

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 790 573**

51 Int. Cl.:

C12P 7/46 (2006.01)

C12N 9/02 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.12.2014 PCT/EP2014/077638**

87 Fecha y número de publicación internacional: **18.06.2015 WO15086839**

96 Fecha de presentación y número de la solicitud europea: **12.12.2014 E 14833137 (4)**

97 Fecha y número de publicación de la concesión europea: **04.03.2020 EP 3080251**

54 Título: **Fumarato reductasas**

30 Prioridad:

12.12.2013 EP 13196950

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

28.10.2020

73 Titular/es:

DSM IP ASSETS B.V. (100.0%)

Het Overloon 1

6411 TE Heerlen, NL

72 Inventor/es:

JONG, DE, RENÉ MARCEL;

ZHAO, ZHENG;

DULK, DEN, BEN y

WINTER, REMKO TSJIBBE

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 790 573 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

Fumarato reductasas

Campo de la invención

5 La presente invención se refiere a un polipéptido variante que tiene actividad fumarato reductasa. La invención también se refiere a un ácido nucleico que comprende una secuencia que codifica un polipéptido variante de este tipo, a una construcción de ácido nucleico que comprende un ácido nucleico de este tipo, a un vector de expresión recombinante que comprende un ácido nucleico o construcción de ácido nucleico de este tipo y a una célula huésped recombinante que comprende un ácido nucleico, construcción de ácido nucleico o vector de expresión de este tipo. La invención se refiere, además, a un método para producir un polipéptido variante que tiene actividad fumarato reductasa y a un método para la producción de un ácido dicarboxílico.

Antecedentes de la invención.

El ácido succínico es un precursor potencial de numerosos productos químicos. Por ejemplo, el ácido succínico se puede convertir en 1,4-butanodiol (BDO), tetrahidrofurano y gamma-butilolactona. Otro producto derivado del ácido succínico es un polímero de poliéster que se prepara uniendo ácido succínico y BDO.

15 El ácido succínico se produce predominantemente a través de procedimientos petroquímicos por hidrogenación de butano. Estos procedimientos se consideran perjudiciales para el medio ambiente y son costosos. La producción fermentativa de ácido succínico puede ser un procedimiento alternativo atractivo para la producción de ácido succínico, en el que se puede utilizar un material de alimentación renovable como fuente de carbono.

20 Un cierto número de diferentes bacterias tales como *Escherichia coli*, y las bacterias del rumen *Actinobacillus*, *Anaerobiospirillum*, *Bacteroides*, *Mannheimia* o *Succinimonas*, sp. son conocidas por producir ácido succínico. La ingeniería metabólica de estas cepas bacterianas ha mejorado el rendimiento y/o la productividad del ácido succínico, o ha reducido la formación de sub-productos.

25 El documento WO2007/061590 describe una levadura negativa a la piruvato descarboxilasa para la producción de ácido málico y/o ácido succínico que se transforma con una enzima piruvato carboxilasa o una fosfoenolpiruvato carboxilasa, una enzima malato deshidrogenasa y una proteína transportadora de ácido málico (MAE).

A pesar de las mejoras que se han hecho en la producción fermentativa de ácido succínico, sigue habiendo una necesidad de microorganismos mejorados para la producción fermentativa de ácido succínico.

30 El documento WO2009/065778 describe que pueden alcanzarse niveles incrementados de producción de ácido succínico con una célula eucariota recombinante seleccionada del grupo que consiste en una levadura y un hongo filamentosos que comprende una secuencia de nucleótidos que codifica la fumarato reductasa dependiente de NAD(H) que cataliza la conversión de ácido fumárico en ácido succínico. Se encontró que la célula eucariota recombinante produce una cantidad incrementada de ácido succínico en comparación con la cantidad de ácido succínico producida por una célula eucariota que no comprende la secuencia de nucleótidos que codifica la fumarato reductasa dependiente de NAD(H).

35 No obstante, sería deseable conseguir niveles incluso más altos de la producción de ácido succínico en un procedimiento fermentativo.

Sumario de la invención

40 La invención se refiere a polipéptidos variantes que tienen actividad de fumarato reductasa (FRD), es decir, a variantes de fumarato reductasa. Una variante de fumarato reductasa de la invención puede tener una o más propiedades modificadas, por ejemplo mejoradas, en comparación con un polipéptido de referencia, teniendo el polipéptido de referencia típicamente actividad de fumarato reductasa, en particular con respecto a actividad dependiente de NADP(H) y/o NAD(H). Esa relación de actividad dependiente de NADP(H):NAD(H) puede modificarse en comparación con un polipéptido de referencia.

45 Un polipéptido de referencia puede ser una fumarato reductasa de tipo salvaje, tal como una fumarato reductasa de tipo salvaje de una fuente de protozoos, tales como *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania braziliensis* o *Leishmania mexicana*. El polipéptido de referencia puede ser una fumarato reductasa (NADH) EC1.3.1.6.

A polipéptidos variantes de la invención puede aludirse como una "variante de fumarato reductasa (FRD)", una "fumarato reductasa mejorada (FRD)" y similares.

De acuerdo con la invención, se proporciona, por tanto, un polipéptido variante que tiene actividad de fumarato reductasa, que tiene actividad dependiente de NADP(H) modificada y/o actividad dependiente de NAD(H) modificada en comparación con un polipéptido de referencia que tiene actividad de fumarato reductasa y que comprende la fumarato reductasa de SEQ ID NO: 33, en donde el polipéptido variante comprende una secuencia de aminoácidos que, cuando se alinea con la fumarato reductasa que comprende la secuencia establecida en SEQ ID NO: 33, comprende al menos una sustitución de un residuo aminoácido correspondiente a cualquiera de los aminoácidos 1042, 1071, 1072, 1082 o 1083, definiéndose dichas posiciones con referencia a SEQ ID NO: 33, y en donde el polipéptido variante tiene al menos 80% de identidad de secuencia con SEQ ID NO: 33. Un polipéptido variante de este tipo es capaz de catalizar la conversión de ácido fumárico en ácido succínico.

Un polipéptido variante puede expresarse en una célula huésped de tal manera que cataliza la conversión de ácido fumárico en ácido succínico. Una célula huésped de este tipo produce una cantidad incrementada de un ácido dicarboxílico, tal como ácido succínico, en comparación con la cantidad de ácido succínico producido por una célula que expresa el polipéptido de referencia.

La invención también proporciona:

- un ácido nucleico que comprende una secuencia que codifica un polipéptido variante de acuerdo con la invención;
- una construcción de ácido nucleico que comprende un ácido nucleico de la invención enlazado operativamente a una o más secuencias de control capaces de dirigir la expresión de una fumarato reductasa en un huésped de expresión adecuado;
- un vector de expresión que comprende un ácido nucleico o una construcción de ácido nucleico de la invención;
- una célula huésped que comprende un ácido nucleico, una construcción de ácido nucleico o un vector de expresión de la invención;
- un método para producir una fumarato reductasa que comprende cultivar una célula huésped de la invención en condiciones adecuadas para la producción de la fumarato reductasa y, opcionalmente, recuperar la fumarato reductasa;
- un método para la producción de un ácido dicarboxílico, tal como ácido succínico, método que comprende fermentar una célula huésped de la invención en condiciones adecuadas para la producción del ácido dicarboxílico, tal como ácido succínico, y, opcionalmente, recuperar el ácido dicarboxílico, tal como ácido succínico.

Breve descripción de los dibujos

La figura 1 presenta una representación esquemática de la integración de los fragmentos 1-8. Las partes rayadas indicadas en los fragmentos 1-8 designan las regiones de solapamiento homólogas únicas que conducen a los eventos de recombinación tal como se indica por las cruces de líneas discontinuas entre las regiones homólogas. El fragmento 1 y el fragmento 8 son homólogos al locus INT59 en el cromosoma XI, la recombinación homóloga da como resultado la integración del fragmento 1-8 en el locus INT59.

La figura 2 presenta una representación esquemática de la integración de los fragmentos 9-12. Las partes rayadas indicadas en los fragmentos 9-12 designan las regiones de solapamiento homólogas únicas que conducen a los eventos de recombinación tal como se indica por las cruces de líneas discontinuas entre las regiones homólogas. El fragmento 9 y el fragmento 12 son homólogos al locus INT1 en el cromosoma XV, la recombinación homóloga da como resultado la integración del fragmento 9-12 en el locus INT1. El fragmento 11 puede reemplazarse por el fragmento 13-22, que contiene las mismas regiones de homología que el fragmento 11.

La Figura 3 muestra la actividad de fumarato reductasa (FRD) dependiente de NADH y NADPH de las variantes de FCC. Se muestra la pendiente del cambio en la absorbancia a 340 nm min^{-1} , que es una medida de la actividad de fumarato reductasa. Se utilizaron cantidades iguales de extractos solubles para testar la actividad dependiente de NADH y NADPH de cada una de las variantes, de modo que el efecto de la mutación sobre la especificidad del cofactor se puede determinar para cada una de las variantes. Está claro que para las variantes de FCC que contenían mutaciones beneficiosas para una producción incrementada de ácido succínico (variantes en el gráfico que figura más adelante), todas exhibieron una especificidad de cofactor significativamente alterada en comparación con la referencia (proteína de tipo salvaje menos SKI C-terminal). En todos los casos, la actividad de FRD dependiente de NADH se reduce significativamente en comparación con la referencia, en la mayoría de los casos más del triple. Sorprendentemente, para un gran número de las variantes de FCC, la especificidad de NADPH aumenta en un 50-100% en comparación con la referencia, lo que indica que los autores de la invención han

producido variantes de fumarato reductasa que exhiben una dependencia significativa de NADPH además de NADH (para la referencia no hay actividad dependiente de NADPH significativa en comparación con la actividad de NADH). La línea de puntos representa la actividad NADPH de fondo de la FRDg de referencia. Figura 3a: figura completa. Figura 3b: el mismo gráfico que en 3a, pero amplía la región inferior del eje y. El marcador corresponde al valor de la actividad de referencia.

La Figura 4 muestra una representación esquemática de la integración de los fragmentos 10, 11, 23 y 24. Las partes rayadas indicadas en los fragmentos 10, 11, 23 y 24 designan las regiones de solapamiento homólogas únicas que conducen a los eventos de recombinación tal como se indica por las cruces de líneas discontinuas entre los regiones homólogas. El fragmento 11 puede sustituirse por el fragmento 13, 16, 17, 19, 21 o 22. El fragmento 23 y el fragmento 24 son homólogos a la región aguas abajo y aguas arriba del locus INT12 en el cromosoma II, 743 pb aguas arriba del ATG del ORF (siglas inglesas de marco de lectura abierto) de YOR071c y 618 pb aguas abajo del ATG de YBL029C-A. La recombinación homóloga da como resultado la integración del fragmento 10, 11 (o 13, 16, 17, 19, 21 o 22), 23 y 24 en el locus INT12.

La Figura 5 muestra la actividad de fumarato reductasa (FRD) dependiente de NADH y NADPH de las variantes de FCC expresadas en la cepa SUC-723. Se muestra la actividad normalizada, determinada como el cambio en la absorbancia a 340 nm (min^{-1}) dividido por la concentración de proteína total (mg mL^{-1}) del extracto celular soluble, que es una medida de la actividad de fumarato reductasa. Se utilizaron cantidades iguales de extractos solubles para testar la actividad dependiente de NADH y NADPH de cada una de las variantes, de modo que el efecto de la mutación sobre la especificidad del cofactor se puede determinar para cada una de las variantes. Está claro que para las variantes de FCC que contenían mutaciones beneficiosas para una producción incrementada de ácido succínico (variantes en el gráfico que figura más adelante), todas exhibieron una especificidad de cofactor significativamente alterada en comparación con la referencia (proteína FRD de tipo salvaje menos SKI C-terminal). Esto se evidencia por el hecho de que en todos los casos la actividad de FRD dependiente de NADH se reduce significativamente en comparación con la referencia, una reducción entre siete y dieciséis veces. Para la variante FCC-034, la especificidad de NADPH aumenta en comparación con la referencia, lo que indica que los autores de la invención han producido una variante de fumarato reductasa en la cepa SUC-723 que exhibe una dependencia significativa de NADPH además de aceptar NADH. FCC_040 ha cambiado completamente la especificidad del cofactor, prefiriendo NADPH frente a NADH, como lo evidencia por su mayor actividad con NADPH en comparación con su actividad con NADH. Sorprendentemente, la alteración de la especificidad del cofactor NAD(P)H de la fumarato reductasa (Figura 5) tiene un efecto beneficioso sobre la producción de ácido succínico de la cepa SUC-723 (Tabla 4). Figura 5a: figura completa. Figura 5b: mismo gráfico que en 5a, pero amplía la región inferior del eje y. El marcador corresponde al valor de la actividad FRD de referencia. La actividad del control SUC-723 (sin presencia de FRD) corresponde a la actividad de fondo en extractos. Todos los valores son medias de mediciones triplo.

La Figura 6 muestra una representación esquemática de la integración de los fragmentos 25 a 28. Las partes rayadas indicadas en los fragmentos 25 a 28 designan las regiones de solapamiento homólogas únicas que conducen a los eventos de recombinación tal como lo indican las cruces de líneas discontinuas entre las regiones homólogas. El fragmento 27 puede sustituirse por el fragmento 29 a 58 individualmente. El fragmento 25 y el fragmento 28 son homólogos a las regiones aguas abajo y aguas arriba del locus INT09.01 en el cromosoma IX, 359 pb aguas abajo del ORF de YIL009W. La recombinación homóloga da como resultado la integración del fragmento 26 y 27 o 29 a 58 individualmente, 23 y 24 en el locus INT09.01.

La Figura 7 muestra la actividad de fumarato reductasa (FRD) dependiente de NADH y NADPH de variantes de FCC expresadas en la cepa CEN.PK113-7D. Se muestra la actividad normalizada, determinada como el cambio en la absorbancia a 340 nm (min^{-1}) dividido por la concentración de proteína total (mg mL^{-1}) del extracto celular soluble, que es una medida de la actividad de fumarato reductasa. Se utilizaron cantidades iguales de extractos solubles para testar la actividad dependiente de NADH y NADPH de cada una de las variantes, de modo que el efecto de la mutación sobre la especificidad del cofactor se puede determinar para cada una de las variantes. Está claro que todas exhibieron una especificidad para el cofactor significativamente alterada en comparación con la referencia (proteína FRD de tipo salvaje menos SKI C-terminal). Esto se evidencia por el hecho de que en todos los casos la actividad de FRD dependiente de NADH se reduce significativamente en comparación con la referencia, una reducción entre cuatro y cincuenta veces. Sorprendentemente, para varias variantes, la especificidad del cofactor se ha cambiado completamente de NADH a NADPH, a saber, FCC_097, 098, 105, 106, 107, 108 y 109. Estas variantes prefieren NADPH frente a NADH como cofactor para la reacción de fumarato reductasa; el incremento de la actividad normalizada específica para NADPH es hasta ~ 11 veces mayor que para la referencia. Sorprendentemente, la posición que parece ser responsable de este cambio completo de especificidad de NADH a NADPH es el residuo 1083. Mutar este residuo, en combinación con los otros residuos, de un residuo aromático voluminoso a un residuo hidrofóbico más pequeño (p. ej., isoleucina o alanina) provoca que la especificidad del

cofactor se cambie por completo. La línea de puntos indica la actividad FRD de referencia (fondo) específica para NADPH.

Descripción de la lista de secuencias

- 5 SEQ ID NO: 1 establece la secuencia de nucleótidos del fragmento 2 (Figura 1), que incluye PEP carboxiquinasa del par de codones de *Actinobacillus succinogenes* optimizado para la expresión en *Saccharomyces cerevisiae*.
 SEQ ID NO: 2 establece la secuencia de nucleótidos del fragmento 3 (Figura 1), que incluye piruvato carboxilasa (PYC2) de un par de codones de *S. cerevisiae* optimizado para la expresión en *S. cerevisiae*.
 SEQ ID NO: 3 establece la secuencia de nucleótidos del fragmento 4 (Figura 1), que incluye un marcador de selección KanMX funcional en *S. cerevisiae*.
- 10 SEQ ID NO: 4 establece la secuencia de nucleótidos del fragmento 5 (Figura 1), que incluye un supuesto transportador de ácido dicarboxílico de un par de codones de *A. niger* optimizado para la expresión en *S. cerevisiae*.
 SEQ ID NO: 5 establece la secuencia de nucleótidos del fragmento 6 (Figura 1), que incluye malato deshidrogenasa (MDH3) de un par de codones de *S. cerevisiae* optimizado para la expresión en *S. cerevisiae*.
 SEQ ID NO: 6 establece la secuencia de nucleótidos del fragmento 7 (Figura 1), que incluye fumarasa (fumB) del par de codones de *Escherichia coli* optimizado para la expresión en *S. cerevisiae*.
- 15 SEQ ID NO: 7 establece la secuencia de nucleótidos del fragmento 10 (Figura 2), que incluye un marcador de selección de nourseotricina funcional en *Saccharomyces cerevisiae*.
 SEQ ID NO: 8 establece la secuencia de nucleótidos del fragmento 11 (Figura 2), que incluye la secuencia codificante para fumarato reductasa del par de codones de *Trypanosoma brucei* (FRDg) optimizado para la expresión en *S. cerevisiae*.
- 20 SEQ ID NO: 9 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 1 (Figura 1).
 SEQ ID NO: 10 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 1 (Figura 1).
 SEQ ID NO: 11 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 2 (Figura 1).
 SEQ ID NO: 12 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 2 (Figura 1).
- 25 SEQ ID NO: 13 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 3 (Figura 1).
 SEQ ID NO: 14 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 3 (Figura 1).
 SEQ ID NO: 15 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 4 (Figura 1).
 SEQ ID NO: 16 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 4 (Figura 1).
 SEQ ID NO: 17 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 5 (Figura 1).
- 30 SEQ ID NO: 18 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 5 (Figura 1).
 SEQ ID NO: 19 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 6 (Figura 1).
 SEQ ID NO: 20 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 6 (Figura 1).
 SEQ ID NO: 21 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 7 (Figura 1).
 SEQ ID NO: 22 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 7 (Figura 1).
- 35 SEQ ID NO: 23 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 8 (Figura 1).
 SEQ ID NO: 24 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 8 (Figura 1).
 SEQ ID NO: 25 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 9 (Figura 2).
 SEQ ID NO: 26 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 9 (Figura 2).
 SEQ ID NO: 27 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 10 (Figura 2).
- 40 SEQ ID NO: 28 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 10 (Figura 2).
 SEQ ID NO: 29 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 12 (Figura 2).
 SEQ ID NO: 30 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 12 (Figura 2).
 SEQ ID NO: 31 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 11 (referencia FRDg) o los fragmentos 13-22 (variantes de FRD) (Figura 2).
- 45 SEQ ID NO: 32 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 11 (referencia FRDg) o los fragmentos 13-22 (variantes de FRD) (Figura 2).
 SEQ ID NO: 33 establece la secuencia de aminoácidos de la proteína fumarato reductasa de *T. brucei* sin la secuencia SKI C-terminal (referencia FRDg).
 SEQ ID NO: 34 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 23.
- 50 SEQ ID NO: 35 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 23.
 SEQ ID NO: 36 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 24.
 SEQ ID NO: 37 establece la secuencia de nucleótidos del cebador utilizado para generar el fragmento 24.
 SEQ ID NO: 38 establece la secuencia de nucleótidos del fragmento 25 de PCR, que consiste en el flanco de integración 5' para fijar como objetivo el locus INT09.01 de CEN.PK113-7D.
- 55 SEQ ID NO: 39 establece la secuencia de nucleótidos del fragmento 26 de PCR, que incluye un marcador de selección de nourseotricina funcional en *Saccharomyces cerevisiae*.
 SEQ ID NO: 40 establece la secuencia de nucleótidos del fragmento 27 de PCR, que incluye la secuencia codificante para fumarato reductasa del par de codones de *Trypanosoma brucei* (FRDg) optimizado para la expresión en *S. cerevisiae*.

SEQ ID NO: 41 establece la secuencia de nucleótidos del fragmento 28 de PCR, que consiste en el flanco de integración 3' para fijar como objetivo el locus INT09.01 de CEN.PK113-7D.

Descripción detallada de la invención

5 A lo largo de la presente memoria descriptiva y las reivindicaciones adjuntas, las palabras "comprenden", "incluyen" y "que tienen" y variaciones tales como "comprende", "que comprende", "incluye" y "que incluye" se han de interpretar inclusivamente. Es decir, estas palabras están destinadas a transmitir la posible inclusión de otros elementos o números enteros no específicamente reseñados, cuando el contexto lo permita.

10 Los artículos "un" y "una" se utilizan en esta memoria para referirse a uno o a más de uno (es decir, a uno o al menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" puede significar un elemento o más de un elemento.

La vía de TCA reductora contiene dos reacciones que requieren el consumo de poder reductor (p. ej., NADH o NADPH): La reacción de malato deshidrogenasa (reducción de oxalacetato a malato) y la reacción de fumarato reductasa (reducción de fumarato a succinato).

15 *Saccharomyces cerevisiae* tiene dos fumarato reductasas endógenas FRD1 y OSM1. Se encuentran tanto en el citosol como en las mitocondrias, y también se han detectado en fracciones de la membrana plasmática. Ambas utilizan FADH2 como cofactor. Sin embargo, no se conoce una fuente citosólica de FADH2. A diferencia de FADH2, los cofactores que contienen nicotinamida, tales como NADH y NADPH, pueden regenerarse, por ejemplo, mediante la gliceraldehído 3-fosfato deshidrogenasa (NADH) en la glucólisis o la rama oxidativa de la vía de la pentosa fosfato (NADPH).

20 Por consiguiente, la introducción de una fumarato reductasa que utiliza un cofactor que contiene nicotinamida puede ser beneficiosa para la producción fermentativa de un ácido dicarboxílico tal como el ácido succínico.

25 En la actualidad, la única clase conocida de fumarato reductasa utilizando un cofactor que contiene nicotinamida son los de especies de *Trypanosoma* y *Leishmania*, tal como el gen fumarato reductasa de *Trypanosoma brucei* y sus homólogos. La combinación de esta FRD y la MDH de levadura en el citosol de la levadura da como resultado una vía reductora de TCA que utiliza NADH como cofactor.

El cambio del cofactor utilizado en la ruta reductora TCA de NADH a NADPH puede tener varios beneficios que incluyen:

1. fuerza motriz termodinámica más fuerte debido a una mayor relación de concentración intracelular de $\frac{[NADPH]}{[NAD]}$ en comparación con $\frac{[NADH]}{[NAD]}$.
- 30 2. mayor concentración de sustrato en función de las mediciones de las concentraciones de metabolitos ($[NADPH] > [NADH]$)
3. regeneración del cofactor desacoplado con glucólisis; y
4. regeneración de cofactor adicional a través de otras vías, tales como la ruta de la pentosa fosfato.

35 De acuerdo con la invención, se proporciona, por lo tanto, un polipéptido variante que tiene actividad de fumarato reductasa (FRD), que tiene la actividad dependiente de NADP(H) modificada y/o la actividad dependiente de NAD(H) modificada en comparación con un polipéptido de referencia que tiene actividad de fumarato reductasa y que comprende la fumarato reductasa de SEQ ID NO: 33, en donde el polipéptido variante comprende una secuencia de aminoácidos que, cuando se alinea con la fumarato reductasa que comprende la secuencia establecida en SEQ ID NO: 33, comprende al menos una sustitución de un residuo aminoácido correspondiente a cualquiera de los aminoácidos

1042, 1071, 1072, 1082 o 1083, estando definidas dichas posiciones con referencia a SEQ ID NO: 33, y

en donde la variante de polipéptido tiene al menos un 80% de identidad de secuencia con SEQ ID NO: 33.

45 Una variante de polipéptido de la invención tiene actividad de fumarato reductasa. La actividad de fumarato reductasa es la actividad que convierte ácido fumárico en ácido succínico:

Succinato + aceptor \Leftrightarrow fumarato + aceptor reducido

Una variante de polipéptido de la invención tiene actividad dependiente de NAD(H) modificada en comparación con un polipéptido de referencia que tiene actividad de FRD.

5 Una variante de polipéptido de este tipo puede tener una actividad dependiente de NAD(H) disminuida en comparación con el polipéptido de referencia.

Una variante de polipéptido de este tipo puede tener una actividad dependiente de NADP(H) incrementada en comparación con el polipéptido de referencia.

La invención también proporciona una variante de polipéptido que tiene actividad de FRD que tiene actividad dependiente de NADP(H) modificada en comparación con un polipéptido de referencia que tiene actividad de FRD.

10 Una variante de polipéptido de este tipo puede tener una actividad dependiente de NADP(H) incrementada en comparación con el polipéptido de referencia.

Una variante de polipéptido de este tipo puede tener una actividad dependiente de NAD(H) incrementada en comparación con el polipéptido de referencia.

15 Una variante de polipéptido de acuerdo con la invención puede ser un polipéptido que se produce de forma no natural y/o que puede ser codificada por una secuencia polinucleotídica que se produce de forma no natural.

20 La expresión "dependiente de NADPH" en esta memoria se refiere típicamente a la propiedad de una enzima de utilizar preferentemente NADPH en lugar de NADH, tal como el cofactor redox. Por lo tanto, una enzima dependiente de NADPH tiene típicamente una actividad más alta, por ejemplo una constante de especificidad más alta (k_{cat}/K_M), con el cofactor NADPH que con el cofactor NADH, por ejemplo según se determina por un ensayo de actividad enzimática tal como se describe en los Ejemplos.

25 La expresión "dependiente de NADH" en esta memoria se refiere típicamente a la propiedad de una enzima de utilizar preferentemente NADH en lugar de NADPH, tal como el cofactor redox. Por lo tanto, una enzima dependiente de NADH tiene típicamente una actividad más alta, por ejemplo una constante de especificidad más alta (k_{cat}/K_M), con el cofactor NADH que con el cofactor NADPH, por ejemplo según se determina por un ensayo de actividad enzimática tal como se describe en los Ejemplos.

30 La actividad de un polipéptido variante de la invención puede determinarse tal como se establece en el Ejemplo 4. $V_{máx}$ y K_M de las enzimas a NADH y NADPH pueden determinarse a partir de v_0 utilizando un gráfico de Lineweaver-Burk (Lineweaver, H y Burk, D. (1934), "The Determination of Enzyme Dissociation Constants". Journal of the American Chemical Society 56 (3): 658-666), en que la concentración de NADH o NADPH varía, por ejemplo, de 25 a 400 μ M.

Una variante de polipéptido de la invención puede mostrar un aumento en la actividad de NADP(H) en función de la actividad de NAD(H) en comparación con un polipéptido de referencia. Es decir, una variante de polipéptido puede mostrar una disminución en la relación de actividad de NADP(H) a NAD(H) en comparación con un polipéptido de referencia.

35 En esta memoria, a las variantes de polipéptido de la invención se las puede aludir como una "variante FRD", "variante de polipéptido FRD", "variante", "variante de polipéptido" o "FCC" o "polipéptido FCC" o similar.

40 Una variante de polipéptido FRD de la invención (por ejemplo, una variante que tiene una o más sustituciones tal como se establece en esta memoria) tiene una identidad de al menos el 80% con el polipéptido FRD de referencia de SEQ ID NO: 33, por ejemplo una identidad de al menos aproximadamente el 85% con el polipéptido de referencia, tal como una identidad de al menos aproximadamente el 90% con el polipéptido de referencia, una identidad de al menos aproximadamente el 95% con el polipéptido de referencia, una identidad de al menos aproximadamente el 98% con el polipéptido de referencia o una identidad de al menos aproximadamente el 99% con el polipéptido de referencia. Una variante de este tipo tendrá típicamente una o más sustituciones o conjuntos de sustituciones como se establecen en una cualquiera de las Tablas 1, 2, 4, 5 o 6.

45 Una variante de FRD de la invención retendrá típicamente la actividad de FRD. Es decir, una variante de FRD de la invención típicamente será capaz de catalizar la reacción arriba establecida, aunque con una especificidad de cofactor modificada en comparación con el polipéptido de referencia.

