

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 396 151**

51 Int. Cl.:

**C12P 7/06** (2006.01)

**C12N 15/53** (2006.01)

**C12N 9/02** (2006.01)

**C12N 1/21** (2006.01)

**C12N 15/90** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.03.2007 E 07732122 (2)**

97 Fecha y número de publicación de la concesión europea: **19.09.2012 EP 2007897**

54 Título: **Potenciamiento de la producción de etanol microbiana**

30 Prioridad:

**24.03.2006 GB 0605890**

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

**19.02.2013**

73 Titular/es:

**BIOCONVERSION TECHNOLOGIES LIMITED  
(100.0%)  
263 FRIMLEY GREEN ROAD  
FRIMLEY GREEN CAMBERLEY SURREY, GB**

72 Inventor/es:

**JAVED, MUHAMMAD y  
BAGHAEI-YAZDI, NAMDAR**

74 Agente/Representante:

**UNGRÍA LÓPEZ, Javier**

**ES 2 396 151 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

Potenciamiento de la producción de etanol microbiana

5 **Campo de la invención**

La presente invención se refiere a procedimientos de fermentación y microorganismos para su uso en ellos y en particular al potenciamiento de la producción de etanol microbiana. Más específicamente, la invención se refiere a la producción de etanol potenciada por bacterias termófilas tales como bacilos de azúcares mixtos derivados de la hidrólisis de biomasa. En particular, la invención prevé una ruta novedosa para la producción de etanol clonando un gen que codifica una enzima formiato deshidrogenasa ligada a NAD en una bacteria termófila del género *Bacillus* que posee un gen funcional que codifica un complejo de enzimas piruvato-formiato liasa pero que carece de actividad de lactato deshidrogenasa.

15 **Antecedentes a la invención**

El bioetanol se prepara actualmente a partir de glucosa, maltosa o sacarosa derivada de almidón de cereal, caña de azúcar o remolacha azucarera, que todos tienen valor alimenticio. Las celulosas y hemicelulosas forman una parte importante de los subproductos agrícolas y podrían ser, en principio, una fuente importante de bioetanol renovable de bajo coste. Sin embargo, es difícil y caro derivar azúcares fermentables de la celulosa. A diferencia, las hemicelulosas son casi tan abundantes como la celulosa y son fáciles de hidrolizar, pero dan una mezcla de azúcares principalmente pentosa que no pueden fermentar levaduras.

Por este motivo, Hartley (véase la publicación internacional número WO 88/09379) propuso la producción de etanol por mutantes de un bacilo termófilo, que fermenta muy rápidamente todos los azúcares derivados de biomasa, a temperaturas de hasta 70°C. Sin embargo, sólo se logra un alto rendimiento de etanol por células estresadas y moribundas.

Muchos microorganismos contienen una ruta de piruvato-formiato liasa (PFL) que convierte piruvato en acetil CoA y formiato (Figura 1A). Los microorganismos fermentativos de heterolactato son una de tal clase. Estos microorganismos convierten primero los azúcares de entrada en piruvato (generalmente por la ruta de EMP de glucólisis), que luego pueden tomar muchas rutas para producir lactato, formiato, acetato, etanol y CO<sub>2</sub>, en diversas proporciones, dependiendo de las condiciones de crecimiento.

En células completamente aerobias, el piruvato es normalmente metabolizado en H<sub>2</sub>O y CO<sub>2</sub> por la ruta de piruvato deshidrogenasa (PDH), ciclo de ácido tri-carboxílico y la cadena de transporte de electrones. Sin embargo, en muchos de estos organismos, particularmente bacilos termófilos, la captación de azúcares y la glucólisis parecen estar sin regular y el lactato es un producto dominante a altas concentraciones de azúcar, incluso bajo condiciones aerobias. Esto sugiere que el flujo de PDH se ha saturado, y que el piruvato en exceso es desviado a una ruta de lactato deshidrogenasa de exceso. Ésta no se usa para el crecimiento, pero produce calor que hace que se eleve la temperatura ambiente y destruya competidores mesófilos, como puede apreciarse cuando césped fresco se pone sobre una pila de compost.

Si el gen *ldh* (que codifica lactato deshidrogenasa) se inactiva como se describe, por ejemplo, en el documento WO 02/29030, se detiene la producción de lactato y el piruvato en exceso es principalmente desviado a la ruta de PFL ligada al crecimiento (Figura 1A). Sin embargo, a concentraciones de azúcar muy altas y/o a pH ácido, el flujo de la ruta de PFL disminuye y entonces el piruvato en exceso se desborda a una ruta de PDH anaerobia, que sólo da etanol y CO<sub>2</sub> (Figura 1B). Por tanto, las condiciones preferidas para obtener altos rendimientos de etanol son aquellas que reducen el flujo por la ruta de PFL y aumentan el flujo por la ruta de PDH (Hartley, B.S. y Shama, G. Proc. Roy. Soc. Lond. 321, 555-568 (1987)). Desafortunadamente, bajo tales condiciones, las células experimentan tensión metabólica, con producción de ATP reducida, y un posible desequilibrio en las relaciones de NAD/NADH y CoA/acetil CoA (Figura 1C).

Se han propuesto diversos protocolos de fermentación para intentar evitar o minimizar este problema tal como el de Hartley, B.S. como se trata anteriormente (véase la publicación internacional número WO 88/09379).

Hay dos clases de formiato deshidrogenasa. Una (codificada por el gen *fdhF*) convierte formiato en CO<sub>2</sub> + H<sub>2</sub> y es típica de enterobacterias tales como *E. coli*. La otra (codificada por el gen *fdh1*) convierte formiato + NAD en CO<sub>2</sub> + NADH<sub>2</sub> y está presente en muchos anaerobios facultativos. Berrios-Rivera y col. (Metabolic Engineering 4, 217-219 (2002)) sustituyeron el gen *fdhF* en *E. coli* con un gen *fdh1* de levadura y encontraron que los productos anaerobios reducidos tales como etanol, lactato y succinato aumentaron con respecto a productos oxidados tales como acetato. Basándose en esta observación, San, K-Y. Berrios-Rivers, S.J. y Bennett, G.N. (véase la publicación internacional número WO 2003/040690) propusieron la introducción de un gen formiato deshidrogenasa ligada a NAD como procedimiento general para aumentar la potencia reductora en células que participaban en un amplio intervalo de bio-transformaciones. Posteriormente, San, K-Y. Bennett, G.N. y Sanchez, A. (solicitud de patente de EE.UU. 2005/0042736 A1) propusieron una aplicación específica de este concepto para la producción de succinato. Estos

estudios se llevaron a cabo en *E. coli*, un ejemplo de un mesófilo en el que la captación de azúcares está regulada. El fin de estos experimentos fue aumentar los niveles de NADH intracelular de manera que se proporcionara potencia reductora potenciada para las diversas bio-transformaciones.

- 5 La formiato deshidrogenasa de levadura recomendada por Sen. y col. (2004) es inactiva a 60°C, que es la temperatura de crecimiento mínima para las bacterias termófilas potencialmente de uso en la producción de bioetanol. Las formiato deshidrogenasas más termoestables descritas hasta la fecha es la enzima 101 de *Pseudomonas sp.* (A. Rojkova, A. Galkin, L. Kulakova, A. Serov, P. Savitsky, V. Fedorchuk, V. Tishkov FEBS Letters, volumen 445, número 1, páginas 183-188, 1999).

10

### **Resumen de la invención**

La presente invención intenta resolver los problemas de producir altos rendimientos de etanol a partir de biomasa. En particular, en el presente documento se ha descrito por primera vez una ruta metabólica novedosa que permite que microorganismos termófilos, especialmente bacterias tales como *Bacillus*, produzcan rendimientos de etanol máximos. La invención se define en las reivindicaciones.

15

La invención se basa en bacterias termófilas del género *Bacillus* que carecen de actividad de lactato deshidrogenasa y, por tanto, requieren una ruta alternativa para la re-oxidación de la NADH en exceso producida por glucólisis. Ésta se proporciona por introducción en las bacterias termófilas del género *Bacillus* de un gen que codifica una formiato deshidrogenasa ligada a NAD, tal como un gen *fdh1*. En microorganismos termófilos, y a diferencia de mesófilos tales como *E. coli*, la captación de azúcares están sin regular y esto conduce a la acumulación de NADH en presencia de altos niveles de azúcares. Esto conduce eventualmente a un colapso metabólico y a la llamada “muerte rédox” como se muestra esquemáticamente en la Figura 1C. La incorporación en el microorganismo de un gen que codifica una formiato deshidrogenasa ligada a NAD ayuda a prevenir la muerte celular a altas concentraciones de azúcar conduciendo a una disminución en niveles de NADH y a un aumento en niveles de NAD. Esto es en parte restaurando un flujo por la ruta de piruvato deshidrogenasa (PDH), pero, y lo que es más importante, la inclusión de un gen que codifica una formiato deshidrogenasa ligada a NAD crea una novedosa ruta de piruvato formiato liasa (PFL)-formiato deshidrogenasa ligada a NAD (FDH) para la producción de etanol. La Figura 1D muestra las posibilidades de esta ruta de PFL-FDH para restaurar el equilibrio rédox convirtiendo todo el piruvato producido por la rápida glucólisis en presencia de altos niveles de azúcar en etanol y CO<sub>2</sub>, especialmente bajo condiciones de pH neutro. Y, lo que es más importante, la ruta opera en condiciones que son opcionales para el crecimiento celular, conduciendo a la rápida producción de etanol y alto rendimiento, ya que la ruta de PFL es la principal ruta anaerobia ligada al crecimiento en microorganismos termófilos.

20

25

30

35

Por consiguiente, en un primer aspecto, la invención proporciona una bacteria termófila del género *Bacillus* que carece de actividad de lactato deshidrogenasa (*ldh*), caracterizada porque la bacteria termófila contiene un gen que codifica una formiato deshidrogenasa ligada a NAD (*fdh*).

40

45

En una realización, el microorganismo carece de actividad de lactato deshidrogenasa en virtud de una delección de gen apropiada u otra mutación que elimina la actividad de lactato deshidrogenasa. Por tanto, preferentemente el gen *ldh* está delecionado o de otro modo se ha convertido en no funcional. Los procedimientos de inactivación y delección de genes son muy conocidos en la técnica y ejemplos preferidos se describen en detalle en el presente documento. Además, cepas conocidas de bacterias que carecen de actividad de lactato deshidrogenasa (tales como TN-T9 depositada bajo el número de acceso NCIMB 41075 y TN-TK depositada bajo el número de acceso NCIMB 41115) pueden ser adecuadas para su uso en la presente invención.

50

La bacteria termófila de la invención normalmente contiene una ruta de piruvato formiato liasa activa. En particular, el microorganismo comprende preferentemente un gen que codifica una piruvato formiato liasa tal como el gen *pf1*. Los microorganismos de la invención también contienen normalmente una ruta de piruvato deshidrogenasa (PDH) activa.

55

60

En una realización preferida, el gen que codifica una formiato deshidrogenasa ligada a NAD está integrado en el genoma de la bacteria termófila. Sin embargo, también es posible lograr la expresión estable sin la integración, por ejemplo, por introducción de un plásmido adecuado. Un procedimiento preferido de integración es por recombinación. El gen que codifica una formiato deshidrogenasa ligada a NAD puede estar operativamente ligado a cualquier elemento regulador adecuado para dirigir la expresión de la formiato deshidrogenasa ligada a NAD. Por “operativamente ligado” se indica que existe un enlace funcional entre el elemento regulador y el gen que codifica una formiato deshidrogenasa ligada a NAD. Por ejemplo, el gen que codifica una formiato deshidrogenasa ligada a NAD puede ligarse a un promotor adecuado que puede ser, por ejemplo, un promotor constitutivo o inducible. Se entiende que “promotor” en el presente documento incluye una región de ADN que participa en la unión de ARN polimerasa para iniciar la transcripción.

