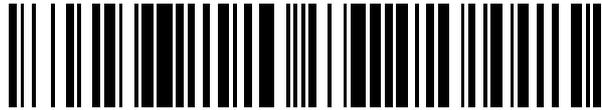


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 462 365**

51 Int. Cl.:

C12N 9/88 (2006.01)

C12N 15/60 (2006.01)

C12P 7/58 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **29.05.2006 E 06743559 (4)**

97 Fecha y número de publicación de la concesión europea: **02.04.2014 EP 1885848**

54 Título: **Conversión microbiana de azúcares ácidos y medios útiles para ello**

30 Prioridad:

30.05.2005 FI 20055263

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

22.05.2014

73 Titular/es:

**TEKNOLOGIAN TUTKIMUSKESKUS VTT
(100.0%)
Vuorimiehentie 3
02044 VTT , FI**

72 Inventor/es:

**KUORELAHTI, SATU;
PENTTILÄ, MERJA y
RICHARD, PETER**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 462 365 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

Conversión microbiana de azúcares ácidos y medios útiles para ello

Campo de la invención

5 La presente invención se refiere a enzimas implicadas en la conversión de azúcares ácidos, y más concretamente, a una proteína enzimática y su uso y producción. También se refiere a moléculas de ADN que codifican dichas enzimas, y a moléculas de ADN genéticamente modificadas y a microorganismos que comprenden dicho ADN. La invención también se refiere a microorganismos genéticamente modificados, en los que el gen codificador de la enzima ha sido inactivado y al uso de dicho microorganismo.

Antecedentes de la invención

10 Los residuos biológicos de la industria, incluyendo la agricultura, contienen azúcares y sus derivados, tales como azúcares ácidos. La conversión de dichos residuos en productos útiles ha suscitado interés y ha representado un reto en el campo de la biotecnología durante mucho tiempo. El ácido D-galacturónico es el principal componente de la pectina, una materia prima de bajo precio enriquecida, por ejemplo, en la pulpa de la remolacha azucarera, y una fuente de carbono para microorganismos que se alimentan de materia vegetal en descomposición.

15 En bacterias se conoce una vía que consiste en 5 enzimas que convierten el ácido D-galacturónico (D-galacturonato) en piruvato y D-gliceraldehído 3-fosfato (figura 1). Los metabolitos intermedios son D-tagaturonato, D-altronato, ácido D-eritro-3-desoxi-hexulosónico (2-ceto-3-desoxi D-gluconato) y ácido D-eritro-3-desoxi-hexulosónico 6-fosfato (2-ceto-3-desoxi D-gluconato 6-fosfato). Las enzimas son uronato isomerasa (EC 5.3.1.12), una D-tagaturonato reductasa que emplea NADH (EC 1.1.1.5), altronato deshidratasa (EC 4.2.1.7), 2-ceto-3-desoxi D-gluconatoquinasa (EC 2.7.1.45) y 2-ceto-3-desoxi D-gluconato 6-fosfatoaldolasa (EC 4.1.2.14).

20 La vía de la figura 1 solo ha sido descrita para organismos procariontes, es decir, no existen informes acerca de una vía similar en microorganismos eucariotes. Debe existir una vía en los microorganismos eucariotes, puesto que muchas especies de levaduras y mohos pueden utilizar y crecen sobre el D-galacturonato, pero se sabe muy poco acerca de esta vía.

25 Solo existen algunos estudios sobre el catabolismo del ácido D-galacturónico en microorganismos eucariotes. Uitzetter *et al.*, 1986 (J. Gen. Microbiol., 132, 1167-1172) mutagenizaron el hongo filamentoso *Aspergillus nidulans* y descubrieron que los mutantes que carecen de actividad piruvato deshidrogenasa o piruvato carboxilasa son incapaces de crecer sobre D-galacturonato, mientras que un mutante de piruvato quinasa fue capaz de crecer sobre D-galacturonato. Se interpretó que esto indica que el D-galacturonato se convierte en piruvato pero no a través de fosfoenolpiruvato, es decir, esto sería similar al caso de las bacterias. Visser *et al.*, 1988 (J. Gen. Microbiol., 134, 655-659) han especulado que, en *A. nidulans*, el ácido D-galacturónico se cataboliza a través de gliceraldehído y piruvato, lo cual se diferencia de la vía bacteriana en que las bacterias lo metabolizan a través de D-gliceraldehído 3-fosfato. También se ha sugerido que el ácido D-galacturónico se metaboliza a través de glicerol, puesto que un mutante de glicerol quinasa presenta un menor crecimiento sobre ácido D-galacturónico (Witteveen, C.F. *et al.*, 30 1990, J. Gen. Microbiol., 136, 1299-1305), y una glicerol deshidrogenasa dependiente de NADP fue inducida por el ácido D-galacturónico (Sealy-Lewis, H.M. y Fairhurst, V., 1992, Curr. Genet., 22, 293-296).

35 No existen informes sobre genes que sean similares a los genes de la vía del ácido D-galacturónico bacteriana, tal como se muestra en la figura 1, en el genoma de ningún microorganismo eucariota del cual se haya secuenciado el genoma. Esto sugiere que existe una vía eucariota para el catabolismo del ácido D-galacturónico que es diferente de la vía bacteriana.

40 En hongos, se ha sugerido que el ácido D-galacturónico es convertido en galactonato por una aldoceto reductasa, tras lo cual una deshidratasa o racemasa modifica el galactonato para producir 2-ceto-3-desoxigalactonato, y una aldolasa divide el 2-ceto-3-desoxigalactonato en piruvato y gliceraldehído. Martens-Uzunova, E. *et al.*, (Fungal Genetics Newsletter, vol. 52, suplemento (185), XXIII Fungal Genetics Conference, 15-20 de marzo, 2005, Pacific Grove, California) han identificado un agrupamiento de genes coexpresados que codifican la aldoceto reductasa, la racemasa y la aldolasa putativas necesarias. No se identificó ninguna deshidratasa, ni los autores explican el papel de la racemasa. De hecho, no mencionan si dicho galactonato o dicho 2-ceto-3-desoxigalactonato o dicho gliceraldehído están en la configuración L o D.

45 La presente invención se basa en el descubrimiento de un nuevo gen y una enzima implicada en el metabolismo fúngico del ácido D-galacturónico. Este descubrimiento revela una vía metabólica putativa del ácido D-galacturónico. El ADN que comprende el gen puede utilizarse para producir microorganismos genéticamente modificados que son capaces de fermentar, de forma eficaz, carbohidratos y sus derivados, tales como azúcares ácidos y sus derivados, a partir de un biomaterial para obtener productos de la fermentación útiles, tales como etanol.

50 Un objetivo de la invención es proporcionar una proteína enzimática, que puede ser expresada por un hospedante para la conversión de azúcares ácidos y sus derivados en productos de la conversión útiles en un medio de fermentación, o que está en forma de una preparación enzimática para la conversión *in vitro* de azúcares ácidos y

sus derivados en productos finales o productos intermedios útiles.

Otro objetivo de la invención es proporcionar un organismo genéticamente modificado en el que se evita la expresión del gen y que, por tanto, es capaz de acumular el sustrato de esta enzima.

