgaatctggt	660
tccatcggtc	gctttctcat	gctacagcct	gcaaccactg	aaattaagca	agcgatcgaa	720
catgcaatac	aatggttcga	acagcatgat	attgaagggt	accggttcga	ccgcatacaa	780
gatagggtga	ctggaaaata	tcaacggcag	cttgtoctcg	atcggacttc	cacgatttgg	840
gcgcgatttt	ataatctcga	agacaaccgt	ccattgtttg	gagatcgggg	caatacaatc	900
aaatacaact	ttgaggaggt	ttcagaggag	cgtagaaatg	gctatgcttg	gttcggcaac	960
tgcccgaaa	agctgatcca	aaaggactat	ccaaaatgga	aaaaacaata	caaaattcaa	1020
taa						1023

<210> 64
<211> 340
<212> PRT
<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(16)

5 <221> DOMINIO

<222> (17)...(340)

<223> Dominio catalítico

<400> 64

10

Met	Leu	Ser	Phe	Ile	Ala	Val	Ser	Val	Phe	His	Asn	Tyr	Cys	Thr	Gly
1				5					10					15	
Gln	Thr	Ala	Ser	Thr	Lys	Asn	Ser	Val	Ala	Glu	Lys	Met	Leu	Gln	Tyr
		20						25					30		
Gln	Leu	Ser	Asn	Gly	Ala	Trp	Pro	Lys	Gln	Leu	Val	Asp	Lys	Ser	Val
	35						40					45			
Val	Asp	Tyr	Ser	Leu	Pro	Leu	Thr	Lys	Glu	Leu	Leu	Gln	Gln	Ile	Lys
	50					55						60			
Lys	Thr	Asp	Ile	Asp	His	Ala	Thr	Leu	Asp	Asn	Ser	Ala	Thr	Thr	Arg
65					70					75					80
Glu	Ile	Thr	Glu	Leu	Ile	Lys	Ala	Phe	Lys	Asp	Thr	Lys	Asn	Lys	Ala
			85						90					95	
Tyr	Leu	Thr	Ala	Ala	Glu	Lys	Gly	Ile	Ala	Tyr	Ile	Leu	Ser	Ala	Gln
			100					105					110		
Tyr	Glu	Asn	Gly	Gly	Phe	Pro	Gln	Tyr	Tyr	Pro	Asn	Lys	Leu	Tyr	Tyr
		115					120					125			
Arg	Ala	Glu	Ile	Thr	Tyr	Asn	Asp	Asp	Ala	Met	Ile	Asn	Ala	Leu	Leu
	130					135					140				
Val	Leu	Tyr	Lys	Val	Ala	Asn	Lys	Arg	Glu	Gly	Phe	Glu	Ala	Ile	Asn
145					150					155					160
Pro	Ile	Phe	Val	Ser	Lys	Ala	Gln	Lys	Ala	Val	Glu	Lys	Gly	Ile	Thr
			165						170					175	
Cys	Ile	Leu	Lys	Thr	Gln	Val	Ile	Gln	Asp	Gly	Lys	Arg	Ser	Ile	Trp
		180						185					190		
Ala	Ala	Gln	Tyr	Asp	Gln	Asn	Thr	Leu	Gln	Pro	Ala	Gln	Ala	Arg	Lys
		195					200					205			
Phe	Glu	Pro	Ala	Ser	Leu	Ser	Thr	Ser	Glu	Ser	Val	Ser	Ile	Val	Arg
	210					215					220				
Phe	Leu	Met	Leu	Gln	Pro	Ala	Thr	Thr	Glu	Ile	Lys	Gln	Ala	Ile	Glu
225					230					235					240
His	Ala	Ile	Gln	Trp	Phe	Glu	Gln	His	Asp	Ile	Glu	Gly	Tyr	Arg	Phe
			245						250					255	
Asp	Arg	Ile	Gln	Asp	Arg	Val	Thr	Gly	Lys	Tyr	Gln	Arg	Gln	Leu	Val
		260						265					270		
Ala	Asp	Arg	Thr	Ser	Thr	Ile	Trp	Ala	Arg	Phe	Tyr	Asn	Leu	Glu	Asp
		275					280					285			
Asn	Arg	Pro	Leu	Phe	Gly	Asp	Arg	Asp	Asn	Thr	Ile	Lys	Tyr	Asn	Phe
	290					295					300				
Glu	Glu	Val	Ser	Glu	Glu	Arg	Arg	Asn	Gly	Tyr	Ala	Trp	Phe	Gly	Asn
305					310					315					320
Trp	Pro	Glu	Lys	Leu	Ile	Gln	Lys	Asp	Tyr	Pro	Lys	Trp	Lys	Lys	Gln
			325						330					335	
Tyr	Lys	Ile	Gln												
			340												

<210> 65

<211> 1311

15 <212> ADN

<213> Bacteria

<400> 65

```

gtgaaccgac gtaccgcgct gggagcggtc gccgcgaccg ccctcgccct gacggtcacc 60
gccccgcgcg ccggtgccc cgcgcgcgct ccccaagcgg cgccacgccc ggtcgccgat 120
ccggctcgcg ccacgctgcc cgcggcgac ggtggggcgt ccgaggggac cggcacgacc 180
ggtggggcgg ccgcccaggg ctcccgggtc ttcaccgctg ccacctggga ggagttccgg 240
gccgcgctcg cgggtgcccg ctccgagccc aggatcgta aggtggtggg cagcgtgaac 300
gccaccgccc ccggtgcccg cgccttcgag gcgcccgggt acgacttcgc ccgctacctc 360
gccgactacg acccggcggt gtgggggtac gagaaggagg tcagcgcccc gcaggaggag 420
ctgcccggcg cgtccgcgac cgcgcagggc caggccatca aggtcaagggt gccggcgaac 480
accacgatcg tcgggggtcg caggcacgcg gggatcacgg gcggcagcct ccagggtgcag 540
ggcgctcgaca acgtcggtgt ccgcaacctg acgctggaga gcccgcctga ctgcttcccg 600
cagtggggac cgaccgacgg cgcgaccggg gcgtggaact ccgagtacga cagcctcgtc 660
gtgtacggct ccacccatgt ctggatcgac cacaacaact tcaccgacgg cgcccacccg 720
gacagttcgc tgcctcgta ctacggcgag gtctaccagc agcacgacgg cgaactggac 780
gtcgtgcggg gcgcccggct cgtcacggtc tctggaacg ccttcaccga ccacgacaag 840
accctgatga tcggcaacag cgacagcgcg ggcgccaccg accggggcaa gctgcgggtc 900
accctgcacc acaacctgtt cgagaacgtc gtgcagcggg cggcccgggt caggttcggg 960
caggtcgacg cgtacaacaa ccacttcgtc gtgcccagtt cggcctacgc gtacagcctg 1020
ggcgtcggcg aggagtcoca gctcttcgcg gagaagaacg cgttcacctt cgccgggggc 1080
gtgcccggcg ggaagatcct caagaagtgg aaggacgccc ccgtcaccac cgtcgggaac 1140
tacgtgaacg gcaggcgggt cgacctgtc gcgttcaca acaccagtt cccggaggag 1200
cagttgcggg ccgacgcccg ctggacccc gtctgcgca ccagggtcga ccacccgagg 1260
gcgtcccccg cgctcgtcga ccaccgcgcg ggcgcggccc gctcctgctg a 1311

```

<210> 66

<211> 436

5 <212> PRT

<213> Bacteria

<220>

<221> SEÑAL

10 <222> (1)...(28)

<221> DOMINIO

<222> (29)...(436)

<223> Dominio catalítico

15

<400> 66

```

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

```


Gly Tyr Asp Phe Ala Arg Tyr Leu Ala Asp Tyr Asp Pro Ala Val Trp
115 120 125
Gly Tyr Glu Lys Glu Val Ser Gly Pro Gln Glu Glu Leu Arg Ala Ala
130 135 140
Ser Ala Thr Ala Gln Gly Gln Ala Ile Lys Val Lys Val Pro Ala Asn
145 150 155 160
Thr Thr Ile Val Gly Val Gly Arg His Ala Gly Ile Thr Gly Gly Ser
165 170 175
Leu Gln Val Gln Gly Val Asp Asn Val Val Arg Asn Leu Thr Leu
180 185 190
Glu Ser Pro Leu Asp Cys Phe Pro Gln Trp Asp Pro Thr Asp Gly Ala
195 200 205
Thr Gly Ala Trp Asn Ser Glu Tyr Asp Ser Leu Val Val Tyr Gly Ser
210 215 220
Thr His Val Trp Ile Asp His Asn Thr Phe Thr Asp Gly Ala His Pro
225 230 235 240
Asp Ser Ser Leu Pro Ser Tyr Tyr Gly Glu Val Tyr Gln Gln His Asp
245 250 255
Gly Glu Leu Asp Val Val Arg Gly Ala Asp Leu Val Thr Val Ser Trp
260 265 270
Asn Ala Phe Thr Asp His Asp Lys Thr Leu Met Ile Gly Asn Ser Asp
275 280 285
Ser Ala Gly Ala Thr Asp Arg Gly Lys Leu Arg Val Thr Leu His His
290 295 300
Asn Leu Phe Glu Asn Val Val Glu Arg Ala Pro Arg Val Arg Phe Gly
305 310 315 320
Gln Val Asp Ala Tyr Asn Asn His Phe Val Val Pro Ser Ser Ala Tyr
325 330 335
Ala Tyr Ser Leu Gly Val Gly Gln Glu Ser Gln Leu Phe Ala Glu Lys
340 345 350
Asn Ala Phe Thr Leu Ala Gly Gly Val Pro Ala Gly Lys Ile Leu Lys
355 360 365
Lys Trp Lys Asp Ala Pro Val Thr Thr Val Gly Asn Tyr Val Asn Gly
370 375 380
Arg Pro Val Asp Leu Leu Ala Val His Asn Thr Gln Phe Pro Glu Glu
385 390 395 400
Gln Leu Arg Ala Asp Ala Gly Trp Thr Pro Val Leu Arg Thr Arg Val
405 410 415
Asp His Pro Arg Ala Val Pro Ala Leu Val Asp His Arg Ala Gly Ala
420 425 430
Gly Arg Ser Cys
435

<210> 67
<211> 1995
5 <212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 67

atgaaaaatt caaaaactgt ttttactgca caaaaaaac tcatgcactc ttgcattgcc 60
gccgctatcg gcttggcgat aagttcaggt gcttgggtcag cttgtactta cactgtcacc 120
aataattggg gttctggctt caccggtgaa atcaaagtta ccaacaacac atcatcggct 180
gtaaattggtt ggtctgtgtc ttggcaggaa tcaggcgcat cagtcaacaa ctcatggaac 240
gcaactctga gcggatcaaa tccttatacg gcagccgcct taggttgga tgcaactctc 300
gcacccggtg cttctgccag ttttggtctt caagcaaagt gcactgctag cgcacctaaa 360
gtgaatggca ctttatgtgg aacagctact tcatcaacac ctgcgtcatc cagcagtgtt 420

```

.gcgagttcgg ttaaatacaag cgcacccggt tcgtccagca gtaaatcatc cagctcaatc 480
actgtgagta gtagttctat cgccagcagc agcgcaacca gtgtttcttc attacaatt 540
caggaagagc aagctggctt ctgtcgtgtt gatggcattg caacagaaag caccaacacc 600
ggctttaccg gcaacggcta taccatgca aacaacgcac aagggtgcagc gattgaatgg 660
gcggtaaatg caccgagcag tggcgcgtac acactcacat tccgttttgc aaatggcggc 720
actgcagcac gcaatgggtc actgttaatt aaogcgcggt gcaatggtaa ctacaccgtg 780
gatttaccac taaccggcgc atgggcgact tggcaaacag cgactgtaga aatcgatttg 840
gtacaaggca ccaacacgct gaaactttct gcattaaccg cagatggctt agctaatac 900
gattcattaa aaattgatgg caaccaaccg aaagcaggca cttgcagcaa tacatcaagc 960
agtgttgcca gcagttcttc atccgttaaa tccagttcaa gttcttcac aagctcatcc 1020
accactgcaa aaatgctgac tcttgatggc aaccccgccg caagttgggt caacaaatcc 1080
agaaccaa at ggaatagcag tcgtgcggac attgtgttgt cttaccagca agctaacggc 1140
ggctggccaa aaaatctgga ttacaactca gtaagcgcag gtaatggtgg tagcgacagc 1200
ggcactatcg acaacggcgc aaccatcacg gaaatgggtt atctcgaga agtttataaa 1260
aacgggggca atacaaaata tcgcgatgca gtacgtaaag cggcaaactt tattgtgagt 1320
tcgcaatata gcactggtgc gttaccacaa ttttaccat tgaaagggtg ttatgcagat 1380
cacgccacct ttaacgataa cggcatggct tacgcattaa cggatttga ttttgcggtg 1440
aacaacggtg cgccgtttga taacgatgta tttctgatg cagaccgcgc aaaattcaaa 1500
actgcggtga ccaaaggtat tgattacatt ttgaaagccc aatggaaaca aaatggaaaa 1560
ctcaccgcat ggtgtgcgca acacggagca aacgactatc aacaaaagc ggcgcgtgct 1620
tatgagttag tatctttaag cggcagcgaa tcggttggca tcatcgctt cctgatgacc 1680
caaccacaaa ctgcgcaaat cgaagcagcg gttaaagccg gtgtaaaactg gttcgctagc 1740
ccgaatacat acttggctaa ttacacctac gactcgtcaa aagcctctac caatccgatt 1800
gtgtacaaat ccggcagcag aatgtgggtat cgcttctacg atctgaacac caatcgcgga 1860
ttcttttagtg atcgcgatgg cagcaaattc tatgacatca ctcaaatgtc tgaagaacgt 1920
cgcaccggct acagctgggg cggttcttac ggtgaatcga ttatcagctt cgcgcaaaaa 1980
gtgggttata tctaa 1995

```

<210> 68
 <211> 664
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(32)

<221> UNIÓN
 15 <222> (33)...(126)
 <223> Módulo de unión a carbohidratos

<221> UNIÓN
 20 <222> (184)...(307)
 <223> Módulo de unión a carbohidratos

<221> DOMINIO
 <222> (308)...(664)
 <223> Dominio catalítico

<400> 68

```

Met Lys Asn Ser Lys Thr Val Phe Thr Ala Gln Lys Lys Leu Met His
  1      5      10      15
Ser Cys Ile Ala Ala Ala Ile Gly Leu Ala Ile Ser Ser Gly Ala Trp
      20      25      30
Ser Ala Cys Thr Tyr Thr Val Thr Asn Asn Trp Gly Ser Gly Phe Thr
      35      40      45
Gly Glu Ile Lys Val Thr Asn Asn Thr Ser Ser Ala Val Asn Gly Trp

```

50	55	60
Ser Val Ser Trp Gln Glu Ser Gly Ala Ser Val Thr Asn Ser Trp Asn		
65	70	75
Ala Thr Leu Ser Gly Ser Asn Pro Tyr Thr Ala Ala Ala Leu Gly Trp		80
	85	90
Asn Ala Thr Leu Ala Pro Gly Ala Ser Ala Ser Phe Gly Phe Gln Ala		95
	100	105
Asn Gly Thr Ala Ser Ala Pro Lys Val Asn Gly Thr Leu Cys Gly Thr		110
	115	120
Ala Thr Ser Ser Thr Pro Ala Ser Ser Ser Ser Val Ala Ser Ser Val		125
	130	135
Lys Ser Ser Ala Pro Val Ser Ser Ser Ser Lys Ser Ser Ser Ser Ile		140
145	150	155
Thr Val Ser Ser Ser Ser Ile Ala Ser Ser Ser Ala Pro Ser Val Ser		160
	165	170
Ser Leu Thr Ile Gln Glu Glu Gln Ala Gly Phe Cys Arg Val Asp Gly		175
	180	185
Ile Ala Thr Glu Ser Thr Asn Thr Gly Phe Thr Gly Asn Gly Tyr Thr		190
	195	200
Asn Ala Asn Asn Ala Gln Gly Ala Ala Ile Glu Trp Ala Val Asn Ala		205
	210	215
Pro Ser Ser Gly Arg Tyr Thr Leu Thr Phe Arg Phe Ala Asn Gly Gly		220
225	230	235
Thr Ala Ala Arg Asn Gly Ser Leu Leu Ile Asn Gly Gly Ser Asn Gly		240
	245	250
Asn Tyr Thr Val Asp Leu Pro Leu Thr Gly Ala Trp Ala Thr Trp Gln		255
	260	265
Thr Ala Thr Val Glu Ile Asp Leu Val Gln Gly Thr Asn Thr Leu Lys		270
	275	280
Leu Ser Ala Leu Thr Ala Asp Gly Leu Ala Asn Ile Asp Ser Leu Lys		285
	290	295
Ile Asp Gly Asn Gln Pro Lys Ala Gly Thr Cys Ser Asn Thr Ser Ser		300
305	310	315
Ser Val Ala Ser Ser Ser Ser Val Lys Ser Ser Ser Ser Ser Ser		320
	325	330
Ser Ser Ser Ser Thr Thr Ala Lys Met Leu Thr Leu Asp Gly Asn Pro		335
	340	345
Ala Ala Ser Trp Phe Asn Lys Ser Arg Thr Lys Trp Asn Ser Ser Arg		350
	355	360
Ala Asp Ile Val Leu Ser Tyr Gln Gln Ala Asn Gly Gly Trp Pro Lys		365
	370	375
Asn Leu Asp Tyr Asn Ser Val Ser Ala Gly Asn Gly Gly Ser Asp Ser		380
385	390	395
Gly Thr Ile Asp Asn Gly Ala Thr Ile Thr Glu Met Val Tyr Leu Ala		400
	405	410
Glu Val Tyr Lys Asn Gly Gly Asn Thr Lys Tyr Arg Asp Ala Val Arg		415
	420	425
Lys Ala Ala Asn Phe Ile Val Ser Ser Gln Tyr Ser Thr Gly Ala Leu		430
	435	440
Pro Gln Phe Tyr Pro Leu Lys Gly Gly Tyr Ala Asp His Ala Thr Phe		445
	450	455
Asn Asp Asn Gly Met Ala Tyr Ala Leu Thr Val Leu Asp Phe Ala Val		460
465	470	475
Asn Lys Arg Ala Pro Phe Asp Asn Asp Val Phe Ser Asp Ala Asp Arg		480
	485	490
Ala Lys Phe Lys Thr Ala Val Thr Lys Gly Ile Asp Tyr Ile Leu Lys		495
	500	505
Ala Gln Trp Lys Gln Asn Gly Lys Leu Thr Ala Trp Cys Ala Gln His		510
	515	520
		525

Gly Ala Asn Asp Tyr Gln Pro Lys Ala Ala Arg Ala Tyr Glu Leu Val
530 535 540
Ser Leu Ser Gly Ser Glu Ser Val Gly Ile Ile Ala Phe Leu Met Thr
545 550 555 560
Gln Pro Gln Thr Ala Gln Ile Glu Ala Ala Val Lys Ala Gly Val Asn
565 570 575
Trp Phe Ala Ser Pro Asn Thr Tyr Leu Ala Asn Tyr Thr Tyr Asp Ser
580 585 590
Ser Lys Ala Ser Thr Asn Pro Ile Val Tyr Lys Ser Gly Ser Arg Met
595 600 605
Trp Tyr Arg Phe Tyr Asp Leu Asn Thr Asn Arg Gly Phe Phe Ser Asp
610 615 620
Arg Asp Gly Ser Lys Phe Tyr Asp Ile Thr Gln Met Ser Glu Glu Arg
625 630 635 640
Arg Thr Gly Tyr Ser Trp Gly Gly Ser Tyr Gly Glu Ser Ile Ile Ser
645 650 655
Phe Ala Gln Lys Val Gly Tyr Leu
660

<210> 69
<211> 1035
5 <212> ADN
<213> Desconocido

<220>
10 <223> Obtenido de una muestra ambiental
<400> 69

atggcgcggtt	tggtccggtg	cgtgtgtgcc	agcctgggag	gatgggcgcg	ggttctggcc	60
gccgcggcgg	gcccgattg	gtccgcctg	ctcgcgcaac	cggacccttg	gtttcgcagt	120
ccggcggggc	aacaggcgg	gacgaacgtt	ttgtcctggc	agagcgcgac	aggcgcctgg	180
ccgaaaaacc	tggacaccac	ccgcgagccg	cgtcggcagg	attccgcccc	gcccaggggc	240
actttcgaca	acggcgccac	caccggcgag	ttgcggtttc	tggcgcgggc	gtttgcggcc	300
accggcgatc	cgcgctgcga	agccgcggtg	ctccgggggc	tggacggcat	cctcgcggcc	360
cagcttccca	gcggcggtg	gccgcagtgt	catectccgc	gcgcgcctta	tcagcgcacc	420
atcaccttca	acgacggtgt	catggtgcgc	atcctggagc	tgctgcgcga	gatagaccgc	480
gcgcgggagt	ttcgttgggt	ggacgaggcg	cggcgcgcg	gggtgcgcgc	ggccttcact	540
cgcgggcttg	agtgcctcct	gcgctgccag	gtggtcgctc	agggcagact	caccgtgtgg	600
tgtgccagc	atgacgcgga	gaactttcaa	ccgcgaccgg	cacgcgccta	cgaactggaa	660
tcgctcagcg	gcgcggaaag	cgccggcatc	ctggtgttcc	tcattgagcct	ggagccgcca	720
accccgga	tcgcgcgcgc	ggtcgaggcc	ggggcgccct	ggttttcggc	ggtaaagctt	780
gaagggttcc	gtctcgaacg	aacggccgac	gacgcgcggg	tgggtggaaga	gccgggcgcg	840
ccgccgctct	gggcgcgggt	ctacgagatc	gggaccaatc	gccccatctt	tgccggctgc	900
gacggtgtca	agaagtaacg	cctgagcgag	atcgagcggg	aacgcccggg	cggctatgcg	960
tggtacggcg	cctgggggtga	accggtcgcc	cgcattatg	cccagtggcg	ggagcggtac	1020
gggacgcgca	aatga					1035

15 <210> 70
<211> 344
<212> PRT
<213> Desconocido

20 <220>
<223> Obtenido de una muestra ambiental

<221> SEÑAL
<222> (1)...(22)

25 <221> DOMINIO
<222> (23)...(344)
<223> Dominio catalítico

<400> 70

```

Met Ala Arg Leu Phe Arg Cys Val Cys Ala Ser Leu Gly Gly Trp Ala
 1           5           10           15
Ala Val Leu Ala Ala Ala Ala Gly Pro Asp Trp Ser Arg Leu Leu Ala
      20           25           30
Gln Pro Asp Pro Trp Phe Arg Ser Pro Ala Gly Gln Gln Ala Val Thr
      35           40           45
Asn Val Leu Ser Trp Gln Ser Ala Thr Gly Ala Trp Pro Lys Asn Leu
      50           55           60
Asp Thr Thr Arg Glu Pro Arg Arg Gln Asp Ser Ala Pro Pro Glu Gly
      65           70           75           80
Thr Phe Asp Asn Gly Ala Thr Thr Gly Glu Leu Arg Phe Leu Ala Arg
      85           90           95
Ala Phe Ala Ala Thr Gly Asp Pro Arg Cys Glu Ala Ala Val Leu Arg
      100           105           110
Gly Leu Asp Gly Ile Leu Ala Ala Gln Leu Pro Ser Gly Gly Trp Pro
      115           120           125
Gln Cys His Pro Pro Arg Ala Pro Tyr Gln Arg His Ile Thr Phe Asn
      130           135           140
Asp Gly Val Met Val Arg Ile Leu Glu Leu Leu Arg Glu Ile Asp Arg
      145           150           155           160
Ala Pro Glu Phe Arg Trp Val Asp Glu Ala Arg Arg Ala Arg Val Arg
      165           170           175
Ala Ala Phe Thr Arg Gly Leu Glu Cys Leu Leu Arg Cys Gln Val Val
      180           185           190
Val Glu Gly Arg Leu Thr Val Trp Cys Ala Gln His Asp Ala Glu Asn
      195           200           205
Phe Gln Pro Arg Pro Ala Arg Ala Tyr Glu Leu Glu Ser Leu Ser Gly
      210           215           220
Ala Glu Ser Ala Gly Ile Leu Val Phe Leu Met Ser Leu Glu Pro Pro
      225           230           235           240
Thr Pro Glu Ile Ala Arg Ala Val Glu Ala Gly Ala Ala Trp Phe Ser
      245           250           255
Ala Val Lys Leu Glu Gly Phe Arg Leu Glu Arg Thr Ala Asp Asp Ala
      260           265           270
Arg Val Val Glu Glu Pro Gly Ala Pro Pro Leu Trp Ala Arg Phe Tyr
      275           280           285
Glu Ile Gly Thr Asn Arg Pro Ile Phe Ala Gly Arg Asp Gly Val Lys
      290           295           300
Lys Tyr Ala Leu Ser Glu Ile Glu Arg Glu Arg Arg Val Gly Tyr Ala
      305           310           315           320
Trp Tyr Gly Ala Trp Gly Glu Pro Val Ala Arg His Tyr Ala Gln Trp
      325           330           335
Arg Glu Arg Tyr Gly Thr Gln Lys
      340

```

5 <210> 71
 <211> 1038
 <212> ADN
 <213> Desconocido

10 <220>
 <223> Obtenido de una muestra ambiental

<400> 71
 gtgactcgtg tcgcccttgc gatggggcctt gttgcatggg ttccggcgct cgcttcagct

60

```

gggcccgcgtg catatttgca gaagccggac gactgggttcg ccagtcocga ggccaggggca 120
atcgcccgga acgtactcgc gcâtccaggcc gatctcggcg ggtggccgaa gaacatcgac 180
acaaccgaagc cgttcaccgg cgaccggacg caaatcaaac cgaccttcga taacagcgcg 240
acaaccgaagc agctccgggtt tctggcgcg atccacaacg cgactcgcga cgagaagtac 300
cgccaccggt tcgagaaggg gctcgattac atcttgaaag cacaatacgc aaacggcggt 360
tgcccgcagt cgcaccgcgc cggcaccggc taccaccggc acatcacctt caacgacaat 420
gccatggtcc gtttgatgga gctcgtgcgc gaagtcgcga cctcgaatcg gtacgacttc 480
ctggacggcg accgcccga ggcctgcgc gccgctttcg atcgcgcat cgaatgcac 540
ctgaagtgcc agatcaaggt cgacagtaag ctgacggcat ggtgcgcca gcacgacgag 600
aaggacctcg ctcccggcc ggccgggacc taagagctcg tctcactcag cggctcggag 660
tcggtcggga tctcgcgc actcatgagc ctcgatcgac caagcccgga ggtcgtcgg 720
gccatcgacg gcgcggtcgc gtggttcag tcggcgaagc tcgaaggcac caaggtcgtt 780
gtcgagcgcg acccgaagta tccggggcgc cgggaacgcg tgggtggtgaa ggatccaaag 840
gcaccgccac tctgggcgcg cttctacgaa atcggcacga atcgcccat cttctccgac 900
cgcgacggca tcaagaagta cgcgctcgc gagatcggcc ccgaacggcg gaatggctat 960
gcctggtatg gcacctggcc gcgcgacctg ctggagaagg aatacccagg gtggaaaaag 1020
aagctggccc ggccgtga 1038

```

<210> 72
 <211> 345
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(20)

<221> DOMINIO
 <222> (21)...(345)
 <223> Dominio catalítico

<400> 72

```

Met Thr Arg Val Ala Leu Ala Met Gly Leu Val Ala Trp Val Pro Ala
1      5      10      15
Leu Ala Ser Ala Gly Pro Ala Ala Tyr Leu Gln Lys Pro Asp Asp Trp
20      25      30
Phe Ala Ser Pro Glu Ala Arg Ala Ile Ala Ala Asn Val Leu Ala His
35      40      45
Gln Ala Asp Leu Gly Gly Trp Pro Lys Asn Ile Asp Thr Thr Lys Pro
50      55      60
Phe Thr Gly Asp Arg Thr Gln Ile Lys Pro Thr Phe Asp Asn Ser Ala
65      70      75      80
Thr Thr Asp Glu Leu Arg Phe Leu Ala Arg Ile His Asn Ala Thr Arg
85      90      95
Asp Glu Lys Tyr Arg Thr Ala Phe Glu Lys Gly Leu Asp Tyr Ile Leu
100     105     110
Lys Ala Gln Tyr Ala Asn Gly Gly Trp Pro Gln Ser His Pro Pro Gly
115     120     125
Thr Gly Tyr His Arg His Ile Thr Phe Asn Asp Asn Ala Met Val Arg
130     135     140
Leu Met Glu Leu Val Arg Glu Val Ala Thr Ser Asn Arg Tyr Asp Phe
145     150     155     160
Leu Asp Ala Asp Arg Arg Lys Ala Cys Arg Ala Ala Phe Asp Arg Gly
165     170     175
Ile Glu Cys Ile Leu Lys Cys Gln Ile Lys Val Asp Ser Lys Leu Thr
180     185     190
Ala Trp Cys Ala Gln His Asp Glu Lys Asp Leu Ala Pro Arg Pro Ala

```

195	200	205
Arg Thr Tyr Glu Leu Val	Ser Leu Ser Gly Ser	Glu Ser Val Gly Ile
210	215	220
Val Arg Leu Leu Met Ser	Leu Asp Arg Pro Ser	Pro Glu Val Ala Arg
225	230	235
Ala Ile Asp Gly Ala Val	Ala Trp Phe Gln Ser	Ala Lys Leu Glu Gly
245	250	255
Thr Lys Val Val Val Glu	Arg Asp Pro Lys Tyr	Pro Gly Gly Arg Glu
260	265	270
Arg Val Val Val Lys Asp	Pro Lys Ala Pro Pro	Leu Trp Ala Arg Phe
275	280	285
Tyr Glu Ile Gly Thr Asn	Arg Pro Ile Phe Ser	Asp Arg Asp Gly Ile
290	295	300
Lys Lys Tyr Ala Leu Ala	Glu Ile Gly Pro Glu	Arg Arg Asn Gly Tyr
305	310	315
Ala Trp Tyr Gly Thr Trp	Pro Arg Asp Leu Leu	Glu Lys Glu Tyr Pro
325	330	335
Gly Trp Lys Lys Lys Leu	Ala Arg Pro	
340	345	

<210> 73
 <211> 1221
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 73

atgctcacca	aaacatcact	acttattgca	ttgctaggca	gttggttgat	cgcaccatta	60
catgcggaca	caccagcaag	caatgcaccg	acaaccaatg	catcaattcc	gctacagcaa	120
actgcgagcg	atgctgccgc	ctggaaaaat	tatctcgcca	aatccaacga	gttgcgcaaa	180
gcagaccagg	cgagctcaa	agccgagctg	aaaaaactcg	ggcaaaaaac	cgcgagtttg	240
cctgagtaca	ccaaagaatt	tggttttgaa	gtgaagcagt	catctgagtg	gtttaaaagc	300
actgaaggta	aacgagtgat	ggatattatc	ctatcgtttc	aaactccttc	tggcggctgg	360
tcaaaacgca	ctgacatgag	caaagcgccg	cgcaaaccgg	gccaggcatt	tgggtgtgaa	420
aaaaattaca	tccccacctt	tgataatggc	gcgaccagca	cacaattaat	gctactggca	480
caggcgcatc	aagccactgg	cgataaacgc	tacagcgatg	catttgcgcg	cgggcttgaa	540
tttatcatca	ccgctcaata	tcccaatggc	ggctggccac	aaaattttcc	attggttggc	600
aagtatcacg	atcacatcac	ttacaacgat	gccctgatgc	gogatttaat	ggtagtgcta	660
cacaagggtg	ccatggccaa	ggatgaattt	gcctttgtat	ccaaggcgca	gcaacaggcc	720
gcacaagcga	gcctcgaacg	cgcgctggac	tgcgttttga	aaaccaggt	gatggccaat	780
ggccaattaa	ctatatgggg	tgcgcgagcac	gatgccaaaa	ccttaaaacc	cgccaaagcg	840
cgcgcctatg	aaatgatttc	actcaccagt	tctgaaagcg	tgtggatgct	cgatttttta	900
atggatttgc	aacagcccag	cgctgacatt	attaaatccg	tgcacgcggc	tgccgcttgg	960
tatgagcaaa	ataaaattat	cggaaaaacc	tggaccgggg	gcgacacagt	tctgaaagac	1020
gataaggatg	caccgccaat	ctgggcgcg	ttttatgaga	taggtacgaa	caaacccttg	1080
tttggcgacc	gcgatgactc	tgtccattac	gatctggcaa	aggtatcgga	agagcgccgc	1140
acgggttatg	cctggtacac	aacctacccc	aatcaggtat	taaaaaagta	cgcgcgctgg	1200
gctaaacaat	atccgcaata	a				1221

<210> 74
 <211> 406
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(22)

<221> DOMINIO

<222> (23)...(406)
 <223> Dominio catalítico

<400> 74

5

```

Met Leu Thr Lys Thr Ser Leu Leu Ile Ala Leu Leu Gly Ser Cys Cys
 1          5          10          15
Ile Ala Pro Leu His Ala Asp Thr Pro Ala Ser Asn Ala Pro Thr Thr
          20          25          30
Asn Ala Ser Ile Pro Leu Gln Gln Thr Ala Ser Asp Ala Ala Ala Trp
          35          40          45
Lys Asn Tyr Leu Ala Lys Ser Asn Glu Leu Arg Lys Ala Asp Gln Ala
          50          55          60
Gln Leu Lys Ala Glu Leu Lys Lys Leu Gly Gln Lys Thr Ala Ser Leu
65          70          75          80
Pro Glu Tyr Thr Lys Glu Phe Gly Phe Glu Val Lys Gln Ser Ser Glu
          85          90          95
Trp Phe Lys Ser Thr Glu Gly Lys Arg Val Met Asp Ile Ile Leu Ser
          100          105          110
Phe Gln Thr Pro Ser Gly Gly Trp Ser Lys Arg Thr Asp Met Ser Lys
          115          120          125
Ala Pro Arg Lys Pro Gly Gln Ala Phe Gly Val Glu Lys Asn Tyr Ile
          130          135          140
Pro Thr Phe Asp Asn Gly Ala Thr Ser Thr Gln Leu Met Leu Leu Ala
145          150          155          160
Gln Ala His Gln Ala Thr Gly Asp Lys Arg Tyr Ser Asp Ala Phe Ala
          165          170          175
Arg Gly Leu Glu Phe Ile Ile Thr Ala Gln Tyr Pro Asn Gly Gly Trp
          180          185          190
Pro Gln Asn Phe Pro Leu Val Gly Lys Tyr His Asp His Ile Thr Tyr
          195          200          205
Asn Asp Ala Leu Met Arg Asp Leu Met Val Val Leu His Lys Val Ala
210          215          220
Met Ala Lys Asp Glu Phe Ala Phe Val Ser Lys Ala Gln Gln Gln Ala
225          230          235          240
Ala Gln Ala Ser Leu Glu Arg Ala Leu Asp Cys Val Leu Lys Thr Gln
          245          250          255
Val Met Ala Asn Gly Gln Leu Thr Ile Trp Gly Ala Gln His Asp Ala
          260          265          270
Lys Thr Leu Lys Pro Ala Lys Ala Arg Ala Tyr Glu Met Ile Ser Leu
          275          280          285
Thr Ser Ser Glu Ser Val Trp Met Leu Asp Phe Leu Met Asp Leu Gln
290          295          300
Gln Pro Ser Ala Asp Ile Ile Lys Ser Val His Ala Ala Ala Ala Trp
305          310          315          320
Tyr Glu Gln Asn Lys Ile Ile Gly Lys Thr Trp Thr Arg Gly Asp Thr
          325          330          335
Val Leu Lys Asp Asp Lys Asp Ala Pro Pro Ile Trp Ala Arg Phe Tyr
          340          345          350
Glu Ile Gly Thr Asn Lys Pro Leu Phe Gly Asp Arg Asp Asp Ser Val
          355          360          365
His Tyr Asp Leu Ala Lys Val Ser Glu Glu Arg Arg Thr Gly Tyr Ala
          370          375          380
Trp Tyr Thr Thr Ser Pro Asn Gln Val Leu Lys Lys Tyr Ala Arg Trp
385          390          395          400
Ala Lys Gln Tyr Pro Gln