- 5 Preferiblemente, un polipéptido variante de FRD de la invención exhibirá típicamente propiedades mejoradas en comparación con el polipéptido de referencia del cual se deriva, típicamente en términos de especificidad de co-factor modificado. Una propiedad mejorada de este tipo será típicamente una que sea relevante si la variante se utilizara como se establece más adelante, por ejemplo, en un método para la producción de un ácido dicarboxílico (mediante la expresión de la FRD).
- 10 Por lo tanto, una variante de FRD de la invención es una que es típicamente capaz de aumentar la producción de un ácido dicarboxílico en un microorganismo recombinante capaz de producir dicho ácido dicarboxílico. Es decir, la sobre-expresión de un polipéptido variante de FRD de la invención en una célula huésped conducirá típicamente a una producción incrementada de un ácido dicarboxílico en comparación con una célula huésped que sobre-expresa el polipéptido de referencia.
- 15 Una variante de FRD que exhibe una propiedad que está mejorado en relación con la FRD de referencia es una que demuestra una reducción o aumento mensurable de la propiedad relevante, es decir, actividad dependiente de NAD(H) o NADP(H), típicamente de modo que la variante de FRD es más adecuada para uso tal como se establece más adelante, por ejemplo, en un método para la producción de un ácido dicarboxílico.
- 20 Por lo tanto, el aminoácido presente en una o más de dichas posiciones será reemplazado por un aminoácido diferente que aparece en esa posición en la secuencia de referencia (estando definidas las posiciones con referencia a SEQ ID NO: 33).
- La sustitución en la posición 1042 (según se define con referencia a SEQ ID NO: 33) será típicamente a un aminoácido polar, por ejemplo un aminoácido no cargado negativamente, tal como un aminoácido cargado positivamente. Aminoácidos cargados positivamente adecuados incluyen arginina (R), lisina (K) e histidina (H). Un aminoácido polar adecuado es glutamina (Q).
- 25 La sustitución en la posición 1071 (según se define con referencia a SEQ ID NO: 33) será típicamente a un pequeño aminoácido. Aminoácidos pequeños adecuados incluyen treonina (T), serina (S), glicina (G), alanina (A), prolina (P) y aspartato (D).
- La sustitución en la posición 1082 (según se define con referencia a SEQ ID NO: 33) será típicamente a un aminoácido cargado positivamente. Aminoácidos cargados positivamente adecuados incluyen arginina (R), lisina (K) e histidina (H).
- 30 La sustitución en la posición 1083 (según se define con referencia a SEQ ID NO: 33) será típicamente a un residuo hidrofóbico más pequeño. Aminoácidos más pequeños adecuados incluyen isoleucina (I) o alanina (A).
- Los diversos tipos de aminoácidos de arriba se clasifican con referencia a, por ejemplo, Betts y Russell, *En Bioinformatics for Geneticists*, Barnes y Gray eds., Wiley 2003.
- En más detalle, un polipéptido variante puede comprender:
- 35 R, K o Q en la posición 1042 tal como se define con referencia a SEQ ID NO: 33;
T o S en la posición 1071 tal como se define con referencia a SEQ ID NO: 33;
K en la posición 1072 tal como se define con referencia a SEQ ID NO: 33;
K o R en la posición 1082 tal como se define con referencia a SEQ ID NO: 33; o
Y, I o A en la posición 1083 tal como se define con referencia a SEQ ID NO: 33.
- 40 Un polipéptido variante de este tipo puede ser modificado de modo que la propiedad modificada es actividad dependiente de NAD(H) modificada, tal como una actividad dependiente de NAD(H) disminuida y/o actividad dependiente de NADP(H) modificada, tal como una actividad dependiente de NAD(H) incrementada.
- Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes, que comprende sustituciones adicionales distintas de las cinco posiciones arriba definidas, por ejemplo, una o más sustituciones, adiciones o deleciones adicionales.
- 45 Una variante de la invención puede comprender una combinación de diferentes tipos de modificación de esta clase. Una variante puede comprender una, dos, tres, cuatro, al menos 5, al menos 10, al menos 15, al menos 20, al menos 25, al menos 30 o más de modificaciones de este tipo (que pueden ser del mismo tipo o pueden ser diferentes tipos de modificación). Típicamente, las modificaciones adicionales pueden ser sustituciones.

Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes tal como se define con referencia a la Tabla 1, 2, 4, 5 o 6. Es decir, un polipéptido variante puede comprender cualquier combinación de sustituciones tal como se establece en la Tabla 1, 2, 4, 5 o 6 en comparación con una secuencia de referencia adecuada tal como la establecida en SEQ ID NO: 33.

- 5 Típicamente, entonces un polipéptido variante puede comprender la secuencia de SEQ ID NO: 33, pero en una o más de las posiciones 1042, 1071, 1072, 1082 o 1083 un aminoácido diferente estará presente en esa posición que está presente en SEQ ID NO: 33 (es decir, el polipéptido variante comprende una sustitución en una o más de las posiciones 1042, 1071, 1072 o 1083). Es decir, el polipéptido variante puede tener un aminoácido diferente al glutamato en la posición 1042 y/o un aminoácido diferente a la asparagina en la posición 1071 y/o un aminoácido diferente a la arginina en la posición 1072 y/o un aminoácido distinto de glicina en la posición 1082 y/o un aminoácido que no sea fenilalanina en la posición 1083.

15 Tabla 1: Fumarato reductasa mutada, que contiene mutaciones en comparación con la secuencia de referencia (SEQ ID NO: 33) en las posiciones de aminoácidos indicadas a continuación. Estas secuencias pueden ser útiles para aumentar adicionalmente los títulos de ácido succínico en comparación con el uso de la secuencia de referencia (SEQ ID NO: 33).

Clon	Mutación en posición				
	1042	1071	1072	1082	1083
Referencia	E	N	R	G	F
FCC_001	E	S	R	G	F
FCC_002	E	S	R	G	Y
FCC_003	E	S	R	K	F
FCC_004	E	S	R	K	Y
FCC_005	E	S	R	R	F
FCC_006	E	S	R	R	Y
FCC_007	E	S	K	G	F
FCC_008	E	S	K	G	Y
FCC_009	E	S	K	K	F
FCC_010	E	S	K	K	Y
FCC_011	E	S	K	R	F
FCC_012	E	S	K	R	Y
FCC_013	E	T	R	G	F
FCC_014	E	T	R	G	Y
FCC_015	E	T	R	K	F
FCC_016	E	T	R	K	Y
FCC_017	E	T	R	R	F
FCC_018	E	T	R	R	Y
FCC_019	E	T	K	G	F
FCC_020	E	T	K	G	Y
FCC_021	E	T	K	K	F
FCC_022	E	T	K	K	Y
FCC_023	E	T	K	R	F
FCC_024	E	T	K	R	Y
FCC_025	R	S	R	G	F
FCC_026	R	S	R	G	Y
FCC_027	R	S	R	K	F
FCC_028	R	S	R	K	Y
FCC_029	R	S	R	R	F
FCC_030	R	S	R	R	Y
FCC_031	R	S	K	G	F
FCC_032	R	S	K	G	Y
FCC_033	R	S	K	K	F
FCC_034	R	S	K	K	Y
FCC_035	R	S	K	R	F
FCC_036	R	S	K	R	Y
FCC_037	R	T	R	G	F
FCC_038	R	T	R	G	Y
FCC_039	R	T	R	K	F

ES 2 790 573 T3

FCC_041	R	T	R	R	F
FCC_042	R	T	R	R	Y
FCC_043	R	T	K	G	F
FCC_044	R	T	K	G	Y
FCC_047	R	T	K	R	F
FCC_049	K	S	R	G	F
FCC_050	K	S	R	G	Y
FCC_051	K	S	R	K	F
FCC_052	K	S	R	K	Y
FCC_053	K	S	R	R	F
FCC_054	K	S	R	R	Y
FCC_055	K	S	K	G	F
FCC_056	K	S	K	G	Y
FCC_057	K	S	K	K	F
FCC_058	K	S	K	K	Y
FCC_059	K	S	K	R	F
FCC_060	K	S	K	R	Y
FCC_061	K	T	R	G	F
FCC_062	K	T	R	G	Y
FCC_063	K	T	R	K	F
FCC_064	K	T	R	K	Y
FCC_066	K	T	R	R	Y
FCC_067	K	T	K	G	F
FCC_068	K	T	K	G	Y
FCC_071	K	T	K	R	F
FCC_072	K	T	K	R	Y
FCC_073	Q	S	R	G	F
FCC_074	Q	S	R	G	Y
FCC_077	Q	S	R	R	F
FCC_079	Q	S	K	G	F
FCC_080	Q	S	K	G	Y
FCC_081	Q	S	K	K	F
FCC_082	Q	S	K	K	Y
FCC_083	Q	S	K	R	F
FCC_084	Q	S	K	R	Y
FCC_085	Q	T	R	G	F
FCC_086	Q	T	R	G	Y
FCC_087	Q	T	R	K	F
FCC_088	Q	T	R	K	Y
FCC_089	Q	T	R	R	F
FCC_090	Q	T	R	R	Y
FCC_091	Q	T	K	G	F
FCC_092	Q	T	K	G	Y
FCC_093	Q	T	K	K	F
FCC_094	Q	T	K	K	Y
FCC_095	Q	T	K	R	F
FCC_096	Q	T	K	R	Y
FCC_097	R	G	R	R	Y
FCC_098	Q	G	R	R	Y
FCC_099	R	G	R	K	Y
FCC_100	Q	G	R	K	Y
FCC_101	R	S	R	R	I
FCC_102	R	G	R	R	I
FCC_103	Q	S	R	R	I
FCC_104	Q	G	R	R	I
FCC_105	R	S	R	K	I
FCC_106	R	G	R	K	I
FCC_107	Q	S	R	K	I
FCC_108	Q	G	R	K	I

FCC_109	R	S	R	R	A
FCC_110	R	G	R	R	A
FCC_111	Q	S	R	R	A
FCC_112	Q	G	R	R	A
FCC_113	R	S	R	K	A
FCC_114	R	G	R	K	A
FCC_115	Q	S	R	K	A
FCC_116	Q	G	R	K	A
FCC_117	R	S	R	R	S
FCC_118	R	G	R	R	S
FCC_119	Q	S	R	R	S
FCC_120	Q	G	R	R	S
FCC_121	R	S	R	K	S
FCC_122	R	G	R	K	S
FCC_123	Q	S	R	K	S
FCC_124	Q	G	R	K	S

5 Un polipéptido variante tendrá típicamente una actividad de FRD modificada en comparación con el polipéptido FRD de referencia. Típicamente, la actividad modificada puede definirse en términos de dependencia de co-factor modificada. Esta actividad dependiente de NAD(H) o NADP(H) puede modificarse, por ejemplo disminuirse (en el caso de la actividad dependiente de NAD(H)), en al menos 10%, al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95% o al menos 99%. Alternativamente, la propiedad puede aumentarse (en el caso de la actividad dependiente de NAHP(H)) en al menos 10%, al menos 25%, al menos 50%, al menos 100%, al menos, 200%, al menos 500 % o al menos 1000%.

10 El porcentaje de disminución o incremento en este contexto representa el porcentaje de disminución o incremento en comparación con el polipéptido de FRD de referencia. Es bien sabido por la persona experta cómo se pueden medir este tipo de cambios porcentuales - es una comparación de la actividad, por ejemplo, actividad dependiente de NAD(H) o NADP(H), de la FRD de referencia y la FRD variante medida como se establece en el Ejemplo o de acuerdo con, por ejemplo, Miura A. et al, *J. Bacteriol* **190**: 7170-717.

15 La actividad modificada puede definirse en términos de un incremento en la producción de un ácido dicarboxílico cuando una FRD variante se sobre-expresa en una célula huésped en comparación con el nivel de producción de una célula huésped equivalente que sobre-expresa el polipéptido de referencia. Una FRD variante puede ser capaz de incrementar los niveles de producción en al menos 5%, al menos 10%, al menos 25%, al menos 50%, al menos 100% o más. El nivel de producción puede expresarse en términos de g/L, por lo que un aumento en el nivel de producción de un ácido dicarboxílico será evidente por un mayor nivel de producción en términos de g/L.

20 La palabra "polipéptido" se utiliza en esta memoria para cadenas que contienen más de aproximadamente siete residuos aminoácidos. Todas las secuencias de polipéptidos en esta memoria se escriben de izquierda a derecha y en la dirección del extremo amino al extremo carboxi. El código de una letra de los aminoácidos utilizados en esta memoria se conoce comúnmente en la técnica y se puede encontrar en Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2ª, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

30 Un polipéptido de variante de FRD de la invención puede estar en forma aislada, tal como en forma sustancialmente aislada. Por polipéptido o proteína "aislado" se entiende un polipéptido o proteína eliminado de su entorno nativo. Por ejemplo, los polipéptidos y proteínas producidos de forma recombinante expresados en células huésped se consideran aislados para el propósito de la invención, al igual que los polipéptidos recombinantes que se han purificado sustancialmente por cualquier técnica adecuada. Un polipéptido variante de FRD de acuerdo con la invención puede recuperarse y purificarse a partir de cultivos de células recombinantes por métodos conocidos en la técnica.

35 Polipéptidos de variantes de FRD de la presente invención incluyen productos de procedimientos químicos sintéticos y productos producidos por técnicas recombinantes de un huésped procarionota o eucariota, que incluyen, por ejemplo, células bacterianas, de levadura, de plantas superiores, de insectos y de mamíferos. Dependiendo del huésped empleado en un procedimiento de producción recombinante, los polipéptidos de la presente invención pueden estar glicosilados o pueden no estar glicosilados. Además, los polipéptidos de la invención también pueden incluir un residuo metionina modificado inicial, en algunos casos como resultado de procesos mediados por el huésped.

La invención también presenta fragmentos biológicamente activos de variantes del polipéptido de FRD de acuerdo con la invención. Se considera que estos fragmentos están incluidos dentro de la expresión "una variante de FRD de la invención".

5 Los fragmentos biológicamente activos de una variante de polipéptido de FRD de la invención incluyen polipéptidos que comprenden secuencias de aminoácidos suficientemente idénticas a o derivadas de la secuencia de aminoácidos de una proteína variante de la invención, que incluyen menos aminoácidos que la proteína de longitud completa, pero que exhiben al menos una actividad biológica de la proteína de longitud completa correspondiente. Típicamente, fragmentos biológicamente activos comprenden un dominio o motivo con al menos una actividad de una proteína variante de la invención. Un fragmento biológicamente activo de una variante FRD de la invención
10 puede ser un polipéptido que tiene, por ejemplo, 10, 25, 50, 100 o más aminoácidos de longitud. Además de ello, otras porciones biológicamente activas, en las que se suprimen otras regiones de la proteína, pueden prepararse mediante técnicas recombinantes y evaluarse para una o más de las actividades biológicas de la forma nativa de un polipéptido de la invención.

15 Típicamente, un fragmento de proteína de una variante de FRD de la invención comprenderá una o más de las sustituciones definidas en esta memoria.

La invención también presenta fragmentos de ácidos nucleicos que codifican los fragmentos biológicamente activos de arriba (fragmentos biológicamente activos que son por sí mismos variantes de la invención).

20 Preferiblemente, una variante de FRD de la invención carece de una señal fijadora de objetivo peroxisomal o mitocondrial para la actividad citosólica de la enzima tras la expresión de la secuencia de nucleótidos codificante en una célula huésped adecuada.

25 La presente invención proporciona polinucleótidos que comprenden la secuencia que codifica un polipéptido variante de FRD de la invención (y fragmentos biológicamente activos de los mismos). La invención también se refiere a un polinucleótido aislado que codifica al menos un dominio funcional de una variante de polipéptido de FRD de la invención. Típicamente, un dominio de este tipo comprenderá una o más de las sustituciones descritas en esta memoria.

30 Una molécula de ácido nucleico de la presente invención se pueden generar utilizando técnicas estándares de biología molecular bien conocidas por los expertos en la técnica, tomadas en combinación con la información de la secuencia proporcionada en esta memoria. Por ejemplo, utilizando técnicas sintéticas estándares, la molécula de ácido nucleico requerida puede generarse por PCR o sintetizarse *de novo*. Un procedimiento sintético de este tipo será típicamente un procedimiento automatizado.

Un ácido nucleico de la invención puede comprender una o más deleciones, es decir, huecos, en comparación con un ácido nucleico que codifica una FRD de referencia. Deleciones/huecos de este tipo también pueden generarse utilizando la mutagénesis dirigida al sitio utilizando oligonucleótidos apropiados. Técnicas para generar este tipo de deleciones son bien conocidas por los expertos en la técnica.

35 Además, los oligonucleótidos que corresponden a o que se hibridan con secuencias de nucleótidos de acuerdo con la invención se pueden preparar mediante técnicas sintéticas estándares, p. ej., utilizando un sintetizador de ADN automatizado.

40 También, ácidos nucleicos complementarios y ácidos nucleicos antisentido se incluyen en la presente invención. Una molécula de ácido nucleico que es complementaria a otra secuencia de nucleótidos es una que es suficientemente complementaria a la otra secuencia de nucleótidos, de modo que pueda hibridarse con la otra secuencia de nucleótidos formando así un dúplex estable.

45 Un aspecto de la invención pertenece a moléculas de ácido nucleico aisladas que codifican un polipéptido variante de la invención, o un fragmento biológicamente activo o dominio del mismo, así como moléculas de ácido nucleico suficientes para uso como sondas de hibridación para identificar moléculas de ácido nucleico que codifican un polipéptido de la invención y fragmentos de este tipo de moléculas de ácido nucleico adecuadas para uso como cebadores de PCR para la amplificación o mutación de moléculas de ácido nucleico, tal como para la preparación de moléculas de ácido nucleico de la invención.

50 Un "ácido nucleico aislado" o "polinucleótido aislado" es un ADN o ARN que no está inmediatamente contiguo a ambas de las secuencias codificantes con las que está inmediatamente contiguo (una en el extremo 5' y una en el extremo 3') en el genoma que se produce de forma natural del organismo del que se deriva. Por lo tanto, en una realización, un ácido nucleico aislado incluye algunas o todas las secuencias 5' no codificantes (p. ej., promotor) que

son inmediatamente contiguas a la secuencia codificante. Por lo tanto, el término incluye, por ejemplo, un ADN recombinante que se incorpora en un vector, en un plásmido o virus que se replica de forma autónoma, o en el ADN genómico de un procarionte o eucarionte, o que existe como una molécula separada (p. ej., un ADNc o un fragmento de ADN genómico producido por PCR o tratamiento de endonucleasa de restricción) independiente de otras secuencias. También incluye un ADN recombinante que es parte de un gen híbrido que codifica un polipéptido adicional que está sustancialmente libre de material celular, material viral o medio de cultivo (cuando se produce mediante técnicas de ADN recombinante), o precursores químicos u otros productos químicos (cuando se sintetizan químicamente). Además de ello, un "fragmento de ácido nucleico aislado" es un fragmento de ácido nucleico que no se produce de forma natural como un fragmento y no se encontraría en el estado natural.

Tal como se utiliza en esta memoria, las expresiones "ácido nucleico", "polinucleótido" o "molécula de ácido nucleico" pretenden incluir moléculas de ADN (p. ej., ADNc o ADN genómico) y moléculas de ARN (p. ej., ARNm) y análogos del ADN o ARN generado utilizando análogos de nucleótidos. La molécula de ácido nucleico puede ser de cadena sencilla o de doble cadena, pero preferiblemente es ADN de doble cadena. El ácido nucleico puede sintetizarse utilizando análogos o derivados de oligonucleótidos (p. ej., nucleótidos de inosina o fosforotioato). Oligonucleótidos de este tipo se pueden utilizar, por ejemplo, para preparar ácidos nucleicos que tienen capacidades de apareamiento de bases alteradas o resistencia incrementada a las nucleasas.

La invención también se refiere a una construcción de ácido nucleico que comprende una secuencia de ácido nucleico que codifica un polipéptido variante de la invención y, enlazadas operativamente a la misma, secuencias de control que permiten la expresión de la secuencia de ácido nucleico en una célula huésped. La construcción de ácido nucleico puede incorporarse en un vector, tal como un vector de expresión y/o en una célula huésped con el fin de efectuar la expresión del polipéptido variante.

A la expresión "construcción de ácido nucleico" se la alude en esta memoria como una molécula de ácido nucleico, ya sea de cadena sencilla o de doble cadena, que se aísla de un gen que se produce de forma natural o, más típicamente, que ha sido modificada para contener segmentos de ácido nucleico que se combinan y yuxtaponen de una manera que de otro modo no existiría en la naturaleza. La expresión construcción de ácido nucleico es sinónima a la expresión "casete de expresión" cuando la construcción de ácido nucleico contiene todas las secuencias de control requeridas para la expresión de una secuencia codificante, en donde dichas secuencias de control están operativamente enlazadas a dicha secuencia codificante.

Tal como se utiliza en esta memoria, la expresión "enlazada operativamente" se refiere a un enlace de elementos de polinucleótidos (o secuencias codificantes o secuencia de ácido nucleico) en una relación funcional. Una secuencia de ácido nucleico está "operativamente enlazada" cuando se dispone en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor o potenciador está operativamente enlazado a una secuencia codificante si afecta a la transcripción de la secuencia codificante.

Tal como se utiliza en esta memoria, el término "promotor" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de uno o más genes, ubicado aguas arriba con respecto a la dirección de la transcripción del sitio de inicio de la transcripción del gen, y es identificado estructuralmente por la presencia de un sitio de unión para ARN polimerasa dependiente de ADN, sitios de inicio de la transcripción y cualesquiera otras secuencias de ADN conocidas por un experto en la técnica. Un promotor "constitutivo" es un promotor que es activo bajo la mayoría de las condiciones medioambientales y de desarrollo. Un promotor "inducible" es un promotor que es activo bajo regulación medioambiental o de desarrollo.

Un promotor que se podría utilizar para lograr la expresión de una secuencia de nucleótidos que codifica una enzima tal como fumarato reductasa dependiente de NAD(H) o cualquier otra enzima introducida en la célula eucariota de la invención, puede ser no nativo a una secuencia de nucleótidos que codifica la enzima que se ha de expresar, es decir, un promotor que es heterólogo a la secuencia de nucleótidos (secuencia codificante) a la que está operativamente enlazada. Preferiblemente, el promotor es homólogo, es decir, endógeno a la célula huésped.

Promotores adecuados en este contexto incluyen promotores naturales, tanto constitutivos como inducibles, así como promotores de ingeniería, que son bien conocidos por la persona experta en la técnica. Promotores adecuados en las células huésped eucariotas pueden ser GAL7, GAL10 o GAL 1, CYC1, HIS3, ADH1, PGL, PH05, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI y AOX1. Otros promotores adecuados incluyen PDC, GPD1, PGK1, TEF1 y TDH.

Habitualmente, una secuencia de nucleótidos que codifica una enzima comprende un terminador. En la presente invención puede utilizarse cualquier terminador que sea funcional en la célula eucariota. Terminadores preferidos se obtienen de genes naturales de la célula huésped. Secuencias de terminadores adecuadas son bien conocidas en la técnica. Preferiblemente, estos terminadores se combinan con mutaciones que evitan la descomposición de ARNm

mediada sin sentido en la célula huésped de la invención (véase, por ejemplo: Shirley et al., 2002, Genetics 161: 1465-1482).

La invención se refiere, además, a un vector, preferiblemente un vector de expresión, que comprende un ácido nucleico o una construcción de ácido nucleico de la invención de la invención (es decir, que comprende la secuencia que codifica un polipéptido variante de FRD de la invención).

Con el fin de facilitar la expresión y/o la traducción de la ISP, la secuencia de ácido nucleico que codifica la ISP puede estar comprendida en un vector de expresión tal que el gen que codifica la ISP está enlazado operativamente a las secuencias de control apropiadas para la expresión y/o traducción *in vitro*, o en una célula huésped de la invención. Es decir, la invención proporciona un vector de expresión que comprende un ácido nucleico o una construcción de ácido nucleico de la invención.

El vector de expresión puede ser cualquier vector (p. ej., un plásmido o virus), que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar la expresión del polinucleótido que codifica el polipéptido variante de FRD. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula en la que se ha de introducir el vector. Los vectores pueden ser plásmidos lineales o circulares cerrados. El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, p. ej., un plásmido, un elemento extracromosómico, un mini-cromosoma o un cromosoma artificial. Si está destinado a utilizarse en una célula huésped de origen fúngico, una construcción de ácido nucleico episomal adecuada puede basarse, p. ej., en los plásmidos de levadura 2 μ o pKD1 (Gleer et al., 1991, Biotechnology 9: 968-975), o los plásmidos AMA (Fierro et al., 1995, Curr Genet. 29:482-489).

Alternativamente, el vector de expresión puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica junto con el o los cromosomas en el o los que ha sido integrado. El vector de clonación integrativo puede integrarse al azar o en un locus diana predeterminado en los cromosomas de la célula huésped. En una realización preferida de la invención, el vector de clonación integrativo comprende un fragmento de ADN, que es homólogo a una secuencia de ADN en un locus diana predeterminado en el genoma de la célula huésped para fijar como objetivo la integración del vector de clonación a este locus predeterminado. Con el fin de fomentar la integración fijada como objetivo, el vector de clonación se lineariza preferiblemente antes de la transformación de la célula. La linearización se realiza preferiblemente de modo que al menos uno, pero preferiblemente cualquier extremo del vector de clonación esté flanqueado por secuencias homólogas al locus diana. La longitud de las secuencias homólogas que flanquean el locus diana es preferiblemente de al menos 20 pb, al menos 30 pb, al menos 50 pb, al menos 0.1 kb, al menos 0,2 kb, al menos 0,5 kb, al menos 1 kb, al menos 2 kb o más larga. La eficacia de la integración fijada como objetivo en el genoma de la célula huésped, es decir, la integración en un locus diana predeterminado, se incrementa por las capacidades aumentadas de recombinación homóloga de la célula huésped.

Las secuencias de ADN flanqueantes homólogas en el vector de clonación, que son homólogas al locus diana, se derivan de un locus altamente expresado, lo que significa que se derivan de un gen que es capaz de un alto nivel de expresión en la célula huésped. Un gen capaz de un alto nivel de expresión, es decir, un gen altamente expresado, se define aquí como un gen, cuyo ARNm puede constituir al menos 0,5% (p/p) del ARNm celular total, p. ej., bajo condiciones inducidas o, alternativamente, un gen cuyo producto génico puede constituir al menos el 1% (p/p) de la proteína celular total o, en el caso de un producto génico secretado, puede secretarse a un nivel de al menos 0,1 g/l.

Una construcción de ácido nucleico o vector de expresión pueden ensamblarse *in vivo* en una célula huésped de la invención y, opcionalmente, integrarse en el genoma de la célula en un solo paso (véase, por ejemplo, el documento WO2013/076280).

Más de una copia de una construcción de ácido nucleico o vector de expresión de la invención puede insertarse en una célula huésped de hongo filamentoso para aumentar la producción del polipéptido variante de FRD (sobre-expresión) codificado por la secuencia de ácido nucleico comprendida dentro de la construcción de ácido nucleico. Esto se puede hacer, preferiblemente, integrando en su genoma dos o más copias del ácido nucleico, más preferiblemente fijando como objetivo la integración del ácido nucleico en un locus altamente expresado definido como se definió arriba.

Los expertos en la técnica apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula huésped a transformar, el nivel de expresión de la proteína deseada, etc. Los vectores de expresión de la invención puede introducirse en células huésped para producir con ello proteínas o péptidos codificados por ácidos nucleicos tal como se describe en esta memoria (p. ej., una variante de FRD de SEQ ID NO:

33, por ejemplo, un equivalente o fragmento funcional, o una proteína de fusión que comprende uno o más de variantes de este tipo).

Las construcciones de ácidos nucleicos y los vectores de la invención pueden diseñarse para la expresión de polipéptidos variantes de FRD de la invención en una célula huésped procariota o célula huésped eucariota.

5 Una construcción de ácido nucleico y/o un vector de expresión de la invención pueden introducirse en células procariotas o eucariotas mediante técnicas de transformación o transfección convencionales. Tal como se utiliza en esta memoria, los términos "transformación" y "transfección" pretenden referirse a una diversidad de técnicas reconocidas en la técnica para introducir ácido nucleico extraño (p. ej., ADN) en una célula huésped bien conocida por los expertos en la técnica. Se pueden encontrar métodos adecuados para transformar o transfectar células huésped en Sambrook, et al. (Molecular Cloning: A Laboratory Manual, 2ª, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., Basic Methods in Molecular Biology (1986) y en otros manuales de laboratorio.

15 Las expresiones "equivalentes funcionales" y "variantes funcionales" se utilizan indistintamente en esta memoria. Equivalentes funcionales de acuerdo con la invención son fragmentos de ácido nucleico aislados que codifican un polipéptido que exhibe una función particular de una variante de FRD tal como se define en esta memoria. Los equivalentes funcionales, por lo tanto, también abarcan fragmentos biológicamente activos y quedan abarcados dentro de la expresión "una variante de FRD" de la invención.

20 Preferiblemente, un equivalente funcional de la invención comprende una o más de las sustituciones descritas en esta memoria. Sin embargo, un equivalente funcional puede comprender una o más modificaciones, además de las sustituciones arriba descritas.

Equivalentes de ácido nucleico funcionales pueden contener típicamente mutaciones silenciosas o mutaciones que no alteran la función biológica del polipéptido variante de FRD codificado. Por consiguiente, la invención proporciona moléculas de ácido nucleico que codifican una proteína variante de FRD que contiene cambios en los residuos aminoácidos que no son esenciales para una actividad biológica particular, es decir, actividad de FRD.

25 Equivalentes funcionales de este tipo de las proteínas variantes de FRD difieren en la secuencia de aminoácidos de la secuencia variante de FRD original de la que se derivan pero conservan al menos una actividad biológica de la misma, preferiblemente retienen al menos actividad de FRD. La persona experta reconocerá que pueden introducirse cambios por mutación en las secuencias de nucleótidos de acuerdo con la invención, lo que conduce a cambios en la secuencia de aminoácidos de la proteína resultante sin alterar sustancialmente la función de dicha proteína.

30 En una realización, la molécula de ácido nucleico aislada comprende una secuencia de nucleótidos que codifica una proteína, en donde la proteína comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% de identidad con la secuencia de aminoácidos de referencia (es decir, SEQ ID NO: 33).

35 Por consiguiente, un equivalente funcional de una variante de FRD de la invención es preferiblemente una proteína que comprende una secuencia de aminoácidos que tiene al menos 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% o más de identidad con la secuencia de polipéptidos de referencia de SEQ ID NO: 33, y típicamente también retiene al menos una actividad funcional del polipéptido de referencia.

40 Polipéptidos de FRD variantes de la invención pueden identificarse, p. ej., mediante el rastreo de colecciones de mutantes, p. ej., mutantes de sustitución, de un polipéptido de referencia adecuado. Los mutantes candidatos pueden rastrearse en función de su capacidad para aumentar la producción de ácido dicarboxílico, tal como la producción de ácido succínico, cuando se expresan en una célula huésped (en comparación con una célula huésped correspondiente que expresa el polipéptido de referencia).