5 La nueva molécula de ADN codifica una azúcar ácido deshidratasa que es activa sobre azúcares ácidos, en los que el grupo hidroxilo de C2 está en la configuración L y el grupo hidroxilo de C3 está en la configuración D en la proyección de Fischer. La enzima no muestra actividad con azúcares ácidos en los que el grupo hidroxilo de C2 está en la configuración D y el grupo hidroxilo de C3 está en la configuración L. Estas deshidratasas fueron previamente descritas, por ejemplo, en Niu *et al.* (J. Am. Chem. Soc., 2003, 125, 12998-12999), que describe una deshidratasa que es activa sobre el ácido L-arabónico y el ácido D-xilónico. Otro ejemplo es la D-gluconato deshidratasa que es activa en la vía de Entner-Doudoroff no fosforilada (véase, por ejemplo, Buchanan *et al.*, Biochem. J., 1999, 343, 563-570).

15 En un extracto bruto de la bacteria *Pseudomonas saccharophila* se ha descubierto actividad enzimática convertidora del ácido D-arabónico, y se cree que el producto de la reacción es el ácido 2-ceto-3-desoxi-D-arabónico (Palleroni, N.J. y Doudoroff, M., 1956, J. Biol. Chem., 223:499-508). Sin embargo, no se aisló ni se expresó ningún gen. El documento EP 1 496 113 describe una gluconato deshidrogenasa bacteriana y el gen codificador. La enzima es capaz de convertir el ácido D-galactónico y el ácido L-arabónico en ácido 2-ceto-3-desoxi-D-galactónico y el ácido 2-ceto-3-desoxi-D-arabónico, respectivamente. Elshafei, A.M. *et al.*, 1995, Antonie van Leeuwenhoek, vol. 67, n.º 2, pp. 211-216 describen la D-galactonato deshidratasa de *Aspergillus terreus*, que deshidrata el D-galactonato para producir 2-ceto-3-desoxi-D-galactonato. Ambas enzimas bacteriana y fúngica descritas deshidratan el D-enantiómero del ácido galactónico.

Resumen de la invención

La invención proporciona una molécula de ADN aislada que comprende un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, en la que dicho ADN codifica una proteína enzimática que comprende una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con SEQ ID NO:2.

La invención proporciona también una molécula de ADN genéticamente modificada, que es un vector que comprende dicha molécula de ADN, y un microorganismo hospedante que la contiene, y que es un microorganismo genéticamente modificado transformado con dicha molécula de ADN genéticamente modificada.

30 La invención también proporciona una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico y que tiene al menos 90% de identidad de secuencia con SEQ ID NO:2, y un método para producir dicha enzima mediante el cultivo del microorganismo hospedante bajo condiciones que permiten la expresión de dicha proteína, y la recuperación de la proteína enzimática.

35 También se proporciona un método para convertir el ácido L-galactónico o el ácido D-arabónico en ácido L-treo-3-desoxi-hexulosónico o ácido D-glicero-3-desoxi-pentulosónico, respectivamente, poniendo en contacto el ácido L-galactónico o el ácido D-arabónico con una proteína enzimática que tiene al menos 30% de identidad de secuencia con SEQ ID NO:2.

La invención incluye también el uso de la anterior proteína enzimática para producir un compuesto deseado a partir de un material que comprende un azúcar ácido o uno de sus derivados, y una preparación enzimática que comprende dicha enzima.

40 La invención incluye también una cepa de *Hypocrea jecorina*, que está genéticamente modificada para inactivar el gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, y un método para producir ácido L-galactónico o ácido D-arabónico proporcionando un microorganismo que tiene dicho gen, inactivando dicho gen, y utilizando el microorganismo genéticamente modificado para la producción de ácido L-galactónico o ácido D-arabónico.

45 Otros objetos, detalles y ventajas de la presente invención serán evidentes a partir de los siguientes dibujos, descripción detallada y ejemplos.

Breve descripción de los dibujos

La figura 1 muestra la vía bacteriana para la utilización del ácido D-galacturónico.

La figura 2 muestra la vía fúngica putativa para la utilización del ácido D-galacturónico.

50 La figura 3 muestra la secuencia de ADN de la región codificadora y la secuencia de aminoácidos para la ácido L-galactónico deshidratasa. La línea superior indica la secuencia de ADN, las letras mayúsculas son la secuencia codificadora, y las letras minúsculas son la secuencia del intrón. La línea inferior muestra la secuencia de aminoácidos.

La figura 4 muestra el plásmido pBluekan7-1.NotI, que se empleó para deleccionar el gen de la ácido L-galactónico deshidratasa.

Descripción detallada de la invención

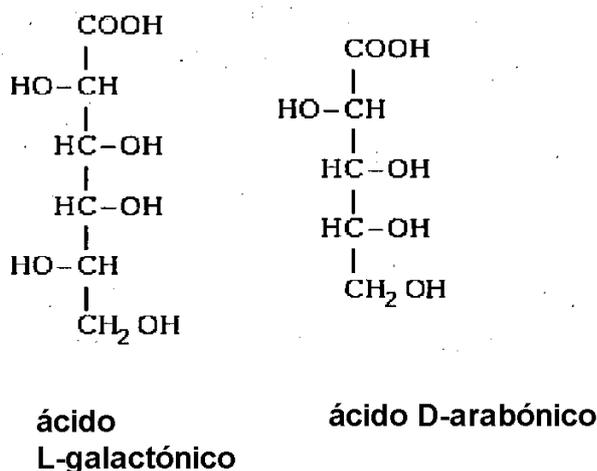
5 Un vía fúngica putativa, que está claramente diferenciada de la vía bacteriana previamente descrita, se resume en la figura 2. En esta vía, el ácido D-galacturónico primero se convierte en ácido L-galactónico mediante una ácido D-galacturónico reductasa. Se ha identificado previamente un gen para la ácido D-galacturónico reductasa en plantas, y se ha descrito la actividad de la enzima en levaduras.

La segunda etapa, en la que el ácido L-galactónico es convertido en ácido L-treo-3-desoxi-hexulosónico por una deshidratasa, es nueva. Los inventores han clonado el gen y expresado la actividad en un hospedante heterólogo.

10 La tercera etapa es una reacción de aldolasa para preparar L-gliceraldehído y piruvato. El L-gliceraldehído puede ser convertido en glicerol por una NADP glicerol deshidrogenasa, puesto que esta enzima es inducida por el ácido D-galacturónico.

La presente invención proporciona, por primera vez, una molécula de ADN aislada, que comprende un gen que codifica una proteína enzimática, que muestra actividad ácido L-galactónico deshidratasa. El aislamiento y el procedimiento de identificación se describen a continuación. La secuencia de ADN de la región codificadora y la secuencia de aminoácidos de la ácido L-galactónico deshidratasa se indica en la figura 3.