```

405

<210> 75
 <211> 1767
 <212> ADN

10

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

5

<400> 75

```

atgaccacaa cccgccgcac tatcctgaaa gccgccgcca gcgccggcgc gatcgccagc      60
accggctggc ccgccttggc cgccgcacag gccgccgaag ccgccgaccc gtgggcccgc      120
gcccagcaga tcacgcgacc ctccgccaag ccgctcagct tcccgaacag ggacttcccg      180
atcaccgagt tcggcgccaa accctgcaag ctggtcaaag cccagggcct ggtcgaagta      240
agagtcaaag gcgaactcga aacgccagca ccgcaagcgc cggacgccta cccggcaatc      300
aaagccgcca tcgccgcagc gagcaaggcc ggaggagggc gctgctgat cccggccggc      360
aactggtact gcaagggcc tctcgtgctg ctgtcgaacg tgcacgtgca ccttgccaag      420
ggcgcgcaag tctacttcag cgccaacgcc aaggacttcg cccgcgacgg cgactacgac      480
tgccggcgcca acggaagct ggtgctctcg cgctggcaag gcaacgattg cctgaacttc      540
tcgcccattg tctacgcgcg cgggcaaaag aatatcgcca ttaccggcga agactggacc      600
agcatcctga acggccaggc cggcgtggcg ttcgaagacg gcagcggcaa tggctgggtg      660
ggcatgaacc ccgccggcgc gccgcccgcc agcaccacgc accaggcgcg agccaatccg      720
aacaacgccg aggagccaat cgccagactg cccacgcgcc acgcgaactg gagcgccgac      780
gacaagtacc tgccgctgct gtccgaagcc ggcggtgccg ccgagcgccg cgtgttcggg      840
ctggggcact acctgcggcc gtcgatggtc gaattcgtcg actgcgggga tgtgctgatg      900
cagggctacc aggtcatcaa cacgccgttc tggattcatc acccgggtcaa ctcacgcaac      960
attcacttct ccaaagtgcg catggaaaag atcgggccga attcggacgg tttcgatccc     1020
gagtccctgc acaccatcct ggtggacggc tgccgtgtca ataccggcga cgactgcac      1080
gccatcaaat ccggcaagaa ccgagactcg caatacggcc caacgcgcaa tatggtggtc     1140
cagaactgca tcatgaaccg cggccacggc ggcgttacgc tgggcagcga aatggcgggt      1200
ggcatcgagc atatctacgc gcagaaaatc gaattccgca acgcgttctg ggaccacgac      1260
ccgctgggca cggccatccg aatgaagacg aacatgaacc gcggcggtta ccttcgtcat      1320
ttctacgtgc gcgacgtgac gctgccgaat ggcggtgcgt ccaagagcgg cttctacaag      1380
acgtgcccgg gatctccgct ggcaggcaag gtctccacca gcggcggcgc tgttatcact      1440
atcgactgcg attacgcgcc gaatgacgac agcgtgcgcg tgcggccgcc gcagggtgtc      1500
gacgtgcata tctcgaacgt ccgcgtcagc aatgtgaaaa cggccgaagg ctcggttctc      1560
tgctaccagg ccatggtgct gctcgggccc gtggcgccca gcttcaacgg cgcgcctggc      1620
acggccatcc tgccgatcac gaatgtcacc gtcagcgatt cggacttcgg cacgccgcgc      1680
aacagcgagc agccctgggt cgcgttcaac gtgcagggac tcaagctgcg caacgtgcgc      1740
atcgatggca aggagtacaa cgtatga

```

<210> 76

10 <211> 588

<212> PRT

<213> Desconocido

<220>

15 <223> Obtenido de una muestra ambiental

<221> SEÑAL

<222> (1)...(34)

20 <221> DOMINIO

<222> (110)...(555)

<223> Dominio catalítico

<400> 76

Met Thr Thr Thr Arg Arg Thr Ile Leu Lys Ala Ala Ala Ser Ala Gly

25

1	5	10	15
Ala Ile Ala Ser Thr Gly Trp Pro Ala Leu Ala Ala Ala Gln Ala Ala			
20	25	30	
Gln Ala Ala Asp Pro Trp Ala Arg Ala Gln Gln Ile Ile Asp Arg Phe			
35	40	45	
Ala Lys Pro Leu Ser Phe Pro Asn Arg Asp Phe Pro Ile Thr Glu Phe			
50	55	60	
Gly Ala Lys Pro Cys Lys Leu Val Lys Ala Gln Gly Leu Val Glu Val			
65	70	75	80
Arg Val Lys Gly Glu Leu Glu Thr Pro Ala Pro Gln Ala Pro Asp Ala			
85	90	95	
Tyr Pro Ala Ile Lys Ala Ala Ile Ala Ala Ala Ser Lys Ala Gly Gly			
100	105	110	
Gly Arg Val Leu Ile Pro Ala Gly Asn Trp Tyr Cys Lys Gly Pro Ile			
115	120	125	
Val Leu Leu Ser Asn Val His Val His Leu Ala Lys Gly Ala Gln Val			
130	135	140	
Tyr Phe Ser Ala Asn Ala Lys Asp Phe Ala Arg Asp Gly Asp Tyr Asp			
145	150	155	160
Cys Gly Ala Asn Gly Lys Leu Val Leu Ser Arg Trp Gln Gly Asn Asp			
165	170	175	
Cys Leu Asn Phe Ser Pro Met Val Tyr Ala Arg Gly Gln Lys Asn Ile			
180	185	190	
Ala Ile Thr Gly Glu Asp Trp Thr Ser Ile Leu Asn Gly Gln Ala Gly			
195	200	205	
Val Ala Phe Glu Asp Gly Ser Gly Asn Gly Trp Trp Gly Met Asn Pro			
210	215	220	
Ala Gly Ala Pro Pro Gly Ser Thr Thr His Gln Gly Ala Ala Asn Pro			
225	230	235	240
Asn Asn Ala Glu Glu Pro Ile Ala Arg Leu Pro Thr Arg His Ala Asn			
245	250	255	
Trp Ser Ala Asp Asp Lys Tyr Leu Pro Leu Leu Ser Glu Ala Gly Val			
260	265	270	
Pro Ala Glu Arg Arg Val Phe Gly Leu Gly His Tyr Leu Arg Pro Ser			
275	280	285	
Met Val Glu Phe Val Asp Cys Gly Asp Val Leu Met Gln Gly Tyr Gln			
290	295	300	
Val Ile Asn Thr Pro Phe Trp Ile His His Pro Val Asn Ser Arg Asn			
305	310	315	320
Ile His Phe Ser Lys Val Arg Met Glu Ser Ile Gly Pro Asn Ser Asp			
325	330	335	
Gly Phe Asp Pro Glu Ser Cys Asp Thr Ile Leu Val Asp Gly Cys Leu			
340	345	350	
Phe Asn Thr Gly Asp Asp Cys Ile Ala Ile Lys Ser Gly Lys Asn Arg			
355	360	365	
Asp Ser Gln Tyr Gly Pro Thr Arg Asn Met Val Val Gln Asn Cys Ile			
370	375	380	
Met Asn Arg Gly His Gly Gly Val Thr Leu Gly Ser Glu Met Ala Gly			
385	390	395	400
Gly Ile Glu His Ile Tyr Ala Gln Lys Ile Glu Phe Arg Asn Ala Phe			
405	410	415	
Trp Asp His Asp Pro Leu Gly Thr Ala Ile Arg Met Lys Thr Asn Met			
420	425	430	
Asn Arg Gly Gly Tyr Leu Arg His Phe Tyr Val Arg Asp Val Thr Leu			
435	440	445	
Pro Asn Gly Val Arg Thr Lys Ser Gly Phe Tyr Lys Thr Leu Pro Gly			
450	455	460	
Ser Pro Leu Ala Gly Lys Val Ser Thr Ser Gly Gly Ala Val Ile Thr			
465	470	475	480

Ile Asp Cys Asp Tyr Ala Pro Asn Asp Asp Ser Val Arg Val Arg Pro
 485 490 495
 Pro Gln Val Ser Asp Val His Ile Ser Asn Val Arg Val Ser Asn Val
 500 505 510
 Lys Thr Ala Glu Gly Ser Phe Ser Cys Tyr Gln Ala Met Val Leu Leu
 515 520 525
 Gly Pro Val Ala Ala Ser Phe Asn Gly Ala Pro Gly Thr Ala Ile Leu
 530 535 540
 Pro Ile Thr Asn Val Thr Val Ser Asp Ser Asp Phe Gly Thr Pro Arg
 545 550 555 560
 Asn Ser Ala Glu Pro Trp Phe Ala Phe Asn Val Gln Gly Leu Lys Leu
 565 570 575
 Arg Asn Val Arg Ile Asp Gly Lys Glu Tyr Asn Val
 580 585

<210> 77

<211> 2043

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 77

atgaaaacct	ccagagcaat	ttttaactaca	tcaacacttt	tacaccgcgc	gcttatcgcg	60
gctagtgtca	gcatggcaat	gagttctgcc	gcatgggcgg	gttgtacct	taccgtcacc	120
aataattggg	gctcaggatt	taccggcgaa	atcaaagtga	ccaacaacac	caccgocagc	180
gtgaacaatt	ggctctgtgc	atggcaggaa	tccgggtgcg	ctatcaccaa	cgcctggaat	240
gcaacgctca	gtggctcaaa	cccttacaca	gocgtatccg	ctggttggaa	tggcacactt	300
gcccccaatg	catcgccac	ttttggtttc	caggcaaacg	gttctgccgg	tgcacctaaa	360
gtgaatggca	gcttgtgcgg	caccaacact	tcatacaacac	cggcatccag	cagtgttgcc	420
agctcgggta	aatcaagcgc	gcccgtatcg	tccagcagca	gatcatccag	ttcaatcgct	480
atcactagca	gctcttttagc	gagaagttct	attgcctcca	gcagctcact	agttagtagc	540
tccagagcga	gcagtagtgc	gccaagcggt	ttctctttta	cgatccagga	agagcaagcg	600
ggcttctgtc	gtgttgatgg	cattgcgcaca	gaaagcacca	acaccgggtt	taccggcaat	660
ggctacacca	atgcgaacaa	cgcgcaaggc	gcagcgattg	aatgggcagt	cagcgcacct	720
agcagtggcc	gttatacagt	agccttcgcg	ttcgccaatg	gcggcacagc	agcgcgcaac	780
ggctcgttgt	taatcaatgg	cggtagcaat	ggtaattaca	ctgtggagtt	acccttgacc	840
ggcgcatggg	caacctggca	aattgccagc	gtggaaattg	atttagtgca	aggcaataat	900
attttaaaac	tctcggcgtt	aaccgctgac	ggtttggcca	atatcgactc	attaaaaata	960
gacggcgcg	aaaccaaagc	aggtacttgc	agcactacat	caagcagcag	cgttgccagc	1020
agctcgtcgt	ccgttaaate	cagcgcaagt	tcttcttcga	gttcatccac	cgtgcacaaa	1080
atactgacat	tagacggtaa	cccgccgcgc	agctggttca	acaaatccag	gaccaagtgg	1140
aatagcagcc	gcgcgatat	tgtgttgtct	taccagcaat	ccaacggcgg	ttggccaaaa	1200
aacctggatt	acaactcagt	gagcgcaggc	aatggcggga	gcgacagcgg	caccatcgac	1260
aatggtgcaa	ccattaccga	aatggtttac	ctcgctgaaa	tttataaaaa	cggcggcaac	1320
accaaataatc	gcgatgcagt	gcgcagagca	gcaaactttt	tagtgagctc	gcaatacagc	1380
acaggcgccct	tgccacaatt	ttatccggtg	aaaggcggct	atgcggatca	tgcgaccttt	1440
aacgataacg	gcatggcgta	cgcgttgacg	gtattggatt	tgcagtaaaa	caaacgcgca	1500
ccgtttgata	acgacatttt	ctctgattct	gatcgggcga	aattcaaaaac	cgtgttgcc	1560
aaaggtgtgg	attacatttt	aaaagcgcag	tggaaacaaa	atggaaaact	cactgcatgg	1620
tgtgcacaac	acggtgctac	ggattaccaa	ccgaaaaaag	cgcgcgctta	tgaattggaa	1680
tcattgagtg	gtagcagtc	ggtcggcatt	ctgccttct	tgatgaccca	accacaaacc	1740
gcgcaaatcg	aagcggcggt	caaggcgggt	gtcaactggt	tgcgcagtc	aaatacttat	1800
ttggctaact	acacttacga	ttcatcaaaa	gcgtctacca	acccgattgt	gtataaatcc	1860
ggaagcagaa	tgtggtatcg	cttctatgac	ctgaacacca	accgtgggtt	ctttagtgat	1920
cgcgatggca	gcaaattcta	tgatatcacc	caaattgtcag	aagagcgtcg	caccggttat	1980
agctgggggtg	gctcttacgg	tgaatctatt	atttccttcg	cgcaaaaagt	gggttatctg	2040
taa						2043

15 <210> 78

<211> 680

<212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

5

<221> SEÑAL

<222> (1)...(33)

<221> UNIÓN

10 <222> (34)...(126)

<223> Módulo de unión a carbohidratos

<221> UNIÓN

<222> (199)...(322)

15 <223> Módulo de unión a carbohidratos

<221> DOMINIO

<222> (323)...(680)

<223> Dominio catalítico

20

<400> 78

Met	Lys	Thr	Ser	Arg	Ala	Ile	Phe	Thr	Thr	Ser	Thr	Leu	Leu	His	Arg
1				5					10					15	
Ala	Leu	Ile	Ala	Ala	Ser	Val	Ser	Met	Ala	Met	Ser	Ser	Ala	Ala	Trp
			20					25					30		
Ala	Gly	Cys	Thr	Tyr	Thr	Val	Thr	Asn	Asn	Trp	Gly	Ser	Gly	Phe	Thr
		35					40					45			
Gly	Glu	Ile	Lys	Val	Thr	Asn	Asn	Thr	Thr	Ala	Ser	Val	Asn	Asn	Trp
	50					55					60				
Ser	Val	Ser	Trp	Gln	Glu	Ser	Gly	Ala	Ala	Ile	Thr	Asn	Ala	Trp	Asn
65				70					75						80
Ala	Thr	Leu	Ser	Gly	Ser	Asn	Pro	Tyr	Thr	Ala	Val	Ser	Ala	Gly	Trp
				85					90					95	
Asn	Gly	Thr	Leu	Ala	Pro	Asn	Ala	Ser	Ala	Thr	Phe	Gly	Phe	Gln	Ala
		100					105						110		
Asn	Gly	Ser	Ala	Gly	Ala	Pro	Lys	Val	Asn	Gly	Ser	Leu	Cys	Gly	Thr
	115						120					125			
Asn	Thr	Ser	Ser	Thr	Pro	Ala	Ser	Ser	Ser	Val	Ala	Ser	Ser	Val	Lys
	130					135					140				
Ser	Ser	Ala	Pro	Val	Ser	Ser	Ser	Ser	Arg	Ser	Ser	Ser	Ser	Ile	Ala
145				150					155						160
Ile	Thr	Ser	Ser	Ser	Leu	Ala	Arg	Ser	Ser	Ile	Ala	Ser	Ser	Ser	Ser
				165					170					175	
Leu	Val	Ser	Ser	Ser	Arg	Ala	Ser	Ser	Ser	Ala	Pro	Ser	Val	Phe	Ser
		180					185						190		
Phe	Thr	Ile	Gln	Glu	Glu	Gln	Ala	Gly	Phe	Cys	Arg	Val	Asp	Gly	Ile
	195						200					205			
Ala	Thr	Glu	Ser	Thr	Asn	Thr	Gly	Phe	Thr	Gly	Asn	Gly	Tyr	Thr	Asn
	210					215					220				
Ala	Asn	Asn	Ala	Gln	Gly	Ala	Ala	Ile	Glu	Trp	Ala	Val	Ser	Ala	Pro
225				230					235						240
Ser	Ser	Gly	Arg	Tyr	Thr	Val	Ala	Phe	Arg	Phe	Ala	Asn	Gly	Gly	Thr
			245						250					255	
Ala	Ala	Arg	Asn	Gly	Ser	Leu	Leu	Ile	Asn	Gly	Gly	Ser	Asn	Gly	Asn
			260					265					270		

Tyr	Thr	Val	Glu	Leu	Pro	Leu	Thr	Gly	Ala	Trp	Ala	Thr	Trp	Gln	Ile
		275					280					285			
Ala	Ser	Val	Glu	Ile	Asp	Leu	Val	Gln	Gly	Asn	Asn	Ile	Leu	Lys	Leu
		290				295					300				
Ser	Ala	Leu	Thr	Ala	Asp	Gly	Leu	Ala	Asn	Ile	Asp	Ser	Leu	Lys	Ile
305					310					315					320
Asp	Gly	Ala	Gln	Thr	Lys	Ala	Gly	Thr	Cys	Ser	Thr	Thr	Ser	Ser	Ser
				325					330						335
Ser	Val	Ala	Ser	Ser	Ser	Ser	Ser	Val	Lys	Ser	Ser	Ala	Ser	Ser	Ser
			340					345					350		
Ser	Ser	Ser	Ser	Thr	Ala	Ala	Lys	Ile	Leu	Thr	Leu	Asp	Gly	Asn	Pro
		355					360					365			
Ala	Ala	Ser	Trp	Phe	Asn	Lys	Ser	Arg	Thr	Lys	Trp	Asn	Ser	Ser	Arg
		370				375					380				
Ala	Asp	Ile	Val	Leu	Ser	Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro	Lys
385					390					395					400
Asn	Leu	Asp	Tyr	Asn	Ser	Val	Ser	Ala	Gly	Asn	Gly	Gly	Ser	Asp	Ser
				405					410					415	
Gly	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu	Ala
			420					425					430		
Glu	Ile	Tyr	Lys	Asn	Gly	Gly	Asn	Thr	Lys	Tyr	Arg	Asp	Ala	Val	Arg
		435					440					445			
Arg	Ala	Ala	Asn	Phe	Leu	Val	Ser	Ser	Gln	Tyr	Ser	Thr	Gly	Ala	Leu
	450				455						460				
Pro	Gln	Phe	Tyr	Pro	Leu	Lys	Gly	Gly	Tyr	Ala	Asp	His	Ala	Thr	Phe
465					470					475					480
Asn	Asp	Asn	Gly	Met	Ala	Tyr	Ala	Leu	Thr	Val	Leu	Asp	Phe	Ala	Val
			485					490					495		
Asn	Lys	Arg	Ala	Pro	Phe	Asp	Asn	Asp	Ile	Phe	Ser	Asp	Ser	Asp	Arg
			500					505					510		
Ala	Lys	Phe	Lys	Thr	Ala	Val	Ala	Lys	Gly	Val	Asp	Tyr	Ile	Leu	Lys
		515					520					525			
Ala	Gln	Trp	Lys	Gln	Asn	Gly	Lys	Leu	Thr	Ala	Trp	Cys	Ala	Gln	His
		530				535					540				
Gly	Ala	Thr	Asp	Tyr	Gln	Pro	Lys	Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Glu
545					550					555					560
Ser	Leu	Ser	Gly	Ser	Glu	Ser	Val	Gly	Ile	Leu	Ala	Phe	Leu	Met	Thr
			565					570					575		
Gln	Pro	Gln	Thr	Ala	Gln	Ile	Glu	Ala	Ala	Val	Lys	Ala	Gly	Val	Asn
			580					585					590		
Trp	Phe	Ala	Ser	Pro	Asn	Thr	Tyr	Leu	Ala	Asn	Tyr	Thr	Tyr	Asp	Ser
		595					600					605			

<210> 79
<211> 1746
<212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 79

```

atgacgacac gacgcgaatt catcaaaggc tttctactta cgggagcagc cgtggcgcgc 60
gctccgcgtt taattgcgtt cggcgcggag gcaagtcgtt gggaaacgat gatgccttcg 120
atcctcgcac gcatcacacc acctcgtttt ccgaaacgca ctttctatct caatcgattc 180
ggcgccaagg gcgatggagt cacagactgc accgcggctt ttcacgcgcg gatcgatgaa 240
tgcaccaaag ccggcgggtg gaaagtcgtc gtgcccggcg gcacttatct caccggcgcg 300
attcatttga agagcaacgt caacctcgaa gtctcggaag gcgcgacgat caagttcagt 360
caggaccgca aacactacct gcctgttgtc ttctcgcgtt gggaaaggtgt cgaagtcttc 420
aactactcgc ctttcattta cgcgttcgaa cagcgaacaa tcgcgatcac cggcaaaggc 480
acgctcgacg gacagagtga ttcggaacac tgggtggcgt ggaacggcgc tccgcagtac 540
ggatggaatg aagggatgaa acagcagcgt cccgatcgca acgcgttggt cacaatggcg 600
gagaaaggcg tgccggtgcg cgagcgcctc tttggcgaag gtcattattt gaggccgcag 660
ttcattcagc cgtaccgctg ccagaacgtg ctgatccagg gcgtgacgat tcggaactcg 720
ccgatgtggg agattcatcc ggtgttggtg cgcaacgtga ctattcacga cgtgcacatc 780
gatagtcgat gaccaaacia cgacggctgc aatcccgaat cgtgcagcga cgtgttgatt 840
aaggatagct acttcgatac cggcgcacgc tgcctcgca tcaaatcggg acgcaacgcc 900
gacgggcggc ggcttaaagc gccgactgag aacatcatcg ttcaaggatg tcgcatgaaa 960
gacggccacg gtggaatcac ggtcggcagc gagatctcgg gcggcgtgcg aaacctgttt 1020
gccgagaatt gccggctcga cagtcctaac ctcgatcacg cctgcgcgt gaagaacaat 1080
gccatgcgcg gcggattact cgagaacttc cacttcgta acatcgaagt cgggcaggtg 1140
gcccattgcc tgattacgat cgacttcaac tacgaagagg gcgcgaaagg gtcgttcacg 1200
ccggtcgttc gcgattacac ggtcgatcgt ttgcgcagca cgaagagcaa gcacgcactc 1260
gacgtccagg gtctgcccg gcgcgcggtc atcaacctgc gattgacaaa ctgcacattc 1320
aacgatgtgc agcaaccgaa cattctcaag aacgtcgaac aatcaacctt tgagaacgtc 1380
acgattaacg gaaagacgat cacacaaaca ggatccatct cagaaagagc ggccacgaca 1440
gcaatgaccg cgttttgcg cgacgcgtcg aggaagaaa acggttatcc cgcgaagtgg 1500
acctatgatc atgggctggt cctgaaagga atcgagcgcg tttggaacaa taccggcgat 1560
aagaagtatc tgaagttcat ccaggacagc atggaccact tcgtcaacga cgacggctcc 1620
attcgcacct acacgatcga cgagtacaac atcgatcacg ttcttcccg acgaaacctc 1680
ctgttccttt acaaaactac cgttcaggaa aagtatcgca aagccgcgcg gttcttgccg 1740
gaacaa

```

<210> 80
 <211> 582
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <221> SEÑAL
 <222> (1)...(28)

15 <221> DOMINIO
 <222> (81)...(458)
 <223> Dominio catalítico

<400> 80

Met	Thr	Thr	Arg	Arg	Glu	Phe	Ile	Lys	Gly	Phe	Leu	Leu	Thr	Gly	Ala
1				5					10					15	
Ala	Val	Ala	Val	Ala	Pro	Arg	Leu	Ile	Ala	Phe	Gly	Ala	Glu	Ala	Ser
			20					25					30		
Pro	Trp	Glu	Thr	Met	Met	Pro	Ser	Ile	Leu	Ala	Arg	Ile	Thr	Pro	Pro
		35					40				45				
Arg	Phe	Pro	Lys	Arg	Thr	Phe	Tyr	Leu	Asn	Arg	Phe	Gly	Ala	Lys	Gly
	50					55					60				
Asp	Gly	Val	Thr	Asp	Cys	Thr	Ala	Ala	Phe	His	Arg	Ala	Ile	Asp	Glu

65					70					75				80	
Cys	Thr	Lys	Ala	Gly	Gly	Gly	Lys	Val	Val	Val	Pro	Ala	Gly	Thr	Tyr
				85						90				95	
Leu	Thr	Gly	Ala	Ile	His	Leu	Lys	Ser	Asn	Val	Asn	Leu	Glu	Val	Ser
			100					105					110		
Glu	Gly	Ala	Thr	Ile	Lys	Phe	Ser	Gln	Asp	Pro	Lys	His	Tyr	Leu	Pro
		115				120						125			
Val	Val	Phe	Ser	Arg	Trp	Glu	Gly	Val	Glu	Val	Phe	Asn	Tyr	Ser	Pro
	130					135					140				
Phe	Ile	Tyr	Ala	Phe	Glu	Gln	Arg	Asn	Ile	Ala	Ile	Thr	Gly	Lys	Gly
145					150					155					160
Thr	Leu	Asp	Gly	Gln	Ser	Asp	Ser	Glu	His	Trp	Trp	Pro	Trp	Asn	Gly
			165						170					175	
Arg	Pro	Gln	Tyr	Gly	Trp	Asn	Glu	Gly	Met	Lys	Gln	Gln	Arg	Pro	Asp
		180						185					190		
Arg	Asn	Ala	Leu	Phe	Thr	Met	Ala	Glu	Lys	Gly	Val	Pro	Val	Arg	Glu
	195						200					205			
Arg	Ile	Phe	Gly	Glu	Gly	His	Tyr	Leu	Arg	Pro	Gln	Phe	Ile	Gln	Pro
210					215						220				
Tyr	Arg	Cys	Gln	Asn	Val	Leu	Ile	Gln	Gly	Val	Thr	Ile	Arg	Asn	Ser
225				230						235					240
Pro	Met	Trp	Glu	Ile	His	Pro	Val	Leu	Cys	Arg	Asn	Val	Thr	Ile	His
			245						250					255	
Asp	Val	His	Ile	Asp	Ser	His	Gly	Pro	Asn	Asn	Asp	Gly	Cys	Asn	Pro
		260						265					270		
Glu	Ser	Cys	Ser	Asp	Val	Leu	Ile	Lys	Asp	Ser	Tyr	Phe	Asp	Thr	Gly
	275						280					285			
Asp	Asp	Cys	Ile	Ala	Ile	Lys	Ser	Gly	Arg	Asn	Ala	Asp	Gly	Arg	Arg
290					295						300				
Leu	Lys	Ala	Pro	Thr	Glu	Asn	Ile	Ile	Val	Gln	Gly	Cys	Arg	Met	Lys
305				310						315					320
Asp	Gly	His	Gly	Gly	Ile	Thr	Val	Gly	Ser	Glu	Ile	Ser	Gly	Gly	Val
			325					330						335	
Arg	Asn	Leu	Phe	Ala	Glu	Asn	Cys	Arg	Leu	Asp	Ser	Pro	Asn	Leu	Asp
		340						345					350		
His	Ala	Leu	Arg	Val	Lys	Asn	Asn	Ala	Met	Arg	Gly	Gly	Leu	Leu	Glu
	355					360						365			
Asn	Phe	His	Phe	Arg	Asn	Ile	Glu	Val	Gly	Gln	Val	Ala	His	Ala	Val
370					375						380				
Ile	Thr	Ile	Asp	Phe	Asn	Tyr	Glu	Glu	Gly	Ala	Lys	Gly	Ser	Phe	Thr
385				390						395					400
Pro	Val	Val	Arg	Asp	Tyr	Thr	Val	Asp	Arg	Leu	Arg	Ser	Thr	Lys	Ser
			405					410						415	
Lys	His	Ala	Leu	Asp	Val	Gln	Gly	Leu	Pro	Gly	Ala	Pro	Val	Ile	Asn
		420						425					430		
Leu	Arg	Leu	Thr	Asn	Cys	Thr	Phe	Asn	Asp	Val	Gln	Gln	Pro	Asn	Ile
	435						440					445			
Leu	Lys	Asn	Val	Glu	Gln	Ser	Thr	Phe	Glu	Asn	Val	Thr	Ile	Asn	Gly
450					455						460				
Lys	Thr	Ile	Thr	Gln	Thr	Gly	Ser	Ile	Ser	Glu	Arg	Ala	Ala	Thr	Thr
465				470						475					480
Ala	Met	Thr	Ala	Leu	Trp	Arg	Asp	Ala	Ser	Arg	Lys	Glu	Asn	Gly	Tyr
			485					490						495	
Pro	Ala	Lys	Trp	Thr	Tyr	Asp	His	Gly	Leu	Val	Leu	Lys	Gly	Ile	Glu
	500							505					510		
Arg	Val	Trp	Asn	Asn	Thr	Gly	Asp	Lys	Lys	Tyr	Leu	Lys	Phe	Ile	Gln
	515					520						525			
Asp	Ser	Met	Asp	His	Phe	Val	Asn	Asp	Asp	Gly	Ser	Ile	Arg	Thr	Tyr
530					535						540				

```

Thr Ile Asp Glu Tyr Asn Ile Asp His Val Leu Pro Gly Arg Asn Leu
545                               550                               555                               560
Leu Phe Leu Tyr Lys Thr Thr Gly Gln Glu Lys Tyr Arg Lys Ala Ala
                               565                               570                               575
Ala Phe Leu Arg Glu Gln
                               580

```

<210> 81
 <211> 1065
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 81

```

atgacgctac ccgttggttc cctgcgcgtg ctgctggcgc tgetggccac gttgccggtc      60
gcctgcgcgg gcgctgcggt atccgcggca gcgaccgacc cggtcgccga gaacatgctg      120
ctgctgcaga ccgcctccgg tggctgggtc aagcaactacc gcgggaagaa ggtcgactac      180
acgcgcaatt acgacaccgc cgagcgcgcc gcgctgcgcg cgcgccggcg gcatgacgac      240
gcgacgatcg acaacaaggc cagcaccagc gagatcgctt acctggtgca ggcacatgcc      300
aggacgggca acccggcgta cctcgacggt gcccgcgcgc gggtcgaata cctgctgcgc      360
gcgcagtacc cgaatggtgg ctggccgcag ttctaccccg accactcgtc ctaccggcac      420
cagatcacgc tcaacgacga cgcgatggtg catgccatca ccgtgctgca ggacatcgcc      480
gccggccgcg acggcatgca ggcgttgacg cccgagttcg gcgtccgcgc cgcgcgcgcc      540
gcgcagcgcg gcatcggaaa cctgctcgag ttgcaggtgc ggatcgacgg cgagccgacc      600
atctgggccc cgcagtacga cgagcatagc ctgcagccgg ccaaggcccg cgcctatgaa      660
ctgcctcgcc tggccgtggc cgaatcggtc ggcgtggtgc gcctgctgat gcgccagccg      720
aggccggatg ccgggaccgt cgcgcgcatc gaatcggcgg cccgctggct ggagggcgcat      780
cgctgcatg acctggcgct cgaacgcgtc gacgcaccgg ccgaggaaac gggcaaggac      840
gtgcgggctg tgaccgggcc cggcgccctg ctgtgggcgc gcttctacga cctggatgga      900
cagcagcctc tgttcgtcga ccgcgacagc aagcccgctc cgttcgccag cctgcccac      960
gagcgcgcga ccggctatgc ctggtacggc acctggccgg agaagctgct ggcgcaggaa     1020
ctccgcgctt ggcgcgaggt ccatgcgcgc ggcgcgcgc cctga                               1065

```

<210> 82
 <211> 354
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(30)

<221> DOMINIO
 <222> (31)...(354)
 <223> Dominio catalítico

<400> 82

```

Met Thr Leu Pro Val Val Ser Leu Arg Val Leu Leu Ala Leu Leu Ala
1      5      10      15
Thr Leu Pro Val Ala Cys Ala Gly Ala Ala Val Ser Ala Ala Ala Thr
20      25      30
Asp Pro Val Ala Glu Asn Met Leu Leu Leu Gln Thr Ala Ser Gly Gly
35      40      45
Trp Ser Lys His Tyr Arg Gly Lys Lys Val Asp Tyr Thr Arg Asn Tyr
50      55      60

```


Asp	Thr	Ala	Glu	Arg	Ala	Ala	Leu	Arg	Ala	Pro	Gly	Arg	His	Asp	Asp	65	70	75	80
Ala	Thr	Ile	Asp	Asn	Lys	Ala	Thr	Thr	Ser	Glu	Ile	Ala	Tyr	Leu	Val	85	90	95	
Gln	Ala	His	Ala	Arg	Thr	Gly	Asn	Pro	Ala	Tyr	Leu	Asp	Gly	Ala	Arg	100	105	110	
Arg	Gly	Val	Glu	Tyr	Leu	Leu	Arg	Ala	Gln	Tyr	Pro	Asn	Gly	Gly	Trp	115	120	125	
Pro	Gln	Phe	Tyr	Pro	Asp	His	Ser	Ser	Tyr	Arg	His	Gln	Ile	Thr	Leu	130	135	140	
Asn	Asp	Asp	Ala	Met	Val	His	Ala	Ile	Thr	Val	Leu	Gln	Asp	Ile	Ala	145	150	155	160
Ala	Gly	Arg	Asp	Gly	Met	Gln	Ala	Leu	Thr	Pro	Glu	Phe	Gly	Val	Arg	165	170	175	
Ala	Ala	Ala	Ala	Ala	Gln	Arg	Gly	Ile	Gly	Asn	Leu	Leu	Glu	Leu	Gln	180	185	190	
Val	Arg	Ile	Asp	Gly	Glu	Pro	Thr	Ile	Trp	Ala	Ala	Gln	Tyr	Asp	Glu	195	200	205	
His	Ser	Leu	Gln	Pro	Ala	Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Pro	Ser	Leu	210	215	220	
Ala	Val	Ala	Glu	Ser	Val	Gly	Val	Val	Arg	Leu	Leu	Met	Arg	Gln	Pro	225	230	235	240
Arg	Pro	Asp	Ala	Arg	Thr	Val	Ala	Ala	Ile	Glu	Ser	Ala	Ala	Arg	Trp	245	250	255	
Leu	Glu	Ala	His	Arg	Leu	His	Asp	Leu	Ala	Leu	Glu	Arg	Val	Asp	Ala	260	265	270	
Pro	Ala	Glu	Glu	Thr	Gly	Lys	Asp	Val	Arg	Val	Val	Thr	Arg	Pro	Gly	275	280	285	
Ala	Ser	Leu	Trp	Ala	Arg	Phe	Tyr	Asp	Leu	Asp	Gly	Gln	Gln	Pro	Leu	290	295	300	
Phe	Val	Asp	Arg	Asp	Ser	Lys	Pro	Val	Pro	Phe	Ala	Ser	Leu	Pro	Asn	305	310	315	320
Glu	Arg	Arg	Thr	Gly	Tyr	Ala	Trp	Tyr	Gly	Thr	Trp	Pro	Glu	Lys	Leu	325	330	335	
Leu	Ala	Gln	Glu	Leu	Pro	Arg	Trp	Arg	Glu	Val	His	Ala	Ala	Gly	Ala	340	345	350	
Ala	Pro																		

<210> 83

<211> 3618

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 83

ttgactgctc	tttcaagaaa	cagtcagggt	gatgtaagct	gggaaccgag	ttccgcccac	60
caggtaacct	acaatctaaa	acgcagtacc	acgaaggagg	gtccttatca	gaccattgct	120
gaaaaaatgg	cagaaaccga	cttccgggat	acagggttag	agaatggcca	gaagtattac	180
tatgttgat	ctgccgaaac	gagtagcggt	gagagtgcag	attcacaagc	tataacggct	240
gtgcctgtag	cgccattgca	agctccgacc	ggcctttcag	caagtcattg	caatggcggg	300
gtaaccattc	attgggaatc	cgtcaatggt	gccgagtctt	accaagtctt	gcgcagtaaa	360
caaaaggggca	tcggctatga	agtcattcaaa	aacgggtgtaa	cggaaaccag	ttatacagat	420
accgggattc	ccgatggcga	gaagtactat	tatgtcgtat	ccgccaagaa	cgatacagct	480
gaaagtgcac	attcccaacc	gattaacggg	gctgctgtat	cgacgagtgg	tgtaccagcc	540
attccaaacg	gtatgaacgc	tactgccggg	gatggcagag	ctgccttaac	ctggctccgct	600
gtatccggcg	cagatttccta	tagcatcaag	cgcgggtgagt	ttaacagtgg	tcaatatgag	660

gtcattgcta	aaaatataca	ctctaccggt	tatcaagata	taggccttac	aaacgggtgat	720
acctatgatt	atgtgatttc	cgctgtcaat	gagcaagggg	aaagtttagg	ctccgaaccc	780
atcgccgta	ctcctgcgaa	agtaacggtt	gtagcgaaag	aaggcggaga	ctttaaaacg	840
attcaagaag	ccattgatgc	tgcacctgat	aacagtacga	aacggcatgt	tatttttata	900
aaaaatggtc	aatatcgtga	aaagcttacg	atccctaaga	gcaaaaaccaa	tctgagtttt	960
gtaggggaaa	gtaaggaagg	gaccgtgctt	gttttttaatg	ataatgcaaa	tacgcctggg	1020
ccagacggca	aaccattagg	cacttccaat	agttcaagta	tctttatcta	tgccaatgat	1080
tttattgccc	aaaatttaac	catccagaac	gactctgggt	aaggaacagg	tcaagcagtt	1140
gccgcttatg	taaggggcga	togtctctac	tttgaaaacg	tgcatgtttt	aggataccag	1200
gatacattat	atgcccatac	gggaagacag	tattataaaa	actgctacgt	agaaggggat	1260
gtggatttca	tttttgccgg	agccacagcc	ttgtttgata	cctgtcacct	tcatagcaag	1320
cgtacaggca	gtaagttaac	cgcagctagt	accgatcaag	tcacaccgta	tgggtatgtt	1380
tttttagatt	caaaaatcac	ctcagatgaa	ggcgtgacca	atgtgcatct	cgggcgacct	1440
tggcgctcct	attcagctgt	cacctatata	aacaccgaaa	tggatgcato	gattgttcct	1500
gacggatggg	ataactgggg	gaaagttgaa	aacgagaaaa	cagccagata	ttcagaatac	1560
aataacatgg	ggccaggggc	agacccgaaa	aagcgggatc	catggaccac	acaattaacg	1620
ccggagggaag	cgaatcaata	cactgtgcaa	aatgtgatga	aaggatctga	cggctgggat	1680
cctgagagaa	tcgggattat	cccattatca	ccactgtcag	caccgattat	ttcacttgat	1740
caacgagatt	ccattgtcaa	tacaccaagc	tttacaatta	caggtcaagt	ggataaagaa	1800
gcagccgttt	ctgtcaatgg	gaaggaaatt	tccttacaaa	aggatggcag	cttcagcacg	1860
acggtgggtgc	tgaatgacgg	tttgaacact	attacagtgg	gagctgtaga	tgcagcaggg	1920
aatcaggcta	ttcctgcagt	gttaaaaatt	gtttatgatc	atgagaaacc	tgctgtatcc	1980
atcgatgatc	ttaaaggaga	aaaaaacggg	aatcactaca	atgtaatcta	caatccgctg	2040
ccgattacag	ggaagctgaa	cgaagcagga	acagttatgg	tgaatggtga	gaaagtaaat	2100
gtatcggaag	agttgacgtt	tagtacaaaa	gtcattttta	agccgggggt	aaataacatt	2160
acgattaccg	ctgttgatca	ggcaggggat	gaagccgaat	ctatcactat	caatgtgggt	2220
ccaaaaggga	atgctgttcc	agacgggtcc	gtcaagatta	tcaaaagtga	aacaacaaat	2280
gcaaataccg	ttgagggttac	ttttaatagc	aagctagaaa	aatttgattc	tagtgatatt	2340
gcattgcaaa	cggctacgaa	cgtttgggca	gctctcaatc	ctgggttgaa	acaattgatg	2400
acagtggaaa	gcattaccac	aaaagtgaat	aaggataacc	aaaccgtagc	ggtgatcaaa	2460
acgaagggaag	cctttcaaga	agatggaacc	attacgctcc	caaaagtga	agatccgttt	2520
catattcaat	atttgaatgc	cgaattattat	accggggatc	gtacgcagga	cattaagcat	2580
gcggatgccc	tcttaacctg	gcagatggat	catggcgggt	ggtttaaaaa	ctgggtggaa	2640
aaatataaac	ggccatggga	tgggaaagaa	ccaaaatctg	aatgggtattc	gactaatcat	2700
ggtgaactag	ggacgattga	taatgacgca	acaacaaacg	agatttctct	tttagctctg	2760
atgtataaag	aaacaggtga	tgctcggtat	aaggattccg	ttttaaaagg	aattgatttt	2820
ttactagaga	tgcagggtga	ttccggcggc	tggccgcagg	tctatcctgc	aagaagcggg	2880
tactcagatt	atgtgacctt	taatgataat	gcgatgggtc	gcgttatgag	tgtattaaacg	2940
atgggttaaag	aaaagaagta	tccgtttaat	tccaacctag	gtgacgagca	actttctgag	3000
cagattgatg	atgcattggg	ccgtgggctg	gattatatgt	taaaatcgca	aattaaggta	3060
gacggtgaag	taaccgcatg	gtgtgctcag	catgacctg	tgacgtatga	accgaaaggg	3120
gctcggtcgt	atgaacatcc	ttcaatctct	ggttcggaat	ctgtagggat	tgtccagtat	3180
ttgatgtcac	tgccgaatcc	ttcaactgag	gttcaggctg	ccattcatgg	agctctaaat	3240
tggtttgaaag	aggcaaaatt	ggcgggaacg	aagtatgtat	caggcgatcc	aaatgggcaa	3300
tattttctacc	cggacgcaa	cagcaatac	tggtaccgct	tctatgaaat	tggcaccaat	3360
cgccgagatt	tctcgggaag	agacgggtgc	attaacaca	acatcttaga	gattgaaaaa	3420
gaaagaagag	ctcgctaccg	ctgggcagga	gaatggccgc	aaaaattatt	aaatatcgcc	3480
aaacacaactg	gtactacga	aaacagagta	tacgtagaag	tcgttgggga	tcagtctaaa	3540
aacgcccgtg	gcgaatcttt	ggaaatagga	aaattatata	gaatagaggc	ctcggcttcc	3600
ggttctacaa	gcaagtaa					3618

