

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 640 947**

51 Int. Cl.:

C12N 15/65 (2006.01)

C12P 21/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.05.2010 PCT/US2010/034201**

87 Fecha y número de publicación internacional: **18.11.2010 WO10132341**

96 Fecha de presentación y número de la solicitud europea: **10.05.2010 E 10775323 (8)**

97 Fecha y número de publicación de la concesión europea: **19.07.2017 EP 2430164**

54 Título: **Producción de proteínas recombinantes utilizando métodos de selección no antibióticos y la incorporación de aminoácidos no naturales en las mismas**

30 Prioridad:

11.05.2009 US 177267 P

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

07.11.2017

73 Titular/es:

**PFENEX, INC. (100.0%)
10790 Roselle St.
San Diego, CA 92121, US**

72 Inventor/es:

**RESTALLACK, DIANE, M.;
CHEW, LAWRENCE, C. y
SQUIRES, CHARLES, H.**

74 Agente/Representante:

ISERN JARA, Jorge

Observaciones :

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 640 947 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

Producción de proteínas recombinantes utilizando métodos de selección no antibióticos y la incorporación de aminoácidos no naturales en las mismas

Referencia cruzada a solicitudes relacionadas

La presente solicitud está relacionada con la Solicitud Provisional de los EE.UU. número 61/177.267 presentada el 11 de mayo de 2009 para la que se reivindica prioridad a tenor de 35 USC 119.

Campo de la invención

La presente invención proporciona un sistema mejorado de expresión para la producción de polipéptidos recombinantes utilizando marcadores seleccionables auxotróficos y para la incorporación de aminoácidos no naturales o compuestos de bases nitrogenadas en el polipéptido recombinante o la secuencia codificante del mismo.

Antecedentes de la invención

El uso de células bacterianas para producir agentes terapéuticos basados en proteínas está aumentando en importancia comercial. Uno de los objetivos del desarrollo de un sistema de expresión bacteriana es la producción de polipéptidos diana de alta calidad de manera rápida, eficiente y abundante. Una célula hospedadora ideal para un sistema de expresión de este tipo sería capaz de utilizar eficientemente una fuente de carbono para la producción de un polipéptido diana, de crecer rápidamente en altas densidades de células en una reacción de fermentación, de expresar el polipéptido diana solamente cuando se induce y de crecer en un medio que no produce preocupaciones normativas y ambientales.

Existen muchos obstáculos para la creación de una célula hospedadora superior. En primer lugar, con el fin de producir una polipéptido recombinante, debe insertarse un vector de expresión que codifique la proteína diana en la célula hospedadora. Muchas bacterias son capaces de revertir de nuevo a un estado no transformado, en el que el vector de expresión es eliminado del hospedador. Dichos revertientes pueden disminuir la eficiencia de la fermentación de la producción del polipéptido recombinante deseado.

Los vectores de expresión que codifican un péptido diana incluyen normalmente un marcador de selección en el vector. Con frecuencia, el marcador de selección es un gen cuyo producto es necesario para la supervivencia durante el proceso de fermentación. Las células que carecen del marcador de selección, tales como los revertientes, son incapaces de sobrevivir. El uso de marcadores de selección durante el proceso de fermentación tiene por objeto asegurar que solo las bacterias que contengan el vector de expresión sobrevivan, eliminar la competencia entre los revertientes y los transformantes y reducir la eficiencia de la fermentación.

Los marcadores de selección utilizados más habitualmente son los genes de resistencia a antibióticos. Las células hospedadoras se cultivan en un medio complementado con un antibiótico capaz de ser degradado por el producto del gen de resistencia a antibiótico seleccionado. Las células que no contienen el vector de expresión con el gen de resistencia a antibiótico son destruidas por el antibiótico. Los genes de resistencia a antibióticos habituales incluyen tetraciclina, neomicina, kanamicina y ampicilina. La presencia de genes de resistencia a antibióticos en una célula hospedadora bacteriana, sin embargo, presenta problemas ambientales, normativos y comerciales. Por ejemplo, los productos que contienen un gen de resistencia a antibiótico (y los productos producidos por el uso de un gen de resistencia a antibiótico) se han identificado como riesgos de bioseguridad potenciales para la salud medioambiental, humana y animal. Por ejemplo, véase M. Droge et al., *Horizontal Gene Transfer as a Biosafety issue: A natural phenomenon of public concern*, *J. Biotechnology*. 64(1): 75-90 (17 de septiembre de 1998); Gallagher, D. M. y D. P. Sinn. 1983. *Penicillin-induced anaphylaxis in a patient under hypotensive anaesthesia*. *Oral Surg. Oral Med. Oral Pathol.* 56:361-364; Jorro, G., C. Morales, J. V. Braso y A. Pelaez. 1996. *Anaphylaxis to erythromycin*. *Ann. Allergy Asthma Immunol.* 77:456-458; F. Gebhard y K. Smalla, *Transformation of Acinetobacter sp. strain BD413 by transgenic sugar beet DNA*, *Appl. & Environ. Microbiol.* 64(4):1550-54 (abril de 1998); T. Hoffmann et al., *Foreign DNA sequences are received by a wild type strain of Aspergillus niger after co-culture with transgenic higher plants*, *Curr. Genet.* 27(1): 70-76 (diciembre de 1994); DK Mercer et al., *Fate of free DNA and transformation of the oral bacterium Streptococcus gordonii DL1 by plasmid DNA in human saliva*, *Appl. & Environ. Microbiol.* 65(1):6-10 (enero de 1999); R. Schubert et al., *Foreign (M13) DNA ingested by mice reaches peripheral leukocytes, spleen, and liver via the intestinal wall mucosa and can be covalently linked to mouse DNA*, *PNAS USA* 94:961-66 (4 de febrero de 1997); y AA Salyers, *Gene transfer in the mammalian intestinal tract*, *Curr. Opin. in Biotechnol.* 4(3):294-98 (junio de 1993).

Como resultado de estas preocupaciones, muchas agencias reguladoras gubernamentales de fármacos, salud y ambientales, así como muchos usuarios finales, requieren que el ácido nucleico del gen de resistencia a antibiótico se retire de los productos o esté ausente de los organismos para su uso en el comercio. Además, deben proporcionarse pruebas que demuestren la eliminación de los antibióticos de selección del producto final con el fin de asegurar la eliminación normativa. El Reino Unido, Canadá, Francia, la Comunidad Europea y los Estados Unidos

han abordado el uso de genes de resistencia a antibióticos en los alimentos, piensos, fármacos y en la producción de fármacos, incluyendo la producción de fármacos recombinantes. La eliminación de estos agentes, y especialmente la demostración de dicha eliminación, es costosa, lleva mucho tiempo y con frecuencia solo mínimamente eficaz.

5 Debido a las preocupaciones inherentes al uso de genes de resistencia a antibióticos para la selección en la producción de polipéptidos recombinantes, se necesitan métodos de selección alternativos.

Sumario de la invención

10 Se ha descubierto que la producción de proteínas recombinantes puede mejorarse mediante la selección como célula hospedadora de un organismo que sea susceptible de selección auxotrófica, no resistente a antibióticos, y la utilización de un análogo de un metabolito para el que la célula hospedadora sea auxotrófica para la expresión de la proteína recombinante. Adicionalmente, los presentes inventores han descubierto que después de la selección de las células hospedadoras adecuadas mediante el procedimiento novedoso "no antibiótico", estas células pueden utilizarse para producir proteínas recombinantes predeterminadas que puedan incorporar aminoácidos no naturales en la proteína expresada. La presente invención se define por las reivindicaciones adjuntas.

20 De acuerdo con un aspecto de la presente invención, se proporciona un método para producir un polipéptido recombinante de interés. El proceso incluye la obtención de una población de células auxotróficas para un primer metabolito y un segundo metabolito. De acuerdo con la invención, el segundo metabolito es un aminoácido natural.

25 Además, el método incluye poner en contacto la población de células con una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica, en el que el marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito y en el que la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito. La población de células se pone en contacto con una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés; y un promotor unido operativamente a la segunda secuencia de ácido nucleico para dirigir la expresión de la segunda secuencia de ácido nucleico. La población de células se somete a un primer medio que carece del primer metabolito en condiciones de manera que se obtengan células transformadas que tengan la prototrofia restablecida para el primer metabolito. Las células transformadas se someten a un segundo medio que comprende un aminoácido no natural correlacionado con el segundo metabolito en condiciones de manera que la segunda secuencia de ácido nucleico se exprese para producir el polipéptido recombinante de interés que tiene el aminoácido no natural incorporado en el mismo. En el método, el primer medio y el segundo medio pueden ser iguales o diferentes. En otras palabras, el primer medio puede contener el aminoácido no natural y, en dicho caso, puede servir como el segundo medio. Como alternativa, el primer medio carece del aminoácido no natural. Una vez que se obtienen las células transformadas, estas células se someten a un segundo medio que contiene un aminoácido no natural.

40 La presente memoria descriptiva describe también un método para producir un polipéptido recombinante de interés. El método comprende: introducir en una célula hospedadora que es auxotrófica para un primer metabolito necesario para la supervivencia de la célula hospedadora, una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica. El marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito y la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito a la célula hospedadora auxotrófica. El método incluye adicionalmente introducir en la célula hospedadora: (i) una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés; (ii) una tercera secuencia de ácido nucleico que codifica una ARNt sintetasa ortogonal; (iii) una cuarta secuencia de ácido nucleico que codifica una ARNt ortogonal que puede interactuar con dicha ARNt sintetasa ortogonal; y (iv) un promotor unido operativamente a la segunda, la tercera y/o la cuarta secuencia de ácidos nucleicos con el fin de dirigir la expresión de la segunda secuencia de ácido nucleico, la tercera secuencia de ácido nucleico y/o la cuarta secuencia de ácido nucleico en la célula hospedadora auxotrófica. La célula hospedadora auxotrófica se somete a un medio que carece del primer metabolito para obtener células transformadas. Las células transformadas se someten a un medio que contiene un aminoácido no natural en condiciones de manera que el segundo ácido nucleico se exprese para producir el polipéptido recombinante que tiene el aminoácido no natural incorporado en el mismo.

60 La presente memoria describe adicionalmente un método para producir un polipéptido recombinante de interés. El método comprende introducir en una célula hospedadora que es auxotrófica para un primer metabolito, una construcción de ácido nucleico. La construcción de ácido nucleico comprende (i) un marcador de selección auxotrófica, en el que el marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito y en el que la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito a la célula hospedadora auxotrófica; (ii) un segundo ácido nucleico que codifica el polipéptido recombinante de interés; (iii) una tercera secuencia de ácido nucleico que codifica una ARNt sintetasa ortogonal; (iv) una cuarta secuencia de ácido nucleico que codifica un ARNt que puede interactuar con la ARNt sintetasa ortogonal; y (v) un promotor unido operativamente a la primera, la

segunda, la tercera y/o la cuarta secuencia de ácido nucleico con el fin de dirigir la expresión de al menos una de las mismas. Además, el método comprende someter la célula hospedadora auxotrófica a un medio que carece del primer metabolito con lo que se obtienen células transformadas con la construcción de ácido nucleico. Las células transformadas se someten a un medio que contiene un aminoácido no natural en condiciones de manera que el segundo ácido nucleico se exprese para producir el polipéptido recombinante que tiene el aminoácido no natural incorporado en el mismo.

Breve descripción de las figuras

La Figura 1 muestra las curvas de crecimiento para las cepas de expresión examinadas en este estudio. Se usaron tres matraces de agitación para cada cepa. Un matraz de cada conjunto recibió metionina como el aditivo para servir como control. Los otros dos matraces del conjunto recibieron AHA para el análisis de la incorporación del análogo de metionina en la proteína IFN-beta expresada. El tiempo de fermentación transcurrido (horas) se indica en el eje X y la densidad óptica a 600 nm en el eje Y. A las 24 horas de incubación, se recogieron las células y se les privaron de nutrientes en un medio sin metionina añadida durante 0,5 horas. Las células se resuspendieron después en un medio que contenía o bien metionina o bien AHA y se indujo la expresión de IFN-beta.