La nueva molécula de ADN codifica una ácido L-galactónico deshidratasa, que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico (también denominado ácido 2-ceto-3-desoxi-L-galactónico). También es activa con el ácido D-arabónico (también denominado ácido D-arabinoico), que se convierte en ácido D-glicero-3-desoxi-pentulosónico (también denominado ácido 2-ceto-3-desoxi-D-arabónico). De forma más general, la enzima es activa sobre azúcares ácidos o sus derivados, en los que el grupo hidroxilo de C2 está en la configuración L y el grupo hidroxilo de C3 está en la configuración D en la proyección de Fischer. Un azúcar ácido es un azúcar oxidado es uno o ambos extremos. Un "derivado de un azúcar ácido" puede ser cualquier compuesto que puede obtenerse a partir de un azúcar ácido o que es un homólogo de un azúcar ácido, y que tiene un grupo carboxilo en C1, un grupo hidroxilo en la configuración L en C2 y en la configuración D en C3. Los otros átomos de C, y en especial el átomo terminal, pueden comprender, por ejemplo, un grupo metilo o éster. Preferiblemente, el azúcar ácido o su derivado comprende de cinco a seis átomos de C, en especial seis átomos de C.



30 Sin embargo, la enzima no muestra actividad con el ácido D-glucónico o el ácido D-xilónico, en los que los grupos hidroxilo de C2 están en la configuración D y los grupos hidroxilo de C3 están en la configuración L en la proyección de Fischer.

Resulta evidente que las expresiones "molécula de ADN", "secuencia de ADN" y "secuencia de ácido nucleico" incluyen ADN genómico y ADNc (ADN complementario).

35 Según una realización de la invención, la secuencia de ADN aislada se deriva de una especie de *Hypocrea* (previamente *Trichoderma*). Según una realización específica, el ADN está en un depósito realizado en the International Depository Authority, Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, Mascheroder Weg 1 b, D-38124 Braunschweig) bajo las condiciones del tratado de Budapest, el 30 de marzo, 2005, con el n.º de registro DSM 17214. Este depósito comprende la secuencia de ADNc que tiene SEQ ID NO:1. El gen

heterólogo se ha denominado en la parte experimental que aparece a continuación el gen *lgd1*, y está presente en un plásmido de múltiples copias bajo el control de un promotor de levadura constitutivo. En esta cepa se expresa la ácido L-galactónico deshidratasa. La secuencia de ácido nucleico depositada se origina de la cepa del moho *Hypocrea jecorina* (*Trichoderma reesei*) Rut C-30 (ATCC 56765). La secuencia de aminoácidos deducida de SEQ ID NO:1 se indica como SEQ ID NO:2. En la base de datos Uniprot, Galagan, J.E. *et al.*, 15 de diciembre, 2003, el n.º de registro Q7S6F9 se refiere a una proteína hipotética que tienen 82,8% de identidad de secuencia con SEQ ID NO:2 en un solapamiento de 450 aminoácidos. No se indica la actividad de la proteína.

Se sabe que genes de diferentes organismos que codifican enzimas con la misma actividad catalítica tienen similitudes de secuencia y que estas similitudes pueden ser aprovechadas de muchas maneras por los expertos en la técnica para clonar otros genes de otros organismos con la misma actividad catalítica o una actividad catalítica similar. Estos genes también son adecuados para practicar la presente invención. Por tanto, las moléculas de ADN aisladas que pueden obtenerse de cualquier organismo, y en especial de organismos eucariotas, tales como hongos, que incluyen levaduras, plantas y animales, que incluyen al ser humano, se incluyen en la invención. Preferiblemente, la molécula de ADN se deriva de un hongo filamentoso.

Las moléculas de ADN de la invención pueden obtenerse, por ejemplo, a través de simulación electrónica, comparando secuencias de nucleótidos. Si estas secuencias no están disponibles se puede identificar una región conservada en la secuencia de nucleótidos o de aminoácidos, y clonar un fragmento de gen utilizando técnicas de PCR. Después de secuenciar el fragmento puede obtenerse el gen completo, por ejemplo, utilizando un banco de ADNc en un vector, según se describe en Richard *et al.* (2001), J. Biol. Chem., 276, 40631-40637. Otra manera para identificar un gen de ácido L-galactónico deshidratasa es mediante hibridación de ácidos nucleicos convencional.

Resulta evidente que muchas pequeñas variaciones en la secuencia de nucleótidos de un gen no cambian significativamente las propiedades catalíticas de la proteína codificada. Por ejemplo, muchos cambios en la secuencia de nucleótidos no cambian la secuencia de aminoácidos de la proteína codificada. Además, una secuencia de aminoácidos puede tener variaciones que no cambian las propiedades funcionales de una proteína, en particular no evitan que una enzima desempeñe su función catalítica. Estas variaciones en la secuencia de nucleótidos de las moléculas de ADN o en una secuencia de aminoácidos se denominan "equivalentes funcionales", porque no cambian significativamente la función del gen para que codifique una proteína con una función concreta, por ejemplo, catalizar una reacción concreta o, respectivamente, cambiar la función concreta de la proteína. Por tanto, estos equivalentes funcionales, que incluyen los fragmentos, de la secuencia de nucleótidos de SEQ ID NO:1 y de la secuencia de aminoácidos de SEQ ID NO:2, respectivamente, se describen en la presente.

Un equivalente funcional de una secuencia de ácido nucleico también incluye secuencias de ácidos nucleicos que son capaces de hibridarse con las secuencias identificadas bajo condiciones de rigurosidad intermedia o alta. Por ejemplo, puede realizarse una hibridación de rigurosidad intermedia en una mezcla de hibridación que contenga 6x SSC (NaCl 0,9 M en citrato de sodio 0,09 M, pH 7), dodecilsulfato de sodio (SDS) al 0,5%, 5x disolución de Denhardt y 100 µg/ml de ADN de espera de arenque a 50 °C. Puede realizarse una hibridación de rigurosidad alta, por ejemplo, en la misma mezcla de hibridación a 68 °C.

En una realización específica de la invención, la proteína enzimática comprende la secuencia de aminoácidos de SEQ ID NO:2 o uno de sus equivalentes funcionales. Los equivalentes funcionales incluyen una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con SEQ ID NO:2.

Además, la invención se dirige a una molécula de ADN genéticamente modificada, es decir, un ADN recombinante, que es un vector, que comprende el gen de la molécula de ADN de la invención, según se describió anteriormente, de modo que puede ser expresado en una célula hospedante, es decir, un microorganismo. En el ADN recombinante, el gen de la invención puede estar unido operablemente a un promotor. El vector puede ser, por ejemplo, un vector convencional, tal como un virus, por ejemplo, un bacteriófago, o un plásmido, preferiblemente un plásmido. La construcción de un vector de expresión está dentro del conocimiento de los expertos en la técnica. A continuación se describe el procedimiento general y ejemplos específicos.