45 Fragmentos de un ácido nucleico de acuerdo con la invención pueden comprender o consistir en secuencias que no codifican polipéptidos funcionales. Dichos ácidos nucleicos pueden funcionar como sondas o cebadores para una reacción PCR.

50 Los ácidos nucleicos de acuerdo con la invención, independientemente de que codifiquen polipéptidos funcionales o no funcionales se pueden utilizar como sondas de hibridación o cebadores de la reacción en cadena de la polimerasa (PCR). Los usos de las moléculas de ácido nucleico de la presente invención que no codifican un polipéptido que tiene actividad FRD incluyen, entre otros, (1) hibridación *in situ* (p. ej., FISH) a extensiones cromosómicas de metafase para proporcionar la ubicación cromosómica precisa de un gen que codifica FRD tal como se describe en Verma *et al.*, *Human Chromosomes: a Manual of Basic Techniques*, Pergamon Press, Nueva

York (1988); (2) análisis de transferencia Northern para detectar la expresión de ARNm de FRD en tejidos y/o células específicos; y (3) sondas y cebadores que pueden utilizarse como una herramienta de diagnóstico para analizar la presencia de un ácido nucleico hibridable con una sonda o cebador de este tipo en una muestra biológica (p. ej., tejido) dada.

- 5 Variantes de una enzima FRD de referencia dada se pueden obtener mediante el siguiente procedimiento estándar:
- Mutagénesis (propensa a errores, oligo dopada, oligo reforzada) o síntesis de variantes
 - Transformación en, por ejemplo, *S. cerevisiae*
 - Cultivo de transformantes, selección de transformantes
 - Expresión en, por ejemplo, *S. cerevisiae*
- 10
- Rastreo primario, por ejemplo sobre la base de la producción de ácido dicarboxílico
 - Identificación de una variante mejorada (por ejemplo, en relación con la especificidad para el co-factor alterada)

En una realización, la invención se refiere a un método para producir una variante de polipéptido de FRD de acuerdo con la invención, método que comprende:

- 15
- a) seleccionar un polipéptido de FRD de referencia que tenga la secuencia establecida en SEQ ID NO: 33;
 - b) sustituir al menos un residuo aminoácido correspondiente a cualquiera de 1042, 1071, 1072, 1082 o 108, definiéndose dichas posiciones con referencia a SEQ ID NO: 33;
 - c) opcionalmente sustituir uno o más aminoácidos adicionales tal como se define en b), en donde el polipéptido variante tiene al menos 80% de identidad de secuencia con SEQ ID NO: 33;
 - d) preparar la variante resultante de las etapas a)-c);
 - e) determinar una propiedad de la variante, por ejemplo tal como se establece en los Ejemplos; y
 - f) seleccionar una variante con una propiedad alterada en comparación con el polipéptido de FRD de referencia.
- 20

- 25 En otra realización, la invención presenta células huésped, p. ej., células huésped transformadas o células huésped recombinantes, que contienen un ácido nucleico, una construcción de ácido nucleico o un vector de la invención. Una "célula huésped" o "célula recombinante" de acuerdo con la invención es típicamente una célula en la que (o en la que un antepasado de la misma) se ha introducido, mediante técnicas de ADN recombinante, un ácido nucleico de acuerdo con la invención, es decir, un ácido nucleico que codifica una FRD de la invención. En el contexto de la presente invención, una "célula huésped" de acuerdo con la invención o un parental de dicha célula huésped puede ser cualquier tipo de célula huésped.
- 30

- Una célula huésped de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde la célula huésped es una célula eucariota o una célula procariota. Por consiguiente, tanto en células procariotas como eucariotas se incluyen, p. ej., bacterias, hongos, levadura, y similares, especialmente preferidas son células de levaduras, por ejemplo, *S. cerevisiae* *K. lactis*. Las células huésped también incluyen, pero no se limitan a líneas celulares de mamíferos, tales como CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38 y líneas celulares del plexo coroideo.
- 35

- La invención proporciona, por lo tanto, un método para producir una FRD, método que comprende cultivar una célula huésped tal como se describe en esta memoria bajo condiciones adecuadas para la producción de la FRD y, opcionalmente, recuperar la FRD. Típicamente, la célula huésped es capaz de producir un ácido dicarboxílico, tal como ácido succínico.
- 40

- La célula huésped puede ser una célula procariota. Preferiblemente, la célula huésped procariota es una célula bacteriana. El término "célula bacteriana" incluye microorganismos tanto Gram-negativos como Gram-positivos. Bacterias adecuadas pueden pertenecer, por ejemplo, a *Mannheimia*, tal como *M. succiniciproducens*, *Actinobacillus*, tal como *A. succinogenes*, *Anaerobiospirillum*, *Bacteroides*, *Succinimonas*, *Escherichia*, tal como *E. coli*.
- 45

- Una célula huésped de acuerdo con la invención puede ser una célula huésped eucariota. Preferiblemente, la célula eucariota es una célula de mamífero, insecto, planta, hongo o alga. Más preferiblemente, la célula eucariota es una célula fúngica. Una célula fúngica adecuada puede pertenecer, por ejemplo, a los géneros *Saccharomyces*, *Aspergillus*, *Penicillium*, *Pichia*, *Kluyveromyces*, *Yarrowia*, *Candida*, *Hansenula*, *Humicola*, *Issatchenkia*, *Torulasporea*, *Trichosporon*, *Brettanomyces*, *Rhizopus*, *Zygosaccharomyces*, *Pachysolen* o *Yamadazyma*. Una célula fúngica puede pertenecer, por ejemplo, a una especie de *Saccharomyces cerevisiae*, *Saccharomyces uvarum*, *Saccharomyces bayanus*, *Aspergillus niger*, *Penicillium chrysogenum*, *Pichia stipidis*, *Kluyveromyces marxianus*, *K. lactis*, *K. thermotolerans*, *Yarrowia lipolyensis*, *Candida sonorensis*, *Candida kruisel*, *C. glabrata*, *Hansenula*
- 50

polymorpha, *Issatchenkia orientalis*, *Torulaspota delbrueckii*, *Brettanomyces bruxellensis*, *Rhizopus oryzae* o *Zygosaccharomyces bailii*. En una realización, una célula fúngica en el procedimiento de la presente invención es una levadura, por ejemplo perteneciente a una *Saccharomyces sp.*, tal como un *Saccharomyces cerevisiae*.

5 Ejemplos de células de levadura huésped específica incluyen *C. sonorensis*, *K. marxianus*, *K. thermotolerans*, *C. methanesorbosa*, *Saccharomyces bulderi* (*S. bulderi*), *I. orientalis*, *C. lambica*, *C. sorboxylosa*, *C. zemplinina*, *C. geochares*, *P. membranifaciens*, *Z. kombuchaensis*, *C. sorbosivorans*, *C. vanderwaltii*, *C. sorbophila*, *Z. bisporus*, *Z. lentus*, *Saccharomyces bayanus* (*S. bayanus*), *D. castellii*, *C. boidinii*, *C. etchellsii*, *K. lactis*, *P. jadinii*, *P. anomala*, *Saccharomyces cerevisiae* (*S. cerevisiae*), *Pichia galeiformis*, *Pichia sp.* YB-4149 (designación NRRL), *Candida ethanolica*, *P. deserticola*, *P. membranifaciens*, *P. fermentans* y *Saccharomycopsis crataegensis* (*S. crataegensis*).
 10 Cepas adecuadas de *K. marxianus* y *C. sonorensis* incluyen las descritas en los documentos WO 00/71738 A1, WO 02/42471 A2, WO 03/049525 A2, WO 03/102152 A2 y WO 03/102201A2. Cepas adecuadas de *I. orientalis* son la cepa ATCC 32196 y la cepa ATCC PTA-6648. En la invención, la célula huésped puede ser un Crabtree negativo como una cepa de tipo salvaje. El efecto Crabtree se define como la aparición de metabolismo fermentativo en condiciones aerobias debido a la inhibición del consumo de oxígeno por parte de un microorganismo cuando se
 15 cultiva a altas tasas de crecimiento específicas (efecto a largo plazo) o en presencia de altas concentraciones de glucosa (efecto a corto plazo). Los fenotipos Crabtree negativos no exhiben este efecto y, por lo tanto, pueden consumir oxígeno incluso en presencia de altas concentraciones de glucosa o con altas tasas de crecimiento.

Además de un ácido nucleico que codifica un polipéptido variante de FRD de la invención, una célula huésped de la invención puede sobre-expresar una secuencia de nucleótidos que comprende una secuencia que codifica una o
 20 más de una piruvato carboxilasa, una fosfoenolpiruvato carboxiquinasa, una malato deshidrogenasa, una fumarasa, una isocitrato liasa, una malato sintasa y un transportador de ácido dicarboxílico. Preferiblemente, cuando una o más de estas enzimas se sobre-expresan, son activas en el citosol.

Por lo tanto, una célula huésped de la invención puede sobre-expresar una secuencia de nucleótidos homóloga o heteróloga adecuada que codifica una enzima endógena y/o heteróloga que cataliza una reacción en la célula que
 25 da como resultado un flujo incrementado hacia un ácido dicarboxílico tal como ácido málico, ácido fumárico y/o ácido succínico.

Una célula huésped de la invención puede sobre-expresar una secuencia de ácido nucleico endógena o heteróloga tal como se describe más adelante en esta memoria.

Una célula huésped puede sobre-expresar una piruvato carboxilasa (PYC) que cataliza la reacción de piruvato a oxaloacetato (EC 6.4.1.1). La piruvato carboxilasa puede, por ejemplo, ser activa en el citosol tras la expresión del gen. La célula huésped puede sobre-expresar una piruvato carboxilasa endógena o heteróloga que se sobre-
 30 expresa.

Preferiblemente, una célula huésped de la invención expresa una secuencia de nucleótidos que codifica una fosfoenolpiruvato (PEP) carboxiquinasa en el citosol. Preferiblemente, se sobre-expresa una secuencia de
 35 nucleótidos que codifica una fosfoenolpiruvato (PEP) carboxiquinasa. La PEP carboxiquinasa (EC 4.1.1.49) es preferiblemente una enzima heteróloga, preferiblemente derivada de bacterias, más preferiblemente la enzima que tiene actividad de PEP carboxiquinasa se deriva de *Escherichia coli*, *Mannheimia sp.*, *Actinobacillus sp.*, o *Anaerobiospirillum sp.*, más preferiblemente *Mannheimia succiniciproducens*. Un gen que codifica una PEP carboxiquinasa puede sobre-expresarse y puede expresarse y ser activo en el citosol de una célula fúngica.
 40 Preferiblemente, una célula de levadura de acuerdo con la presente invención se modifica genéticamente con una PEP carboxiquinasa que tiene al menos 80, 85, 90, 95, 99 o 100% de identidad de secuencia con la secuencia de aminoácidos codificada por la secuencia de ácido nucleico de SEQ ID NO: 1).

En otra realización, la célula huésped de la invención sobre-expresa una piruvato carboxilasa (PYC), que cataliza la reacción de piruvato a oxaloacetato (EC 6.4.1.1). Preferiblemente, la piruvato carboxilasa es activa en el citosol tras
 45 la expresión del gen. Preferiblemente, se sobre-expresa una piruvato carboxilasa endógena u homóloga. Preferiblemente, una célula huésped de acuerdo con la presente invención se modifica genéticamente con una piruvato carboxilasa que tiene al menos 80, 85, 90, 95, 99 o 100% de identidad de secuencia con la secuencia de aminoácidos codificada por la secuencia de ácido nucleico de SEQ ID NO: 2).

En una realización, una célula huésped se modifica con un ácido nucleico que comprende una secuencia que
 50 codifica una malato deshidrogenasa (MDH) activa en el citosol tras la expresión del ácido nucleico. La expresión citosólica se puede obtener por delección de una señal de fijación de objetivo peroxisomal. La malato deshidrogenasa puede ser sobre-expresada. Una MDH citosólica puede ser cualquier malato deshidrogenasa homóloga o heteróloga adecuada, catalizando la reacción de oxaloacetato a malato (EC 1.1.1.37), por ejemplo derivada de *S. cerevisiae*.

Preferiblemente, la MDH es MDH3 de *S. cerevisiae*, más preferiblemente una que tiene una delección SKL C-terminal de manera que es activa en el citosol. Preferiblemente, una célula huésped de acuerdo con la presente invención comprende una secuencia de nucleótidos que codifica una malato deshidrogenasa que tiene al menos 70%, preferiblemente al menos 75, 80, 85, 90, 92, 94, 95, 96, 97, 98, 99% o 100% de identidad de secuencia con la secuencia de aminoácidos codificada por la secuencia de ácido nucleico de SEQ ID NO: 4.

En otra realización, una célula huésped de la presente divulgación se modifica con un gen que codifica una fumarasa, que cataliza la reacción de ácido málico para dar ácido fumárico (EC 4.2.1.2). Un ácido nucleico que comprende una secuencia que codifica una fumarasa puede derivarse de cualquier origen adecuado, preferiblemente de origen microbiano, por ejemplo, una levadura tal como *Saccharomyces* o un hongo filamentoso, tal como *Rhizopus oryzae*, o una bacteria tal como *Escherichia coli*. Una célula huésped de la presente invención puede sobre-expresar una secuencia de nucleótidos que codifica una fumarasa. La fumarasa puede ser activa en el citosol tras la expresión de la secuencia de nucleótidos, por ejemplo eliminando una señal de fijación de objetivo peroxisomal.

Preferiblemente, una célula huésped de la presente invención sobre-expresa una secuencia de nucleótidos que codifica una fumarasa que tiene al menos 70%, preferiblemente al menos 75, 80, 85, 90, 92, 94, 95, 96, 97, 98 o 99% o 100% de identidad de secuencia con la secuencia de aminoácidos codificada por la secuencia de ácido nucleico de SEQ ID NO: 6.

Una célula huésped de acuerdo con la invención puede expresar una secuencia de nucleótidos que codifica una proteína transportadora de ácido dicarboxílico, preferiblemente una proteína transportadora de ácido málico (MAE) en el citosol. Preferiblemente, la proteína transportadora de ácido dicarboxílico se sobre-expresa. Una proteína transportadora de ácido dicarboxílico puede ser cualquier proteína homóloga o heteróloga adecuada. Preferiblemente, la proteína transportadora de ácido dicarboxílico es una proteína heteróloga. Una proteína transportadora de ácido dicarboxílico puede derivarse de cualquier organismo adecuado, preferiblemente de levadura u hongos, tales como *Schizosaccharomyces pombe* o *Aspergillus niger*. Preferiblemente, una proteína transportadora de ácido dicarboxílico es una proteína transportadora de ácido málico (MAE) que tiene al menos 80, 85, 90, 95 o 99% o 100% de identidad de secuencia con el aminoácido codificado por la secuencia de ácido nucleico de SEQ ID NO: 5).

Una célula huésped de la invención puede sobre-expresar un ácido nucleico que comprende una secuencia que codifica una isocitrato liasa (EC 4.1.3.1), que puede ser cualquier enzima heteróloga u homóloga adecuada. La isocitrato liasa puede obtenerse, por ejemplo, de *Kluyveromyces lactis* o *Escherichia coli*.

Una célula fúngica modificada genéticamente puede sobre-expresar adicionalmente un ácido nucleico que comprende una secuencia que codifica una malato sintasa (EC 2.3.3.9). La malato sintasa puede ser sobre-expresada y/o activa en el citosol, por ejemplo, por delección de una señal de fijación de objetivo peroxisomal. En el caso de la malato sintasa es una malato sintasa de *S. cerevisiae*, por ejemplo, la malato sintasa nativa se altera por la delección de la secuencia de SKL carboxi-terminal.

En otra realización, una célula huésped de la invención puede comprender una disrupción de un gen que codifica una enzima de la vía de fermentación de etanol. Un gen que codifica una enzima de una vía de fermentación de etanol puede ser piruvato descarboxilasa (EC 4.1.1.1), que cataliza la reacción de piruvato para dar acetaldehído, o alcohol deshidrogenasa (EC 1.1.1.1), que cataliza la reacción de acetaldehído para dar etanol. Preferiblemente, una célula huésped de la invención comprende una disrupción de uno, dos o más genes que codifican una alcohol deshidrogenasa. En el caso de que la célula fúngica sea una levadura, p. ej., *S. cerevisiae*, la levadura comprende preferiblemente una disrupción de uno o más genes de alcohol deshidrogenasa (*adh1*, *adh2*, *adh3*, *adh4*, *adh5*).

Alternativamente o además, la célula huésped de la invención puede comprender al menos un gen que codifica la glicerol-3-fosfato deshidrogenasa que no es funcional. Un gen glicerol-3-fosfato deshidrogenasa que no es funcional se utiliza en esta memoria para describir una célula eucariota, que comprende una actividad reducida de glicerol-3-fosfato deshidrogenasa, por ejemplo por mutación, disrupción o delección del gen que codifica glicerol-3-fosfato deshidrogenasa, lo que resulta en una formación disminuida de glicerol en comparación con una célula de tipo salvaje.

La expresión citosólica de las enzimas arriba descritas puede obtenerse mediante la delección de una señal de fijación de objetivo peroxisomal o mitocondrial. La presencia de una señal de fijación de objetivo peroxisomal o mitocondrial puede determinarse, por ejemplo, mediante el método descrito por Schlüter *et al.*, Nucleic Acid Research 2007, 35, D815-D822.

5 Tal como se utiliza en esta memoria, una levadura genéticamente modificada de acuerdo con la presente invención se define como una célula que contiene, o se transforma o modifica genéticamente con o una secuencia de nucleótidos o un polipéptido que no se produce de forma natural en la célula de levadura, o contiene una copia adicional o copias adicionales de una secuencia de ácido nucleico endógeno, o contiene una delección o disrupción de una secuencia de nucleótidos endógena u homóloga. Una célula eucariota de tipo salvaje se define en esta memoria como la célula parental de la célula recombinante.

10 Tal como se utilizan en esta memoria, el término "gen" y la expresión "gen recombinante" se refieren a moléculas de ácido nucleico que incluyen un marco de lectura abierto que codifica una variante de FRD u otra enzima tal como se describe en esta memoria. Un gen puede incluir secuencias codificantes, secuencias no codificantes, intrones y secuencias reguladoras. Es decir, un "gen", tal como se utiliza en esta memoria, puede referirse a una molécula de ácido nucleico aislada tal como se define en esta memoria. Por consiguiente, el término "gen", en el contexto de la presente solicitud, no se refiere únicamente a secuencias que se producen de forma natural.

15 El término "endógeno", cuando se utiliza para indicar la relación entre una molécula de ácido nucleico (recombinante) o polipéptido dado y un organismo huésped o célula huésped dado, significa que en la naturaleza la molécula de ácido nucleico o polipéptido es producida por una célula huésped u organismo de la misma especie, preferiblemente de la misma variedad o cepa.

20 El término "heterólogo" cuando se utiliza con respecto a un ácido nucleico (ADN o ARN) o una proteína se refiere a un ácido nucleico o una proteína que no se produce de forma natural como parte del organismo, célula, genoma o secuencia de ADN o ARN en el cual está presente, o que se encuentra en una célula o ubicación o ubicaciones en el genoma o la secuencia de ADN o ARN que difieren de aquellos en la que se encuentra en la naturaleza. Los ácidos nucleicos o las proteínas heterólogos no son endógenos a la célula en la que se introducen, sino que se han obtenido de otra célula o se han producido de forma sintética o recombinante.

25 La invención se refiere a un procedimiento para la producción de un ácido dicarboxílico. La expresión "ácido dicarboxílico" y el término "dicarboxilato", al igual que "ácido succínico" y "succinato", tienen el mismo significado en esta memoria y se utilizan indistintamente, siendo el primero la forma hidrogenada de este último.

De acuerdo con la invención, se proporciona, por lo tanto, un método para la producción de un ácido dicarboxílico, tal como ácido succínico, método que comprende fermentar una célula huésped según se describe en esta memoria bajo condiciones adecuadas para la producción de un ácido dicarboxílico, tal como ácido succínico y, opcionalmente, recuperar el ácido dicarboxílico, tal como el ácido succínico.

30 En el procedimiento, una célula huésped se fermenta en un recipiente que comprende un medio de fermentación adecuado. El término fermentar, fermentación o fermentado, y similares, tal como se utiliza en esta memoria, se refiere a la producción microbiana de compuestos, aquí ácidos dicarboxílicos a partir de hidratos de carbono.

Preferiblemente, el producto de fermentación es un ácido dicarboxílico, preferiblemente ácido málico, ácido fumárico o ácido succínico o ácido adípico, preferiblemente ácido succínico.

35 Una fermentación por lotes se define en esta memoria como una fermentación en la que todos los nutrientes se añaden al comienzo de una fermentación.

Una fermentación semicontinua es una fermentación por lotes en la que se añaden los nutrientes durante la fermentación. Los productos en una fermentación por lotes y semicontinua pueden cosecharse en un momento adecuado, por ejemplo, cuando uno o más nutrientes se agotan.

40 Una fermentación continua es una fermentación en la que los nutrientes se añaden continuamente a la fermentación y en la que los productos se separan continuamente de la fermentación.

45 En una realización, la fermentación de la célula huésped en el procedimiento de la invención se lleva a cabo bajo condiciones limitantes de hidratos de carbono. Tal como se utiliza en esta memoria, las condiciones limitantes de hidratos de carbono se definen como mantener la concentración de hidratos de carbono por debajo de 10 g/l, por ejemplo, aproximadamente 5 g/l.

50 El procedimiento para la producción de ácido dicarboxílico de acuerdo con la presente invención puede llevarse a cabo en cualquier volumen y escala adecuados, preferiblemente a escala industrial. La escala industrial se define en esta memoria como un volumen de al menos 10 o 100 litros, preferiblemente al menos 1 metro cúbico, preferiblemente al menos 10 o 100 metros cúbicos, preferiblemente al menos 1000 metros cúbicos, habitualmente por debajo de 10.000 metros cúbicos.

- 5 La fermentación de la célula huésped en el procedimiento de la invención se puede llevar a cabo en cualquier medio de fermentación adecuado que comprenda una fuente de nitrógeno adecuada, hidratos de carbono y otros nutrientes requeridos para el crecimiento y la producción de un ácido dicarboxílico en el procedimiento de la invención. Un hidrato de carbono adecuado en el procedimiento de fermentación de acuerdo con la invención puede ser glucosa, galactosa, xilosa, arabinosa, sacarosa o maltosa.
- En una realización, el procedimiento de fermentación se lleva a cabo bajo una presión parcial de CO₂ de entre 5% y 60%, preferiblemente de aproximadamente 50%.
- 10 El pH durante el procedimiento para la producción de ácido dicarboxílico disminuye habitualmente durante la producción del ácido dicarboxílico. Preferiblemente, el pH en el procedimiento para la producción de ácido dicarboxílico varía entre 1 y 5, preferiblemente entre 1,5 y 4,5, más preferiblemente entre 2 y 4.
- 15 En otra realización preferida, el procedimiento de acuerdo con la presente invención comprende una etapa de precultivar la célula huésped bajo condiciones aerobias en presencia de un hidrato de carbono. Preferiblemente, la fermentación de la célula huésped durante el precultivo se lleva a cabo a un pH de entre 4 y 6. Preferiblemente, el hidrato de carbono durante el precultivo es un hidrato de carbono no represivo, preferiblemente galactosa. Se ha encontrado ventajoso precultivar células huésped en un hidrato de carbono no represivo, ya que esto evita que se produzca la represión de la glucosa, lo cual puede influir negativamente en la cantidad de biomasa producida. Además, se ha encontrado que una etapa de precultivar células huésped en condiciones aerobias da como resultado un mayor rendimiento de biomasa y un crecimiento más rápido. Preferiblemente, el precultivo se lleva a cabo en modo por lotes.
- 20 Típicamente se lleva a cabo una etapa de propagación para producir biomasa incrementada, preferiblemente bajo condiciones limitantes de hidratos de carbono.
- Se puede llevar a cabo un procedimiento para producir un ácido dicarboxílico a cualquier temperatura adecuada. Una temperatura adecuada puede ser, por ejemplo, entre aproximadamente 10 y aproximadamente 40 grados centígrados, por ejemplo, entre aproximadamente 15 y aproximadamente 30 grados centígrados.
- 25 El procedimiento para la producción de un ácido dicarboxílico puede comprender, además, recuperar el ácido dicarboxílico. La recuperación del ácido dicarboxílico puede llevarse a cabo por cualquier método adecuado.
- 30 En una realización, un ácido dicarboxílico que se produce en un procedimiento tal como se describe en esta memoria se recupera del medio de fermentación. La recuperación de un ácido dicarboxílico puede llevarse a cabo mediante cualquier método adecuado conocido en la técnica, por ejemplo mediante cristalización, precipitación con amonio, tecnología de intercambio iónico, centrifugación o filtración o cualquier combinación adecuada de estos métodos.
- En una realización preferida, la recuperación de ácido dicarboxílico comprende cristalizar el ácido dicarboxílico y formar cristales de ácido dicarboxílico. Preferiblemente, la cristalización del ácido dicarboxílico comprende separar parte del medio de fermentación, preferiblemente por evaporación, para obtener un medio concentrado.
- 35 De acuerdo con la presente invención, un ácido dicarboxílico, tal como ácido succínico, se puede recuperar mediante cristalización del ácido dicarboxílico, tal como ácido succínico, a partir de una solución acuosa que tiene un pH de entre 1 y 5 y que comprende ácido succínico, que comprende evaporar parte de la solución acuosa para obtener una solución concentrada, reducir la temperatura de la solución concentrada a un valor de entre 5 y 35 grados centígrados, en donde se forman cristales de ácido succínico. Preferiblemente, la cristalización comprende llevar la temperatura del medio concentrado a una temperatura de entre 10 y 30 grados centígrados, preferiblemente entre 15 y 25 grados centígrados. Preferiblemente, el medio de fermentación tiene un pH de entre 1,5 y 4,5, preferiblemente entre 2 y 4.
- 40 Se ha encontrado que la cristalización de un ácido dicarboxílico, tal como ácido succínico, a temperaturas más altas, tales como entre 10 y 30 grados centígrados da como resultado cristales de un ácido dicarboxílico, tal como el ácido succínico, con una menor cantidad de impurezas tales como ácido orgánico, proteína, color y/u olor, que cristales de un ácido dicarboxílico, tal como ácido succínico, que se cristalizaron a una temperatura baja por debajo de 10 grados.
- 45 Otra ventaja de cristalizar ácido succínico a una temperatura más alta es que requiere una menor cantidad de energía para enfriar la solución acuosa en comparación con un procedimiento en el que la cristalización del ácido dicarboxílico se lleva a cabo por debajo de 10 o 5 grados centígrados, dando como resultado un procedimiento más económico y sostenible.
- 50

Preferiblemente, la cristalización del ácido dicarboxílico, tal como ácido succínico, comprende una etapa de lavar los cristales de ácido dicarboxílico. El ácido dicarboxílico, tal como ácido succínico, puede cristalizarse directamente a partir del medio de fermentación que tiene un pH entre 1 y 5 a una pureza de al menos 90% p/p, preferiblemente al menos 95, 96, 97 o al menos 98%, o 99 a 100% p/p.

- 5 Preferiblemente, la recuperación del ácido dicarboxílico, preferiblemente ácido succínico, comprende separar la biomasa del medio de fermentación y cristalizar el ácido dicarboxílico, preferiblemente cristalizar como se describe arriba en esta memoria. Preferiblemente, la separación de biomasa se lleva a cabo por filtración.

10 En una realización preferida, el procedimiento para la producción de un ácido dicarboxílico comprende, además, utilizar el ácido dicarboxílico en un proceso industrial. Un proceso industrial para un ácido dicarboxílico puede ser la aplicación como un aditivo cosmético, agente de deshielo, aditivo alimentario o como un bloque de construcción para (bio)polímeros.

15 En una realización preferida, el medio de fermentación comprende una cantidad de ácido succínico de entre 1 y 150 g/l, preferiblemente entre 5 y 100 g/l, más preferiblemente entre 10 y 80 g/l o entre 15 y 60 g/l de ácido succínico. En cualquier caso, una célula huésped de la invención será capaz típicamente de acumular más ácido succínico en el medio de fermentación en comparación con una célula huésped equivalente que exprese el polipéptido de referencia.

20 En otro aspecto, la presente invención se refiere a un procedimiento para cristalizar ácido succínico a partir de una solución acuosa que tiene un pH entre 1 y 5 y que comprende ácido succínico, que comprende separar parte de la solución acuosa por evaporación para obtener una solución concentrada, y llevar la temperatura de la solución concentrada a un valor de entre 10 y 30 grados centígrados, en donde se forman cristales de ácido succínico. Preferiblemente, la cristalización comprende llevar la temperatura de la solución concentrada entre 15 y 25 grados centígrados, preferiblemente entre 18 y 22 grados centígrados. Preferiblemente, la solución acuosa tiene un pH de entre 1,5 y 4,5, preferiblemente entre 2 y 4. La solución acuosa puede ser cualquier solución adecuada que comprenda ácido succínico. La solución acuosa puede comprender constituyentes solubles y constituyentes insolubles, tales como (fragmentos de) células microbianas, proteínas, biomasa vegetal, lignocelulosa, celulosa y similares. Preferiblemente, la solución acuosa es un medio de fermentación, preferiblemente un medio de fermentación obtenible por un procedimiento para la producción de un ácido dicarboxílico tal como se describe en esta memoria.