La Figura 2 muestra análisis por SDS-PAGE (A) y transferencia Western (B) de la proteína soluble (S) e insoluble (I) expresada en la cepa auxotrófica para metionina PD-17 (DC485) que llevaba pIFN-001. Las muestras recogidas en I0, I6, I12 y I24 se normalizaron a DO600 = 20. Se cargaron 15 µl de muestras normalizadas diluidas 1:2 en cada gel. Se cargó proteína patrón [400 ng (A) y 50 ng (B)] como referencia. Las muestras de un matraz complementado con metionina se muestran a la izquierda, mientras que las muestras de los matraces complementados con AHA se muestran en el centro y a la derecha.

Descripción detallada de la invención

Visión de conjunto

En el presente documento se proporcionan métodos para la expresión eficiente de polipéptidos recombinantes en un sistema de expresión. Los métodos comprenden el uso de marcadores de selección auxotrófica en lugar de marcadores de selección con antibióticos para el mantenimiento de plásmidos y la selección de transformantes. A partir de entonces, las células seleccionadas, que también puede incluir una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés, pueden expresar el polipéptido recombinante de interés con un aminoácido no natural incorporado en el mismo.

El sistema de selección auxotrófica comprende una célula hospedadora auxotrófica que se transforma con una o más construcciones de expresión que codifican un polipéptido de interés y un polipéptido capaz de restablecer la prototrofia a la célula hospedadora auxotrófica. La "célula hospedadora auxotrófica" es cualquier célula hospedadora que es deficiente en uno o más metabolitos necesarios para la supervivencia de la célula hospedadora. En una realización, la auxotrofia es el resultado de modificaciones genéticas de al menos un gen de biosíntesis de un compuesto de bases nitrogenadas o al menos un gen de biosíntesis de aminoácidos. Un "polipéptido capaz de restablecer la prototrofia" es cualquier polipéptido activo en la biosíntesis del metabolito que es necesario para la supervivencia de dicha célula hospedadora. De este modo, las células hospedadoras auxotróficas se vuelven prototróficas mediante la expresión de la construcción heteróloga que comprende el polipéptido de restablecimiento de la prototrofia.

La presente invención comprende adicionalmente la incorporación de aminoácidos o análogos de compuestos de bases nitrogenadas en el polipéptido recombinante codificado de interés o una secuencia codificante que codifica el polipéptido de interés, respectivamente. Los métodos comprenden la adición de uno o más análogos de un metabolito para el que las células hospedadoras auxotróficas son deficientes en el medio de crecimiento para la célula hospedadora de una manera en la que el análogo se incorpora en la proteína recombinante o el compuesto de bases nitrogenadas que codifica la proteína.

"Incorporación", como se usa en el presente documento se refiere a cualquier adición, sustitución, reemplazo, mutación u otra modificación en la que se introducen uno o más aminoácidos análogos o compuestos de bases nitrogenadas en la molécula diana además de o como sustituto de un aminoácido de origen natural o compuesto de bases nitrogenadas.

En diversas realizaciones, el uso del análogo de metabolito en presencia de un sistema de selección auxotrófica da como resultado una expresión de proteínas mejorada con respecto a la expresión en un sistema de selección con antibióticos, o con respecto a la expresión en un sistema de selección auxotrófica o con antibióticos en ausencia de un análogo.

La ingeniería de proteínas por medio de la introducción de aminoácidos no naturales es un enfoque importante para la investigación del plegamiento, la estructura y la función de las proteínas, así como el diseño de una reactividad de proteínas novedosa (Dougherty (2000) *Curr Opin Chem Biol* 4:645-652). Los inventores se han dado cuenta de que

las técnicas utilizadas actualmente para la expresión de proteínas recombinantes en presencia de aminoácidos no naturales pueden dar como resultado un menor rendimiento global de la proteína recombinante en comparación con la expresión en ausencia del aminoácido no natural. Aunque sin quedar ligados a ninguna teoría o mecanismo particular, los inventores suponen que el rendimiento reducido puede estar relacionado con el uso de un marcador de selección con antibióticos en presencia del aminoácido no natural. Después de haberse dado cuenta de un problema inherente a las técnicas actuales, los inventores han determinado que sería deseable desarrollar técnicas mejoradas que sean más fáciles para controlar la selección de las células transformadas, controlar la incorporación de análogos y que puedan lograr un mayor rendimiento de los polipéptidos de interés. En la presente invención, los inventores han desarrollado sistemas, materiales y métodos para producir polipéptidos recombinantes que utilizan marcadores de selección auxotrófica, en lugar de marcadores de selección con antibióticos, para el mantenimiento del plásmido con el fin de mejorar el rendimiento de proteínas recombinantes en presencia de aminoácidos no naturales.

En una primera realización, la presente invención comprende obtener una población de células auxotróficas para un primer metabolito y un segundo metabolito. En el presente documento se proporciona una lista no limitante de metabolitos. En una realización, el segundo metabolito es un aminoácido natural.

La población de células se pone en contacto con una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica. El marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito. La expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito.

En algunas realizaciones, el sistema de traducción comprende adicionalmente un medio de cultivo que contiene uno o más aminoácidos no naturales. En otras realizaciones más, dichos uno o más aminoácidos no naturales se seleccionan entre el grupo que consiste en: azidonorleucina, 3-(1-naftil)alanina, 3-(2-naftil)alanina, p-etinil-fenilalanina, p-propargil-oxi-fenilalanina, m-etinil-fenilalanina, 6-etinil-triptofano, 5-etinil-triptofano, ácido (R)-2-amino-3-(4-etinil-1H-pirrol-3-il)propánico, p-bromofenilalanina, p-idiofenilalanina, p-azidofenilalanina, 3-(6-cloroindolil)alanina, 3-(6-bromoindolil)alanina, 3-(5-bromoindolil)alanina, azidohomoalanina y p-clorofenilalanina. En otras realizaciones más, dicho AARS modificado se selecciona entre el grupo que consiste en: un PheRS modificado, un TrpRS modificado, un TyrRS modificado y un MetRS modificado.

Además, la población de células se pone en contacto con una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés. La segunda secuencia de ácido nucleico también comprende un promotor unido operativamente a la segunda secuencia de ácido nucleico para dirigir la expresión de la segunda secuencia de ácido nucleico. Normalmente, pero no necesariamente, la expresión unido operativamente, cuando se describe un promotor, se refiere a que está dispuesto sobre la construcción de manera que está aguas arriba hacia el extremo 5' de la construcción con respecto a la secuencia de ácido nucleico.

Para seleccionar células transformadas de un medio, el método comprende someter la población de células a un primer medio que carece del primer metabolito en condiciones de manera que se obtengan células transformadas que tengan la prototrofia restablecida para el primer metabolito. De esta manera, cualquier célula que no haya tenido la prototrofia restablecida para el primer metabolito no sobrevive en el medio.

Una vez que se obtienen las células que tienen la prototrofia restablecida para el primer metabolito, la población de células se somete a un segundo medio que comprende un aminoácido no natural correlacionado con el segundo metabolito en condiciones de manera que la segunda secuencia de ácido nucleico se exprese para producir el polipéptido recombinante de interés que tiene el aminoácido no natural incorporado en el mismo. Las expresiones "correlacionado con" o "que se correlaciona" como se usan en el presente documento con respecto a la descripción de un aminoácido no natural significan que el aminoácido no natural se incorpora en una secuencia de polipéptido en un codón relacionado con el segundo metabolito, que es un aminoácido natural. Por ejemplo, el aminoácido no natural (segundo metabolito) que puede unirse a un ARNt correspondiente a un aminoácido natural se correlaciona con el aminoácido natural.

En el método, el primer medio y el segundo medio pueden ser iguales o diferentes. En una realización, el primer medio y el segundo medio son iguales. Por "iguales", se entiende que ambos medios contienen el aminoácido no natural deseado. En una realización, el primer medio y el segundo medio son diferentes. Por "diferentes", se entiende que uno de los medios contiene el aminoácido no natural deseado y el otro medio no.

En otra realización, hay un método para producir un polipéptido recombinante de interés. El método comprende introducir en una célula hospedadora que es auxotrófica para un primer metabolito necesario para la supervivencia de la célula hospedadora, una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica. El marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito. La expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito a la célula hospedadora auxotrófica.

El método incluye también la introducción de una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés en la célula hospedadora.

Adicionalmente, una tercera secuencia de ácido nucleico que codifica una ARNt sintetasa ortogonal se introduce en la célula hospedadora, junto con un promotor unido operativamente a las secuencias de ácido nucleico segunda y tercera con el fin de dirigir la expresión de la segunda secuencia de ácido nucleico y la tercera secuencia de ácido nucleico en la célula hospedadora auxotrófica. Además, la segunda construcción de ácido nucleico comprende una cuarta secuencia de ácido nucleico que codifica un ARNt ortogonal que puede interactuar con la ARNt sintetasa ortogonal.

La célula hospedadora auxotrófica se somete a un medio que carece del primer metabolito permitiendo de este modo la selección de células transformadas. Las células transformadas se cultivan en condiciones de manera que las secuencias de ácido nucleico tercera y cuarta se expresen para producir una ARNt sintetasa ortogonal y ARNt ortogonal, respectivamente. Las células transformadas se someten a un medio que comprende un aminoácido no natural que puede interactuar con el ARNt ortogonal. La ARNt sintetasa ortogonal y el ARNt ortogonal interactúan para facilitar la incorporación del aminoácido no natural que puede interactuar en el polipéptido recombinante de interés durante la expresión para producir el polipéptido recombinante que tiene el aminoácido no natural incorporado en el mismo.

De acuerdo con la invención, las secuencias de ácido nucleico primera y segunda descritas en los párrafos anteriores se proporcionan en la misma construcción.

Definiciones

Como se usa en el presente documento, la expresión "porcentaje de proteína celular total" se refiere a la cantidad de proteína o péptido en la célula hospedadora como porcentaje de la proteína celular general.

La expresión "unido operativamente", como se usa en el presente documento, se refiere a cualquier configuración en la que los elementos reguladores transcripcionales y cualesquiera elementos reguladores traduccionales se unen covalentemente a la secuencia codificante en dicha disposición o disposiciones, con respecto a la secuencia codificante, de manera que en y por acción de la célula hospedadora, los elementos reguladores puedan dirigir la expresión de la secuencia codificante.

El término "auxotrófico", como se usa en el presente documento, se refiere a una célula que ha sido modificada para eliminar o reducir su capacidad de producir una sustancia específica necesaria para el crecimiento y el metabolismo.

El término "prototrofia", como se usa en el presente documento, se refiere a una célula que es capaz de producir una sustancia específica necesaria para el crecimiento y el metabolismo.

La expresión "aminoácido no natural" como se usa en el presente documento, se refiere a un aminoácido que es diferente de los veinte aminoácidos de origen natural (alanina, arginina, glicina, asparragina, ácido aspártico, cisteína, glutamina, ácido glutámico, serina, treonina, histidina, lisina, metionina, prolina, valina, isoleucina, leucina, tirosina, triptófano, fenilalanina) en su funcionalidad de cadena lateral. Dichas funcionalidades de cadena lateral diferentes pueden incluir, pero no se limitan a, halógenos, hidrocarburos insaturados, heterociclos, silicio, unidades organometálicas. Estas cadenas laterales adicionales pueden mejorar la estabilidad de la estructura plegada de las proteínas sin necesidad de modificaciones de la secuencia. Una lista no limitante de aminoácidos no naturales que pueden usarse de acuerdo con el contenido del presente documento incluye, pero no se limita a, azidonorleucina, 3-(1-naftil)alanina, 3-(2-naftil)alanina, p-etinil-fenilalanina, p-propargil-oxi-fenilalanina, m-etinil-fenilalanina, 6-etinil-triptofano, 5-etinil-triptofano, ácido (R)-2-amino-3-(4-etinil-1H-pirrol-3-il)propánico, p-bromofenilalanina, p-idiofenilalanina, p-azidofenilalanina, 3-(6-cloroindolil)alanina, 3-(6-bromoindolil)alanina, 3-(5-bromoindolil)alanina, azidohomoalanina y p-clorofenilalanina. En otras realizaciones más, dicho AARS modificado se selecciona entre el grupo que consiste en: un PheRS modificado, un TrpRS modificado, un TyrRS modificado y un MetRS modificado.