La presente invención también hace posible generar un organismo genéticamente modificado en el que la actividad ácido L-galactónico deshidratasa está ausente. En este organismo se acumula el ácido L-galactónico, es decir, este organismo sería adecuado para producir ácido L-galactónico a partir del ácido D-galacturónico o a partir de otros sustratos a partir de los cuales puede derivarse el ácido L-galactónico. En consecuencia, puede acumularse ácido D-arabónico a partir de D-arabinosa. El conocimiento de la secuencia de ADN de la ácido L-galacturónico deshidratasa puede utilizarse para inactivar el correspondiente gen o genes en un microorganismo adecuado. El gen puede ser inactivado, por ejemplo, evitando su expresión o mediante mutación o delección del gen o de parte de este. Existen diversas técnicas para inactivar un gen. Estas técnicas emplean la secuencia de nucleótidos del gen o la secuencia de nucleótidos en la proximidad del gen. La construcción de un microorganismo en el que se ha evitado la expresión del gen de la ácido L-galactónico deshidratasa, o este haya sido mutado o delecionado está dentro del conocimiento de los expertos en la técnica. Naturalmente, el gen puede ser inactivado en cualquier microorganismo que tenga dicho gen, y la invención hace posible identificar estos microorganismos. A continuación se describe el procedimiento general y ejemplos específicos.

El ácido L-galactónico puede utilizarse, por ejemplo, como acidificante en la industria alimentaria, o puede utilizarse en cosméticos o en la industria del hormigón.

La molécula de ADN que codifica una ácido L-galactónico deshidratasa puede transferirse a cualquier microorganismo adecuado o el gen que codifica una ácido L-galactónico deshidratasa puede deleccionarse en cualquier microorganismo adecuado. Un microorganismo adecuado puede ser adecuado para la producción de los productos de la conversión deseados o puede ser adecuado para acceder a los sustratos requeridos. Un ejemplo es un microorganismo fúngico, que utilice de forma eficaz el ácido D-galacturónico. En este microorganismo, la deleción de la ácido L-galactónico deshidratasa conduciría a una acumulación de ácido L-galactónico durante el proceso de fermentación. Otro ejemplo es un microorganismo en el que se está acumulando el ácido D-arabónico o el ácido L-galactónico, y la expresión de la ácido L-galactónico deshidratasa facilita su conversión en los productos de reacción deseados.

Naturalmente, cualquiera que sea el material que vayan a utilizar dichos microorganismos de la invención, este comprenderá el azúcar ácido que es convertible en presencia de la ácido L-galactónico deshidratasa, o el microorganismo será capaz de expresar otros genes para producir las enzimas necesarias para la conversión del material de partida en un azúcar ácido utilizable por dicha deshidratasa expresada por el gen de la invención. El material de partida es preferiblemente de origen natural, es decir, un biomaterial, por ejemplo, una biomasa que comprende azúcares, azúcares ácidos o sus derivados. Un ejemplo de un biomaterial adecuado es la pulpa de la remolacha azucarera, que comprende pectina, que consiste principalmente en ácido D-galacturónico. Además, pueden utilizarse otros materiales que comprendan pectina.

Según una realización de la invención, la biomasa que comprende un azúcar ácido o uno de sus derivados es fermentada por un microorganismo transformado con una molécula de ADN que comprende un gen que codifica una proteína enzimática capaz de convertir el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, y el compuesto deseado producido se recupera. Si el microorganismo transformado también expresa una aldolasa capaz de convertir el ácido L-treo-3-desoxi-hexulosónico en L-gliceraldehído y piruvato, y el L-gliceraldehído después se convierte, por ejemplo, en glicerol, estos metabolitos pueden ser convertidos por el microorganismo en etanol, ácido láctico o cualquier otro compuesto metabólicamente derivable de estos metabolitos, utilizando la vía metabólica de ese microorganismo. Dicho piruvato también puede ser convertido después por el microorganismo en etanol a través de la piruvato descarboxilasa y la alcohol deshidrogenasa, en ácido láctico a través de la lactato deshidrogenasa, o en cualquier otro compuesto que puede derivarse metabólicamente del piruvato.

La invención no se limita a modificar genéticamente el moho o la levadura. Los genes que codifican la ácido L-galactónico deshidratasa pueden expresarse en cualquier organismo, tales como bacterias, plantas o eucariotas superiores, aplicando las herramientas genéticas adecuadas y conocidas en la técnica para ese organismo concreto. Por tanto, el término "microorganismo" debe interpretarse en un sentido amplio e incluye también líneas celulares de organismos superiores.

De manera conveniente, la ácido L-galactónico deshidratasa es producida mediante la tecnología recombinante. Esto se refiere al aislamiento de un fragmento que comprende el gen de la deshidratasa mediante amplificación en una reacción de PCR (Coen, D.M., 2001, The polymerase chain reaction, publicado en: Ausubel, F.M., Brent, R., Kingston, R.E., More, D.D., Seidman, J.G., Smith, K. y Struhl, K. (eds.), Current protocols in molecular biology, John Wiley & Sons. Inc., Hoboken, EEUU), u otros métodos de ADN recombinante (Sambrook, J., Fritsch, E.F. y Maniatis, T., 1989, Molecular cloning: A laboratory manual, 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), la inserción del gen bajo el control de un promotor fuerte en un vector de expresión, la transferencia del vector a células hospedantes adecuadas, y el cultivo de las células hospedantes en condiciones que provocan la expresión de dicha enzima. Los métodos para la producción de proteínas mediante la tecnología recombinante en diferentes sistemas de hospedantes son muy conocidos por los expertos en la técnica (Gellissen, G. (ed.), 2005, Production of recombinant proteins. Novel microbial and eukaryotic expression systems, Wiley-VCH Verlag GmbH & Co., Weinheim, Alemania). Como alternativa, solo el promotor fuerte está unido operablemente al gen de la deshidratasa en el cromosoma del hospedante, con lo cual la expresión de dicho gen está sobreexpresada. La proteína expresada puede aislarse y purificarse mediante métodos de purificación de proteínas convencionales.

También se describe una preparación de enzimas que comprende la ácido L-galactónico deshidratasa. Esta preparación puede ser un extracto de células bruto del organismo genéticamente modificado, o la enzima puede ser purificada de este, con lo cual la preparación comprende al menos la ácido L-galactónico deshidratasa en una forma purificada. La preparación también puede comprender otras enzimas que participan en el catabolismo de los azúcares o los azúcares ácidos o sus derivados.

Además, la invención proporciona el uso de una ácido L-galactónico deshidratasa para la conversión del ácido L-galactónico o del ácido D-arabónico o, más en general, para la conversión de azúcares ácidos o sus derivados, en los que los grupos hidroxilo de C2 y C3 están en la configuración L y D, respectivamente, para producir los productos descritos previamente.

Es evidente que en todos los casos en que se describe un azúcar ácido, tal como el ácido D-galacturónico, el ácido L-galactónico, el ácido L-treo-3-desoxi-hexulosónico, el ácido D-arabónico y todos los demás ácidos, la descripción

también incluye la forma aniónica del azúcar ácido, es decir, D-galacturonato, L-galactonato, L-treo-3-desoxi-hexulosoato, D-arabonato o la correspondiente forma aniónica del ácido, porque, en la práctica, a menudo es difícil distinguir entre la forma disociada y no disociada del ácido.

- 5 La invención se ilustra mediante los siguientes ejemplos no limitantes. Sin embargo, debe entenderse que las realizaciones indicadas en la anterior descripción y en los ejemplos solo tienen fines ilustrativos, y que son posibles diversos cambios y modificaciones.