30 Preferiblemente, el ácido dicarboxílico, tal como ácido succínico, que se prepara en el procedimiento de acuerdo con la presente invención se convierte adicionalmente en un producto deseable. Un producto deseable puede ser, por ejemplo, un polímero, tal como ácido polibutileno succínico (PBS), un agente de deshielo o un tensioactivo. Es decir, la invención proporciona un método para la producción de un producto, por ejemplo un polímero, tal como ácido polibutileno succínico (PBS), un agente de deshielo o un tensioactivo, método que comprende: producir un ácido carboxílico tal como se describe en esta memoria; y utilizar dicho ácido dicarboxílico en la producción de dicho producto.

40 Para el propósito de esta invención, se define aquí que, con el fin de determinar el porcentaje de identidad de secuencia de dos secuencias de aminoácidos o de dos secuencias de ácidos nucleicos, las secuencias se alinean para propósitos de comparación óptima. Con el fin de optimizar el alineamiento entre las dos secuencias, se pueden introducir huecos en cualquiera de las dos secuencias que se comparan. Un alineamiento de este tipo puede llevarse a cabo a lo largo de toda la longitud de las secuencias que se comparan. Alternativamente, el alineamiento puede llevarse a cabo a lo largo de una longitud más corta, por ejemplo, más de aproximadamente 20, aproximadamente 50, aproximadamente 100 o más ácidos nucleicos/bases o aminoácidos. La identidad de secuencia es el porcentaje de coincidencias idénticas entre las dos secuencias a lo largo de la región alineada reseñada.

45 Técnicas genéticas estándares, tales como la sobre-expresión de enzimas en las células huésped, la modificación genética de células huésped o técnicas de hibridación, son métodos conocidos en la técnica tales como los descritos en Sambrook y Russel (2001) "Molecular Cloning: A Laboratory Manual (3ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, o F. Ausubel et al., Eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, Nueva York (1987). Métodos para la transformación, la modificación genética, etc. de células huésped fúngicas se conocen, p. ej., de los documentos EP-A-0 635 574, WO 98/46772, WO 99/60102 y WO 00/37671, WO 90/14423, EP-A-0481008, EP-A- 0 635 574 y US 6.265.186.

Una comparación de secuencias y la determinación del porcentaje de identidad de secuencia entre dos secuencias se puede lograr utilizando un algoritmo matemático. La persona experta será consciente del hecho de que están disponibles varios programas de computadora diferentes para alinear dos secuencias y determinar la identidad entre

dos secuencias (Kruskal, J. B. (1983) Una visión general de la comparación de secuencias en D. Sankoff y J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, págs. 1-44 Addison Wesley). El porcentaje de identidad de secuencia entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos puede determinarse utilizando el algoritmo de Needleman y Wunsch para el alineamiento de dos secuencias. (Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). El algoritmo puede alinear tanto las secuencias de aminoácidos como las secuencias de nucleótidos. El algoritmo de Needleman-Wunsch se ha implementado en el programa informático NEEDLE. Para el propósito de esta invención, se utilizó el programa NEEDLE del paquete EMBOSS (versión 2.8.0 o superior, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice, P. Longden, I. y Bleasby, A. Trends in Genetics 16, (6) págs. 276-277, <http://emboss.bioinformatics.nl/>). Para secuencias de proteínas se utiliza EBLOSUM62 para la matriz de sustitución. Para la secuencia de nucleótidos se utiliza EDNAFULL. Los parámetros opcionales utilizados son una penalización de apertura de hueco de 10 y una penalización de extensión de hueco de 0,5. La persona experta apreciará que todos estos parámetros diferentes producirán resultados ligeramente diferentes, pero que el porcentaje de identidad global de dos secuencias no se altera significativamente cuando se utilizan algoritmos diferentes.

Después del alineamiento por el programa NEEDLE tal como se describe arriba, el porcentaje de identidad de secuencia entre una secuencia de consulta y una secuencia de la invención se calcula de la siguiente manera: número de posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico o un nucleótido idéntico en ambas secuencias, dividido por la longitud total del alineamiento después de restar el número total de huecos en el alineamiento. La identidad definida en esta memoria se puede obtener de NEEDLE utilizando la opción NOBRIEF y se marca en la salida del programa como "identidad más larga".

Las secuencias de ácido nucleico y proteínas de la presente invención pueden utilizarse adicionalmente como una "secuencia de consulta" para realizar un rastreo frente a bases de datos públicas, por ejemplo, para identificar otros miembros de la familia o secuencias relacionadas. Dichos rastreos se pueden realizar utilizando los programas NBLAST y XBLAST (versión 2.0) de Altschul, et al. (1990) J. Mol. Biol. 215: 403-10. Los rastreos de nucleótidos BLAST se pueden realizar con el programa NBLAST, puntuación = 100, longitud de palabra = 12 para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico de la invención. Los rastreos de proteínas BLAST se pueden realizar con el programa XBLAST, puntuación = 50, longitud de palabra = 3 para obtener secuencias de aminoácidos homólogas a las moléculas de proteína de la invención. Para obtener alineamientos separados para fines de comparación, se puede utilizar BLAST con huecos tal como se describe en Altschul et al., (1997) Nucleic Acids Res. 25(17): 3389-3402. Cuando se utilizan los programas BLAST y XBLAST con huecos, se pueden utilizar los parámetros por defecto de los respectivos programas (p. ej., XBLAST y NBLAST). Véase la página web del Centro Nacional de Información Biotecnológica en <http://www.ncbi.nlm.nih.gov/>.

Una referencia en esta memoria a un documento de patente u otra materia que se proporciona como técnica anterior no debe tomarse como una admisión de que ese documento o materia era conocido o que la información que contiene era parte del conocimiento general común como la fecha de prioridad de cualquiera de las reivindicaciones.

La divulgación de cada una de las referencias recogidas en esta memoria se incorpora en esta memoria como referencia en su totalidad.

La presente invención se ilustra adicionalmente mediante los siguientes Ejemplos:

EJEMPLOS

Ejemplo 1: Construcción de la cepa SUC-1099

Generación de fragmentos de PCR

Los fragmentos de PCR se generaron utilizando ADN polimerasa Phusion (New England Biolabs, EE.UU.) de acuerdo con las instrucciones del fabricante.

Las secuencias de cebador descritas en SEQ ID NO: 9 y SEQ ID NO: 10 se utilizaron para generar el fragmento 1 de PCR que consiste en el sitio de integración INT59 5', utilizando ADN genómico de la cepa *Saccharomyces cerevisiae* cepa CEN.PK 113-7D (MATa HIS3 LEU2 TRP1 MAL2-8 SUC2), descrito por Daran-Lapujade *et al.*, (FEMS Yeast Res (2003) 4: 285-296) como molde.

El fragmento 2 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 11 y SEQ ID NO: 12, utilizando SEQ ID NO: 1 como molde. La SEQ ID NO: 1 codifica fosfoenolpiruvato carboxiquinasa (PCKa) de *Actinobacillus succinogenes* tal como se describe en la solicitud de patente WO2009/065780. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos los sitios de restricción apropiados, fue

5 sintetizada por DNA 2.0 (Menlo Park, California, EE.UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor TPI1 controla la expresión del gen PCKa. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador GND2.

10 El fragmento 3 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 13 y SEQ ID NO: 14, utilizando SEQ ID NO: 2 como molde. La SEQ ID NO: 2 codifica piruvato carboxilasa (PYC2) de *Saccharomyces cerevisiae* tal como se describe en la solicitud de patente WO2009/065780. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos los sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE. UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor PGK1 controla la expresión del gen PYC2. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador ADH1.

15 El fragmento 4 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 15 y SEQ ID NO: 16, utilizando SEQ ID NO: 3 como molde. La SEQ ID NO: 3 codifica un marcador de selección KanMX funcional en *Saccharomyces cerevisiae* que se amplificó a partir del plásmido pUG7-EcoRV. pUG7-EcoRV es una variante del plásmido pUG6 descrito por Gueldener *et al.*, (Nucleic Acids Res. 1 de julio de 1996; 24(13):2519-24), en el que los sitios loxP presentes en pUG6 se cambiaron a sitios lox66 y lox71 (Lambert *et al.*, Appl. Environ. Microbiol. Febrero de 2007; 73(4):1126-35. Epub 1 de diciembre de 2006).

25 El fragmento 5 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 17 y SEQ ID NO: 18, utilizando SEQ ID NO: 4 como molde. La SEQ ID NO: 4 codifica un supuesto transportador de ácido dicarboxílico de *Aspergillus niger* tal como se describe en el documento EP2495304. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos los sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE.UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor ENO1 controla la expresión del gen DCT_02. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador TEF2.

30 El fragmento 6 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 19 y SEQ ID NO: 20, utilizando SEQ ID NO: 5 como molde. La SEQ ID NO: 5 codifica la malato deshidrogenasa (MDH3) de *Saccharomyces cerevisiae* tal como se describe en la solicitud de patente WO2009/065778. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE.UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor FBA1 controla la expresión del gen MDH3. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador GPM1.

40 El fragmento 7 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 21 y SEQ ID NO: 22, utilizando SEQ ID NO: 6 como molde. La SEQ ID NO: 6 codifica fumarasa (fumB) de *Escherichia coli* (EC 4.2.1.2, número de acceso UniProt P14407). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. La secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE.UU.). El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *Kluyveromyces lactis*, es decir, el promotor de la gliceraldehído 3-fosfato deshidrogenasa controla la expresión del gen fumB. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador TDH1.

50 Las secuencias de cebador descritas en SEQ ID NO: 23 y SEQ ID NO: 24 se utilizaron para generar el fragmento 8 de PCR que consiste en el sitio de integración INT59 3', utilizando ADN genómico de la cepa CEN.PK 113-7D como molde.

Los fragmentos 1 a 8 de PCR se purificaron utilizando el kit de ADN Clean & Concentrator™-25 (Zymo Research, Irvine, CA, EE.UU.) de acuerdo con las instrucciones del fabricante.

Transformación en CEN.PK113-7D con el fin de construir la cepa SUC-1099

- La transformación en levadura se realizó por un método conocido por los expertos en la técnica. La cepa de *S. cerevisiae* CEN.PK113-7D se transformó con fragmentos 1 a 8 de PCR purificados. Los fragmentos 2 a 7 de PCR contenían solapamientos en sus extremos 5' y 3' y fragmentos de PCR 1 y 8 en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los ocho fragmentos de PCR. El extremo 5' del fragmento 1 de PCR y el extremo 3' del fragmento 8 de PCR eran homólogos al locus INT59 y permitieron la integración de los ocho fragmentos de PCR en el locus INT59 (véase la Figura 1). Esto dio como resultado un fragmento lineal que consistía en los fragmentos 1 a 8 de PCR integrados en el locus INT59. Este método de integración se describe en la solicitud de patente WO2013076280. El locus INT59 se encuentra en el cromosoma XI, 923 pb aguas abajo de YKR092C y 922 pb aguas arriba de YKR093W.
- Mezclas de transformación se colocaron en placas en agar YEPH (por litro: 10 gramos de extracto de levadura, 20 gramos de PhytonePeptone, 20 gramos de galactosa, 20 gramos de agar) que contenían 100 µg de G418 (Sigma Aldrich, Zwijndrecht, Países Bajos) por ml. Después de tres días de crecimiento a 30°C, los transformantes individuales se volvieron a sembrar en placas de agar YEPH que contenían 20 gramos de galactosa por litro y 100 µg de G418 por ml. La presencia de todos los genes introducidos se confirmó mediante el uso de PCR utilizando secuencias de cebadores que pueden reasociar las secuencias codificantes de los ORF codificados por SEQ ID NO: 2 a SEQ ID NO: 6. La cepa resultante se denominó SUC-1099.

Ejemplo 2: Transformación de fumarato reductasa en la cepa SUC-1099 y producción de ácido succínico en transformantes resultantes

Generación de fragmentos de PCR

- Se generaron fragmentos de PCR utilizando ADN polimerasa Phusion (New England Biolabs, EE.UU.) de acuerdo con las instrucciones del fabricante.

Las secuencias de cebador descritas en SEQ ID NO: 25 y SEQ ID NO: 26 se utilizaron para generar el fragmento 9 de PCR que consistía en el sitio de integración INT1 5', utilizando ADN genómico de la cepa CEN.PK 113-7D como molde.

- El fragmento 10 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 27 y SEQ ID NO: 28, utilizando SEQ ID NO: 7 como plantilla. La SEQ ID NO: 7 codifica un marcador de selección de nourseotricina funcional en *Saccharomyces cerevisiae* que se amplificó a partir de una versión modificada del plásmido pUG7-Nat. pUG7-Nat es una variante del plásmido pUG6 descrito por Geldener *et al.*, (Nucleic Acids Res. 1 de julio de 1996; 24(13):2519-24), en el que los sitios loxP presentes en pUG6 se cambiaron a sitios lox66 y lox71 (Lambert *et al.*, Appl. Environ. Microbiol. febrero de 2007; 73(4):1126-35. Epub 1 de diciembre de 2006) y en el que el marcador KanMX fue reemplazado por un marcador de nourseotricina (Goldstein y McCusker, Yeast. Octubre de 1999; 15(14):1541-53).

- El fragmento 11 de PCR se generó utilizando las secuencias de cebador descritas en SEQ ID NO: 31 y SEQ ID NO: 32, utilizando SEQ ID NO: 8 como molde. La SEQ ID NO: 8 codifica fumarato reductasa (FRDg) de *Trypanosoma brucei*, tal como se describe en la solicitud de patente WO2009/065778. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, incluidos los sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE.UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor TDH3 controla la expresión del gen FRDg. La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador TAL1. Las secuencias de cebador descritas en SEQ ID NO: 29 y SEQ ID NO: 30 se utilizaron para generar el fragmento 12 de PCR que consistía en el sitio de integración INT1 3', utilizando ADN genómico de la cepa CEN.PK 113-7D como molde.

- Los fragmentos 9 a 12 de PCR se purificaron utilizando el kit de ADN Clean & Concentrator™-25 (Zymo Research, Irvine, CA, EE.UU.) de acuerdo con las instrucciones del fabricante.

Transformación en SUC-1099

- La transformación de la levadura se realizó por un método conocido por personas expertas en la técnica. La cepa de *S. cerevisiae* CEN.PK113-7D se transformó con los fragmentos 9 a 12 de PCR purificados. Los fragmentos 10 y 11 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 9 y 12 de PCR contenían solapamientos en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR. El extremo 5' del fragmento 9 de PCR y el extremo 3' del fragmento 12 de PCR eran homólogos al locus INT1 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT1 (véase la Figura 2). Esto

5 dio como resultado un fragmento lineal que consistía en los fragmentos 9 a 12 de PCR integrados en el locus INT1. Este método de integración se describe en la solicitud de patente WO2013076280. El locus INT1 se encuentra en el cromosoma XV, 659 pb aguas abajo de YOR071c y 998 pb aguas arriba de YOR070c. Este enfoque dio como resultado la expresión de la proteína fumarato reductasa de 1139 aminoácidos tal como se indica en SEQ ID NO: 33, que carece del SKI de aminoácido C-terminal en comparación con la secuencia nativa de *T. brucei*.

10 Mezclas de transformación se colocaron en placas en agar YEPH (por litro: 10 gramos de extracto de levadura, 20 gramos de PhytonePeptone, 20 gramos de galactosa, 20 gramos de agar) que contenían 100 µg de nourseotricina (Jena Bioscience, Alemania) por ml. Después de tres días de crecimiento a 30°C, los transformantes individuales se volvieron a sembrar en placas de agar YEPH que contenían 20 gramos de galactosa por litro y 100 µg de nourseotricina por ml. La presencia de los genes introducidos se confirmó mediante el uso de PCR utilizando secuencias de cebador que pueden reasociarse con las secuencias codificantes de los ORF codificados por SEQ ID NO: 7 y SEQ ID NO: 8. Tres colonias individuales resultantes se denominaron PS107n⁰¹_n⁰², PS107n⁰¹_n⁰³ y PS107n⁰¹_n⁰⁴.

Producción de ácido succínico

15 Para determinar la producción de ácido succínico, la cepa SUC-1099 se cultivó por triplicado y los transformantes PS107n⁰¹_n⁰², PS107n⁰¹_n⁰³ y PS107n⁰¹_n⁰⁴ se cultivaron en placas de microtitulación utilizando galactosa como fuente de C. Después de una fase de crecimiento para generar biomasa, se inició un experimento de producción resuspendiendo las células en medio de cultivo. Se añadieron 30 g/L de ácido succínico al medio de cultivo al comienzo del experimento de producción. Se tomaron muestras de sobrenadante después de 96 horas de cultivo. Las muestras para la RMN de flujo se prepararon como sigue: de cada uno de los matraces con pocillos, se tomaron cultivos de 600 microlitros y se centrifugaron durante 1 minuto a 14.000 rpm. El sobrenadante de 50 microlitros se transfirió a una placa MTP de 96 pocillos. Se añadieron 450 microlitros de patrón interno (20 g/L de ácido maleico, 40 g/l de EDTA en D2O) y 500 microlitros de 80:20 H2O/D2O a cada una de las muestras. El ácido dicarboxílico, incluidas concentraciones de ácido succínico y otros compuestos tales como glucosa en el sobrenadante de fermentación se determinaron con un espectrómetro Bruker BEST avance II de 500 MHz. Los espectros de RMN se registraron con un programa de pulso de supresión de agua a 27 grados centígrados con un retraso de relajación de 30 s.

En el sobrenadante de SUC-1099, se midió un título medio de 30,9 g/L de ácido succínico. Cuando se introdujo FRDg y se sobre-expresó en SUC-1099, se midió un título medio de 53,5 g/L de ácido succínico.

Ejemplo 3: Transformación de variantes de fumarato reductasa en la cepa SUC-1099 y producción de ácido succínico en transformantes resultantes

Generación de fragmentos de PCR.

20 Los fragmentos 9, 10 y 12 de PCR se generaron tal como se describe en el Ejemplo 2. Secuencias de nucleótidos sintéticas que codifican diferentes variantes de proteínas de la secuencia de fumarato reductasa de referencia que se describe en SEQ ID NO: 33 se sintetizaron mediante DNA 2.0 (Menlo Park, California, EE.UU.). Las secuencias de nucleótidos sintéticos codifican un aminoácido mutado con relación a la secuencia de FRDg de referencia (SEQ ID NO: 33) en las posiciones indicadas en la Tabla 2. Además de codificar los aminoácidos mutados indicados en la Tabla 2, las variantes de la secuencia de nucleótidos sintéticos son idénticas a SEQ ID NO: 8. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor TDH3 controla la expresión del gen FRD mutado (FCC). La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador TAL1.

Las secuencias de genes sintéticas que contienen, entre otros, un promotor TDH3 - FRD mutado - terminador TAL1 se amplificaron por PCR utilizando las secuencias de cebador descritas en SEQ ID NO: 31 y SEQ ID NO: 32, para generar fragmentos 13 a 22 de PCR (véase la Tabla 2).

45 Los fragmentos 9, 10, 12 y 13 a 22 de PCR se purificaron utilizando el kit de ADN Clean & Concentrator™-25 (Zymo Research, Irvine, CA, EE.UU.) de acuerdo con las instrucciones del fabricante.

Transformación en SUC-1099

50 La cepa SUC-1099 se transformó con los fragmentos 9, 10 y 12 de PCR purificados en combinación con los fragmentos 13 a 22 de PCR individualmente. El fragmento 10 de PCR y los fragmentos 13 a 22 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 9 y 12 de PCR contenían solapamientos en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR. El

extremo 5' del fragmento 9 de PCR y el extremo 3' del fragmento 12 de PCR eran homólogos al locus INT1 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT1 (Figura 2). La transformación y selección de transformantes se describe en el Ejemplo 2.

Producción de ácido succínico

5 Para determinar la producción de ácido succínico, se cultivaron cuatro transformantes independientes de SUC-1099 que expresaban secuencias de fumarato reductasa mutadas en placas de microtitulación y se midieron los títulos de ácido succínico tal como se describe en el Ejemplo 2. Los títulos medios de ácido succínico se representan en la Tabla 2. La producción media de ácido succínico de varios transformantes SUC-1099 que expresan secuencias de fumarato reductasa mutadas excedió los 55 g/L de ácido succínico. Esto es significativamente más que el título medio de ácido succínico de SUC-1099 transformado con la secuencia de FRDg de referencia. Significativamente más significa que no se solapan los intervalos de confianza del 95% de los títulos succínicos para cepas de fumarato reductasa con secuencias de referencia y mutadas mejoradas.

15 Tabla 2: Títulos medios de ácido succínico medidos en el sobrenadante del medio de producción después de 4 días de cultivo de la cepa SUC-1099, que expresa fosfoenolpiruvato carboxiquinasa (PCKa), piruvato carboxilasa (PYC2), malato deshidrogenasa (MDH3), fumarasa (fumB), transportador ácido dicarboxílico (DCT 02) transformado con fumarato reductasa de referencia (SEQ ID NO: 33) o fumarato reductasa mutada, que contiene mutaciones en comparación con la secuencia de referencia en las posiciones de aminoácidos indicadas a continuación.

Fragmento de PCR	Clon	Mutación en posición					Título medio de ácido succínico (g/L)
		1042	1071	1072	1082	1083	
11	Referencia	E	N	R	G	F	53,5
13	FCC_040	R	T	R	K	Y	58,2
14	FCC_045	R	T	K	K	F	57,1
15	FCC_046	R	T	K	K	Y	58,2
16	FCC_048	R	T	K	R	Y	58,9
17	FCC_065	K	T	R	R	F	59,9
18	FCC_069	K	T	K	K	F	58,5
19	FCC_070	K	T	K	K	Y	59,7
20	FCC_075	Q	S	R	K	F	56,3
21	FCC_076	Q	S	R	K	Y	61,3
22	FCC_078	Q	S	R	R	Y	64,5

20 La introducción de un residuo cargado positivamente (lisina o arginina; K o R) en la posición de aminoácido 1082 es única, debido a que la secuencia de referencia de fumarato reductasa contiene una pequeña glicina sin carga (G) en esta posición. Además, ninguna variante natural previamente descrita o mutante de fumarato reductasa disponible públicamente contiene un residuo cargado en la posición número 1082. Como se representa en la Tabla 2, reemplazar la glicina en la secuencia de FRDg de referencia por un residuo cargado es ventajoso para lograr títulos de ácido succínico incrementados.

25 **Ejemplo 4: medición de la actividad específica de NADH y NADPH de las variantes de fumarato reductasa (FRD)**

30 Transformantes generados en el Ejemplo 3 se cultivaron tal como se describe en el Ejemplo 2. La biomasa se recolectó por centrifugación (4000 rpm, 10 min, 4°C) y se lavó dos veces con PBS (solución salina tamponada con fosfato, Sigma Aldrich), después de lo cual los sedimentos celulares se congelaron a -20°C. La disrupción celular se logró en placas de microtitulación de 96 pocillos profundos (MTP) con perlas de vidrio lavadas con ácido de 0,5 mm en combinación con el TissueLyser II de Qiagen (3000 rpm durante 2x 10 s). Se añadieron perlas de vidrio de un volumen de 600 µl al sedimento celular antes de la adición de 1 ml de medio de ensayo *in vivo* descrito en van Eunen et al. (*FEBS Journal* **277**: 749-760) que contenía DTT 0,5 mM (ditiotreitól, Sigma-Aldrich) y PMSF (fluoruro de fenilmetanosulfonilo, Amresco) 0,1 mM. Las perlas de vidrio se añadieron invirtiendo las MTP de pocillos profundos que contenían los sedimentos congelados sobre una MTP estándar, en que cada uno de los pocillos se llenó por completo con perlas de vidrio (= un volumen de 300 µl) y luego invirtiendo ambas placas, de modo que las perlas de vidrio cayeran sobre los sedimentos celulares. Este proceso se repitió para obtener 600 µl de perlas de vidrio en los sedimentos celulares. Después de la disrupción celular, los desechos celulares se sedimentaron por centrifugación (4000 rpm, 30 min, 4°C). El sobrenadante (extractos celulares solubles) se recogió y se almacenó en

hielo. La concentración de proteínas de los extractos se determinó mediante Bradford, utilizando albúmina de suero bovino (BSA) como patrón.

La actividad de fumarato reductasa (FRD) se ensayó por espectrofotometría siguiendo la disminución de la absorbancia a 340 nm provocada por la oxidación de NADH o NADPH a NAD⁺ o NADP⁺. Las mezclas de ensayo contenían NADH o NADPH 150 μ M, ácido fumárico 1 mM, 0,5 mg de proteína mL⁻¹ extractos celulares solubles en medio de ensayo *in vivo* en un volumen final de 200 μ l. Las reacciones se iniciaron mediante la adición de ácido fumárico, se siguieron durante 9 minutos a 30 grados centígrados y la pendiente se utilizó como una medida de la actividad de FRD dependiente de NADH o NADPH. La absorbancia se midió utilizando un lector de placas Tecan Infinite M1000. La actividad dependiente de NADH de cada una de las variantes se comparó con la actividad de NADPH. Se calculó la relación de actividad dependiente de NADPH:NADH para cada una de las variantes de modo que las variantes pudieran clasificarse. Las Figuras 3a y 3b muestran las actividades de FRD dependientes de NADPH y NADH para todas las variantes que contienen mutaciones beneficiosas para una producción de succínico incrementada descrita en el Ejemplo 3.

Ejemplo 5: Construcción de la cepa SUC-501

La cepa SUC-501 se construyó reemplazando los dos transportadores de ácido dicarboxílico SpMAE1 en la cepa SUC-401, descrita en el documento WO 2013/004670, con transportadores de ácido dicarboxílico DCT_02, similares a los descritos para SUC-489 en el documento WO 2013/004670. SUC-401 se transformó con el fragmento purificado de 7,7 kb del plásmido pSUC174 restringido con *Bsu36I* y *FseI*. El plásmido pSUC174 se describe en el documento WO 2013/004670. El fragmento de 7,7 kB contiene en el extremo 5' el gen sintético FUMR, el transportador DCT_02 y un marcador de selección KanMX flanqueado por sitios lox66/lox71 (Lambert JM, Bongers RS, Kleerebezem M., Appl Environ Microbiol. Febrero de 2007; 73(4):1126-35), y en el extremo 3' el gen sintético MDH3.

Transformantes correctos se seleccionaron inicialmente en cuanto a su resistencia contra G418, debido a la integración del marcador de resistencia KanMX. A continuación, se realizó una PCR de diagnóstico en la cepa intermedia SUC-461 para confirmar el reemplazo del gen sintético SpMAE1 por el gen sintético DCT_02. El marcador KanMX flanqueado por los sitios lox66 y lox71 se separó de la cepa SUC-461 por transformación de Cre-recombinasa (Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH., Nucleic Acids Res. 1 de julio de 1996; 24(13):2519-2524) utilizando el plásmido pSH65 que contenía un marcador de resistencia a la fleomicina. Posteriormente, el plásmido pSH65 se curó de las células mediante crecimiento en medio no selectivo (YEP 2% de galactosa), quedando un lox72 en el ADN genómico. La cepa resultante se designó SUC-464.

A continuación, se reemplazó una segunda copia del gen SpMAE1 presente en el ADN genómico de SUC-401. SUC-464 se transformó con el fragmento purificado de 7,7 kb del plásmido pSUC174 restringido con *Bsu36I* y *FseI*. Los transformantes correctos se seleccionaron inicialmente en cuanto a su resistencia contra G418, debido a la integración del marcador de resistencia KanMX. El fragmento de pSUC174 puede reemplazar el gen SpMAE1 restante o el gen DCT_02 introducido. Se realizó una PCR de diagnóstico en los transformantes para confirmar el reemplazo de ambos genes sintéticos SpMAE1 por el gen sintético DCT_02. Un transformante que contenía dos copias del gen DCT_02 se denominó SUC-467. El marcador KanMX flanqueado por los sitios lox66 y lox71 se separó de la cepa SUC-467 por transformación de Cre-recombinasa (Güldener U, Heck S, Fielder T, Beinhauer J, Hegemann JH., Nucleic Acids Res. 1 de julio de 1996; 24(13):2519-2524) utilizando el plásmido pSH65 que contenía un marcador de resistencia a la fleomicina. Posteriormente, el plásmido pSH65 se curó de las células mediante crecimiento en medio no selectivo (YEP 2% de galactosa), quedando un lox72 en el ADN genómico.

La cepa resultante se designó SUC-501 (*MATa ura3,52 HIS3 LEU2 TRP1 sit2::TPI1p-PCKa-PMA1t;TDH3p-FRDg-TDH3t sit4::TDH3p-MDH3-TDH3t;ENO1p-DCT_02-ENO1t;TPI1p-FUMR-PMA1t;lox72 adh1::PGK1p-PYC2-PGK1t;URA3p-URA3-URA3t MAL2-8 SUC2*).

Ejemplo 6: Selección de la cepa SUC-723 a partir de la evolución adaptativa de SUC-501

Se cultivó SUC-501 en un cultivo de auxóstato de 1,5 litros en un fermentador de 2 litros que contenía medio con una concentración de ácido succínico de 50 g/L a pH 3 a 30°C. La velocidad del agitador se ajustó a 150 rpm y el flujo de aire se ajustó a 2 NI/h. El tamaño del inóculo fue de 150 g de biomasa (cultivada en un matraz de agitación en medio Verduyn (Verduyn C, Postma E, Scheffers WA, Van Dijken JP. Yeast, julio de 1992; 8(7):501-517)) con galactosa como fuente de carbono). La tasa de alimentación del medio al cultivo de auxóstatos se incrementó gradualmente a lo largo de un período de 60 días. La tasa de alimentación inicial se ajustó a una tasa de dilución de 0,05 h⁻¹, la tasa de alimentación final se ajustó a una tasa de dilución de 0,16 h⁻¹. La tasa de alimentación se

maximizó para mantener el cultivo a la tasa de crecimiento máxima posible mientras se evita la formación de etanol y la separación por lavado del cultivo.