Los términos "introduce", "introducen", "introducido" o "que introduce", como se usan en el presente documento en relación con secuencias de ácido nucleico o construcciones que comprenden las mismas, se refieren al contacto y la captación de una secuencia exógena y/o construcción en una célula. "Introducción" se refiere al acto de introducir.

"Transformación" es un tipo de introducción y normalmente se refiere a la captación y la replicación estable de un plásmido en una célula, o a la captación de una secuencia o construcción de ácido nucleico lineal en una célula con lo que la secuencia o construcción exógena se inserta de manera estable en el genoma de la célula transformada.

La expresión "construcción de ácido nucleico" como se usa en el presente documento se refiere a un polinucleótido que tiene dos o más elementos de secuencia de ácido nucleico que tienen propósitos o funciones separados. En ciertas realizaciones, una construcción de ácido nucleico puede proporcionarse en forma de un plásmido u otro vector adecuado para la introducción a una célula.

Sistemas de selección auxotrófica

Se han utilizado marcadores de selección auxotrófica como alternativa a la selección con antibióticos en algunos sistemas. Por ejemplo, se han utilizado marcadores auxotróficos ampliamente en la levadura, en gran parte debido a la ineficiencia de los marcadores de selección de resistencia a antibióticos en estas células hospedadoras. Véase, por ejemplo, JT Pronk, (2002) "*Auxotrophic yeast strains in fundamental and applied research*", *App. & Environ. Micro.* 68(5): 2095-2100; Boeke et al., (1984) "*A positive selection for mutants lacking orotidine-5'-phosphate decarboxylase activity in yeast; 5-fluoro-orotic acid resistance*", *Mol. Gen. Genet.* 197: 345-346; Botstein y Davis, (1982) "*Principles and practice of recombinant DNA research with yeast*", págs. 607-636, en J N Strathern, E W Jones y JR Broach (ed.), *The molecular biology of the yeast Saccharomyces cerevisiae, Metabolism and gene expression*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Cost y Boeke, (1996) "*A useful colony color phenotype associated with the yeast selectable/counter selectable marker MET15*", *Yeast* 12: 939-941.

También se ha descrito anteriormente la selección de marcadores auxotróficos en bacterias. Véase, por ejemplo, las Patentes de los EE.UU. N.º 4.920.048, 5.691.185, 6.291.245, 6.413.768 y 6.752.994; la Publicación de Patente de los EE.UU. N.º 20050186666; Struhl et al. (1976) *PNAS USA* 73; 1471-1475; MacCormick et al., (1995) *FEMS Microbiol. Lett.* 127:105-109; Dickely et al. (1995) *Mol. Microbiol.* 15:839-847; Sorensen et al. (2000) *Appl. Environ. Microbiol* 66:1253-1258; y Fiedler y Skerra (2001), *Gene* 274:111-118.

En un aspecto de la presente invención, una población de células hospedadoras se obtiene o se proporciona de otro modo, que es auxotrófica para al menos un metabolito. Se contempla que las células hospedadoras auxotróficas puedan proporcionarse de una fuente comercial adecuada o, en una realización, puedan haber sido modificadas genéticamente para inducir la auxotrofia para al menos un metabolito. La modificación genética puede ser la de un gen o genes que codifican una enzima que es operativa en una vía metabólica, tal como una vía biosintética anabólica o una vía de utilización catabólica. Preferentemente, la célula hospedadora tiene todos los genes operativos que codifican una actividad biocatalítica dada eliminados o inactivados con el fin de asegurar la eliminación de la actividad biocatalítica.

Puede seleccionarse una o más de una actividad metabólica para la supresión o reemplazo. En el caso de la auxotrofia o las auxotrofias nativas, pueden proporcionarse supresiones o reemplazos metabólicos adicionales. Cuando se seleccionan múltiples actividades, los marcadores de selección por restablecimiento de la auxotrofia pueden ser de un tipo biosintético (anabólico), de un tipo de utilización (catabólico) o pueden elegirse entre ambos tipos. Por ejemplo, puede suprimirse una o más de una actividad en una vía biosintética propuesta para el compuesto seleccionado; o puede suprimirse más de una actividad, cada una de diferentes vías biosintéticas. La actividad o actividades correspondientes entonces se proporcionan por al menos un vector recombinante que, tras la transformación en la célula, restablece la prototrofia a la célula.

Los compuestos y moléculas cuya biosíntesis o utilización pueden dirigirse para producir células hospedadoras auxotróficas incluyen: lípidos, incluyendo, por ejemplo, ácidos grasos; mono y disacáridos y derivados sustituidos de los mismos, incluyendo, por ejemplo, glucosa, fructosa, sacarosa, glucosa-6-fosfato y ácido glucurónico, así como intermedios y productos de la vía de Entner-Doudoroff y de las pentosas fosfato; nucleósidos, nucleótidos, dinucleótidos, incluyendo, por ejemplo, ATP, dCTP, FMN, FAD, NAD, NADP, bases nitrogenadas, incluyendo, por ejemplo, piridinas, purinas, pirimidinas, pterinas e hidro-, deshidro-derivados de bases nitrogenadas y/o derivados de bases nitrogenadas sustituidos, tales como cofactores, por ejemplo, biotina, cobamamida, riboflavina, tiamina; ácidos orgánicos e intermedios y productos de la glicólisis y del ciclo del ácido cítrico, incluyendo, por ejemplo, hidroxiaácidos y aminoácidos; hidratos de carbono de almacenamiento y polímeros de poli(hidroxialcanoato) de almacenamiento, incluyendo, por ejemplo, celulosa, almidón, amilosa, amilopectina, glicógeno, poli-hidroxibutirato y pollactato.

En una realización, la actividad o actividades biocatalíticas suprimidas para producir la célula hospedadora auxotrófica se selecciona entre el grupo que consiste en: los lípidos; los nucleósidos, nucleótidos, dinucleótidos, bases nitrogenadas y derivados de bases nitrogenadas; y los ácidos orgánicos e intermedios y productos de la glucólisis y del ciclo del ácido cítrico. Preferentemente, la actividad o actividades biocatalíticas suprimidas se seleccionan entre el grupo que consiste en: los nucleósidos, nucleótidos, dinucleótidos, bases nitrogenadas y derivados de bases nitrogenadas; y los ácidos orgánicos e intermedios y productos de la glucólisis y del ciclo del ácido cítrico. Más preferentemente, la actividad o actividades biocatalíticas suprimidas se seleccionan entre el grupo que consiste en: los nucleósidos de pirimidina, nucleótidos, dinucleótidos, bases nitrogenadas y derivados de bases nitrogenadas; y los aminoácidos.

Una célula hospedadora transgénica dada puede usar uno o más de un marcador de selección o sistema marcador de selección. Por ejemplo, uno o más marcadores de selección de biosíntesis o sistemas marcadores de selección de acuerdo con la presente invención pueden usarse junto entre sí, y/o puede usarse en combinación con un marcador de selección de tipo de utilización o sistema marcador de selección de acuerdo con la presente invención.

En cualquiera de estas realizaciones que permiten la prototrofia, la célula hospedadora puede contener también uno o más marcadores de selección no auxotrófica o sistemas marcadores de selección. Los ejemplos de marcadores y sistemas de selección no auxotrófica incluyen, por ejemplo: genes marcadores de resistencia a toxinas tales como

genes de resistencia a antibióticos que codifican una actividad enzimática que degrada un antibiótico; genes marcadores de resistencia a toxinas, tales como, por ejemplo, mutantes resistentes a la imidazolinona de la acetolactato sintasa ("ALS"; CE 2.2.1.6) en los que se expresan una mutación o mutaciones que hacen a la enzima insensible a la inhibición por toxinas que presentan las versiones de la enzima que no contienen dicha mutación o mutaciones. El compuesto o compuestos pueden ejercer este efecto directamente; o el compuesto o compuestos pueden ejercer este efecto indirectamente, por ejemplo, como resultado de la acción metabólica de la célula que convierte el compuesto o compuestos en forma de toxina o como resultado de la combinación del compuesto o compuestos con al menos un compuesto o compuestos adicionales.

Pueden obtenerse genes operativos para células hospedadoras que codifican dichas enzimas marcadoras de la cepa de la célula hospedadora elegida para la construcción de la célula suprimida, de cepas relacionadas o de otros organismos, y pueden usarse en forma nativa o forma modificada (por ejemplo, mutada o de secuencia recombinada). Por ejemplo, puede obtenerse una secuencia de ADN que codifica una enzima que presenta la actividad biocatalítica suprimida de uno o más organismos y después unirse operativamente a elementos reguladores del ADN operativos dentro de la célula hospedadora. En concreto, todos los genes intracelulares de la célula hospedadora elegida que codifican una actividad enzimática seleccionada se suprimen; la célula hospedadora suprimida después se transforma con un vector que contiene al menos una copia operativa de un gen nativo o no nativo que codifica una enzima que presenta la actividad perdida mediante la supresión.

Los genes que codifican dichas enzimas pueden seleccionarse y obtenerse a través de diversos recursos disponibles para un experto habitual en la materia. Estos incluyen las secuencias de nucleótidos de secuencias que codifican enzimas y elementos reguladores del ADN operativos en la especie. Los recursos en línea de Internet incluyen, por ejemplo: (1) la instalación de proteómica ExpASy (véanse las características de ENZYME and BIOCHEMICAL PATHWAYS MAPS) del Instituto Suizo de Bioinformática (Bâtiment Ecole de Pharmacie, Aula 3041; Universidad de Lausana; 1015 Lausana-Dorigny; Suiza) disponible en, por ejemplo, us.expasy.org/; y (2) la instalación de GenBank y otros recursos de Entrez (véase las características de PUBMED, PROTEIN, NUCLEOTIDE, STRUCTURE, GENOME, et al.) ofrecido por el Centro Estadounidense de Información Biotecnológica (NCBI por sus siglas en inglés, Biblioteca Nacional de Medicina, Institutos Estadounidenses de Salud, Departamento de Salud y Servicios Sociales de los EE.UU.; Edificio 38A, Bethesda, Maryland, EE.UU.) y disponible en www.ncbi.nlm.nih.gov/entrez/query.fcgi.

La secuencia codificante seleccionada puede modificarse alterando el código genético de la misma para que coincida con la empleada por la célula hospedadora utilizada en el sistema, y la secuencia de codón de la misma puede potenciarse para que se aproxime mejor a la empleada por el hospedador. La selección del código genético y la potenciación de la frecuencia de codones pueden realizarse de acuerdo con cualquiera de los diversos métodos conocidos para un experto habitual en la materia, por ejemplo, mutagénesis dirigida por oligonucleótidos. Los recursos en línea de Internet para ayudar en este proceso incluyen, por ejemplo: (1) la Base de Datos de Uso de Codones del Instituto de Investigación del ADN Kazusa (2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818 Japón) y disponible en www.kazusa.or.jp/codon/; y (2) las tablas de Códigos Genéticos disponibles de la base de datos de Taxonomía del NCBI en www.ncbi.nlm.nih.gov/Taxonomy/Utils/wprintgc.cgi?mode=c. Por ejemplo, se notifica que especies de *Pseudomonas* utilizan la Tabla de Traducción del Código Genético 11 del sitio de Taxonomía del NCBI y en el sitio de Kazusa como que presentan la frecuencia de uso de codones de la tabla que se muestra en www.kazusa.or.jp/codon/cgibin/.