Ejemplo 1

Clonación de la ácido L-galactónico deshidratasa

10 El genoma de *Hypocrea jecorina* (*Trichoderma reesei*) se seleccionó para buscar genes con homología con las deshidratasas. Después, los marcos de lectura abiertos se amplificaron mediante PCR y se acoplaron con un vector de expresión de levadura. Para este fin se diseñaron cebadores de PCR que contenían sitios de restricción *Bam*HI para amplificar los marcos de lectura abiertos. El molde de PCR fue un banco de ADNc de *H. jecorina*. El producto de la PCR se acopló a un vector TOPO (Invitrogen). A partir del vector TOPO se liberó el fragmento *Bam*HI y se acopló en un vector de expresión de levadura. El vector de expresión se derivó del plásmido pYX212 (R&D Systems) mediante su digestión con *Eco*RI y *Xho*I para eliminar el ATG y HA-marcador del sitio de clonación múltiple y la introducción de un sitio de restricción *Bam*HI en el sitio de clonación insertando un fragmento cortado con *Eco*RI y *Sal*I del plásmido pUC19 (Norlander, J., Kempe, T. y Messing, J., 1983, *Gene*, 26, 101-106).

20 El vector resultante después se transformó en una cepa de *Saccharomyces cerevisiae*. La cepa de *S. cerevisiae* después de disgregó mediante agitación en vórtice con esferas de vidrio, y el extracto de levadura se analizó para la actividad ácido L-galactónico deshidratasa. Para ensayar la actividad ácido L-galactónico deshidratasa se mezcló ácido L-galactónico con el extracto de levadura y se siguió la formación de azúcares reductores utilizando protocolos convencionales (Bemfeld, P., 1955, *Amylases, α and β* , en: *Methods in enzymology*, vol. 1, Colovick S.P., Kaplan, N.O., eds., Academic Press, NY, pp. 149-158).

25 Utilizando los cebadores 5'-GGATCCACCATGTCTGAAGTCACCAT-3' en la dirección sentido, y el cebador 5'-GGATCCTCAGATCTTCTCTCCGTTCA-3' en la dirección antisentido se obtuvo una ácido L-galactónico deshidratasa activa después de la expresión en *S. cerevisiae*. El gen se denominó *ldg1*.

La cepa de *S. cerevisiae* que sobreexpresa la ácido L-galactónico deshidratasa se denomina H3350 y está depositada con el n.º de depósito DSM 17214.

Ejemplo 2

30 Identificación del producto de la reacción y ensayo de la especificidad de la ácido L-galactónico deshidratasa

El ácido L-galactónico se mezcló con el extracto de levadura de la cepa H3350, según se describe en el ejemplo 1. El producto de la reacción se identificó como un 2-ceto-3-desoxi azúcar ácido en un ensayo químico y se cuantificó como se describe en Buchanan *et al.*, 1999, *Biochem. J.*, 343, 563-570. La concentración de proteínas del extracto de levadura en el medio de reacción fue de 0,15 g/l, y la concentración inicial de ácido L-galactónico de 10 mM. Después de 21 horas se formó 1,04 mM de ácido L-treo-3-desoxi-hexulosónico.

40 Para ensayar la especificidad de la enzima, el extracto de levadura, según se describió anteriormente, se mezcló con los azúcares ácidos ácido D-glucónico, ácido D-arabónico, ácido D-xilónico, ácido L-gulónico y ácido L-galactónico. Se siguió la formación de azúcares reductores según se describe en el ejemplo 1. Se observó actividad con los azúcares ácidos ácido L-galactónico y ácido D-arabónico.

Ejemplo 3

Delección de la ácido L-galactónico deshidratasa en *H. jecorina*

45 Para la delección del gen *ldg1* se construyó un módulo de delección. Para el módulo de delección, se clonaron áreas de 1,5 kb de ambos lados del gen de la ácido L-galactónico deshidratasa y se acoplaron al plásmido pBluekan7-1.NotI (figura 4). La parte cadena arriba de *ldg1* se clonó utilizando los oligonucleótidos 5'-GAGCTCAAGCTTCCACGCAGTTGCTACTTCTA-3' y 5'-GAGCTCTGGTTATTTGGCAGAGCGAC-3' introduciendo sitios de restricción *Sac*I y *Hind*III. El fragmento *Sac*I se acopló en el sitio de clonación *Sac*I de pBluekan7-1.NotI. La parte cadena abajo de *ldg1* se clonó con los oligonucleótidos 5'-ACTAGTGGGGCAAAGTTGGACATGAT-3' and 5'-ACTAGTAAGCTTGC AATACCTGGACCAAGCTA-3' introduciendo sitios de restricción *Spe*I y *Hind*III. El fragmento *Spe*I se acopló en el sitio *Spe*I de pBluekan7-1.NotI. Los dos fragmentos de ADN después se acoplaron al vector pBluekan7-1.NotI de tal forma que la orientación de los dos fragmentos de ADN entre sí no cambia, y se colocó un gen para la resistencia a higromicina entre los dos fragmentos. El módulo de delección que comprende los dos fragmentos de ADN y el gen de resistencia a higromicina después fue liberado mediante digestión con *Hind*III, se transformó en la cepa Rut C-30 de *H. jecorina* y se seleccionó para la resistencia a higromicina utilizando protocolos

convencionales. La cepa resultante fue capaz de convertir el ácido D-galacturónico en ácido L-galactónico.

Ejemplo 4

Análisis de RMN del producto de la reacción

5 El ácido L-galactónico se mezcló con el extracto de levadura de la cepa H3350 según se describió en el ejemplo 1. El producto de la reacción se identificó mediante RMN.

10 Los experimentos de RMN se realizaron a 23 °C en un espectrómetro Varian Inova que funciona con una frecuencia de protón de 500 MHz. Los anchos espectrales de los espectros de ^1H y ^{13}C unidimensionales fueron de 5000 Hz y 30675 Hz, respectivamente. En experimentos DQFCOSY y TOCSY, el ancho espectral fue de 3400 Hz y se adquirieron matrices de 1024 x 128 puntos de datos complejos. El tiempo de bloqueo de los espines en el TOCSY fue de 80 ms. En HSQC, los anchos espectrales en las dimensiones de ^1H y ^{13}C fueron de 1654 Hz y 10000 Hz, respectivamente, y se adquirió una matriz de 1024 x 256 puntos de datos complejos. Todas las matrices de datos bidimensionales se rellenaron en cero en F1 y se aplicó una función de ponderación Cosine Bell en ambas dimensiones antes de la transformación de Fourier.