<210> 84
 <211> 1205
 5 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <221> DOMINIO
 <222> (268)...(556)
 <223> Dominio de pectin metil esterasa

15 <221> DOMINIO

<222> (782)...(1164)
 <223> Dominio catalítico

<400> 84

5

```

Met Thr Ala Leu Ser Arg Asn Ser Gln Val Asp Val Ser Trp Glu Pro
 1      5      10      15
Ser Ser Ala Gln Gln Val Thr Tyr Asn Leu Lys Arg Ser Thr Thr Lys
      20      25      30
Glu Gly Pro Tyr Gln Thr Ile Ala Glu Lys Met Ala Glu Thr Asp Phe
      35      40      45
Arg Asp Thr Gly Leu Glu Asn Gly Gln Lys Tyr Tyr Tyr Val Val Ser
      50      55      60
Ala Glu Thr Ser Ser Gly Glu Ser Ala Asp Ser Gln Ala Ile Thr Ala
      65      70      75      80
Val Pro Val Ala Pro Leu Gln Ala Pro Thr Gly Leu Ser Ala Ser His
      85      90      95
Gly Asn Gly Gly Val Thr Ile His Trp Glu Ser Val Asn Gly Ala Glu
      100      105      110
Ser Tyr Gln Val Leu Arg Ser Lys Gln Lys Gly Ile Gly Tyr Glu Val
      115      120      125
Ile Lys Asn Gly Val Thr Glu Thr Ser Tyr Thr Asp Thr Gly Ile Pro
      130      135      140
Asp Gly Glu Lys Tyr Tyr Tyr Val Val Ser Ala Lys Asn Asp Thr Ala
      145      150      155      160
Glu Ser Ala Asn Ser Gln Pro Ile Asn Gly Ala Ala Val Ser Thr Ser
      165      170      175
Gly Val Pro Ala Ile Pro Asn Gly Met Asn Ala Thr Ala Gly Asp Gly
      180      185      190
Arg Ala Ala Leu Thr Trp Ser Ala Val Ser Gly Ala Asp Ser Tyr Ser
      195      200      205
Ile Lys Arg Gly Glu Phe Asn Ser Gly Gln Tyr Glu Val Ile Ala Lys
      210      215      220
Asn Ile His Ser Thr Gly Tyr Gln Asp Ile Gly Leu Thr Asn Gly Asp
      225      230      235      240
Thr Tyr Asp Tyr Val Ile Ser Ala Val Asn Glu Gln Gly Glu Ser Leu
      245      250      255
Gly Ser Glu Pro Ile Ala Val Thr Pro Ala Lys Val Thr Val Val Ala
      260      265      270
Lys Glu Gly Gly Asp Phe Lys Thr Ile Gln Glu Ala Ile Asp Ala Ala
      275      280      285
Pro Asp Asn Ser Thr Lys Arg His Val Ile Phe Ile Lys Asn Gly Gln
      290      295      300
Tyr Arg Glu Lys Leu Thr Ile Pro Lys Ser Lys Thr Asn Leu Ser Phe
      305      310      315      320
Val Gly Glu Ser Lys Glu Gly Thr Val Leu Val Phe Asn Asp Asn Ala
      325      330      335
Asn Thr Pro Gly Pro Asp Gly Lys Pro Leu Gly Thr Ser Asn Ser Ser
      340      345      350
Ser Ile Phe Ile Tyr Ala Asn Asp Phe Ile Ala Gln Asn Leu Thr Ile
      355      360      365
Gln Asn Asp Ser Gly Gln Gly Thr Gly Gln Ala Val Ala Ala Tyr Val
      370      375      380
Arg Ala Asp Arg Leu Tyr Phe Glu Asn Val Gln Phe Leu Gly Tyr Gln
      385      390      395      400

```

Asp Thr Leu Tyr Ala His Thr Gly Arg Gln Tyr Tyr Lys Asn Cys Tyr
 405 410 415
 Val Glu Gly Asp Val Asp Phe Ile Phe Gly Gly Ala Thr Ala Leu Phe
 420 425 430
 Asp Thr Cys His Leu His Ser Lys Arg Thr Gly Ser Lys Leu Thr Ala
 435 440 445
 Ala Ser Thr Asp Gln Val Thr Pro Tyr Gly Tyr Val Phe Leu Asp Ser
 450 455 460
 Lys Ile Thr Ser Asp Glu Gly Val Thr Asn Val His Leu Gly Arg Pro
 465 470 475 480
 Trp Arg Pro Tyr Ser Ala Val Thr Tyr Ile Asn Thr Glu Met Asp Ala
 485 490 495
 Ser Ile Val Pro Asp Gly Trp Asp Asn Trp Gly Lys Val Glu Asn Glu
 500 505 510
 Lys Thr Ala Arg Tyr Ser Glu Tyr Asn Asn Met Gly Pro Gly Ala Asp
 515 520 525
 Pro Lys Lys Arg Asp Pro Trp Thr Thr Gln Leu Thr Pro Glu Glu Ala
 530 535 540
 Asn Gln Tyr Thr Val Gln Asn Val Met Lys Gly Ser Asp Gly Trp Asp
 545 550 555 560
 Pro Glu Arg Ile Gly Ile Ile Pro Leu Ser Pro Leu Ser Ala Pro Ile
 565 570 575
 Ile Ser Leu Asp Gln Arg Asp Ser Ile Val Asn Thr Pro Ser Phe Thr
 580 585 590
 Ile Thr Gly Gln Val Asp Lys Glu Ala Ala Val Ser Val Asn Gly Lys
 595 600 605
 Glu Ile Ser Leu Gln Lys Asp Gly Ser Phe Ser Thr Thr Val Val Leu
 610 615 620
 Asn Asp Gly Leu Asn Thr Ile Thr Val Gly Ala Val Asp Ala Ala Gly
 625 630 635 640
 Asn Gln Ala Ile Pro Ala Val Leu Lys Ile Val Tyr Asp His Glu Lys
 645 650 655
 Pro Val Val Ser Ile Asp Asp Leu Lys Gly Glu Lys Asn Gly Asn His
 660 665 670
 Tyr Asn Val Ile Tyr Asn Pro Leu Pro Ile Thr Gly Lys Leu Asn Glu
 675 680 685
 Ala Gly Thr Val Met Val Asn Gly Glu Lys Val Asn Val Ser Glu Lys
 690 695 700
 Leu Thr Phe Ser Thr Lys Val Ile Leu Lys Pro Gly Leu Asn Asn Ile
 705 710 715 720
 Thr Ile Thr Ala Val Asp Gln Ala Gly Asn Glu Ala Glu Ser Ile Thr
 725 730 735
 Ile Asn Val Val Pro Lys Gly Asn Ala Val Pro Asp Gly Pro Val Lys
 740 745 750
 Ile Ile Lys Ser Glu Thr Thr Asn Ala Asn Thr Val Glu Val Thr Phe
 755 760 765
 Asn Ser Lys Leu Glu Lys Phe Asp Ser Ser Asp Ile Ala Leu Gln Thr
 770 775 780
 Ala Thr Asn Val Trp Ala Ala Leu Asn Pro Gly Leu Lys Gln Leu Met
 785 790 795 800
 Thr Val Glu Ser Ile Thr Thr Lys Val Asn Lys Asp Asn Gln Thr Val
 805 810 815
 Ala Val Ile Lys Thr Lys Glu Ala Phe Gln Glu Asp Gly Thr Ile Thr
 820 825 830
 Leu Pro Lys Val Glu Asp Pro Phe His Ile Gln Tyr Leu Asn Ala Asp
 835 840 845
 Tyr Tyr Thr Gly Asp Arg Thr Gln Asp Ile Lys His Ala Asp Ala Leu
 850 855 860
 Leu Thr Trp Gln Met Asp His Gly Gly Trp Phe Lys Asn Trp Val Glu

865		870		875		880
Lys Tyr Lys Arg Pro Trp Asp Gly Lys Glu Pro Lys Ser Glu Trp Tyr						
		885		890		895
Ser Thr Asn His Gly Glu Leu Gly Thr Ile Asp Asn Asp Ala Thr Thr						
		900		905		910
Asn Glu Ile Leu Phe Leu Ala Leu Met Tyr Lys Glu Thr Gly Asp Ala						
		915		920		925
Arg Tyr Lys Asp Ser Val Leu Lys Gly Ile Asp Phe Leu Leu Glu Met						
		930		935		940
Gln Val Asp Ser Gly Gly Trp Pro Gln Val Tyr Pro Ala Arg Ser Gly						
		945		950		955
Tyr Ser Asp Tyr Val Thr Phe Asn Asp Asn Ala Met Val Arg Val Met						
		965		970		975
Ser Val Leu Thr Met Val Lys Glu Lys Lys Tyr Pro Phe Asn Ser Asn						
		980		985		990
Leu Gly Asp Glu Gln Leu Ser Glu Gln Ile Asp Asp Ala Leu Gly Arg						
		995		1000		1005
Gly Leu Asp Tyr Met Leu Lys Ser Gln Ile Lys Val Asp Gly Glu Val						
		1010		1015		1020
Thr Ala Trp Cys Ala Gln His Asp Pro Val Thr Tyr Glu Pro Lys Gly						
		1025		1030		1035
Ala Arg Ala Tyr Glu His Pro Ser Ile Ser Gly Ser Glu Ser Val Gly						
		1045		1050		1055
Ile Val Gln Tyr Leu Met Ser Leu Pro Asn Pro Ser Thr Glu Val Gln						
		1060		1065		1070
Ala Ala Ile His Gly Ala Leu Asn Trp Phe Glu Glu Ala Lys Leu Ala						
		1075		1080		1085
Gly Thr Lys Tyr Val Ser Gly Asp Pro Asn Gly Gln Tyr Phe Tyr Pro						
		1090		1095		1100
Asp Ala Asn Ser Asn Thr Trp Tyr Arg Phe Tyr Glu Ile Gly Thr Asn						
		1105		1110		1115
Arg Pro Ile Phe Ser Gly Arg Asp Gly Val Ile Lys His Asn Ile Leu						
		1125		1130		1135
Glu Ile Glu Lys Glu Arg Arg Asp Gly Tyr Arg Trp Ala Gly Glu Trp						
		1140		1145		1150
Pro Gln Lys Leu Leu Asn Ile Ala Asn Thr Thr Gly Tyr Tyr Glu Asn						
		1155		1160		1165
Arg Val Tyr Val Glu Val Val Gly Asp Gln Ser Lys Asn Ala Ala Gly						
		1170		1175		1180
Glu Ser Leu Glu Ile Gly Asn Leu Tyr Arg Ile Glu Ala Ser Ala Ser						
		1185		1190		1195
Gly Ser Thr Ser Lys						1200
		1205				

<210> 85
 <211> 1152
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 85

atgtcgggtg gaccaggtgc taatccgaaa gctcgtgttc catggtccaa acagttatcg	60
ggtgttgagg caaagctgtt tcagcgcgag cggttcttca gcctcgtgc ggagcacact	120
tctaagaaaa atgatcagga agtcggcgcg atcgcgtgga aagatgcaca tggaaagccg	180
gatgagtggg atgcgagtgt tgaggcactg cggatggccg ataacgtcgt tctctatcaa	240
cgcgactcag gtggttgcc caagaacatc gacatggcga aggcactcaa cgatcgtgag	300
caggctgcga tctccgcca gaagaaaaag aacgactcca cgatcgacaa tgggtgcgact	360

```

cacacacagt tatcctttct ggcgcgcgtc tatacagcac agcgtcagga gcgacatcgc 420
gagtcgtttt tcaaaggatt ggattactta ctgaatgcgc agtatccaaa tggaggctgg 480
ccgcagtttt atccgaacc cgcgggctat cacaagcaca ttacttacaa cgcaggtgcg 540
atgattggtg tgatgaaggt gctgcgcgat atcgtgcgcg cgaagccttt gtacgctttt 600
gtcgcacgaag ctgcgcgcgc gaaggcgacg agtgcagttg aaaaagggat cgagtgcatt 660
ttgaaaacgc aggtggtggt agatgggcgt cgcactgtgt ggagtgcgca acatgatgaa 720
gtaacgttag cgcagctcc tgcgcgaacc ttgcagttaa cttcgttgag cggcgggtgag 780
agcgtagata tcgttcgatt tttaatgtcg atcaaggatc cgtcgcctaa agtagttgat 840
gcggttgaat cggcgggttaa gtggtttgag caatcggagt taaaaggcgt gaagtgggtt 900
aagaaggcgg acgcttctaa acctggcggg tttgattgcg tcgtagttaa ggatccggag 960
agctcggttt ggcgcgcgtt ttatgagatt ggcacgaacc ggcgcatctt ttccgggcgc 1020
gatggagtgg tcaaatacga cgtggcgcag atcgaacacg agcggcggac gaattatgag 1080
tggtacgttg atgaagcagc caagctgctg aagaaagagt atccggcctg gcggaaaaga 1140
acatctctgt ga 1152

```

<210> 86

<211> 383

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<221> DOMINIO

<222> (1)...(383)

<223> Dominio catalítico

15 <400> 86

```

Met Ser Val Gly Pro Gly Ala Asn Pro Lys Ala Arg Val Pro Trp Ser
1      5      10      15
Lys Gln Leu Ser Gly Val Glu Ala Lys Leu Phe Gln Arg Glu Arg Phe
20      25      30
Phe Ser Leu Ala Ala Glu His Thr Ser Lys Lys Asn Asp Gln Glu Val
35      40      45
Gly Ala Ile Ala Trp Lys Asp Ala His Gly Lys Pro Asp Glu Trp Tyr
50      55      60
Ala Ser Val Glu Ala Leu Arg Met Ala Asp Asn Val Val Leu Tyr Gln
65      70      75      80
Arg Asp Ser Gly Gly Trp Pro Lys Asn Ile Asp Met Ala Lys Ala Leu
85      90      95
Asn Asp Arg Glu Gln Ala Ala Ile Leu Arg Gln Lys Lys Lys Asn Asp
100     105     110
Ser Thr Ile Asp Asn Gly Ala Thr His Thr Gln Leu Ser Phe Leu Ala
115     120     125
Arg Val Tyr Thr Ala Gln Arg Gln Glu Arg His Arg Glu Ser Phe Phe
130     135     140
Lys Gly Leu Asp Tyr Leu Leu Asn Ala Gln Tyr Pro Asn Gly Gly Trp
145     150     155     160
Pro Gln Phe Tyr Pro Asn Pro Thr Gly Tyr His Lys His Ile Thr Tyr
165     170     175
Asn Asp Gly Ala Met Ile Gly Val Met Lys Val Leu Arg Asp Ile Ala
180     185     190
Ala Ala Lys Pro Leu Tyr Ala Phe Val Asp Glu Ala Arg Arg Ala Lys
195     200     205
Ala Thr Ser Ala Val Glu Lys Gly Ile Glu Cys Ile Leu Lys Thr Gln
210     215     220
Val Val Val Asp Gly Arg Arg Thr Val Trp Ser Ala Gln His Asp Glu
225     230     235     240
Val Thr Leu Ala Pro Ala Pro Ala Arg Thr Phe Glu Leu Thr Ser Leu

```

				245				250					255				
Ser	Gly	Gly	Glu	Ser	Val	Asp	Ile	Val	Arg	Phe	Leu	Met	Ser	Ile	Lys		
			260					265					270				
Asp	Pro	Ser	Pro	Lys	Val	Val	Asp	Ala	Val	Glu	Ser	Ala	Val	Lys	Trp		
		275					280					285					
Phe	Glu	Gln	Ser	Glu	Leu	Lys	Gly	Val	Lys	Trp	Val	Lys	Lys	Ala	Asp		
	290					295					300						
Ala	Ser	Lys	Pro	Gly	Gly	Phe	Asp	Cys	Val	Val	Val	Lys	Asp	Pro	Glu		
305				310						315					320		
Ser	Ser	Val	Trp	Ala	Arg	Phe	Tyr	Glu	Ile	Gly	Thr	Asn	Arg	Pro	Ile		
				325						330				335			
Phe	Ser	Gly	Arg	Asp	Gly	Val	Val	Lys	Tyr	Asp	Val	Ala	Gln	Ile	Glu		
			340					345					350				
His	Glu	Arg	Arg	Thr	Asn	Tyr	Glu	Trp	Tyr	Val	Asp	Glu	Ala	Ala	Lys		
		355				360						365					
Leu	Leu	Lys	Lys	Glu	Tyr	Pro	Ala	Trp	Arg	Lys	Arg	Thr	Ser	Leu			
	370					375					380						

<210> 87
 <211> 1698
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 87

atgtctacta	caaaatgttt	taacacagcc	ccaggtttta	ccctgaaagc	agtcgcagca	60
gcagtggcga	tgtttgccag	ttcttcagta	ttcgcagcgg	ctacaggtgg	tttttccacg	120
actgatggcg	gtgcggcaag	cggctcgcaa	tccttcacgg	cggccaacct	tgaccagctc	180
aacaccattg	ttgccaatgc	gaagagtggc	ggttaccocg	ttgtgattac	ctataccggg	240
aatgaagaca	gcttgattaa	ccagatgatc	aaagaaccaca	ccgtggattc	ttcaggcaac	300
tgcccgaaac	cacgctggag	tgaaacctac	cgcaaggtag	aaattaagga	gatgaccaa	360
ggtgtcacca	tcacgggtgc	caatggttct	tcggcaaaact	tcggtattgt	ggtgaacaag	420
tcacgcaatg	tgattatccg	caacatgaaa	atcgggtgoc	tgcccggtgc	cagcaacgac	480
gcggatatga	ttcgtatcga	tagcggcact	aacgtatggg	ttgaccacaa	cgaattgttc	540
gcggtgaaca	acgaatgtaa	aggttcaccc	gatggcgatt	tgaccttcga	aagcgccatc	600
gacatcaaga	aagattcaca	caacatcacc	gtgtcttaca	acctgattcg	cgacagcaaa	660
aaagtggggc	ttgatgggtc	cagcagcagc	gatatcgccg	gtggccgcga	gatcactttc	720
caccacaaca	tttacaacaa	cgtgaatgca	cgtttgcccgt	tgcaacgcgg	tggttggaag	780
cacatgtata	acaacctgta	cgacggcatt	accggttcog	gtatcaacgt	acgtcaggcc	840
ggttatgcgt	tgattgaaag	caactgggtc	caaaatgcgg	ttaaccoggt	gacttgccgt	900
tacgacagca	gcaactgcgg	tttctgggat	ctgcgcaata	acaacgtgaa	gtcgccagca	960
gatttcgcga	cctataacat	cacctggacc	agcggcgcca	ctattgatgc	aaccaactgg	1020
acgaccaccg	ctccgttccc	gatcagcatt	ccttacagct	actcgccggg	gtctccacag	1080
tgcgatgaag	acaagttggc	cagcgttgcg	ggtgtgggta	aaaacgggtg	agttctgaac	1140
tcacagtggt	gtggtggaag	cagctctgtt	ccatcatcaa	gctcagtcgc	tactacttcc	1200
aaatcatcca	gctcggtagc	aaccagcaag	tcagctccg	tcgctacgac	gtccagtaag	1260
tcacccagct	cggtagtgc	atcatcatca	agctcaagtt	cagtgggttaa	taacggcagc	1320
atcgcggtta	ccgccactgc	taccggcaat	agcattgtcc	tgagctggtc	gccgaacaa	1380
ctgacactgg	gcacccagga	ggtgtatcgc	gataccgatt	cagacccaag	tgcccggtgtg	1440
cgtattgctg	ccctgagttc	cagcactcgc	atgtacaccg	atgccactgc	atcgccgggg	1500
caaacgttct	actactggat	caaaaacacc	accaacgggtg	taaccaccaa	ttccaatgcg	1560
gcttcggcgg	caattggcga	tcagctcgc	gccattcgcg	catgcgcagg	aaaccgagga	1620
agtggcgctc	gcaccagtcg	cgcagtttcg	actgggtcaa	atcctcgtgg	gcctgcccgt	1680
agccatccca	gagcttga					1698

<210> 88
 <211> 565
 <212> PRT
 <213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

<221> SEÑAL

5 <222> (1)...(32)

<221> DOMINIO

<222> (33)...(375)

<223> Dominio catalítico

10

<400> 88

```

Met Ser Thr Thr Lys Cys Phe Asn Thr Ala Pro Gly Phe Thr Leu Lys
 1      5      10      15
Ala Val Ala Ala Val Ala Met Phe Ala Gly Ser Ser Val Phe Ala
      20      25      30
Ala Ala Thr Gly Gly Phe Ser Thr Thr Asp Gly Gly Ala Ala Ser Gly
      35      40      45
Ser Gln Ser Phe Thr Ala Ala Asn Leu Asp Gln Leu Asn Thr Ile Val
      50      55      60
Ala Asn Ala Lys Ser Gly Gly Tyr Pro Val Val Ile Thr Tyr Thr Gly
      65      70      75      80
Asn Glu Asp Ser Leu Ile Asn Gln Met Ile Lys Asp His Thr Val Asp
      85      90      95
Ser Ser Gly Asn Cys Pro Asn Pro Arg Trp Ser Glu Thr Tyr Arg Lys
      100      105      110
Val Glu Ile Lys Glu Met Thr Lys Gly Val Thr Ile Ile Gly Ala Asn
      115      120      125
Gly Ser Ser Ala Asn Phe Gly Ile Val Val Asn Lys Ser Ser Asn Val
      130      135      140
Ile Ile Arg Asn Met Lys Ile Gly Ala Leu Ala Gly Ala Ser Asn Asp
      145      150      155      160
Ala Asp Met Ile Arg Ile Asp Ser Gly Thr Asn Val Trp Val Asp His
      165      170      175
Asn Glu Leu Phe Ala Val Asn Asn Glu Cys Lys Gly Ser Pro Asp Gly
      180      185      190
Asp Leu Thr Phe Glu Ser Ala Ile Asp Ile Lys Lys Asp Ser His Asn
      195      200      205
Ile Thr Val Ser Tyr Asn Leu Ile Arg Asp Ser Lys Lys Val Gly Leu
      210      215      220
Asp Gly Ser Ser Ser Ser Asp Ile Ala Gly Gly Arg Glu Ile Thr Phe
      225      230      235      240
His His Asn Ile Tyr Lys Asn Val Asn Ala Arg Leu Pro Leu Gln Arg
      245      250      255
Gly Gly Trp Thr His Met Tyr Asn Asn Leu Tyr Asp Gly Ile Thr Gly
      260      265      270
Ser Gly Ile Asn Val Arg Gln Ala Gly Tyr Ala Leu Ile Glu Ser Asn
      275      280      285
Trp Phe Gln Asn Ala Val Asn Pro Val Thr Cys Arg Tyr Asp Ser Ser
      290      295      300
Asn Cys Gly Phe Trp Asp Leu Arg Asn Asn Asn Val Lys Ser Pro Ala
      305      310      315      320
Asp Phe Ala Thr Tyr Asn Ile Thr Trp Thr Ser Gly Gly Thr Ile Asp
      325      330      335
Ala Thr Asn Trp Thr Thr Thr Ala Pro Phe Pro Ile Ser Ile Pro Tyr
      340      345      350
Ser Tyr Ser Pro Val Ser Pro Gln Cys Val Lys Asp Lys Leu Ala Ser

```


Val	Ala	Gly	Val	Gly	Lys	Asn	Gly	Ala	Val	Leu	Asn	Ser	Ser	Val	Cys
370						375					380				
Gly	Gly	Ser	Ser	Ser	Val	Pro	Ser	Ser	Ser	Ser	Val	Ala	Thr	Thr	Ser
385					390					395					400
Lys	Ser	Ser	Ser	Ser	Val	Ala	Thr	Ser	Lys	Ser	Ser	Ser	Val	Ala	Thr
				405					410					415	
Thr	Ser	Ser	Lys	Ser	Ser	Ser	Ser	Val	Val	Pro	Ser	Ser	Ser	Ser	Ser
			420					425					430		
Ser	Ser	Val	Val	Asn	Asn	Gly	Ser	Ile	Ala	Leu	Thr	Ala	Thr	Ala	Thr
		435					440					445			
Gly	Asn	Ser	Ile	Val	Leu	Ser	Trp	Ser	Pro	Asn	Asn	Leu	Thr	Leu	Gly
450					455					460					
Thr	Gln	Glu	Val	Tyr	Arg	Asp	Thr	Asp	Ser	Asp	Pro	Ser	Gly	Arg	Val
465				470						475					480
Arg	Ile	Ala	Ala	Leu	Ser	Ser	Ser	Thr	Arg	Met	Tyr	Thr	Asp	Ala	Thr
				485					490					495	
Ala	Ser	Ala	Gly	Gln	Thr	Phe	Tyr	Tyr	Trp	Ile	Lys	Asn	Thr	Thr	Asn
			500				505						510		
Gly	Val	Thr	Thr	Asn	Ser	Asn	Ala	Ala	Ser	Ala	Ala	Ile	Gly	Asp	Ala
		515					520					525			
Ala	Arg	Ala	Ile	Arg	Ala	Cys	Ala	Gly	Asn	Arg	Gly	Ser	Gly	Ala	Arg
530					535					540					
Thr	Ser	Arg	Ala	Val	Ser	Thr	Gly	Ser	Asn	Pro	Arg	Gly	Pro	Ala	Gly
545					550					555					560
Ser	His	Pro	Arg	Ala											
				565											

<210> 89
 <211> 1377
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 89

atgacgacgc	gacgcgaatt	catttcgagat	cttttggttg	gcggcgtagt	ggtcgctggt	60
gcaccgcgtt	tcttgccgtt	ttcttcgggtg	gcgagtcctg	gggaaacggt	gatgccttcg	120
atcctcgaac	gcatacaagcc	accgcgtttt	ccgaaacgca	cgtgctatct	caaccgggtt	180
ggagcaaaa	gcgacgggca	aactgattgc	acttcagctt	ttcgacgcgc	aatcgatcag	240
tggtcgaaa	cgggcggttg	caaagtgatc	gttccgcagg	gaatgtatct	caccggcgca	300
attcacttga	agagcaacgt	caatctcgag	atctccgaag	gcgcgacgat	caagttcagt	360
caaaaccgca	aagactatct	cccgttggtt	ttttcgcggt	gggaaggcgt	cgaagtattc	420
aactactcac	ctttcatcta	cgcatttgaa	cagcagaaca	tcgcgatcac	gggcaagggc	480
acgctcgatg	ggcagagtga	taacgaacac	tggtggccat	ggaacggacg	cgccaggtac	540
ggttggaag	aagggatgag	ccaccagcgt	ccggatcgaa	acgcgctctt	tgcatggcg	600
gaaaaaggtg	tttcggttcg	cgaacgtggt	ttcggcgagg	gtcattactt	aaggccgcag	660
ttcattcagc	cgtatcgctg	ccagaacgta	ttgatcgacg	gagttacgat	acgaaactcg	720
ccgatgtggg	aaattcatcc	ggtgctgtgc	cggaatgtca	tcgtgcaaaa	cgtgcacatt	780
aacagtcattg	gaccaaacia	cgatggctgc	aatcccaggt	cgtgcactga	tgtgctgatt	840
aagaactggt	acttcgacac	tggcgacgac	tgtatcgcg	tcaaatacagg	acgcaacgcg	900
gacggccggc	ggcttaaagc	gccgacagag	aacgtgatcg	tgcaagactg	tcaaataaaa	960
gatggacacg	gcgggatcac	tgtcggcagt	gagatctcag	gcgggtgtgag	aaatctgttt	1020
gcggagaact	gccggcttga	tagtccaaac	ctggaccatg	ctttgcgggt	taagaacaac	1080
gcgatgcgtg	gagggctgct	cgagaatttg	cacttccgaa	acatcgaagt	tggtcaggtg	1140
gcgatgcag	tgatcacgat	cgattttaat	tacgaggaag	gcgcgaaaag	atcggtcacg	1200
ccggtggttc	gtgactacac	tgtcgatggt	ttgcgcagca	cgcgaaagca	atcgcgctc	1260
gacgttcaag	gtctgtcggg	cgcgccgatc	gtaaatctgc	gtctgacgaa	ttgcacgttc	1320
gacaatgttg	ccgaagggaa	cgtcgtgaag	aatgttaagg	acgcgacaat	tcaaaaa	1377

<210> 90

<211> 459
 <212> PRT
 <213> Desconocido

5 <220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(31)

10 <221> DOMINIO
 <222> (32)...(459)
 <223> Dominio catalítico

15 <400> 90

Met	Thr	Thr	Arg	Arg	Glu	Phe	Ile	Arg	Asp	Leu	Leu	Val	Gly	Gly	Val
1				5					10					15	
Val	Val	Ala	Val	Ala	Pro	Arg	Phe	Leu	Ala	Phe	Ser	Ser	Val	Ala	Ser
		20						25					30		
Pro	Trp	Glu	Thr	Val	Met	Pro	Ser	Ile	Leu	Glu	Arg	Ile	Lys	Pro	Pro
		35					40					45			
Arg	Phe	Pro	Lys	Arg	Thr	Cys	Tyr	Leu	Asn	Arg	Phe	Gly	Ala	Lys	Gly
	50					55					60				
Asp	Gly	Gln	Thr	Asp	Cys	Thr	Ser	Ala	Phe	Arg	Arg	Ala	Ile	Asp	Gln
65					70					75				80	
Cys	Ser	Lys	Ala	Gly	Gly	Gly	Lys	Val	Ile	Val	Pro	Gln	Gly	Met	Tyr
				85					90				95		
Leu	Thr	Gly	Ala	Ile	His	Leu	Lys	Ser	Asn	Val	Asn	Leu	Glu	Ile	Ser
			100					105					110		
Glu	Gly	Ala	Thr	Ile	Lys	Phe	Ser	Gln	Asn	Pro	Lys	Asp	Tyr	Leu	Pro
		115					120					125			
Val	Val	Phe	Ser	Arg	Trp	Glu	Gly	Val	Glu	Val	Phe	Asn	Tyr	Ser	Pro
		130				135					140				
Phe	Ile	Tyr	Ala	Phe	Glu	Gln	Gln	Asn	Ile	Ala	Ile	Thr	Gly	Lys	Gly
145					150					155				160	
Thr	Leu	Asp	Gly	Gln	Ser	Asp	Asn	Glu	His	Trp	Trp	Pro	Trp	Asn	Gly
				165					170					175	
Arg	Ala	Arg	Tyr	Gly	Trp	Lys	Glu	Gly	Met	Ser	His	Gln	Arg	Pro	Asp
			180					185					190		
Arg	Asn	Ala	Leu	Phe	Ala	Met	Ala	Glu	Lys	Gly	Val	Ser	Val	Arg	Glu
		195					200					205			
Arg	Val	Phe	Gly	Glu	Gly	His	Tyr	Leu	Arg	Pro	Gln	Phe	Ile	Gln	Pro
		210				215					220				
Tyr	Arg	Cys	Gln	Asn	Val	Leu	Ile	Asp	Gly	Val	Thr	Ile	Arg	Asn	Ser
225					230					235				240	
Pro	Met	Trp	Glu	Ile	His	Pro	Val	Leu	Cys	Arg	Asn	Val	Ile	Val	Gln
				245					250					255	
Asn	Val	His	Ile	Asn	Ser	His	Gly	Pro	Asn	Asn	Asp	Gly	Cys	Asn	Pro
			260					265					270		
Glu	Ser	Cys	Thr	Asp	Val	Leu	Ile	Lys	Asn	Cys	Tyr	Phe	Asp	Thr	Gly
		275					280					285			
Asp	Asp	Cys	Ile	Ala	Val	Lys	Ser	Gly	Arg	Asn	Ala	Asp	Gly	Arg	Arg
		290				295					300				
Leu	Lys	Ala	Pro	Thr	Glu	Asn	Val	Ile	Val	Gln	Asp	Cys	Gln	Met	Lys
305					310					315				320	
Asp	Gly	His	Gly	Gly	Ile	Thr	Val	Gly	Ser	Glu	Ile	Ser	Gly	Gly	Val

Arg	Asn	Leu	Phe	Ala	Glu	Asn	Cys	Arg	Leu	Asp	Ser	Pro	Asn	Leu	Asp
			340					345					350		
His	Ala	Leu	Arg	Val	Lys	Asn	Asn	Ala	Met	Arg	Gly	Gly	Leu	Leu	Glu
		355					360					365			
Asn	Leu	His	Phe	Arg	Asn	Ile	Glu	Val	Gly	Gln	Val	Ala	His	Ala	Val
	370					375					380				
Ile	Thr	Ile	Asp	Phe	Asn	Tyr	Glu	Glu	Gly	Ala	Lys	Gly	Ser	Phe	Thr
385					390					395					400
Pro	Val	Val	Arg	Asp	Tyr	Thr	Val	Asp	Gly	Leu	Arg	Ser	Thr	Arg	Ser
				405					410					415	
Lys	Tyr	Ala	Leu	Asp	Val	Gln	Gly	Leu	Ser	Gly	Ala	Pro	Ile	Val	Asn
			420					425					430		
Leu	Arg	Leu	Thr	Asn	Cys	Thr	Phe	Asp	Asn	Val	Ala	Glu	Gly	Asn	Val
		435					440					445			
Val	Lys	Asn	Val	Lys	Asp	Ala	Thr	Ile	Gln	Lys					
	450					455									