Marcadores de selección de bases nitrogenadas y nucleósidos biosintéticos

En una realización, puede elegirse una enzima biosintética implicada en el metabolismo anabólico como el marcador de selección auxotrófica. En particular, la enzima biosintética puede seleccionarse entre las que implicadas en la biosíntesis de los nucleósidos, nucleótidos, dinucleótidos, bases nitrogenadas y derivados de bases nitrogenadas, que se denominan colectivamente en el presente documento como marcadores de selección de bases.

En una realización particular puede elegirse al menos una enzima biosintética de tipo purina como un marcador de selección auxotrófica. Dichas enzimas biosintéticas de purina incluyen, por ejemplo, adenina fosforribosiltransferasas, adenilosuccinato liasas, adenilosuccinato sintasas, GMP sintasas, IMP ciclohidrolasas, IMP deshidrogenasas, fosforribosilamina-glicina ligasas, fosforribosil-aminoimidazolcarboxamida formiltransferasas, fosforribosilaminoimidazol carboxilasas, fosforribosil aminoimidazolsuccinocarboxamida sintasas, fosforribosil-formilglicinamidina ciclo ligasas, fosforribosil-formilglicinamidina sintasas, fosforribosil-glicinamida formil-transferasas, ribosa-fosfato difosfocinasas y ribosa-5-fosfato-amoniaco ligasas.

En otra realización particular, puede elegirse una enzima biosintética de tipo pirimidina como un marcador de selección auxotrófica. Dicha enzima biosintética de tipo pirimidina incluye enzimas implicadas en la biosíntesis de UMP, tales como carbamato cinasa (EC 2.7.2.2), carbamoil-fosfato sintasa (EC 6.3.5.5), aspartato carbamoiltransferasa (EC 2.1.3.2), dihidroorotasa (EC 3.5.2.3), dihidroorotato deshidrogenasa (EC 1.3.3.1), orotato fosforribosiltransferasa ("OPRT"; EC 2.4.2.10) y orotidina-5'-fosfato descarboxilasa ("ODCasa"; CE 4.1.1.23).

65

Se conocen bien ejemplos de genes que codifican enzimas biosintéticas de tipo pirimidina. En el caso de la síntesis bacteriana de UMP, los ejemplos de genes útiles incluyen: genes *arcC*, que codifican carbamato cinasas; genes *carA* y *carB*, que codifican colectivamente carbamoil-fosfato sintasas; genes *pyrB*, que codifican aspartato carbamoiltransferasas; genes *pyrC*, que codifican dihidroorotasas; genes *pyrD*, que codifican individualmente o colectivamente dihidroorotato deshidrogenasas; genes *pyrE*, que codifican orotato fosforribosiltransferasas; y genes *pyrF*, que codifican orotidina-5'-fosfato descarboxilasas.

En una realización particular, un sistema de expresión de acuerdo con la presente invención utilizará un gen de marcador de selección auxotrófica *pyrF*. Los genes *pyrF* codifican ODCasa, una enzima necesaria para la vía de biosíntesis de nucleótidos de pirimidina bacteriana, mediante la cual la célula realiza la síntesis de novo de nucleótidos de pirimidina apropiados (UTP, CTP), así como desoxinucleótidos de pirimidina (dTTP, dCTP). Los reactivos iniciales de la vía son ATP, una fuente de grupos amino (es decir, iones amonio o L-glutamina) y una fuente de grupos carboxilo (es decir, dióxido de carbono o de iones bicarbonato); el producto final de la vía es dTTP, formándose también dCTP, UTP y CTP en el proceso. Específicamente, la vía de biosíntesis de nucleótidos de pirimidina de novo bacteriana comienza con la formación de fosfato de carbamoilo. El fosfato de carbamoilo se sintetiza ya sea: (a) por acción de la carbamato cinasa (EC 2.7.2.2), codificada por el gen *arcC*; o, más habitualmente, (b) por la acción de la carbamoil-fosfato sintasa (EC 6.3.5.5), que hidroliza glutamina, cuyas subunidades pequeñas y grandes están codificadas por los genes *carA* y *carB*, respectivamente. El fosfato de carbamoilo después se convierte en UDP mediante la siguiente vía de seis etapas: 1) conversión del fosfato de carbamoilo en N-carbamoil-L-aspartato, por la aspartato carbamoiltransferasa (EC 2.1.3.2), codificada por *pyrB*; después, 2) la conversión del mismo en (S)-dihidroorotato, por la dihidroorotasa (EC 3.5.2.3), codificada por *pyrC*; después, 3) la conversión del mismo en orotato, por la dihidroorotato deshidrogenasa (EC 1.3.3.1), codificada por el gen o genes *pyrD*; después, 4) la conversión del mismo en orotidina-5'-monofosfato ("OMP"), por la orotato fosforribosiltransferasa ("OPRT"; EC 2.4.2.10), codificada por *pyrE*; y después 5) la conversión del mismo en uridina-5'-monofosfato ("UMP"), por la orotidina-5'-fosfato descarboxilasa ("ODCasa"; EC 4.1.1.23), codificada por *pyrF*. El UMP después es utilizado por una variedad de vías para la síntesis de nucleótidos de pirimidina (UTP, CTP, dTTP, dCTP), ácidos nucleicos, nucleoproteínas y otros metabolitos celulares.

En las bacterias en las que uno o más de los genes *carA*, *carB* o *pyrB-pyrF* se han vuelto inactivos o perdidos, o mutaron para codificar una enzima no funcional, la célula todavía puede prosperar si se añade uracilo al medio, a condición de que la célula contenga una vía de rescate de uracilo funcional. La mayoría de las bacterias contienen una vía de rescate de uracilo nativa, incluyendo las pseudomonas y especies relacionadas. En una vía de rescate de uracilo, la célula importa y convierte el uracilo exógeno en UMP, para sintetizar los nucleótidos de pirimidina necesarios. En esta, el uracilo se hace reaccionar con 5-fosforribosil-1-pirofosfato para formar UMP, por la acción ya sea de la uracilo fosforribosiltransferasa (EC 2.4.2.9), codificada por el gen *upp*, o por la proteína reguladora del operón de pirimidina bifuncional ("proteína bifuncional *PyrR*"), codificada por *PyrR*. El UMP resultante se convierte después en UDP y después en los subsiguientes nucleótidos de pirimidina, como se ha descrito anteriormente.

En consecuencia, una célula hospedadora *pyrF*(-) puede mantenerse en un medio que contiene uracilo. Después de que una construcción de ADN que contiene el gen *pyrF* se transfecta en la célula *pyrF*(-) y se expresa para formar una enzima ODCasa funcional, el sistema plásmido de *pyrF*(+)-célula hospedadora combinado resultante puede mantenerse en un medio que carece de uracilo.

La secuencia codificante del gen *pyrF* para su uso en una célula hospedadora de interés puede ser proporcionada por cualquier gen que codifique una enzima orotidina-5'-fosfato descarboxilasa ("ODCasa") (u homólogo de la misma), a condición de que la secuencia codificante pueda ser transcrita, traducida y procesada de otro modo por la célula hospedadora seleccionada para formar una ODCasa funcional. La secuencia que codifica *pyrF* puede ser una secuencia nativa o puede ser una secuencia modificada por ingeniería genética que sea resultado de, por ejemplo, la aplicación de una o más técnicas conocidas en la técnica que alteren la secuencia, que combinen la secuencia y/o que generen la secuencia. Antes de su uso como parte de un gen marcador de la selección de *pyrF*, la secuencia codificante seleccionada primero puede mejorarse u optimizarse de acuerdo con el código genético y/o la frecuencia de uso de codones de una célula hospedadora seleccionada. Las secuencias codificantes expresables se unirán operativamente a un promotor de la transcripción capaz de funcionar en la célula hospedadora elegida, así como todos los demás elementos reguladores de la transcripción y la traducción necesarios. Una secuencia codificante nativa para un gen *pyrF* como se ha descrito anteriormente puede obtenerse de una bacteria o de cualquier otro organismo, a condición de que cumpla los requisitos descritos anteriormente.

En una realización, la secuencia codificante de *pyrF* se aísla de la célula hospedadora en la que se pretende usar como marcador de selección. Puede aislarse el gen *pyrF* entero (incluyendo la secuencia codificante y las regiones reguladoras circundantes) de la misma.

En una realización alternativa, un sistema de expresión de acuerdo con la presente invención utilizará gen marcador de selección auxotrófica de *thyA*. Los genes *thyA* codifican la timidilato sintasa (EC 2.1.1.45), una enzima necesaria para la vía de biosíntesis de nucleótidos de pirimidina bacteriana. Puesto que el ADN contiene timina (5-metiluracilo) como una base principal en lugar del uracilo, la síntesis de monofosfato de timidina (dTMP o timidilato) es esencial para proporcionar el dTTP (trifosfato de timidina) necesario para la replicación del ADN junto con dATP, dGTP y

dCTP. La metilación de dUMP por la timidilato sintasa utilizando 5,10-metilentetrahidrofolato como la fuente del grupo metilo genera timidilato. La síntesis de timidilato puede interrumpirse y, por consiguiente, la síntesis de ADN detenerse, mediante la eliminación, la inhibición o la alteración de la timidilato sintasa.

- 5 En las bacterias en las que el gen *thyA* se ha vuelto inactivo o perdido, o se ha mutado para codificar una enzima no funcional, la célula todavía puede prosperar si se añade timidina exógena al medio.

En *Pseudomonas fluorescens*, la adición de un gen *tdk* de *E. coli*, que codifica la timidina cinasa, es necesaria para la supervivencia con timidina exógena. Por tanto, antes de la selección, puede usarse un plásmido que comprenda un gen *tdk* para transformar células hospedadoras de *P. fluorescens thyA(-)*, generando una célula *thyA(-)/ptdk*, permitiendo la supervivencia en un medio que contiene timidina. Como alternativa, puede insertarse en el genoma un gen *tdk* que produce una enzima timidilato sintasa funcional capaz de utilizar timidina exógena en *Pseudomonas fluorescens*, produciendo una célula hospedadora *thyA(-)/tdk(+)*. Después de que se transfecta una construcción de ADN que contiene el gen *thyA* en la célula *thyA(-)/ptdk* y se expresa para formar una enzima timidilato sintasa funcional, el sistema plásmido *thyA(+)*-célula hospedadora puede mantenerse en un medio que carece de timidina.

Marcadores de selección de aminoácidos sintéticos

20 En una realización alternativa, la enzima biosintética implicada en el metabolismo anabólico elegida como el marcador de selección auxotrófica puede seleccionarse entre las implicadas en la biosíntesis de aminoácidos. En realizaciones particulares, las enzimas de aminoácidos biosintéticos se seleccionan entre el grupo que consiste en enzimas activas en la biosíntesis de: la familia del glutamato (Glu; Gln, Pro y Arg); la familia del aspartato (Asp; Asn, Met, Thr, Lys e Ile); la familia de la serina (Ser; Gly y Cys); la familia del piruvato (Ala, Val, Leu e); la familia aromática (Trp, Phe y Tyr); y la Familia de la histidina (His). Los ejemplos de genes y enzimas que participan en estas vías biosintéticas incluyen: los genes *arg*, *gdh*, *gln* y *pro* miembros de la familia del glutamato, incluyendo, por ejemplo, *argA-argH*, *gdhA*, *glnA*, *proA*, *proC*; los genes miembros de la familia del aspartato *asd*, *asn*, *asp*, *dap*, *lys*, *met* y *thr*, incluyendo, por ejemplo, *asnA*, *asnB*, *aspC*, *dapA*, *dapB*, *dapD-dapF*, *lysA*, *lysC*, *metA-metC*, *metE*, *metH*, *metL*, *thrA-thrC*; los genes miembros de la familia de la serina *cys*, *gly* y *ser*, incluyendo, por ejemplo, *cysE*, *cysK*, *glyA*, *Sera-serC*; los genes miembros de la familia aromática *aro*, *phe*, *trp* y *Tyr*, incluyendo, por ejemplo, *aroA-aroH*, *aroK*, *aroL*, *trpAtrpE*, *tyrA* y *tyrB*; y los genes miembros de la familia de la histidina, incluyendo *hisA- hisD*, *hisF-hisH*.