15 La estructura del producto de la reacción se verificó mediante espectroscopía de RMN, y los desplazamientos químicos de ^1H y ^{13}C del producto se indican en la tabla 1. A partir del espectro de ^1H unidimensional de la mezcla de reacción, son claramente visibles las señales del producto, y a partir de los experimentos DQFCOSY bidimensional y [^1H , ^{13}C]HSQC resultó evidente que el producto tenía un sistema de espín de protón CH₂-CH-CH-CH₂, en el que uno de los CH₂ tiene desplazamientos químicos típicos de un grupo hidroximetilo y el segundo tiene unos desplazamientos químicos bastante exclusivos hacia los grupos CH₂ cercanos a un grupo cetona o a una estructura de hemiacetal. El espectro DEPT también confirmó que la molécula tiene dos átomos de carbono de tipo CH₂ y dos de tipo CH. Además de estos cuatro carbonos, el espectro de ^{13}C reveló dos señales de carbono adicionales. Una está en el área del carboxilo cerca de la señal del carbono carboxílico del sustrato, y la otra (97,84 ppm) es típica de un carbono cuaternario en una estructura de hemiacetal, como las señales de C2 en ácidos siálicos. Los resultados de la RMN demuestran que el producto de la reacción es ácido 2-ceto-3-desoxi-galactónico y que existe predominantemente como un anillo de piranosa. Solo se detectaron señales de un anómero, pero no fue posible determinar cuál de los dos anómeros era.

Tabla 1. Desplazamientos químicos de RMN de ^1H y ^{13}C del producto de ácido 2-ceto-3-desoxi-galactónico

	δ		δ
	(ppm) ^a		(ppm) ^b
H3	1,789	C1	177,53
H3'	2,162	C2	97,84
H4	3,859	C3	40,22
H5	3,604	C4	70,13
H6	3,606	C5	71,92
H6'	3,801	C6	64,18

^a referenciado al TSP interno (0 ppm)

^b referenciado a la acetona externa (31,5 ppm)

30

LISTADO DE SECUENCIAS

<110> Valtion teknillinen tutkimuskeskus

<120> Conversión microbiana de azúcares ácidos y medios útiles para ello

5 <130> 2050724EP

<160> 8

<170> PatentIn versión 3.3

<210> 1

<211> 1353

10 <212> ADN

<213> Hypocrea jecorina

<400> 1

```

atgtctgaag tcaccatcac aggcttcagg agcccgcatg tgcggttccc cacgtcccta      60
gacaagacgg gctcggatgc gatgaacgct gcggcgact attcagcggc atactgcatc      120
ctcgagactg attcagcgc cagtggatcat ggcatgacat tcaccattgg acgcggaaac      180
gacatcgtct gcgccgccat caaccacgtc gcggaccgac tcaagggcaa gaaactgtca      240
tcactagtgg ccgactgggg caagaectgg cggtatctgg tcaacgacag ccagctgcgg      300
tggattggcc ccgaaaaggg cgtcatccat cttgcgctcg gagccgtcgt caacgcctc      360
tgggacctgt gggcaaagac gctcaacaag ccggtttggc gcatcgttgc cgacatgacg      420
cccaggagat atgtccgctg catcgacttc cgctacatta ccgacgcaat caccoccgag      480
gaagccgtgg cgatgctgcg cgagcaggag gccggcaagg ccaagcgcac cgaggaggct      540
ctccagaacc gagcggtgcc tgcatacaca acaagtgccg gttggctggg atacggagag      600
gacaagatga agcagctcct gagagagacg ctggctgccg gatacagaca cttcaaggtc      660
aaggttggcg gcagcgtcga ggaggaccga aggcgcctcg gcattgctcg cgaaattctt      720
ggtttcgaca agggcaacgt tctcatggtc gatgccaacc aggtctggtc cgttcccga      780
gcgatcgact acatgaagca gctcagcag tacaagccct ggttcattga ggagcccacc      840
tcaccgcagc acatcatggg ccacaaggcc attcgcgatg ccctcaagcc ctatggcatc      900
ggcgtcgtca ccggcgagat gtgccagaac cgcgtcatgt tcaagcagct gatcatgacg      960
ggcgccatcg acatctgcca gattgatgcc tgccgcctcg gcggcgtcaa cgaagtgctg     1020
gcogtcctgc tcatggccaa gaagtacggg gtgccattg tgccgcattc cggcgggcgtg     1080
ggccttcccg agtacacca gcatctgagc accatcgact acgtggctcgt cagcggcaag     1140
ctttccgtct tggagtttgt agaccacctc cacgagcact tcttgcattc ttcagtcac     1200
aaggacggat actaccagac accaaccgag gccggctaca gcgcttgagat gaagccggag     1260
agcatggaca agtatgagta tcccggaag aaggcgtaa gttggtggac gaccgacgag     1320
gctctgccca tcttgaacgg agagaagatc tga                                     1353
    
```