<210> 91
<211> 1125
<212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<400> 91

gtgttcctag	gtaataacgg	cggcagcttg	agttgcgtcc	aatataattgt	gatttgtgaaa	60
ggaccccggtg	gacctcgacc	gccggtgaaa	ccggccgtcc	aggcgcccg	tagggttacc	120
tggagcgc	ccctagtcca	gcgcccgaa	tggtagcgga	gtgacgaagc	gatccgcac	180
gcggacaacg	tcctcctcta	ccagcgcaac	accggcggtt	ggccgaagga	catagatatg	240
gccgagccca	tcccggaaaca	caggaaagtcc	tttttctca	ccgagaagga	gcggaccgat	300
gactcgacca	tcgacaacgg	tgccaccgtg	accagctca	agtatctcgc	ccgcgtctac	360
aaggcgacca	ggctggaacg	gttcaaggag	ggcttctca	aaggctctga	ctacctcttg	420
gccgcccgagt	accggaacgg	cggctggccc	cagtattatc	ctaacttgag	gggctactac	480
gcgaacatca	cttataacga	caatgccatg	gtgaacgtgc	tcacctctct	ccagagcatc	540
gccaaaaagg	ccccggagta	cgacttcgtc	gacccggcgc	gccgggagaa	ggccgcccgg	600
gccgtggcga	aagggatcga	ctgcactctc	aagaccaga	tccgtgtcaa	tggaaaactt	660
accgcctgg	gcgccagca	tgaacccaag	acgttggcgc	ccgcgccggc	ccgttcgtat	720
gagcttgagt	ccatcagcgg	tttcgagagc	gtcgggatcg	tccggttctt	aatgagcctc	780
gagaatccga	gcccgaaagt	catcgaggcg	gtagaggccg	ccgtgaaatg	gttcgaggag	840
gtcaagctta	ccgggatcaa	ggtggtcgag	aaaccgcacc	cgtcccttcc	gggcggttac	900
gaccgcgtg	tggtcgaaga	ccccaacggc	ccgccatct	gggcccgggt	ctacgagatc	960
ggccaacaac	gtcccttctt	ctcgccggcg	gatggtatca	aaaaatacag	cctggcggag	1020
atcgaacaag	aacgcgggtt	cgtttaactc	tggtagacca	atgccccggc	ctacctcatc	1080
gagaaggagt	atccgtctctg	gcggggccaaa	cacctacca	agtaa		1125

<210> 92
<211> 374
<212> PRT
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental
<221> DOMINIO
<222> (1)...(374)

<400> 92

Met Val Leu Gly Asn Asn Gly Gly Ser Leu Ser Cys Val Gln Tyr Ile
1 5 10 15
Val Ile Val Lys Gly Pro Gly Gly Pro Arg Pro Pro Val Lys Pro Ala
20 25 30
Val Gln Ala Pro Val Arg Val Thr Trp Ser Ala Ser Leu Val Gln Arg
35 40 45
Pro Glu Trp Tyr Gly Ser Asp Glu Ala Ile Arg Ile Ala Asp Asn Val
50 55 60
Leu Leu Tyr Gln Arg Asn Thr Gly Gly Trp Pro Lys Asp Ile Asp Met
65 70 75 80
Ala Glu Pro Ile Pro Glu His Arg Lys Ser Phe Phe Leu Thr Glu Lys
85 90 95
Glu Arg Thr Asp Asp Ser Thr Ile Asp Asn Gly Ala Thr Val Thr Gln
100 105 110
Leu Lys Tyr Leu Ala Arg Val Tyr Lys Ala Thr Arg Leu Glu Arg Phe
115 120 125
Lys Glu Gly Phe Leu Lys Gly Leu Asp Tyr Leu Leu Ala Ala Gln Tyr
130 135 140
Pro Asn Gly Gly Trp Pro Gln Tyr Tyr Pro Asn Leu Arg Gly Tyr Tyr
145 150 155 160
Ala Asn Ile Thr Tyr Asn Asp Asn Ala Met Val Asn Val Leu Thr Leu
165 170 175
Leu Gln Ser Ile Ala Lys Lys Ala Pro Glu Tyr Asp Phe Val Asp Pro
180 185 190
Ala Arg Arg Glu Lys Ala Ala Arg Ala Val Ala Lys Gly Ile Asp Cys
195 200 205
Ile Leu Lys Thr Gln Ile Arg Val Asn Gly Lys Leu Thr Ala Trp Cys
210 215 220
Ala Gln His Asp Ala Lys Thr Leu Ala Pro Ala Pro Ala Arg Ser Tyr
225 230 235 240
Glu Leu Glu Ser Ile Ser Gly Phe Glu Ser Val Gly Ile Val Arg Phe
245 250 255
Leu Met Ser Leu Glu Asn Pro Ser Pro Lys Val Ile Glu Ala Val Glu
260 265 270
Ala Ala Val Lys Trp Phe Glu Glu Val Lys Leu Thr Gly Ile Lys Val
275 280 285
Val Glu Lys Pro Asp Pro Ser Leu Pro Gly Gly Tyr Asp Arg Val Val
290 295 300
Val Glu Asp Pro Asn Ala Pro Pro Ile Trp Ala Arg Phe Tyr Glu Ile
305 310 315 320
Gly Thr Asn Arg Pro Phe Phe Cys Gly Arg Asp Gly Ile Lys Lys Tyr
325 330 335
Ser Leu Ala Glu Ile Glu His Glu Arg Arg Val Gly Tyr Ser Trp Tyr
340 345 350
Thr Asn Ala Pro Ala Tyr Leu Ile Glu Lys Glu Tyr Pro Leu Trp Arg
355 360 365
Ala Lys His Pro Thr Lys
370

<210> 93
<211> 1062
5 <212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

10 <400> 93

gtggatccaa agaattggaa cccgaaaaaa gccgacgatt catggctcga aaagacgaaa 60

```

cccgattacc ggctgggtctc ctggcgcgac gtttttagatc aaactcagct ctgggtacgcg 120
gtcgcacgaag cgacgcgcac cgccaaccag gttttgctct ttcagcgcgga taacggcggc 180
tgggaaaaaa acgtcgacat ggcggcgatg ctactcaag ccgaacgaga aaaactcgtc 240
aaagaaaaat ctacacacga tacgaccatc gacaacggcg cgacgaccac gcagctgcgt 300
tatctggcaa aagtcacac ggcgaaaaac atcgaagctc ataaacagtc gtttctcaag 360
ggattggatt ttctgctcgc gatgcagtat gaaaacggag gatttccgca atattatcct 420
ttgaaaaacg attattcgcg cgagattact ttcaacgacg acgcgatgat caatgttctt 480
aaattgctgc gcgacgtggc aaaaaagaag gaagattatt tattcgtcga cgaagaccgg 540
cgcgccagag cggaaggcgc ggtcgaaaaa ggcgtccgcc tgatcttgaa aacacaggtc 600
gccatcgacg gcaaaaaaac gatctgggcg gcgcagtacg acgaaaaacac tttgaaacgg 660
gcaaatcgca gaaagtttga gccgcctcgc ctgcgttcgc gcgaatcggg cagcgtgggc 720
agatttttga tgctcgacgc caaacccgac gaggaaaaaa tcggagcgat cgaatcggcg 780
atcgaatggg ttcaaaaaaa caaactgagc ggcattcgct gggaatcgaa aagcggagaa 840
aacctggctg tcaaagacaa agcggcgccg ccgatctggg gaagggttta tcaattcgaa 900
accatgcgcc ccatttttat cgggcgcgac gcggtgatto gctacgatgt catgcaaadc 960
gaagccgaac gccgcaacgg ctacggctgg tacacgaacg agccgaacga gcttttggac 1020
aaagattatc cgaaatggaa agageaaaatt aagaaaaatt ag 1062

```

<210> 94

<211> 353

5 <212> PRT

<213> Desconocido

<220>

10 <223> Obtenido de una muestra ambiental

<221> DOMINIO

<222> (1)...(353)

<400> 94

15

```

Met Asp Pro Lys Asn Trp Asn Pro Lys Lys Ala Asp Asp Ser Trp Leu
 1      5      10      15
Glu Lys Thr Lys Pro Asp Tyr Arg Leu Val Ser Trp Arg Asp Val Leu
      20      25      30
Asp Gln Thr Gln Leu Trp Tyr Ala Val Asp Glu Ala Thr Arg Ile Ala
      35      40      45
Asn Gln Val Leu Leu Phe Gln Arg Asp Asn Gly Gly Trp Glu Lys Asn
      50      55      60
Val Asp Met Ala Ala Met Leu Thr Gln Ala Glu Arg Glu Lys Leu Val
      65      70      75      80
Lys Glu Lys Ser His Thr Asp Thr Thr Ile Asp Asn Gly Ala Thr Thr
      85      90      95
Thr Gln Leu Arg Tyr Leu Ala Lys Val Ile Thr Ala Lys Asn Ile Glu
      100      105      110
Ala His Lys Gln Ser Phe Leu Lys Gly Leu Asp Phe Leu Leu Ala Met
      115      120      125
Gln Tyr Glu Asn Gly Gly Phe Pro Gln Tyr Tyr Pro Leu Lys Asn Asp
      130      135      140
Tyr Ser Arg Glu Ile Thr Phe Asn Asp Asp Ala Met Ile Asn Val Leu
      145      150      155      160
Lys Leu Leu Arg Asp Val Ala Lys Lys Lys Glu Asp Tyr Leu Phe Val
      165      170      175
Asp Glu Asp Arg Arg Ala Arg Ala Glu Gly Ala Val Glu Lys Gly Val
      180      185      190
Arg Leu Ile Leu Lys Thr Gln Val Ala Ile Asp Gly Lys Lys Thr Ile
      195      200      205
Trp Ala Ala Gln Tyr Asp Glu Asn Thr Leu Lys Pro Ala Asn Ala Arg
      210      215      220
Lys Phe Glu Pro Ala Ser Leu Ala Ser Arg Glu Ser Val Ser Val Val

```

225	Arg	Phe	Leu	Met	Leu	Asp	Ala	Lys	Pro	Asp	Glu	Glu	Lys	Ile	Gly	Ala
					245					250					255	
	Ile	Glu	Ser	Ala	Ile	Glu	Trp	Phe	Gln	Lys	Asn	Lys	Leu	Ser	Gly	Ile
				260					265					270		
	Arg	Trp	Glu	Ser	Lys	Ser	Gly	Glu	Asn	Leu	Val	Val	Lys	Asp	Lys	Ala
			275					280					285			
	Ala	Pro	Pro	Ile	Trp	Gly	Arg	Phe	Tyr	Gln	Phe	Glu	Thr	Met	Arg	Pro
		290					295					300				
	Ile	Phe	Ile	Gly	Arg	Asp	Ala	Val	Ile	Arg	Tyr	Asp	Val	Met	Gln	Ile
305					310					315					320	
	Glu	Ala	Glu	Arg	Arg	Asn	Gly	Tyr	Gly	Trp	Tyr	Thr	Asn	Glu	Pro	Asn
				325					330					335		
	Glu	Leu	Leu	Asp	Lys	Asp	Tyr	Pro	Lys	Trp	Lys	Glu	Lys	Ile	Lys	Lys
			340					345					350			
	Asn															

<210> 95
 <211> 1074
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 95

atgacgctac	ccgttggttc	cctgcgcgta	ctgctggcgc	tgctggccac	gtcgccggtc	60
gcctgcgcgg	gcgcgcgcgc	acccgcgact	gcgaccgatc	cggtcgccga	gaacatgctg	120
cttctgcaga	ccgcctccgg	cggctggctc	aagcactacc	gcgagaagaa	ggctcgactac	180
gcgcgcgact	acgacgccgc	cgagcgcgcc	gcgctgcgcg	cgcccgaccg	gcatgacgat	240
gcgacgatcg	acaacaaggc	cacgaccacc	gagatcgcc	acctgggtga	ggcacatgcc	300
aggacgggca	atccggccta	cctcgacggc	gcgcgcgcgc	gcgtcgagta	cctgctgcgc	360
gcccagttac	cgaacggcgc	ctggccgcag	ttctaccccg	accattcgtc	ctaccggcac	420
cagatcacgc	tcaacgaaga	tgcgatgggt	cacgccaatc	ccgtgctgca	ggacatcgcc	480
gcgggccgca	acggcatgca	ggtgctggcg	ccggagttcg	gcgtccgcgc	cgccgcggcc	540
gcgcagcgcg	gcatcggaag	cctgctcgag	ttgcaggtgc	ggatcgccgg	ggtgccgacg	600
atatgggccc	cgcagtagca	cgagaccagc	ctgcaaccgg	ccaaggcccg	cgcgtacgaa	660
ctgccttggc	tggccgtggc	cgaatcggtc	ggcgtgggtc	gcctgctgat	gcgccagccg	720
gcgcctgatg	cgcgcacggc	cgccgcgata	gaggcgggcg	ccgactggct	ggaggcgcac	780
cgcctgccgg	acctcgccct	ggaacgcata	gaagcccccg	ccgaggaaac	cggcaaggac	840
gtccgcgtcg	tggccagacc	gggcgcgctc	ttgtggggcg	gcttctacga	cctcgagcgg	900
caggtgccgc	tggtcgctga	tgcgaacagc	cgcccggtgc	ccttcgccga	gcttcccaac	960
gagcgtcgta	ccggctatgg	ctggtatggc	acctggccgg	aaaagctgct	ggcacaggaa	1020
ctcccgcgct	ggcgcaaggc	ccatgcggcc	agcgcgggcg	ctccggcccg	ttga	1074

15 <210> 96
 <211> 357
 <212> PRT
 <213> Desconocido

20 <220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(31)

25 <221> DOMINIO
 <222> (32)...(357)
 <223> Dominio catalítico

30 <400> 96

```

Met Thr Leu Pro Val Val Ser Leu Arg Val Leu Leu Ala Leu Leu Ala
 1          5          10          15
Thr Ser Pro Val Ala Cys Ala Gly Ala Ala Ala Pro Ala Thr Ala Thr
          20          25          30
Asp Pro Val Ala Glu Asn Met Leu Leu Leu Gln Thr Ala Ser Gly Gly
          35          40          45
Trp Ser Lys His Tyr Arg Glu Lys Lys Val Asp Tyr Ala Arg Asp Tyr
          50          55          60
Asp Ala Ala Glu Arg Ala Ala Leu Arg Ala Pro Asp Arg His Asp Asp
65          70          75          80
Ala Thr Ile Asp Asn Lys Ala Thr Thr Thr Glu Ile Ala Tyr Leu Val
          85          90          95
Gln Ala His Ala Arg Thr Gly Asn Pro Ala Tyr Leu Asp Gly Ala Arg
          100          105          110
Arg Gly Val Glu Tyr Leu Leu Arg Ala Gln Tyr Pro Asn Gly Gly Trp
          115          120          125
Pro Gln Phe Tyr Pro Asp His Ser Ser Tyr Arg His Gln Ile Thr Leu
          130          135          140
Asn Asp Asp Ala Met Val His Ala Ile Thr Val Leu Gln Asp Ile Ala
145          150          155          160
Ala Gly Arg Asn Gly Met Gln Val Leu Ala Pro Glu Phe Gly Val Arg
          165          170          175
Ala Ala Ala Ala Ala Gln Arg Gly Ile Gly Asn Leu Leu Glu Leu Gln
          180          185          190
Val Arg Ile Ala Gly Val Pro Thr Ile Trp Ala Ala Gln Tyr Asp Glu
          195          200          205
Thr Ser Leu Gln Pro Ala Lys Ala Arg Ala Tyr Glu Leu Pro Ser Leu
          210          215          220
Ala Val Ala Glu Ser Val Gly Val Val Arg Leu Leu Met Arg Gln Pro
225          230          235          240
Ala Pro Asp Ala Arg Thr Val Ala Ala Ile Glu Ala Ala Ala Asp Trp
          245          250          255
Leu Glu Ala His Arg Leu Pro Asp Leu Ala Leu Glu Arg Ile Glu Ala
          260          265          270
Pro Ala Glu Glu Thr Gly Lys Asp Val Arg Val Val Ala Arg Pro Gly
          275          280          285
Ala Ser Leu Trp Ala Arg Phe Tyr Asp Leu Glu Arg Gln Val Pro Leu
          290          295          300
Phe Val Asp Arg Asn Ser Arg Pro Val Pro Phe Ala Glu Leu Pro Asn
305          310          315          320
Glu Arg Arg Thr Gly Tyr Gly Trp Tyr Gly Thr Trp Pro Glu Lys Leu
          325          330          335
Leu Ala Gln Glu Leu Pro Arg Trp Arg Lys Val His Ala Ala Ser Ala
          340          345          350
Gly Ala Pro Ala Arg
          355

```

<210> 97

<211> 2097

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<400> 97

```

ttgaacgcgc cgggcagccg gcggttcgcg caactcgctg tcgcgggatct gcggcggtctg      60
gtgcccgcgc tggcgccctt ctttcgcgac gagccgctgg cgggaggagt cgcgcgcctc      120
cagcgcagcg tcgatgcgat cgtcgccgcg gacggcaccg gacagtttgc gacgggtgcag      180
gaggcgatca acgcccgcgc gcagaacacc agcagacca gccgctggat catcctcgtc      240
aaaccaggca cgtatcgcgga ggtcgctctac gtgcagcggt agaagcgctt cgtcacgctg      300
atcggcgaag acccggcacg gacgacgac acgtaccacc tcaaagcgctc tgacgtgggg      360
ctcgacggca agcccatcgg caggtttcgc acgcccagca tgggtggtgga tgccgacgat      420
ttcacgatcg agaacctcac catcgagaac ggggcagggc cggtcggtca agcgtgccc      480
ttgcgagtgg acggcgatcg cgtgacggtg aggaacagcc gcctgctggg ctggcaggac      540
acgatctttc tcaaccgtgg gcgcactac ttcgaggact cgttcacgcy cgggcacgtg      600
gatttcattt tcggcggcgc gacggcggtg ttcgagcgat gccatcttcg cgcctggcgg      660
gacggtacc tcacggcgcc gtccacgccc gcggagcaac gattcggtt cgtgttcctg      720
aacagcatcg tcagtggaga agctggcgcc cgcacgtacc tcggtcgacc gtggcgggcg      780
ttcgcgacgc tggccttcac caagacgacg atgggcgagg tggtgccccc ggtgggctgg      840
aacaactggg accggccgga gcgtgagaag accgtgcgtt ttctogaagc aggcaccagc      900
ggcgcgggcg gcagcgctgc tgcgcgcgtc gcctgggcgc gcgtcgccac gccagccgag      960
ctcgctgatc tgacgacgga ggtggtgctt ggcggcaccg acggtggga cccgcgtcgc      1020
gtcgcgccgt acccgctcgg cgttcgcgcc aacgcggcgc cgtgcgcgc gccgcgccgg      1080
ccgacgtcgt ctggcccgca gagcccgccc gccttgacgt gggaccaggt cgcgcgccag      1140
ccagctcgtt ggctggccac accggaagcg ctgcggtatt ccgagaacgt gcgcctctat      1200
caacggcaca ctggcggtg gcccaaaac ctcgacatgg cgcagcgtt gacggacgcc      1260
gatcgcgcg gtctcacggc cgatcgcgcg ctcgacgact cgaccatcga caatggcgcc      1320
acgacgcggc agatcgagtt tctcgcccg atcgcccgcc ccaaccgga cgagcgcgcg      1380
caggcgctga tgctggctgg gatcgactac ctgctcgcg cccagtatcc aaacggcgcg      1440
tgccgcgagt atttcccgct ccgcaacgac tactcgcgcc acatcacgtt caacgacgac      1500
gcgatgatcg cggccgcgac gatcctgcag tcggtcgcg tcggccgtcc gccgttcgcc      1560
ggcgtcgacg cgactcgccg ccggcgggcg gcggaggccg tcgcgcgcgc ccacgcgtg      1620
attctggcct cgcagattcg cgtcaacggc cagctcactg gctggtgcca gcagcacgat      1680
gcacgcacgc tggagccagc gcgcggcgcg acctacgagc atccatcgat cagtggccgc      1740
gaaaccgtga cgatcgtaa tttcctgcg tcgatcgaac cgcgcgaccg ccagacccaa      1800
gccgccatcg atgcgcgat ggagtggctc aaggcgtgc agatccgcg ctggcgacg      1860
gagcgggcgc ccgatccctc aggaccgggc ggttacgac tggatggtt ggaggacccc      1920
aacgcggcgc cgtctgggc ccgcttctac gagattggca ccaatcgctc gatctactcg      1980
ggccgggacg gcgtcatcaa gtaccggctc gccgagatcg aaattgaac gcggaccggc      2040
tacagctggg tcggaccgta cgcgcaggcg ctgctcgatg aagagcgag gaagtaa      2097

```

<210> 98

<211> 698

<212> PRT

5 <213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10 <221> DOMINIO

<222> (45)...(333)

<223> Dominio de pectin metil esterase

<221> DOMINIO

15 <222> (336)...(698)

<223> Dominio catalítico

<400> 98

```

Met Asn Ala Ala Gly Ser Arg Arg Phe Ala Gln Leu Val Val Ala Asp
 1           5           10           15
Leu Arg Arg Leu Val Pro Ala Leu Ala Pro Phe Phe Arg Asp Glu Pro
           20           25           30
Leu Ala Gly Gly Val Ala Ala Leu Gln Arg Ser Val Asp Ala Ile Val
           35           40           45

```

20

Ala	Ala	Asp	Gly	Thr	Gly	Gln	Phe	Ala	Thr	Val	Gln	Glu	Ala	Ile	Asn	50	55	60
Ala	Ala	Pro	Gln	Asn	Thr	Ser	Thr	Thr	Ser	Arg	Trp	Ile	Ile	Leu	Val	65	70	75
Lys	Pro	Gly	Thr	Tyr	Arg	Glu	Val	Val	Tyr	Val	Gln	Arg	Glu	Lys	Arg	85	90	95
Phe	Val	Thr	Leu	Ile	Gly	Glu	Asp	Pro	Ala	Arg	Thr	Thr	Ile	Thr	Tyr	100	105	110
His	Leu	Lys	Ala	Ser	Asp	Val	Gly	Leu	Asp	Gly	Lys	Pro	Ile	Gly	Thr	115	120	125
Phe	Arg	Thr	Pro	Thr	Met	Val	Val	Asp	Ala	Asp	Asp	Phe	Thr	Ile	Glu	130	135	140
Asn	Leu	Thr	Ile	Glu	Asn	Gly	Ala	Gly	Pro	Val	Gly	Gln	Ala	Leu	Ala	145	150	155
Leu	Arg	Val	Asp	Gly	Asp	Arg	Val	Thr	Val	Arg	Asn	Ser	Arg	Leu	Leu	165	170	175
Gly	Trp	Gln	Asp	Thr	Ile	Phe	Leu	Asn	Arg	Gly	Arg	His	Tyr	Phe	Glu	180	185	190
Asp	Ser	Phe	Ile	Gly	Gly	His	Val	Asp	Phe	Ile	Phe	Gly	Gly	Ala	Thr	195	200	205
Ala	Val	Phe	Glu	Arg	Cys	His	Leu	Arg	Ala	Trp	Arg	Asp	Gly	Tyr	Leu	210	215	220
Thr	Ala	Ala	Ser	Thr	Pro	Ala	Glu	Gln	Arg	Phe	Gly	Phe	Val	Phe	Leu	225	230	235
Asn	Ser	Ile	Val	Ser	Gly	Glu	Ala	Gly	Ala	Arg	Thr	Tyr	Leu	Gly	Arg	245	250	255
Pro	Trp	Arg	Ala	Phe	Ala	His	Val	Ala	Phe	Ile	Lys	Thr	Thr	Met	Gly	260	265	270
Glu	Val	Val	Arg	Pro	Val	Gly	Trp	Asn	Asn	Trp	Asp	Arg	Pro	Glu	Arg	275	280	285
Glu	Lys	Thr	Val	Arg	Phe	Leu	Glu	Ala	Gly	Thr	Ser	Gly	Ala	Gly	Gly	290	295	300
Ser	Val	Ala	Ala	Arg	Val	Ala	Trp	Ala	Arg	Val	Ala	Thr	Pro	Ala	Glu	305	310	315
Leu	Ala	Asp	Leu	Thr	Thr	Glu	Val	Val	Leu	Gly	Gly	Thr	Asp	Gly	Trp	325	330	335
Asp	Pro	Arg	Arg	Val	Ala	Pro	Tyr	Pro	Ser	Ala	Val	Arg	Ala	Asn	Ala	340	345	350
Ala	Pro	Leu	Pro	Arg	Pro	Pro	Gly	Pro	Asp	Val	Ala	Gly	Pro	Gln	Ser	355	360	365
Pro	Pro	Ala	Leu	Thr	Trp	Asp	Gln	Val	Ala	Arg	Gln	Pro	Ala	Ser	Trp	370	375	380
Leu	Ala	Thr	Pro	Glu	Ala	Leu	Arg	Ile	Ala	Glu	Asn	Val	Arg	Leu	Tyr	385	390	395
Gln	Arg	His	Thr	Gly	Gly	Trp	Pro	Lys	Asn	Leu	Asp	Met	Ala	Gln	Pro	405	410	415
Leu	Thr	Asp	Ala	Asp	Arg	Ala	Arg	Leu	Thr	Ala	Asp	Arg	Ala	Leu	Asp	420	425	430
Asp	Ser	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Thr	Arg	Gln	Ile	Glu	Phe	Leu	435	440	445
Ala	Arg	Ile	Ala	Ala	Ala	Asn	Arg	Asp	Glu	Arg	Ala	Gln	Ala	Ser	Met	450	455	460
Leu	Ala	Gly	Ile	Asp	Tyr	Leu	Leu	Ala	Ala	Gln	Tyr	Pro	Asn	Gly	Gly	465	470	475
Trp	Pro	Gln	Tyr	Phe	Pro	Leu	Arg	Asn	Asp	Tyr	Ser	Arg	His	Ile	Thr	485	490	495
Phe	Asn	Asp	Asp	Ala	Met	Ile	Ala	Ala	Ala	Thr	Ile	Leu	Gln	Ser	Val	500	505	510
Ala	Leu	Ala	Arg	Pro	Pro	Phe	Ala	Gly	Val	Asp	Ala	Thr	Arg	Arg	Arg			

Arg	Ala	Ala	Glu	Ala	Val	Ala	Arg	Ala	His	Arg	Val	Ile	Leu	Ala	Ser
530						535					540				
Gln	Ile	Arg	Val	Asn	Gly	Gln	Leu	Thr	Gly	Trp	Cys	Gln	Gln	His	Asp
545					550					555					560
Ala	Arg	Thr	Leu	Glu	Pro	Ala	Arg	Gly	Arg	Thr	Tyr	Glu	His	Pro	Ser
				565					570					575	
Ile	Ser	Gly	Arg	Glu	Thr	Val	Thr	Ile	Val	Asn	Phe	Leu	Arg	Ser	Ile
			580				585					590			
Glu	Pro	Arg	Asp	Arg	Gln	Thr	Gln	Ala	Ala	Ile	Asp	Ala	Ala	Met	Glu
595							600					605			
Trp	Leu	Lys	Ala	Val	Gln	Ile	Arg	Gly	Trp	Arg	Thr	Glu	Arg	Arg	Pro
610						615					620				
Asp	Pro	Ser	Gly	Pro	Gly	Gly	Tyr	Asp	Val	Val	Met	Val	Glu	Asp	Pro
625					630					635					640
Asn	Ala	Ala	Pro	Leu	Trp	Ala	Arg	Phe	Tyr	Glu	Ile	Gly	Thr	Asn	Arg
			645						650					655	
Pro	Ile	Tyr	Ser	Gly	Arg	Asp	Gly	Val	Ile	Lys	Tyr	Arg	Leu	Ala	Glu
			660						665				670		
Ile	Glu	Ile	Glu	Arg	Arg	Thr	Gly	Tyr	Ser	Trp	Val	Gly	Pro	Tyr	Ala
675							680					685			
Gln	Ala	Leu	Leu	Asp	Glu	Glu	Arg	Arg	Lys						
690						695									

<210> 99
 <211> 1782
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 99

atgttcacta	ctaacagctc	tatttgcgcc	cggaaatccg	cgcgtttttc	actgactgcc	60
atggctgctg	cggtggctat	gatcgcgggc	acctctgcct	ttgcggcctc	taccggtggc	120
tttctgacca	cggtatggcg	caatgtgtca	gggtcaaaat	cctttaccgc	ctcaagccac	180
acccaaatcc	agcaaatcct	tgaggatgcc	aaagatggca	attatccggt	ggtgatcacc	240
tacaccggca	atgaggattc	actgattaac	caagtcgtcc	gggatcacac	cgtcgattct	300
tcaggcaact	gccctaaagc	gcgttggaat	gatgcctacc	gcaaagtcga	aatcaaagaa	360
atgaccaagg	gtgtcaccat	tcagggtgcc	aatgggtcgt	cggcgaattt	cggaatcgtg	420
gtgaataaat	ccagcaacgt	gattattcgc	aacatgaaga	ttgggtgact	gggcggcgct	480
aataacgatg	cggatatgat	ccgtgtggac	agcgggtgtg	acgtctggat	cgatcacaac	540
gaattattcg	ccgtgaacaa	cgagtgtgag	ggttcacccg	atggcgatct	gacctttgaa	600
agcgcgattg	atatcaaaaa	agcctcgcaa	gatatacccg	tgtcctacaa	cgtgattcgc	660
gacagtaaaa	aagtcggttt	ggatggctcc	agcagcagcg	atatcgccgg	cggccgcaaa	720
attactttcc	accacaatat	ctaccgcaac	gtagggtgcg	gcttaccttt	gcagcgcggc	780
ggttggaacg	acatgtacaa	caacctgtac	gacggcatta	ccagctcggg	catcaacgtg	840
cgccaaaaag	ggtatgcgtt	aattgaaagc	aactgggtcc	aaaacgcggt	taaccgggtc	900
acctgccgtt	ttgacagcag	caactgcggc	aagtgggata	tgcgcaacaa	taacatccgc	960
aaccggggtg	atgttgcgac	ttacaacatc	acctggacca	gtggcggcac	catcgacgcc	1020
accaactgga	ccaccactgc	gcccttcctt	atcagcattc	cctacagcta	ttcaccgggt	1080
actccgcaat	gtgtgaaaga	tcgtctggcg	agttacgcgg	gtgtgggtaa	aaacggcgcg	1140
cagctgactg	cctcggcctg	cgggtggtgc	gcacgtgcca	cacctgcata	gtccacacct	1200
gcaagttcca	gctctgcggc	aaacagttcc	gctgcatcag	gcagtgtgag	tttggtgggc	1260
agtgcgggta	atgcatcggt	tgcacttaac	tggaccgiga	atgccaacat	taatgcgctg	1320
gaaatttata	aggatacggg	ttctgatccc	gccggacgtg	tgcgcattgc	gtcgtctgca	1380
accagcgcg	ccaactacac	cgcaacaggt	ctgagcaacg	gcactaccta	ttacttctgg	1440
gtgaaatata	gcaccaacca	taatgtgtgg	agcaactcca	atgtgttcag	cgccaagcca	1500
agttcaggta	caaccccgct	atcatccagc	agcgcggctt	catcaacgcc	aagtgggtgca	1560

```

ccggtgttaa gtggtacagg tgattaccca agcgggttct ccaagtgtgc tgatctgggt 1620
ggcacctgct cagtcgcctc gggcgatggt tgggttgctt ttggtcgcaa aggcaagtgg 1680
gtcaccaaaa aagtgtcagt cggtagctct attgcctgta ccgttgccgc gtttggatct 1740
gatccacaag gcaatcccaa taagtgttct tataaaaagt aa 1782

```

<210> 100
 <211> 593
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(35)

<221> DOMINIO
 <222> (36)...(593)
 <223> Dominio catalítico

<400> 100

```

Met Phe Thr Thr Asn Ser Ser Ile Cys Ala Arg Lys Ser Ala Arg Phe
 1      5      10      15
Ser Leu Thr Ala Met Ala Ala Val Ala Met Ile Ala Gly Thr Ser
      20      25      30
Ala Phe Ala Ala Ser Thr Gly Gly Phe Ser Thr Thr Asp Gly Gly Asn
 35      40      45
Val Ser Gly Ser Lys Ser Phe Thr Ala Ser Ser His Thr Gln Ile Gln
 50      55      60
Gln Ile Leu Glu Asp Ala Lys Asp Gly Asn Tyr Pro Val Val Ile Thr
 65      70      75      80
Tyr Thr Gly Asn Glu Asp Ser Leu Ile Asn Gln Val Val Arg Asp His
      85      90      95
Thr Val Asp Ser Ser Gly Asn Cys Pro Lys Ala Arg Trp Asn Asp Ala
      100      105      110
Tyr Arg Lys Val Glu Ile Lys Glu Met Thr Lys Gly Val Thr Ile Gln
      115      120      125
Gly Ala Asn Gly Ser Ser Ala Asn Phe Gly Ile Val Val Asn Lys Ser
      130      135      140
Ser Asn Val Ile Ile Arg Asn Met Lys Ile Gly Ala Leu Gly Gly Ala
      145      150      155      160
Asn Asn Asp Ala Asp Met Ile Arg Val Asp Ser Gly Val Asn Val Trp
      165      170      175
Ile Asp His Asn Glu Leu Phe Ala Val Asn Asn Glu Cys Lys Gly Ser
      180      185      190
Pro Asp Gly Asp Leu Thr Phe Glu Ser Ala Ile Asp Ile Lys Lys Ala
      195      200      205
Ser Gln Asp Ile Thr Val Ser Tyr Asn Val Ile Arg Asp Ser Lys Lys
      210      215      220
Val Gly Leu Asp Gly Ser Ser Ser Ser Asp Ile Ala Gly Gly Arg Lys
      225      230      235      240
Ile Thr Phe His His Asn Ile Tyr Arg Asn Val Gly Ala Arg Leu Pro
      245      250      255
Leu Gln Arg Gly Trp Thr His Met Tyr Asn Asn Leu Tyr Asp Gly
      260      265      270
Ile Thr Ser Ser Gly Ile Asn Val Arg Gln Asn Gly Tyr Ala Leu Ile
      275      280      285
Glu Ser Asn Trp Phe Gln Asn Ala Val Asn Pro Val Thr Cys Arg Phe
      290      295      300

```

Asp Ser Ser Asn Cys Gly Lys Trp Asp Leu Arg Asn Asn Asn Ile Arg
 305 310 315 320
 Asn Pro Gly Asp Phe Ala Thr Tyr Asn Ile Thr Trp Thr Ser Gly Gly
 325 330 335
 Thr Ile Asp Ala Thr Asn Trp Thr Thr Ala Pro Phe Pro Ile Ser
 340 345 350
 Ile Pro Tyr Ser Tyr Ser Pro Val Thr Pro Gln Cys Val Lys Asp Arg
 355 360 365
 Leu Ala Ser Tyr Ala Gly Val Gly Lys Asn Gly Ala Gln Leu Thr Ala
 370 375 380
 Ser Ala Cys Gly Gly Ala Ala Ser Ser Thr Pro Ala Ser Ser Thr Pro
 385 390 395 400
 Ala Ser Ser Ser Ser Ala Ala Asn Ser Ser Ala Ala Ser Gly Ser Val
 405 410 415
 Ser Leu Gly Gly Ser Ala Gly Asn Ala Ser Val Ala Leu Asn Trp Thr
 420 425 430
 Val Asn Ala Asn Ile Asn Ala Leu Glu Ile Tyr Gln Asp Thr Asp Ser
 435 440 445
 Asp Pro Ala Gly Arg Val Arg Ile Ala Ser Leu Pro Thr Ser Ala Thr
 450 455 460
 Asn Tyr Thr Ala Thr Gly Leu Ser Asn Gly Thr Thr Tyr Tyr Phe Trp
 465 470 475 480
 Val Lys Tyr Arg Thr Thr Asn Asn Val Trp Ser Asn Ser Asn Val Phe
 485 490 495
 Ser Ala Lys Pro Ser Ser Gly Thr Thr Pro Ser Ser Ser Ser Ala
 500 505 510
 Ala Ser Ser Thr Pro Ser Gly Ala Pro Val Leu Ser Gly Thr Gly Asp
 515 520 525
 Tyr Pro Ser Gly Phe Ser Lys Cys Ala Asp Leu Gly Gly Thr Cys Ser
 530 535 540
 Val Ala Ser Gly Asp Gly Trp Val Ala Phe Gly Arg Lys Gly Lys Trp
 545 550 555 560
 Val Thr Lys Lys Val Ser Val Gly Ser Ser Ile Ala Cys Thr Val Ala
 565 570 575
 Ala Phe Gly Ser Asp Pro Gln Gly Asn Pro Asn Lys Cys Ser Tyr Lys
 580 585 590
 Lys

<210> 101
 <211> 1404
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 101

atgactatag	accgtcgaga	attccttata	gacctcatta	tgggcaccgc	cggtctcgca	60
atcgaccgga	gtgatgcgtt	cggccaagct	gatccatgga	aaaccgtcta	tccgcaaata	120
ctcgctcgca	tacggccacc	gaaatttccg	aagcgagatt	tcattcatcac	tagattcggc	180
gcaaagccgg	gaaccgacag	cgccgctgcg	atcgcaaaag	ctatcacccg	gtgcagcaag	240
gcaggcggag	gacgtgttct	cgttcccgca	ggagagtttc	tcaccggagc	gatccatctg	300
aaatcgaacg	taaactttca	cgtgtcaaaa	ggcgcgacgc	tgaaattctc	gaccgacccg	360
aaggcatatc	tcccgattgt	acatacgcga	tgggaaggaa	tggagctgat	gcattctgtca	420
cggttcatct	acgcttatga	gcagacgaac	atcgctatca	cgggtcaggg	aacgctcgac	480
ggccaggga	aatcattctt	ctggaaatgg	catggcaatc	cggcttatgg	cggcgatccg	540
aacacgctca	gccaacggcc	cgctcgtgcg	cggctttacg	agatgatgga	taagaatgtg	600
cgggtcgccg	aacgtgtctt	cggctctcgga	cattatctgc	ggccgcagtt	tattcagccg	660

```

tacaatgca ggaacgtttt gatcgaagat gtgacgatcg togattcgcc gatgtgggaa 720
gttcatccgg tgctttgcga gaacgtcacg gtccgaaatg ttcacatttc atcgcatggt 780
ccgaacaatg acggatgcga tccggagtcg tgcaaggacg tactgatcga caactgtttt 840
ttcgacaacg gcgacgattg catcgcgatc aagtcgggcc gcaacaatga cggtcgctcg 900
atcaatgtcc cgaccgagaa catcatcgtc cgcaactgca caatgaaaga cggtcatggc 960
ggcatcacgg tcggcagtga gatttcggga ggcggtcgaa atctttttgc gcacgattgt 1020
cgactcgaca gtgcggatct ctggaccgcg cttcgcgta agaacaatgc gtcgcgaggc 1080
ggcaagctcg agaattttta ttttcggaat ataacggtcg gccaggtcgc acgcgctgtg 1140
gtcgagatcg attttaatta cgaggaaggc gcgaaaggct cgtatatattcc tgcggtcga 1200
aattatgttg ttgaaggact gacatgcgcc acaggcaatc gcgcgctcga tctgcaagga 1260
ttggacaacg cgccgatcta caatgtaacg ctgcgaaact gtacgtttgg ttctgtccga 1320
aatcgtagtg ttgtgaaaaa cgttcgtgga cttcggctcg agaatgtgaa gatcggcggc 1380
aggatcgtaa acgaactggt atga 1404