65

Normalmente, el promotor es un promotor procarionta y, por tanto, incluye las secuencias -10 y -35 apropiadas, las secuencias consenso del cual están bien definidas en la materia. El gen también puede estar operativamente ligado a otras secuencias reguladoras apropiadas tales como, por ejemplo, terminadores. “Terminador” se define como una

- 5 secuencia de nucleótidos que produce ARN polimerasa para terminar la transcripción. En una realización, el gen que codifica una formiato deshidrogenasa ligada a NAD se expresa a partir de su propio promotor. En una realización alternativa, el gen que codifica una formiato deshidrogenasa ligada a NAD se expresa a partir de un promotor de la bacteria termófila (debido a la integración en una localización apropiada en el genoma). También pueden preverse construcciones en las que la expresión del gen que codifica una formiato deshidrogenasa ligada a NAD es accionada por un promotor extraño. Esto puede hacerse para lograr, por ejemplo, niveles de expresión máximos o expresión inducible. Como un ejemplo, promotores de fago tales como T7 pueden utilizarse conjuntamente con una polimerasa de fago adecuada (que puede proporcionarse en una construcción de ADN separada o la misma).
- 10 En una realización particularmente preferida, el gen que codifica una formiato deshidrogenasa ligada a NAD está operativamente ligado a las regiones reguladoras apropiadas de un gen que codifica una lactato deshidrogenasa, en particular las regiones reguladoras en la dirección 5'. La región reguladora comprende preferentemente el promotor de un gen que codifica una lactato deshidrogenasa. Puede entenderse que el promotor incluye como unidad funcional mínima las secuencias -10 y -35 apropiadas. Por tanto, según una realización preferida de la invención, el
- 15 gen que codifica una formiato deshidrogenasa ligada a NAD se inserta en el gen lactato deshidrogenasa de la bacteria termófila, inactivándose así la actividad de lactato deshidrogenasa de la bacteria termófila. Esta realización es particularmente preferida, ya que ambas modificaciones requeridas para producir una bacteria termófila de la invención se producen en la misma etapa. Construcciones adecuadas para lograr esto se describen en detalle en el presente documento.
- 20 La producción de etanol por bacterias termófilas es ventajosa, ya que puede llevarse a cabo a altas temperaturas. Mientras que los microorganismos termófilos tienen menor tolerancia al etanol que las levaduras, el etanol puede eliminarse continuamente y convenientemente de la fermentación a altas temperatura por membrana y/o evaporación a vacío suave. En condiciones anaerobias óptimas de crecimiento, la cepa LLD-R de *Bacillus* crece
- 25 muy rápidamente a 70°C casi exclusivamente por la ruta de PFL (Hartley y Shama, 1982). Puede imaginarse que el crecimiento por la novedosa ruta de PFL-FDH sería igualmente vigoroso, pero la temperatura de crecimiento máxima estaría limitada por la termoestabilidad de la formiato deshidrogenasa ligada a NAD introducida en el microorganismo termófilo. La bacteria termófila de la invención puede incorporar un gen que codifica una formiato deshidrogenasa ligada a NAD termostable y/o un gen cuya secuencia de nucleótidos ha sido optimizada por
- 30 codones para facilitar la expresión por un microorganismo termófilo. La producción de una formiato deshidrogenasa ligada a NAD termostable se describe en detalle en el presente documento. En una realización específica, el gen que codifica una formiato deshidrogenasa ligada a NAD comprende o consiste en la secuencia de nucleótidos expuesta como SEC ID NO: 1. En otra realización, la bacteria termófila de la invención puede incorporar un gen que está optimizado por codones para la expresión en *Bacillus*, que codifica una formiato deshidrogenasa ligada a NAD
- 35 termostable que comprende o que consiste en la secuencia de nucleótidos expuesta como SEC ID N°: 2. Esta secuencia incluye, además de la secuencia de deshidrogenasa ligada a NAD termostable básica, regiones promotoras y terminadoras y también sitios Xba1 para facilitar la clonación del gen en una construcción de ADN adecuada.
- 40 El gen que codifica una formiato deshidrogenasa ligada a NAD puede ser el gen *fdh1*. El gen *fdh1* puede derivarse de cualquier fuente adecuada y está preferentemente optimizado por codones para la expresión en el microorganismo termófilo relevante.
- 45 La bacteria termófila de la invención puede producirse por transformación con cualquiera de las construcciones de ADN de la invención como se describe en más detalle en el presente documento. Por consiguiente, la discusión aquí proporcionada se aplica, cambiando lo que se deba cambiar, a esta realización de la invención.
- 50 El microorganismo termófilo de la invención es una bacteria termófila del género *Bacillus* y más preferentemente *Bacillus stearothermophilus*. El microorganismo termófilo de la invención puede derivarse de la cepa LLD-R o LLD-15 (de *Bacillus stearothermophilus*) conocida. La bacteria termófila descrita en este documento puede ser *Geobacillus thermoglucosidasius*.
- 55 Los procesos de fermentación facilitados por la presente invención utilizan preferentemente una formiato deshidrogenasa ligada a NAD sintética diseñada para expresar una secuencia de aminoácidos termostable debido al uso de las preferencias de codón del microorganismo termófilo apropiado tal como la cepa LLD-R de *Bacillus*. El gen sintético contiene preferentemente sitios de restricción manipulados para ayudar en la inserción en el gen lactato deshidrogenasa. De este modo, la delección de la ruta de LDH y la creación de la ruta de PFL-FDH se consiguen en una única operación. Por consiguiente, en el presente documento se describe una formiato deshidrogenasa ligada a NAD termostable. Preferentemente, la formiato deshidrogenasa ligada a NAD termostable sigue siendo funcional a
- 60 o por debajo de una temperatura de 60°C. Preferentemente, la enzima termostable está codificada por una secuencia de nucleótidos que ha sido optimizada por codones para la expresión en un microorganismo termófilo. La formiato deshidrogenasa puede comprender, consistir esencialmente en o consistir en la secuencia de aminoácidos expuesta como SEC ID N°: 3.
- 65 Se ha diseñado una formiato deshidrogenasa ligada a NAD termostable específica basándose en la secuencia de aminoácidos de la formiato deshidrogenasa de 101 de *Pseudomonas sp* (SEC ID N°: 3) y mediante el uso de

codones optimizados para *Geobacillus thermoglucosidasius* como se trata en más detalle en la descripción detallada más adelante. El experto apreciará que derivados de esta secuencia básica retendrán la funcionalidad. Por ejemplo, sustituciones conservativas y semi-conservativas pueden producir formiato deshidrogenasas ligadas a NAD termoeestables que retienen actividad catalítica y termoeestabilidad eficaces de forma que son útiles en la producción de etanol usando microorganismos termófilos. Similarmente, deleciones menores y/o adiciones de aminoácidos pueden producir derivados que retienen la funcionalidad apropiada.

En un segundo aspecto, la invención se refiere a un gen sintético que codifica una formiato deshidrogenasa ligada a NAD termoeestable. El gen comprende o consiste en la secuencia de nucleótidos expuesta como SEC ID N°: 1. Esta secuencia representa una secuencia del gen *fdh* novedosa en la que los codones están optimizados para la producción de una formiato deshidrogenasa ligada a NAD termoeestable. En una realización más específica, el gen que codifica una formiato deshidrogenasa ligada a NAD termoeestable comprende o consiste en la secuencia de nucleótidos expuesta como SEC ID N°: 2. Esta secuencia incorpora la región codificante para la formiato deshidrogenasa ligada a NAD termoeestable junto con un promotor de *Bacillus* adecuado y terminador independiente de rho. La secuencia también incorpora sitios de restricción adecuados para ayudar en la clonación, en particular sitios *Xba*1. El experto apreciará fácilmente que pueden hacerse modificaciones menores a la secuencia de nucleótidos sin alterar la funcionalidad o termoeestabilidad de la enzima resultante, por ejemplo, sustituyendo codones optimizados con otros codones que se prefieren en los sistemas de traducción del microorganismo termófilo apropiado.

En el presente documento también se describe una construcción de ADN que contiene un gen que codifica una formiato deshidrogenasa ligada a NAD, en particular una formiato deshidrogenasa ligada a NAD termoeestable, en la que el gen está flanqueado por sitios de restricción para facilitar la clonación del gen en una construcción de ADN adecuada, tal como un vector de expresión o plásmido.

En un aspecto relacionado, en el presente documento se describe una construcción de ADN que comprende una secuencia reguladora operativamente ligada a un gen que codifica una formiato deshidrogenasa ligada a NAD termoeestable. Por tanto, esta construcción de ADN facilita la transformación de bacterias termófilas del género *Bacillus*, en particular aquellas que carecen de actividad de lactato deshidrogenasa, con el fin de producir microorganismos termófilos que pueden fermentar eficientemente dando rendimientos de etanol máximos. Como se ha mencionado anteriormente, el término “operativamente ligado” como se usa en el presente documento se refiere a un enlace funcional entre la secuencia reguladora y el gen que codifica la formiato deshidrogenasa ligada a NAD, de forma que la secuencia reguladora puede influir en la expresión génica. Por ejemplo, una secuencia reguladora preferida es un promotor. Como se ha mencionado anteriormente, el gen que codifica una formiato deshidrogenasa ligada a NAD puede comprender preferentemente, consistir esencialmente en o consistir en la secuencia de nucleótidos expuesta como SEC ID N°: 1. Una secuencia reguladora preferida es un promotor, aunque la construcción de ADN puede incorporar adicionalmente secuencias terminadoras. El promotor puede comprender la secuencia de nucleótidos expuesta como SEC ID N°: 4. Otros promotores, como se trata anteriormente, pueden utilizarse para altos niveles y/o expresión inducible.

En el presente documento se describe una construcción de ADN que comprende un gen que codifica una formiato deshidrogenasa ligada a NAD y una secuencia de inserción, en la que la secuencia de inserción facilita la integración del gen que codifica una formiato deshidrogenasa ligada a NAD en el genoma de una bacteria termófila transformada con la construcción de ADN. Por “secuencia de inserción” se indica un elemento de ADN transponible que puede integrarse en el genoma del microorganismo termófilo apropiado. Las secuencias de inserción también pueden denominarse elementos de secuencia de inserción (EI) y pueden producirse naturalmente. La secuencia de inserción puede derivarse de la cepa LLD-R o LLD-15 de *Bacillus stearothermophilus*.

La secuencia de inserción puede comprender, consistir esencialmente en o consistir en la secuencia de nucleótidos expuesta como SEC ID N°: 5 (Figura 3). La secuencia de inserción preferida puede generarse por amplificación usando cebadores que comprenden, consisten esencialmente en o que consiste en la secuencia de nucleótidos expuesta como SEC ID N°: 6 y 7. En este caso, el ADN genómico de la cepa LLD-15 de *Bacillus stearothermophilus* conocida puede usarse como molde. Una construcción de ADN particularmente preferida es el plásmido pUB-ISF1 (como se describe en la sección experimental más adelante y en la Figura 5).

En el presente documento se describe una construcción de ADN que comprende un gen (*fdh*) que codifica una formiato deshidrogenasa ligada a NAD operativamente ligada a regiones reguladoras apropiadas de un gen que codifica una lactato deshidrogenasa, en particular las regiones reguladoras en la dirección 5'. Las regiones reguladoras preferentemente comprenden el promotor de un gen que codifica una lactato deshidrogenasa (*ldh*). Puede entenderse que el promotor incluye como unidad funcional mínima las secuencias -10 y -35 apropiadas para permitir la unión de ARN polimerasa eficaz. El promotor del gen lactato deshidrogenasa es adecuado para accionar altos niveles de expresión en bacterias termófilas tales como bacilos y también puede usarse ventajosamente como parte de la estrategia de clonación para lograr tanto la deleción de actividad de lactato deshidrogenasa como la introducción de actividad de formiato deshidrogenasa ligada a NAD en la misma etapa. Por tanto, también puede definirse que la construcción de ADN comprende un gen que codifica una formiato deshidrogenasa ligada a NAD operativamente ligada a una molécula de ácido nucleico que comprende el promotor de un gen que codifica una

lactato deshidrogenasa (ldh).