En una realización particular adicional, el marcador de selección auxotrófica puede seleccionarse entre enzimas implicadas en la biosíntesis de los miembros de la familia del glutamato. Los ejemplos de marcadores de selección auxotrófica de la familia del glutamato útiles incluyen los siguientes, enumerados con ejemplos representativos de sus genes codificantes: *arg*, que codifica las N-acetilglutamato sintetas, aminoácido acetiltransferasas; *argB*, que codifica acetilglutamato cinasas; *argC*, que codifica N-acetil-gammaglutamilfosfato reductasas; *argD*, que codifica acetilornitina delta-aminotransferasas; *argE*, que codifica acetilornitina desacetilasas; *argF* y *argI*, que codifican ornitina carbamoiltransferasas; *argG*, que codifica argininosuccinato sintetetas; *argH*, que codifica argininosuccinato liasas; *gdhA*, que codifica glutamato deshidrogenasas; *glnA*, que codifica glutamina sintetetas; *proA*, que codifica gamma-glutamylfosfato reductasas; *proB*, que codifica gamma-glutamato cinasas; y *proC*, que codifica pirrolina-5-carboxilato reductasas.

45 En una realización, un gen marcador de selección de la biosíntesis de aminoácidos puede ser al menos un miembro de la familia de la biosíntesis de prolina, en particular, *proA*, *proB* o *Proc*. En una realización particular, el gen marcador de selección de la biosíntesis de prolina puede comprender un gen *proC*. Los genes *Proc* codifican una enzima que cataliza la etapa final de la vía de biosíntesis de prolina. En las bacterias, la vía de biosíntesis de prolina (es decir, L-prolina) comprende un proceso de tres enzimas, comenzando por ácido L-glutámico. Las etapas de este proceso son: 1) la conversión del ácido L-glutámico en L-glutamyl-5-fosfato, por la glutamato-5-cinasa ("GK;" EC 2.7.2.11), codificada por *proB*; después, 2a) la conversión del mismo en L-glutamato-5-semialdehído, por la glutamato-5-semialdehído deshidrogenasa (EC 1.2.1.41), también conocida como glutamyl-5-fosfato reductasa ("GPR"), codificada por *proA*, seguida de 2b) la ciclación espontánea del mismo para formar 1-pirrolina-5-carboxilato; y después 3) la conversión del mismo en L-prolina, por la Δ 1-pirrolina-5-carboxilato reductasa ("P5CR"; EC 1.5.1.2), codificada por *proC*. En la mayoría de las bacterias, *proC* codifica la subunidad P5CR, siendo la enzima P5CR activa un homo-multímero de la misma.

Marcadores de selección de utilización

60 En una realización, una enzima implicada en la utilización catabólica de los metabolitos puede elegirse como el marcador de selección auxotrófica. En particular, las enzimas pueden seleccionarse entre aquellas implicadas en la utilización de una fuente de carbono. Los ejemplos de dichas enzimas incluyen, por ejemplo, sacarasas, lactasas, maltasas, enzimas catabólicas de almidón, enzimas catabólicas de glucógeno, celulasas y poli(hidroxicanoato) despolimerasas. Si la célula hospedadora bacteriana presenta actividad catabólica nativa del tipo seleccionado, ésta puede suprimirse antes de la transformación con el vector de restablecimiento de la prototofía. Las bacterias que presentan auxotrofia nativa para estos compuestos también pueden utilizarse en su estado nativo para dicha transformación. En aquellas realizaciones en las que un compuesto no importable o difusible en la célula puede

seleccionarse y suministrarse al medio, la enzima o enzimas que restablecen la prototrofia o permiten la prototrofia pueden secretarse para su uso. En ese caso, la enzima o enzimas secretadas pueden degradar el compuesto extracelularmente para producir compuestos más pequeños, por ejemplo glucosa, que son difusibles o importables en la célula, mediante la selección o el diseño de la secuencia codificante de la enzima o enzimas para incluir una

5 secuencia codificante para un péptido señal de secreción operativo dentro de la célula hospedadora elegida. En estas realizaciones, el gen que restablece la prototrofia puede seleccionarse o modificarse mediante ingeniería genética para incluir una secuencia codificante para un péptido señal de secreción operativo dentro de la célula hospedadora elegida para la obtención de transporte de la enzima a través de la membrana citoplasmática. En cualquiera de estas realizaciones, o aquellas en las que el compuesto seleccionado es importable o difusible en la

10 célula, la célula se cultiva en un medio que no suministra ninguna otra fuente de carbono aparte del compuesto seleccionado.

En un sistema marcador basado en la utilización de una fuente de carbono, cada enzima de restauración de la prototrofia o enzima de utilización de una fuente de carbono que permite la prototrofia puede estar implicada en la

15 utilización de solo una fuente de carbono. Por ejemplo, dos genes de la misma vía catabólica pueden expresarse juntos en un vector o pueden coexpresarse por separado en diferentes vectores con el fin de proporcionar la prototrofia. Los ejemplos específicos de dichos sistemas marcadores basados en la utilización una fuente de carbono de múltiples genes incluyen, por ejemplo, el uso de glucógeno como única fuente de carbono con la expresión transgénica de tanto una glucógeno fosforilasa como una (alfa-1,4)glucanotransferasa; y el uso de almidón como única fuente de carbono con la expresión transgénica de tanto una alfa-amilasa como una alfa(1->6)

20 glucosidasa. Sin embargo, el sistema marcador de utilización de una fuente de carbono de uno o varios genes seleccionado puede utilizarse simultáneamente con otro tipo de sistema o sistemas marcadores en la misma célula hospedadora, a condición de que la única fuente de carbono proporcionada a la célula sea el compuesto seleccionado para su uso en el sistema marcador de selección catabólica de una fuente de carbono.

Otros ejemplos de enzimas útiles para las actividades de tipo de utilización bioquímica son bien conocidos en la técnica, y pueden incluir racemasas y epimerasas que son capaces de convertir una fuente de D-carbono no utilizable, suministrada a la célula, en una fuente de L-carbono nutritivo. Los ejemplos de estos sistemas incluyen, por ejemplo: un compuesto de D-ácido o D-acilo utilizado con la expresión transgénica de la racemasa correspondiente; y lactato utilizado con lactato racemasa expresada transgénicamente.

25

30

De forma similar, cuando una actividad biosintética de aminoácidos se ha seleccionado para su uso en el sistema marcador, la auxotrofia también puede superarse mediante el suministro a la célula de tanto un R-aminoácido no utilizable como una racemasa de R-aminoácido o epimerasa (EC 5.1.1) que convierte el R-aminoácido en el L-aminoácido correspondiente para el que la célula es auxotrófica.

35

Apilamiento de rasgos

También puede hacerse una pluralidad de cambios fenotípicos a una célula hospedadora, antes o después de la inserción de un gen marcador de selección auxotrófica, para la expresión del gen diana, de acuerdo con la presente invención. Por ejemplo, la célula puede modificarse por ingeniería genética, ya sea simultáneamente o secuencialmente, para presentar una diversidad de rasgos fenotípicos de potenciación. Este proceso se conoce como "apilamiento de rasgos". Por ejemplo, una delección de pyrF puede estar presente como uno de dichos rasgos fenotípicos. En dicha cepa, un gen pyrF, de acuerdo con la presente invención, puede usarse en un vector suicida como un marcador seleccionable y un marcador contraseleccionable (en presencia de ácido 5'-fluoroorótico) con el fin de efectuar un intercambio de alelos cruzado/tachado de otros rasgos deseables, por tanto, un gen pyrF de acuerdo con la presente invención puede usarse en un proceso para el "apilamiento de rasgos" de una célula hospedadora. En dicho proceso, un vector suicida que contiene un gen pyrF de este tipo puede transformarse en la cepa de célula hospedadora en una pluralidad de transformaciones separadas; en cada procedimiento de este tipo el restablecimiento del fenotipo pyrF puede usarse para crear, hasta el infinito, un cambio fenotípico genéticamente potenciador posterior. Por tanto, el gen pyrF no solo puede proporcionar un rasgo por sí mismo, sino que puede usarse para obtener rasgos fenotípicos adicionales en un proceso de apilamiento de rasgos.

40

45

50

En una realización, la presente invención proporciona células hospedadoras auxotróficas que han sido modificadas genéticamente adicionalmente para inducir auxotrofias adicionales. Por ejemplo, un auxótrofo para pyrF(-) puede modificarse adicionalmente para inactivar otra enzima biosintética presente en una vía anabólica o catabólica, tales como a través de la inactivación de un gen proC o un gen thyA. De esta manera, pueden producirse múltiples auxotrofias en la célula hospedadora.

55

En otra realización, pueden hacerse alteraciones genéticas a la célula hospedadora con el fin de mejorar la expresión de polipéptidos recombinantes en la célula hospedadora. Modificaciones adicionales pueden incluir alteraciones genéticas que permiten una utilización más eficiente de una fuente de carbono particular, optimizando de este modo la eficiencia global de toda la fermentación.

60

En una realización particular, las células hospedadoras auxotróficas se modifican adicionalmente mediante la inserción de un transgén que contiene lacl en el cromosoma hospedador. Preferentemente, el transgén lacl, o un

65

derivado del mismo, es distinto de parte de un gen estructural entero o truncado que contiene la construcción PlacI-lacI-lacZYA.

5 Análogos de aminoácidos y bases nitrogenadas

La presente invención abarca la incorporación en el polipéptido recombinante expresado por la célula hospedadora auxotrófica de uno o más análogos del metabolito para el que la célula hospedadora es auxotrófica. En diversas realizaciones, el análogo del metabolito es un aminoácido no natural o un compuesto de bases nitrogenadas no natural. Un experto en la materia reconocería qué análogos podrían sustituir a los diversos metabolitos descritos en el presente documento y son conocidos en la técnica.

"Análogo de aminoácido", "aminoácido no canónico", "aminoácido no natural", "aminoácido modificado", "sustrato AARS no natural", "sustrato AARS no natural", "aminoácido no convencional", "aminoácido no natural" y similares pueden usarse indistintamente y se entiende que incluyen todos los compuestos de tipo aminoácido que son similares en estructura y/o forma general a uno o más de los veinte L-aminoácidos que se encuentran habitualmente en proteínas de origen natural (Ala o A, Cys o C, Asp o D, Glu o E, Phe o F, Gly o G, His o H, Ile o I, Lys o K, Leu o L, Met o M, Asn o N, Pro o P, Gln o Q, Arg o R, Ser o S, Thr o T, Val o V, Trp o W, Tyr o Y, como se han definido y se enumeran en la Norma WIPO ST.25 (1998), Apéndice 2, Tabla 3). Los análogos de aminoácidos también pueden ser aminoácidos naturales con cadenas laterales o cadenas principales modificadas. Los aminoácidos también pueden ser aminoácidos de origen natural en forma D, en lugar de forma L. En algunas realizaciones, el aminoácido no natural comprende un grupo carbonilo, un grupo acetilo, un grupo aminooxi, un grupo hidrazina, un grupo hidrazida, un grupo semicarbazida, un grupo azida o un grupo alquino.

Análogamente, "análogo de base nitrogenada", "base nitrogenada no natural", "base nitrogenada modificada", "base nitrogenada no convencional", "base nitrogenada no natural" y similares, pueden usarse indistintamente y se entiende que incluyen todos los compuestos similares a bases nitrogenadas que son similares en estructura y/o forma general a una o más de las cinco bases nitrogenadas que se encuentran habitualmente en moléculas de ácido nucleico de origen natural.