```

<210> 102

<211> 467

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<221> SEÑAL

<222> (1)...(28)

<221> DOMINIO

15 <222> (78)...(459)

<223> Dominio catalítico

<400> 102

```

Met Thr Ile Asp Arg Arg Glu Phe Leu Ile Asp Leu Ile Ile Gly Thr
 1           5           10           15
Ala Gly Phe Ala Ile Ala Pro Ser Asp Ala Phe Gly Gln Ala Asp Pro
          20           25           30
Trp Lys Thr Val Tyr Pro Gln Ile Leu Ala Arg Ile Arg Pro Pro Lys
          35           40           45
Phe Pro Lys Arg Asp Phe Ile Ile Thr Arg Phe Gly Ala Lys Pro Gly
 50           55           60
Thr Asp Ser Ala Ala Ala Ile Ala Lys Ala Ile Thr Ala Cys Ser Lys
65           70           75           80
Ala Gly Gly Gly Arg Val Leu Val Pro Ala Gly Glu Phe Leu Thr Gly
          85           90           95
Ala Ile His Leu Lys Ser Asn Val Asn Phe His Val Ser Lys Gly Ala
          100          105          110
Thr Leu Lys Phe Ser Thr Asp Pro Lys Ala Tyr Leu Pro Ile Val His
          115          120          125
Thr Arg Trp Glu Gly Met Glu Leu Met His Leu Ser Pro Phe Ile Tyr
          130          135          140
Ala Tyr Glu Gln Thr Asn Ile Ala Ile Thr Gly Gln Gly Thr Leu Asp
145          150          155          160
Gly Gln Gly Lys Ser Phe Phe Trp Lys Trp His Gly Asn Pro Ala Tyr
          165          170          175
Gly Gly Asp Pro Asn Thr Leu Ser Gln Arg Pro Ala Arg Ala Arg Leu
          180          185          190
Tyr Glu Met Met Asp Lys Asn Val Pro Val Ala Glu Arg Val Phe Gly
          195          200          205
Leu Gly His Tyr Leu Arg Pro Gln Phe Ile Gln Pro Tyr Lys Cys Arg
          210          215          220
Asn Val Leu Ile Glu Asp Val Thr Ile Val Asp Ser Pro Met Trp Glu

```

20

225		230		235		240
Val His Pro Val	Leu Cys Glu Asn Val Thr	Val Arg Asn Val His Ile				
	245	250		255		
Ser Ser His Gly	Pro Asn Asn Asp Gly Cys Asp Pro Glu Ser Cys Lys					
	260	265		270		
Asp Val Leu Ile	Asp Asn Cys Phe Phe Asp Thr Gly Asp Asp Cys Ile					
	275	280		285		
Ala Ile Lys Ser Gly	Arg Asn Asn Asp Gly Arg Arg Ile Asn Val Pro					
	290	295		300		
Thr Glu Asn Ile Ile	Val Arg Asn Cys Thr Met Lys Asp Gly His Gly					
305	310	315		320		
Gly Ile Thr Val Gly	Ser Glu Ile Ser Gly Gly Val Arg Asn Leu Phe					
	325	330		335		
Ala His Asp Cys Arg	Leu Asp Ser Ala Asp Leu Trp Thr Ala Leu Arg					
	340	345		350		
Val Lys Asn Asn Ala	Ser Arg Gly Gly Lys Leu Glu Asn Phe Tyr Phe					
	355	360		365		
Arg Asn Ile Thr Val	Gly Gln Val Ala Arg Ala Val Val Glu Ile Asp					
	370	375		380		
Phe Asn Tyr Glu Glu	Gly Ala Lys Gly Ser Tyr Ile Pro Val Val Arg					
385	390	395		400		
Asn Tyr Val Val Glu	Gly Leu Thr Cys Ala Thr Gly Asn Arg Ala Val					
	405	410		415		
Asp Leu Gln Gly Leu	Asp Asn Ala Pro Ile Tyr Asn Val Thr Leu Arg					
	420	425		430		
Asn Cys Thr Phe Gly	Ser Val Arg Asn Arg Ser Val Val Lys Asn Val					
	435	440		445		
Arg Gly Leu Arg Leu	Glu Asn Val Lys Ile Gly Gly Arg Ile Val Asn					
	450	455		460		
Glu Leu Val						
465						

<210> 103
 <211> 1101
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 103

atgaacacccg	cactgcacccg	cgatcatccgc	ctgccgctgc	tgctggcgct	gtgcctgccc	60
gcgctgcagg	cacaggccac	gcagaccgag	cccgctgccg	agaacatgct	gctgctgcag	120
accgcgtccg	gcggctggtc	caagcaccac	cagggcaagg	cggtcgacta	cggccacacg	180
ttcacccgatg	ccgaacgtgc	ggcgctgcgc	gcgcccgaacc	gcagggacga	tgccagcatc	240
gacaacaagg	cgaccacgct	tgagatcgtc	gcgctgctgg	aagcccacca	gcgcaccggc	300
aatgccgcct	atctggcggc	tgcgcagcgc	ggcgctggact	acctgctggc	cgcgcagtac	360
ccgaacggcg	gctggccgca	gtactaccgc	gacogttcgc	tgtaccggca	ccaggtcacc	420
ttcaacgatg	atgcgatgac	ccgcgtgctg	gagctgctgc	aggacatcgt	cgagggcaag	480
ggcgcgctgg	cgcagctgac	acccacgcac	ggcgaacgcg	ccagggccgc	gctcgacagg	540
ggcatcgccct	gcgtgctcgc	cacccaggta	cgcatcgatg	gcgagctcac	gctctggggc	600
gcgcagtaag	acgaagccac	gctgcagccg	gcgaaggcgc	gctcctacga	gctgccatcg	660
ctggcggtcg	ccgaatcggt	cggggtgatg	cggtgctga	tgccgccagc	acagccgtcg	720
ccgcaggtgc	tgacggcggt	cgaggccggc	gcacgctggc	tggaggcgca	ccgcatgccc	780
gacctggccc	ggcgaaagat	cgacgcggcc	ggcgaaagaa	ccggccagga	cgtgggtgatc	840
gtcgccgagc	ccggcgcgct	gctgtgggca	cgcttctacg	acctgcagca	ccagcagccg	900
atgttcgtga	accgcgaagg	cgagcaggtg	gcccgtttcg	ccgacatgcc	caacgaacgc	960
cgcgtcggtc	acgcctggta	tgccgtgtgg	ccggagaagc	tgctgcagca	ggagctgcca	1020
cgctggtaca	acacccatgc	cgaggcattg	cgcgcgatta	cgctgcgca	tgccgagcca	1080
aggccgccga	agcggccctg	a				1101

<210> 104

<211> 366
<212> PRT
<213> Desconocido

5 <220>
<223> Obtenido de una muestra ambiental

<221> SEÑAL
<222> (1)...(26)

10 <221> DOMINIO
<222> (27)...(366)
<223> Dominio catalítico

15 <400> 104

```

Met Asn Thr Ala Leu His Arg Val Ile Arg Leu Pro Leu Leu Leu Ala
 1      5      10      15
Leu Cys Leu Pro Ala Leu Gln Ala Gln Ala Thr Gln Thr Glu Pro Val
      20      25      30
Ala Glu Asn Met Leu Leu Leu Gln Thr Ala Ser Gly Gly Trp Ser Lys
      35      40      45
His His Gln Gly Lys Ala Val Asp Tyr Gly His Thr Phe Thr Asp Ala
 50      55      60
Glu Arg Ala Ala Leu Arg Ala Pro Asp Arg Arg Asp Ala Thr Ile
 65      70      75      80
Asp Asn Lys Ala Thr Thr Leu Glu Ile Val Ala Leu Leu Glu Ala His
      85      90      95
Gln Arg Thr Gly Asn Ala Ala Tyr Leu Ala Ala Ala Gln Arg Gly Val
      100      105      110
Asp Tyr Leu Leu Ala Ala Gln Tyr Pro Asn Gly Gly Trp Pro Gln Tyr
      115      120      125
Tyr Pro Asp Arg Ser Leu Tyr Arg His Gln Val Thr Phe Asn Asp Asp
      130      135      140
Ala Met Thr Arg Val Leu Glu Leu Leu Gln Asp Ile Val Glu Gly Lys
      145      150      155      160
Gly Ala Leu Ala Gln Leu Thr Pro Thr His Gly Glu Arg Ala Arg Ala
      165      170      175
Ala Leu Asp Arg Gly Ile Ala Cys Val Leu Ala Thr Gln Val Arg Ile
      180      185      190
Asp Gly Glu Leu Thr Leu Trp Ala Ala Gln Tyr Asp Glu Ala Thr Leu
      195      200      205
Gln Pro Ala Lys Ala Arg Ser Tyr Glu Leu Pro Ser Leu Ala Val Ala
      210      215      220
Glu Ser Val Gly Val Met Arg Leu Leu Met Arg Gln Pro Gln Pro Ser
      225      230      235      240
Pro Gln Val Leu Thr Ala Val Glu Ala Gly Ala Arg Trp Leu Glu Ala
      245      250      255
His Arg Met Arg Asp Leu Ala Arg Arg Lys Ile Asp Ala Pro Gly Glu
      260      265      270
Glu Thr Gly Gln Asp Val Val Ile Val Ala Glu Pro Gly Ala Ser Leu
      275      280      285
Trp Ala Arg Phe Tyr Asp Leu Gln His Gln Gln Pro Met Phe Val Asn
      290      295      300
Arg Glu Gly Glu Gln Val Ala Arg Phe Ala Asp Met Pro Asn Glu Arg
      305      310      315      320
Arg Val Gly Tyr Ala Trp Tyr Gly Val Trp Pro Glu Lys Leu Leu Gln
      325      330      335
Gln Glu Leu Pro Arg Trp Tyr Asn Thr His Ala Glu Ala Leu Arg Ala
      340      345      350
Ile Thr Pro Ala His Ala Glu Pro Arg Pro Pro Lys Arg Pro
      355      360      365

```

<210> 105
<211> 1203
<212> ADN
<213> Desconocido

5

<220>
<223> Obtenido de una muestra ambiental

<400> 105

10

```

atgcaattca tcgaaacaca gcaattgggg acogccgcga aaccgcgtggc gggacgagga      60
ggcgacaggc gctttccgcg ggtcatgccc gccgtttgcg cgggccttgc cctcgccgtg      120
tcgtcgcccg agccggtccg ggcgcagggc gcggatgcgg atgcggatgg cccactgccc      180
aggtggaaca ggaggctggt ggatcgcccc gaggactggt tcgcctccga cgagggacag      240
cgcgttgccg ccaacgtcct ccgtaccaa tcggcggaag gagcctggcc caaaaacacc      300
aatctggccg ccaactccct tcgcccgcag gacattccct cctcgacctc cggggtggcc      360
aacacgatcg acaatgaagc caccaccgtg cccattcggt ttttgcccg tttcgcgcaa      420
atcaacgagg acacggccag ccgcgaggcg gtccagcgcg gattggacta tctcctcaag      480
gcgcaatata cgaacggtgg ctggccgcag tatttccgcg tcgcgcgcgg ctaccactcg      540
cacatcacct acaacgacga cgcctatggt aatgtgctcg acctgctgct ggacgtgtcg      600
ctgggcgagg agccgttcga ttttgtggac gaggatcgcc gccagcgggc cgcgaccgcc      660
gtggagcggg ggatcgaatg catcctccgc acccaaatcc ggcaggagga ccaaccacc      720
ggctggtgcg cgcagtatga ccccgaaacc ttggccccgg cgtggggacg ggcgtacgag      780
ccgcgctcga tttccggagc cgagaccgtc ggcgtggcgc ggtttctgat gcggctggag      840
tcgccatgcg cgggaagccgt cgaagccatc gagggcgcca tcgcctggct cgacacggtg      900
ggcatcgagg aattgcgtct cgaatggttc accaacagcg agggcaagcg tgaccggcgc      960
gtggtcgagg acgcttccgt gggcacccct tgggcgcgct tttacgaact cgaaacgaac     1020
cgcccccttg tcgtggaccg cgacggggtg ctccgctacg acttcgcgga actgacggcg     1080
gagcgccgcc aaggttacag ctactacggc acttggccgg cgcctattgt ggccacggaa     1140
tatccgcgct ggcgaggat gaacgagtcg gcctgctcg agtcgtcctt catctcgcat     1200
tga                                                                    1203

```

<210> 106
<211> 400
<212> PRT
<213> Desconocido

15

<220>
<223> Obtenido de una muestra ambiental

20

<221> SEÑAL
<222> (1)...(43)

<221> DOMINIO
<222> (44)...(400)
<223> Dominio catalítico

25

<400> 106

30

```

Met  Gln  Phe  Ile  Glu  Thr  Gln  Gln  Leu  Gly  Thr  Ala  Ala  Lys  Pro  Val
 1          5          10          15
Ala  Gly  Arg  Gly  Gly  Asp  Arg  Arg  Phe  Pro  Arg  Val  Met  Pro  Ala  Val
          20          25          30
Cys  Ala  Gly  Leu  Ala  Leu  Ala  Val  Ser  Ser  Ala  Glu  Pro  Val  Arg  Ala
          35          40          45

```


Gln	Gly	Ala	Asp	Ala	Asp	Ala	Asp	Gly	Pro	Leu	Pro	Arg	Trp	Asn	Arg
50						55					60				
Arg	Leu	Val	Asp	Arg	Pro	Glu	Asp	Trp	Phe	Ala	Ser	Asp	Glu	Gly	Gln
65					70					75					80
Arg	Val	Ala	Ala	Asn	Val	Leu	Arg	Tyr	Gln	Ser	Ala	Glu	Gly	Ala	Trp
				85				90					95		
Pro	Lys	Asn	Thr	Asn	Leu	Ala	Ala	Thr	Pro	Leu	Arg	Pro	Glu	Asp	Ile
			100					105					110		
Pro	Ser	Ser	Thr	Ser	Gly	Val	Ala	Asn	Thr	Ile	Asp	Asn	Glu	Ala	Thr
		115					120				125				
Thr	Val	Pro	Ile	Arg	Phe	Leu	Ala	Arg	Phe	Ala	Gln	Ile	Asn	Glu	Asp
	130					135					140				
Thr	Ala	Ser	Arg	Glu	Ala	Val	Gln	Arg	Gly	Leu	Asp	Tyr	Leu	Leu	Lys
145					150					155					160
Ala	Gln	Tyr	Pro	Asn	Gly	Gly	Trp	Pro	Gln	Tyr	Phe	Pro	Leu	Arg	Arg
				165				170						175	
Gly	Tyr	His	Ser	His	Ile	Thr	Tyr	Asn	Asp	Asp	Ala	Met	Val	Asn	Val
			180					185					190		
Leu	Asp	Leu	Leu	Leu	Asp	Val	Ser	Leu	Gly	Glu	Glu	Pro	Phe	Asp	Phe
	195						200					205			
Val	Asp	Glu	Asp	Arg	Arg	Gln	Arg	Ala	Ala	Thr	Ala	Val	Glu	Arg	Gly
	210					215					220				
Ile	Glu	Cys	Ile	Leu	Arg	Thr	Gln	Ile	Arg	Gln	Glu	Asp	Gln	Pro	Thr
225					230					235					240
Gly	Trp	Cys	Ala	Gln	Tyr	Asp	Pro	Glu	Thr	Leu	Ala	Pro	Ala	Trp	Gly
			245					250						255	
Arg	Ala	Tyr	Glu	Pro	Pro	Ser	Ile	Ser	Gly	Ala	Glu	Thr	Val	Gly	Val
		260						265					270		
Ala	Arg	Phe	Leu	Met	Arg	Leu	Glu	Ser	Pro	Ser	Pro	Glu	Ala	Val	Glu
	275						280					285			
Ala	Ile	Glu	Gly	Ala	Ile	Ala	Trp	Leu	Asp	Thr	Val	Gly	Ile	Glu	Glu
	290					295					300				
Leu	Arg	Leu	Glu	Trp	Phe	Thr	Asn	Ser	Glu	Gly	Lys	Arg	Asp	Arg	Arg
305					310					315					320
Val	Val	Glu	Asp	Ala	Ser	Val	Gly	Thr	Leu	Trp	Ala	Arg	Phe	Tyr	Glu
			325					330						335	
Leu	Glu	Thr	Asn	Arg	Pro	Leu	Phe	Val	Asp	Arg	Asp	Gly	Val	Leu	Arg
			340					345					350		
Tyr	Asp	Phe	Ala	Glu	Leu	Thr	Ala	Glu	Arg	Arg	Gln	Gly	Tyr	Ser	Tyr
	355						360					365			
Tyr	Gly	Thr	Trp	Pro	Ala	Pro	Leu	Leu	Ala	Thr	Glu	Tyr	Pro	Arg	Trp
	370					375					380				
Arg	Arg	Met	Asn	Glu	Ser	Ala	Leu	Leu	Glu	Ser	Ser	Phe	Ile	Ser	His
385					390					395					400

<210> 107
 <211> 1074
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 107

atgaagctac	cagttgtttc	cctgcgcgta	ctgctggcgc	tgctggccac	gtcgccggtc	60
gcctgcgcgg	gcgctgcggc	accgcgcact	gcgaccgatc	cggtcgccga	gaacatgctg	120
cttctgcaga	cgcctccgg	cggtcgttcc	aagcactacc	gcgagaagaa	ggctcgactac	180
gcgcgcgact	acgacgcgc	cgagcgcgcc	gcgctgcgcg	cgcccgaccg	gcatgacgat	240
gccacgatcg	acaacaaggc	cacgaccacc	gagatcgcat	acctggtgca	ggcacatgcc	300

```

aggacgggca atccggccta cctcgacggc gcgcgccgcg gcgtcgagta cctgctgcgc 360
gcgcagtagc cgaacggcgg ctggccgcag ttctaccccg accattcgtc ctaccggcac 420
cagatcacgc tcaacgacga tgcgatggcg cacgccatca ccgtgctgca ggacatcgcc 480
gcggggccga acggcatgca ggtgctggcg ccggagttcg gcgtccgcgc gccgcgggcc 540
gcgcagcgcg gcatcggaaa cctgctcgag ttgcaggtgc ggatcgacgg ggtgccgacg 600
atctggggcg cgagtagca cgagaccacc ctgcaaccgg ccaaggcccg tgcgtacgag 660
ttgccctcgc tggccgtggc cgaatcggcg ggcgtgatgc gcctgctgat gcgccagccg 720
gggcctgatg cgcgcacgat cgcgcgcgac gaggcggcgg cggactggct ggaggcgcac 780
cgctgcccgg acctcgccct ggaacgcata gaagcccccg ccgaggaaac cggcaaggac 840
gtccgcgctc tggccagacc gggcgcgctc ttgtggggcg gcttctacga cctcgagcgg 900
caggtgccgc tgttcgctga tcgcaacagc cgtccggttc cattcgccga gcttcccaac 960
gagcgtcgta ccggtatgg ctggtatggc acctggccgg aaaagctgct ggcacaggaa 1020
ctcccgcgct ggcgcaaggc ccatgcggcc agcgcgggcg ctccggcccg ttga 1074

```

<210> 108

<211> 357

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental

10

<221> SEÑAL

<222> (1)...(31)

<221> DOMINIO

15 <222> (32)...(357)

<223> Dominio catalítico

<400> 108

```

Met Thr Leu Pro Val Val Ser Leu Arg Val Leu Leu Ala Leu Leu Ala
.1          5          10          15
Thr Ser Pro Val Ala Cys Ala Gly Ala Ala Ala Pro Ala Thr Ala Thr
          20          25          30
Asp Pro Val Ala Glu Asn Met Leu Leu Leu Gln Thr Ala Ser Gly Gly
          35          40          45
Trp Ser Lys His Tyr Arg Glu Lys Lys Val Asp Tyr Ala Arg Asp Tyr
          50          55          60
Asp Ala Ala Glu Arg Ala Ala Leu Arg Ala Pro Asp Arg His Asp Asp
65          70          75          80
Ala Thr Ile Asp Asn Lys Ala Thr Thr Thr Glu Ile Ala Tyr Leu Val
          85          90          95
Gln Ala His Ala Arg Thr Gly Asn Pro Ala Tyr Leu Asp Gly Ala Arg
          100          105          110
Arg Gly Val Glu Tyr Leu Leu Arg Ala Gln Tyr Pro Asn Gly Gly Trp
          115          120          125
Pro Gln Phe Tyr Pro Asp His Ser Ser Tyr Arg His Gln Ile Thr Leu
130          135          140
Asn Asp Asp Ala Met Val His Ala Ile Thr Val Leu Gln Asp Ile Ala
145          150          155          160
Ala Gly Arg Asn Gly Met Gln Val Leu Ala Pro Glu Phe Gly Val Arg
          165          170          175
Ala Ala Ala Ala Ala Gln Arg Gly Ile Gly Asn Leu Leu Glu Leu Gln
          180          185          190
Val Arg Ile Asp Gly Val Pro Thr Ile Trp Ala Ala Gln Tyr Asp Glu
195          200          205
Thr Thr Leu Gln Pro Ala Lys Ala Arg Ala Tyr Glu Leu Pro Ser Leu
210          215          220
Ala Val Ala Glu Ser Val Gly Val Met Arg Leu Leu Met Arg Gln Pro

```

20

225		230		235		240									
Gly	Pro	Asp	Ala	Arg	Thr	Ile	Ala	Ala	Ile	Glu	Ala	Ala	Ala	Asp	Trp
				245					250					255	
Leu	Glu	Ala	His	Arg	Leu	Pro	Asp	Leu	Ala	Leu	Glu	Arg	Ile	Glu	Ala
			260					265					270		
Pro	Ala	Glu	Glu	Thr	Gly	Lys	Asp	Val	Arg	Val	Val	Ala	Arg	Pro	Gly
		275					280					285			
Ala	Ser	Leu	Trp	Ala	Arg	Phe	Tyr	Asp	Leu	Glu	Arg	Gln	Val	Pro	Leu
	290					295					300				
Phe	Val	Asp	Arg	Asn	Ser	Arg	Pro	Val	Pro	Phe	Ala	Glu	Leu	Pro	Asn
305					310					315					320
Glu	Arg	Arg	Thr	Gly	Tyr	Gly	Trp	Tyr	Gly	Thr	Trp	Pro	Glu	Lys	Leu
			325						330					335	
Leu	Ala	Gln	Glu	Leu	Pro	Arg	Trp	Arg	Lys	Val	His	Ala	Ala	Ser	Ala
		340						345					350		
Gly	Ala	Pro	Ala	Arg											
		355													

<210> 109
 <211> 1422
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 109

atgacgacac	gacgcgaatt	catcaaaggc	tttctactta	ccggagcagc	cgtggccgctc	60
gctccgcgtt	tgcttgcggt	cgccgcggag	gcaagtccgt	gggaaacgat	gatgccttcg	120
atcctcgcac	gcatcagacc	acctcgtttt	ccgaaacgca	ccttctatct	caatcgatcc	180
ggcgccaagg	gtgatggagt	cacagactgc	accgcggctt	ttcatcgcg	gatcgatgaa	240
tgcaccaaag	ccggcggtgg	gaaagtcgtc	gtgccggcgg	gcacttatct	caccggcgcg	300
attcatttga	agagcaacgt	caacctcgaa	gtctcggaag	gcgcgacgat	caagttcagt	360
caggacccga	aacactacct	gcctgttgtc	ttctcgcggt	gggaagggtg	cgaagtcttc	420
aactactcgc	ctttcattta	cgcgttcgaa	cagcgaacaa	tcgcgatcac	cggcaaaagg	480
acgctcgacg	gacagagtga	ttcggaacac	tggtggccgt	ggaacggccg	tcgcgagtac	540
ggatggaaaag	aagggatgaa	acagcagcgt	cccgatcgca	acgcgttggt	cacaatggcg	600
gagaaaggcg	tgccgggtgcg	cgagcgcac	tttggcgaa	gtcattatct	gaggccgcag	660
ttcattcagc	cgtaccgctg	ccagaacgtg	ctgatccagg	gcgtgacgat	tcggaactcg	720
ccgatgtggg	agattcatcc	ggtgttggtc	cgtaacgtga	ctattcacga	cgtgcacatc	780
gatatgcatg	gaccaaacaa	cgacggctgc	aatcccgaat	cgtgcagcga	cgtgttgatt	840
aaggtagctg	acttcgatac	cggcgacgac	tgcacgcga	tcaaactcgg	acgcaacgcc	900
gacggggcgg	ggcttaaagc	gccgactgag	aacatcatcg	ttcaaggatg	tcgcatgaaa	960
gacggccacg	gtggaatcac	ggtcggcagc	gagatctcgg	gcggcggtcg	aaacctgttt	1020
gccgagaatt	gccggctcga	cagtccaaac	ctcgatcacg	ccctgcgcgt	gaagaacaat	1080
gccatgcgcg	gcggattact	cgagaacttc	cacttccgta	acatcgaagt	cgggcagggtg	1140
gcccattgccg	tgattacgat	cgacttcaac	tacgaagagg	gcgcgaaagg	gtcgttcacg	1200
ccggtcgctt	gcgattacac	ggtcgatcgt	ttgcgcagca	cgaagagcaa	gcacgcactc	1260
gacgtccagg	gtctgcccgg	cgcgcgggtc	atcaacctgc	gattgacgaa	ctgcacattc	1320
aacgatgtgc	agcaaccgaa	cattctcaag	aacgtcgaac	aatcaacctt	tgaaaacgtc	1380
acgattaacg	gaaagacgat	cacacaaaca	ggatccaaag	aa		1422

<210> 110
 <211> 474
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(21)

<221> DOMINIO
 <222> (28)...(308)
 <223> Dominio de pectin metil esterasa

5 <221> DOMINIO
 <222> (309)...(637)
 <223> Dominio catalítico

<400> 110

10

```

Met Thr Thr Arg Arg Glu Phe Ile Lys Gly Phe Leu Leu Thr Gly Ala
 1      5      10      15
Ala Val Ala Val Ala Pro Arg Leu Leu Ala Phe Ala Ala Glu Ala Ser
      20      25      30
Pro Trp Glu Thr Met Met Pro Ser Ile Leu Ala Arg Ile Arg Pro Pro
      35      40      45
Arg Phe Pro Lys Arg Thr Phe Tyr Leu Asn Arg Phe Gly Ala Lys Gly
 50      55      60
Asp Gly Val Thr Asp Cys Thr Ala Ala Phe His Arg Ala Ile Asp Glu
 65      70      75      80
Cys Thr Lys Ala Gly Gly Gly Lys Val Val Val Pro Ala Gly Thr Tyr
      85      90      95
Leu Thr Gly Ala Ile His Leu Lys Ser Asn Val Asn Leu Glu Val Ser
      100      105      110
Glu Gly Ala Thr Ile Lys Phe Ser Gln Asp Pro Lys His Tyr Leu Pro
      115      120      125
Val Val Phe Ser Arg Trp Glu Gly Val Glu Val Phe Asn Tyr Ser Pro
      130      135      140
Phe Ile Tyr Ala Phe Glu Gln Arg Asn Ile Ala Ile Thr Gly Lys Gly
 145      150      155      160
Thr Leu Asp Gly Gln Ser Asp Ser Glu His Trp Trp Pro Trp Asn Gly
      165      170      175
Arg Pro Gln Tyr Gly Trp Lys Glu Gly Met Lys Gln Gln Arg Pro Asp
      180      185      190
Arg Asn Ala Leu Phe Thr Met Ala Glu Lys Gly Val Pro Val Arg Glu
      195      200      205
Arg Ile Phe Gly Glu Gly His Tyr Leu Arg Pro Gln Phe Ile Gln Pro
      210      215      220
Tyr Arg Cys Gln Asn Val Leu Ile Gln Gly Val Thr Ile Arg Asn Ser
 225      230      235      240
Pro Met Trp Glu Ile His Pro Val Leu Cys Arg Asn Val Thr Ile His
      245      250      255
Asp Val His Ile Asp Ser His Gly Pro Asn Asn Asp Gly Cys Asn Pro
      260      265      270
Glu Ser Cys Ser Asp Val Leu Ile Lys Asp Ser Tyr Phe Asp Thr Gly
      275      280      285
Asp Asp Cys Ile Ala Ile Lys Ser Gly Arg Asn Ala Asp Gly Arg Arg
      290      295      300
Leu Lys Ala Pro Thr Glu Asn Ile Ile Val Gln Gly Cys Arg Met Lys
 305      310      315      320
Asp Gly His Gly Gly Ile Thr Val Gly Ser Glu Ile Ser Gly Gly Val
      325      330      335
Arg Asn Leu Phe Ala Glu Asn Cys Arg Leu Asp Ser Pro Asn Leu Asp
      340      345      350
His Ala Leu Arg Val Lys Asn Asn Ala Met Arg Gly Gly Leu Leu Glu
      355      360      365

```

Asn Phe His Phe Arg Asn Ile Glu Val Gly Gln Val Ala His Ala Val
 370 375 380
 Ile Thr Ile Asp Phe Asn Tyr Glu Glu Gly Ala Lys Gly Ser Phe Thr
 385 390 395 400
 Pro Val Val Arg Asp Tyr Thr Val Asp Arg Leu Arg Ser Thr Lys Ser
 405 410 415
 Lys His Ala Leu Asp Val Gln Gly Leu Pro Gly Ala Pro Val Ile Asn
 420 425 430
 Leu Arg Leu Thr Asn Cys Thr Phe Asn Asp Val Gln Gln Pro Asn Ile
 435 440 445
 Leu Lys Asn Val Glu Gln Ser Thr Phe Glu Asn Val Thr Ile Asn Gly
 450 455 460
 Lys Thr Ile Thr Gln Thr Gly Ser Lys Glu
 465 470

<210> 111
 <211> 1440
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<400> 111
 atgcaaaatc gtcgagaatt tttaacaact ttattttgccg gtgccgggtgc cggacttgtt 60
 ttgccgcaga tttcttttcgg gcagactaaa caagccgacg cctggacgac cgagtatccg 120
 aagatttttag ccagaatcaa accgcgcgaaa tttcgcaaaa aagattttcc gatcaccaaaa 180
 tatggagccg ttgcggacgg gaaaaccctg ggcacgcgaaa gcatcaaaaa agccatcgaa 240
 gcggtgcgcca aatcgggcgg cgggcgcgctc gtcgtgcccc agggagaatt tttgaccggc 300
 gcgattcatt tgaaatcaaaa cgtcaatctg cacatcacga aaggcgcgac cgtcaaattt 360
 tccaccaacc cgaaagatta tctgccgatc gttcacacgc gctgggaagg gatggaattg 420
 atgcatattt cgcctttaat ttatgcctac gagcaaaacca acatcgccgt caccggcgag 480
 ggaacgctcg acgggcaggg caaggctttt ttctggaaat ggcacggaaa cccgcgctac 540
 ggcggaaatc cggatgtgat cagccagcgt cgggcgcgcg ccgggctgta tgaaatgatg 600
 gaaaaaggcg tgcctgtggc ggagcggatt ttccggcgaat ctcatgtatc tcgcccgcag 660
 tttatccagc cctataaatg caaaaatgtt ttgatcgaag gcgttaaaat catcgattcg 720
 cogatgtggg aagttcaccc cgttttgtgc gaaaacgtga cgatccgaaa acttcatatt 780
 tctacccacg gaccgaacaa cgacgggtgc gatccggaaa gctgcaagga cgttttgatc 840
 gaagactgct atttcgacac cggcgacgat tgcattgcca tcaaggcggg gcgcaatgaa 900
 gacgggcgac gcatcaatgt tccgaccgaa aacgtcgtcg tgcgcgggtg cgtgatgaag 960
 gacggtcacg gcggaatcac catcggaagc gagatttccg gcggcgtgcg aaatgttttc 1020
 gcggaaaaca accggctcga cagcgcggat ttgtggactg cgtgagagt gaaaaacaac 1080
 gcttcgcgcg gcggaaaact ggagaatttt tacttcgcg atatcacgt cgggcaggtc 1140
 tcgcgcgccc tcgtcgaaat agatttttaac tacgaggaag gcgctaaagg aaaacacacg 1200
 ccggtcgttc gcaattacgt ggtcgaaaat ctaacctgca ataaaggcaa tcgagcggtc 1260
 gatctgcagg gcttgacaa cgccccgatt tacgacatca cgatgaaaaa ctgtacgttt 1320
 aacgtggtcg aaaagccgag cgtcgtgaaa aacgtcaaag gcgtcaaact ggaaaacgtg 1380
 aagattaacg gcaaagtcgt cgagagtctg gaaaatgctg caacgacggc taaaaataa 1440

<210> 112
 <211> 479
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

<221> SEÑAL
 <222> (1)...(27)

<221> DOMINIO
 <222> (82)...(461)

<223> Dominio catalítico

<400> 112

```

Met Gln Asn Arg Arg Glu Phe Leu Gln Leu Leu Phe Ala Gly Ala Gly
 1           5           10           15
Ala Gly Leu Val Leu Pro Gln Ile Ser Phe Gly Gln Thr Lys Gln Ala
 20           25           30
Asp Ala Trp Thr Thr Glu Tyr Pro Lys Ile Leu Ala Arg Ile Lys Pro
 35           40           45
Pro Lys Phe Arg Lys Lys Asp Phe Pro Ile Thr Lys Tyr Gly Ala Val
 50           55           60
Ala Asp Gly Lys Thr Leu Ala Thr Glu Ser Ile Lys Lys Ala Ile Glu
 65           70           75           80
Ala Cys Ala Lys Ser Gly Gly Gly Arg Val Val Val Pro Gln Gly Glu
 85           90           95
Phe Leu Thr Gly Ala Ile His Leu Lys Ser Asn Val Asn Leu His Ile
100           105           110
Thr Lys Gly Ala Thr Val Lys Phe Ser Thr Asn Pro Lys Asp Tyr Leu
115           120           125
Pro Ile Val His Thr Arg Trp Glu Gly Met Glu Leu Met His Ile Ser
130           135           140
Pro Leu Ile Tyr Ala Tyr Glu Gln Thr Asn Ile Ala Val Thr Gly Glu
145           150           155           160
Gly Thr Leu Asp Gly Gln Gly Lys Ala Phe Phe Trp Lys Trp His Gly
165           170           175
Asn Pro Arg Tyr Gly Gly Asn Pro Asp Val Ile Ser Gln Arg Pro Ala
180           185           190
Arg Ala Arg Leu Tyr Glu Met Met Glu Lys Gly Val Pro Val Ala Glu
195           200           205
Arg Ile Phe Gly Glu Thr Gln Tyr Leu Arg Pro Gln Phe Ile Gln Pro
210           215           220
Tyr Lys Cys Lys Asn Val Leu Ile Glu Gly Val Lys Ile Ile Asp Ser
225           230           235           240
Pro Met Trp Glu Val His Pro Val Leu Cys Glu Asn Val Thr Ile Arg
245           250           255
Lys Leu His Ile Ser Thr His Gly Pro Asn Asn Asp Gly Cys Asp Pro
260           265           270
Glu Ser Cys Lys Asp Val Leu Ile Glu Asp Cys Tyr Phe Asp Thr Gly
275           280           285
Asp Asp Cys Ile Ala Ile Lys Ala Gly Arg Asn Glu Asp Gly Arg Arg
290           295           300
Ile Asn Val Pro Thr Glu Asn Val Val Val Arg Gly Cys Val Met Lys
305           310           315           320
Asp Gly His Gly Gly Ile Thr Ile Gly Ser Glu Ile Ser Gly Gly Val
325           330           335
Arg Asn Val Phe Ala Glu Asn Asn Arg Leu Asp Ser Ala Asp Leu Trp
340           345           350
Thr Ala Leu Arg Val Lys Asn Asn Ala Ser Arg Gly Gly Lys Leu Glu
355           360           365
Asn Phe Tyr Phe Arg Asp Ile Thr Val Gly Gln Val Ser Arg Ala Val
370           375           380
Val Glu Ile Asp Phe Asn Tyr Glu Glu Gly Ala Lys Gly Lys His Thr
385           390           395           400
Pro Val Val Arg Asn Tyr Val Val Glu Asn Leu Thr Cys Asn Lys Gly
405           410           415
Asn Arg Ala Val Asp Leu Gln Gly Leu Asp Asn Ala Pro Ile Tyr Asp

           420           425           430
Ile Thr Met Lys Asn Cys Thr Phe Asn Val Val Glu Lys Pro Ser Val
           435           440           445
Val Lys Asn Val Lys Gly Val Lys Leu Glu Asn Val Lys Ile Asn Gly
           450           455           460
Lys Val Val Glu Ser Leu Glu Asn Ala Ala Thr Thr Ala Lys Lys
465           470           475