<210> 2

<211> 450

<212> PRT

15 <213> Hypocrea jecorina

ES 2 462 365 T3

<400> 2

Met Ser Glu Val Thr Ile Thr Gly Phe Arg Ser Arg Asp Val Arg Phe
 1 5 10 15

Pro Thr Ser Leu Asp Lys Thr Gly Ser Asp Ala Met Asn Ala Ala Gly
 20 25 30

Asp Tyr Ser Ala Ala Tyr Cys Ile Leu Glu Thr Asp Ser Ala His Ser
 35 40 45

Gly His Gly Met Thr Phe Thr Ile Gly Arg Gly Asn Asp Ile Val Cys
 50 55 60

Ala Ala Ile Asn His Val Ala Asp Arg Leu Lys Gly Lys Lys Leu Ser
 65 70 75 80

Ser Leu Val Ala Asp Trp Gly Lys Thr Trp Arg Tyr Leu Val Asn Asp
 85 90 95

Ser Gln Leu Arg Trp Ile Gly Pro Glu Lys Gly Val Ile His Leu Ala
 100 105 110

Leu Gly Ala Val Val Asn Ala Val Trp Asp Leu Trp Ala Lys Thr Leu
 115 120 125

Asn Lys Pro Val Trp Arg Ile Val Ala Asp Met Thr Pro Glu Glu Tyr
 130 135 140

Val Arg Cys Ile Asp Phe Arg Tyr Ile Thr Asp Ala Ile Thr Pro Glu
 145 150 155 160

Glu Ala Val Ala Met Leu Arg Glu Gln Glu Ala Gly Lys Ala Lys Arg
 165 170 175

Ile Glu Glu Ala Leu Gln Asn Arg Ala Val Pro Ala Tyr Thr Thr Ser
 180 185 190

Ala Gly Trp Leu Gly Tyr Gly Glu Asp Lys Met Lys Gln Leu Leu Arg
 195 200 205

Glu Thr Leu Ala Ala Gly Tyr Arg His Phe Lys Val Lys Val Gly Gly
 210 215 220

Ser Val Glu Glu Asp Arg Arg Arg Leu Gly Ile Ala Arg Glu Ile Leu
 225 230 235 240

Gly Phe Asp Lys Gly Asn Val Leu Met Val Asp Ala Asn Gln Val Trp
 245 250 255

ES 2 462 365 T3

Ser Val Pro Glu Ala Ile Asp Tyr Met Lys Gln Leu Ser Glu Tyr Lys
 260 265 270

Pro Trp Phe Ile Glu Glu Pro Thr Ser Pro Asp Asp Ile Met Gly His
 275 280 285

Lys Ala Ile Arg Asp Ala Leu Lys Pro Tyr Gly Ile Gly Val Ala Thr
 290 295 300

Gly Glu Met Cys Gln Asn Arg Val Met Phe Lys Gln Leu Ile Met Thr
 305 310 315 320

Gly Ala Ile Asp Ile Cys Gln Ile Asp Ala Cys Arg Leu Gly Gly Val
 325 330 335

Asn Glu Val Leu Ala Val Leu Leu Met Ala Lys Lys Tyr Gly Val Pro
 340 345 350

Ile Val Pro His Ser Gly Gly Val Gly Leu Pro Glu Tyr Thr Gln His
 355 360 365

Leu Ser Thr Ile Asp Tyr Val Val Val Ser Gly Lys Leu Ser Val Leu
 370 375 380

Glu Phe Val Asp His Leu His Glu His Phe Leu His Pro Ser Val Ile
 385 390 395 400

Lys Asp Gly Tyr Tyr Gln Thr Pro Thr Glu Ala Gly Tyr Ser Val Glu
 405 410 415

Met Lys Pro Glu Ser Met Asp Lys Tyr Glu Tyr Pro Gly Lys Lys Gly
 420 425 430

Val Ser Trp Trp Thr Thr Asp Glu Ala Leu Pro Ile Leu Asn Gly Glu
 435 440 445

Lys Ile
 450

5 <210> 3
 <211> 26
 <212> ADN
 <213> Artificial

<220>
 <223> Cebador

10 <400> 3
 ggatccacca tgtctgaagt caccat

<210> 4
 <211> 26
 <212> ADN
 <213> Artificial

15

ES 2 462 365 T3

	<220>	
	<223> Cebador	
	<400> 4	
	ggatcctcag atcttctctc cgttca	26
5	<210> 5	
	<211> 32	
	<212> ADN	
	<213> Artificial	
10	<220>	
	<223> Cebador	
	<400> 5	
	gagctcaagc ttccacgcag ttgctacttc ta	32
15	<210> 6	
	<211> 26	
	<212> ADN	
	<213> Artificial	
20	<220>	
	<223> Cebador	
	<400> 6	
	gagctctggt tattggcag agcgac	26
25	<210> 7	
	<211> 26	
	<212> ADN	
	<213> Artificial	
30	<220>	
	<223> Cebador	
	<400> 7	
	actagtgggg caaagttgga catgat	26
35	<210> 8	
	<211> 32	
	<212> ADN	
	<213> Artificial	
35	<220>	
	<223> Cebador	
	<400> 8	
	actagtaagc ttgcaatacc tggaccaagc ta	32

REIVINDICACIONES

- 1.- Una molécula de ADN aislada que comprende un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, en la que dicho ADN codifica una proteína enzimática que comprende una secuencia de aminoácidos que tiene al menos 90% de identidad de secuencia con SEQ ID NO:2.
- 5 2.- La molécula de ADN según la reivindicación 1, en la que dicho ADN codifica una proteína enzimática que comprende la secuencia de aminoácidos de SEQ ID NO:2.
- 3.- La molécula de ADN según la reivindicación 1, en la que dicho ADN codifica una ácido L-galactónico deshidratasa de origen fúngico.
- 10 4.- La molécula de ADN según la reivindicación 1, en la que dicho gen comprende la secuencia de ácido nucleico de SEQ ID NO:1.
- 5.- La molécula de ADN según la reivindicación 1, que comprende una secuencia de ácido nucleico capaz de hibridarse con SEQ ID NO:1 bajo condiciones intermedias o rigurosas.
- 6.- Un vector que comprende la molécula de ADN según una cualquiera de las reivindicaciones 1 a 5.
- 7.- Un microorganismo hospedante que contiene el vector de la reivindicación 6.
- 15 8.- El microorganismo según la reivindicación 7, estando dicho microorganismo depositado con el n.º de registro DSM 17214.
- 9.- Una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, en la que dicha proteína enzimática tiene al menos 90% de identidad de secuencia con SEQ ID NO:2.
- 10.- La proteína enzimática según la reivindicación 9 que comprende la secuencia de aminoácidos de SEQ ID NO:2.
- 20 11.- La proteína enzimática según la reivindicación 9, en la que dicha proteína convierte también el ácido D-arabónico en ácido D-glicero-3-desoxi-pentulosónico.
- 12.- Un método para producir una proteína enzimática que convierte el el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, comprendiendo dicho método cultivar el microorganismo de la reivindicación 7 bajo condiciones que permiten la expresión de dicha proteína, y recuperar la proteína enzimática.
- 25 13.- Un método para convertir el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, que comprende poner en contacto el ácido L-galactónico con una proteína enzimática que tiene al menos 30% de identidad de secuencia con SEQ ID NO:2.
- 14.- Un método para producir un compuesto deseado a partir de biomasa que comprende un azúcar ácido o uno de sus derivados, comprendiendo dicho método las etapas de:
- 30 transformar un microorganismo hospedante con una molécula de ADN que comprende un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico;
fermentar dicha biomasa con el microorganismo transformado; y
recuperar el compuesto deseado producido.
- 35 15.- El método según la reivindicación 14, que comprende fermentar la pulpa de la remolacha azucarera u otro material que comprende pectina, y recuperar de la fermentación etanol, ácido láctico u otros compuestos derivados del ácido pirúvico.
- 16.- Un método para convertir el ácido D-arabónico en ácido D-glicero-3-desoxi-pentulosónico, que comprende poner en contacto el ácido D-arabónico con una proteína enzimática que tiene al menos 30% de identidad de secuencia con SEQ ID NO:2.
- 40 17.- El uso de una proteína enzimática según la reivindicación 9 para producir un compuesto deseado a partir de un material que comprende un azúcar ácido o uno de sus derivados.
- 18.- Una preparación enzimática que comprende la proteína enzimática de una cualquiera de las reivindicaciones 9 a 11.
- 45 19.- Una cepa de *Hypocrea jecorina* que está genéticamente modificada para inactivar el gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico.
- 20.- Un método para producir ácido L-galactónico proporcionando un microorganismo que tiene un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, inactivando

dicho gen, y utilizando el microorganismo genéticamente modificado obtenido para la producción de ácido L-galactónico.

5 21.- Un método para producir ácido D-arabónico proporcionando un microorganismo que tiene un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, inactivando dicho gen, y utilizando el microorganismo genéticamente modificado obtenido para la producción de ácido D-arabónico.

22.- Un método para modificar un microorganismo que tiene un gen que codifica una proteína enzimática que convierte el ácido L-galactónico en ácido L-treo-3-desoxi-hexulosónico, en el que dicho método comprende inactivar dicho gen.