La construcción de ADN también contiene preferentemente parte de la secuencia codificante del gen lactato deshidrogenasa huésped en la dirección 3' del gen que codifica una formiato deshidrogenasa ligada a NAD. Esto facilita la integración del gen en un microorganismo transformado con la construcción de ADN. Por "al menos parte" se indica una parte del gen de longitud suficiente para permitir la integración del gen en el genoma de un microorganismo que contiene el gen lactato deshidrogenasa por recombinación (preferentemente por cruce doble). La parte de la secuencia codificante incorpora preferentemente el extremo del gen lactato deshidrogenasa. Al menos 100, 200, 300, 400, 500, 600, 700 ó 750 nucleótidos del gen lactato deshidrogenasa pueden incorporarse en la dirección 3' del gen que codifica una formiato deshidrogenasa ligada a NAD. Por tanto, la construcción de ADN puede comprender un gen que codifica una formiato deshidrogenasa ligada a NAD en la que el gen que codifica una formiato deshidrogenasa ligada a NAD está flanqueado por la secuencia de nucleótidos de un gen que codifica una lactato deshidrogenasa (derivado del microorganismo termófilo de interés). Las secuencias flanqueantes son de longitud suficiente para permitir la integración del gen que codifica una formiato deshidrogenasa ligada a NAD en el gen huésped que codifica una lactato deshidrogenasa para de esta forma introducir actividad de formiato deshidrogenasa ligada a NAD e inactivar la actividad de lactato deshidrogenasa en una única etapa de clonación. Preferentemente, el gen que codifica una formiato deshidrogenasa ligada a NAD está flanqueado en la dirección 5' por al menos la región promotora del gen que codifica una lactato deshidrogenasa, de manera que tras la integración por recombinación el gen que codifica una formiato deshidrogenasa ligada a NAD es operativamente ligado al promotor. La porción en la dirección 3' del gen lactato deshidrogenasa puede ser una obtenible por amplificación del gen ldh usando cebadores que comprenden, consisten esencialmente en o que consisten en la secuencia de nucleótidos expuesta como SEC ID N<sup>o</sup>: 8 y 9, usando la cepa LLD-R como molde. La región flanqueante en la dirección 5' que incorpora preferentemente el promotor ldh comprende preferentemente al menos 100, 200, 300, 400, 500, 600, 700 ó 750 nucleótidos de las regiones en la dirección 5' de ldh apropiadas para maximizar la eficiencia de integración por recombinación con el genoma del huésped. Esta región en la dirección 5' puede ser dependiente tras el contexto de la secuencia del gen ldh en el microorganismo termófilo específico de interés, como sería fácilmente determinado por un experto habitual. Por tanto, el experto con conocimiento de la secuencia del gen ldh determinaría fácilmente regiones flanqueantes apropiadas para permitir la integración por recombinación. Por ejemplo, pueden estudiarse secuencias genómicas publicadas, reacciones de secuenciación llevadas a cabo o regiones flanqueantes amplificadas por PCR usando cebadores derivados de la secuencia del gen. Por tanto, el gen fdh se interpone entre dos secuencias de nucleótidos derivadas del gen ldh de forma que el gen fdh sustituye, en el marco, al menos parte del gen ldh.

Por tanto, la construcción de ADN generalmente comprende un casete de integración del gen en el que el gen de interés (gen que codifica una formiato deshidrogenasa ligada a NAD) se inserta dentro de la secuencia codificante (ORF) de un gen que va a inactivarse durante la integración (gen lactato deshidrogenasa en este caso). Tras la integración mediante recombinación, la expresión del gen de interés está en efecto bajo el control del gen inactivado. Una construcción tal puede ser de aplicabilidad general en la circunstancia en la que un gen necesite inactivarse a favor de la expresión de un gen heterólogo. En una realización preferida, el gen que codifica una formiato deshidrogenasa ligada a NAD codifica una formiato deshidrogenasa ligada a NAD termoestable. Por tato, la discusión de las formiato deshidrogenasas ligadas a NAD termoestables proporcionadas en el presente documento se aplica, cambiando lo que se deba cambiar, a este aspecto de la invención. En particular, en una realización, el gen que codifica una formiato deshidrogenasa ligada a NAD termoestable comprende o consiste en la secuencia de nucleótidos expuesta como SEC ID N<sup>o</sup>: 1 ó 2. Una construcción de ADN particularmente preferida es el plásmido pUCK-LF1 (como se describe en la sección experimental más adelante y en la Figura 8).

Para todas las construcciones de ADN de la invención, una forma preferida es un vector de expresión. Por tanto, las construcciones de ADN permiten la expresión fidedigna del gen que codifica una formiato deshidrogenasa ligada a NAD termoestable en una bacteria termófila del género *Bacillus* transformada con la construcción.

La construcción de ADN puede ser un plásmido. Preferentemente, la construcción de ADN sólo puede replicarse en el microorganismo termófilo huésped mediante recombinación con el genoma del microorganismo termófilo huésped.

Las construcciones de ADN descritas en el presente documento también incorporan preferentemente un gen indicador adecuado como un indicador de transformación satisfactoria. El gen indicador puede ser un gen de resistencia a antibióticos tal como un gen de resistencia a kanamicina o ampicilina. Pueden utilizarse otros indicadores, tales como la proteína fluorescente verde (GFP) y beta-galactosidasa (lacZ), según convenga. Las construcciones de ADN pueden incorporar múltiples genes indicadores, según convenga. La pérdida de la función indicadora es, en generaciones posteriores, indicativa de la integración del gen que codifica una formiato deshidrogenasa ligada a NAD termoestable, junto con regiones flanqueantes apropiadas.

En todavía otro aspecto, la invención se refiere a una bacteria termófila del género *Bacillus* que comprende una construcción de ADN de la invención. Microorganismos receptores preferidos son microorganismo fermentativo de heterolactato. En particular, la invención se refiere preferentemente a *Bacillus stearothermophilus*. La bacteria puede derivarse, por ejemplo, de la cepa LLD-R o LLD-15. El microorganismo termófilo puede ser *Geobacillus thermoglucosidasius*.

En el presente documento se desvela adicionalmente el uso de una bacteria termófila del género *Bacillus* de la invención en fermentación, y en particular para la producción de etanol.

5 Similarmente, la invención se refiere a un proceso de fermentación para la producción de etanol que comprende suministrar a una bacteria termófila del género *Bacillus* de la invención azúcares. Los microorganismos construidos según la presente invención son particularmente adecuados para el alto rendimiento y productividad volumétrica de etanol bajo condiciones de crecimiento óptimas. Por consiguiente, cualquier microorganismo de la invención puede usarse en cualquier configuración de fermentador tal como procesos de fermentación por lotes, por lotes alimentados o continuos. El proceso de fermentación puede ser un procedimiento por lotes alimentados.

15 Uno de los principales beneficios del uso de microorganismos tales como bacterias termófilas para producir bioetanol es que, a diferencia de las levaduras, pueden fermentar una amplia gama de azúcares derivados de productos de residuos agrícolas tales como hemicelulosas. Los azúcares usados en los procedimientos de fermentación de la invención pueden derivarse de biomasa. La fermentación puede ser de azúcares mixtos. Los azúcares mixtos pueden incluir azúcares de pentosa, preferentemente una mayoría de azúcares de pentosa.

20 El proceso de fermentación puede mantenerse en equilibrio rédox. Esto es particularmente crítico con termófilos ya que, a diferencia de los mesófilos, la captación de azúcares parece estar sin regular en estos microorganismos. Preferentemente, esto se logra mediante el uso de sensores de retroalimentación.

25 Aunque las bacterias termófilas tienen baja tolerancia al etanol, esto puede vencerse convenientemente en los procesos de fermentación de la invención por eliminación regular o continua de etanol. Esto garantiza que la concentración de etanol en la fermentación se mantenga por debajo de la tolerancia del etanol de las bacterias termófilas de la invención. El etanol puede eliminarse continuamente y convenientemente de la fermentación a alta temperatura mediante evaporación o destilación tal como, por ejemplo, evaporación en membrana y/o a vacío suave.

30 La invención se describirá ahora con referencia a la siguiente descripción no limitante y figuras.

### **Breve descripción de las figuras**

35 La Figura 1 muestra el efecto de diversas condiciones sobre las rutas metabólicas en un microorganismo termófilo en el que ha sido inactivada la ruta de lactato deshidrogenasa. La sombra y el espesor de las flechas indican la dominancia relativa de las rutas metabólicas respectivas.

- 40 A. Rutas metabólicas activas a pH neutro y en presencia de bajos azúcares. Aquí domina la ruta de piruvato formiato liasa.
- B. Rutas metabólicas activas a bajo pH y en presencia de bajos azúcares. Aquí domina una ruta de piruvato deshidrogenasa aerobia.
- C. Rutas metabólicas activas a bajo pH y en presencia de altos azúcares. Aquí, las células experimentan estrés metabólico y rompen con el equilibrio rédox conduciendo a la llamada "muerte rédox".
- 45 D. Rutas metabólicas activas en microorganismos termófilos de la presente invención, a pH neutro y en presencia de altos azúcares. Aquí, la ruta de PFL-FDH novedosa es dominante y los únicos productos anaerobios son etanol y CO<sub>2</sub>.

50 La Figura 2 expone la secuencia de nucleótidos del gen *fdh* sintético producido por optimización de codones para maximizar la termoestabilidad. La secuencia de ADN del marco de lectura abierto de *fdh* está flanqueada por regiones promotoras y terminadoras (cursiva). Los recuadros de -35 y -10 del promotor están subrayados. Para clonar la construcción en un vector adecuado, los sitios de *Xba*1 se introdujeron en ambos lados de la secuencia.

55 La Figura 3 muestra la secuencia de nucleótidos de la secuencia de inserción (SI) de la cepa LLD-15 de *B. stearrowthermophilus*. Los extremos de repeticiones invertidas de 9 pb se muestran en negrita.

60 La Figura 4 es una representación esquemática del plásmido derivado de pCR-Blunt pCR-F1. El plásmido incluye un gen *fdh1* optimizado por codones bajo el control del promotor *ldh*, clonado en pCR-Blunt en el único sitio de *Xba*1.

La Figura 5 es una representación esquemática del plásmido derivado de pUB110 pUB-ISF1. El plásmido incluye un gen *fdh1* optimizado por codones bajo el control del promotor *ldh*, derivado del plásmido pCR-F1, y también una secuencia de inserción (SI) derivada de la cepa LLD-15 de *Bacillus* conocida.

65 La Figura 6 es una representación esquemática del plásmido derivado de pUC18 llamado pUCK. El plásmido incluye un gen de resistencia a kanamicina clonado a partir del plásmido pUB110 en el único sitio

de restricción *Zra1* en pUC18.

La Figura 7 es una representación esquemática del derivado de pUCK pUCK-LC. El plásmido lleva un gen *ldh* con una delección de 363 pb en el centro del ORF.

La Figura 8 es una representación esquemática del derivado de pUCK pUCK-ldhB. El plásmido contiene 750 pb del gen *ldh*, que incluye la región en la dirección 3' del gen.

La Figura 9 es una representación esquemática del plásmido pUCK-LF1. Este plásmido es un derivado de pUCK que incorpora un casete integrante del gen que contiene el gen *fdh* bajo el control del promotor de *ldh*.

## **Descripción de la invención**

### **Materiales**

#### **Medios y tampones**

**Medio LB:** Triptona 10 g; extracto de levadura 5 g; NaCl 10 g; agua desionizada hasta 1 l

Ajustado a pH 7 y esterilizado en autoclave

Para el medio de la placa, 20 g/l de agar se añadieron al medio antes de la esterilización en autoclave, se enfriaron a 55°C y se vertieron en placas de Petri estériles (aprox. 20 ml/placa).

Para placas de LB-amp, disolución de ampicilina esterilizada por filtración a concentración final de 50 µg/ml se añadió antes de verter las placas de Petri.

**Medio SOC:** Triptona 2,0 g; extracto de levadura 0,5 g; NaCl 0,05 g; MgCl<sub>2</sub>·6H<sub>2</sub>O 0,204 g;

MgSO<sub>4</sub>·7H<sub>2</sub>O 0,247 g; glucosa 0,36 g; H<sub>2</sub>O desionizada hasta 100 ml. Disuelto, ajustado a pH 7,0 y esterilizado en autoclave.

**Medio TGP:** Triptona 17 g; peptona de soja 3 g; K<sub>2</sub>HPO<sub>4</sub> 2,5 g; NaCl 5 g; piruvato de Na 4 g; glicerol 4 ml; agua desionizada hasta 1 l. Ajustado a pH 7 y esterilizado en autoclave.

Para el medio de la placa, 20 g/l de agar se añadieron al medio antes de la esterilización en autoclave, se enfriaron a 55°C y se vertieron en placas de Petri estériles (aprox. 25 ml/placa).

Para placas de TGP-kan, la disolución de kanamicina esterilizada por filtración a concentración final de 10 µg/ml se añadió antes de verter las placas de Petri.

**Tampón TH:** Trehalosa 272 mM; HEPES (pH 7,5 con KOH) 8 mM; H<sub>2</sub>O doblemente destilada hasta 1 l

#### **Cepas microbianas**

Células químicamente competentes DH5-alfa de *E. coli* se compraron de Invitrogen (Cat. 18265-017).

Colección de cultivos alemana de *Bacillus subtilis* subsp. *subtilis*, DSMZ (DSM nº 10)

*Bacillus stearothermophilus*. Cepa LLD-R - Depositada como NCIMB 12403

*Bacillus stearothermophilus*. Cepa LLD-15 - Depositada como NCIMB 12428

#### **Plásmidos**

Se obtuvieron plásmido pCR-Blunt y pCR-TOP02 de Invitrogen

Plásmido pUB110 – La cepa BD170 de *Bacillus subtilis* que aloja este plásmido se obtuvo de la Colección de cultivos alemana, DSMZ (DSM nº 4514).

El plásmido pUC18 se obtuvo de Sigma-Aldrich.