```

<210> 113
 <211> 1017
 <212> ADN
 <213> Desconocido

5

<220>
 <223> Obtenido de una muestra ambiental

<400> 113

10

```

atgaagatat ttttaacaat attgctctcg gcattattca gcatttcaaa tgcacaggtg      60
ctatcggatc ctggtgcgga tcgtatgacc agctaccaac ttaaaaaacgg aggctggccg      120
aagcacttgg ccgataaatc tgttggttaac tattcaaaac ctctctcacc tgctttgcaa      180
aaagtcacgc atcaatcgac cgaaaagtct gcgacaattg ataataatgc aaccacacgt      240
gagataaacc atcttctcct cgcttattcc aaaaccaaca atgacaagta tcttcaagcg      300
gcgacaaaag gtggttgagta tatcctgagt gctcaaaatg acaaaggagg atggcctcaa      360
tattatccag acagtagctc atatcggtgt cagatcacct acaatgacgg cgcgatgatt      420
aatgtattgg aaattttact ttccatatca acaaaacaag agccctatgc tgttctaacg      480
aataaattta acgaaagaat agaaagggcc ttaacacgag ggattcactg catcttacia      540
acccagggtta aacaaggaga taaactaacc atctggggccg cacagtacga tcagaaaaca      600
atggaacctg ctcaagccag actggttgaa ccggtagcgt tagcgacagc ggaatcggcg      660
ggcattctcc gctttttaat gcgtcttgac catcctactc ccgaaataaa aaatgcaatc      720
aaccacgctg tagaatggtt ttccctccat aaagaggtag gctatgatta cgttaaaacg      780
gaaaaaaacg gaaaactttt gcgggatttg gtttcttcgc cggcctctac cgtatgggca      840
agattttatg acatcaggac gaatcaacct atctttggtg atcgcgataa tacgataaag      900
tattcgctga atgaaataag cgaggaacga caaatggct actcttggtg tggtaactgg      960
ccagaaaaga taattacaaa agaatatgaa aaatggctta agaaggtaaa tgaataa      1017
    
```

<210> 114
 <211> 338
 <212> PRT
 <213> Desconocido

15

<220>
 <223> Obtenido de una muestra ambiental

20

<221> SEÑAL
 <222> (1)...(18)

<221> DOMINIO
 <222> (19)...(388)
 <223> Dominio catalítico

25

<400> 114

30

```

Met Lys Ile Phe Leu Thr Ile Leu Leu Ser Ala Leu Phe Ser Ile Ser
 1          5          10          15
Asn Ala Gln Val Leu Ser Asp Pro Val Ala Asp Arg Met Thr Ser Tyr
          20          25          30
Gln Leu Lys Asn Gly Gly Trp Pro Lys His Leu Ala Asp Lys Ser Val
          35          40          45
Val Asn Tyr Ser Lys Pro Leu Ser Pro Ala Leu Gln Lys Val Ile Asp
          50          55          60
    
```

Gln Ser Thr Glu Lys Ser Ala Thr Ile Asp Asn Asn Ala Thr Thr Arg
65 70 75 80
Glu Ile Asn His Leu Leu Leu Ala Tyr Ser Lys Thr Asn Asn Asp Lys
85 90 95
Tyr Leu Gln Ala Ala Thr Lys Gly Val Glu Tyr Ile Leu Ser Ala Gln
100 105 110
Asn Asp Lys Gly Gly Trp Pro Gln Tyr Tyr Pro Asp Ser Ser Ser Tyr
115 120 125
Arg Gly Gln Ile Thr Tyr Asn Asp Gly Ala Met Ile Asn Val Leu Glu
130 135 140
Ile Leu Leu Ser Ile Ser Thr Lys Gln Glu Pro Tyr Ala Val Leu Thr
145 150 155 160
Asn Lys Phe Asn Glu Arg Ile Glu Arg Ala Leu Thr Arg Gly Ile His
165 170 175
Cys Ile Leu Gln Thr Gln Val Lys Gln Gly Asp Lys Leu Thr Ile Trp
180 185 190
Ala Ala Gln Tyr Asp Gln Lys Thr Met Glu Pro Ala Gln Ala Arg Leu
195 200 205
Phe Glu Pro Val Ala Leu Ala Thr Ala Glu Ser Ala Gly Ile Leu Arg
210 215 220
Phe Leu Met Arg Leu Asp His Pro Thr Pro Glu Ile Lys Asn Ala Ile
225 230 235 240
Asn His Ala Val Glu Trp Phe Ser Ser His Lys Glu Val Gly Tyr Asp
245 250 255
Tyr Val Lys Thr Glu Lys Asn Gly Lys Leu Leu Arg Asp Leu Val Ser
260 265 270
Ser Pro Ala Ser Thr Val Trp Ala Arg Phe Tyr Asp Ile Arg Thr Asn
275 280 285
Gln Pro Ile Phe Gly Asp Arg Asp Asn Thr Ile Lys Tyr Ser Leu Asn
290 295 300
Glu Ile Ser Glu Glu Arg Gln Asn Gly Tyr Ser Trp Tyr Gly Asn Trp
305 310 315 320
Pro Glu Lys Ile Ile Thr Lys Glu Tyr Glu Lys Trp Leu Lys Lys Val
325 330 335
Asn Glu

<210> 115
<211> 996
5 <212> ADN
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

10 <400> 115

gtggccaagg	cgatcggcgg	tccgttgccg	ccggcaccag	ggcagggatc	gccggtaacg	60
tgggcgacga	ttctccggca	gccatcgccg	tggtacgcgt	ccgcggacgc	gaaggcgggt	120
gccgaaaccg	tgcgcgcgag	ccagagagcc	accggcggct	ggccgaagaa	cacggattgg	180
acggcgctcc	agagcgacgc	tgagcggcag	gcgctgcgaa	atgccgcgcg	cgagaccgat	240
tcgacgatcg	acaatggcgc	cacggtcacc	gagettccgt	ttctcaccgc	cgtgtatgtc	300
gccacgcgcg	acgagctttt	acgggaggcc	gtgcttcgcg	gcctcgacta	cctgctggcg	360
tcgcagtaca	gcaacggcgg	ctggccacaa	tactttccgt	tgccgaccga	ttactcgccg	420
gacatcacgt	tcaacgacga	cgcgatgacc	ggcgtgggtg	tgctgctgaa	ggatgccgcg	480
gacgggtcag	caggtttcga	attcgtcgac	aaggcgagac	gtgaccgcgc	tgccgcggcc	540
gtgacgcgcg	ccatcgcggt	gatcctccgc	acgcagattc	gggtcaacgg	tacgctgacc	600
ggctgggtgcc	agcagtacga	cgccgacgcg	ctgacgcgcg	cgcgcgggcg	ctcgtacgag	660
catcgcgcga	ttgcgagccg	cgagacggtc	gggatcgcg	ggctgctgat	gggcgtgccg	720
aatcgcgcgc	cagagatcgt	ggctgcgcgt	gacgcggcgt	ccgcatgggt	gggtaaatcg	780

gaactgaagg gtgtgcccga ggcgacggcg ccaggacttt gggcgcgctt ctacgacatc 840
gctacgaatc ggcgatctta ttccggccgc gacggcgctca tcaagtaccg gctcgacgag 900
atcgagctcg agcggcgcac aggtacagc tgggttgcc cgtacgccgc ggcatttctg 960
acgaccgaat atccgaaatg gcgggcggca cgatga 996

<210> 116
<211> 331
<212> PRT
<213> Desconocido

<220>
<223> Obtenido de una muestra ambiental

<221> DOMINIO
<222> (1)...(331)
<223> Dominio catalítico

<400> 116

Met	Ala	Lys	Ala	Ile	Gly	Gly	Pro	Leu	Pro	Pro	Ala	Pro	Gly	Gln	Gly
1				5					10					15	
Ser	Pro	Val	Thr	Trp	Ala	Thr	Ile	Leu	Arg	Gln	Pro	Ser	Pro	Trp	Tyr
			20					25					30		
Ala	Ser	Ala	Asp	Ala	Lys	Ala	Val	Ala	Glu	Thr	Val	Arg	Ala	Ser	Gln
		35					40					45			
Arg	Ala	Thr	Gly	Gly	Trp	Pro	Lys	Asn	Thr	Asp	Trp	Thr	Ala	Leu	Gln
	50					55				60					
Ser	Asp	Ala	Glu	Arg	Gln	Ala	Leu	Arg	Asn	Ala	Arg	Ala	Glu	Thr	Asp
65				70					75					80	
Ser	Thr	Ile	Asp	Asn	Gly	Ala	Thr	Val	Thr	Glu	Leu	Arg	Phe	Leu	Thr
			85					90					95		
Arg	Val	Tyr	Val	Ala	Thr	Arg	Asp	Glu	Leu	Leu	Arg	Glu	Ala	Val	Leu
		100						105					110		
Arg	Gly	Leu	Asp	Tyr	Leu	Leu	Ala	Ser	Gln	Tyr	Ser	Asn	Gly	Gly	Trp
		115					120					125			
Pro	Gln	Tyr	Phe	Pro	Leu	Arg	Thr	Asp	Tyr	Ser	Arg	Asp	Ile	Thr	Phe
	130					135					140				
Asn	Asp	Asp	Ala	Met	Thr	Gly	Val	Val	Leu	Leu	Lys	Asp	Ala	Ala	
145				150					155					160	
Asp	Gly	Ser	Ala	Gly	Phe	Glu	Phe	Val	Asp	Lys	Ala	Arg	Arg	Asp	Arg
			165					170					175		
Ala	Ala	Ala	Ala	Val	Thr	Arg	Ala	Ile	Ala	Val	Ile	Leu	Arg	Thr	Gln
		180					185						190		
Ile	Arg	Val	Asn	Gly	Thr	Leu	Thr	Gly	Trp	Cys	Gln	Gln	Tyr	Asp	Ala
	195						200					205			
Asp	Ala	Leu	Thr	Pro	Ala	Arg	Gly	Arg	Ser	Tyr	Glu	His	Pro	Ser	Ile
210					215						220				
Ala	Ser	Arg	Glu	Thr	Val	Gly	Ile	Ala	Arg	Leu	Leu	Met	Gly	Val	Pro
225				230						235				240	
Asn	Pro	Ser	Pro	Glu	Ile	Val	Ala	Ala	Val	Asp	Ala	Ala	Ala	Ala	Trp
			245					250					255		
Leu	Gly	Lys	Ser	Glu	Leu	Lys	Gly	Val	Pro	Glu	Ala	Thr	Ala	Pro	Gly
		260					265						270		
Leu	Trp	Ala	Arg	Phe	Tyr	Asp	Ile	Ala	Thr	Asn	Arg	Pro	Ile	Tyr	Ser
	275						280					285			
Gly	Arg	Asp	Gly	Val	Ile	Lys	Tyr	Arg	Leu	Asp	Glu	Ile	Glu	Leu	Glu
290					295					300					
Arg	Arg	Thr	Gly	Tyr	Ser	Trp	Val	Gly	Pro	Tyr	Ala	Ala	Ala	Phe	Leu
305				310						315				320	
Thr	Thr	Glu	Tyr	Pro	Lys	Trp	Arg	Ala	Ala	Arg					

<210> 117
 <211> 1725
 <212> ADN
 <213> Desconocido

5

<220>
 <223> Obtenido de una muestra ambiental.

<400> 117

10

```

atgaagaatt ttgggtttgg taactacaag tttttttagt cggcaatgtc tgtcgcgtct 60
ttttcgtatg cggcaagcta tacacccccg tcaacagcag tttegaaaat caacagctat 120
cgaggctatt cggagctgac ttcagctgca tccggcatgg atatcgacca gtacacctac 180
aacatgacca cttggcfaat cgcaaacggc ggtttttaca aagccatggc cgacaagtat 240
aaaagcgcggt atggcgggcg tcaaaaatcc gaatggcaag ctaaaggcgg tggcgacctc 300
ggcactatag acaacaacgc caccatccag gaaatgcgtt tgcgcgcgt gcgttacaaa 360
gaaacgacga acaacaatta caaatccgca ttttaagacaa gtttcaacaa ggcggtcaat 420
ttttttttga ccatgcagcg ctccaaaggc ggactccac aagtttgcc caaacgcggc 480
aactattctg accaaatcac gctaaatgac aacgccatga tccgcgccat gggtcacgatg 540
atggatatcg ccaacaagac gagtccattt gattcggata tcatcgacga cgcacccgc 600
agcaaatga aatcggtctt cgacaaagcg gtcgattact tgcacaaggc gcaaatcgtg 660
aacgacggaa aggtcacggt atggtgcgcc cagcacgaca ccaacagcct cgcgcccgta 720
ggcgacgag cctacgaact cccgagcaaa tccggcaacg aatccatggg cgttgtgtgg 780
tttttgatga actggccaga ccqaaacgaa gcaatccaga aggcgggtcaa aggcgcaatc 840
gcttggtaca aaaagaataa actaaaagac aaggcggtta gcaagaccgc aggcgttgtg 900
gacaaggcgg gttcatcgct gtggttccgc ttttacgaag tcaacaacga caactacttt 960
ttctgcgacc gcgatggtgc tagcaccaag acgcaggact tcatgaaaat cagcgaagaa 1020
cgtcgaagg gctaccagtg ggcaggcgat tacggctctg caattctagg caccgaaaat 1080
gcataccttg aagcactcgc caagatggac gacaactatg ttccacctcc gccagacca 1140
gctatgtgcg gaaacgacac ttgcaaaacg tacatcgatg gcgttgactt tattgacatt 1200
caaggcgtca aggaaacaac caacacggga ttcggtggcg aaggttacgc caacgttgac 1260
aactccaccg gaagctatgt gacctacggc gtcaccgcgt tcaaggaagg caaatacact 1320
ttgttcacga gctttgcaaa cggcggtggt tccgcacgcg gttacagcgt ttctgcagga 1380
gacaagacgt tacttgaga cggcagcatg gaatctacag ccgcattggac cacttgaaa 1440
atgcaatcca tcgaaatcga attgccaatg ggctatagcg aactcaagtt cacaagcctt 1500
tcgaaagacg gtatggcgaa catcgattac atcggctgga tgaacgatga tttgaaagtt 1560
ggcgaagttg aagtaccacg ctcatccatt gaagcaatac gcgccatccg caaagcccag 1620
caggacaacc gctactttgt ggactttggc ggcaacaata atagcgcagg ggcttacttt 1680
aagcgtggca tcaacacgtt ccgcgtgaat gggaagatga ggtaa 1725
    
```

<210> 118
 <211> 574
 <212> PRT
 <213> Desconocido

15

<220>
 <223> Obtenido de una muestra ambiental.

20

<221> SEÑAL
 <222> (1)...(24)

25

<221> DOMINIO
 <222> (25)...(574)
 <223> Dominio catalítico

<400> 118

Met Lys Asn Phe Gly Phe Gly Asn Tyr Lys Phe Phe Val Ala Ala Met
 1 5 10 15

Ser	Val	Ala	Ser	Phe	Ser	Tyr	Ala	Ala	Ser	Tyr	Thr	Pro	Pro	Ser	Thr
			20					25					30		
Ala	Val	Ser	Lys	Ile	Asn	Ser	Tyr	Arg	Gly	Tyr	Ser	Glu	Leu	Thr	Ser
		35					40					45			
Ala	Ala	Ser	Gly	Met	Asp	Ile	Asp	Gln	Tyr	Thr	Tyr	Asn	Met	Thr	Thr
		50				55						60			
Trp	Gln	Ile	Ala	Asn	Gly	Gly	Phe	Tyr	Lys	Ala	Met	Ala	Asp	Lys	Tyr
65					70					75					80
Lys	Ser	Ala	Tyr	Gly	Gly	Gly	Gln	Lys	Ser	Glu	Trp	Gln	Ala	Lys	Gly
				85					90					95	
Gly	Gly	Asp	Leu	Gly	Thr	Ile	Asp	Asn	Asn	Ala	Thr	Ile	Gln	Glu	Met
			100					105					110		
Arg	Leu	Leu	Ala	Val	Arg	Tyr	Lys	Glu	Thr	Thr	Asn	Asn	Asn	Tyr	Lys
		115					120					125			
Ser	Ala	Phe	Lys	Thr	Ser	Phe	Asn	Lys	Ala	Val	Asn	Phe	Leu	Leu	Thr
		130					135				140				
Met	Gln	Arg	Ser	Lys	Gly	Gly	Leu	Pro	Gln	Val	Trp	Pro	Lys	Arg	Gly
145					150					155					160
Asn	Tyr	Ser	Asp	Gln	Ile	Thr	Leu	Asn	Asp	Asn	Ala	Met	Ile	Arg	Ala
				165					170					175	
Met	Val	Thr	Met	Met	Asp	Ile	Ala	Asn	Lys	Thr	Ser	Pro	Phe	Asp	Ser
			180					185					190		
Asp	Ile	Ile	Asp	Asp	Ala	Thr	Arg	Ser	Lys	Met	Lys	Ser	Ala	Leu	Asp
		195					200					205			
Lys	Ala	Val	Asp	Tyr	Leu	Leu	Lys	Ala	Gln	Ile	Val	Asn	Asp	Gly	Lys
		210					215				220				
Val	Thr	Val	Trp	Cys	Ala	Gln	His	Asp	Thr	Asn	Ser	Leu	Ala	Pro	Val
225					230					235					240
Gly	Ala	Arg	Ala	Tyr	Glu	Leu	Pro	Ser	Lys	Ser	Gly	Asn	Glu	Ser	Met
				245					250					255	
Gly	Val	Val	Trp	Phe	Leu	Met	Asn	Trp	Pro	Asp	Gln	Asn	Glu	Ala	Ile
			260					265					270		
Gln	Lys	Ala	Val	Lys	Gly	Ala	Ile	Ala	Trp	Tyr	Lys	Lys	Asn	Lys	Leu
		275					280						285		
Lys	Asp	Lys	Ala	Phe	Ser	Lys	Thr	Ala	Gly	Val	Val	Asp	Lys	Ala	Gly
		290				295					300				
Ser	Ser	Leu	Trp	Phe	Arg	Phe	Tyr	Glu	Val	Asn	Asn	Asp	Asn	Tyr	Phe
305					310					315					320
Phe	Cys	Asp	Arg	Asp	Gly	Ala	Ser	Thr	Lys	Thr	Gln	Asp	Phe	Met	Lys
				325					330					335	
Ile	Ser	Glu	Glu	Arg	Arg	Lys	Gly	Tyr	Gln	Trp	Ala	Gly	Asp	Tyr	Gly
			340					345					350		
Ser	Ala	Ile	Leu	Gly	Thr	Glu	Asn	Ala	Tyr	Leu	Glu	Ala	Leu	Ala	Lys
		355					360					365			
Met	Asp	Asp	Asn	Tyr	Val	Pro	Pro	Pro	Pro	Ala	Pro	Ala	Met	Cys	Gly
		370				375					380				
Asn	Asp	Thr	Cys	Lys	Thr	Tyr	Ile	Asp	Gly	Val	Asp	Phe	Ile	Asp	Ile
385					390					395					400
Gln	Gly	Val	Lys	Glu	Thr	Thr	Asn	Thr	Gly	Phe	Val	Gly	Glu	Gly	Tyr
				405					410					415	
Ala	Asn	Val	Asp	Asn	Ser	Thr	Gly	Ser	Tyr	Val	Thr	Tyr	Gly	Val	Thr
			420					425					430		
Ala	Phe	Lys	Glu	Gly	Lys	Tyr	Thr	Leu	Phe	Ile	Ser	Phe	Ala	Asn	Gly
		435					440					445			
Gly	Gly	Ser	Ala	Arg	Gly	Tyr	Ser	Val	Ser	Ala	Gly	Asp	Lys	Thr	Leu
		450				455					460				
Leu	Ala	Asp	Gly	Ser	Met	Glu	Ser	Thr	Ala	Ala	Trp	Thr	Thr	Trp	Lys
465					470					475					480
Met	Gln	Ser	Ile	Glu	Ile	Glu	Leu	Pro	Met	Gly	Tyr	Ser	Glu	Leu	Lys

				485					490					495					
Phe	Thr	Ser	Leu	Ser	Lys	Asp	Gly	Met	Ala	Asn	Ile	Asp	Tyr	Ile	Gly				
			500					505					510						
Trp	Met	Asn	Asp	Asp	Leu	Lys	Val	Gly	Glu	Val	Glu	Val	Pro	Arg	Ser				
		515					520					525							
Ser	Ile	Glu	Ala	Ile	Arg	Ala	Ile	Arg	Lys	Ala	Gln	Gln	Asp	Asn	Arg				
	530					535					540								
Tyr	Phe	Val	Asp	Phe	Gly	Gly	Asn	Asn	Asn	Ser	Ala	Gly	Ala	Tyr	Phe				
545					550					555					560				
Lys	Arg	Gly	Ile	Asn	Thr	Phe	Arg	Val	Asn	Gly	Lys	Met	Arg						
				565					570										

<210> 119
 <211> 1848
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental.

<400> 119

gtgtcatggc	aggaatccgg	tgcggctatc	accaaagcct	ggaatgcaac	gctcagtggc	60
tcaaaccctt	acacagccgt	atccgctggg	tggaaatggca	cacttgcccc	caatgcatcg	120
gccacttttg	gtttccaggc	aaacggttct	gccgggtgcac	ctaaagtga	tggcagcttg	180
tgcggcacca	acacttccat	aacacgggca	tccagcagtg	ttgccagctc	gggttaaata	240
agcgcgcgcc	tatcgctccg	cagcagatca	tccagttcaa	tcgctatcac	tagcagctct	300
ttagcgagaa	gttctattgc	ctccagcagc	tactagttta	gtagctccag	agcgagcagt	360
agtgcgccaa	gcgtttttct	ttttacgata	caggaagagc	aagcgggctt	ctgtcgtgtt	420
gatggcattg	cgacagaaag	caccaacacc	ggttttaccg	gcaatggcta	caccaatgcg	480
aacaacgcgc	aaggcgcagc	gattgaatgg	gcagtcagcg	cacctagcag	tggccggttat	540
acagtagcct	tccgcttcgc	caatggcggc	acagcagcgc	gcaacggctc	gttggttaata	600
aatggcggtg	gcaatggtaa	ttacactgtg	gagttacccc	tgaccggcgc	atgggcaacc	660
tggcaaattg	ccagcgtgga	aattgattta	gtgcaaggca	ataatatttt	aaaactctcg	720
gcgttaaccg	ctgacggttt	ggccaatatc	gactcattaa	aaatagacgg	cgcgcaaac	780
aaagcaggtg	cttgacgac	tacatcaagc	agcagcgttg	ccagcagctc	gtcgtccgtt	840
aaatccagcg	caagttcttc	ttcgagttca	tccacgctg	caaaaatact	gacattagac	900
ggtaaccggg	ccgccagctg	gttcaacaaa	tccaggacca	agtgggaatg	cagccgcgcc	960
gatattgtgt	tgtcttacc	gcaatccaac	ggcgggtggc	caaaaaacct	ggattacaac	1020
tcagttagcg	caggcaatgg	cgggagcgac	agcggcacca	tcgacaatgg	tgcaaccatt	1080
accgaaatgg	tttacctcgc	tgaattttat	aaaaacggcg	gcaacaccaa	atatcgcgat	1140
gcagtgcgca	gagcagcaaa	cttttttagt	agctcgcaat	acagcacagg	cgccttgcca	1200
caattttatc	cgttgaaagg	cggctatgcg	gatcatgcga	cctttaacga	taacggcatg	1260
gcgtacgcgt	tgacggtatt	ggatttcgca	gtaaacaaac	gcgcacogtt	tgataacgac	1320
attttctctg	attctgatcg	ggcgaaattc	aaaacgcgtg	ttgccaaagg	tgtggattac	1380
attttaaaag	cgcagtggaa	acaaaatgga	aaactcactg	catgggtgtg	acaacacggt	1440
gctacggatt	accaaccgaa	aaaagcgcgc	gcttatgaat	tggaaatcatt	gagtggtagc	1500
gagtcggctg	gcattctcgc	cttcttgatg	acccaaccac	aaaccgcgca	aatcgaagcg	1560
gcgggtcaagg	cgggtgtcaa	ctgggttcgcc	agtccaaata	cttatttggc	taactacact	1620
tacgattcat	caaaagcgtc	taccaaccgg	attgtgtata	aatccggaag	cagaatgtgg	1680
tatcgcttct	atgacctgaa	caccaaccgt	ggtttcttta	gtgatcgoga	tggcagcaaa	1740
ttctatgata	tcaccaaat	gtcagaagag	cgtcgaccgg	gttatagctg	gggtggctct	1800
tacgggtgaat	ctattatttc	cttcgcgcaa	aaagtgggtt	atctgtaa		1848

<210> 120
 <211> 615
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental.

<221> UNIÓN
 <222> (1)...(61)

<223> Módulo de unión a carbohidratos

<221> UNIÓN

<222> (139)...(257)

5 <223> Módulo de unión a carbohidratos

<221> DOMINIO

<222> (258)...(615)

<223> Dominio catalítico

10

<400> 120

Met	Ser	Trp	Gln	Glu	Ser	Gly	Ala	Ala	Ile	Thr	Asn	Ala	Trp	Asn	Ala
1				5					10					15	
Thr	Leu	Ser	Gly	Ser	Asn	Pro	Tyr	Thr	Ala	Val	Ser	Ala	Gly	Trp	Asn
			20					25					30		
Gly	Thr	Leu	Ala	Pro	Asn	Ala	Ser	Ala	Thr	Phe	Gly	Phe	Gln	Ala	Asn
		35					40					45			
Gly	Ser	Ala	Gly	Ala	Pro	Lys	Val	Asn	Gly	Ser	Leu	Cys	Gly	Thr	Asn
	50					55					60				
Thr	Ser	Ser	Thr	Pro	Ala	Ser	Ser	Ser	Val	Ala	Ser	Ser	Val	Lys	Ser
65				70					75					80	
Ser	Ala	Pro	Val	Ser	Ser	Ser	Ser	Arg	Ser	Ser	Ser	Ser	Ile	Ala	Ile
			85					90					95		
Thr	Ser	Ser	Ser	Leu	Ala	Arg	Ser	Ser	Ile	Ala	Ser	Ser	Ser	Ser	Leu
			100					105					110		
Val	Ser	Ser	Ser	Arg	Ala	Ser	Ser	Ser	Ala	Pro	Ser	Val	Phe	Ser	Phe
			115					120				125			
Thr	Ile	Gln	Glu	Glu	Gln	Ala	Gly	Phe	Cys	Arg	Val	Asp	Gly	Ile	Ala
	130					135					140				
Thr	Glu	Ser	Thr	Asn	Thr	Gly	Phe	Thr	Gly	Asn	Gly	Tyr	Thr	Asn	Ala
145				150						155				160	
Asn	Asn	Ala	Gln	Gly	Ala	Ala	Ile	Glu	Trp	Ala	Val	Ser	Ala	Pro	Ser
			165					170						175	
Ser	Gly	Arg	Tyr	Thr	Val	Ala	Phe	Arg	Phe	Ala	Asn	Gly	Gly	Thr	Ala
		180					185					190			
Ala	Arg	Asn	Gly	Ser	Leu	Leu	Ile	Asn	Gly	Gly	Ser	Asn	Gly	Asn	Tyr
	195						200					205			
Thr	Val	Glu	Leu	Pro	Leu	Thr	Gly	Ala	Trp	Ala	Thr	Trp	Gln	Ile	Ala
	210					215					220				
Ser	Val	Glu	Ile	Asp	Leu	Val	Gln	Gly	Asn	Asn	Ile	Leu	Lys	Leu	Ser
225				230						235				240	
Ala	Leu	Thr	Ala	Asp	Gly	Leu	Ala	Asn	Ile	Asp	Ser	Leu	Lys	Ile	Asp
			245					250					255		
Gly	Ala	Gln	Thr	Lys	Ala	Gly	Thr	Cys	Ser	Thr	Thr	Ser	Ser	Ser	Ser
		260					265					270			
Val	Ala	Ser	Ser	Ser	Ser	Ser	Val	Lys	Ser	Ser	Ala	Ser	Ser	Ser	Ser
	275						280					285			
Ser	Ser	Ser	Thr	Ala	Ala	Lys	Ile	Leu	Thr	Leu	Asp	Gly	Asn	Pro	Ala
	290					295					300				
Ala	Ser	Trp	Phe	Asn	Lys	Ser	Arg	Thr	Lys	Trp	Asn	Ser	Ser	Arg	Ala
305				310						315				320	
Asp	Ile	Val	Leu	Ser	Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro	Lys	Asn
			325					330					335		
Leu	Asp	Tyr	Asn	Ser	Val	Ser	Ala	Gly	Asn	Gly	Gly	Ser	Asp	Ser	Gly
			340					345					350		

Thr	Ile	Asp	Asn	Gly	Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu	Ala	Glu
		355					360					365			
Ile	Tyr	Lys	Asn	Gly	Gly	Asn	Thr	Lys	Tyr	Arg	Asp	Ala	Val	Arg	Arg
	370					375					380				
Ala	Ala	Asn	Phe	Leu	Val	Ser	Ser	Gln	Tyr	Ser	Thr	Gly	Ala	Leu	Pro
385					390					395					400
Gln	Phe	Tyr	Pro	Leu	Lys	Gly	Gly	Tyr	Ala	Asp	His	Ala	Thr	Phe	Asn
				405					410					415	
Asp	Asn	Gly	Met	Ala	Tyr	Ala	Leu	Thr	Val	Leu	Asp	Phe	Ala	Val	Asn
			420					425					430		
Lys	Arg	Ala	Pro	Phe	Asp	Asn	Asp	Ile	Phe	Ser	Asp	Ser	Asp	Arg	Ala
		435					440						445		
Lys	Phe	Lys	Thr	Ala	Val	Ala	Lys	Gly	Val	Asp	Tyr	Ile	Leu	Lys	Ala
	450					455					460				
Gln	Trp	Lys	Gln	Asn	Gly	Lys	Leu	Thr	Ala	Trp	Cys	Ala	Gln	His	Gly
465					470					475					480
Ala	Thr	Asp	Tyr	Gln	Pro	Lys	Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Glu	Ser
				485					490					495	
Leu	Ser	Gly	Ser	Glu	Ser	Val	Gly	Ile	Leu	Ala	Phe	Leu	Met	Thr	Gln
			500					505					510		
Pro	Gln	Thr	Ala	Gln	Ile	Glu	Ala	Ala	Val	Lys	Ala	Gly	Val	Asn	Trp
		515					520						525		
Phe	Ala	Ser	Pro	Asn	Thr	Tyr	Leu	Ala	Asn	Tyr	Thr	Tyr	Asp	Ser	Ser
	530					535					540				
Lys	Ala	Ser	Thr	Asn	Pro	Ile	Val	Tyr	Lys	Ser	Gly	Ser	Arg	Met	Trp
545					550					555					560
Tyr	Arg	Phe	Tyr	Asp	Leu	Asn	Thr	Asn	Arg	Gly	Phe	Phe	Ser	Asp	Arg
				565					570					575	
Asp	Gly	Ser	Lys	Phe	Tyr	Asp	Ile	Thr	Gln	Met	Ser	Glu	Glu	Arg	Arg
			580					585					590		
Thr	Gly	Tyr	Ser	Trp	Gly	Gly	Ser	Tyr	Gly	Glu	Ser	Ile	Ile	Ser	Phe
		595					600						605		
Ala	Gln	Lys	Val	Gly	Tyr	Leu									
	610					615									

<210> 121
 <211> 1047
 5 <212> ADN
 <213> Bacteria

<400> 121

atgatgagat	caagcatcgt	caagctagtt	gctttcagtg	ttgtgggttat	gttatggctc	60
ggtgtatcct	ttcaaacggc	agaagcgaat	acgccaaatt	tcaacttaca	aggctttgcc	120
acgttaaattg	ggggaacaac	tggtgggtgca	ggtggagatg	tagtgacggt	tcgtacaggg	180
aatgaattaa	taaacgcttt	gaagtccaaa	aaccctaatac	gtccgttaac	aatttatgta	240
aacggtagca	taacacctag	taatacgtct	gatagtaaga	tcgatattaa	ggatgtttcc	300
aatgtatcga	ttttaggggg	tggtacaaat	ggacgattaa	atgggatcgg	tattaaagta	360
tggcgagcga	ataatatcat	cattcgcaac	ttgacgatcc	atgaagtcca	tacaggtgat	420
aaagatgcga	ttagcattga	agggccctct	cggaacattt	ggattgacca	taacgagctt	480
tatgccagct	tgaacgttca	taaagaccac	tatgacggct	tgtttgacgt	aaagcgcgat	540
gcttacaata	ttaccttctc	ttggaattat	gtccatgatg	gctggaaagc	gatgctcatg	600
gggaactctg	atagtataa	ctacgaccga	aacataacat	tccaccataa	ctacttcaaa	660
aacttaaaact	ctcgcgtacc	tgcgtaccgt	tttggaaagg	cgcacttggt	tagcaattac	720
tttgagaaca	ttttagaac	aggcattaat	tcacggatgg	gagcggaaat	gctcgttgaa	780
cataacgttt	ttgagaatgc	caccaaccgg	ttaggattct	ggcatagcag	tcgaacaggt	840
tattggaatg	ttgccataa	ccgctatata	aatagcacgg	gtagcatgcc	gaccacttcc	900
acgaccaatt	atcgacctcc	ttatccctat	acggtcacac	cagttggtga	tgtgaaatcg	960
gttgtcacac	gttatgcggg	agttggtgtc	atccagccgt	atgcaagaaa	gocatccgag	1020
cgattgctct	ggtggctttt	tgcataa				1047

<210> 122
 <211> 348

<212> PRT
<213> Bacteria

<220>
<221> SEÑAL
<222> (1)...(29)

<221> DOMINIO
<222> (30)....(348)

10 <223> Dominio catalítico

<400> 122

```

Met Met Arg Ser Ser Ile Val Lys Leu Val Ala Phe Ser Val Val Val
 1      5      10      15
Met Leu Trp Leu Gly Val Ser Phe Gln Thr Ala Glu Ala Asn Thr Pro
 20      25      30
Asn Phe Asn Leu Gln Gly Phe Ala Thr Leu Asn Gly Gly Thr Thr Gly
 35      40      45
Gly Ala Gly Gly Asp Val Val Thr Val Arg Thr Gly Asn Glu Leu Ile
 50      55      60
Asn Ala Leu Lys Ser Lys Asn Pro Asn Arg Pro Leu Thr Ile Tyr Val
 65      70      75      80
Asn Gly Thr Ile Thr Pro Ser Asn Thr Ser Asp Ser Lys Ile Asp Ile
 85      90      95
Lys Asp Val Ser Asn Val Ser Ile Leu Gly Val Gly Thr Asn Gly Arg
100     105     110
Leu Asn Gly Ile Gly Ile Lys Val Trp Arg Ala Asn Asn Ile Ile Ile
115     120     125
Arg Asn Leu Thr Ile His Glu Val His Thr Gly Asp Lys Asp Ala Ile
130     135     140
Ser Ile Glu Gly Pro Ser Arg Asn Ile Trp Ile Asp His Asn Glu Leu
145     150     155     160
Tyr Ala Ser Leu Asn Val His Lys Asp His Tyr Asp Gly Leu Phe Asp
165     170     175
Val Lys Arg Asp Ala Tyr Asn Ile Thr Phe Ser Trp Asn Tyr Val His
180     185     190
Asp Gly Trp Lys Ala Met Leu Met Gly Asn Ser Asp Ser Asp Asn Tyr
195     200     205
Asp Arg Asn Ile Thr Phe His Asn Tyr Phe Lys Asn Leu Asn Ser
210     215     220
Arg Val Pro Ala Tyr Arg Phe Gly Lys Ala His Leu Phe Ser Asn Tyr
225     230     235     240
Phe Glu Asn Ile Leu Glu Thr Gly Ile Asn Ser Arg Met Gly Ala Glu
245     250     255
Met Leu Val Glu His Asn Val Phe Glu Asn Ala Thr Asn Pro Leu Gly
260     265     270
Phe Trp His Ser Ser Arg Thr Gly Tyr Trp Asn Val Ala Asn Asn Arg
275     280     285
Tyr Ile Asn Ser Thr Gly Ser Met Pro Thr Thr Ser Thr Thr Asn Tyr
290     295     300
Arg Pro Pro Tyr Pro Tyr Thr Val Thr Pro Val Gly Asp Val Lys Ser
305     310     315     320
Val Val Thr Arg Tyr Ala Gly Val Gly Val Ile Gln Pro Tyr Ala Arg
325     330     335
Lys Pro Ser Glu Arg Leu Leu Trp Trp Leu Phe Ala

```

340

345

15

<210> 123
<211> 1830
<212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental.