10

Figura 1

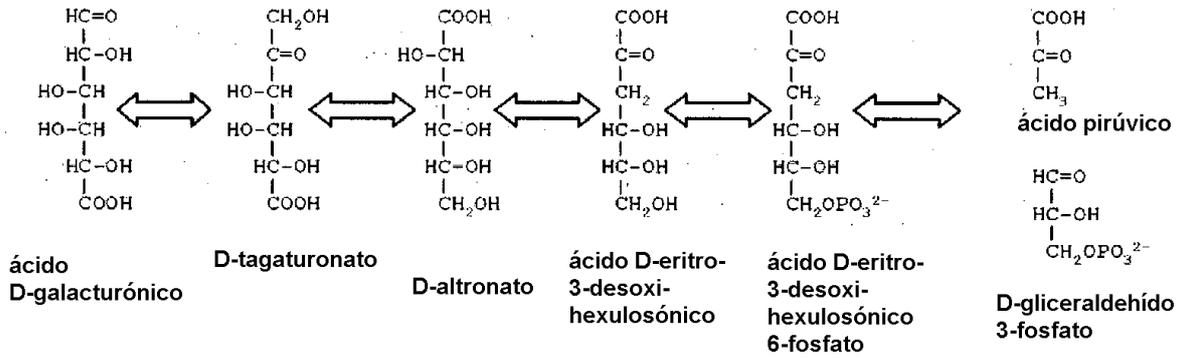


Figura 2

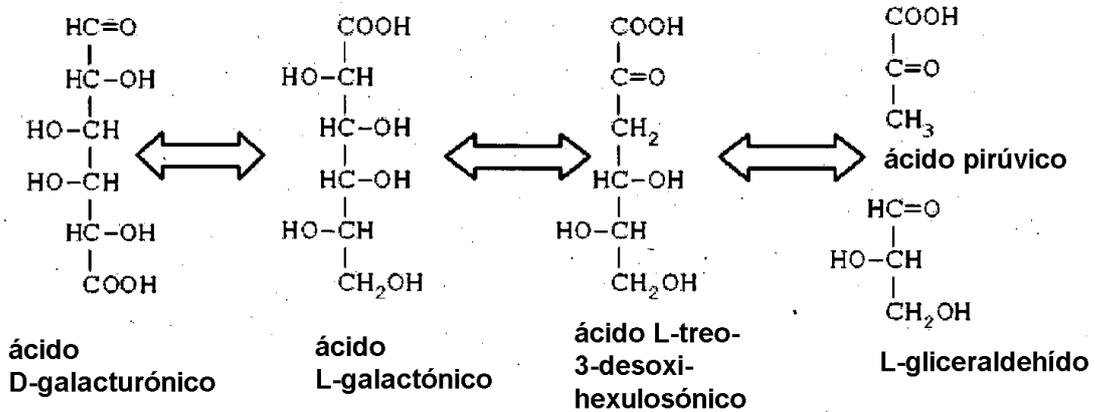


Figura 3

```

1      ATGTCTGAAGTCACCATCACAGGCTTCAGGAGCCGCGATGTGCGGTTCCCCACGTC CCTA      1
1      M S E V T I T G F R S R D V R F P T S L

61     GACAAGACGGGCTCGGATGCGATGAACGCTGCGGGCGACTATTCAGCGGCATACTGCATC      61
21     D K T G S D A M N A A G D Y S A A Y C I

121    CTCGAGACTGATTTCAGCGCACAGTGGTCATGGCATGgtcagccccctttgaagcactct      121
41     L E T D S A H S G H G M

157    gcgaagctgaagttccaagttgacagctttgctatagACATTTCACCATTGGACGCGGAAA      181
53                                     T F T I G R G N

180    CGACATCGTCTGCGCCGCCATCAACCACGTCGCGGACCGACTCAAGGGCAAGAACTGTC      241
61     D I V C A A I N H V A D R L K G K K L S

240    ATCACTAGTGGCCGACTGGGGCAAGACCTGGCGGTATCTGGTCAACGACAGCCAGCTGCG      301
81     S L V A D W G K T W R Y L V N D S Q L R

300    GTGGATTGGCCCCGAAAAGGGCGTCATCCATCTTGGCGCTCGGAGCCGTCGTCAACGCCGT      361
101    W I G P E K G V I H L A L G A V V N A V

360    CTGGGACCTGTGGGCAAAGACGCTCAACAAGCCGGTTTGGCGCATCGTTGCCGACATGAC      421
121    W D L W A K T L N K P V W R I V A D M T

420    GCCCGAGGAGTATGTCCGCTGCATCGACTTCCGCTACATTACCGACGCAATCACCCCCGA      481
141    P E E Y V R C I D F R Y I T D A I T P E

480    GGAAGCCGTGGCGATGCTGCGCGAGCAGGAGGCCGGCAAGGCCAAGCCGATCGAGGAGGC      541
161    E A V A M L R E Q E A G K A K R I E E A

540    TCTCCAGAACCAGCGGTGCCCTGCATACACAACAAGTGCCGGTTGGCTGGGATACGGAGA      601
181    L Q N R A V P A Y T T S A G W L G Y G E

600    GGACAAGATGAAGCAGCTCCTGAGAGAGACGCTGGCTGCCGATACAGACACTTCAAGGT      661
201    D K M K Q L L R E T L A A G Y R H F K V

660    CAAGGTTGGCGGCAGCGTCGAGGAGGCCGAAGGCCGCTCGGCATTGCTCGCGAAATTCT      721
221    K V G G S V E E D R R R L G I A R E I L

720    TGGTTTCGACAAGGGCAACGTTCTCATGGTCGATGCCAACCAGGTCTGGTCCGTTCCCGA      781
241    G F D K G N V L M V D A N Q V W S V P E

780    AGCGATCGACTACATGAAGCAGCTCAGCGAGTACAAGCCCTGGTTCATTGAGGAGCCCAC      841
261    A I D Y M K Q L S E Y K P W F I E E P T

840    CTCACCCGACGACATCATGGGCCACAAGGCCATTGCGGATGCCCTCAAGCCCTATGGCAT      901
281    S P D D I M G H K A I R D A L K P Y G I

900    CGGCGTCGCTACCGCGAGATGTGCCAGAACC GCGTCATGTTCAAGCAGCTGATCATGAC      961
301    G V A T G E M C Q N R V M F K Q L I M T

960    GGGCGCCATCGACATCTGCCAGATTGATGCCTGCCGCTCGGGCGGTCAACGAAGTGTCT      1021
321    G A I D I C Q I D A C R L G G V N E V L

1020   GGCCGTCCTGCTCATGGCCAAGAAGTACGGTGTGCCCATTTGTGCCGATTCCGGCGGCGT      1081
341    A V L L M A K K Y G V P I V P H S G G V

1080   GGGCCTTCCCGAGTACACCCAGCATCTGAGCACCATCGACTACGTGGTCGTCAGCGGCAA      1141
361    G L P E Y T Q H L S T I D Y V V V S G K
    
```

Figura 3 (continuación)

1140	GCTTCCGTCCTGGAGTTTGTAGACCACCTCCACGAGCACTTCTTGCATCCTTCAGTCAT	1201
381	L S V L E F V D H L H E H F L H P S V I	
1200	CAAGGACGGATACTACCAGACACCAACCGAGGCCGGCTACAGCGTTGAGATGAAGCCGGA	1261
401	K D G Y Y Q T P T E A G Y S V E M K P E	
1260	GAGCATGGACAAGTATGAGTATCCCGCAAGAAGGGCGTAAGTTGGTGGACGACCGACGA	1321
421	S M D K Y E Y P G K K G V S W W T T D E	
1320	CGCTCTGCCCATCTTGAACGGAGAGAAGATCTGA	1381
441	A L P I L N G E K I *	

Figura 4