#### **Ejemplo 1. Construcción de una formiato deshidrogenasa sintética (Fig. 2)**

Una secuencia de aminoácidos (nº de acceso de la base de datos de proteínas NCBI P33160 - SEC ID Nº: 3) de formiato deshidrogenasas de 101 de *Pseudomonas sp* se retrotradujo en la secuencia de ADN con codones optimizados para *Geobacillus thermoglucosidasius*. Se añadieron una región promotora y terminadora independiente de rho de *Bacillus*, en la dirección 5' y en la dirección 3' de la secuencia traducida, respectivamente (Figura 2). La secuencia novedosa mostró menos del 40% de similitud con secuencias del gen *fdh* conocidas (37% de identidad con el gen *fdh1* conocido). Se diseñaron sitios *Xba1* en ambos lados de la construcción para facilitar su clonación en vectores adecuados.

La secuencia deseada se sintetizó usando el procedimiento de Gao y col. (véase Xinxin Gao, Peggy Yo, Andrew Keith, Timothy J. Ragan y Thomas K. Harris (2003). *Nucleic Acids Research*, 31 (22), e143) y se clonaron en pCR-Blunt en su única posición *Xba1*. El vector pCR-F1 resultante (Figura 4) se introdujo en células DH5-alfa de *E. coli* y los clones positivos se confirmaron por PCR y análisis de restricción.

Están disponibles dos estrategias alternativas para insertar y expresar este gen *fdh* sintético en el genoma de bacilos diana como se muestra en los siguientes ejemplos.

Ejemplo 2. Inserción del gen *fdh* en múltiples sitios (SI)

5 Esta estrategia se aplica a cepas tales como la cepa LLD-R de *Bacillus stearothermophilus* que contienen una secuencia de inserción (SI) que frecuentemente se recombina en múltiples sitios de inserción. Se espera que un vector que lleva el gen *fdh* y esta secuencia SI se integre establemente en una o más de tales localizaciones.

Construcción del plásmido pUB-ISF1 (Figura 5)

10 En primer lugar, la secuencia de inserción conocida de la cepa LLD-R (SEC ID N°: 5 y Figura 3) se amplifica por PCR usando un cebador directo (AGTACTGAAATCCGGATTTGATGGCG - SEC ID N°: 6) y un cebador inverso (AGTACTGCTAAATTTCCAAGTAGC - SEC ID N°: 7) con la cepa LLD-15 de *B. stearothermophilus* como molde. Los sitios de restricción *Sca1* se introducen en ambos extremos de la secuencia. El producto de PCR se clona primero en el plásmido pCR-TOPO2.1 y el plásmido pCR-IS resultante se introduce luego en células DH5-alfa de *E. coli* y se usa para aislar la región de SI por digestión con restricción de *Sca1*. Entonces, la SI aislada se clona en pUB110 en su único sitio *Sca1* y el plásmido resultante pUB-IS se introduce en *Bacillus subtilis*.

20 Entonces, un fragmento de 1,5 kb que contiene el promotor *ldh* y el gen *fdh* se digieren a partir del plásmido pCR-F1 usando la enzima de restricción *Xba1*, y se clona en el plásmido pUB-IS que ya se había linealizado con la misma enzima. Entonces, el plásmido pUB-ISF1 resultante (Figura 5) se introduce en *B. subtilis* y los clones positivos se seleccionan sobre placas de TGP-kan y se confirman por PCR y análisis de restricción.

Integración del gen *fdh* en la cepa LLD-R

25 Entonces, el plásmido pUB-ISF1 se metila *in vitro* con la enzima metilasa *HaeIII* y luego la cepa LLD-R de *Bacillus stearothermophilus* o sus células de cepas delecionadas en *ldh* (véase el Ejemplo 3) se transforman con el plásmido pUB-ISF1 metilado. Los clones positivos se seleccionan después de 48 horas sobre placas de TGP-kan a 50°C y se analizan por amplificación por PCR del gen *fdh*.

30 Entonces, el gen *fdh* se integra en el cromosoma cultivando un clon transformado en medio TGP-kan a 60-65°C para algunas generaciones y seleccionando sobre placas de TGP-kan. Los clones positivos se analizan para la presencia del gen *fdh* y luego se criban para la producción de etanol y utilización de azúcares C5 (pentosa) y C6 (hexosa) en matraces agitados y fermentadores.

35 Ejemplo 3. Construcción de cepas delecionadas en *ldh*

La primera etapa es clonar marcador de resistencia de kanamicina de *Bacillus* (*kan*) y un casete que lleva el gen *ldh* de la cepa LLD-R de *B. stearothermophilus* en el plásmido pUC18, que sólo puede replicarse en microorganismos Gram-negativos.

40 Construcción de un vector de clonación de *Bacillus*. Plásmido pUCK (Figura 6)

45 Se clonó un gen de resistencia a kanamicina (*kan*) en el plásmido pUC18 en su único sitio *Zra1* que está fuera de cualquier región codificante y del gen indicador (*lacZ*) en el plásmido. Para clonar el gen *kan*, un fragmento de 1,13 kb que contiene el gen de resistencia a kanamicina se purificó por PCR con los cebadores:

kan-BsZ-F (ACACAGACGTCGGCGATTTGATTCATAC - SEC ID N°: 10) y  
kan-BsZ-R (CGCCATGACGTCCATGATAATTACTAATACTAGG - SEC ID N°: 11)

50 usando el plásmido pUB110 como molde. Los sitios *Zra1* se introdujeron en ambos extremos del gen *kan* mediante los cebadores. Entonces, el producto de PCR se digirió con la enzima endonucleasa de restricción *Zra1* y se ligó con el plásmido pUC18 previamente digerido con *Zra1* y desfosforilado. Entonces, el plásmido pUCK resultante (Figura 6) se introdujo en células DH5-alfa de *E. coli*. Se seleccionaron clones positivos sobre placas de LB-amp y se confirmaron por PCR y análisis de restricción.

55 Construcción del plásmido pUCK-LC (Fig. 7) que lleva un gen *ldh* delecionado

60 Se diseñó un casete de *ldh* de 1,36 kb para contener el gen *ldh* completo de la cepa LLD-R de la que se delecionaron 363 pb de su ORF más sus flancos. El casete se construyó por amplificación por PCR de las regiones superiores e inferiores del gen *ldh* usando la cepa LLD-R como molde. Entonces, estas regiones se ligaron y se clonaron en el plásmido pUCK. Los sitios *BglII* se introdujeron en los cebadores internos. La región superior se amplificó por PCR usando los siguientes cebadores:

65 LC-U-F1 (AGGGCAATCTGAAAGGAAGGGAAAATTCC - SEC ID N°: 12) y  
LC-UB-R1 TGCACAGATCTCCACCAAATCGGCGTC - SEC ID N°: 13).

La región inferior se amplificó por PCR usando los siguientes cebadores:

LC-DB-F1 (TTGAGCAGATCTTGATGCAAAACGATAAC - SEC ID N°: 14) y  
LC-D-R1 (TAAAGCCGATGAGCAGCAGTTGAAG - SEC ID N°: 15).

Los productos de PCR se digirieron con la enzima endonucleasa de restricción BgIII y se ligaron usando la enzima T4 ADN ligasa. Usando el ligando como molde, el casete de *ldh* se amplificó luego por PCR usando como cebadores:

LC-UX-F2 (ATATTATCTAGACATTACGGAAATGATAATGGC - SEC ID N°: 16) y  
LC-DX-R2 (TCACAATCTAGACAATCGGCCATAAAC - SEC ID N°: 17).

Los sitios *Xba*I se introdujeron en ambos extremos del casete mediante los cebadores. Entonces, el producto de PCR se digirió con la enzima *Xba*I y se clonó en el plásmido pUCK previamente digerido con la misma enzima y se desfosforiló. Entonces, el plásmido pUCK-LC resultante se introdujo en DH5-alfa de *E. coli*. Los clones positivos se seleccionaron sobre placas de LB-amp y se confirmaron por PCR y análisis de restricción.

#### Construcción de cepas delecionadas en *ldh*

El plásmido pUCK-LC se metila *in vitro* con la enzima metilasa *Hae*III y células termófilas naturales (por ejemplo, cepa LLD-R) se transforman en el plásmido metilado por electroporación (1000 V, 201 ohmios, 25 micro-faradios y 5 mili-segundos). Se seleccionan clones positivos sobre placas de TGP-kan a 65°C y se confirman como acontecimientos de cruces individuales por amplificación por PCR del gen *kan*.

Para lograr la deleción de genes por cruce doble, los clones positivos se cultivan en medio TGP durante algunas generaciones (aproximadamente 5 sub-cultivos) y se seleccionan los clones que pueden cultivarse sobre placas de TGP, pero no sobre placas de TGP-kan. Entonces, los clones positivos se confirman como deleciones del gen *ldh* y para la ausencia del gen kanamicina por análisis por PCR. Entonces, los clones se caracterizan para la producción de etanol y la utilización de azúcares C5 y C6 en matraces agitados y en fermentadores.

#### Ejemplo 4. Inserción simultánea del gen *fdh* y deleción del gen *ldh*

Esta estrategia alternativa es ampliamente aplicable a una amplia clase de microorganismos fermentativos de heterolactato, además de bacilos termófilos, aunque los últimos se usarán como ilustración.

#### Construcción del plásmido pUCK-LF (Fig. 9)

Un casete integrante del gen que contiene el gen *fdh* más el gen *ldh* completo y aproximadamente 300 pb de regiones flanqueantes en la dirección 5' y en la dirección 3' se clona en el plásmido pUCK. En esta construcción, los primeros 450 pb del marco de lectura abierto de *ldh* se sustituyen con el gen *fdh* de tal forma que la expresión génica esté bajo el control del promotor del gen *ldh*.

Para lograr esto, un fragmento de ADN de aproximadamente 750 pb que contiene la región en la dirección 3' del gen *ldh* se amplifica por PCR usando;

LDHB-X-F1 (GAACGATTCTAGATACAGCAAGATTCCGC - SEC ID N°: 8) y  
LDHB-E-R1 (GTTTGCGAATTCATAGACGGACGCAG - SEC ID N°: 9) como cebadores y la cepa LLD-R de *Bacillus stearothermophilus* como molde. Por tanto, los sitios *Xba*I y *Eco*R1 se introducen en los extremos del fragmento del PCR. Entonces, el fragmento de PCR se digiere y se clona direccionalmente en el plásmido pUCK entre los sitios *Xba*I y *Eco*R1. El plásmido resultante, pUCK-*ldh*B (Figura 8) se introduce en DH5-alfa de *E. coli* y los clones positivos se seleccionan sobre placas de LB-amp y se confirman por PCR y restricción.

Entonces, un fragmento de 1,5 kb que contiene el promotor de *ldh* y el gen *fdh* se digieren fuera del plásmido pCR-F1 usando la enzima de restricción *Xba*I y se clonan en el plásmido pUCK-*ldh*B (Figura 8) que ya se había linealizado con la misma enzima. El plásmido pUCK-LF1 resultante (Figura 9) se introduce en células DH5-alfa de *E. coli* y los clones se seleccionan sobre placas de LB-Amp. Los clones positivos y la orientación correcta de la construcción se confirman por PCR y análisis de restricción.

#### Construcción de cepas que producen etanol por la novedosa ruta de PFL-FDH

El plásmido pUCK-LF1 se metila *in vitro* con la enzima metilasa *Hae*II y células naturales diana (por ejemplo, células de la cepa LLD-R) se transforman con el plásmido metilado y se seleccionan sobre placas de TGP-kan a 60°C. Los clones positivos representan acontecimientos de cruce individuales y se analizan por amplificación por PCR del gen *fdh*.

Para lograr la integración de cruce doble se seleccionan clones que se cultivan sobre placas de TGP, pero no sobre placas de TGP-kan. Entonces, los clones positivos se confirman para la presencia del gen *fdh* y la ausencia del gen

# ES 2 396 151 T3

kanamicina. Finalmente, los clones se criban para la producción de etanol y la utilización de azúcares C5 y C6 en matraces agitados y en fermentadores.