5 El medio de alimentación se basaba en Verduyn *et al.* (Verduyn C, Postma E, Scheffers WA, Van Dijken JP. Yeast, julio de 1992; 8(7):501-517), con modificaciones en las fuentes de carbono y nitrógeno, tal como se describe a continuación (véase la Tabla 3). El pH fue ajustado a 3 (utilizando KOH 1M/H₂SO₄ 1M).

10 La producción de ácido succínico se midió en varios momentos durante los 60 días de crecimiento del cultivo de auxóstatos, utilizando el método descrito en el Ejemplo 2 en el documento WO 2013/004670. Después de 30 días, los títulos de ácido succínico producidos por el cultivo de auxóstatos fueron similares a los de SUC-501 no evolucionada. Después de 60 días de evolución adaptativa, el cultivo de auxóstatos mostró una disminución en el título de ácido succínico de más del 50% en comparación con la SUC-501.

15 Un aislado de colonia único obtenido del cultivo de SUC-501 después de 60 días de crecimiento en el cultivo de auxóstatos se denominó SUC-723. El análisis por PCR, utilizando cebadores específicos para los marcos de lectura abiertos integrados utilizando un método conocido por una persona experta en la técnica, de los genes de ácido succínico que estaban presentes en la cepa SUC-501 (*FUMR*, *MDH3*, *PCKa*, *FRDg*, *PYC2*, *DCT02*) reveló que el gen *FRDg* no está presente en el SUC-723. La presencia de todos los otros genes de ácido succínico que están presentes en la SUC-501 podría demostrarse en la cepa SUC-723.

Tabla 3. Composición media del medio de alimentación utilizado en la evolución adaptativa de SUC-501 descrita en el Ejemplo 6. El medio de alimentación se basó en Verduyn *et al.* (Verduyn C, Postma E, Scheffers WA, Van Dijken JP. Yeast, julio de 1992; 8(7):501-517).

Materia prima		Concentración (g/kg)
Dextrosa	C ₆ H ₁₂ O ₆ . H ₂ O	27,5
Sulfato de amonio	(NH ₄) ₂ SO ₄	5,0
Dihidrógeno fosfato de potasio	KH ₂ PO ₄	3,0
Sulfato de magnesio	MgSO ₄ 7H ₂ O	0,5
Solución de elementos traza		1
Solución de vitaminas		1
Ácido succínico	C ₄ H ₆ O ₄	50

20 **Ejemplo 7: Transformación de fumarato reductasa en cepa SUC-723 y producción de ácido succínico en transformantes resultantes**

25 Para testar si la producción reducida de ácido succínico de SUC-723 se debe únicamente a la ausencia del gen *FRDg* (y que la SUC-723 puede utilizarse en consecuencia para evaluar la funcionalidad de variantes del gen *FRDg* en la producción de ácido succínico) una construcción de expresión de *FRDg* se integró en la cepa SUC-723.

Generación de fragmentos de PCR.

Los fragmentos de PCR se generaron utilizando ADN polimerasa Phusion (New England Biolabs, EE.UU.) de acuerdo con las instrucciones del fabricante.

30 Las secuencias de cebador descritas en SEQ ID NO: 34 y SEQ ID NO: 35 se utilizaron para generar el fragmento 23 de PCR que consistía en el sitio de integración INT12 5', utilizando ADN genómico de la cepa CEN.PK113-7D como molde.

Los fragmentos 10 y 11 de PCR se generaron tal como se describe en el Ejemplo 2.

35 Las secuencias de cebador descritas en SEQ ID NO: 36 y SEQ ID NO: 37 se utilizaron para generar el fragmento 24 de PCR que consistía en el sitio de integración INT12 3', utilizando ADN genómico de la cepa CEN.PK 113-7D como molde.

Los fragmentos 10, 11, 23 y 24 de PCR se purificaron utilizando el kit DNA Clean & Concentrator™-25 (Zymo Research, Irvine, CA, EE. UU.) de acuerdo con las instrucciones del fabricante.

Transformación en SUC-723

La transformación de levadura se realizó por un método conocido por los expertos en la técnica. La cepa SUC-723 de *S. cerevisiae* se transformó con fragmentos 10, 11, 23 y 24 de PCR purificados. Los fragmentos 10 y 11 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 23 y 24 de PCR contenían solapamientos en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR (véase la Figura 4). El extremo 5' del fragmento 23 de PCR y el extremo 3' del fragmento 24 de PCR eran homólogos al locus INT12 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT12 (véase la Figura 4). Esto dio como resultado un fragmento lineal que consistía en los fragmentos 10, 11, 23 y 24 de PCR integrados en el locus INT12. Este método de integración se describe en la solicitud de patente WO2013076280. El locus INT12 se encuentra en el cromosoma II, 743 pb aguas arriba del ATG del ORF YOR071c y 618 pb aguas abajo del ATG de YBL029C-A. Este enfoque resultó en la expresión de la proteína fumarato reductasa de 1139 aminoácidos tal como se indica en SEQ ID NO: 33, que carece del SKI de aminoácido C-terminal en comparación con la secuencia nativa de *T. brucei*, en la cepa SUC-723.

Mezclas de transformación se colocaron en placas en agar YEPH (por litro: 10 gramos de extracto de levadura, 20 gramos de PhytonePeptone, 20 gramos de galactosa, 20 gramos de agar) que contenían 100 µg de nourseotricina (Jena Bioscience, Alemania) por ml. Después de tres a cuatro días de crecimiento a 30°C, los transformantes individuales se volvieron a sembrar en placas de agar YEPH que contenían 20 gramos de galactosa por litro y 100 µg de nourseotricina por ml. La presencia de los genes introducidos se confirmó mediante el uso de PCR utilizando secuencias de cebadores que se pueden reasociar a las secuencias codificantes de los ORF codificados por SEQ ID NO: 7 (marcador de nourseotricina) y SEQ ID NO: 8 (gen fumarato reductasa). La cepa resultante se denominó SUC-723-FRDg.

Producción de ácido succínico de SUC-723 y SUC-723-FRDg

Para determinar la producción de ácido succínico, las cepas SUC-723 y SUC-723-FRDg se cultivaron por triplicado en placas de microtitulación utilizando galactosa como fuente de C. Después de una fase de crecimiento para generar biomasa, se inició un experimento de producción resuspendiendo las células en medio de cultivo. Se añadieron 30 g/L de ácido succínico al medio de cultivo al comienzo del experimento de producción. Se tomaron muestras de sobrenadante después de 96 horas de cultivo.

Las muestras para la RMN de flujo se prepararon como sigue: A partir de cada uno de los matraces con pocillos, se tomaron 600 microlitros de cultivo y se centrifugaron durante 1 minuto a 14.000 rpm. El sobrenadante de 50 microlitros se transfirió a una placa MTP de 96 pocillos profundos. Se añadieron 450 microlitros de patrón interno (20 g/L de ácido maleico, 40 g/l de EDTA en D2O) y 500 microlitros de 80:20 H2O/D2O a cada una de las muestras. El ácido dicarboxílico, incluidas las concentraciones de ácido succínico y otros compuestos tales como glucosa en el sobrenadante de fermentación se determinaron con un espectrómetro Bruker BEST avance II de 500 MHz. Los espectros de RMN se registraron con un programa de pulso de supresión de agua a 27 grados centígrados con un retraso de relajación de 30 s.

En el sobrenadante de SUC-723, se midió un título medio de 37,4 g/L de ácido succínico. Cuando se introdujo FRDg y se sobre-expresó en SUC-723, como en SUC-723-FRDg, se midió un título medio de 49,5 g/L de ácido succínico.

Ejemplo 8: Transformación de variantes de fumarato reductasa en la cepa SUC-723 y producción de ácido succínico en transformantes resultantes

Generación de fragmentos de PCR

El fragmento 10 de PCR se generó tal como se describe en el Ejemplo 2.
Los fragmentos 23 y 24 de PCR se generaron tal como se describe en el Ejemplo 7.

Las secuencias de nucleótidos sintéticas que codifican diferentes variantes de proteínas de la secuencia de fumarato reductasa de referencia que se describe en SEQ ID NO: 33 fueron sintetizadas por DNA 2.0 (Menlo Park, California, EE. UU.). Las secuencias de nucleótidos sintéticos codifican un aminoácido mutado en relación con la secuencia de FRDg de referencia (SEQ ID NO: 33) en las posiciones indicadas en la Tabla 4. Aparte de codificar los aminoácidos mutados indicados en la Tabla 4, las variantes de la secuencia de nucleótidos sintéticos son idénticas a SEQ ID NO: 8. El gen sintético está bajo control de (o está enlazado operativamente a) un promotor de *S. cerevisiae*, es decir, el promotor TDH3 controla la expresión del gen FRDg mutado (FCC). La terminación adecuada se controla mediante una secuencia de terminador de *S. cerevisiae*, es decir, el terminador TAL1.

Las secuencias de genes sintéticas que contienen entre otros un promotor TDH3 – FRD mutada - terminador TAL1 se amplificaron por PCR utilizando las secuencias de cebador descritas en SEQ ID NO: 31 y SEQ ID NO: 32, para generar fragmentos 13, 16, 17, 19, 21 y 22 de PCR (véase la Tabla 4).

Los fragmentos 10, 12, 13, 16, 17, 19, 21, 22, 23 y 24 de PCR se purificaron utilizando el kit DNA Clean & Concentrator™-25 (Zymo Research, Irvine, CA, EE. UU.) de acuerdo con las instrucciones del fabricante.

Transformación en SUC-723

5 La cepa SUC-723 se transformó con fragmentos 23, 10 y 24 de PCR purificados en combinación con los fragmentos 13, 16, 17, 19, 21 y 22 de PCR individualmente. Los fragmentos 10, 13, 16, 17, 19, 21 y 22 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 23 y 24 de PCR contenían solapamientos en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR. El extremo 5' del fragmento 23 de PCR y el extremo 3' del fragmento 24 de PCR eran homólogos al locus INT12 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT12 (Figura 4). La transformación y selección de transformantes se describe en el Ejemplo 2.

15 De cada una de las transformaciones se verificaron tres colonias por PCR utilizando cebadores que se hibridaban en el gen FRDg y cebadores que se hibridaban en el ADN genómico 5' de la región homóloga al fragmento 23 de PCR y 3' de la región homóloga al fragmento 24 de PCR. De todas las transformaciones se identificó al menos un transformante correcto. Una colonia de las cepas transformadas con fragmentos 13, 16, 17, 19, 21 y 22 de PCR se denominaron SUC-723-FCC_40, SUC-723-FCC_48, SUC-723-FCC_65, SUC-723-FCC_70, SUC-723-FCC_76 y SUC-723-FCC_78, respectivamente.

Producción de ácido succínico

20 Para determinar la producción de ácido succínico, SUC-723-FFC_40, SUC-723-FFC_48, SUC-723-FFC_65, SUC-723-FFC_70, SUC-723-FFC_76 y SUC-723-FFC_78 se cultivaron en triplicado en placas de microtitulación y los títulos de ácido succínico se midieron tal como se describe en el Ejemplo 2. Los títulos medios de ácido succínico se representan en la Tabla 4. La producción media de ácido succínico de varios transformantes SUC-723 que expresan secuencias de fumarato reductasa mutadas excedió de 54 g/L de ácido succínico. Esto es significativamente más que el título de ácido succínico medio de SUC-723 transformado con la secuencia de FRDg de referencia que produjo 49,5 g/L de ácido succínico. Significativamente más significa que no se solapan los intervalos de confianza del 95% de los títulos succínicos para cepas con secuencias de fumarato reductasa de referencia y mutadas mejoradas.

30 Tabla 4: Títulos medios de ácido succínico medidos en el sobrenadante del medio de producción después de 4 días de cultivo de la cepa SUC-723, que expresa fosfoenolpiruvato carboxiquinasa (PCKa), piruvato carboxilasa (PYC2), malato deshidrogenasa (MDH3), fumarasa (fumR), transportador de ácido dicarboxílico (DCT_02) transformado con fumarato reductasa de referencia (SEQ ID NO: 33) o fumarato reductasa mutada, que contiene mutaciones en comparación con la secuencia de referencia en las posiciones de aminoácidos indicadas a continuación.

Fragmento de PCR	Clon	Mutación en posición					Título medio de ácido succínico (g/L)
		1042	1071	1072	1082	1083	
-	SUC-723 Control						37,4
11	Referencia	E	N	R	G	F	49,5
13	FCC_040	R	T	R	K	Y	55,4
16	FCC_048	R	T	K	R	Y	54,9
17	FCC_065	K	T	R	R	F	53,7
19	FCC_070	K	T	K	K	Y	54,9
21	FCC_076	Q	S	R	K	Y	58,1
22	FCC_078	Q	S	R	R	Y	54,6

35 La introducción de un residuo cargado positivamente (lisina o arginina; K o R) en la posición de aminoácido 1082 es única, porque la secuencia de fumarato reductasa de referencia contiene una pequeña glicina sin carga (G) en esta posición. Además, ninguna variante natural previamente descrita o mutante de fumarato reductasa disponible públicamente contiene un residuo cargado en la posición número 1082. Tal como se representa en la Tabla 4, reemplazar la glicina en la posición 1082 en la secuencia de FRDg de referencia por un residuo cargado es

ventajoso para lograr títulos de ácido succínico incrementados. En el Ejemplo 3 esto se muestra en la cepa SUC-1099, en este ejemplo se muestra el mismo efecto en un fondo de cepa diferente, es decir, la cepa SUC-723.

Ejemplo 9: Medición de la actividad específica de NADH y NADPH de variantes de fumarato reductasa (FRD) expresadas en la cepa SUC-723

- 5 Transformantes generados en los Ejemplos 7 y 8 se cultivaron tal como se describe en el Ejemplo 2. Se recolectó biomasa, se rompieron las células y se determinó la concentración de proteína como en el Ejemplo 4. La actividad de fumarato reductasa (FRD) se ensayó por espectrofotometría, en triplicado, como en el Ejemplo 5. Las Figuras 5a y 5b muestran las actividades de FRD dependientes de NADPH y NADH para las variantes FCC_40, FCC_48, FCC_65, FCC_70, FCC_76, FCC_78 y FRD de referencia que se expresaron en la cepa SUC-723. Se tomó como control una cepa SUC-723 no transformada con FRD. Las seis variantes de FRD testadas fueron aquellas variantes que contenían mutaciones beneficiosas para una producción de succínico incrementada descrita en el Ejemplo 3 y el Ejemplo 8.

Ejemplo 10: Transformación de fumarato reductasa en cepa CEN.PK113-7D

Generación de fragmentos de PCR

- 15 El fragmento 25 de PCR se generó mediante amplificación por PCR de SEQ ID NO: 38 utilizando cebadores que amplifican la secuencia de nucleótidos completa de SEQ ID NO: 38. La SEQ ID NO: 38 describe el flanco de integración 5' INT09.01. La SEQ ID NO: 38 incluye una región de 50 pb de homología con el fragmento 26 de PCR ubicado en el extremo 3' de SEQ ID NO: 38. El sitio de integración INT09.01 se encuentra 359 pb aguas abajo del ORF YIL009W en el cromosoma IX.

- 20 El fragmento 26 de PCR se generó mediante amplificación por PCR de SEQ ID NO: 39 utilizando cebadores que amplificaban la secuencia de nucleótidos completa de SEQ ID NO: 39. La SEQ ID NO: 39 codifica un marcador de selección de nourseotricina funcional en *Saccharomyces cerevisiae* que se amplificó a partir de una versión modificada del plásmido pUG7-Nat. pUG7-Nat es una variante del plásmido pUG6 descrito por Gueldener *et al.* (Nucleic Acids Res. 1 de julio de 1996; 24(13):2519-24), en el que los sitios loxP presentes en pUG6 se cambiaron a sitios lox66 y lox71 (Lambert *et al.*, Appl. Environ. Microbiol. febrero de 2007; 73(4):1126-35. Epub 1 de diciembre de 2006) y en el que el marcador KanMX fue reemplazado por un marcador de nourseotricina (Goldstein y McCusker, Yeast. Octubre de 1999; 15(14):1541-53).

El fragmento 27 de PCR se generó mediante amplificación por PCR de SEQ ID NO: 40 utilizando cebadores que amplifican la secuencia de nucleótidos completa de SEQ ID NO: 40.

- 30 La SEQ ID NO: 40 codifica la fumarato reductasa (FRDg) de *Trypanosoma brucei*, tal como se describe en la solicitud de patente WO2009/065778 y se describe en SEQ ID NO: 33. Esta secuencia sintética, que incluye la secuencia del promotor-gen-terminador, que incluye sitios de restricción apropiados, fue sintetizada por DNA 2.0 (Menlo Park, California, EE. UU.). La secuencia del gen fue optimizada en un par de codones para la expresión en *S. cerevisiae* tal como se describe en la solicitud de patente WO2008/000632. El gen sintético está bajo el control de (o está enlazado operativamente a) un promotor de *Kluyveromyces lactis*, es decir, el promotor de la quinasa 3-fosfoglicerato PGK1 (número de acceso uniprot P14828) controla la expresión del gen FRDg. La terminación adecuada es controlada por una secuencia de terminador de *S. cerevisiae*, es decir, el terminador ADH1.

- 40 El fragmento 28 de PCR se generó mediante amplificación por PCR de SEQ ID NO: 41 utilizando cebadores que amplifican la secuencia de nucleótidos completa de SEQ ID NO: 41. La SEQ ID NO: 41 describe el flanco de integración 3' INT09.01. El extremo 5' del fragmento 28 de PCR consiste en una región de 50 pb homóloga al extremo 3' del fragmento 27 de PCR y los fragmentos 29 a 58.

- Los fragmentos de PCR 29 a 58 se generaron de manera similar a la generación del producto de PCR 27. En lugar de SEQ ID NO: 40 como diana para el fragmento 27 de la PCR, se utilizaron secuencias de nucleótidos sintéticas que codifican diferentes variantes de proteínas de la secuencia de fumarato reductasa de referencia que se describe en SEQ ID NO: 33. Las secuencias de nucleótidos sintéticas fueron sintetizadas por DNA 2.0 (Menlo Park, California, EE. UU.). Las secuencias de nucleótidos sintéticas codifican un aminoácido mutado con relación a la secuencia de FRDg de referencia (SEQ ID NO: 33) en las posiciones indicadas en la Tabla 5.

- 50 Aparte de codificar los aminoácidos mutados indicados en la Tabla 5, las variantes de la secuencia de nucleótidos sintéticos son idénticas a SEQ ID NO: 40. Las secuencias se amplificaron por PCR utilizando los cebadores utilizados para la amplificación del fragmento 27 de PCR para generar fragmentos 29 a 58 de PCR (véase la Tabla 5).

Tabla 5: Mutaciones presentes en las variantes de fumarato reductasa mutadas presentes en los fragmentos 29 a 58 de PCR en comparación con la fumarato reductasa de referencia en el fragmento 27 de PCR (SEQ ID NO: 33). En los fragmentos 29 a 60 de PCR están presentes mutaciones en las posiciones de aminoácidos 1042, 1071, 1072, 1082 y 1083 tal como se indica a continuación.

Fragmento de PCR	clon	1042	1071	1072	1082	1083
27	FRDg	E	N	R	G	F
29	FCC_097	R	G	R	R	Y
30	FCC_098	Q	G	R	R	Y
31	FCC_099	R	G	R	K	Y
32	FCC_100	Q	G	R	K	Y
33	FCC_101	R	S	R	R	I
34	FCC_102	R	G	R	R	I
35	FCC_103	Q	S	R	R	I
36	FCC_104	Q	G	R	R	I
37	FCC_105	R	S	R	K	I
38	FCC_106	R	G	R	K	I
39	FCC_107	Q	S	R	K	I
40	FCC_108	Q	G	R	K	I
41	FCC_109	R	S	R	R	A
42	FCC_110	R	G	R	R	A
43	FCC_111	Q	S	R	R	A
44	FCC_112	Q	G	R	R	A
45	FCC_113	R	S	R	K	A
46	FCC_114	R	G	R	K	A
47	FCC_115	Q	S	R	K	A
48	FCC_116	Q	G	R	K	A
49	FCC_117	R	S	R	R	S
50	FCC_118	R	G	R	R	S
51	FCC_119	Q	S	R	R	S
52	FCC_120	Q	G	R	R	S
53	FCC_121	R	S	R	K	S
54	FCC_122	R	G	R	K	S
55	FCC_123	Q	S	R	K	S
56	FCC_124	Q	G	R	K	S
57	FCC_030	R	S	R	R	Y
58	FCC_028	R	S	R	K	Y

5

Transformación en CEN.PK113-7D para construir la cepa CPK-FRDg

La transformación de levadura se realizó por un método conocido por los expertos en la técnica. La cepa de *S. cerevisiae* CENPK.113-7D (*MATa HIS3 LEU2 TRP1 MAL2-8 SUC2*) se transformó con fragmentos 25 a 28 de PCR purificados. Los fragmentos 26 y 27 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 25 y 26 de PCR en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR. El extremo 5' del fragmento 25 de PCR y el extremo 3' del fragmento 28 de PCR eran homólogos al locus INT09.01 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT09.01 (véase la Figura 6).

10

Las mezclas de transformación se sembraron en placas en agar YEPH (por litro: 10 gramos de extracto de levadura, 20 gramos de PhytonePeptone, 20 gramos de galactosa, 20 gramos de agar) que contenían 100 µg de nourseotricina (Jena Bioscience, Alemania) por ml. Después de tres días de crecimiento a 30°C, los transformantes individuales se volvieron a sembrar en placas de agar YEPH que contenían 20 gramos de galactosa por litro y 100 µg de nourseotricina por ml. La presencia de los genes introducidos se confirmó utilizando PCR utilizando secuencias de cebadores que se pueden reasociar con las secuencias codificantes de los ORF codificados por SEQ ID NO: 39 y SEQ ID NO: 40. Se seleccionaron colonias individuales y se denominaron CPK-FRDg.

15

Transformación en CEN.PK113-7D con el fin de construir la cepa CPK-FCC 097 a CPK-FCC 124 y CPK-FCC 030 y CPK-FCC 028.

20

La transformación de la levadura se realizó por un método conocido por personas expertas en la técnica. La cepa de *S. cerevisiae* CENPK.113-7D (*MATa HIS3 LEU2 TRP1 MAL2-8 SUC2*) se transformó con los fragmentos 25, 26, 28 de PCR purificados y los fragmentos 29 a 58 individualmente. El fragmento 26 y los fragmentos 29 a 58 de PCR contenían solapamientos en sus extremos 5' y 3' y los fragmentos 25 y 28 de PCR en sus extremos 3' y 5', respectivamente, de modo que esto permitió la recombinación homóloga de los cuatro fragmentos de PCR tal como se indica en la Figura 6 El extremo 5' del fragmento 25 de PCR y el extremo 3' del fragmento 28 de PCR eran homólogos al locus INT09.01 y permitieron la integración de los cuatro fragmentos de PCR en el locus INT09.01 (véase la Figura 6).

Mezclas de transformación se sembraron en placas en agar YEPH (por litro: 10 gramos de extracto de levadura, 20 gramos de PhytonePeptone, 20 gramos de galactosa, 20 gramos de agar) que contenían 100 µg de nourseotricina (Jena Bioscience, Alemania) por ml. Después de tres días de crecimiento a 30°C, los transformantes individuales se volvieron a sembrar en placas de agar YEPH que contenían 20 gramos de galactosa por litro y 100 µg de nourseotricina por ml. La presencia de los genes introducidos se infirió de la medición de la actividad de la enzima fumarato reductasa.

Ejemplo 11: Medición de la actividad específica de NADH y NADPH de las variantes de fumarato reductasa (FRD) expresadas en la cepa CEN.PK113-7D

Transformantes generados en el Ejemplo 10 se cultivaron en placas de 24 pocillos profundos (Axygen) en medio Verduyn (Verduyn C, Postma E, Scheffers WA, Van Dijken JP. Yeast, julio de 1992; 8(7):501-517) con galactosa como fuente de carbono y que contenían 150 µg de nourseotricina (Jena Bioscience, Alemania) por ml. Se recolectó biomasa, se rompieron las células y se determinó la concentración de proteína como en el Ejemplo 4. La actividad de fumarato reductasa (FRD) se ensayó por espectrofotometría como en el Ejemplo 4, con las siguientes modificaciones menores: se utilizaron NADH o NADPH 400 µM en lugar de 150 µM. Las actividades medidas se normalizaron para la proteína total presente, dividiendo por la concentración de proteína total en el ensayo, según lo determinado por Bradford utilizando albúmina de suero bovino como patrón. La Tabla 6 enumera la relación de especificidad para cada una de las variantes; la relación de actividad dependiente de NADPH:NADH = (pendiente a 340 nm con NADPH) / (pendiente a 340 nm con NADH). Esta relación se calculó de modo que las variantes pudieran clasificarse según la especificidad de cofactor alterada (véase: Tabla 6). Un valor aumentado para la relación en comparación con el clon de referencia indicó que la especificidad del cofactor había sido alterada. Esto puede indicar una menor especificidad de NADH, una actividad incrementada de NADPH o una combinación de ambos. La Figura 7 muestra las actividades para variantes de FRD dependientes de NADPH y NADH para las variantes de FRD expresadas en CEN.PK113-7D. Como control negativo se tomó una cepa CEN.PK113-7D no transformada con FRD.

Tabla 6. Relación de especificidad NADPH:NADH para variantes de FRD expresadas en la cepa CEN.PK113-7D. Al calcular la relación entre la actividad de FRD específica para NADPH y NADH, las variantes podrían clasificarse y compararse fácilmente con la secuencia de referencia. Sorprendentemente, todas las variantes mostraron una especificidad de cofactor alterada, muchas muy drásticamente, en comparación con la FRD de referencia, como lo demuestra una relación de especificidad diferente.

Fragmento de PCR	Clon	Mutación en posición					Actividad dependiente de la relación NADH:NADPH
		1042	1071	1072	1082	1083	
-	CEN.PK113-7D control						n.d.
27	Referencia	E	N	R	G	F	0,01
59	FCC_028	R	S	R	K	Y	0,6
57	FCC_030	R	S	R	R	Y	0,8
29	FCC_097	R	G	R	R	Y	3,9
30	FCC_098	Q	G	R	R	Y	5,5
31	FCC_099	R	G	R	K	Y	0,04
32	FCC_100	Q	G	R	K	Y	0,5
33	FCC_101	R	S	R	R	I	0,2
34	FCC_102	R	G	R	R	I	0,1
35	FCC_103	Q	S	R	R	I	2,3
36	FCC_104	Q	G	R	R	I	0,4
37	FCC_105	R	S	R	K	I	4,0
38	FCC_106	R	G	R	K	I	12,7

ES 2 790 573 T3

39	FCC_107	Q	S	R	K	I	2,2
40	FCC_108	Q	G	R	K	I	1,5
41	FCC_109	R	S	R	R	A	1,8
42	FCC_110	R	G	R	R	A	0,7
43	FCC_111	Q	S	R	R	A	0,2
44	FCC_112	Q	G	R	R	A	0,4
45	FCC_113	R	S	R	K	A	0,9
46	FCC_114	R	G	R	K	A	0,2
47	FCC_115	Q	S	R	K	A	0,2
48	FCC_116	Q	G	R	K	A	0,2
49	FCC_117	R	S	R	R	S	0,8
50	FCC_118	R	G	R	R	S	0,2
51	FCC_119	Q	S	R	R	S	0,5
52	FCC_120	Q	G	R	R	S	0,2
53	FCC_121	R	S	R	K	S	0,6
54	FCC_122	R	G	R	K	S	0,5
55	FCC_123	Q	S	R	K	S	0,2
56	FCC_124	Q	G	R	K	S	0,1

LISTADO DE SECUENCIAS

<110> DSM IP Assets B.V.