Se conocen en la técnica análogos de aminoácidos y compuestos de bases nitrogenadas. Los ejemplos no limitantes incluyen pirazolo[3,4-d]pirimidinas, 5-metilcitosina (5-me-C), 5-hidroximetil citosina, xantina, hipoxantina, 2-aminoadenina, 6-metil y otros derivados de alquilo de adenina y guanina, 2-propil y otros derivados de alquilo de adenina y guanina, 2-tiouracilo, 2-tiotimina y 2-tiocitosina, 5-propinil uracilo y citosina, 6-azo uracilo, citosina y timina, 5-uracilo (pseudouracilo), 4-tiouracilo, 8-halo (por ejemplo, 8-bromo), 8-amino, 8-tiol, 8-tioalquilo, 8-hidroxilo y otras adeninas y guaninas 8-sustituidas, 5-halo en particular 5-bromo, 5-trifluorometil y otros uracilos y citosinas 5-sustituidos, 7-metilguanina y 7-metiladenina, 8-azaguanina y 8-azaadenina, desazaguanina, 7-desazaguanina, 3-desazaguanina, desazaadenina, 7-desazaadenina, 3-desazaadenina, pirazolo[3,4-d]pirimidina, imidazo[1,5-a]1,3,5 triazinonas, 9-desazapurinas, imidazo[4,5-d]pirazinas, tiazolo[4,5-d]pirimidinas, pirazin-2-onas, 1,2,4-triazina, piridazina, 1,3,5-triazina, yodotirosina, azidohomoalanina, homopropargilglicina, para-bromofenilalanina, para-yodofenilalanina, azidofenilalanina, acetilfenilalanina, etinilfenilalanina, azidonorleucina, 3-(1-naftil)alanina, 3-(2-naftil)alanina, p-etinil-fenilalanina, p-propargil-oxi-fenilalanina, m-etinil-fenilalanina, 6-etinil-triptófano, 5-etinil-triptófano, ácido (R)-2-amino-3-(4-etinil-1H-pirol-3-il)propánico, p-bromofenilalanina, p-idiofenilalanina, p-azidofenilalanina, 3-(6-clorindolil)alanina, 3-(6-bromoindolil)alanina, 3-(5-bromoindolil)alanina, azidohomoalanina y p-clorofenilalanina.

45 Modificaciones para inducir la auxotrofia

Una célula hospedadora seleccionada para su uso en un sistema de expresión de acuerdo con la presente invención puede ser deficiente en su capacidad para expresar cualquier biocatalizador funcional que presente la actividad auxotrófica seleccionada. Por ejemplo, cuando se selecciona una actividad orotidina-5'-fosfato descarboxilasa, la célula hospedadora puede ser deficiente en su capacidad para expresar a) cualquier producto del gen pyrF (es decir, cualquier enzima ODCasa funcional) y b) cualquier reemplazo eficaz del mismo (es decir, cualquier otro biocatalizador que tenga actividad ODCasa). En una realización, la célula hospedadora se hará biocatalíticamente deficiente para la actividad seleccionada mediante la alteración de su gen o genes genómicos de manera que la célula no pueda expresar, a partir de su genoma, una enzima funcional implicada en la auxotrofia objetivo (es decir, ODCasa). En otras palabras, la célula prototrófica (célula actividad(+)) se convertirá en auxotrófica a través de la "supresión" de un gen de codificación enzimática funcional implicado en la vía prototrófica objetivo (es decir, una célula actividad(-)). Esta alteración puede hacerse mediante la alteración de la secuencia codificante genómica de la célula o células del gen o genes que codifican la actividad o actividades seleccionadas. En una realización, la alteración o alteraciones de secuencia codificante se realizará mediante la introducción de: mutación o mutaciones de inserción o delección que cambian la secuencia codificante del marco o marcos de lectura; mutaciones de sustitución o inversión que alteran un número suficiente de codones; y/o mutaciones de delección que eliminan un grupo suficientemente grande de codones contiguos de la misma capaz de producir una enzima no funcional.

En una realización en la que la cepa de célula hospedadora también ha proporcionado el gen o genes auxotróficos para su uso como marcador o marcadores de selección en la misma, preferentemente cada uno de los elementos

promotores de la transcripción y/o terminadores de la transcripción del gen seleccionado también puede inactivarse mediante la introducción de mutación o mutaciones, incluyendo mutaciones de delección. Por ejemplo, la inactivación del elemento de transcripción puede realizarse opcionalmente además de la alteración o alteraciones de la secuencia codificante descrita anteriormente. En una realización en la que la cepa de célula hospedadora también ha proporcionado el gen o genes marcadores de selección auxotrófica, todo el ADN del gen o genes seleccionados puede ser eliminarse del genoma de la célula hospedadora.

Dichas cepas suprimidas pueden prepararse de acuerdo con cualquiera de los diversos métodos conocidos en la técnica como eficaces. Por ejemplo, pueden transformarse en la célula hospedadora vectores de recombinación homóloga que contienen secuencias 5' y 3' de genes homólogos dirigidos de la secuencia de delección de ácido nucleico deseada. Idealmente, tras la recombinación homóloga, puede producirse una supresión de gen enzimático dirigida deseada. Un experto en la materia reconocerá adicionalmente que hay disponible en el mercado una diversidad de estirpes celulares auxotróficas.

Los ejemplos específicos de metodologías de supresión génica incluyen, por ejemplo: se ha descrito anteriormente la inactivación génica mediante la inserción de un polinucleótido. Véase, por ejemplo, D L Roeder y A Collmer, *Marker-exchange mutagenesis of a pectate lyase isozyme gene in Erwinia chrysanthemi*, *J Bacteriol.* 164(1):51-56 (1985). Como alternativa, la mutagénesis de transposón y la selección para el fenotipo deseado (tal como la incapacidad de metabolizar benzoato o antranilato) pueden usarse para aislar cepas bacterianas en las que se han inactivado los genes diana mediante inserción. Véase, por ejemplo, K Nida y P P Cleary, *Insertional inactivation of streptolysin S expression in Streptococcus pyogenes*, *J Bacteriol.* 155(3):1156-61 (1983). Pueden construirse mutaciones o delecciones específicas en un gen particular usando mutagénesis de casete, por ejemplo, como se describe en J A Wells et al., *Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites*, *Gene* 34(2-3):315-23 (1985); por la cual se hacen mutaciones directas o aleatorias en una porción seleccionada de un gen y después se incorporan en la copia cromosómica del gen mediante recombinación homóloga.

En una realización, tanto el organismo del que se obtiene el gen o genes marcadores de selección como la célula hospedadora en la que se utiliza el gen o genes marcadores de selección pueden seleccionarse entre un procarionte.

En una realización particular, tanto el organismo del que se obtiene el gen o genes marcadores de selección como la célula hospedadora en la que se utiliza un gen o genes marcadores de selección pueden seleccionarse entre una bacteria. En otra realización, tanto las bacterias de las que se obtiene el gen o genes marcadores de selección como la célula hospedadora bacteriana en la que se utiliza un gen o genes marcadores de selección, se seleccionará entre las Proteobacterias. En otra realización más, tanto las bacterias de las que se obtiene el gen o genes marcadores de selección como las células hospedadoras bacterianas en las que se utiliza un gen o genes marcadores de selección, pueden seleccionarse entre las Pseudomonas y bacterias estrechamente relacionadas o entre un subgrupo de las mismas, como se define a continuación.

En una realización particular, tanto el organismo fuente del gen o genes marcadores de selección como la célula hospedadora pueden seleccionarse entre la misma especie. Preferentemente, la especie será un procarionte; más preferentemente una bacteria, aún más preferentemente una proteobacteria. En otra realización particular, tanto el organismo fuente del gen o genes marcadores de selección como la célula hospedadora pueden seleccionarse entre la misma especie en un género seleccionado entre las Pseudomonas y bacterias estrechamente relacionadas o entre un subgrupo de las mismas, como se define a continuación. En una realización, tanto el organismo fuente del gen o genes marcadores de selección como la célula hospedadora pueden seleccionarse entre una especie del género Pseudomonas, en particular la especie Pseudomonas fluorescens y preferentemente la especie Pseudomonas fluorescens biotipo A.

Construcciones de ácidos nucleicos

En otro aspecto más de la presente invención, se proporcionan construcciones de ácido nucleico para su uso en la producción mejorada de péptidos. En una realización, se proporciona una construcción de ácido nucleico para su uso en la transformación de una célula hospedadora auxotrófica que comprende a) una secuencia de ácido nucleico que codifica un polipéptido recombinante y b) una secuencia de ácido nucleico que codifica una enzima que permite la prototofia.

En una realización de la presente invención, se proporciona una construcción de ácido nucleico que comprende ácidos nucleicos que codifican al menos una enzima biosintética capaz de transformar una célula hospedadora auxotrófica a la prototofia. La enzima biosintética puede ser cualquier enzima capaz de permitir que una célula hospedadora auxotrófica sobreviva en un medio de selección en el que, sin la expresión de la enzima biosintética, la célula hospedadora sería incapaz de sobrevivir debido a la deficiencia metabólica auxotrófica. Como tal, la enzima biosintética puede ser una enzima que complemente la deficiencia metabólica del hospedador auxotrófico mediante el restablecimiento de la capacidad prototrófica para crecer en medios complementados con el metabolito no auxotrófico.

En una realización alternativa, la presente invención proporciona una construcción de ácido nucleico que codifica al menos una enzima biosintética capaz de transformar una célula hospedadora auxotrófica a la prototrofia y un marcador de selección no auxotrófico adicional. Los ejemplos de marcadores de selección no auxotróficos son bien conocidos en la técnica y pueden incluir marcadores que dan origen a una reacción colorimétrica/cromogénica o luminiscente tal como el gen lacZ, el gen GUS, el gen CAT, el gen luxAB, marcadores de selección de resistencia a antibióticos tales como genes de resistencia a anfotericina B, bacitracina, carbapenem, cefalosporina, etambutol, fluoroquinolonas, isonizid, cefalosporina, meticilina, oxacilina, vancomicina, estreptomycin, quinolinas, rifampina, rifampicina, sulfonamidas, ampicilina, tetraciclina, neomicina, cefalotina, eritromicina, estreptomycin, kanamicina, gentamicina, penicilina y cloranfenicol, u otros marcadores de selección no auxotróficos utilizados habitualmente. En diversas realizaciones, el sistema de expresión carece específicamente de un marcador de selección de antibióticos.

En otra realización, el vector de expresión puede comprender más de una enzima biosintética capaz de transformar una célula hospedadora auxotrófica a la prototrofia. Las enzimas biosintéticas pueden ser cualesquiera enzimas capaces de permitir que una célula hospedadora auxotrófica sobreviva en un medio de selección en el que, sin la expresión de la enzima biosintética, la célula hospedadora sería incapaz de sobrevivir debido a la deficiencia metabólica auxotrófica. Como tales, las enzimas biosintéticas pueden ser enzimas que complementan las deficiencias metabólicas del hospedador auxotrófico mediante el restablecimiento de la capacidad prototrófica para crecer en medios complementados con el metabolito no auxotrófico. Por ejemplo, un vector de expresión comprende un primer y un segundo gen marcador de selección que permite la prototrofia, permitiendo que la célula hospedadora que contiene la construcción se mantenga en una o ambas de las condiciones en las que la supervivencia de la célula hospedadora necesita la presencia del gen o genes marcadores de selección. Cuando solo una de las condiciones de supervivencia dependientes del gen marcador está presente, el gen marcador correspondiente debe expresarse y el otro gen o genes marcadores pueden entonces estar activos o inactivos, aunque todos los nutrientes necesarios para los cuales la célula permanece auxotrófica todavía serán suministrados por el medio. Esto permite que el mismo gen diana, o el mismo conjunto de genes diana unidos covalentemente, que codifica el producto o productos transgénicos deseados y/o la actividad o actividades transgénicas deseadas, se mantenga en la célula hospedadora de forma continua a medida que la célula hospedadora se somete a la transición entre o entre diferentes condiciones. La secuencia codificante de cada uno de los genes marcadores de selección elegidos puede estar unida operativamente independientemente a un promotor ya sea constitutivo o regulado.

Promotores

En un proceso de fermentación, una vez que se induce la expresión del polipéptido recombinante diana, es ideal tener un alto nivel de producción con el fin de maximizar la eficiencia del sistema de expresión. El promotor inicia la transcripción y generalmente se coloca 10-100 nucleótidos corriente arriba del sitio de unión al ribosoma. Idealmente, un promotor será lo suficientemente fuerte para permitir una acumulación de polipéptido recombinante de aproximadamente el 50 % de la proteína celular total de la célula hospedadora, estará sujeto a una regulación estricta y se inducirá fácilmente (y económicamente).