## LISTADO DE SECUENCIAS

5	<110> Bioconversion Technologies Limited	
	<120> Potenciamiento de la producción de etanol microbiana	
10	<130> P86806WO00	
	<150> GB 0605890.3	
	<151> 24/03/2006	
15	<160> 17	
	<170> PatentIn versión 3.3	
20	<210> 1	
	<211> 1170	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> gen fdh sintético	
	<400> 1	
	atggcaaaag tactttgCGT tctttatgat gatccggtag atggctatcc gaagacgtat	60
	gcccGagatg atttaccGaa gatagatcac tatccaggag ggcaaacatt gccGacgccg	120
	aaagcaattg acttcacgcc tgggcaattg ttaggaagcg tatctggcga gctgggactt	180
	agaaagtatc ttgagtcaa tggacatacg ttagtggtaa ctagcgataa ggatgggcca	240
	gactcagtgt ttgaacggga gttagtggat gccgatgttg tcattagtca accgtttctgg	300
	cctGcatatc ttacgccgga aagaatcgca aaagcgaaga acttgaaact agccctgaca	360
	gcaggaattg gaagcgatca tgtggatttg caaagcgcta ttgatcgcaa tgttaccgtg	420
	gcagaggtga catattgtaa ttctattagt gtagctgagc atgtggtaat gatgatttta	480
	tccttagtta gaaattactt gccgagccac gaatgggctc gtaaaggcgg gtggaatatt	540
	gcagattgCG tttctcatgc ttatgattta gaggcgatgc atgttggcac ggttgccggcg	600
	ggacgtatag gcttggcagt cttgcgtcga ctagcaccgt ttgacgttca tttacactat	660
	actgatcgac atcgtcttcc agagtccgtt gagaaagaac ttaacttaac ctggcatgca	720
	acgCGtgaag atatgtatcc ggtgtgtgac gtggtaacat tgaattgccc gttacatcct	780
	gaaactgaac acatgatcaa tgacgaaacg ttgaaactgt ttaaacgagg cgttatatc	840
	gtaaacacag ccagagggaa actttgtgat cgggatgctg tagccagagc acttgagagc	900
	ggacgcttag ccgggtatgc aggcgacgtg tggtttccac aacctgcccc gaaagatcat	960
30	ccgtggagaa cgatgccgta taatggaatg acgccacata tttcaggcac tacgttaaca	1020
	gcacaagcac gttatgCGgc cggcaccCGt gaaattcttg agtgcttctt cgaagccCGt	1080
	ccgatccgag atgaatattt gattgtacaa ggtggcgcGat tagctgggac aggagcacat	1140
	agttatagca aaggcaatgc tacgggaggc	1170

ES 2 396 151 T3

<210> 2  
 <211> 1558  
 <212> ADN  
 <213> secuencia artificial

5

<220>  
 <223> gen fdh sintético que incluye promotor y terminador

<400> 2

10

```

gcctctagaa gggcaatctg aaaggaaggg aaaattcctt tcggattgtc cttttagtta      60
tttttatggg gagtgaatat tatataggca ttacggaaat gataatggca gagttttttc     120
atattattaga ctgcttgatg taattggatg tgatgataca aaaataatgt tgtgtaaaca     180
aaatgttaac aaaaaagaca aatttcattc atagttgata cttgataaaag attgtgaaat     240
aatgcacaat atatcaatgt atgagcagtt tcacaaatc attttttggg aaggatgaca     300
gacagcgatg gcaaaaagtac tttgcgttct ttatgatgat ccggtagatg gctatccgaa     360
gacgtatgcc cgagatgatt taccgaagat agatcactat ccaggagggc aaacattgcc     420
gacgccgaaa gcaattgact tcacgcctgg gcaattgtta ggaagcgtat ctggcgagct     480
gggacttaga aagtatcttg agtccaatgg acatacgtta gtggtaacta gcgataagga     540
tgggccagac tcagtgtttg aacgggagtt agtggatgcc gatgttgtca ttagtcaacc     600
gttctggcct gcatacttta cgcggaaaag aatcgcaaaa gcgaagaact tgaaactagc     660
cctgacagca ggaattggaa gcgatcatgt ggatttgcaa agcgctattg atcgcaatgt     720
taccgtggca gaggtgacat attgtaattc tattagtgtg gctgagcatg ttgtaatgat     780
gattttatcc ttagttagaa attacttgcc gagccacgaa tgggctcgta aaggcgggtg     840
gaatattgca gattgcgttt ctcatgctta tgatttagag gcgatgcatg ttggcacggt     900
tgccggcgga cgtataggct tggcagttct gcgtcgacta gcaccgtttg acgttcattt     960
acactatact gatcgacatc gtcttcaga gtccgttgag aaagaactta acttaacctg    1020
gcatgcaacg cgtgaagata tgtatccggt gtgtgacgtg gtaacattga attgcccgtt    1080
acatcctgaa actgaacaca tgatcaatga cgaaacgttg aaactgttta aacgaggcgc    1140
ttatatcgta aacacagcca gagggaaact ttgtgatcgg gatgctgtag ccagagcaact    1200
tgagagcggg cgcttagccg ggtatgcagg cgacgtgtgg tttccacaac ctgccccgaa    1260
agatcatccg tggagaacga tgccgtataa tggaatgacg ccacatattt caggcactac    1320
gttaacagca caagcacggt atgcggccgg caccctgaa attcttgagt gcttcttcga    1380
aggccgtccg atccgagatg aatatttgat tgtacaaggt ggcgcattag ctgggacagg    1440
agcacatagt tatagcaaag gcaatgctac gggaggcagc gaggaagcag ctaaatttaa    1500
gaaagcgggt taacacagca ggggctgac ggcccctgtt atgtttcatt ctagagcc     1558
    
```

<210> 3  
 <211> 401  
 <212> PRT  
 <213> 101 de Pseudomonas sp.

15

<400> 3

20

ES 2 396 151 T3

Met Ala Lys Val Leu Cys Val Leu Tyr Asp Asp Pro Val Asp Gly Tyr  
 1 5 10 15

Pro Lys Thr Tyr Ala Arg Asp Asp Leu Pro Lys Ile Asp His Tyr Pro  
 20 25 30

Gly Gly Gln Thr Leu Pro Thr Pro Lys Ala Ile Asp Phe Thr Pro Gly  
 35 40 45

Gln Leu Leu Gly Ser Val Ser Gly Glu Leu Gly Leu Arg Lys Tyr Leu  
 50 55 60

Glu Ser Asn Gly His Thr Leu Val Val Thr Ser Asp Lys Asp Gly Pro  
 65 70 75 80

Asp Ser Val Phe Glu Arg Glu Leu Val Asp Ala Asp Val Val Ile Ser  
 85 90 95

Gln Pro Phe Trp Pro Ala Tyr Leu Thr Pro Glu Arg Ile Ala Lys Ala  
 100 105 110

Lys Asn Leu Lys Leu Ala Leu Thr Ala Gly Ile Gly Ser Asp His Val  
 115 120 125

Asp Leu Gln Ser Ala Ile Asp Arg Asn Val Thr Val Ala Glu Val Thr  
 130 135 140

Tyr Cys Asn Ser Ile Ser Val Ala Glu His Val Val Met Met Ile Leu  
 145 150 155 160

ES 2 396 151 T3

Ser Leu Val Arg Asn Tyr Leu Pro Ser His Glu Trp Ala Arg Lys Gly  
 165 170 175

Gly Trp Asn Ile Ala Asp Cys Val Ser His Ala Tyr Asp Leu Glu Ala  
 180 185 190

Met His Val Gly Thr Val Ala Ala Gly Arg Ile Gly Leu Ala Val Leu  
 195 200 205

Arg Arg Leu Ala Pro Phe Asp Val His Leu His Tyr Thr Asp Arg His  
 210 215 220

Arg Leu Pro Glu Ser Val Glu Lys Glu Leu Asn Leu Thr Trp His Ala  
 225 230 235 240

Thr Arg Glu Asp Met Tyr Pro Val Cys Asp Val Val Thr Leu Asn Cys  
 245 250 255

Pro Leu His Pro Glu Thr Glu His Met Ile Asn Asp Glu Thr Leu Lys  
 260 265 270

Leu Phe Lys Arg Gly Ala Tyr Ile Val Asn Thr Ala Arg Gly Lys Leu  
 275 280 285

Cys Asp Arg Asp Ala Val Ala Arg Ala Leu Glu Ser Gly Arg Leu Ala  
 290 295 300

Gly Tyr Ala Gly Asp Val Trp Phe Pro Gln Pro Ala Pro Lys Asp His  
 305 310 315 320

Pro Trp Arg Thr Met Pro Tyr Asn Gly Met Thr Pro His Ile Ser Gly  
 325 330 335

Thr Thr Leu Thr Ala Gln Ala Arg Tyr Ala Ala Gly Thr Arg Glu Ile  
 340 345 350

Leu Glu Cys Phe Phe Glu Gly Arg Pro Ile Arg Asp Glu Tyr Leu Ile  
 355 360 365

Val Gln Gly Gly Ala Leu Ala Gly Thr Gly Ala His Ser Tyr Ser Lys  
 370 375 380

Gly Asn Ala Thr Gly Gly Ser Glu Glu Ala Ala Lys Phe Lys Lys Ala  
 385 390 395 400

Val

5 <210> 4  
 <211> 307  
 <212> ADN  
 <213> Cepa LLD-R de Bacillus stearothermophilus

10 <400> 4

ES 2 396 151 T3

gcctctagaa gggcaatctg aaaggaaggg aaaattcctt tcggattgtc cttttagtta 60  
 tttttatggg gagtgaatat tatataggca ttacggaaat gataatggca gagttttttc 120  
 atttattaga ctgcttgatg taattggatg tgatgataca aaaataatgt tgtgtaaaca 180  
 aatgttaac aaaaaagaca aatttcattc atagttgata cttgataaag attgtgaaat 240  
 aatgcacaat atatcaatgt atgagcagtt tcacaaattc attttttggg aaggatgaca 300  
 gacagcg 307

<210> 5  
 <211> 3006  
 <212> ADN  
 <213> Cepa LLD-R de Bacillus stearothermophilus

5

<400> 5

tgtaaattat cactttattt cgcacaaaa aagactcttt tttgcacatt ctttcggaat 60  
 atccctctcc ccctttccga aagaatgtgc taaatTTTTT gtgaattatt tcggaatggg 120  
 acatgggtga tttccagagg ggggacgagg acgtgattcg gtttcatgac tttcaggtcg 180  
 atgtccagac atatgcccag cggggaaaac aaaacgactt tccccttctt aagcgggtcc 240  
 ctcatgcca ggcaaaaacgc cctctttatc gccatgggta ctacgaacga aatgcogtga 300  
 cgtcgcacatca gtcttatcgc atttggatcg ctcggtatcg ctgcccggag tgcaggagga 360  
 cggtggccgt gttgccttca tttcttctcc cttatTTTca gtatacgttg cccaccatat 420  
 ggagagtggg gaaagaacgg ttgggcctga ctccgaaacg ggggatggag gaggctccac 480  
 tccttcctac ggatgaaggg gttttatTTT atgtcccagc gtttattccg aaatttgaac 540  
 caccttcatt ggtTTTTTgc ggagcgtgg agaaaaattg gtctgcat cgcceaagcc 600  
 gagagaacga gccctatggg ggatccagac gatggaggag atcggcctct ttttcgcat 660  
 ccaagagata tgggagcacc gatcgacgca tctTTTTTca cgtacattca gttcctgatt 720  
 tacttatatc cccttatatg gaatcattta tagattccca aacctttcct ctcgacggtc 780  
 gggggaatga tccgatagga tagagacagg atggaccgat aaggtcctag aatgggatga 840