5

<400> 123

ttgagtctac	ttagtgtaat	gaccottttg	cctgtaatgg	caagtaacaa	cgtagctccc	60
tggggctggg	ccacctgctc	cgatgagtc	gcgacagctt	atactctgaa	cggaggttgc	120
ttttctgatg	catcttcogt	tactctgaaa	gctcttggca	atgaacaaac	agatgacaaa	180
caaatacaac	aggctatcgc	tcagaaagac	atcattatct	tagatggttc	caatggcgat	240
ttcatcctta	atgaatacat	caagatttcg	acaaaaaaca	aaaccatcat	tggtatcaac	300
aacgcccggc	tgtgtacaaa	gttctaccta	accgctgatg	atattacgta	ccttaaagca	360
caaggactgg	agggactgag	tagtacaaat	caacatacag	gaactctgcc	tgatggcaca	420
acagtgaact	gtgacgagcg	tgcctttttc	accaagaaag	ccatcatgga	actccaatat	480
cagaaaaacg	gacccataac	cctacccaat	aaatcaggta	tcttttattt	agatgcccgt	540
tctgagaata	tcacatcccg	aaatatctcg	ctgataggcg	caggagccgt	agatatagac	600
ggagctgacc	tgattacca	tcagggttaag	cacgtctgga	ttgaccattg	cacgtttgtg	660
gactctcaag	atggtgccc	ggacagcaag	gtatgcgact	gggccaccta	tacctataac	720
cacttctact	atacagaccg	cagttactca	catgcctaca	cttgcggttg	cggatgggtc	780
agcaatcatg	aaatgggtgat	tcacatgacc	tttgcctgta	atatctgggg	agcaaaatgt	840
atgogtctgc	tgccgcaagc	agatgactgt	ttcatacacc	ttgtgaacaa	ctatcacaac	900
tgtcctggca	atagtgtcgg	tatgaccatt	aacagttaca	gcaaagcatt	ggttgagggg	960
aactatgctg	ctgcaggtgt	caacaagcca	ttagatggca	gtggggccaa	ccgtaatgta	1020
acagctaagg	ataatagttt	tgcaaaactca	caagccggtt	ctggttggtc	tgtgccatac	1080
gactatacca	agattgcagc	cgccgacgtt	ccagctacgc	tgactggaac	agagggtgca	1140
ggcgccacat	taggcaacga	tgcaacatac	attctgtcta	ctattccaac	tgctgaccga	1200
caagaaggcg	aatcttcaat	ctactatttc	attgatggcc	tggtgggaac	taatagtga	1260
ggctattcca	ttatagagtt	taatgatggc	gcaacattgc	tgctgaacaa	taaagagaaa	1320
gcatggtcta	atggttagtgc	aattcaactt	ggtgacgata	attatacgag	tattaaactt	1380
tctaattggag	cagaaaaacat	cctcacagca	cctactggca	aaaaagtaag	tggtattacc	1440
ttctattctt	atatcaatat	aaaagaagaa	aaactcgact	tcaccaaata	tccagaatat	1500
ggtttccgca	cctgtttctg	gcagaaagtt	gccaaactca	cttattctgc	gacttctgat	1560
gacgtacaaa	tcttgaaatc	tcgtgatcca	cagaatactg	acgtggcatc	attccatttc	1620
actccaacaa	atggttgaag	tttcaaaaat	tcagggtgaac	agctttgttt	cttaatgaaa	1680
gtcacctata	gtgatgaaag	cacaggtatc	tctgctatcc	agaaaaaat	gcctatcgat	1740
ggcgttacct	ataaccttca	aggtatccgt	atagataatc	ccaccaaggg	aatctatatt	1800
cagaacggaa	agaaaatcat	tatcaaataa				1830

10 <210> 124

<211> 609

<212> PRT

<213> Desconocido

15 <220>

<223> Obtenido de una muestra ambiental.

<221> SEÑAL

<222> (1)...(21)

20

<221> DOMINIO

<222> (22)...(390)

<223> Dominio catalítico

25 <400> 124

Leu	Ser	Leu	Leu	Ser	Val	Met	Thr	Leu	Leu	Pro	Val	Met	Ala	Ser	Asn
1				5				10					15		
Asn	Val	Ala	Pro	Trp	Gly	Trp	Ala	Thr	Cys	Ser	Asp	Glu	Ser	Ala	Thr
		20						25					30		
Ala	Tyr	Thr	Leu	Asn	Gly	Gly	Cys	Phe	Ser	Asp	Ala	Ser	Ser	Val	Thr
		35					40					45			
Leu	Lys	Ala	Leu	Gly	Asn	Glu	Gln	Thr	Asp	Asp	Lys	Gln	Ile	Lys	Gln
	50				55					60					
Ala	Ile	Ala	Gln	Lys	Asp	Ile	Ile	Ile	Leu	Asp	Gly	Ser	Asn	Gly	Asp
65				70					75					80	
Phe	Ile	Leu	Asn	Glu	Tyr	Ile	Lys	Ile	Ser	Thr	Lys	Asn	Lys	Thr	Ile
			85					90					95		
Ile	Gly	Ile	Asn	Asn	Ala	Arg	Leu	Cys	Thr	Lys	Phe	Tyr	Leu	Thr	Ala
			100					105					110		
Asp	Asp	Ile	Thr	Tyr	Leu	Lys	Ala	Gln	Gly	Leu	Glu	Gly	Leu	Ser	Ser
		115					120					125			
Thr	Asn	Gln	His	Thr	Gly	Thr	Leu	Pro	Asp	Gly	Thr	Thr	Val	Thr	Cys
	130				135					140					
Asp	Glu	Arg	Ala	Phe	Phe	Thr	Lys	Lys	Ala	Ile	Met	Glu	Leu	Gln	Tyr
145				150					155					160	
Gln	Lys	Thr	Gly	Ser	Tyr	Thr	Leu	Pro	Asn	Lys	Ser	Gly	Ile	Phe	Tyr
			165					170					175		
Leu	Asp	Ala	Ala	Ser	Glu	Asn	Ile	Ile	Ile	Arg	Asn	Ile	Ser	Leu	Ile
		180					185					190			
Gly	Pro	Gly	Ala	Val	Asp	Ile	Asp	Gly	Ala	Asp	Leu	Ile	Thr	Asn	Gln
	195						200					205			
Gly	Lys	His	Val	Trp	Ile	Asp	His	Cys	Thr	Phe	Val	Asp	Ser	Gln	Asp
	210				215					220					
Gly	Ala	Leu	Asp	Ser	Lys	Val	Cys	Asp	Trp	Ala	Thr	Tyr	Thr	Tyr	Asn
225				230					235					240	
His	Phe	Tyr	Tyr	Thr	Asp	Arg	Ser	Tyr	Ser	His	Ala	Tyr	Thr	Cys	Gly
			245					250					255		
Cys	Gly	Trp	Val	Ser	Asn	His	Glu	Met	Val	Ile	His	Met	Thr	Phe	Ala
		260					265					270			
Cys	Asn	Ile	Trp	Gly	Ala	Lys	Cys	Met	Arg	Arg	Leu	Pro	Gln	Ala	Asp
	275						280					285			
Asp	Cys	Phe	Ile	His	Leu	Val	Asn	Asn	Tyr	His	Asn	Cys	Pro	Gly	Asn
	290				295					300					
Ser	Val	Gly	Met	Thr	Ile	Asn	Ser	Tyr	Ser	Lys	Ala	Leu	Val	Glu	Gly
305				310					315					320	
Asn	Tyr	Ala	Ala	Ala	Gly	Val	Asn	Lys	Pro	Leu	Asp	Gly	Ser	Gly	Ala
			325					330					335		
Asn	Arg	Asn	Val	Thr	Ala	Lys	Asp	Asn	Ser	Phe	Ala	Asn	Ser	Gln	Ala
		340					345						350		
Gly	Ser	Val	Val	Ser	Val	Pro	Tyr	Asp	Tyr	Thr	Lys	Ile	Ala	Ala	Ala
	355					360						365			
Asp	Val	Pro	Ala	Thr	Leu	Thr	Gly	Thr	Glu	Gly	Ala	Gly	Ala	Thr	Leu
	370				375					380					
Gly	Asn	Asp	Ala	Thr	Tyr	Ile	Leu	Ser	Thr	Ile	Pro	Thr	Val	Asp	Arg
385				390					395					400	
Gln	Glu	Gly	Glu	Ser	Ser	Leu	Tyr	Tyr	Phe	Ile	Asp	Gly	Leu	Val	Gly
			405					410					415		
Thr	Asn	Ser	Glu	Gly	Tyr	Ser	Ile	Ile	Glu	Phe	Asn	Asp	Gly	Ala	Thr
		420					425						430		
Leu	Leu	Leu	Asn	Asn	Lys	Glu	Lys	Ala	Trp	Ser	Asn	Gly	Ser	Ala	Ile
	435					440						445			
Gln	Leu	Gly	Asp	Asp	Asn	Tyr	Thr	Ser	Ile	Lys	Leu	Ser	Asn	Gly	Ala
	450				455					460					
Glu	Asn	Ile	Phe	Thr	Ala	Pro	Thr	Gly	Lys	Lys	Val	Ser	Gly	Ile	Thr

465 470 475 480
 Phe Tyr Ser Tyr Ile Asn Ile Lys Glu Glu Lys Leu Asp Phe Thr Lys
 485 490 495
 Tyr Pro Glu Tyr Gly Phe Arg Thr Cys Phe Trp Gln Lys Val Ala Asn
 500 505 510
 Leu Thr Tyr Ser Ala Thr Ser Asp Asp Val Gln Ile Leu Lys Ser Arg
 515 520 525
 Asp Pro Gln Asn Thr Asp Val Ala Ser Phe His Phe Thr Pro Thr Asn
 530 535 540
 Val Val Ser Phe Lys Asn Ser Gly Glu Gln Leu Cys Phe Leu Met Lys
 545 550 555 560
 Val Thr Tyr Ser Asp Glu Ser Thr Gly Ile Ser Ala Ile Gln Lys Lys
 565 570 575
 Met Pro Ile Asp Gly Val Thr Tyr Asn Leu Gln Gly Ile Arg Ile Asp
 580 585 590
 Asn Pro Thr Lys Gly Ile Tyr Ile Gln Asn Gly Lys Lys Ile Ile Ile
 595 600 605
 Lys

<210> 125
 <211> 1170
 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental.

<400> 125

atgagggtcta	aaatcatcag	cgccataaat	aattatagtg	ttattattct	cgatggctcg	60
aatggcgatt	tcactattag	tgctacaatg	agtttcagta	gcaaatcaaa	caaaaccata	120
gttggtgtaa	ataatgctcg	cctatgcacc	aagttctatc	taaccgatga	aataaagact	180
gcgctcgatg	ctgctaattg	aaaatcagca	agttcaacca	gtggagggtg	tacactctca	240
aatgggaaat	cagtgtcaga	acaacgtgaa	taccttactc	gtcaaacaat	tatcgatcta	300
actggcgatg	cttcggaatc	gtgtcagaaa	gcgggcatct	ttagcttcag	tagttgtacc	360
aatatcatca	tgcgaaacct	cgttttgggt	ggccctggcc	catgcatgtg	aggtggcaac	420
gatttgcttt	cgctcaactg	ttctaagcat	ttttgggtcg	atcactgtga	gttaaccgat	480
ggtatagatg	gcaattttcg	tattaccaag	agtagcgatt	tcaatactgt	tacttggtgt	540
atattcaatt	ataccgatcg	tgcatcgcac	cacatgaact	ccaatcttat	tggtagctcc	600
gatagcgaag	atgctgccta	tttgaacact	actatggcat	gcaatatttg	gggctacaag	660
tgcaatcagc	gaatgccaat	ggctcgtgct	ggtaatatcc	accttggtgaa	caacttttac	720
gattgcgctg	gcaatagtg	ggctgttaac	cctcgtaaaa	attctgagtt	cttagtcgag	780
aactgctact	ttgccacggg	tgtgaagcca	ttctcgcaga	gtggtgcgtt	gggatacaac	840
tttattgatt	gctatacaga	agattcatac	acttttcagc	agagtgggtac	agtgtctgtg	900
ccatacgttt	actctaagtt	tgatgtgcaa	ttagtaccgg	agcaactcaa	taaatatgct	960
ggcgcaacgc	ttacttctcc	gcttgtcata	ggtcgggaag	aggggtgtgt	tactcctatt	1020
agtgtgtgt	ctgttgatag	cgatgttggt	ttggtcgaat	actattcgct	gactggtaat	1080
cgtgttaaca	cgctcaatag	aggcatcaat	atcgtagtaa	ctatttacgc	caacggcaaa	1140
gtaaccacac	aaaagggttt	ggtgaaatag				1170

<210> 126
 <211> 389
 <212> PRT
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental.

<221> DOMINIO
 <222> (24)...(325)
 <223> Dominio catalítico

<400> 126

```

Met Arg Ser Lys Ile Ile Ser Ala Ile Asn Asn Tyr Ser Val Ile Ile
1      5      10      15
Leu Asp Gly Ser Asn Gly Asp Phe Thr Ile Ser Ala Thr Met Ser Phe
20      25      30
Ser Ser Lys Ser Asn Lys Thr Ile Val Gly Val Asn Asn Ala Arg Leu
35      40      45
Cys Thr Lys Phe Tyr Leu Thr Asp Glu Ile Lys Thr Ala Leu Asp Ala
50      55      60
Ala Asn Val Lys Ser Ala Ser Ser Thr Ser Gly Gly Thr Leu Ser
65      70      75      80
Asn Gly Lys Ser Val Ser Glu Gln Arg Glu Tyr Leu Thr Arg Gln Thr
85      90      95
Ile Ile Asp Leu Thr Gly Asp Ala Ser Glu Ser Cys Gln Lys Ala Gly
100     105     110
Ile Phe Ser Phe Ser Ser Cys Thr Asn Ile Ile Met Arg Asn Leu Val
115     120     125
Leu Val Gly Pro Gly Pro Cys Asp Val Gly Gly Asn Asp Leu Leu Ser
130     135     140
Leu Thr Gly Ser Lys His Phe Trp Val Asp His Cys Glu Leu Thr Asp
145     150     155     160
Gly Ile Asp Gly Asn Phe Asp Ile Thr Lys Ser Ser Asp Phe Asn Thr
165     170     175
Val Thr Trp Cys Ile Phe Asn Tyr Thr Asp Arg Ala Tyr Asp His Met
180     185     190
Asn Ser Asn Leu Ile Gly Ser Ser Asp Ser Glu Asp Ala Ala Tyr Leu
195     200     205
Asn Thr Thr Met Ala Cys Asn Ile Trp Gly Tyr Lys Cys Asn Gln Arg
210     215     220
Met Pro Met Ala Arg Ala Gly Asn Ile His Leu Val Asn Asn Phe Tyr
225     230     235     240
Asp Cys Ala Gly Asn Ser Val Ala Val Asn Pro Arg Lys Asn Ser Glu
245     250     255
Phe Leu Val Glu Asn Cys Tyr Phe Ala Thr Gly Val Lys Pro Phe Ser
260     265     270
Gln Ser Gly Ala Leu Gly Tyr Asn Phe Ile Asp Cys Tyr Thr Glu Asp
275     280     285
Ser Tyr Thr Phe Gln Gln Ser Gly Thr Val Ser Val Pro Tyr Val Tyr
290     295     300
Ser Lys Phe Asp Val Gln Leu Val Pro Glu Gln Leu Asn Lys Tyr Ala
305     310     315     320
Gly Ala Thr Leu Thr Ser Pro Leu Val Ile Gly Arg Glu Glu Gly Val
325     330     335
Val Thr Pro Ile Ser Ala Val Ser Val Asp Ser Asp Val Val Leu Val
340     345     350
Glu Tyr Tyr Ser Leu Thr Gly Asn Arg Val Asn Thr Leu Asn Arg Gly
355     360     365
Ile Asn Ile Val Arg Thr Ile Tyr Ala Asn Gly Lys Val Thr Thr Gln
370     375     380
Lys Val Leu Val Lys
385

```

<210> 127

<211> 1449

5 <212> ADN

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental.

10

<400> 127

```

atgcaatatg gcaaattagt acgcttgtcg gcaactgacaa cagcgctggc attcagcgcc 60
ctggcacagg caaataacct ggcaattaca ggccccggag ccggggctga tggttccagc 120
aaagccagtg gcagtagcta cggcgatgta aaagacggcg atctgcaaag ctactggcaa 180
ccgctgtcta ataacggcca aagagtgtcg gtttaagtga gcagcgctat cagcgttaat 240
caggtaatat tgcgtgaaca gggcagtaat gtaaccagct ggcggtggt aaataatgac 300
aacggcgtag tattggcaac cggcaccagc attggcagca acagaacggg taacttcagc 360
actgtaagca cgaaaaaact caatctggaa atactaactg ccagcggtgc cccgcgcatt 420
gctgagtttg aagtttattt aaataccaat ggcggaacc cgccaaatcc tactgacccg 480
gaaccaggcc cggtaacttc ttgcgcagcg tctccacagg gctatgcctc gcttaacggt 540
ggcactaccg gcggcagtg gacgaacgg gtcacggtaa cggtaagcac cggcgctcaa 600
atggtatcgg cgctacaaaa ccgcatcta aaccggccgc tcatatccg ggtaaatggc 660
actatcacac cgggtaattc tggcggtgtc agtaagtttg acattaaaga tatggataat 720
gtcagcatta ttggtgtagg caacaatgcg ttgtttgacg gtatcggtat taaaatctgg 780
cgggccaata acgttattat ccgcaacctt acaatgcgtt atgttaacac cggcgataaa 840
gacgtatta ccattgaagg ccggcgcggt aatatctgga ttgaccaca cgaatatctat 900
aacagcctga atgtgggtaa agatttttac gacgagctta taagcggtaa aaaagacgta 960
gataacgtaa ctatctctta caactacctg cagcagagct ggaaaacctc gctgtggggc 1020
agcagtgatt ccgacaacta caaccggcgt attacctttc accataacca ctggcataag 1080
gtaaattcac gcccgccact gttccgtttt ggccaggggc atattttaca taactattac 1140
aacgaatttc aggcacccgg tattaacagc cggatgggtg cggtaattcg tattgaaaac 1200
aatgtgtttg aaaacgcgaa aaaccggata gtgtcgtttt attccagcgg ctacggttac 1260
tgggacaccc gcggtaatat ctttagcaat attacctggc aggaataccc cagcgacggc 1320
attatcgccg ggccaaatgt acaaccaca gcggtgctaa acctgcccta cagctttaac 1380
ctgttaccca ccaaccaggt aaaagcccac gtactggcca acgcggcggt gaataaatgt 1440
agtttctaa 1449

```

<210> 128

<211> 482

5 <212> PRT

<213> Desconocido

<220>

<223> Obtenido de una muestra ambiental.

10

<221> SEÑAL

<222> (1)...(24)

<221> DOMINIO

15 <222> (5)...(482)

<223> Dominio catalítico

<400> 128

```

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

```

20

Val	Asn	Asn	Asp	Asn	Gly	Ala	Val	Leu	Ala	Thr	Gly	Thr	Ser	Ile	Gly
			100					105						110	
Ser	Asn	Arg	Thr	Val	Asn	Phe	Ser	Thr	Val	Ser	Thr	Lys	Lys	Leu	Asn
		115						120				125			
Leu	Glu	Ile	Leu	Thr	Ala	Ser	Gly	Ala	Pro	Arg	Ile	Ala	Glu	Phe	Glu
		130				135					140				
Val	Tyr	Leu	Asn	Thr	Asn	Gly	Gly	Asn	Pro	Pro	Asn	Pro	Thr	Asp	Pro
145					150					155					160
Glu	Pro	Gly	Pro	Val	Thr	Ser	Cys	Ala	Ala	Ser	Pro	Gln	Gly	Tyr	Ala
				165					170					175	
Ser	Leu	Asn	Gly	Gly	Thr	Thr	Gly	Gly	Ser	Gly	Ser	Asn	Ala	Val	Thr
			180					185					190		
Val	Thr	Val	Ser	Thr	Gly	Ala	Gln	Met	Val	Ser	Ala	Leu	Gln	Asn	Arg
		195					200					205			
Asp	Leu	Asn	Arg	Pro	Leu	Thr	Ile	Arg	Val	Asn	Gly	Thr	Ile	Thr	Pro
	210					215					220				
Gly	Asn	Ser	Gly	Gly	Val	Ser	Lys	Phe	Asp	Ile	Lys	Asp	Met	Asp	Asn
225					230					235					240
Val	Ser	Ile	Ile	Gly	Val	Gly	Asn	Asn	Ala	Leu	Phe	Asp	Gly	Ile	Gly
				245					250					255	
Ile	Lys	Ile	Trp	Arg	Ala	Asn	Asn	Val	Ile	Ile	Arg	Asn	Leu	Thr	Met
			260					265					270		
Arg	Tyr	Val	Asn	Thr	Gly	Asp	Lys	Asp	Ala	Ile	Thr	Ile	Glu	Gly	Pro
		275					280					285			
Ala	Arg	Asn	Ile	Trp	Ile	Asp	His	Asn	Glu	Ile	Tyr	Asn	Ser	Leu	Asn
		290				295					300				
Val	Gly	Lys	Asp	Phe	Tyr	Asp	Glu	Leu	Ile	Ser	Gly	Lys	Lys	Asp	Val
305					310					315					320
Asp	Asn	Val	Thr	Ile	Ser	Tyr	Asn	Tyr	Leu	His	Asp	Ser	Trp	Lys	Thr
				325					330					335	
Ser	Leu	Trp	Gly	Ser	Ser	Asp	Ser	Asp	Asn	Tyr	Asn	Arg	Arg	Ile	Thr
			340					345					350		
Phe	His	His	Asn	His	Trp	His	Lys	Val	Asn	Ser	Arg	Leu	Pro	Leu	Phe
		355					360					365			
Arg	Phe	Gly	Gln	Gly	His	Ile	Tyr	Asn	Asn	Tyr	Tyr	Asn	Asp	Ile	Gln
		370				375					380				
Asp	Thr	Gly	Ile	Asn	Ser	Arg	Met	Gly	Ala	Val	Ile	Arg	Ile	Glu	Asn
385					390					395					400
Asn	Val	Phe	Glu	Asn	Ala	Lys	Asn	Pro	Ile	Val	Ser	Phe	Tyr	Ser	Ser
				405					410					415	
Gly	Tyr	Gly	Tyr	Trp	Asp	Thr	Arg	Gly	Asn	Ser	Phe	Ser	Asn	Ile	Thr
			420					425					430		
Trp	Gln	Glu</													

<210> 129
<211> 1173
<212> ADN
<213> *Bacillus halodurans* ATCC 27557

<220>
<400> 129

```

atgagttcga aaatcaaaaa tgctatcaat aactatagtg ttattattct cgatggctcg      60
aatggcgatt ttacagtcaa tgctacaatg agttttcagtg gcaagtccaa taaaactatt    120
gtgggtgtga acaatgctcg cctatgcacc aaattctaca ttacgcccga gataaaagaa    180
gccctcgatg ctgccgatgt gaaatctaag agctcaagta gtggcactgg tggaaactctt    240
tctaattgta cgtcgggtcag tgaggctcgc gaattggcta ctcgtcaaac gttgattgat    300
tatctcggcg atagctcaga atcgtatcag aaagctggta tctttggctt tagcaactgc    360
actaatatta ttatgcgcaa cattgttttc gttggccctg gtccatgcga tgtagggtggc    420
aacgacttgc tttcgcctcg tggttcgaag catttctggg tcgaccactg cgagtttacc    480
gatggcatcg atggcaactt cgacatcacc aagagtagcg acttcaacac cgtttcgtgg    540
tgcactttca gctataccga ccgcgcatac gaccacatga attccaacct tattggtagc    600
tccgattcag agaattgcggc ttaccttaat actactatgg cttccaacgt ctggggcaat    660
aagtgcatac agcgtatgcc tatggctcgt gccggtaata ttcacctcgt aaataattat    720
tacaactgcc ctggcaatag cgtggctgtg aatcctcgca aaaactcaga atttttggtg    780
gagaattgct atttcgcaag tggcgtaaag cctttctcgc agagcggcgc tcttagctat    840
ctatttatcg attgctacac cgaagatact tacaccttcc agaaatctgg ctctactacg    900
gtgccatata catatagcaa attcgatgct cagcttggtc ccgagcaact caccacaattc    960
gctggcgcaa cattgacttc gccgcttggt attggtaggg aatctgagaa tgttacacca   1020
gtctcagtcg ttgctgcaaa tagcgatgtc atatctgtag aatactattc gctcactggc   1080
aagcgcatac gcgaaccaac taaaggcatc aatatcgta gaactattta tactaacggc   1140
aacgtgacca caaaaaggt cttggtgaaa taa                                1173

```

<210> 130
 <211> 390
 5 <212> PRT
 <213> Bacillus halodurans ATCC 27557

<220>
 <222> DOMINIO
 10 <222> (38)...(326)
 <223> Dominio catalítico

<400> 130

```

Met Ser Ser Lys Ile Lys Asn Ala Ile Asn Asn Tyr Ser Val Ile Ile
 1                    5                    10                    15
Leu Asp Gly Ser Asn Gly Asp Phe Thr Val Asn Ala Thr Met Ser Phe
                20                25                30
Ser Gly Lys Ser Asn Lys Thr Ile Val Gly Val Asn Asn Ala Arg Leu
                35                40                45
Cys Thr Lys Phe Tyr Ile Thr Pro Glu Ile Lys Glu Ala Leu Asp Ala
 50                55                60
Ala Asp Val Lys Ser Lys Ser Ser Ser Gly Thr Gly Gly Thr Leu
65                70                75                80
Ser Asn Gly Thr Ser Val Ser Glu Ala Arg Glu Leu Ala Thr Arg Gln
                85                90                95
Thr Leu Ile Asp Tyr Leu Gly Asp Ser Ser Glu Ser Tyr Gln Lys Ala
                100                105                110
Gly Ile Phe Gly Phe Ser Asn Cys Thr Asn Ile Ile Met Arg Asn Ile
                115                120                125
Val Phe Val Gly Pro Gly Pro Cys Asp Val Gly Gly Asn Asp Leu Leu
130                135                140
Ser Leu Val Gly Ser Lys His Phe Trp Val Asp His Cys Glu Phe Thr
145                150                155                160
Asp Gly Ile Asp Gly Asn Phe Asp Ile Thr Lys Ser Ser Asp Phe Asn
                165                170                175
Thr Val Ser Trp Cys Thr Phe Ser Tyr Thr Asp Arg Ala Tyr Asp His
                180                185                190
Met Asn Ser Asn Leu Ile Gly Ser Ser Asp Ser Glu Asn Ala Ala Tyr
195                200                205

```

15

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

<210> 131
 <211> 972
 5 <212> ADN
 <213> Desconocido

<220>
 <223> Obtenido de una muestra ambiental

10 <400> 131

atggcaaaaa	tactgacatt	agacggtaac	ccggccgcca	gctgggttcaa	caaattccagg	60
accaagtgga	atagcagccg	cgccgatatt	gtgttgtctt	accagcaatc	caacggcggt	120
tggccaaaaa	acctggatta	caactcagtg	agcgcaggca	atggcgggag	cgacagcggc	180
accatcgaca	atggtgcaac	cattaccgaa	atggtttacc	tcgctgaaat	ttataaaaaac	240
ggcggcaaca	ccaaatatcg	cgatgcagtg	cgcagagcag	caaacttttt	agtgagctcg	300
caatacagca	caggcgcctt	gccacaattt	tatccgttga	aaggcggcta	tgccgatcat	360
gcgaccttta	acgataacgg	catggcgtac	gcgttgacgg	tattggattt	cgcagtaaac	420
aaacgcgcac	cgtttgataa	cgacattttc	tctgattctg	atcgggcgaa	attcaaaacc	480
gctgttgcca	aaggtgtgga	ttacatttta	aaagcgcagt	ggaaacaaaa	tggaaaactc	540
actgcatggt	gtgcacaaca	cgtgctacg	gattaccaac	cgaaaaaagc	gcgcgcttat	600
gaattggaat	cattgagtgg	tagcgagtcg	gtcggcattc	tcgccttctt	gatgacccaa	660
ccacaaaccg	cgcaaatcga	agcggcggtc	aaggcgggtg	tcaactgggt	cgccagtcca	720
aatacttatt	tggctaacta	cacttacgat	tcatcaaaaag	cgtctaccaa	cccgaattgtg	780
tataaatccg	gaagcagaat	gtggtatcgc	ttctatgacc	tgaacaccaa	ccgtgggttc	840
tttagtgatc	gcgatggcag	caaattctat	gatatcacc	aaatgtcaga	agagcgtcgc	900
accggttata	gctgggggtg	ctcttacggg	gaatctatta	tttccttcgc	gcaaaaagtg	960
ggttatctgt	ag					972

<210> 132
 15 <211> 323
 <212> PRT
 <213> Desconocido

<220>
 20 <223> Obtenido de una muestra ambiental

<400> 132

Met	Ala	Lys	Ile	Leu	Thr	Leu	Asp	Gly	Asn	Pro	Ala	Ala	Ser	Trp	Phe
1				5					10					15	
Asn	Lys	Ser	Arg	Thr	Lys	Trp	Asn	Ser	Ser	Arg	Ala	Asp	Ile	Val	Leu
			20				25						30		
Ser	Tyr	Gln	Gln	Ser	Asn	Gly	Gly	Trp	Pro	Lys	Asn	Leu	Asp	Tyr	Asn
		35				40						45			
Ser	Val	Ser	Ala	Gly	Asn	Gly	Gly	Ser	Asp	Ser	Gly	Thr	Ile	Asp	Asn
	50				55						60				
Gly	Ala	Thr	Ile	Thr	Glu	Met	Val	Tyr	Leu	Ala	Glu	Ile	Tyr	Lys	Asn
65					70					75					80
Gly	Gly	Asn	Thr	Lys	Tyr	Arg	Asp	Ala	Val	Arg	Arg	Ala	Ala	Asn	Phe
			85						90					95	
Leu	Val	Ser	Ser	Gln	Tyr	Ser	Thr	Gly	Ala	Leu	Pro	Gln	Phe	Tyr	Pro
			100					105					110		
Leu	Lys	Gly	Gly	Tyr	Ala	Asp	His	Ala	Thr	Phe	Asn	Asp	Asn	Gly	Met
	115					120						125			
Ala	Tyr	Ala	Leu	Thr	Val	Leu	Asp	Phe	Ala	Val	Asn	Lys	Arg	Ala	Pro
	130					135					140				
Phe	Asp	Asn	Asp	Ile	Phe	Ser	Asp	Ser	Asp	Arg	Ala	Lys	Phe	Lys	Thr
145					150					155					160
Ala	Val	Ala	Lys	Gly	Val	Asp	Tyr	Ile	Leu	Lys	Ala	Gln	Trp	Lys	Gln
			165						170					175	
Asn	Gly	Lys	Leu	Thr	Ala	Trp	Cys	Ala	Gln	His	Gly	Ala	Thr	Asp	Tyr
			180					185					190		
Gln	Pro	Lys	Lys	Ala	Arg	Ala	Tyr	Glu	Leu	Glu	Ser	Leu	Ser	Gly	Ser
		195					200					205			
Glu	Ser	Val	Gly	Ile	Leu	Ala	Phe	Leu	Met	Thr	Gln	Pro	Gln	Thr	Ala
	210					215					220				
Gln	Ile	Glu	Ala	Ala	Val	Lys	Ala	Gly	Val	Asn	Trp	Phe	Ala	Ser	Pro
225					230					235					240
Asn	Thr	Tyr	Leu	Ala	Asn	Tyr	Thr	Tyr	Asp	Ser	Ser	Lys	Ala	Ser	Thr
			245						250					255	
Asn	Pro	Ile	Val	Tyr	Lys	Ser	Gly	Ser	Arg	Met	Trp	Tyr	Arg	Phe	Tyr
			260					265					270		
Asp	Leu	Asn	Thr	Asn	Arg	Gly	Phe	Phe	Ser	Asp	Arg	Asp	Gly	Ser	Lys
	275					280					285				
Phe	Tyr	Asp	Ile	Thr	Gln	Met	Ser	Glu	Glu	Arg	Arg	Thr	Gly	Tyr	Ser
	290				295						300				
Trp	Gly	Gly	Ser	Tyr	Gly	Glu	Ser	Ile	Ile	Ser	Phe	Ala	Gln	Lys	Val
305					310					315					320
Gly	Tyr	Leu													

<210> 133

<211> 972

5 <212> ADN

<213> Artificial Sequence

<220>

<223> Polinucleótido generado sintéticamente

10

<400> 133

atggcaaaaa	tactgacatt	agacggtaac	ccggccgcca	gctggttcaa	caaattccagg	60
accaagtgga	atagcagccg	cgccgatatt	gtgttgtctt	accagcaatc	caacggcggt	120
tggccaaaaa	acctggatta	caactcagtg	agcgcaggca	atggcgggag	cgacagcggc	180
accatcgaca	atggtgcaac	cattaccgaa	atggtttacc	tcgctgaaat	ttataaaaaa	240
ggcggcaaca	ccaaatatcg	cgatgcagtg	cgcagagcag	caaacttttt	agtgagctcg	300


```

caatacagca caggcgcctt gccacaattt tatccgttga aaggcggcta tcatgatcat 360
gcgaccttta acgataacgg catggcgtac gcgttgacgg tattggattt cgcagtaaac 420
aaacgcgcac cgtttgataa cgacattttc tctgattctg atcgggcgaa attcaaaacc 480
gctgttgcca aagggtgtgga ttacatttta aaagcgcagt ggaaacaaaa tggaaaactc 540
actgcatggt gtgcacaaca cgggtgctttg gattaccaac cgaaaaaagg tgcgcgttat 600
gaattggaat cattgagtgg taaggagtgc gtcggcattc tgccttcttt gatgacccaa 660
ccacaaaccg cgcaaatcga agcggcggtc aaggcgggtg tcaactgggt cgccagtcca 720
aatacttatt tggctaacta cacttacgat tcatcaaaag cgtctaccaa cccgattgtg 780
tataaaaagg gaagcagaat gtggtatcgc ttctatgacc tgtataccaa ccgtgggtttc 840
tttagtgatc gcgatggcag caaattctat gatatcacc aaatgtcaga agagcgtcgc 900
accggttata gctgggggtg ctcttggggg gaagttatta tttccttcgc gcaaaaagtg 960
ggttatctgt ag 972

```

<210> 134

<211> 323

<212> PRT

5 <213> Artificial Sequence

<220>

<223> Polinucleótido generado sintéticamente

10 <400> 134

```

Met Ala Lys Ile Leu Thr Leu Asp Gly Asn Pro Ala Ala Ser Trp Phe
 1           5           10           15
Asn Lys Ser Arg Thr Lys Trp Asn Ser Ser Arg Ala Asp Ile Val Leu
          20           25           30
Ser Tyr Gln Gln Ser Asn Gly Gly Trp Pro Lys Asn Leu Asp Tyr Asn
          35           40           45
Ser Val Ser Ala Gly Asn Gly Gly Ser Asp Ser Gly Thr Ile Asp Asn
 50           55           60
Gly Ala Thr Ile Thr Glu Met Val Tyr Leu Ala Glu Ile Tyr Lys Asn
65           70           75           80
Gly Gly Asn Thr Lys Tyr Arg Asp Ala Val Arg Arg Ala Ala Asn Phe
          85           90           95
Leu Val Ser Ser Gln Tyr Ser Thr Gly Ala Leu Pro Gln Phe Tyr Pro
          100          105          110
Leu Lys Gly Gly Tyr His Asp His Ala Thr Phe Asn Asp Asn Gly Met
          115          120          125
Ala Tyr Ala Leu Thr Val Leu Asp Phe Ala Val Asn Lys Arg Ala Pro
130          135          140
Phe Asp Asn Asp Ile Phe Ser Asp Ser Asp Arg Ala Lys Phe Lys Thr
145          150          155          160
Ala Val Ala Lys Gly Val Asp Tyr Ile Leu Lys Ala Gln Trp Lys Gln
          165          170          175
Asn Gly Lys Leu Thr Ala Trp Cys Ala Gln His Gly Ala Leu Asp Tyr
          180          185          190
Gln Pro Lys Lys Gly Arg Ala Tyr Glu Leu Glu Ser Leu Ser Gly Lys
          195          200          205
Glu Ser Val Gly Ile Leu Ala Phe Leu Met Thr Gln Pro Gln Thr Ala
210          215          220
Gln Ile Glu Ala Ala Val Lys Ala Gly Val Asn Trp Phe Ala Ser Pro
225          230          235          240
Asn Thr Tyr Leu Ala Asn Tyr Thr Tyr Asp Ser Ser Lys Ala Ser Thr
          245          250          255
Asn Pro Ile Val Tyr Lys Lys Gly Ser Arg Met Trp Tyr Arg Phe Tyr
          260          265          270
Asp Leu Tyr Thr Asn Arg Gly Phe Phe Ser Asp Arg Asp Gly Ser Lys
          275          280          285
Phe Tyr Asp Ile Thr Gln Met Ser Glu Glu Arg Arg Thr Gly Tyr Ser

```

ES 2 545 639 T3

	290		295		300										
Trp	Gly	Gly	Ser	Trp	Gly	Glu	Val	Ile	Ile	Ser	Phe	Ala	Gln	Lys	Val
305					310					315					320
Gly	Tyr	Leu													

REIVINDICACIONES

1. Un ácido nucleico aislado, sintético, o recombinante que comprende

- 5 (a) una secuencia de ácidos nucleicos que tiene una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o superior o que es idéntica en un 100 % a la SEC ID N°: 77, sobre una región de al menos 950, 1000, 1050, 1100, 1150 o más restos, o la longitud completa de un gen o de un transcrito, en la que el ácido nucleico codifica al menos un polipéptido que tiene una actividad de pectato liasa;
- 10 (b) una secuencia de ácidos nucleicos que comprende una secuencia como se establece en la SEC ID N°: 77;
- (c) una secuencia de ácidos nucleicos que codifica un polipéptido que tiene una secuencia como se establece en las SEC ID N°: 78, SEC ID N°: 132 o SEC ID N°: 134;
- (d) una secuencia que se hibrida en condiciones rigurosas con un ácido nucleico que comprende la SEC ID N°: 77, SEC ID N°: 131 o SEC ID N°: 133, en la que el ácido nucleico codifica un polipéptido que tiene una actividad de pectato liasa,
- 15 en la que las condiciones rigurosas incluyen una etapa de lavado que comprende un lavado en 0,2X SSC a una temperatura de aproximadamente 65 °C durante aproximadamente 15 minutos;
- (e) el ácido nucleico de (a), (b), (c) o (d) que codifica un pectato liasa pero que carece de una secuencia señal, un dominio prepro, o una combinación de los mismos;
- (f) el ácido nucleico de (a), (b), (c), (d) o (e) que codifica un pectato liasa pero que también tiene una secuencia heteróloga; o
- 20 (g) una secuencia complementaria con (a), (b), (c), (d), (e), o (f).