Los promotores utilizados de acuerdo con la presente invención pueden ser promotores constitutivos o promotores regulados. Los ejemplos de promotores inducibles utilizados habitualmente y sus posteriores inductores incluyen lac (IPTG), lacUV5 (IPTG), tac (IPTG), trc (IPTG), P.sub.syn (IPTG), trp (privación de triptófano), araBAD (1-arabinosa), 1pp.sup.a (IPTG), 1pp-lac (IPTG), phoA (privación de fosfato), recA (ácido nalidixico), proU (osmolaridad), cst-1 (privación de glucosa), teta (tetraciclina), cada (pH), nar (condiciones anaeróbicas), PL (desplazamiento térmico a 42 grados C.), cspA (desplazamiento térmico a 200 °C.), T7 (inducción térmica), operador T7-lac (IPTG), operador T3-lac (IPTG), operador T5-lac (IPTG), T4gene32 (infección por T4), operador nprM-lac (IPTG), Pm (alquil o halo-benzoatos), Pu (alquil o halo-toluenos), Psa1 (salicilatos) y VHb (oxígeno). Véase, por ejemplo, Makrides, S. C. (1996) *Microbiol. Rev.* 60, 512-538; Hannig G. y Makrides, S. C. (1998) *TIBTECH* 16, 54-60; Stevens, R. C. (2000) *Structures* 8, R177-R185. Véase, por ejemplo: J. Sanchez-Romero y V. De Lorenzo, *Genetic Engineering of Nonpathogenic Pseudomonas strains as Biocatalysts for Industrial and Environmental Processes*, en *Manual of Industrial Microbiology and Biotechnology* (A. Demain y J. Davies, ed.) págs.460-74 (1999) (ASM Press, Washington, D.C.); H. Schweizer, *Vectors to express foreign genes and techniques to monitor gene expression for Pseudomonads*, *Current-Opinion in Biotechnology*, 12:439-445 (2001); y R. Slater y R. Williams, *The Expression of Foreign DNA in Bacteria*, en *Molecular Biology and Biotechnology* (J. Walker y R. Rapley, ed.) págs. 125-54 (2000) (The Royal Society of Chemistry, Cambridge, RU).

Un promotor que tiene la secuencia de nucleótidos de un promotor nativo para la célula hospedadora auxotrófica seleccionada también puede usarse para controlar la expresión del transgén que codifica el polipéptido diana, por ejemplo, un antranilato de *Pseudomonas* o un promotor del operón de benzoato (Pant, Pben). También pueden usarse promotores en tándem en los que más de un promotor están unidos covalentemente entre sí, ya sean iguales o diferentes en secuencia, por ejemplo, un promotor en tándem Pant-Pben (interpromotor híbrido) o un promotor en tándem Plac-Plac. Véase, por ejemplo, la Publicación de Patente de los EE.UU. N.º 20050202544.

Los promotores regulados utilizan proteínas reguladoras de promotores con el fin de controlar la transcripción del gen del que el promotor es una parte. Cuando se usa un promotor regulado en el presente documento, una proteína reguladora del promotor correspondiente también será parte de un sistema de expresión de acuerdo con la presente

invención. Los ejemplos de proteínas reguladoras de promotores incluyen: proteínas activadoras, por ejemplo, proteína activadora de catabolitos de *E. coli*, proteína MalT; activadores transcripcionales de la familia AraC; proteínas represoras, por ejemplo, proteínas LacI de *E. coli*; y proteínas reguladoras de doble facción, por ejemplo, la proteína NagC de *E. coli*. Se conocen en la técnica muchos pares promotor regulado/proteína reguladora del promotor.

Las proteínas reguladoras de promotores interactúan con un compuesto efector, es decir, un compuesto que se asocia de forma reversible o irreversible a la proteína reguladora con el fin de permitir que la proteína ya sea se libere o se una a al menos una región reguladora de la transcripción de ADN del gen que está bajo el control del promotor, permitiendo o bloqueando de este modo la acción de una enzima transcriptasa en la iniciación de la transcripción del gen. Los compuestos efectores se clasifican ya sea como inductores o co-represores y estos compuestos incluyen compuestos efectores nativos y compuestos inductores gratuitos. Se conocen en la técnica muchos tríos promotor regulado/proteína reguladora del promotor/compuesto efector. Aunque puede usarse un compuesto efector en todo el cultivo celular o la fermentación, en una realización particular en la que se usa un promotor regulado, después del crecimiento de una cantidad o densidad deseada de la biomasa de células hospedadoras, se añade un compuesto efector apropiado al cultivo con el fin de dar como resultado, directa o indirectamente, la expresión del gen o genes diana deseados.

A modo de ejemplo, cuando se utiliza un promotor de la familia lac, un gen lacI o un derivado del mismo tal como un gen lacI^{sup.Q} o lacI^{sup.Q1}, también puede estar presente en el sistema. El gen lacI, que es (normalmente) un gen expresado constitutivamente, codifica la proteína represora de Lac (proteína LacI), que se une al operador lac de estos promotores. De este modo, cuando se utiliza un promotor de la familia lac, el gen lacI también puede incluirse y expresarse en el sistema de expresión. En el caso de los miembros de la familia del promotor lac, por ejemplo, el promotor tac, el compuesto efector es un inductor, preferentemente un inductor gratuito tal como IPTG (isopropil-beta-D-1-tiogalactopiranosido, también denominado "isopropiltiogalactósido").

En una realización particular, se utiliza un promotor de la familia lac o tac en la presente invención, incluyendo Plac, Ptac, Ptrc, PtaclI, PlacUV5, lpp-PlacUV5, lpp-lac, nprM-lac, T7lac, T5lac, T3lac y Pmac.

Otros elementos

Pueden incluirse otros elementos reguladores en una construcción de expresión, incluyendo secuencias lacO y derivados, como se analiza en la Publicación de Patente de los EE.UU. N.º 20050186666. Dichos elementos incluyen, pero no se limitan a, por ejemplo, secuencias potenciadoras transcripcionales, secuencias potenciadoras traduccionales, otros promotores, activadores, señales de inicio y final de la traducción, terminadores de la transcripción, reguladores cistrónicos, reguladores policistrónicos, secuencias marcadoras, tales como "marcadores" de secuencias de nucleótidos y secuencias codificantes de péptido "marcador", que facilita la identificación, separación, purificación o aislamiento de un polipéptido expresado, incluyendo marcador His, marcador Flag, marcador T7, marcador S, marcador HSV, marcador B, marcador Strep, poliarginina, policisteína, polifenilalanina, ácido poliaspártico, (Ala-Trp-Trp-Pro)_n, tioredoxina, beta-galactosidasa, cloranfenicol acetiltransferasa, ciclomaltodextrina gluconotransferasa, CTP:CMP-3-desoxi-D-mano-octulosonato citidiltransferasa, trpE o el trpLE, avidina, estreptavidina, gen 10 de T7, gp55 T4, proteína estafilocócica A, proteína estreptocócica G, GST, DHFR, CBP, MBP, dominio de unión a galactosa, dominio de unión a calmodulina, GFP, KSI, c-myc, ompT, ompA, pelB, NusA, ubiquitina y hemolisina A.

Como mínimo, un gen que codifica proteína de acuerdo con la presente invención puede incluir, además de la secuencia codificante de la proteína, los siguientes elementos reguladores unidos operativamente al mismo: un promotor, un sitio de unión al ribosoma (RBS), un terminador de la transcripción, señales de inicio y de terminación de la traducción. Pueden obtenerse RBS útiles de cualquiera de las especies útiles como células hospedadoras en sistemas de expresión de acuerdo con la presente invención, preferentemente de la célula hospedadora seleccionada. Se conocen muchos RBS específicos y una diversidad de RBS de consenso, por ejemplo, aquellos descritos y referenciados por D. Frishman et al., *Starts of bacterial genes: estimating the reliability of computer predictions*, *Gene* 234(2):257-65 (8 de julio de 1999); y B. E. Suzek et al., *A probabilistic method for identifying start codons in bacterial genomes*, *Bioinformatics* 17(12):1123-30 (diciembre de 2001). Además, pueden usarse RBS ya sea nativos o sintéticos, por ejemplo, los descritos en: el documento EP 0207459 (RBS sintético); O. Ikehata et al., *Primary structure of nitrile hydratase deduced from the nucleotide sequence of a Rhodococcus species and its expression in Escherichia coli*, *Eur. J. Biochem.* 181(3):563-70 (1989) (secuencia de AAGGAAG de RBS nativo). Se describen ejemplos adicionales de métodos, vectores y elementos de traducción y transcripción, y otros elementos útiles en la presente invención en, por ejemplo: la Patente de los EE.UU. N.º 5.055.294 para Gilroy y la Patente de los EE.UU. N.º 5.128.130 para Gilroy et al.; la Patente de los EE.UU. N.º 5.281.532 para Rammler et al.; las Patentes de los EE.UU. N.º 4.695.455 y 4.861.595 para Barnes et al.; la Patente de los EE.UU. N.º 4.755.465 para Gray et al.; y la patente de los EE.UU. N.º 5.169.760 para Wilcox.

Vectores

La transcripción del ADN que codifica las enzimas de la presente invención por un hospedador *Pseudomonas* puede aumentarse adicionalmente insertando una secuencia potenciadora en el vector o plásmido. Los potenciadores típicos son elementos del ADN que actúan en cis, generalmente de aproximadamente 10 a 300 pb de tamaño, que actúan sobre el promotor para aumentar su transcripción.

Generalmente, los vectores de expresión recombinantes incluirán orígenes de replicación y marcadores seleccionables que permiten la transformación de la célula hospedadora *Pseudomonas*, por ejemplo, los genes de restablecimiento de la prototrofia de la presente invención y un promotor derivado de un gen altamente expresado para dirigir la transcripción de una secuencia estructural corriente abajo. Dichos promotores se han descrito anteriormente. La secuencia estructural heteróloga se ensambla en la fase apropiada con secuencias de iniciación y de terminación de la traducción y, en ciertas realizaciones, una secuencia líder capaz de dirigir la secreción del polipéptido traducido. Opcionalmente, y de acuerdo con la presente invención, la secuencia heteróloga puede codificar un polipéptido de fusión que incluye un péptido de identificación N-terminal que transmite las características deseadas, por ejemplo, la estabilización o la purificación simplificada del producto recombinante expresado.

Se construyen vectores de expresión útiles para su uso en la presente invención insertando una secuencia de ADN estructural que codifica un polipéptido diana deseada junto con señales de iniciación y de terminación de la traducción adecuadas en fase de lectura operable con un promotor funcional. El vector comprenderá uno o más marcadores seleccionables fenotípicos y un origen de replicación para asegurar el mantenimiento del vector y para, si se desea, proporcionar amplificación dentro del hospedador.

Los vectores son conocidos en la técnica como útiles para expresar proteínas recombinantes en células hospedadoras, y cualquiera de éstos pueden modificarse y utilizarse para expresar los genes de acuerdo con la presente invención. Dichos vectores incluyen, por ejemplo, plásmidos, cósmidos y vectores de expresión de fagos.