<120> FUMARATO REDUCTASAS

<130> 29929-WO-PCT

5 <150> EP13196950.3

< 151> 02-12-2013

<160> 41

<170> PatentIn version 3.5

<210> 1

10 < 211> 2626

< 212> ADN

< 213> Actinobacillus succinogenes

<400> 1

ES 2 790 573 T3

ttgccatcg aacgtacaag tactcctctg ttctctcctt cctttgcttt gtgcgacacc	60
taactacata gtgtttaaag attacggata ttttaacttac ttagaataat gccatTTTTT	120
tgagttataa taatcctacg ttagtgtgag cgggatttaa actgtgagga ccttaataca	180
ttcagacact tctgcggtat caccctactt attcccttcg agattatata taggaacca	240
tcaggttggg ggaagattac cggttctaag acttttcagc ttctcttatt gatgttacac	300
ctggacaccc cttttctggc atccagtttt taatcttcag tggcatgtga gattctccga	360
aattaattaa agcaatcaca caattctctc ggataccacc tcggttgaaa ctgacaggtg	420
gtttgttacg catgctaagc caaaggagcc tataatcctt tggctcggct gctgtaacag	480
ggaatataaa gggcagcata attaggagt ttagtgaact tgcaacattt actatTTTcc	540
cttcttacgt aaatTTTTT ctttttaatt ctaaataaat ctttttcaat tttttgtttg	600
tattcttttc ttgcttaaata ctataactac aaaaaacaca tacataaact aaaaatgact	660
gatttgaaca aattgggtcaa ggaattgaat gatttggggt tgactgacgt caaggaaatt	720
gtctacaacc catcttacga acaattattc gaagaagaaa ccaagccagg tttggaagg	780
ttcgacaagg gtactttgac cactttaggt gctgttgctg ttgacaccgg tattttcacc	840
ggtcgttctc caaaggacaa atacattggt tgtgatgaaa ccaccaagga caccgtctgg	900
tggaactctg aagctgccaa gaacgataac aagccaatga ctcaagaaac ctggaatct	960
ttgagagaat tggttgccaa gcaattgtct ggtaagagat tattcgttgt tgacgctttc	1020
tgtggtgctt ctgaaaagca cagaattggg gtcagaatgg tcaactgaagt tgcttgccaa	1080
gctcatttcg tcaagaacat gttcatcaga ccaactgacg aagaattgaa gaacttcaag	1140
gctgacttca ccgTTTTgaa tggtgccaag tgtaccaacc caaactggaa ggaacaagg	1200
ttgaactctg aaaactttgt tgctttcaac atcaactgaag gtatccaatt gattggtgg	1260

ES 2 790 573 T3

acttggtacg gtggtgaaat gaagaagggt atgttctcca tgatgaacta tttcttgcca 1320
 ttgaaagggt ttgcttccat gcactgttct gccaatgtcg gtaaggatgg tgacgttgcc 1380
 atcttcttcg gtctatccgg tactggtgaag accactctat ccaactgaccc aaagagacaa 1440
 ttgattgggt atgacgaaca cggttgggac gaatctgggt tctttaactt tgaaggtggt 1500
 tgttacgcca agaccatcaa cttatctcaa gaaaacgaac cagatatcta cggtgccatc 1560
 cgtcgtgatg ctttgttgga aaacgttggt gtcagagctg acggttctgt tgacttcgac 1620
 gacggttcca agactgaaaa caccagagtt tcttaccocaa tctaccacat tgacaacatt 1680
 gtcagacctg tttccaaggc tggtcacgct accaaggtta tcttcttgac tgctgatgct 1740
 ttcggtgtct tgccacctgt ttccaaattg actccagaac aaaccgaata ctacttcttg 1800
 tccggtttca ctgccaaatt ggctggtact gaaagagggt tcaactgaacc aactccaact 1860
 ttctctgctt gtttcggtgc tgctttctta tctttgcacc caatccaata cgctgatgctc 1920
 ttggttgaaa gaatgaaggc ttctggtgct gaagcttact tggccaacac cggttggaac 1980
 ggtactggta agagaatctc catcaaggat accagaggta tcattgatgc tatcttggac 2040
 ggttccattg aaaaggctga aatgggtgaa ttgccaatct tcaacttggc cattccaaag 2100
 gctttgccag gtgttgacc agccatctta gatccaagag acacctacgc tgacaaggct 2160
 caatggcaag tcaaggctga agatttggct aacagattcg tcaagaactt tgtcaaatac 2220
 actgctaacc cagaagctgc caaattggtt ggtgctggtc caaaggctta aaggagttaa 2280
 aggcaaagtt ttcttttcta gagccgttcc cacaaataat tatacgtata tgcttctttt 2340
 cgtttactat atatctatat ttacaagcct ttattcactg atgcaatttg tttccaaata 2400
 cttttttgga gatctcataa ctagatatca tgatggcgca acttggcgct atcttaatta 2460
 ctctggctgc caggcccgtg tagagggccg caagacctc tgtacgcat atagtctcta 2520
 agaacttgaa caagtttcta gacctattgc cgccttctcg atcgctattg ttcctccgga 2580
 tcgatgtaca caaccgactg cacccaaacg aacacaaatc ttagca 2626

<210> 2

< 211> 4552

5 < 212> ADN

< 213> *Saccharomyces cerevisiae*

<400> 2

ES 2 790 573 T3

```
cggatcgatg tacacaaccg actgcaccca aacgaacaca aatcttagca gtgcgggcca      60
gaaaaaggaa gtgtttccct ccttcttgaa ttgatgttac ctcataaag cacgtggcct      120
cttatcgaga aagaaattac cgtcgctcgt gattgtttg caaaaagaac aaaactgaaa      180
aaaccagac acgctcgact tcctgtcttc ctattgattg cagcttcaa ttctgcaca      240
caacaaggtc ctagcgacgg ctcacaggtt ttgtaacaag caatcgaagg ttctggaatg      300
```

ES 2 790 573 T3

gcgggaaag gtttagtacc acatgctatg atgcccactg tgatctccag agcaaagtcc 360
 gttcgatcgt actgttactc tctctctttc aaacagaatt gtccgaatcg tgtgacaaca 420
 acagcctggt ctcacacact cttttcttct aaccaagggg gtggtttagt ttagtagaac 480
 ctctgtaaac ttacatttac atatatataa acttgcataa attggtcaat gcaagaaata 540
 catatttggt cttttctaata tctgtagttt tcaagtctt agatgctttc tttttctctt 600
 ttttacagat catcaaggaa gtaattatct actttttaca acaaatataa aacaatgtcc 660
 tcttccaaga tcttggtggtg tttgagagac aacttttctt tgttggtgga aaagaacaag 720
 attttggtcg ccaacagagg tgaaatccca atcagaattt tcagatctgc tcacgaattg 780
 tctatgagaa ctatcgccat ctactctcac gaagatagat tatccatgca cagattgaag 840
 gctgatgaag cctacgttat cgggtgaagaa ggtcaatata ccccagtcgg tgcttacttg 900
 gccatggacg aatcatcga aattgccaag aagcacaagg tcgatttcat ccaccaggt 960
 tacggtttct tgtctgaaaa ctctgaattt gctgacaagg ttgttaaggc tggattacc 1020
 tggattggtc caccagctga agtcattgaa tctgttggtg acaaggtttc tgccagacat 1080
 ttggctgctc gtgccaacgt tccaactgtc ccaggctactc caggctctat cgaaaccgtt 1140
 caagaagctc tagatttctg caatgaatac ggttaccag ttatcatcaa ggctgcttcc 1200
 ggtggtggtg gtcgtggtat gagagtgtgc agagaagggtg acgatgtcgc tgatgcttcc 1260
 caaagagcca cttctgaagc tagaactgct ttcggtaacg gtacttgttt cgtcgaaaga 1320
 ttcttggaag agccaaagca cattgaagtt caattattag ctgacaacca cggtaacgtt 1380
 gtccacttgt tcgaaagaga ctgttccgctc caaagacgctc accaaaagggt tgtogaagtt 1440
 gctccagcta agactttacc aagagaagtt agagatgcta tcttgaccga tgccgttaag 1500
 ttggctaagg tttgtggtta cagaaacgct ggtactgctg aattcttgggt tgacaaccaa 1560
 aacagacatt acttcattga aatcaaccca agaattcaag tcgaacacac catcactgaa 1620
 gaaatcactg gtattgacat tgtctccgct caaatccaaa tcgccgctgg tgctactttg 1680
 actcaattag gtctattaca agacaaaatc accaccagag gtttctctat ccaatgtcgt 1740
 atcaccactg aagatccatc caagaacttc caaccagaca ctggctgctt ggaagtctac 1800
 agatccgctg gtggtaacgg tgtcagattg gacggtggtg acgcctacgc tggtgctacc 1860
 atctctccac actacgactc catgttggtt aagtgttctt gttctggttc tacctacgaa 1920
 attgtcagaa gaaagatgat cagagctttg attgaattca gaatcagagg tgtcaagacc 1980
 aacatccat tcttggtgac tttgttgacc aaccagttt tcattgaagg tacctactgg 2040
 accactttca tcgatgacac tccacaattg ttccaaatgg tttcctctca aaacagagct 2100
 caaaaattgt tgcactactt ggctgacttg gccgtcaacg gttcctctat caagggtcaa 2160
 atcggtttac caaagttgaa gtccaaccct tccgttccac atttgcaaga tgetcaaggt 2220

ES 2 790 573 T3

aatgtcatca acgttaccaa atctgcccc	ccatccggtt ggagacaagt cttggtgga	2280
aagggtccat ccgaatttgc caagcaagtc	agacaattca acggtacttt gttgatggac	2340
accacctgga gagatgctca ccaatctttg	ctagctacca gagtcagaac tcacgatttg	2400
gccaccattg ctccaaccac tgctcacgct	ttggctggtg cctttgcttt ggaatggttg	2460
ggtggtgcta ctttcgatgt cgccatgaga	ttcttgcatg aggacccatg ggaaagattg	2520
agaaaattga gatctttggt cccaaacatt	ccattccaaa tgttggtgag aggtgctaac	2580
ggtggtgctt actcctcttt gccagacaac	gccattgacc atttcgtaa gcaagccaag	2640
gacaatggtg ttgacatttt cagagtcttt	gacgctttga acgacttgga acaattgaag	2700
gttggtgtta atgctgtcaa gaaggctgg	ggtggtgtcg aagctaccgt ttgttactct	2760
ggtgacatgt tgcaaccagg taagaaatac	aacttgact actacttaga agttgtcgaa	2820
aagatcgttc aaatgggtac tcacatcttg	ggtatcaagg acatggctgg taccatgaag	2880
ccagctgctg ccaaattggt gattggttct	ttacgtacca gataccaga cttgccaatc	2940
cacgttcact ctcatgactc cgctggact	gctggtgctt ccatgactgc ttgtgctttg	3000
gccggtgctg atgttgttga cgttgccatt	aactccatgt ccggtttgac ctctcaacca	3060
tctattaacg ctttgttggc ctcttgtaa	ggtaacattg aactggtat caactcgaa	3120
cacgttagag aattggacgc ttactgggct	gaaatgagat tattatactc ttgtttcgaa	3180
gctgacttga agggctcaga ccctgaagtt	taccaacacg aaattccagg tggcaattg	3240
accaacttgt tgttccaagc tcaacaatta	ggtctaggtg aacaatgggc tgaaaccaag	3300
agagcttaca gagaagctaa ctactgtttg	ggtgacattg ttaaggtcac cccaacttct	3360
aaggtcgttg gtgatttggc tcaattcatg	gtttctaaca aattgacttc tgatgacatc	3420
agaagattag ctaactcttt ggacttccca	gactccgtaa tggacttctt cgaaggtttg	3480
atcggtaaac catacgggtg tttcccagaa	ccattgagat ccgatgtttt gagaaacaag	3540
cgtcgtaaat tgactttag accaggttta	gaattggaac cattcgattt ggaaaagatc	3600
agagaagatt tgcaaaacag attcggtgat	atcgatgaat gtgatggtgc ctcctacaac	3660
atgtatcctc gtgtctacga agatttccaa	aagattagag aaacttacgg tgacttgtct	3720
gtcttaccia ccaagaactt cttggctcca	gctgaaccag acgaagaaat cgaagtcacc	3780
attgaacaag gtaagacttt gattatcaaa	ttacaagctg ttggtgattt gaacaagaaa	3840
accggtcaaa gagaagtcta ctctgaattg	aacggatgaat tgagaaagat cagagttgct	3900
gacaaatctc aaaacattca atctgttgcc	aagccaaagg ctgatgtcca cgacaccac	3960
caaatcgggtg ctccaatggc tgggtgcatc	attgaagtca aggttcacaa gggttctttg	4020
gtcaagaagg gtgaatctat cgccgttttg	tctgctatga agatggaaat ggttgtttcc	4080

ES 2 790 573 T3

```
tctccagctg atggtcaagt caaagatgtc tttatccgtg acggtgaatc cgtc gatgct 4140
tctgacttgt tggttgtttt ggaagaagaa actctaccac cttctcaaaa gaaataaagc 4200
gaatttctta tgatttatga tttttattat taaataagtt ataaaaaaaa taagtgtata 4260
caaattttaa agtgactcct aggtttttaa acgaaaattc ttattcttga gtaactcttt 4320
cctgtaggtc aggttgcttt ctcaggata gcatgaggtc gctcttattg accacacctc 4380
taccgcatg ccgagcaaat gcctgcaaat cgctcccat ttcaccaat tntagatag 4440
ctaactccag caatgagttg atgaatctog gtgtgtattt tatgtcctca gaggacaacc 4500
tcacgctttc cggcatcttc cagaccacag tatatccatc cgctcctgt tg 4552
```

<210> 3

< 211> 1572

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Marcador de selección KanMX

<400> 3

ES 2 790 573 T3

tcgtacgctg caggtcgacg aattctaccg ttcgtataat gtatgctata cgaagttata 60
gatctgttta gcttgcctcg tccccgccgg gtcaccggc cagcgacatg gagggccaga 120
ataccctoct tgacagtctt gacgtgcgca gctcaggggc atgatgtgac tgtcgcccggt 180
acatttagcc catacatccc catgtataat catttgcac catacatttt gatggccgca 240
cggcgcgaag caaaaattac ggctcctcgc tgcagacctg cgagcagggga aacgctcccc 300
tcacagacgc gttgaattgt cccacgcgog cggccctgta gagaaatata aaaggttagg 360
atltgccaact gaggttcttc tttcatatac ttccttttaa aatcttgcta ggatacagtt 420
ctcacatcac atccgaacat aaacaacat gggttaaggaa aagactcacg tttcgaggcc 480
gcgattaaat tccaacatgg atgctgattt atatgggtat aaatgggctc gcgataatgt 540
cgggcaatca ggtgcgacaa tctatcgatt gtatgggaag cccgatgcgc cagagttggt 600
tctgaaacat ggcaaaggtg gcggtgcaa tgatgttaca gatgagatgg tcagactaaa 660
ctggctgacg gaatttatgc ctctccgac catcaagcat tttatccgta ctctgatga 720
tgcattggtta ctaccactg cgatccccgg caaacagca ttccaggtat tagaagaata 780
tcctgattca ggtgaaaata ttgttgatgc gctggcagtg ttctgccc ggttgcattc 840
gattcctggt tgtaattgtc cttttaacag cgatcgcgta tttcgtctcg ctccggcgca 900
atcacgaatg aataacggtt tgggtgatgc gaggatgtt gatgacgagc gtaatggctg 960
gcctgttgaa caagtctgga aagaaatgca taagcttttg ccattctcac cggattcagt 1020
cgtcactcat ggtgatttct cacttgataa ccttattttt gacgagggga aattaatagg 1080
ttgtattgat gttggacgag tcggaatcgc agaccgatac caggatcttg ccacctatg 1140

gaactgcctc ggtgagtttt ctcttcatt acagaaacgg ctttttcaa aatatggtat 1200
tgataatcct gatatgaata aattgcagtt tcatltgatg ctogatgagt ttttctaac 1260
agtactgaca ataaaaagat tcttgttttc aagaacttgt catttgtata gttttttat 1320
attgtagttg ttctatttta atcaaatggt agcgtgattt atattttttt tcgcctcgac 1380
atcatctgcc cagatgcgaa gttaaagtgc cagaaagtaa tatcatgcgt caatcgtatg 1440
tgaatgctgg tcgctatact gctgtcgatt cgatactaac gccgccatcc agtgtcgaaa 1500
acgagctcat aacttcgtat aatgtatgct atacgaacgg tagaattcga tatcagatcc 1560
actagtggcc ta 1572

- 5 <210> 4
- < 211> 2260
- < 212> ADN
- < 213> Aspergillus niger

<400> 4

ES 2 790 573 T3

aacgttgacc aggtttgtat ccacgtgtgt ccgttccgcc aatattccgc gtgcccggg	60
aaccgccaga tattcattac ttgacgcaaa agcgtttgaa ataatgacga aaaagaagga	120
agaaaaaaaa agaaaaatac cgcttctagg cgggttatct actgatccga gcttccacta	180
ggatagcacc caaacacctg catatttggga cgacctttac ttacaccacc aaaaaccact	240
ttcgctctc ccgccctga taacgtccac taattgagcg attacctgag cggctctctt	300
ttgtttgcag catgagactt gcatactgca aatcgtaagt agcaacgtgt caaggtcaaa	360
actgtatgga aaccttgtca cctcacttaa ttctagctag cctaccctgc aagtcaagag	420
gtgtccgga ttcttagcca cctcaaggta tgcctctccc cggaaactgt ggccttttct	480
ggcacacatg atctccacga tttcaacata taaatagctt ttgataatgg caatattaat	540
caaatttatt ttacttcttt cttgtaacat ctctcttgta atcccttatt ccttctagct	600
atthttcata aaaaaccaag caactgctta tcaacacaca aacactaaat caaaatgaac	660
gttgaaactt ctttgccagg ttcttctggt tctgacttgg aaactttcca ccacgaacc	720
aagaagcatg ccaaccacga ctctggtatt tccgtcaacc atgaagctga aattggtggt	780
aaccacactt tcgaaaagcc aggtccagtt ggtatcagag aaagattacg tcacttcacc	840
tgggcttggg acactttgac catgtcctgt ggtggtttgg ctttgttgat tgtcaaccaa	900
ccacacgact tcaagggttt gaaagatatt gccagagttg tctactgttt gaacttggct	960
ttctttgtta tcgttacctc tttgatggcc atcagattca tcttgcaaaa gaacatgtgg	1020
gaatccttgg gtcacgacag agaaggtttg tttttccaa ctttctggtt atccattgct	1080
accatgatca ctggtttgta caagtgttc ggtgatgatg ctaacgaaaa gttcaccaag	1140
tgtttgcaag ttttgttctg gatctactgt ggttgtagca tgatcactgc tgtcgggtcaa	1200

ES 2 790 573 T3

tactctttcg tctttgctac ccacaaatac gaattgcaca ccatgatgcc atcctggatc 1260
 ttgccagctt tcccagttat gttgtctggt actatcgcct ccgtcatcgg ttctggtaa 1320
 ccagcttccg atggtattcc aattattatt gctggtatca ctttccaagg tttaggtttc 1380
 tccatctcct tcatgatgta cgctcactac attggtagat tgatggaagt tggtttacca 1440
 tctccagaac acagaccagg tatgttcatc tgtgttggtc ctocagcttt caccgctttg 1500
 gctttggtcg gtatggccaa ggctttacca gacgacttcc aaattgtcgg tgaccctcac 1560
 gctgtcattg acggtcgtgt tatgttggtc ttggctgtct ctgctgccat cttcttatgg 1620
 gctttgtcct tctggttctt ctgtatcgtt gttgttgctg ttgtcagatc tccaccaaa 1680
 ggtttccatt tgaactggtt tgccatggtt ttcccaaaca ctggtttcac cttggtacc 1740
 atcactttgg ctaacatggt cgaatctcca ggtgtcaagg gtgttgccac tgctatgtcc 1800
 ctatgtgtca tcatcatggt tattttcgtc ttggtttctg ccatcagagc tgatcatcaga 1860
 aaggacatca tgtggccagg tcaagatgaa gatgtttctg aataaagagt aataattatt 1920
 gcttccatat aatattttta tatacctctt atttttatgt attagttaat taagtatttt 1980
 tatctatctg cttatcattt tcttttcata tagggggggg tgggtgtttc ttgcccatca 2040
 gattgatgtc ctccaactcg gcactatttt acaaaggggt tttttgtaag agaaggagaa 2100
 gacagatact aaaccatcag ttactcgaaa caaaaaaaaa aaaaatggaa aaagctgcta 2160
 tcaacaaaag acggcctcat caaacctaaa gaaaccatgt cagcgtcctc aaataaccac 2220
 aaacatcctt cccatagctt cggtcgtgct tgtgttacct 2260

<210> 5
 <211> 2032
 <212> ADN
 <213> *Saccharomyces cerevisiae*

5

<400> 5
 aaataaccac aaacatcctt cccatagctt cggtcgtgct tgtgttacct gtgcctactt 60
 ggcttcacat acgttgcata cgtcgatata gataataatg ataatgacag caggattatc 120
 gtaatacgtg atagttgaaa atctcaaaaa tgtgtgggtc attacgtaaa taatgatagg 180
 aatgggattc ttctattttt cttttttcca ttctagcagc cgtcgggaaa acgtggcatc 240
 ctctctttcg ggctcaattg gagtcacgct gccgtgagca tcctctcttt ccatatctaa 300
 caactgagca cgtaaccaat ggaaaagcat gagcttagcg ttgctccaaa aaagtattgg 360
 atggtaata ccatttgtct gttctcttct gactttgact cctcaaaaa aaaaaatcta 420
 caatcaacag atcgttcaa ttacgcctc acaaaaactt ttttccttct tcttcgcca 480
 cgtaaattt tatccctcat gttgtctaac ggatttctgc acttgattta ttataaaaag 540
 acaaagacat aatacttctc tatcaatttc agttattggt cttccttgcg ttattcttct 600

ES 2 790 573 T3

gttcttcttt ttcttttgtc atatataacc ataaccaagt aatacatatt caaaatggtt 660
aaggttgcca tcttaggtgc ttctgggtgt gtcggtcaac cattatctct attattgaaa 720
ttgtctccat acgtttctga attggctttg tacgatatca gagctgctga aggtattggt 780
aaggatttgt cccacatcaa caccaactcc tcttgtgttg gttacgacaa ggattccatc 840
gaaaacactt tgtccaatgc tcaagttgtc ttgattccag ctggtggtcc aagaaagcca 900
ggtttgacca gagatgattt gttcaagatg aacgctggta tegttaagtc tttggttact 960
gctgtcggta aatttgcccc aaacgctcgt atcttagtca tctccaacc tgtaactct 1020
ttggttccaa ttgccgttga aactttgaag aagatgggta agttcaagcc aggtaacggt 1080
atgggtgtca ccaacttga tttggtcaga gctgaaactt tcttggttga ctactgatg 1140
ttgaagaacc caaagatcgg tcaagaacaa gacaagacca ccatgcacag aaagtcacc 1200
gtcatcgggtg gtcactctgg tgaaacatc attccaatca tcaactgacaa atccttggtt 1260
ttccaattgg acaagcaata cgaacatttc atccacagag tccaattcgg tggtagcгаа 1320
attgtcaagg ccaagcaagg tgccggttct gctacctgt ccatggcttt cgctggtgcc 1380
aaatttgctg aagaagtctt acgttctttc cacaacgaaa agccagaaac tgaatctttg 1440
tctgctttcg tctacttgcc aggtttgaag aacggtaaga aggctcaaca attagtcggt 1500
gacaactcca ttgaatactt ctctttgcca attgttttga gaaacggttc cgttgtttcc 1560
attgacactt ctgttttggа aaaattgtct ccaagagaag aacaattggt caacactgct 1620
gtcaaggaat tgagaaagaa cattgaaaag ggtaagtctt tcatcttggа cagttaaagt 1680
ctgaagaatg aatgatttga tgatttcttt ttccctccat ttttcttact gaatatatca 1740
atgatataga cttgtatagt ttattatttc aaattaagta gctatatata gtcaagataa 1800
cgtttgtttg acacgattac attattcgtc gacatctttt ttcagcctgt cgtggtagca 1860
atttgaggag tattattaat tgaataggtt cattttgctc tcgcataaac agttttcgtc 1920
agggacagta tgttggaatg agtggttaatt aatggtgaca tgacatgta tagcaatacc 1980
tcgaaacctt cgaatccagc cagcatgtcg acaccacaa gatgtagtgc ac 2032

<210> 6
< 211> 3056
5 < 212> ADN
< 213> Escherichia coli

<400> 6
gaaaccttgc aatccagcca gcatgtcgac acccacaaga tgtagtgac gtgccgtaaa 60
aactaaaacg agccccacc aaagaacaaa aaagaagggtg ctgggcccc actttcttcc 120
cttgcacgtg ataggaagat ggctacagaa acaagaagat ggaaatcgaa ggaaagaggg 180
agactggaag ctgtaaaaac tgaaatgaaa aaaaaaaaaa aaaaaaaaaa caagaagctg 240

ES 2 790 573 T3

aaaatggaag actgaaattht gaaaaatggt aaaaaaaaa aagaaacacg aagctaaaaa	300
cctggattcc atthttgagaa gaagcaagaa aggtaagtat ggtaacgacc gtacaggcaa	360
gcgcgaaggc aaatggaaaa gctggagtcc ggaagataat catttcatct tctthttgta	420
gaacagaaca gtggatgtcc ctcatctcgg taacgtattg tccatgccct agaactctct	480
gtccctaaaa agaggacaaa aaccaatgg thtccccagc thccagtga gccaccgatc	540
ccactggaaa ccactggaca ggaagagaaa atcacggact tcctctattg aaggataatt	600
caacactthc accagatccc aaatgtcccg cccctattcc cgtgttccat cacgtaccat	660
aacttaocat thcatcacgt tctctatggc aactgttac tgcttcgact gctthtgctc	720
atcttctcta tgggccaatg agctaatag cacaatgtgc tgcgaaataa agggatatct	780
aatthtatatt attacattat aatatgtact agtgtggta thggaattg tactthaatt	840
tgatatataa aggggtgatc ththtcattt tgaatcagaa thggaattg aacttgtctc	900
thgtcactat tactthaatag taatthatt thttattaac cththththta agtcaaaaa	960
ccaaggacaa gaactactct tcaaaggat thcaagtht catacgtgc acacacgctt	1020
cacagthtca agtaaaaaaa aagaatatta cacaatgtcc aacaagcctt tcatctacca	1080
agctccattc ccaatgggta aggacaacac tgaatactac thgttgactt ctgactacgt	1140
thccgttgct gatttcgatg gtgaaacct cthgaaggth gaaccagaag ccttgactth	1200
gthgtgctcaa caagcctthc acgatgctc thtcatgthg cgtccagctc accaaaagca	1260
agthgtgctc atthttgcacg acccagaagc ctccgaaaa gacaaatacg thgctthtga	1320
atthttgaga aactctgaaa thgtgctcaa ggggtgctta ccaactgtc aagacactgg	1380
tactgccatc atthgtcggta agaagggtca aagagtctgg accggtggth gtgacgaaga	1440
aactctatcc aagggtgtht acaacactta cattgaagat aatthacgth actctcaaaa	1500
tgctgctthg gacatgtaca aggaagtcaa cactgttact aacttgccag ctcaaatcga	1560
ctthatacgt gthgtacggth acgaatacaa gthctthgtg gthgtccaag gthgtggthc	1620
tgctaaacaag acctactgtt accaagaaac caaggctthg thgactccag gthaaattgaa	1680
gaactthctg gtcgaaaaaga tgagaactth gggthactgt cthgtgtccac cataccacat	1740
tgctthctgt atcggthgta cthccgctga aaccaactg aaaaaccgthc aatthggctc	1800
cgtcactac tacgatgaat tgccaactga aggtaacgaa cacggtcaag ccttcagaga	1860
tgtccaattg gaacaagaat thgttgaaga agctcaaaaa thaggththg thgtctcaatt	1920
thgttggtaaa tactthtgctc acgatatcag agthtatcaga thaccaagac atggthgctc	1980
thgtccagth ggtatggthg thtctgthc thgtgacaga aacatcaag ccaagatcaa	2040
cagagaaggt atctggatthg aaaaattgga acacaacca gthcaataca thccacaaga	2100

ES 2 790 573 T3

attgagacaa gctggtgaag gtgaagctgt caaggttgac ttgaacagac caatgaagga 2160
aatcttggt caattatctc aataccaggt ttccaccaga ttatcttga ccggtactat 2220
cattgtcggg cgtgacattg ctcatgcca gttgaaggaa ttgattgatg ctggttaagga 2280
attgcctcaa tacatcaagg accatccaat ctactacgct ggtccagcca agaccccagc 2340
tggttaccca tctggttctt tgggtccaac caccgctggg agaatggact cttacgttga 2400
cttgctacaa tctcacgggt gttccatgat catgttggct aagggttaaca gatctcaaca 2460
agtcaccgat gcttgtcaca agcacgggtg tttctatctt ggttccattg gtggtccagc 2520
tgctgtcttg gctcaacaat ctatcaagca cttggaatgt gttgcttacc cagaattggg 2580
tatggaagcc atctggaaga ttgaagtcga agatttccca gctttcatct tagtcgatga 2640
caagggtaac gacttcttcc aacaaattgt caacaagcaa tgtgccaaact gtaccaagta 2700
aaataaagca atcttgatga ggataatgat ttttttttga atatacataa atactaccgt 2760
ttttctgcta gatcttctga agacgtaaat aagtacatat tactttttaa gccaaagacaa 2820
gattaagcat taactttacc cttttctctt ctaagtttca atactagtta tcaactgttta 2880
aaagttatgg cgagaacgtc ggcgggttaa atatattacc ctgaacgtgg tgaattgaag 2940
ttctaggatg gtttaaagat ttttcctttt tgggaaataa gtaacaata tattgctgcc 3000
ttcctcaaag ccaaagttcg cgttccgacc ttgcctccca aatccgagtt gcgatt 3056

<210> 7

< 211> 1335

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> marcador de selección nurseotricina

<400> 7

tcgtacgctg caggtcgacg aattctaccg ttctgataat gtatgctata cgaagttata 60
gatctgttta gcttgcctcg tccccgccgg gtcaccggc cagcgacatg gaggccaga 120
ataccctcct tgacagtctt gacgtgcgca gctcagggc atgatgtgac tgtcgccgt 180
acatttagcc catacatccc catgtataat catttgcac catacatttt gatggccgca 240
cggcgcgaag caaaaattac ggctcctcgc tgcagacctg cgagcagga aacgctcccc 300
tcacagagcg gttgaattgt cccacgccc cgcccctgta gagaaatata aaaggttagg 360
atttgccact gaggttcttc tttcatatac ttcttttaa aatcttgcta ggatacagtt 420
ctcacatcac atccgaacat aaacaacat ggtaccact cttgacgaca cggcttaccg 480
gtaccgcacc agtgtcccgg gggacgccga ggccatcgag gcaactggatg ggtccttcac 540
caccgacacc gtcttccgcg tcaccgccac cggggacggc ttcaccctgc gggaggtgcc 600
ggtggacccg cccctgacca aggtgttccc cgacgacgaa tcggacgacg aatcggaaga 660