2. Una sonda de ácido nucleico para identificar un ácido nucleico que codifica un polipéptido con una actividad de pectato liasa, en el que la sonda comprende de aproximadamente 60 a 100, o de aproximadamente 60 a 150 bases consecutivas de una secuencia que comprende la SEC ID N°: 77, SEC ID N°: 131 o SEC ID N°: 133, en el que la sonda identifica al ácido nucleico mediante unión o hibridación en condiciones rigurosas,

25 en la que las condiciones rigurosas incluyen una etapa de lavado que comprende un lavado en 0,1x SSC, SDS al 0,5 % a una temperatura de 68 °C durante aproximadamente 15 minutos.

30 3. Un ácido nucleico aislado, sintético o recombinante

(A) que tiene una secuencia que comprende una modificación de la secuencia de la SEC ID N°: 131, en la que la modificación de la SEC ID N°: 131 comprende uno o más de los siguientes cambios:

- 35 los nucleótidos en los restos 352 a 354 son CAT o CAC,
los nucleótidos en los restos 544 a 546 son GTG, GTT, GTC, o GTA,
los nucleótidos en los restos 568 a 570 son TTG, TTA, CTT, CTC, CTA, o CTG
los nucleótidos en los restos 589 a 591 son GGT, GGC, GGA, o GGG,
los nucleótidos en los restos 622 a 624 son AAG o AAA,
- 40 los nucleótidos en los restos 655 a 657 son ATG,
los nucleótidos en los restos 667 a 669 son GAG o GAA,
los nucleótidos en los restos 763 a 765 son CGG, CGT, CGC, CGA, AGA, AGG,
los nucleótidos en los restos 787 a 789 son AAG o AAA,
los nucleótidos en los restos 823 a 825 son TAT o TAC,
- 45 los nucleótidos en los restos 925 a 927 son TGG, o
los nucleótidos en los restos 934 a 936 son GTT, GTG, GTC, o GTA; o

(B) que tiene la secuencia de la reivindicación 1(a), que comprende adicionalmente una modificación de secuencia, en el que la modificación de la secuencia comprende uno o más de los siguientes cambios:

- 50 los nucleótidos en la posición equivalente de los restos 352 a 354 de la SEC ID N°: 131 se cambian a CAT o CAC,
los nucleótidos en la posición equivalente de los restos 544 a 546 de la SEC ID N°: 131 se cambian a GTG, GTT, GTC, o GTA,
- 55 los nucleótidos en la posición equivalente de los restos 568 a 570 de la SEC ID N°: 131 se cambian a TTG, TTA, CTT, CTC, CTA, o CTG
los nucleótidos en la posición equivalente de los restos 589 a 591 de la SEC ID N°: 131 se cambian a GGT, GGC, GGA, o GGG,
los nucleótidos en la posición equivalente de los restos 622 a 624 de la SEC ID N°: 131 se cambian a AAG o AAA,
- 60 los nucleótidos en la posición equivalente de los restos 655 a 657 de la SEC ID N°: 131 se cambian a ATG,
los nucleótidos en la posición equivalente de los restos 667 a 669 de la SEC ID N°: 131 son GAG o GAA,
los nucleótidos en la posición equivalente de los restos 763 a 765 de la SEC ID N°: 131 se cambian a CGG, CGT, CGC, CGA, AGA, AGG,
- 65 los nucleótidos en la posición equivalente de los restos 787 a 789 de la SEC ID N°: 131 se cambian a AAG o AAA,

los nucleótidos en la posición equivalente de los restos 823 a 825 de la SEC ID N°: 131 se cambian a TAT o TAC,
los nucleótidos en la posición equivalente de los restos 925 a 927 de la SEC ID N°: 131 se cambian a TGG, o los nucleótidos en la posición equivalente de los restos 934 a 936 de la SEC ID N°: 131 se cambian a GTT, GTG, GTC, o GTA.

4. Un método para generar una variante de un ácido nucleico que codifica un polipéptido con una actividad de pectato liasa que comprende

(A)

(a) proporcionar un ácido nucleico molde que comprende una secuencia como se establece en la SEC ID N°: 77; y
(b) modificar, suprimir o añadir, o una combinación de los mismos, uno o más nucleótidos en la secuencia molde para generar una variante del ácido nucleico molde,

en el que la variante tiene una identidad de al menos un 90 % con la secuencia de ácidos nucleicos de la reivindicación 1; o

(B)

(a) proporcionar un ácido nucleico que codifica un polipéptido con una actividad de pectato liasa que comprende una secuencia como se establece en la reivindicación 1; y,
(b) identificar un codón en el ácido nucleico de la etapa (a) y sustituirlo con un codón diferente que codifica el mismo aminoácido que el del codón sustituido, modificando de este modo los codones en un ácido nucleico que codifica una pectato liasa.

5. Un casete de expresión, un vector o un vehículo de clonación que comprende un ácido nucleico que comprende una secuencia como se establece en la reivindicación 1 o 3.

6. Una célula hospedadora que comprende un casete de expresión, un vector o un vehículo de clonación como se establece en la reivindicación 5.

7. Un animal no humano transgénico o planta transgénica o semilla transgénica que comprende una secuencia como se establece en la reivindicación 1 o 3, o que comprende un casete de expresión, un vector o un vehículo de clonación como se establece en la reivindicación 5.

8. Un método para producir un polipéptido recombinante que comprende las etapas de: (a) proporcionar un ácido nucleico, en el que el ácido nucleico comprende una secuencia como se establece en la reivindicación 1 o 3; y (b) expresar el ácido nucleico de la etapa (a) en condiciones que permiten la expresión del polipéptido, produciendo de este modo un polipéptido recombinante, en el que opcionalmente el método es para sobreexpresión del polipéptido mediante el uso de un promotor de actividad elevada, un vector dicistrónico o mediante la amplificación genética del vector.

9. Un polipéptido aislado, sintético o recombinante

(a) que tiene una identidad de secuencia de al menos un 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 %, o superior o que es idéntico en un 100 % a la SEC ID N°: 78, sobre una región de al menos 300, 350, 400, 450, 500, 550, 600, 650 o más restos, o la longitud completa y en el que el polipéptido tiene actividad de pectato liasa,
(b) codificado por un ácido nucleico que tiene una secuencia como se establece en la reivindicación 1 o 3,
(c) el polipéptido de (a) o (b), pero que carece de una secuencia señal, un dominio prepro, o una combinación de los mismos; o
(d) el polipéptido de (a), (b) o (c), pero que también tiene una secuencia heteróloga; en el que el polipéptido has actividad de pectato liasa; o
(e) el polipéptido de (a), (b), (c), o (d), en el que el polipéptido comprende al menos un sitio de glicosilación o está glicosilado.

10. Un polipéptido aislado, sintético o recombinante

(A) que tiene una secuencia que comprende una modificación de la secuencia de la SEC ID N°: 132, en el que la modificación de la SEC ID N°: 132 comprende una o más de las siguientes mutaciones: la alanina en la posición del aminoácido 118 es histidina, la alanina en la posición del aminoácido 182 es valina, la treonina en la posición del aminoácido 190 es leucina, la alanina en la posición del aminoácido 197 es glicina, la serina en la posición del aminoácido 208 es lisina, la treonina en la posición del aminoácido 219 es metionina, la treonina en la posición del aminoácido 223 es ácido glutámico, la serina en la posición del aminoácido 255 es arginina, la serina en la posición del aminoácido 263 es lisina, la asparagina en la posición del aminoácido 275 es tirosina, la

tirosina en la posición del aminoácido 309 es triptófano, o, la serina en la posición del aminoácido 312 es valina;
o

(B) que tiene la secuencia de la reivindicación 9 y que comprende una modificación de la secuencia, en el que la modificación de la secuencia comprende uno o más de los siguientes cambios:

el aminoácido en la posición equivalente de la alanina en el resto 118 de la SEC ID N°: 132 se cambia por una histidina,
el aminoácido en la posición equivalente de la alanina en el resto 182 de la SEC ID N°: 132 se cambia por una valina,
el aminoácido en la posición equivalente de la treonina en el resto 190 de la SEC ID N°: 132 se cambia por una leucina,
el aminoácido en la posición equivalente de la alanina en el resto 197 de la SEC ID N°: 132 se cambia por una glicina,
el aminoácido en la posición equivalente de la serina en el resto 208 de la SEC ID N°: 132 se cambia por una lisina,
el aminoácido en la posición equivalente de la treonina en el resto 219 de la SEC ID N°: 132 se cambia por una metionina,
el aminoácido en la posición equivalente de la treonina en el resto 223 de la SEC ID N°: 132 se cambia por un ácido glutámico,
el aminoácido en la posición equivalente de la serina en el resto 255 de la SEC ID N°: 132 se cambia por una arginina,
el aminoácido en la posición equivalente de la serina en el resto 263 de la SEC ID N°: 132 se cambia por una lisina,
el aminoácido en la posición equivalente de la asparagina en el resto 275 de la SEC ID N°: 132 se cambia por una tirosina,
el aminoácido en la posición equivalente de la tirosina en el resto 309 de la SEC ID N°: 132 se cambia por un triptófano, o,
el aminoácido en la posición equivalente de la serina en el resto 312 de la SEC ID N°: 132 se cambia por una valina.

11. Un heterodímero o un homodímero que comprende un polipéptido como se establece en cualquiera de las reivindicaciones 9 o 10.

12. Un polipéptido inmovilizado, en el que el polipéptido comprende una secuencia como se establece en cualquiera de las reivindicaciones 9 o 10.

13. Un anticuerpo aislado, sintético o recombinante que se une de forma específica a un polipéptido como se establece en la reivindicación 9 o 10.

14. Un método para hidrolizar, licuar, retirar, descomponer o alterar una composición que comprende pectina, pectato (ácido poligalacturónico), homogalacturonano, ramnogalacturonano o pared de célula vegetal que comprende las siguientes etapas:

(a) proporcionar un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;
(b) proporcionar una composición que comprende una pectina, un pectato, homogalacturonano, ramnogalacturonano o pared de célula vegetal; y
(c) poner en contacto el polipéptido de la etapa (a) con la composición de la etapa (b) en condiciones en las que el polipéptido se hidroliza, licúa, retira, descomponen o altera dicha composición.

15. Una composición que comprende un polipéptido como se establece en cualquiera de las reivindicaciones 9 o 10, o un ácido nucleico como se establece en cualquiera de las reivindicaciones 1 o 3.

16. Un método para lavar un objeto que comprende las siguientes etapas:

(a) proporcionar una composición que comprende un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;
(b) proporcionar un objeto; y
(c) poner en contacto el polipéptido de la etapa (a) y el objeto de la etapa (b) en condiciones en las que la composición puede lavar el objeto.

17. Un método para limpieza de material de fibra, hilo, textil, tejido o celulósico que comprende las siguientes etapas:

(a) proporcionar un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;
(b) proporcionar una fibra, un hilo, un textil, un tejido o un material celulósico; y

(c) poner en contacto el polipéptido de la etapa (a) y el material de fibra, hilo, textil, tejido o celulósico de la etapa (b) en condiciones en las que la pectato liasa puede limpiar el material de fibra, hilo, textil, tejido o celulósico.

18. Un método para mejorar la extracción de aceite de un material vegetal rico en aceite que comprende las siguientes etapas:

(a) proporcionar un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;

(b) proporcionar un material vegetal rico en aceite; y

(c) poner en contacto el polipéptido de la etapa (a) y el material vegetal rico en aceite.

19. Un método para preparar un zumo, jarabe, puré o extracto de fruta o vegetal que comprende las siguientes etapas:

(a) proporcionar un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;

(b) proporcionar una composición o un líquido que comprende un material de fruta o vegetal; y

(c) poner en contacto el polipéptido de la etapa (a) y la composición, preparando de este modo el zumo, jarabe, puré o extracto de fruta o vegetal.

20. Un método para tratar un papel, un producto de papel, o una pasta de madera que comprende las siguientes etapas:

(a) proporcionar un polipéptido que tiene una actividad de pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;

(b) proporcionar una composición que comprende un papel, un producto de papel, o una pasta de madera; y

(c) poner en contacto el polipéptido de la etapa (a) y la composición de la etapa (b) en condiciones en las que la pectato liasa puede tratar el papel, producto de papel, pasta de papel, o pasta de madera.

21. Un proceso de biolimpieza que comprende las siguientes etapas:

(a) proporcionar una pectato liasa como se establece en cualquiera de las reivindicaciones 9 o 10;

(b) proporcionar un material que comprende pectina, ácido poligalacturónico (pectato), homogalacturonano, ramnogalacturonano o pared de célula vegetal;

(c) poner en contacto la pectato liasa de (a) con el material de (b) en condiciones alcalinas en un tampón de bicarbonato o equivalente.

22. El ácido nucleico aislado, sintético o recombinante de la reivindicación 1, en la que el ácido nucleico comprende la SEC ID N°: 133.

23. El polipéptido aislado, sintético o recombinante de la reivindicación 9, en la que el polipéptido comprende la SEC ID N°: 134.

24. El polipéptido aislado, sintético o recombinante de la reivindicación 9, en la que el polipéptido comprende la SEC ID N°: 132.

25. Una formulación para tratar un material, comprendiendo dicha formulación al menos un polipéptido de la reivindicación 9 o 10 a una dosificación seleccionada entre:

i. entre 1 y 100 g por ton de material a tratar;

ii. entre 1 µg y 100 µg por g de material a tratar;

iii. entre 0,5 mg y 50 mg por 0,45 kg de material a tratar

iv. una dosificación que comprende una potencia enzimática entre 100 y 40.000 unidades/ml.

FIGURA 1

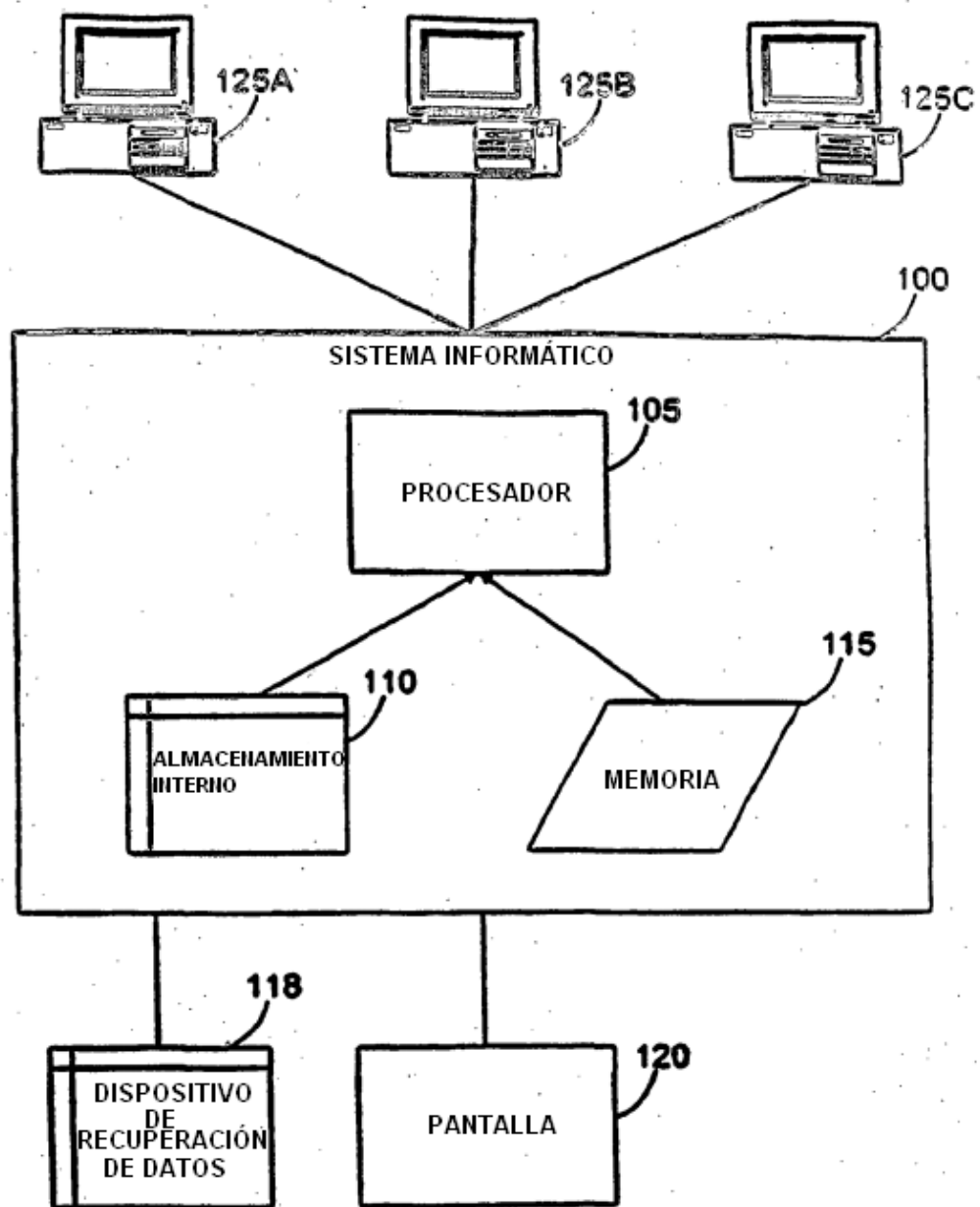


FIGURA 2

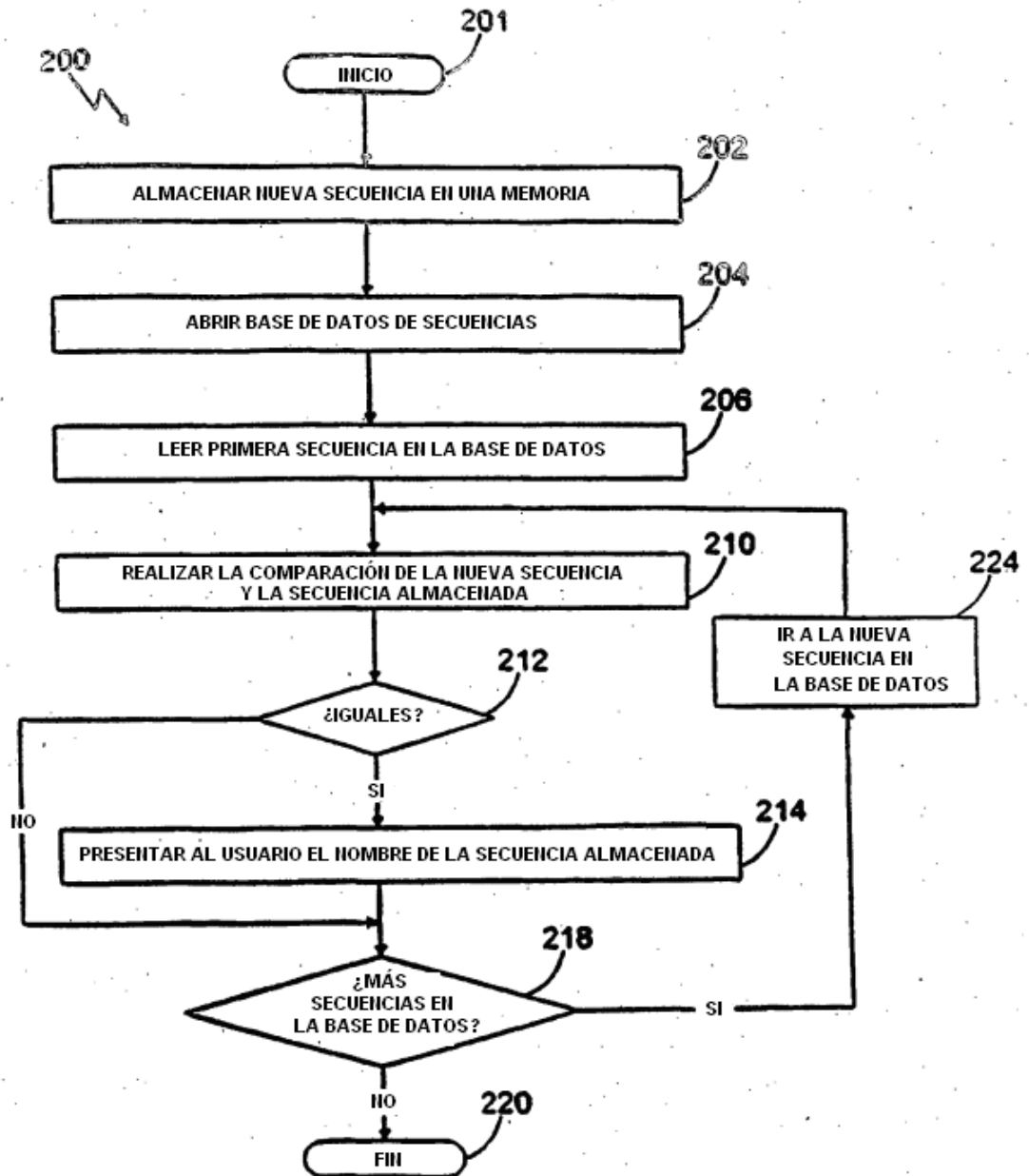


FIGURA 3

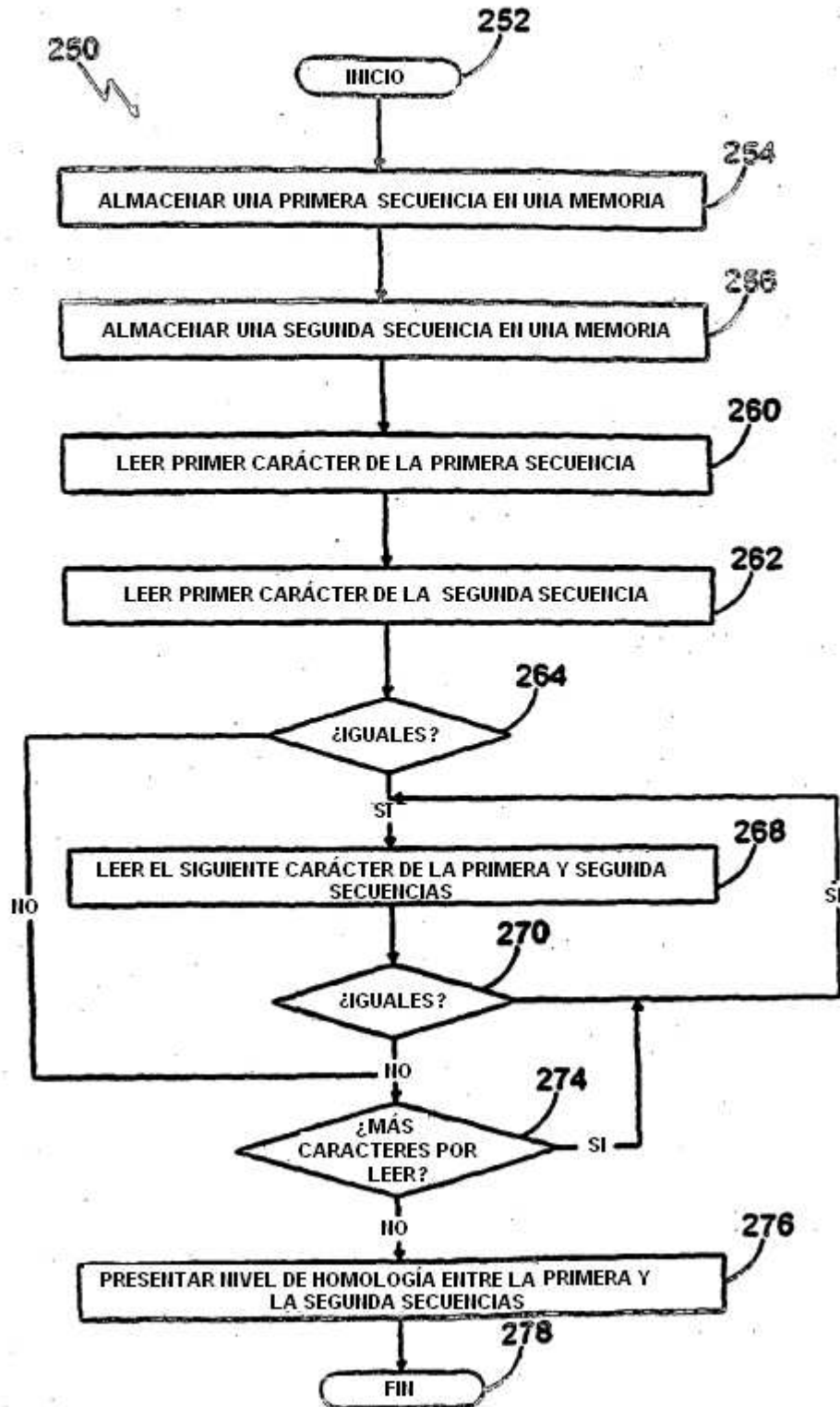


FIGURA 4

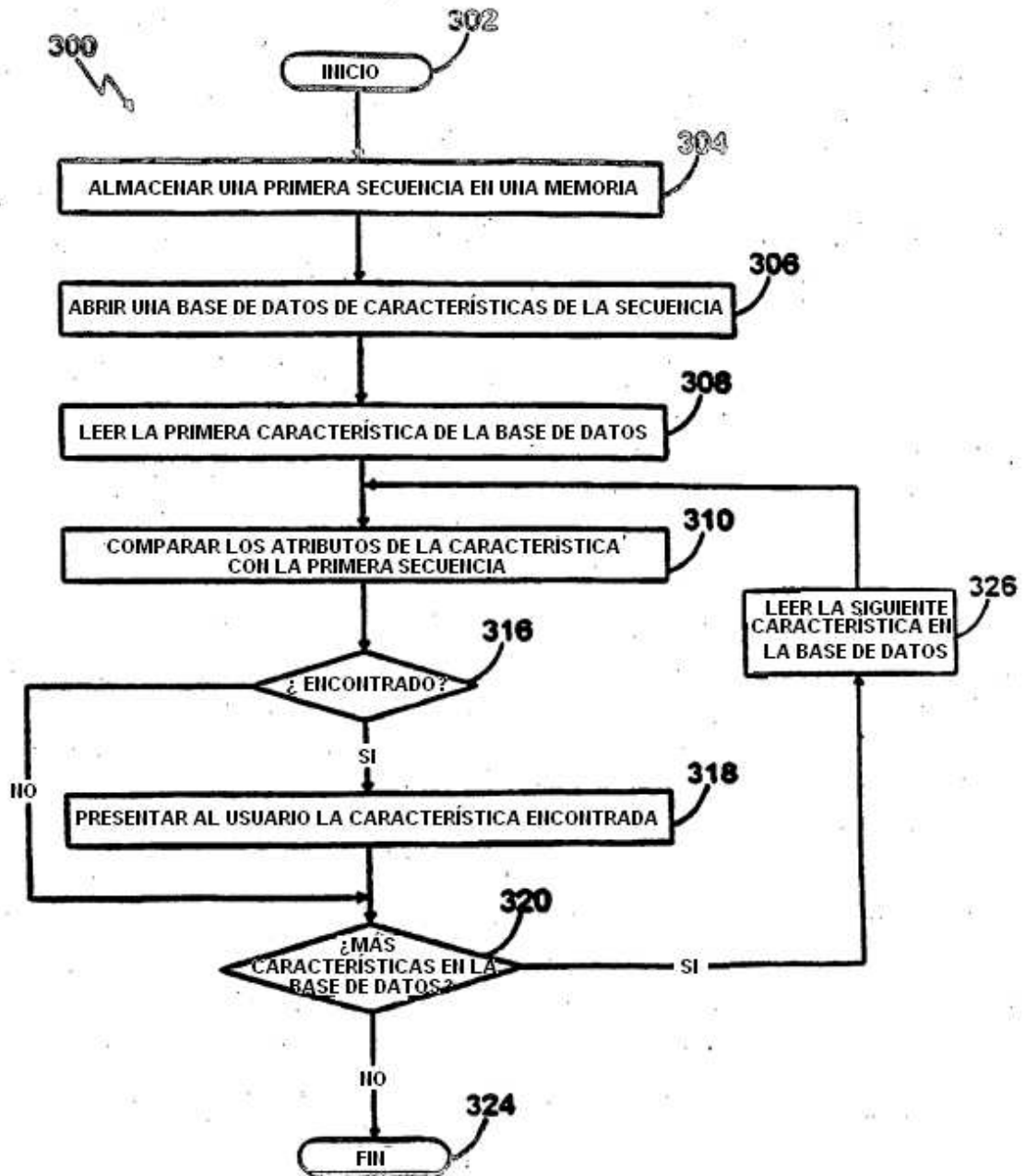


FIGURA 5A

SEC ID Nº: en provisional	Especificidad Relativa Sustrato	Valor Especificidad Relativa Sustrato	Temp Actividad Caracterización	pH Actividad Caracterización	Enzima	Descripción Caracterización	Sustrato Caracterización
101, 102					Pectinasa		
1, 2			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
103, 104					Pectato liasa		
105, 106	Algodón	4,8	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
107, 108	Algodón	1,5	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
107, 108			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
109, 110					Pectato liasa		
11, 12					Pectato liasa		
111, 112					Pectinasa		
113, 114					Pectato liasa		
115, 116			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
117, 118	Algodón	3	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
119, 120			temp ambiente	9	Pectato liasa		
121, 122	Algodón	7,8	50	9	Pectato liasa	Application Biolumpieza	Algodón
121, 122	Algodón	10,3	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
123, 124	Algodón	3,6	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
125, 126	Algodón	2,5	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
127, 128			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
127, 128			temp ambiente	11	Pectato liasa	ensayo de pga	Acido poligalacturónico
129, 130	Algodón	1,8	40	9	Pectato liasa	Aplicación de Biolumpieza	Algodón
13, 14			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
15, 16					Pectato liasa		
17, 18				9	Pectato liasa	ensayo de pga	Acido poligalacturónico
19, 20					Pectato liasa		
21, 22			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
23, 24					Pectato liasa		
25, 26			temp ambiente	10	Pectato liasa	ensayo de pga	Acido poligalacturónico
25, 26			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
27, 28					Pectato liasa		
29, 30					Pectinasa		
3, 4					Pectinasa		

FIGURA 5B

31, 32	Algodón	4,3	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
31, 32			temp ambiente	8	Pectato liasa	essay de pga	Acido poligalaturónico
33, 34			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
35, 36	Algodón	1,6	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
35, 36			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
37, 38	Algodón	11,2	50	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
39, 40			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
41, 42	Algodón	2	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
43, 44				9	Pectato liasa		
45, 46			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
47, 48			temp ambiente	1	Pectato liasa	ensayo de pga	Acido poligalaturónico
49, 50					Pectato liasa		
5, 6					Pectato liasa		
51, 52					Pectato liasa		
53, 54	Algodón	13,4	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
55, 56					Pectato liasa		
55, 56			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
57, 58				9	Pectato liasa	ensayo de pga	Acido poligalaturónico
59, 60			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
61, 62	Algodón	12,5	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
61, 62			temp ambiente	1	Pectato liasa	ensayo de pga	Acido poligalaturónico
63, 64					Pectato liasa		
65, 66					Pectato liasa		
67, 68	Algodón	9,8	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
69, 70			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
7, 8			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
71, 72			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
73, 74	Algodón	0,8	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
73, 74			temp ambiente	1	Pectato liasa	ensayo de pga	Acido poligalaturónico
75, 76					Pectinasa		
77, 78	Algodón	16,2	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
79, 80					Pectinasa		
81, 82	Algodón	16	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
81, 82			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalaturónico
83, 84	Algodón	3,2	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón

FIGURA 5C

85, 86	Algodón	0,1	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
85, 86			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
87, 88	Algodón	9,4	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
89, 90					Pectinasa		
9, 10			temp ambiente	10	Pectato liasa	ensayo de pga	Acido poligalacturónico
91, 92					Pectato liasa		
93, 94					Pectato liasa		
95, 96	Algodón	5,4	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
95, 96			temp ambiente	9	Pectato liasa	ensayo de pga	Acido poligalacturónico
97, 98					Pectato liasa		
99, 100	Algodón	16	50	9	Pectato liasa	Aplicación de Biolimpieza	Algodón
99, 100	Algodón	12,2	40	9	Pectato liasa	Aplicación de Biolimpieza	Algodón

FIGURA 6

Mutantes positivos GSSM™ de Pectato liasa													
Mutante positivo	Mutación												
	A118H	A182V	T190L	A197G	S208K	T219M	T223E	S255R	S263K	N275Y	Y309W	S312V	
A	X			X	X	X				X	X	X	
B	X		X		X	X				X	X	X	
C	X	X			X		X		X	X	X		
D	X	X	X		X			X	X	X	X		
E	X				X				X	X	X		
F	X	X			X		X	X		X	X	X	
G	X		X		X		X	X	X	X	X	X	
H	X	X	X	X	X		X			X	X	X	
I	X			X	X		X		X		X		
J	X			X	X				X	X	X	X	
K	X		X	X	X	X	X		X		X	X	
L	X	X	X		X		X	X	X	X	X		
M	X	X		X	X		X		X	X	X		
N	X		X	X	X				X	X	X	X	
O	X			X	X				X		X	X	
P	X	X		X	X	X		X		X	X	X	
Q	X		X		X				X		X		
R	X		X	X	X				X	X	X		
S	X		X	X		X	X	X	X	X	X		

FIGURA 7

Mutante positivo	Temperatura de fusión (°C)	Actividad específica (SA)				
		a 30 °C	SA a 40 °C	SA a 50 °C	SA a 60 °C	SA a 70 °C
AA	60,6	370,3	582,1	848,2	ND	ND
BB	58,3	366,8	542,4	648,6	ND	ND
CC	57	331,1	439,7	713,6	ND	ND
DD	58	468,6	595,8	714,6	ND	ND
EE	60	102,2	202,5	378,1	ND	ND
FF	58,5	585,5	744,5	955,8	ND	ND
GG	57,2	323,8	590,4	909,3	ND	ND
HH	58,7	267,6	425,5	706,3	ND	ND
II	57,8	357,1	527,3	875,1	ND	ND
JJ	57,5	372,2	537,6	834,9	ND	ND
KK	58,9	444,1	678,9	859,4	ND	ND
LL	58,4	375,2	557,6	1007,1	ND	ND
A	71	ND	ND	311,1	483,3	777,7
B	70,4	ND	ND	317,0	432,4	628,0
C	72,5	ND	ND	377,6	468,6	849,1
D	73,25	ND	ND	323,8	352,2	926,4
E	71	ND	ND	340,4	557,6	641,7
F	71,75	ND	ND	389,3	550,8	438,3
G	73	ND	ND	297,4	545,9	790,4
H	72,25	ND	ND	356,1	480,3	980,2
I	70,4	ND	ND	363,9	701,4	666,2
J	72,8	ND	ND	241,6	622,2	694,6
K	71,5	ND	ND	298,4	701,4	700,4
L	73	ND	ND	245,5	947,0	1003,7
M	73	ND	ND	304,2	547,8	309,1
N	73	ND	ND	685,8	1010,5	1284,5
O	71,3	ND	ND	268,0	517,5	330,7
P	72	ND	ND	398,2	542,9	689,7
Q	70,8	ND	ND	354,1	382,5	526,3
R	71,9	ND	ND	361,0	408,9	722,0
S	70,9	ND	ND	753,3	1186,6	579,1

FIGURA 8

1020												
	1	2	3	4	5	6	7	8	9	10	11	12
A	31.031	49.339	NoFl	10.691	380.25	NoFl	-21.00	35.459	70.312	20.312	89.575	27.906
B	357.39	13.439	0.789	59.177	NoFl	NoFl	-2.769	-2.240	9.632	-0.259	391.55	NoFl
C	2.797	29.459	20.911	341.11	45.819	9.219	NoFl	75.481	71.450	134.58	NoFl	35.029
D	0.853	24.278	33.829	40.309	25.291	12.819	37.999	59.704	51.230	33.400	31.101	39.549
E	19.999	3.173	11.325	NoFl	37.245	-18.36	5.922	NoFl	21.055	NoFl	4.394	NoFl
F	10.769	13.045	NoFl	NoFl	9.395	25.479	-7.969	4.155	715.69	1.097	0.253	-4.394
G	0.478	4.149	-0.394	5.544	31.309	7.839	7.145	57.099	7.839	19.154	-32.02	10.394
H	9.684	51.590	26.901	NoFl	1.185	3.019	49.200	0.000	9.823	39.539	399.79	9.919

Cinética
Tiempo: 2:00
Interval: 0:11
Lecturas: 11

Unit: 250

Automatic: Onco
Calibration: Onco

Tiempo de espera: 0:00
Tiempo final: 2:00
DO Mín: 0
DO Máx: 0,5
Vmax Pts: 11/11

Plate Last Read:
3:59 PM 12/22/2002

Combinación de longitud de onda: Lmt
Modo de Datos: Absorbancia