10

ES 2 396 151 T3

acgaaggagg agatcgaaat gaatgagtcg atgagacagg agatcgcttt atttcggat 900  
 ggattgatcg ctccattggt gaatggacaa gtcgatccaa aacgtacttg aaggaagtag 960  
 cggaacggat ccatcaagt cccaccatg gagagaaacg catcgccgcc aaaacgatcc 1020  
 tcgactggtg cacgcagtac aaaaaagggg gctttgaggc gctgaagccg aaacgacggt 1080  
 cggaccgtgg ccattcccgg aggctgtcac ctgaagaaga ggatcacatt ttagccctga 1140  
 gaaaaaaca cccccatg cccgtgacgg tgttttacca acaccttacc gagcaggggg 1200  
 aatccaatc catctcttat ttcaactat accgactttt aaaaaatac aacctcgtgg 1260  
 ggaaagaaat tttaccgatt cctgaacgaa aacgattcgc gtacgatcag gtcaatgagc 1320  
 tctggcaagg tgatttgtcc catggccggt tgattcgcgt gaatggcaaa acgcaaaaaa 1380  
 cgtttttgat tgcctatata gatgactgct cgcgggctgt gccgtacgct cagtttttct 1440  
 cttccgagaa atttgacggg ttgcgatcg taaccgcgcg gagttagga tcaccttgat 1500  
 ccatacccag cgtacgatc cgcaaacgaa agggaaaatc gaacgatttt tccgcaccgt 1560  
 acagacgcgg ttttaccggt tgctcgaat gaattcaccg aagtcgctcg aagagctaaa 1620  
 cgagcgattt tgggaagtgg tggaggaaga ttaccatcga aaaccgcatg cctcgttgaa 1680  
 cgggaagacg ccacatgaag tgtttcaatc gcaagtccat ttggtgctgt tcgtcgagga 1740  
 ttcggattgg ctcgactcga tctttttgaa acgcgaatac cgtaaagtga aggccgatgg 1800  
 tacggtcacg ttgaacaagc agctgtatga agttccgccc cggttcatcg gacaatcgat 1860  
 cgaactcctg tatgatgaac aaggcgtgta tgtgtacgaa gacggtaac gggtcgccga 1920  
 agcggctcct gttcgttcg aggacaatgc ctatgtgaaa cgccatcggg caccgtttgc 1980  
 ggcggttccg gtagaggag gcgaaaacga tgtataaac gttttattcc ctttcccag 2040  
 agccgttttc gaaggagacg aatccaccag aggcttatca aggggcctcg tatcaagagg 2100  
 ccctcgccgc tttgactac gtgaaacgaa caagaggat cgggctattg atcggtgaa 2160  
 caggggccgg caagacattc gcccttcggg cgtttaagga atccctgaat ccgtcactgt 2220  
 atcacgtcgt ttatttcca ttgtcaacgg gaagcgtgat ggacttttat cgcggccttg 2280  
 cttcgggct cggggaagag ccgaaatacc gcaaggctga ctgtttttat caaatccaac 2340  
 aagggatcga gcgcttgat catgaacaac ggtaacgtc agtgttcatc ctcgatgaaa 2400  
 tgcatttagc gaaggatgcc tttctgcagg atatcgcat cctgttcaac tttcacatgg 2460  
 actcaacaaa tccgtttgc ttgattttgg cgggctgcc ccatttacag gcaaaactac 2520

# ES 2 396 151 T3

```

ggttgaatca acaccgtccg cttaccaaac gaatcatcat gcgataccag atggggcctc 2580
ttgataagga agaagtggta ggatataatcg aacaccgcct gaaacaggcg ggggcgaaac 2640
acccgatttt taccocagct gccttagaag cgatcgccct gcagtcgcag ggggtggccgc 2700
ggatcatcaa caacctcgcc accacttgcc tgttatacgg cgctcaatta aaaaaacata 2760
tgattgacga agacattgtg cgtatggcag ccgaagaaat ggggtactga cacagcaggg 2820
gctgatcggc ccctgttatg tttcatcccg atccatcctc attctagtta atcatccgaa 2880
ataatgtgca aatgttcgga aataatctgc aaaacctgga ataattcgca aagattttgc 2940
acattatttc cgaatccgtc cgaataaatt tgaaaaaggg attctgaaat aatgtgctaa 3000
tttaca 3006

```

```

5 <210> 6
  <211> 26
  <212> ADN
  <213> Secuencia artificial

10 <220>
  <223> cebador

  <400> 6

  agtactgaaa tccgatttg atggcg 26

15 <210> 7
  <211> 24
  <212> ADN
  <213> Secuencia artificial

20 <220>
  <223> cebador

  <400> 7

25 agtactgcta aattccaag tagc 24

  <210> 8
  <211> 29
  <212> ADN
  <213> Secuencia artificial

30 <220>
  <223> cebador

  <400> 8

  gaacgattct agatacagca agattccgc 29

40 <210> 9
  <211> 26
  <212> ADN
  <213> Secuencia artificial

45 <220>
  <223> cebador

  <400> 9

50 gttgcgaat tcatagacgg acgcag 26
  <210> 10

```

# ES 2 396 151 T3

<211> 28  
 <212> ADN  
 <213> Secuencia artificial

5

<220>  
 <223> cebador

<400> 10

10

acacagacgt cggcgatttg attcatac 28

<210> 11  
 <211> 34  
 <212> ADN  
 <213> Secuencia artificial

15

<220>  
 <223> cebador

20

<400> 11

cgccatgacg tccatgataa ttactaatac tagg 34

25

<210> 12  
 <211> 29  
 <212> ADN  
 <213> Secuencia artificial

30

<220>  
 <223> cebador

<400> 12

35

agggcaatct gaaaggaagg gaaaattcc 29

<210> 13  
 <211> 27  
 <212> ADN  
 <213> Secuencia artificial

40

<220>  
 <223> cebador

<400> 13

45

tgacagatc tccaccaaat cggcgtc 27

<210> 14  
 <211> 29  
 <212> ADN  
 <213> Secuencia artificial

50

<220>  
 <223> cebador

55

<400> 14

ttgacagat cttgatgcaa aacgataac 29

60

<210> 15  
 <211> 25  
 <212> ADN  
 <213> Secuencia artificial

65

<220>  
 <223> cebador

# ES 2 396 151 T3

<400> 15  
5 taaagccgat gagcagcagt tgaag 25  
<210> 16  
<211> 33  
<212> ADN  
<213> Secuencia artificial  
10  
<220>  
<223> cebador  
15  
<400> 16  
atattatcta gacattacgg aaatgataat ggc 33  
<210> 17  
<211> 27  
20 <212> ADN  
<213> Secuencia artificial  
<220>  
25 <223> cebador  
<400> 17  
tcacaatcta gacaatcggc cataaac 27  
30

**REIVINDICACIONES**

- 5 1. Una bacteria termófila del género *Bacillus* que carece de actividad de lactato deshidrogenasa y que tiene actividad de piruvato formiato liasa, caracterizada porque la bacteria termófila contiene un gen que codifica una formiato deshidrogenasa ligada a NAD.
2. La bacteria termófila de la reivindicación 1, en la que el gen que codifica una formiato deshidrogenasa ligada a NAD está integrado en el genoma de la bacteria termófila.
- 10 3. La bacteria termófila de la reivindicación 1 ó 2, en la que el gen que codifica una formiato deshidrogenasa ligada a NAD se expresa a partir de su propio promotor o a partir de un promotor de la bacteria termófila.
- 15 4. La bacteria termófila de una cualquiera de las reivindicaciones 1 a 3, en la que el gen que codifica una formiato deshidrogenasa ligada a NAD se inserta en el gen lactato deshidrogenasa de la bacteria termófila, inactivándose así la actividad de lactato deshidrogenasa de la bacteria termófila.
- 20 5. La bacteria termófila de cualquier reivindicación precedente, en la que el gen que codifica una formiato deshidrogenasa ligada a NAD comprende la secuencia de nucleótidos expuesta como SEC ID N°: 1 ó 2.
- 25 6. La bacteria termófila de cualquier reivindicación precedente que ha sido transformada con una construcción de ADN que comprende un gen que codifica una formiato deshidrogenasa ligada a NAD operativamente ligado a una región en la dirección 5' de un gen que codifica una lactato deshidrogenasa, en la que la región en la dirección 5' incluye el promotor y que comprende además al menos parte del gen lactato deshidrogenasa en la dirección 3' del gen que codifica una formiato deshidrogenasa ligada a NAD de forma que el gen que codifica una formiato deshidrogenasa ligada a NAD está interpuesto entre una porción suficiente del gen lactato deshidrogenasa sobre cualquier lado para facilitar la integración del gen que codifica una formiato deshidrogenasa ligada a NAD por recombinación con un gen lactato deshidrogenasa en el genoma de la bacteria termófila.
- 30 7. Un gen que codifica una formiato deshidrogenasa ligada a NAD termoestable que comprende la secuencia de nucleótidos expuesta como SEC ID N°: 1.
- 35 8. Una construcción de ADN que comprende una secuencia reguladora operativamente ligada a un gen que codifica una formiato deshidrogenasa ligada a NAD termoestable que comprende la secuencia de nucleótidos expuesta como SEC ID N°: 1.
- 40 9. Una construcción de ADN que comprende un gen que codifica una formiato deshidrogenasa ligada a NAD, opcionalmente una formiato deshidrogenasa ligada a NAD termoestable, y un elemento de ADN transponible que puede integrarse en el genoma de una bacteria termófila del género *Bacillus* que facilita la integración del gen que codifica una formiato deshidrogenasa ligada a NAD en el genoma de la bacteria termófila transformado con la construcción de ADN.
- 45 10. Una construcción de ADN que comprende un gen que codifica una formiato deshidrogenasa ligada a NAD, opcionalmente una formiato deshidrogenasa ligada a NAD termoestable, operativamente ligado a una región en la dirección 5' de un gen que codifica una lactato deshidrogenasa, en la que la región en la dirección 5' incluye el promotor.
- 50 11. Una bacteria termófila del género *Bacillus* que comprende la construcción de ADN como se define en una cualquiera de las reivindicaciones 8 a 10.
- 55 12. Uso de una bacteria termófila como se reivindica en una cualquiera de las reivindicaciones 1 a 6 o según la reivindicación 11 para la producción de etanol.
13. Un proceso de fermentación para la producción de etanol que comprende suministrar una bacteria termófila como se reivindica en una cualquiera de las reivindicaciones 1 a 6 o según la reivindicación 11 con azúcares.