ES 2 790 573 T3

cggggaggac ggcgacccgg actcccggac gttcgtcgcg tacggggacg acggcgacct 720
 ggcgggcttc gtggtcgtct cgtactccgg ctggaaccgc cggctgaccg tcgaggacat 780
 cgaggtcgcc ccggagcacc gggggcacgg ggtcggggcg gcgttgatgg ggctcgcgac 840
 ggagttcgcc cgcgagcggg gcgcccggca cctctggctg gaggtcacca acgtcaacgc 900
 accggcgatc cacgcgtacc ggcggatggg gttcacctc tgcggcctgg acaccgcctt 960
 gtacgacggc accgcctcgg acggcgagca ggcgctctac atgagcatgc cctgccccta 1020
 atcagtactg acaataaaaa gattcttgtt ttcaagaact tgcatttgt atagtttttt 1080
 tatattgtag ttgttctatt ttaatcaaat gttagcgtga tttatatattt ttttcgcctc 1140
 gacatcatct gccagatgc gaagttaagt gcgagaaaag taatatcatg cgtcaatcgt 1200
 atgtgaatgc tggtcgctat actgctgtcg attcgatact aacgccgcca tccagtgtcg 1260
 aaaacgagct cataacttgc tataatgtat gctatacгаа cggtagaatt cgatatcaga 1320
 tccactagtg gccta 1335

<210> 8
 <211> 4429
 <212> ADN
 <213> Trypanosoma brucei

5

<400> 8
 aaagccaaag ttcgcgttcc gaccttgctt cccaaatccg agttgcgatt gtgcttagtc 60
 aaaaaattag ctttttaatt ctgctgtaac ccgtacatgc ccaaaatagg gggcgggtta 120
 cacagaatat ataacatcgt aggtgtctgg gtgaacagtt tattcctggc atccactaaa 180
 tataatggag cccgcttttt aagctggcat ccagaaaaaa aaagaatccc agcaccaaaa 240
 tattgttttc ttcaccaacc atcagttcat aggtccattc tcttagcgca actacagaga 300
 acaggggcac aaacaggcaa aaaacgggca caacctcaat ggagtgatgc aacctgcctg 360
 gagtaaatga tgacacaagg caattgacct acgcatgtat ctatctcatt ttcttacacc 420
 ttctattacc ttctgctctc tctgatttgg aaaaagctga aaaaaaggt tgaaaccagt 480
 tccctgaaat tattccccta cttgactaat aagtatataa agacggtagg tattgattgt 540
 aattctgtaa atctatttct taaacttctt aaattctact tttatagtta gtcttttttt 600
 tagttttaa acaccaagaa cttagtttct aataaacaca cataaaciaa caaatggtt 660
 gatgtagat cttctgcttc cattgttgcc gttgaccag aaagagctgc cagagaaaga 720
 gatgctgctg ccagagcttt gttgcaagac tctccattgc acaccacat gcaatcgcct 780
 acctctggtt tggaaattgac tgttcatac gctttgaagg ttggtgcttc tgctgacact 840
 ttcgacagag ccaaggaagt tgctgatgaa gtcttgagat gtgcctggca attggctgac 900
 accgttttga actctttcaa cccaaactct gaagtctctt tagtcggtag attaccagtc 960

ES 2 790 573 T3

ggtcaaaagc atcaaatgtc tgctccattg aaacgtgtca tggcttggtg tcaaagagtc 1020
 tacaactcct ctgctgggtg tttcgaccca tccactgctc cagttgcca ggctttgaga 1080
 gaaattgctt tgggtaagga aagaacaat gcttgtttg aagctttgac tcaagcttgt 1140
 accttgccaa actctttcgt cattgatttc gaagctggta ctatctccag aaagcacgaa 1200
 cacgcttctt tggatttggg tgggttttcc aagggttaca tcgtcgatta cgtcattgac 1260
 aacatcaatg ctgctggttt caaaacggtt ttctttgact ggggtgggta ctgctgtgcc 1320
 tccggtatga acgccagaaa cactccatgg gttgtcggta tcaactagacc tccttccttg 1380
 gacatggtgc caaacctcc aaaggaagct tottacatct ccgtcatctc tttggacaat 1440
 gaagctttgg ctacctctgg tgattacgaa aacttgatct aactgctga cgataaacca 1500
 ttgacctgta cctacgattg gaaaggtaag gaattgatga agccatctca atccaatctc 1560
 gctcaagttt ccgtcaagtg ttactctgcc atgtacgctg acgctttggc taccgcttgt 1620
 ttcatcaagc gtgaccagc caaggtcaga caattgttg atggttgag atacgttaga 1680
 gacaccgtca gagattaccg tgtctacgtc agagaaaacg aaagagttgc caagatgttc 1740
 gaaattgcca ctgaagatgc tgaatgaga aagagaagaa tttccaacac tttaccagct 1800
 cgtgtcattg ttggttggtg tggtttggct ggtttgtccg ctgccattga agctgtcggg 1860
 tgtggtgctc aagttgttt gatgaaaag gaagccaagt tgggtggtaa ctctgccaag 1920
 gctacctctg gtatcaacgg ttgggtact agagcccaag ctaaggcttc cattgtcgat 1980
 ggtggttaagt acttcgaaag agatacctac aagtctggta tcggtggtaa caccgatcca 2040
 gctttggta agactttgtc catgaaatct gctgacgcta tcggttggtt gacttctcta 2100
 ggtgttccat tgactgtttt gtccaatta ggtggtcact ccagaaagag aactcacaga 2160
 gccccagaca agaaggatgg tactccattg ccaattggtt tcaccatcat gaaaacttta 2220
 gaagatcatg ttagaggtaa cttgtccggt agaatacca tcatggaaaa ctgttccggt 2280
 acctctttgt tgtctgaaac caaggaaaga ccagacggtta ctaagcaaat cagagttacc 2340
 ggtgtcgaat tcaactcaagc tggttctggt aagaccacca ttttggtga tgctgttacc 2400
 ttggccaccg gtggtttctc caacgacaag actgctgatt ctttggtgag agaacatgcc 2460
 ccacacttg ttaacttccc aaccaccaac ggtccatggg ctactggtga tgggtgcaag 2520
 ttggctcaa gattaggtgc tcaattggtc gatatggaca aggttcaatt gcaccaact 2580
 ggtttgatca acccaaagga ccagccaac ccaaccaa atcttgggtcc agaagctcta 2640
 agaggttctg gtggtgtttt gttgaacaaa caagtaaga gatttgtcaa cgaattggat 2700
 ttgagatctg ttgtttccaa ggccatcatg gaacaagggt ctgaatcccc aggttctggt 2760
 ggttccatgt ttgcttactg tgtcttgaac gctgctgctc aaaaattggt tgggtgttcc 2820

ES 2 790 573 T3

tctcacgaat tctactggaa gaagatgggt ttgttcgtca aggctgacac catgagagac 2880
 ttggctgctt tgattgggtt tccagttgaa tccgttcaac aaactttaga agaatacgaa 2940
 agattatcca tctctcaaag atcttgcctc attaccagaa aatctgttta cccatgtggt 3000
 ttgggtacta aaggtccata ctatgtcgcc tttgtcactc catctatcca ctacaccatg 3060
 ggtggttgtt tgatttctcc atctgctgaa atccaaatga agaacacttc ttccagagcc 3120
 ccattgtccc actccaaccc aatcttgggt ttattcgggt ctggtgaagt caccgggtgt 3180
 gtccacgggt gtaacagatt aggtggtaac tctttgttgg aatgtgttgt tttcggtaga 3240
 attgccgggt acagagcttc taccattttg caaagaaagt cctctgcttt gtctttcaag 3300
 gtctggacca ctgttgtttt gagagaagtc agagaagggt gtgtctacgg tgctggttcc 3360
 cgtgtcttga gattcaactt accaggtgct ctacaaagat ctggtctatc cttgggtcaa 3420
 ttcattgcca tcagaggtga ctgggacgggt caacaattga ttggttacta ctctccaatc 3480
 actttgccag acgatttggg tatgattgac attttggcca gatctgacaa gggacttta 3540
 cgtgaatgga tctctgcttt ggaaccaggt gacgctgtcg aatgaaggc ttgtggtggt 3600
 ttggtcatcg aaagaagatt atctgacaag cacttcgctt tcatgggtca cattatcaac 3660
 aagctatggt tgattgctgg tggtactggt gttgctccaa tgttgcaa atcaagccc 3720
 gctttcatga agccattcat cgacactttg gaatcogtcc acttgatcta cgctgctgaa 3780
 gatgtcactg aattgactta cagagaagtt ttggaagaac gtcgtcgtga atccagaggt 3840
 aaattcaaga aaactttcgt tttgaacaga cctcctccat tatggactga cgggtcgggt 3900
 ttcacgcacc gtggtatcct gaccaaccac gttcaaccac catctgacaa cttattggtt 3960
 gccatctgtg gtccaccagt tatgcaaaga attgtcaagg ccactttaa gactttaggt 4020
 tacaacatga acttggctcag aaccgttgac gaaactgaac catctggaag ttaaaggaag 4080
 tatctcggaa atattaattht aggccatgct cttatgcacg tttcttttga tacttacggg 4140
 tacatgtaca caagtatctc tatatatata aattaatgaa aatcccctat ttatatatat 4200
 gactttaacg agacagaaca gttttttatt tttatccta tttgatgaat gatacagttt 4260
 cttattcacg tgttatacc acaccaaatc caatagcaat accggccatc acaatcactg 4320
 tttcggcagc ccctaagatc agacaaaaca tccggaacca ccttaaatca acgtccctca 4380
 gaaagcctgt atgcgaagcc acaatccttt ccaacagacc atactaagt 4429

<210> 9
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

5

<220>
 < 223> cebador

<400> 9
 cattatatcg aggaaagccc 20

ES 2 790 573 T3

<210> 10
 < 211> 75
 < 212> ADN
 < 213> Secuencia artificial

5 <220>
 < 223> cebador

<400> 10
 aaagcaaagg aaggagagaa cagaggagta cttgtacggt cgatgggcaa agaaagagac 60
 acaaaactac gtggg 75

10 <210> 11
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebadorr

15 <400> 11
 ttgccatcg aacgtacaag 20

20 <210> 12
 < 211> 23
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

<400> 12
 tgctaagatt tgtgtcgtt tgg 23

25 <210> 13
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

30 <220>
 < 223> cebador

<400> 13
 cggatcgatg tacacaaccg 20

35 <210> 14
 < 211> 23
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

40 <400> 14
 caacaggagg cggatggata tac 23

ES 2 790 573 T3

<210> 15
 < 211> 79
 < 212> ADN
 < 213> Secuencia artificial

5 <220>
 < 223> cebador

<400> 15
 acgctttccg gcatcttcca gaccacagta tatccatccg cctcctgttg tcgtacgctg 60
 caggtcgcgcg aattctacc 79

10 <210> 16
 < 211> 75
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

15 <400> 16
 gcggaatatt ggcggaacgg acacacgtgg atacaaacct ggacaacggt taggccacta 60
 gtggatctga tatcg 75

<210> 17
 < 211> 22
 < 212> ADN
 < 213> Secuencia artificial

20 <220>
 < 223> cebador

<400> 17
 aacgtgtcc aggtttgtat cc 22

25 <210> 18
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

30 <400> 18
 aggtacaaca agcagaccg 20

<210> 19
 < 211> 23
 < 212> ADN
 < 213> Secuencia artificial

35 <220>
 < 223> cebador

ES 2 790 573 T3

<400> 19
 aaataaccac aaacatcctt ccc 23

5 <210> 20
 < 211> 23
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

10 <400> 20
 gtgcactaca tcttgggt gtc 23

<210> 21
 < 211> 22
 < 212> ADN
 < 213> Secuencia artificial

15 <220>
 < 223> cebador

<400> 21
 gaaacctcg aatccagcca gc 22

20 <210> 22
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

25 <400> 22
 aatcgcaact cggattggg 20

30 <210> 23
 < 211> 73
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> cebador

<400> 23
 aaagccaaag ttcgcgttcc gaccttgcct cccaaatccg agttgcgatt tcttggcggc 60

cgtacgcata tag 73

35 <210> 24
 < 211> 20
 < 212> ADN
 < 213> Secuencia artificial

40 <220>
 < 223> cebador

ES 2 790 573 T3

<220>
< 223> cebador

<400> 29
agaaagcctg tatgcgaagc cacaatcctt tccaacagac cataactaagt attttatttt 60
acttttttta gaatgacctg ttcccgcacac 90

5 <210> 30
< 211> 24
< 212> ADN
< 213> Secuencia artificial

10 <220>
< 223> cebador

<400> 30
cacaagctta ttctcctaaa aatc 24

15 <210> 31
< 211> 20
< 212> ADN
< 213> Secuencia artificial

<220>
< 223> cebador

20 <400> 31
aaagccaaag ttcgcgttcc 20

<210> 32
< 211> 24
< 212> ADN
< 213> Secuencia artificial

25 <220>
< 223> cebador

<400> 32
acttagtatg gtctgttggg aagg 24

30 <210> 33
< 211> 1139
< 212> PRT
< 213> Secuencia artificial

<220>
< 223> secuencia de proteína FRDg (menos SKI C-terminal)

ES 2 790 573 T3

<400> 33

Met Val Asp Gly Arg Ser Ser Ala Ser Ile Val Ala Val Asp Pro Glu
1 5 10 15

Arg Ala Ala Arg Glu Arg Asp Ala Ala Ala Arg Ala Leu Leu Gln Asp
20 25 30

Ser Pro Leu His Thr Thr Met Gln Tyr Ala Thr Ser Gly Leu Glu Leu
35 40 45

ES 2 790 573 T3

Thr Val Pro Tyr Ala Leu Lys Val Val Ala Ser Ala Asp Thr Phe Asp
50 55 60

Arg Ala Lys Glu Val Ala Asp Glu Val Leu Arg Cys Ala Trp Gln Leu
65 70 75 80

Ala Asp Thr Val Leu Asn Ser Phe Asn Pro Asn Ser Glu Val Ser Leu
85 90 95

Val Gly Arg Leu Pro Val Gly Gln Lys His Gln Met Ser Ala Pro Leu
100 105 110

Lys Arg Val Met Ala Cys Cys Gln Arg Val Tyr Asn Ser Ser Ala Gly
115 120 125

Cys Phe Asp Pro Ser Thr Ala Pro Val Ala Lys Ala Leu Arg Glu Ile
130 135 140

Ala Leu Gly Lys Glu Arg Asn Asn Ala Cys Leu Glu Ala Leu Thr Gln
145 150 155 160

Ala Cys Thr Leu Pro Asn Ser Phe Val Ile Asp Phe Glu Ala Gly Thr
165 170 175

Ile Ser Arg Lys His Glu His Ala Ser Leu Asp Leu Gly Gly Val Ser
180 185 190

Lys Gly Tyr Ile Val Asp Tyr Val Ile Asp Asn Ile Asn Ala Ala Gly
195 200 205

Phe Gln Asn Val Phe Phe Asp Trp Gly Gly Asp Cys Arg Ala Ser Gly
210 215 220

Met Asn Ala Arg Asn Thr Pro Trp Val Val Gly Ile Thr Arg Pro Pro
225 230 235 240

Ser Leu Asp Met Leu Pro Asn Pro Pro Lys Glu Ala Ser Tyr Ile Ser
245 250 255

Val Ile Ser Leu Asp Asn Glu Ala Leu Ala Thr Ser Gly Asp Tyr Glu
260 265 270

Asn Leu Ile Tyr Thr Ala Asp Asp Lys Pro Leu Thr Cys Thr Tyr Asp
275 280 285

Trp Lys Gly Lys Glu Leu Met Lys Pro Ser Gln Ser Asn Ile Ala Gln

ES 2 790 573 T3

Leu Leu Ser Glu Thr Lys Glu Arg Pro Asp Gly Thr Lys Gln Ile Arg
 545 550 555 560
 Val Thr Gly Val Glu Phe Thr Gln Ala Gly Ser Gly Lys Thr Thr Ile
 565 570 575
 Leu Ala Asp Ala Val Ile Leu Ala Thr Gly Gly Phe Ser Asn Asp Lys
 580 585 590
 Thr Ala Asp Ser Leu Leu Arg Glu His Ala Pro His Leu Val Asn Phe
 595 600 605
 Pro Thr Thr Asn Gly Pro Trp Ala Thr Gly Asp Gly Val Lys Leu Ala
 610 615 620
 Gln Arg Leu Gly Ala Gln Leu Val Asp Met Asp Lys Val Gln Leu His
 625 630 635
 Pro Thr Gly Leu Ile Asn Pro Lys Asp Pro Ala Asn Pro Thr Lys Phe
 645 650 655
 Leu Gly Pro Glu Ala Leu Arg Gly Ser Gly Gly Val Leu Leu Asn Lys
 660 665 670
 Gln Gly Lys Arg Phe Val Asn Glu Leu Asp Leu Arg Ser Val Val Ser
 675 680 685
 Lys Ala Ile Met Glu Gln Gly Ala Glu Tyr Pro Gly Ser Gly Gly Ser
 690 695 700
 Met Phe Ala Tyr Cys Val Leu Asn Ala Ala Ala Gln Lys Leu Phe Gly
 705 710 715 720
 Val Ser Ser His Glu Phe Tyr Trp Lys Lys Met Gly Leu Phe Val Lys
 725 730 735
 Ala Asp Thr Met Arg Asp Leu Ala Ala Leu Ile Gly Cys Pro Val Glu
 740 745 750
 Ser Val Gln Gln Thr Leu Glu Glu Tyr Glu Arg Leu Ser Ile Ser Gln
 755 760 765
 Arg Ser Cys Pro Ile Thr Arg Lys Ser Val Tyr Pro Cys Val Leu Gly
 770 775 780
 Thr Lys Gly Pro Tyr Tyr Val Ala Phe Val Thr Pro Ser Ile His Tyr
 785 790 795 800

ES 2 790 573 T3

Thr Met Gly Gly Cys Leu Ile Ser Pro Ser Ala Glu Ile Gln Met Lys
 805 810 815
 Asn Thr Ser Ser Arg Ala Pro Leu Ser His Ser Asn Pro Ile Leu Gly
 820 825 830
 Leu Phe Gly Ala Gly Glu Val Thr Gly Gly Val His Gly Gly Asn Arg
 835 840 845
 Leu Gly Gly Asn Ser Leu Leu Glu Cys Val Val Phe Gly Arg Ile Ala
 850 855 860
 Gly Asp Arg Ala Ser Thr Ile Leu Gln Arg Lys Ser Ser Ala Leu Ser
 865 870 875 880
 Phe Lys Val Trp Thr Thr Val Val Leu Arg Glu Val Arg Glu Gly Gly
 885 890 895
 Val Tyr Gly Ala Gly Ser Arg Val Leu Arg Phe Asn Leu Pro Gly Ala
 900 905 910
 Leu Gln Arg Ser Gly Leu Ser Leu Gly Gln Phe Ile Ala Ile Arg Gly
 915 920 925
 Asp Trp Asp Gly Gln Gln Leu Ile Gly Tyr Tyr Ser Pro Ile Thr Leu
 930 935 940
 Pro Asp Asp Leu Gly Met Ile Asp Ile Leu Ala Arg Ser Asp Lys Gly
 945 950 955 960
 Thr Leu Arg Glu Trp Ile Ser Ala Leu Glu Pro Gly Asp Ala Val Glu
 965 970 975
 Met Lys Ala Cys Gly Gly Leu Val Ile Glu Arg Arg Leu Ser Asp Lys
 980 985 990
 His Phe Val Phe Met Gly His Ile Ile Asn Lys Leu Cys Leu Ile Ala
 995 1000 1005
 Gly Gly Thr Gly Val Ala Pro Met Leu Gln Ile Ile Lys Ala Ala
 1010 1015 1020
 Phe Met Lys Pro Phe Ile Asp Thr Leu Glu Ser Val His Leu Ile
 1025 1030 1035
 Tyr Ala Ala Glu Asp Val Thr Glu Leu Thr Tyr Arg Glu Val Leu
 1040 1045 1050

ES 2 790 573 T3

Glu Glu Arg Arg Arg Glu Ser Arg Gly Lys Phe Lys Lys Thr Phe
 1055 1060 1065

Val Leu Asn Arg Pro Pro Pro Leu Trp Thr Asp Gly Val Gly Phe
 1070 1075 1080

Ile Asp Arg Gly Ile Leu Thr Asn His Val Gln Pro Pro Ser Asp
 1085 1090 1095

Asn Leu Leu Val Ala Ile Cys Gly Pro Pro Val Met Gln Arg Ile
 1100 1105 1110

Val Lys Ala Thr Leu Lys Thr Leu Gly Tyr Asn Met Asn Leu Val
 1115 1120 1125

Arg Thr Val Asp Glu Thr Glu Pro Ser Gly Ser
 1130 1135

<210> 34

<211> 30

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Cebador

<400> 34

ataaaaattg catatatgtg gtagaaacc 30

10 <210> 35

<211> 78

<212> ADN

<213> Secuencia artificial

<220>

15 <223> Cebador

<400> 35

tatagcatac attatacga cggtagaatt cgtcgacctg cagcgtacga tcaaagcaga 60

aatctgatgg gttcgaac 78

<210> 36

<211> 75

20 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador

<400> 36

25 agaaaacctg tatgcgaagc cacaatcctt tccaacagac cataactaagt acaggtgatt 60

gtatgtgggc ttatg 75

ES 2 790 573 T3

<210> 37
 < 211> 29
 < 212> ADN

5 <220>
 < 223> Cebador

<400> 37
 acattattgt aaaaacggag tagaaaggg 29

10 <210> 38
 < 211> 595
 < 212> ADN
 < 213> Secuencia artificial

15 <220>
 < 223> Fragmento 25 de PCR que consiste en el flanco 5' de integración para fijar como objetivo el locus INT09.01 de CEN.PK113-7D

<400> 38
 gggaaaggaac ataatccggt atgtgaaatc aattgaaatct atcctctgtc tcggtataat 60
 attatacaag ttaggagaat agttatataat aaatatatgt gaactttagtag aactgatggc 120
 atttgatggc cgtcacatgg tacgccacta aaaccttgat taagtagatt agattagact 180
 agcaagtgtt ggtaacgttt tgtaggaaag agtctcaaag taaccagcgc caacttttgc 240
 taaagtcaat cataaaatat atgtttttta aatatgtagg tatgccgtgc atttttattc 300
 ttttttaatg tttagctttc aaaaaagcta aaataaacca taagcgttat tcttattata 360
 ttgatagtag aagaggaaga gtggttatct ttactcaaat taacgattgc ttttatttct 420
 tacattaagg tcgaaactat ttgctactat gaaagccaag cagaaacaaa caagacctta 480
 ttaagattga tgtttcaagc tatgcattcc caattcagaa gaacaagga agagcataag 540
 acggctcctc ggatcggcca gttgggagca agtggtact tctgtcgtcg tattt 595

20 <210> 39
 < 211> 1381
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Fragmento 26 de PCR que incluye un marcador de selección de nourseotricina funcional en Saccharomyces cerevisiae.

25 <400> 39
 tcctcggatc ggccagttgg gagcaagggt gtacttctgt cgtcgtattt gtgctaccgt 60
 tcgtataatg tatgctatac gaagttatgt ccccgccggg tcaccggcc agcgacatgg 120
 aggccagaa taccctcctt gacagtcttg acgtgcgcag ctcaggggca tgatgtgact 180
 gtcgcccgtc catttagccc atacatcccc atgtataatc atttgcatcc atacattttg 240

ES 2 790 573 T3

atggccgcac ggcgcgaagc aaaaattacg gtcctctgct gcagacctgc gagcagggaa 300
 acgctcccct cacagacgcg ttgaattgtc cccacgccgc gccctgtag agaaatataa 360
 aaggtagga tttgccactg aggttcttct tcatatact tccttttaaa atcttgctag 420
 gatacagttc tcaatcaca tccgaacata aacaacaatg aatgggtacc actcttgacg 480
 acacggctta ccggtaccgc accagtgtcc cgggggacgc cgaggccatc gaggcactgg 540
 atgggtcctt caccaccgac accgtcttcc gcgtcaccgc caccggggac ggcttcaccc 600
 tgcgggaggt gccggtggac ccgcccctga ccaaggtgtt ccccgacgac gaatcggacg 660
 acgaatcgga cgacggggag gacggcgacc cggactcccg gacgttcgtc gcgtacgggg 720
 acgacggcga cctggcgggc ttcgtggtcg tctcgtactc cggctggaac cgccggctga 780
 ccgtcgagga catcgaggtc gccccggagc accgggggca cggggtcggg cgcgcgttga 840
 tggggctcgc gacggagttc gcccgcgagc gggcgcccg gcacctctgg ctggaggtca 900
 ccaacgtcaa cgcaccggcg atccacgcgt accggcggat ggggttcacc ctctgcggcc 960
 tggacaccgc cctgtacgac ggcaccgcct cggacggcga gcaggcgctc tacatgagca 1020
 tgcctgccc ctaaataaat cagtactgac aataaaaaga ttcttgtttt caagaacttg 1080
 tcatttgat agttttttta tattgtagtt gttctatttt aatcaaagt tagcgtgatt 1140
 tatatttttt ttgcctcga catcatctgc ccagatgca agttaagtgc gcagaaagta 1200
 atatcatgcy tcaatcgtat gtgaatgctg gtcgctatac tgctgctgat tcgatactaa 1260
 cgccgccatc cagtgtcgaa aacgagctca taacttcgta taatgtatgc tatacgaacg 1320
 gtacctcoct catttggaac agcaggcgag gtgtattggt gcgagaagag gtgtttccga 1380
 g 1381

<210> 40
 < 211> 4838
 < 212> ADN
 < 213> Secuencias artificial

<220>
 < 223> Fragmento 27 de PCR que incluye la secuencia codificante de fumarato reductasa de *Trypanosoma brucei* (FRDg) optimizada en un par de codones para la expresión en *S. cerevisiae*.