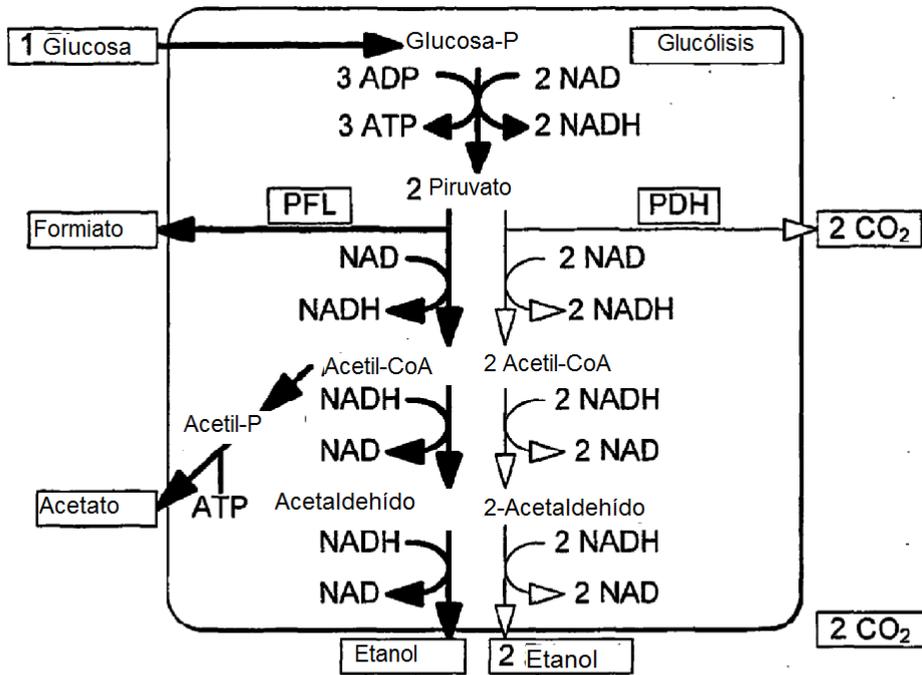


FIG. 1A

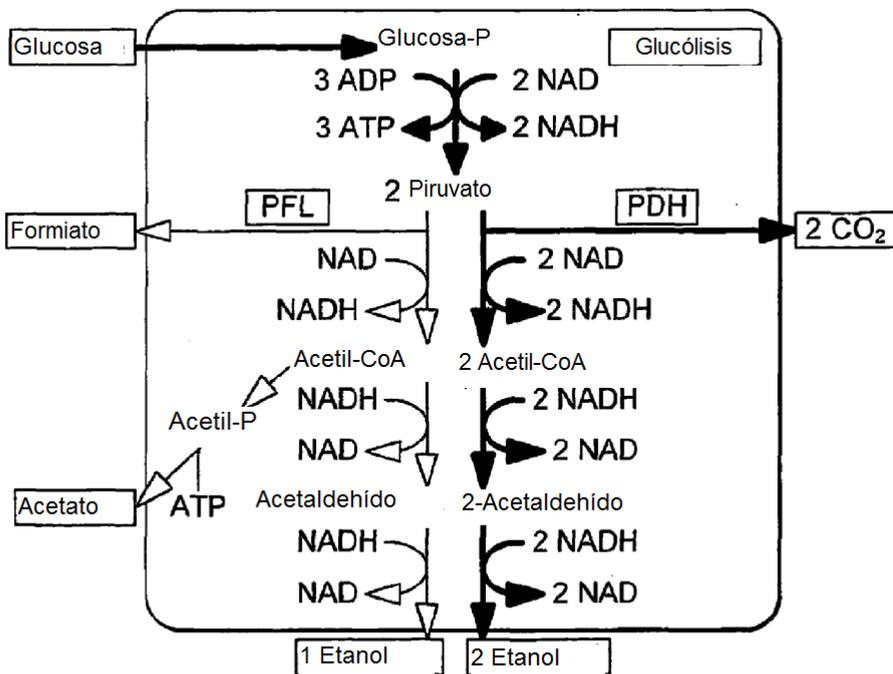


FIG. 1B

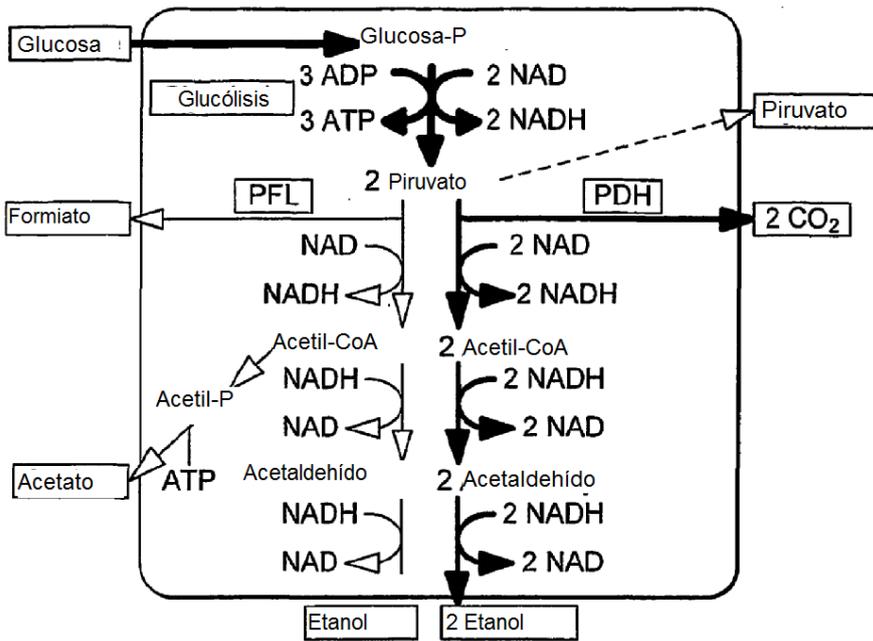


FIG. 1C

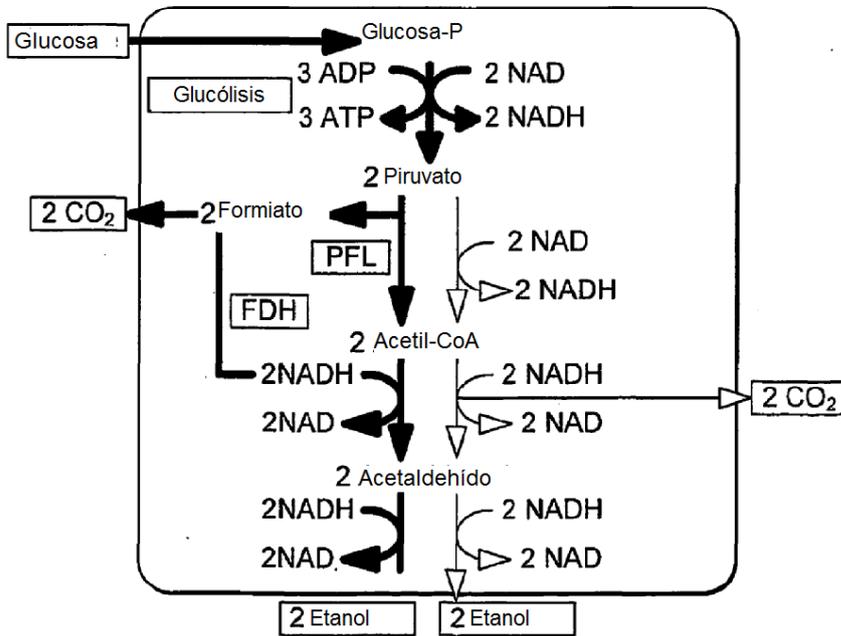


FIG. 1D

GCCTCTAGA  
 AGGGCAATCTGAAAGGAAGGGAAAATTCCTTTCCGGATTGTCCTTTTAGTTATTTT  
 TATGGGGAGTGAATATTATATAGGCATTACGGAAATGATAATGGCAGAGTTTTTT  
 CATTATTAGACTGCTTGATGTAATTGGATGTGATGATACAAAAATAATGTTGTGT  
 AAACAAAATGTTAACAAAAAGACAAATTTCAATCATAGTTGATACTTGATAAAGA  
 TTGTGAAATAATGCACAATATATCAATGTATGAGCAGTTTCACAAATTCATTTTT  
 GGAAAGGATGACAGACAGCG  
 ATGGCAAAAGTACTTTGCGTTCCTTTATGATGATCCGGTAGATGGCTATCCG  
 AAGACGTATGCCCGAGATGATTTACCGAAGATAGATCACTATCCAGGAGG  
 GCAAACATTGCCGACGCCGAAAGCAATTGACTTCACGCCTGGGCAATTGT  
 TAGGAAGCGTATCTGGCGAGCTGGGACTTAGAAAGTATCTTGAGTCCAAT  
 GGACATACGTTAGTGGTAACTAGCGATAAGGATGGGCCAGACTCAGTGTT  
 TGAACGGGAGTTAGTGGATGCCGATGTTGTCATTAGTCAACCGTTCTGGC  
 CTGCATATCTTACGCCGGAAAGAATCGCAAAAGCGAAGAAGACTTGAAACTA  
 GCCCTGACAGCAGGAATTGGAAGCGATCATGTGGATTTGCAAAGCGCTAT  
 TGATCGCAATGTTACCGTGGCAGAGGTGACATATTGTAATTCTATTAGTGT  
 AGCTGAGCATGTGGTAATGATGATTTTATCCTTAGTTAGAAATTACTTGCC  
 GAGCCACGAATGGGCTCGTAAAGGCGGGTGGAATATTGCAGATTGCGTTT  
 CTCATGCTTATGATTTAGAGGCGATGCATGTTGGCACGGTTGCGGCCGGA  
 CGTATAGGCTTGGCAGTCTTGCGTGCGACTAGCACCGTTTGACGTTTCATTTA  
 CACTATACTGATCGACATCGTCTTCCAGAGTCCGTTGAGAAAGAACTTAA  
 CTTAACCTGGCATGCAACGCGTGAAGATATGTATCCGGTGTGTGACGTGG  
 TAACATTGAATTGCCCGTTACATCCTGAAACTGAACACATGATCAATGAC  
 GAAACGTTGAAACTGTTTAAACGAGGCGCTTATATCGTAAACACAGCCAG  
 AGGGAAACTTTGTGATCGGGATGCTGTAGCCAGAGCACTTGAGAGCGGAC  
 GCTTAGCCGGGTATGCAGGCGACGTGTGGTTTCCACAACCTGCCCCGAAA  
 GATCATCCGTGGAGAACGATGCCGTATAATGGAATGACGCCACATATTT  
 AGGCACTACGTTAACAGCACAAAGCACGTTATGCGGCCGGCACCCGTGAAA  
 TTCTTGAGTGCTTCTTCGAAGGCGTCCGATCCGAGATGAATATTTGATTG  
 TACAAGGTGGCGCATTAGCTGGGACAGGAGCACATAGTTATAGCAAAGG  
 CAATGCTACGGGAGGC  
 AGCGAGGAAGCAGCTAAATTTAAGAAAGCGGTTTAA  
 CACAGCAGGGGCTGATCGGCCCTGTTATGTTTCAT  
TCTAGAGCC

FIG. 2

**TGTA**AATTATCACTTTATTTCCGCACAAAAAAGACTCTTTTTTGCACATTCCCTTCGGAAT  
 ATCCCTCTCCCCCTTTCCGAAAGAATGTGCTAAATTTTTTGTGAATTATTCGGAATGGG  
 ACATGGGTGATTTCAGAGGGGGGACGAGGACGTGATTCCGGTTTCATGACTTTCAGGTC  
 GATGTCCAGACATATGCCAGCGGGGAAAAACAAACGACTTTCCCCTTCTTAAGCGGT  
 GCCCTCATTGCCAGGCAAAACGCCCTCTTTATCGCCATGGGTACTACGAACGAAATGCC  
 GTGACGTGCGATCAGTCTTATCGCATTTGGATCGCTCGGTATCGCTGCCCGGAGTGCAG  
 GAGGACGGTGGCCGTGTTGCCTTCATTTCTTCTCCCTATTTTTCAGTATACGTTGCCAC  
 CATATGGAGAGTGGTGAAGAACGGTTGGGCTGACTCCGAAACGGGGGATGGAGGA  
 GGCTCCACTCCTTCTACGGATGAAGGGGTTTATTTATGTCCCGACGTTTATTCCGAA  
 ATTTGAACCACCTTCATTGGTTTTTTGCGGAGCGCTGGAGAAAAATTGGTCTGCCATC  
 GCCAAGCCGAGAGAACGAGCCCTATGGTGGATCCAGACGATGGAGGAGATCGGCCTC  
 TTTTCGTACATCCAAGAGATATGGGAGCACCGATCGACGCATCTTTTGCACGTACATT  
 CAGTTCCTGATTTACTTATATCCCCTTATATGGAATCATTATAGATTCCCAAACCTTTC  
 CTCGACGGTGGGGGAATGATCCGATAGGATAGAGACAGGATGGACCGATAAGGTC  
 CTAGAATGGGATGAACGAAGGAGGAGATCGAAATGAATGAGTCGATGAGACAGGAGA  
 TCGCTTATTTCCGGTATGGATTGATCGTCCATTGGTGAATGGACAAGTCGATCCAAAA  
 CGTACTTGAAGGAAGTAGCGGAACGGATCCATCAAGTTCCCACCATGGAGAGAAACG  
 CATCGCCGCAAAAACGATCCTCGACTGGTGCACGCAGTACAAAAAAGGGGGCTTTGAG  
 CGCTGAAGCCGAAACGACGGTCCGACCGTGGCCATTCCCAGGCTGTCACCTTGAAG  
 AAGAGGATCACATTTTAGCCCTGAGAAAAAAACACCCCCACATGCCCGTGACGGTGT  
 TTACCAACACCTTATCGAGCAGGGGGAAATCCAATCCATCTCTTATTTACTATATAACC  
 GACTTTTAAAAAATAACAACCTCGTGGGGAAAGAAATTTTACCGATTCTGAACGAAA  
 ACGATTCCGCTACGATCAGGTCAATGAGCTCTGGCAAGGTGATTTGTCCCATGGCCCGT  
 TGATTCCGCTGAATGGCAAAACGCAAAAAACGTTTTTGGATTGCCTATATCGATGACTGC  
 TCGCGGGTCTGTGCCGTACGCTCAGTTTTTCTCTCCGAGAAATTTGACGGGTTGCGGAT  
 CGIAACCGCGGGAGTTAGGGATCACCTTGATCCATACCCAGCCGTACGATCCGCAAA  
 GCAAAGGGAAAATCGAACGATTTTTCCGCACCGTACAGACGCGGTTTTACCCGTTGCTC  
 GAAATGAATTCACCGAAGTCGCTCGAAGAGCTAAACGAGCGATTTTGGAAAGTGGTTGG  
 AGGAAGATTACCATCGAAAAACCGCATGCCTCGTTGAACGGGAAGACGCCACATGAAGT  
 GTTCAATCGCAAGTCCATTTGGTGTCTGTTGAGGATTTCGATTGGCTCGATCGATCGA  
 TCTTTTTGAAACGCGAATACCGTAAAGTGAAGGCCGATGGTACGGTACGTTGAACAA  
 GCAGCTGTATGAAGTTCGCCCCCGGTTTCATCGGACAATCGATCGAACTCCGTTATGATG  
 AACAAAGGCGTGTATGTGTACGAAGACGGTCAACGGGTCGCCGAAGCGGTCCTTGTTCG  
 CTTGAGGACAATGCCTATGTGAAACGCCATCGGTACCGTTTTGCGGCGGTTCCGGTAG  
 AGGGAGGCGAAAACGATGTATAAAACGTTTTTATCCCTTTCCCGAGAGCCGTTTTCGAA  
 GGAGACGAATCCACCAGAGGCTTATCAAGGGGCCCTCGTATCAAGAGGCCCTCGCCGCT  
 TTGGACTACGTGAAACGAACAAGAGGGATCGGGCTATTGATCGGTGAACCAGGGGCCG  
 GCAAGACATTCGCCCTTCGGGCGTTAAGGAATCCCTGAATCCGTCACGTGTATCACGTC  
 GTTATTTTCCATTTGCAACGGGAAGCGTGATGGACTTTTATCGCGGCTTGCCTTCGGG  
 CTCGGGGAAGAGCCGAAATACCGCAAGGTCGACTTGTTTTATCAAATCCAACAAGGGA  
 TCGAGCGTTGTATCATGAACAACGGGTAACTGAGTTCATCCTCGATGAAATGCAT  
 TTAGCGAAGGATGCCCTTCTGCGAGGATATCGCGATCCTGTTCAACTTTCACATGGACTC  
 AACAAATCCGTTTGTCTTGATTTTGGCGGGGCTGCCCATTTACAGGCAAAACTACGGT  
 TGAATCAACACCGTCCGCTTACCAACGAATCATCATGCGATACCAGATGGGGCCTCTT  
 GATAAGGAAGAAGTGGTAGGATATATCGAACACCGCCTGAAACAGGCGGGGGCGAAA  
 CACCCGATTTTTACCCAGCTGCCTTAGAAGCGATCGCCCTGCAGTCGACAGGGGTGGCC  
 GCGGATCATCAACAACCTCGCCACCACTTGCTTGTATACGGCGCTCAATTAATAAAC  
 ATATGATTGACGAAGACATTGTGCGTATGGCAGCCGAAGAAATGGGGTACTGACACAG  
 CAGGGGCTGATCGGCCCTGTTATGTTTCATCCCGATCCATCCTCATTCTAGTTAATCAT  
 CCGAAATAATGTGCAAAATGTTGCGAAATAATCTGCAAAACCTGGAATAATTCGCAAAG  
 ATTTTGCACATTATTTCCGAATCCGTCCGAAATAATTTGAAAAAGGGATTCTGAAATAA  
 TGTGCTAATTTACA