10 <400> 40
 atttgaaca gcaggcgagg tgtattggtg cgagaagagg tgtttccgag gtgcgttccct 60
 catcactaga agccgaactg ttgtcttcag tggggattgg ttogacattt tgccaattgc 120
 tgtcgatgta ccctttcaa gccatgtacc ttaaactctc atccttgca agtagattca 180
 tcgggtgtgt ttgaagtaag aatattgct tgttttatg gtatcaaagg tatatgtgt 240
 agaagacaat ttccggtaat ccaattgtct gtctgctcag tttagcacat gtatagtacg 300

ES 2 790 573 T3

ttgcacatag tctacaatat tcagcattca gcattcagta tacagcatat ggctaaatga 360
 tcacaaatgt gattgatgat ttgacacgac tagaaaagag aacgaaaaag ggaaattcca 420
 tgtcacgtgc gttggcacgt gacatggaat atcgaagaaa gaaaaaaaaa acgatctcgt 480
 cctagtggaa gcccagagtc tgggtccccc ggagtcttcc caaaacaaga agctgacaca 540
 tgttgacaca gaacacccca cagcaaatgc accacgctac gtagatcagg aagcttaact 600
 ctagcgacct gtcgctcgcc ccacagaacc tcacccgaga accacacatt acacgccgcc 660
 agctcccact ataactcatct tgcttccctt aagcgttctc acgattcgtt cgctgccctt 720
 cttcaagagt ottctgattc taattctcat tcgaaatcct ctacagttaa tgaattgott 780
 gacatgacat tcattgtctc atggttttgg ctttttggct tttgtctttt aaagctatat 840
 caactttaca tataaatata cgtcaaaagg ggattcatta attagaaaat tctctttttc 900
 aatagtgtct attcattatc aatctattca actcaattgg ttattatttt catctttttg 960
 tcatcctaaa ccatcaacaa tatttaaata tatctgttgc tacattaaga gttacttcag 1020
 aaataacaaa aaaatcgatc aagaattaat aaaaatgaat ggtcgacggt agatcctctg 1080
 cttctatcgt cgctgtcgac ccagaaagag ccgctcgtga acgtgacgct gctgctcgtg 1140
 ctttgctaca agattctcca ttgcacacta ccatgcaata tgctacttcc ggtttggaat 1200
 taaccgtccc atacgccttg aaggttgtcg ottctgctga caccttcgat cgtgctaagg 1260
 aagttgctga cgaagtttta agatgtgctt ggcaattagc cgacactggt ttaaactctt 1320
 ttaacccaaa ctccgaagtc tctttggttg gtagattgcc agttggtcaa aaacacccaa 1380
 tgtccgctcc attgaagaga gtcatggcct gttgtcaacg tgtttacaac tcttctgctg 1440
 gttgcttoga tccatctact gctccagtcg ccaaagcttt gagagaaatc gctttaggta 1500
 aggaaacgtaa caacgcttgc ttggaagctt tgacccaagc ttgtactcta ccaaactctt 1560
 tcgttatcga cttcgaagct ggcaccattt ctagaaagca cgaacatgct tctttggact 1620
 tgggtgggtg ctctaagggt tacattggtg actacgttat cgacaacatc aacgctgccg 1680
 gtttccaaaa cgtcttcttc gactgggggtg gcgactgtag agcttctggt atgaacgcta 1740
 gaaacacccc atgggttggt ggtattacca gaccaccatc tttggacatg ctaccaaacc 1800
 caccaaagga agcttctctac atctctgtta tttctttgga caacgaagct ttggccactt 1860
 ccggtgacta cgaaaacttg atttacaccg ctgatgacaa gccattgact tgtacttacg 1920
 actggaaggg taaggaatta atgaagccat ctcagtccaa cattgctcaa gtctctgtta 1980
 agtgttactc cgctatgtac gctgatgctt tagccactgc ttgtttcatt aagagagatc 2040
 cagctaaggt tcgtcaattg ttggacgggt ggagatacgt ccgtgatact gttcgtgact 2100
 acagagttta cgttcgtgaa aacgaaagag tcgctaagat gtttgaaatc gctaccgaag 2160
 acgctgaaat gagaaagcgt agaatctcta acactttgcc agctagagtt attgtcgtcg 2220

ES 2 790 573 T3

gtggtggttt	ggccggtttg	tccgccgcta	togaagctgc	cggttggtgt	gctcaggtcg	2280
tcttgatgga	aaaagaagct	aagttgggtg	gtaacagcgc	taaggctact	tctggtatta	2340
acggttgggg	caccagagcc	caagccaaag	cctctatcgt	cgacgggtgt	aaatacttcg	2400
aacgtgacac	ttacaagtct	ggtattggtg	gtaacactga	cccagctttg	gtcaagacct	2460
tgtccatgaa	gtctgctgat	gccatcgggt	ggttaacctc	cttgggtgtc	ccattaaccg	2520
ttttgtctca	attgggtggt	cactctagaa	agagaaccca	ccgtgctcca	gacaagaaag	2580
atggtactcc	attaccaatc	ggtttacta	tcatgaagac	tttgaagac	cacgttcgtg	2640
gtaactgtgc	tggtcgtatc	actatcatgg	aaaactgtag	cgtcacctct	ctattatctg	2700
aaaccaagga	acgtccagat	ggtactaaac	aatccgtgt	cactggtgtt	gaatttacc	2760
aagctggttc	cggttaagact	actatcttgg	ctgatgctgt	cattttggct	actggtggtt	2820
tcagtaatga	caagaccgct	gactccttgt	tgagagaaca	cgctccacac	ttggtcaact	2880
tcccaactac	caacggccca	tgggctaccg	gtgacggtgt	taagttggcc	caaagattgg	2940
gtgctcaatt	ggttgacatg	gataaggttc	aattgcatcc	aactggtttg	attaacccaa	3000
aggatccagc	taaccaaac	aagtcttgg	gtccagaggc	cttgagaggt	tccggtggtg	3060
tcttgttgaa	caagcaaggt	aagagattcg	tcaacgaatt	agatctaaga	tccgttggtt	3120
ccaagctat	catggagcaa	ggtgccgaat	accaggttc	cggtggttcc	atgttcgctt	3180
actgtgtttt	gaacgctgcc	gctcaaaagt	tattcgggtg	ctcttctcat	gaattctatt	3240
ggaagaagat	gggtctattt	gttaaggctg	ataccatgag	agatctagct	gctttgatcg	3300
gctgtccagt	tgaatctgtc	caacaaacct	tggaagaata	cgaaogtttg	tctatttctc	3360
aacgttcttg	tccaatcacc	cgtaagtctg	tctaccatg	tgtcttaggg	accaagggtc	3420
catactacgt	tgctttcgtc	accccatcta	ttcactaac	tatgggcggt	tgtttgatct	3480
ccccatccgc	cgaaattcaa	atgaaaaaca	cttcctctcg	tgctccattg	tctcactcta	3540
atccaatctt	gggtttgttc	ggtgctggcg	aagttaccgg	tggtgttcat	ggtggttaacc	3600
gtttgggtgg	taactccttg	ttggagtgtg	ttgtcttcgg	tagaatcgct	ggtgacagag	3660
cctccaccat	cttgcaacgt	aagtcttctg	ccttgcctt	caaggtttg	actaccgttg	3720
tcttgagaga	agttagagaa	ggtggcgtct	acggtgccgg	ttctagagtt	ttgagattca	3780
acttgccctg	tgctttgcaa	cgttccggtt	tgtctttagg	tcaattcatc	gctatcagag	3840
gtgattggga	cggtcaacaa	ctaattggtt	actactccc	aattactttg	ccagatgact	3900
tgggtatgat	cgatatctta	gctcgttctg	acaaaggtac	tttgagagaa	tggatttctg	3960
ctttggaacc	tggtgatgct	gttgaaatga	aggcctgtgg	tggtttggtt	attgaaagaa	4020
gattgtccga	caagcacttc	gtcttcatgg	gtcacatcat	caacaaattg	tgtttgattg	4080

ES 2 790 573 T3

```

ccggtggtac tgggtgctgcc ccaatgttgc aaattatcaa ggctgctttc atgaagcctt      4140
tcatcgatac cttggaatct gttcacttga tttacgctgc tgaggatgtt accgaattga      4200
cctacagaga agtcttggaa gaaagaagaa gagaatccag aggtaagttc aagaagactt      4260
ttgtcttaaa cagaccacca ccattgtgga cgcgaggtgt tggtttcatc gatagaggta      4320
tcttgactaa ccacgttcaa cctccatccg acaacttggt ggctgctatc tgtggtcctc      4380
cagtcattga aagaattggt aaggctactt tgaaaacctt gggttacaat atgaacttgg      4440
ttagaactgt tgatgaaacc gaaccttccg gttaaataaa gcgaatttct tatgatttat      4500
gatttttatt attaaataag ttataaaaaa aataagtgtg tacaattttt aaagtgactc      4560
ttaggtttta aaacgaaaat tcttattctt gagtaactct ttctgttagg tcaggttgcct      4620
ttctcaggta tagcatgagg tcgctcttat tgaccacacc tctaccgga tgccgagcaa      4680
atgctgcaa atcgctcccc atttcacca attgtagata tgctaactcc agcaatgagt      4740
tgatgaatct cgggtgtgat tttatgtcct cagaggacaa cctccctctt tccttgtacc      4800
aaagaacgag tatataccca ggtatccaga gcgcttgt      4838

```

<210> 41

< 211> 586

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Fragmento 28 de PCR que consiste en el flanco 3' de integración para fijar como objetivo el locus INT09.01 de CEN.PK113-7D

<400> 41

```

tttccttgta ccaaagaacg agtatatacc caggtatcca gagcgcttgt atctgcattt      60
aacagtagtg gtcactccta cagtagtcat acttctcaag cctactttcg gtatggcttt      120
taaagtcata aaataacgaa tagttcgtat accgctcatg tggcaaactc aatagcaaaa      180
tgtgatttca tatcacgaag ttttttata tttttaatta tacaattgta atttgtatta      240
cggaaatgaa acatgtttgt atgttatgta gtatatcaat aattgattta ccgcatctgt      300
aaaagagaac catttatcac gcacatagtt ttagtattca ctaatcggtg ctctatggaa      360
gtacgagaca ccgtgcatag atggaagggg aaggtaataa gcttctctct aagaattttc      420
tttgtaaacc ctttgcactt ctgatttaac agcgggtaag atttctcttc tctttaattt      480
ttgagcagat gtgacaaagc cattttcagg tgtccattct tcatcaaaga aaacaatacc      540
acataatagt tcaataccaa ccaaaccttg agatttggct gttgcg      586

```

10

15

REIVINDICACIONES

1. Un polipéptido variante que tiene actividad de fumarato reductasa, que tiene actividad dependiente de NADP(H) modificada y/o actividad dependiente de NAD(H) modificada en comparación con un polipéptido de referencia que tiene actividad de fumarato reductasa y que comprende la fumarato reductasa de SEQ ID NO: 33,
 5 en donde el polipéptido variante comprende una secuencia de aminoácidos que, cuando se alinea con la fumarato reductasa que comprende la secuencia establecida en SEQ ID NO: 33, comprende al menos una sustitución de un residuo aminoácido correspondiente a cualquiera de los aminoácidos
 1042, 1071, 1072, 1082 o 1083
 estando definidas dichas posiciones con referencia a SEQ ID NO: 33, y
 10 en donde el polipéptido variante tiene al menos 80% de identidad de secuencia con SEQ ID NO: 33.
2. Un polipéptido variante de acuerdo con la reivindicación 1, que tiene actividad dependiente de NADP(H) incrementada en comparación con el polipéptido de referencia.
3. Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes, que tiene un aumento en la actividad dependiente de NADP(H) con respecto a la actividad dependiente de NAD(H) en comparación con el polipéptido de referencia.
 15
4. Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde el polipéptido variante es un polipéptido que no se produce de forma natural.
5. Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes, que comprende sustituciones adicionales distintas de las definidas en la reivindicación 1, en donde el polipéptido variante tiene al menos 80% de identidad de secuencia con SEQ ID No. 33.
 20
6. Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes según se define en la Tabla 1.
7. Un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes, que tiene al menos 85%, al menos 90%, al menos 95%, al menos 97%, al menos 98% o al menos 99% de identidad de secuencia con SEQ ID NO: 33).
 25
8. Un ácido nucleico que comprende una secuencia que codifica un polipéptido variante de acuerdo con una cualquiera de las reivindicaciones precedentes.
9. Una construcción de ácido nucleico que comprende la secuencia de ácido nucleico de la reivindicación 8 enlazada operativamente a una o más secuencias de control capaces de dirigir la expresión de una fumarato reductasa en un huésped de expresión adecuado.
 30
10. Un vector de expresión que comprende un ácido nucleico de acuerdo con la reivindicación 8 o una construcción de ácido nucleico de acuerdo con la reivindicación 9.
11. Una célula huésped que comprende un ácido nucleico de acuerdo con la reivindicación 8, una construcción de ácido nucleico de acuerdo con la reivindicación 9 o un vector de expresión de acuerdo con la reivindicación 10.
 35
12. Una célula huésped de acuerdo con la reivindicación 11, que es una célula procariota, tal como una célula bacteriana, o una célula eucariota, tal como una célula de levadura o una célula fúngica filamentosa.
13. Una célula huésped de acuerdo con la reivindicación 12, en donde la célula de levadura es una célula huésped de *Saccharomyces cerevisiae*.
14. Un método para producir una fumarato reductasa, que comprende cultivar una célula huésped de acuerdo con una cualquiera de las reivindicaciones 11 a 13 bajo condiciones adecuadas para la producción de fumarato reductasa y, opcionalmente, recuperar la fumarato reductasa.
 40
15. Un método para la producción de un ácido dicarboxílico, tal como ácido succínico, método que comprende fermentar una célula huésped de acuerdo con una cualquiera de las reivindicaciones 11 a 13 bajo condiciones adecuadas para la producción de ácido succínico y, opcionalmente, recuperar el ácido succínico.
 45

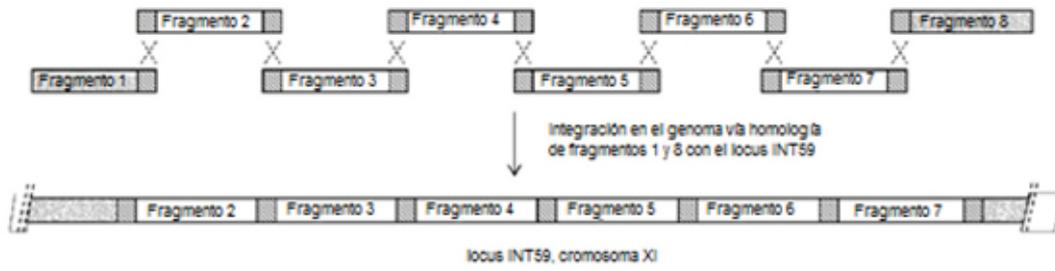


Fig. 1

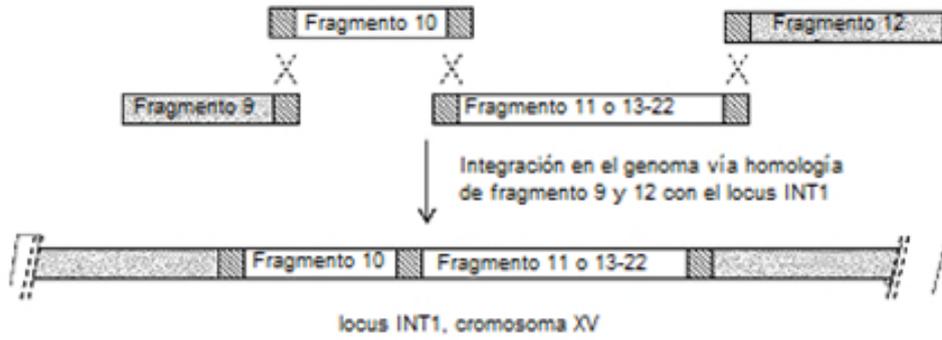


Fig. 2

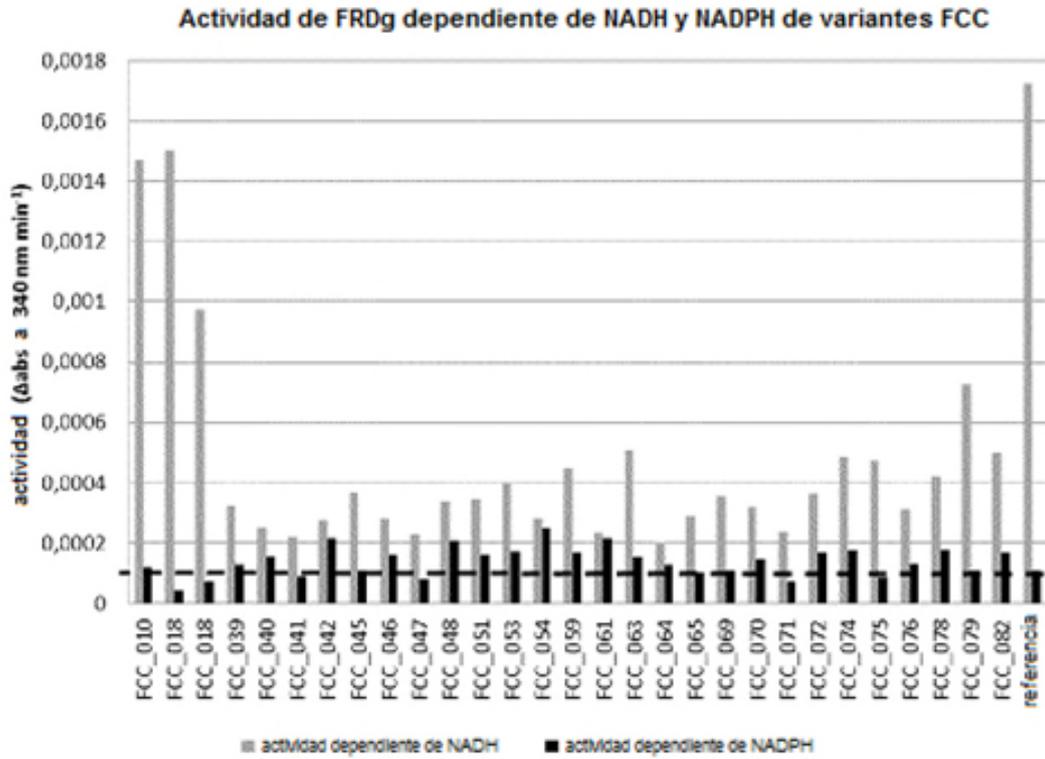


Fig. 3a

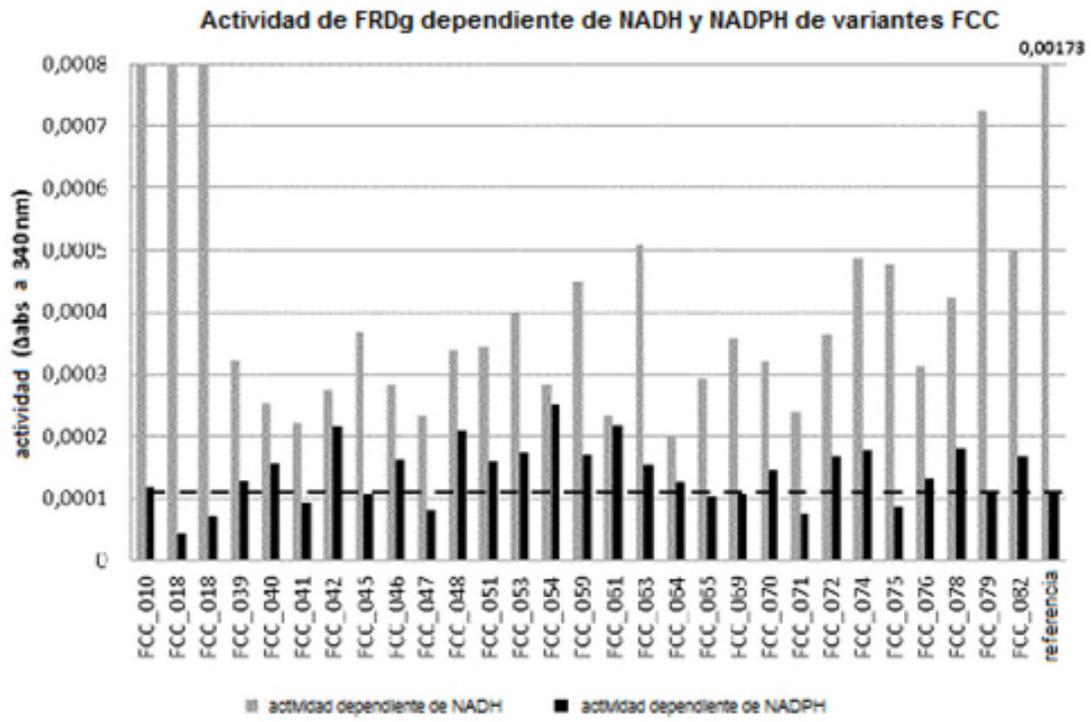


Fig. 3b

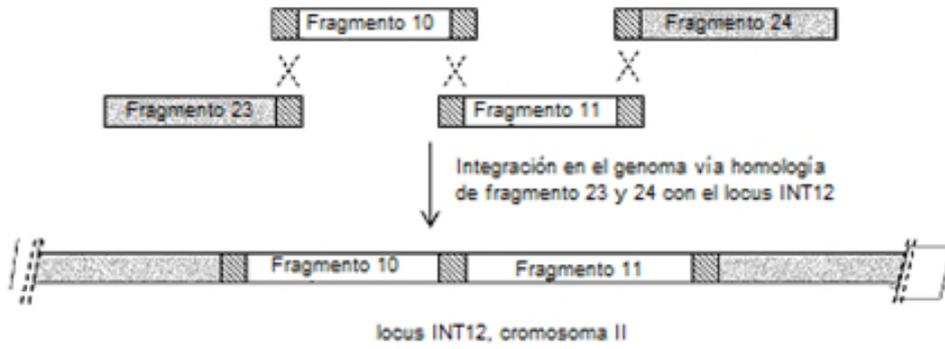


Fig. 4

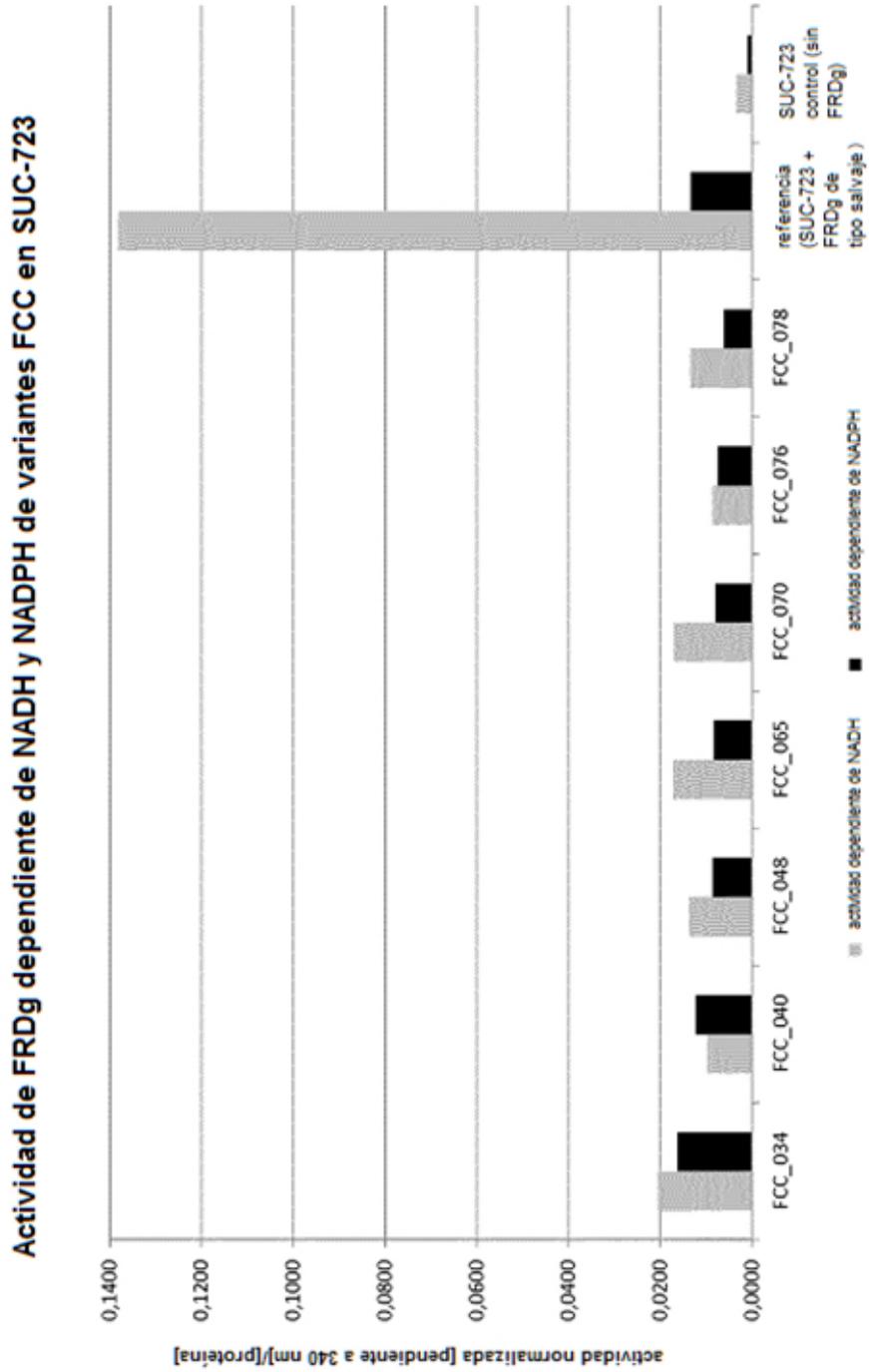


Fig. 5a

Actividad de FRDg dependiente de NADH y NADPH de variantes FCC en SUC-723

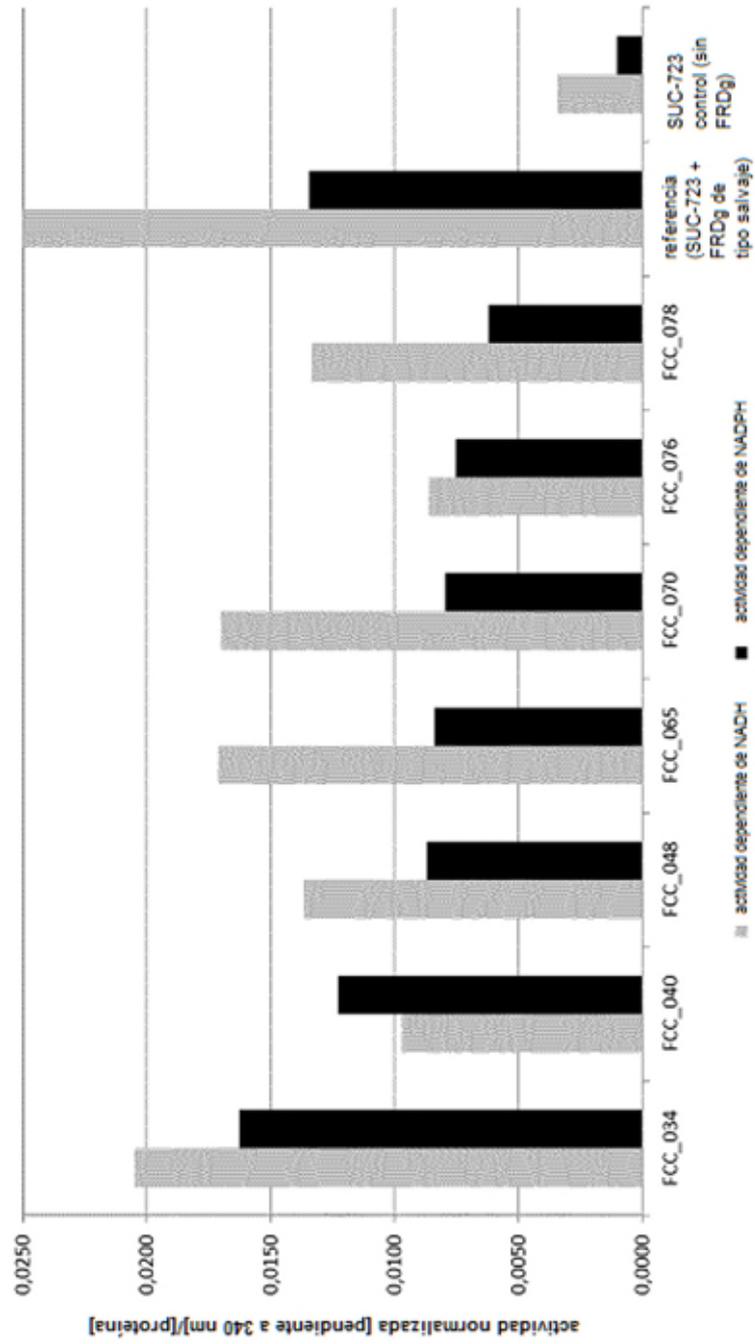


Fig. 5b

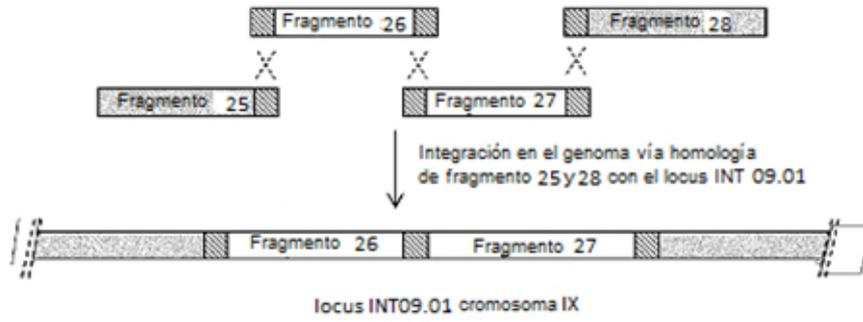


Fig. 6

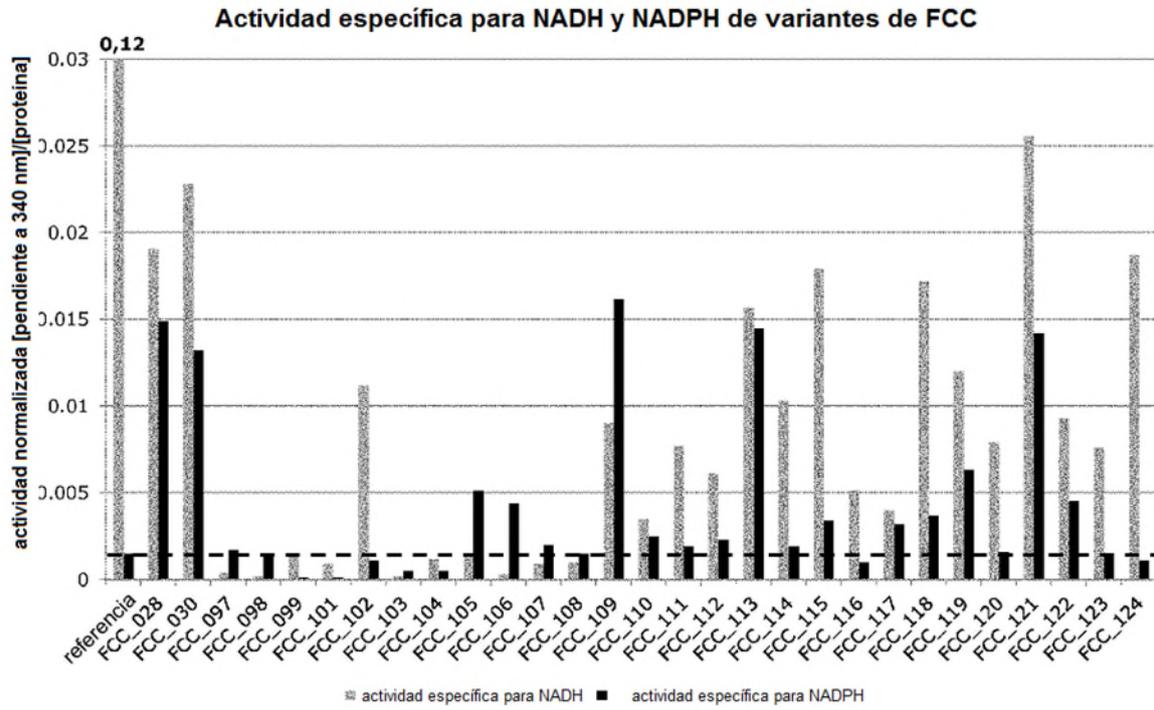


Fig. 7