FIG. 3

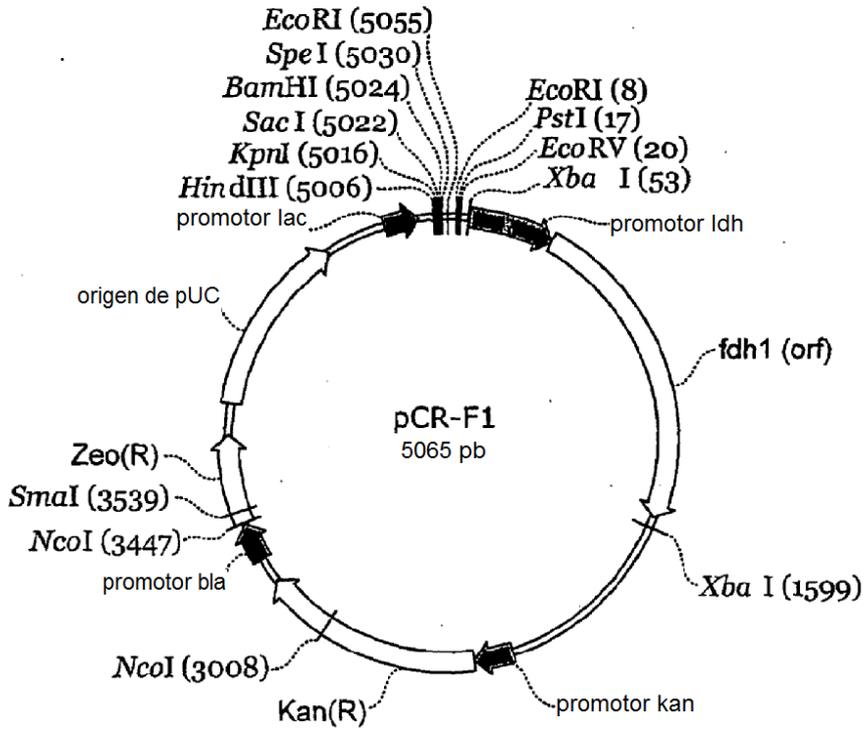


FIG. 4

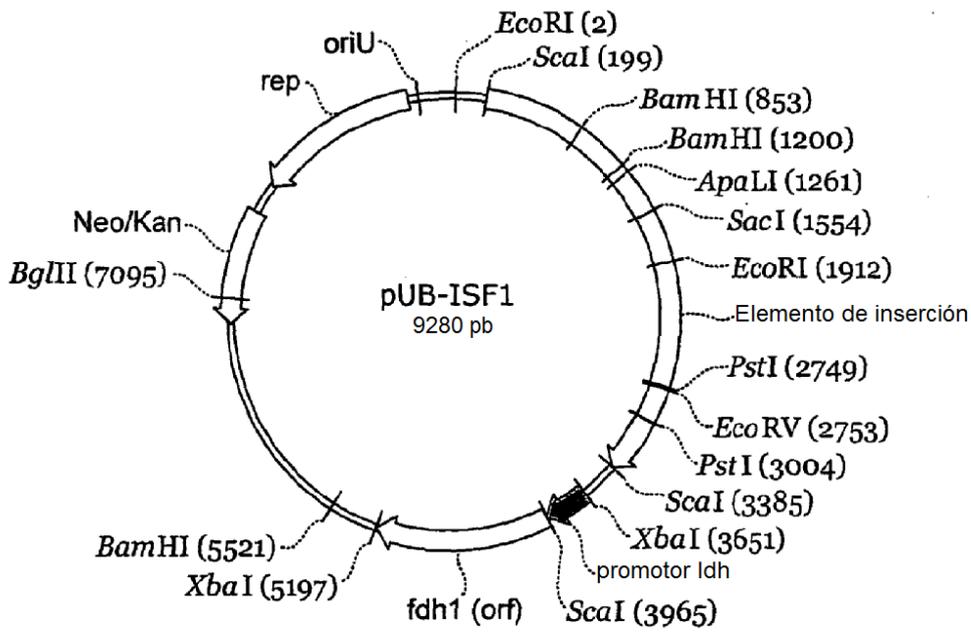


FIG. 5

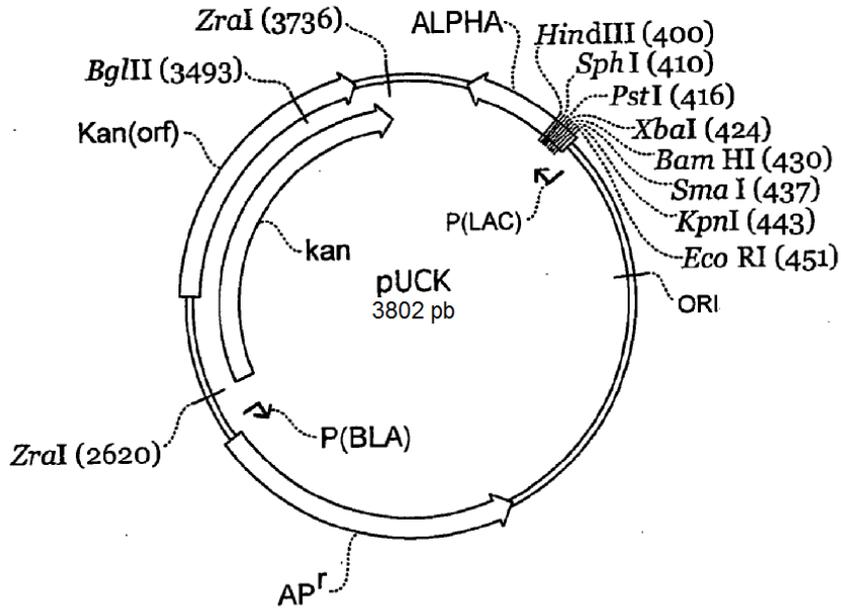


FIG. 6

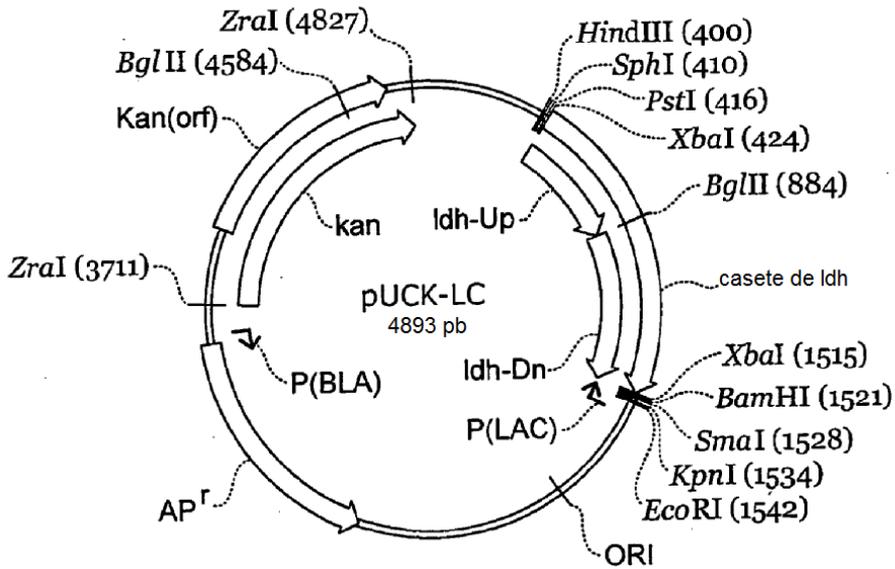


FIG. 7

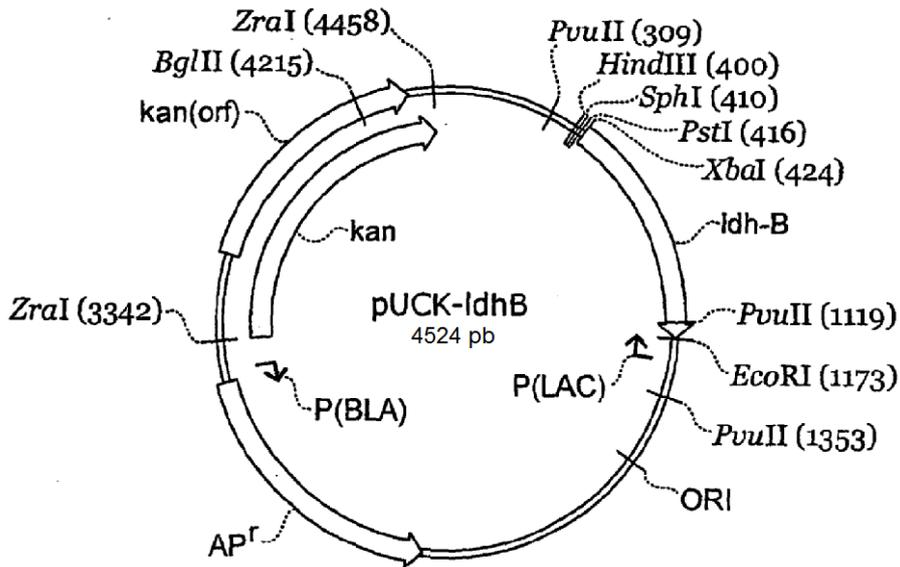


FIG. 8

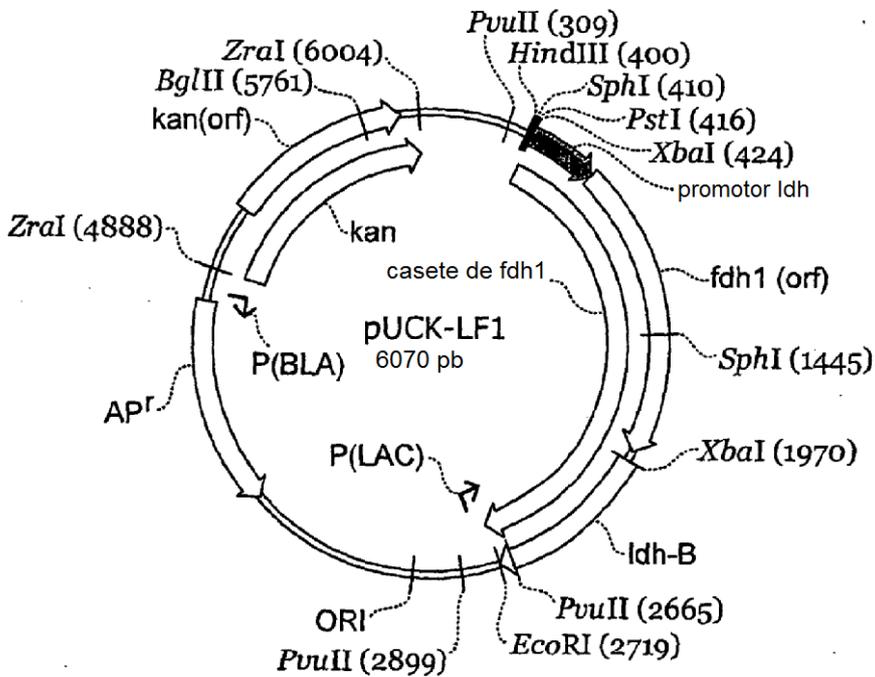


FIG. 9