Los ejemplos de vectores de plásmidos útiles que pueden modificarse para su uso en la presente invención incluyen, pero no se limitan a, los plásmidos de expresión pBBR1MCS, pDSK519, pKT240, pML122, pPS10, RK2, rK6, pRO1600 y RSF1010. Ejemplos adicionales pueden incluir pALTER-Ex1, pALTER-Ex2, pBAD/His, pBAD/MycHis, pBAD/gIII, pCal-n, pCal-n-EK, pCal-c, pCal-Kc, pcDNA 2.1, pDUAL, pET-3a-c, pET 9a-d, pET-11a-d, pET-12a-c, pET-14b, pET15b, pET-16b, pET-17b, pET-19b, pET-20b(+), pET-21a-d(+), pET-22b(+), pET-23a-d(+), pET24a-d(+), pET-25b(+), pET-26b(+), pET-27b(+), pET28a-c(+), pET-29a-c(+), pET-30a-c(+), pET31b(+), pET-32a-c(+), pET33b(+), pET-34b(+), pET35b(+), pET-36b(+), pET-37b(+), pET-38b(+), pET-39b(+), pET-40b(+), pET41la-c(+), pET42a-c(+), pET43a-c(+), pETBlue-1, pETBlue-2, pETBlue-3, pGEMEX-1, pGEMEX-2, pGEX1.lambda.T, pGEX-2T, pGEX-2TK, pGEX-3X, pGEX4T, pGEX-5X, pGEX-6P, pHAT10/11/12, pHAT20, pHAT-GFPuv, pKK223-3, pLEX, pMAL-c2X, pMAL-c2E, pMAL-c2g, pMAL-p2X, pMAL-p2E, pMAL-p2G, pProEX HT, pPROLar.A, pPROTet.E, pQE-9, pQE-16, pQE30/31/32, pQE40, pQE-50, pQE-70, pQE-80/81/82L, pQE-100, pRSET, y pSE280, pSE380, pSE420, pThioHis, pTrc99A, pTrcHis, pTrcHis2, pTriEx-1, pTriEx-2, pTrxFus. Otros ejemplos de dichos vectores útiles incluyen aquellos descritos por, por ejemplo: N. Hayase, en *Appl. Envir. Microbiol.* 60(9):3336-42 (septiembre de 1994); A. A. Lushnikov et al., en *Basic Life Sci.* 30:657-62 (1985); S. Graupner y W. Wackernagel, en *Biomolec. Eng.* 17(1):11-16. (Octubre de 2000); H. P. Schweizer, en *Curr. Opin. Biotech.* 12(5):439-45 (Octubre de 2001); M. Bagdasarian y K. N. Timmis, en *Curr. Topics Microbiol. Immunol.* 96:47-67 (1982); T. Ishii et al., en *FEMS Microbiol. Lett.* 116(3):307-13 (1 de marzo de 1994); I. N. Olekhovich y Y. K. Fomichev, en *Gene* 140(1):63-65 (11 de marzo de 1994); M. Tsuda y T. Nakazawa, en *Gene* 136(1-2):257-62 (22 de diciembre de 1993); C. Nieto et al., en *Gene* 87(1):145-49 (1 de marzo de 1990); J. D. Jones y N. Gutterson, en *Gene* 61(3):299-306 (1987); M. Bagdasarian et al., en *Gene* 16(1-3):237-47 (diciembre de 1981); H. P. Schweizer et al., en *Genet. Eng.* (NY) 23:69-81 (2001); P. Mukhopadhyay et al., en *J. Bact.* 172(1):477-80 (enero de 1990); D. O. Wood et al., en *J. Bact.* 145(3):1448-51 (marzo de 1981); Holtwick et al., en *Microbiology* 147(Pt 2):337-44 (febrero de 2001).

Ejemplos adicionales de vectores de expresión que pueden ser útiles en células hospedadoras auxotróficas incluyen aquellos enumerados en la Tabla 1 como derivados de los replicones indicados.

Replicón	Vector o vectores
PPS10	PCN39, PCN51
RSF1010	PKT261-3
	PMMB66EH
	PEB8
	PPLGN1
	PMYC1050
RK2/RP1	PRK415
	PJB653

PRO1600	PUCP
	PBSP

El plásmido de expresión, RSF1010, se describe, por ejemplo, por F. Heffron et al., en *Proc. Nat'l Acad. Sci. USA* 72(9):3623-27 (septiembre de 1975), y por K. Nagahari y K. Sakaguchi, en *J. Bact.* 133(3):1527-29 (marzo de 1978). El plásmido RSF10100 y derivados del mismo son vectores particularmente útiles en la presente invención. Los ejemplos de derivados útiles de RSF1010, que son conocidos en la técnica, incluyen, por ejemplo, pKT212, pKT214, pKT231 y plásmidos relacionados, y pMYC1050 y plásmidos relacionados (véanse, por ejemplo, las Patentes de los EE.UU. N.º 5.527.883 y 5.840.554 para Thompson et al.), tales como, por ejemplo, pMYC1803. El plásmido pMYC1803 deriva del plásmido pTJS260 basado en RSF1010 (véase la Patente de los EE.UU. N.º 5.169.760 para Wilcox), que lleva un marcador de resistencia a tetraciclina regulado y los loci de replicación y movilización del plásmido RSF1010. Otros vectores útiles de ejemplo incluyen aquellos descritos en la Patente de los EE.UU. N.º 4.680.264 para Pühler et al.

En una realización, un plásmido de expresión se usa como vector de expresión. En otra realización, RSF1010 o un derivado del mismo se usa como vector de expresión. En otra realización más, pMYC1050 o un derivado del mismo, o pMYC1803 o un derivado del mismo, se usa como vector de expresión.

Célula hospedadora

De acuerdo con la presente invención, la célula hospedadora puede seleccionarse entre el "Subgrupo 18 de Proteobacterias gramnegativas". El "subgrupo 18 de Proteobacterias gramnegativas" se define como el conjunto de todas las subespecies, variedades, cepas y otras unidades de subespecies de la especie *Pseudomonas fluorescens*, incluyendo aquellas que pertenecen, por ejemplo, a las siguientes (mostrándose el número de depósito de ATCC u otros de la cepa o cepas de ejemplo entre paréntesis): *Pseudomonas fluorescens* biotipo A, también denominada biovariedad 1 o biovariedad I (ATCC 13.525); *Pseudomonas fluorescens* biotipo B, también denominada biovariedad 2 o biovariedad II (ATCC 17816); *Pseudomonas fluorescens* biotipo C, también denominada biovariedad 3 o biovariedad III (ATCC 17.400); *Pseudomonas fluorescens* biotipo F, también denominada biovariedad 4 o biovariedad IV (ATCC 12.983); *Pseudomonas fluorescens* biotipo G, también denominada biovariedad 5 o biovariedad V (ATCC 17.518); *Pseudomonas fluorescens* biovariedad VI; *Pseudomonas fluorescens* Pf0-1; *Pseudomonas fluorescens* Pf-5 (ATCC BAA-477); *Pseudomonas fluorescens* SBW25; y *Pseudomonas fluorescens* subespecie celulosa (NCIMB 10462).

La célula hospedadora puede seleccionarse entre el "Subgrupo 19 de Proteobacterias gramnegativas". El "subgrupo 19 de Proteobacterias gramnegativas" se define como el conjunto de todas las cepas de *Pseudomonas fluorescens* biotipo A. Se ha descubierto que *Pseudomonas fluorescens* (Pf) no presenta los problemas inherentes asociados a la alimentación cruzada observados en otros sistemas de células hospedadoras, por ejemplo, *E. coli* y levadura.

Aunque no se desea quedar ligado a ninguna teoría particular, se cree que la *Pseudomonas fluorescens* auxotrófica es un organismo particularmente adecuado para su uso como célula hospedadora debido a la incapacidad observada de una célula Pf auxotrófica para superar competitivamente a una célula auxotrófica que contiene un plásmido que permite la prototrofia en un medio complementado que contiene el metabolito auxotrófico, indicando una dificultad innata de una Pf auxotrófica para importar el metabolito necesario. Por tanto, las células Pf auxotróficas que pierden el plásmido marcador de selección no ganan una ventaja selectiva sobre las células auxotróficas Pf que contienen el marcador de selección, incluso en presencia de un metabolito complementario, reduciendo en gran medida cualquier efecto potencial de alimentación cruzada. Debido a los efectos reducidos de alimentación cruzada, los rendimientos de la producción del polipéptido recombinante en una operación de fermentación (en particular ejecuciones de fermentación de lote grande) no se reducen debido a la presencia de células productoras de polipéptidos no recombinantes.

Una cepa particularmente preferida de este biotipo es la cepa de *P. fluorescens* MB101 (véase la Patente de los EE.UU. N.º 5.169.760 para Wilcox) y derivados de la misma. Un ejemplo de un derivado preferido de la misma es la cepa de *P. fluorescens* MB214, construida mediante la inserción en el locus *asd* cromosómico MB101 (gen de la deshidrogenasa aspartato), una construcción *PlacI-lacI-lacZYA* nativa de *E. coli* (es decir, en la que *PlacZ* se ha eliminado).

Las cepas adicionales de *P. fluorescens* que pueden usarse en la presente invención incluyen *Pseudomonas fluorescens* Migula y *Pseudomonas fluorescens* Loitokitok, que tienen las designaciones de ATCC siguientes: [NCIB 8286]; NRRL B-1244; NCIB 8865 cepa CO1; NCIB 8866 cepa CO2; 1291 [ATCC 17458; IFO 15837; NCIB 8917; LA; NRRL B-1864; pirrolidina; PW2 [ICMP 3966; NCPPB 967; NRRL B-899]; 13475; NCTC 10038; NRRL B-1603 [6; IFO 15840]; 52-1C; CCEB 488A [BU 140]; CCEB 553 [EM 15/47]; IAM 1008 [AHH-27]; IAM 1055 [AHH-23]; 1 [IFO 15842]; 12 [ATCC 25323; NIH 11; den Dooren de Jong 216]; 18 [IFO 15833; WRRL P-7]; 93 [TR-10]; 108 [52-22; IFO 15832]; 143 [IFO 15836; PL]; 149 [2-40-40; IFO 15838]; 182 [IFO 3081; PJ 73]; 184 [IFO 15830]; 185 [W2 L-1]; 186 [IFO 15829; PJ 79]; 187 [NCPPB 263]; 188 [NCPPB 316]; 189 [PJ227; 1208]; 191 [IFO 15834; PJ 236; 22/1]; 194 [Klinge R-60; PJ 253]; 196 [PJ 288]; 197 [PJ 290]; 198 [PJ 302]; 201 [PJ 368]; 202 [PJ 372]; 203 [PJ 376]; 204

[IFO 15835; PJ 682]; 205 [PJ 686]; 206 [PJ 692]; 207 [PJ 693]; 208 [PJ 722]; 212. [PJ 832]; 215 [PJ 849]; 216 [PJ 885]; 267 [B-9]; 271 [B-1612]; 401 [C71A; IFO 15831; PJ 187]; NRRL B-3178 [4; IFO. 15841]; KY 8521; 3081; 30-21; [IFO 3081]; N; PYR; PW; D946-B83 [BU 2183; FERMP 3328]; P-2563 [FERM-P 2894; IFO 13658]; IAM-1126 [43F]; M-1; A506 [A5-06]; A505 [A5-05-1]; A526 [A5-26]; B69; 72; NRRL B-4290; PMW6 [NCIB 11615]; SC 12936; AI [IFO 15839]; F 1847 [CDC-EB]; F 1848 [CDC 93]; NCIB 10586; P17; F-12; AmMS 257; PRA25; 6133D02; 6519E01; Ni; SC15208; BNL-WVC; NCTC 2583 [NCIB 8194]; H13; 1013 [ATCC 11251; CCEB 295]; IFO 3903; 1062; o Pf-5.

De acuerdo con la invención, el hospedador será un miembro de los hospedadores Gamma proteobacterianos, del taxón Pseudomonadalesandin, en particular un miembro de la familia Pseudomonadaceae, del género Pseudomonas de la especie Pseudomonas fluorescens.

Otros organismos Pseudomonas también pueden ser útiles. Las Pseudomonas y las especies estrechamente relacionadas incluyen el Subgrupo 1 de Proteobacterias gramnegativas, que incluyen el grupo de Proteobacteria pertenecientes a las familias y/o géneros descritos como "Bacilos y cocos aeróbicos gramnegativos" por R. E. Buchanan y N. E. Gibbons (ed.), *Bergey's Manual of Determinative Bacteriology*, págs. 217-289 (octava ed., 1974) (The Williams & Wilkins Co., Baltimore, Md., EE.UU.) (en lo sucesivo en el presente documento "Bergey (1974)"). La Tabla 2 presenta estas familias y géneros de organismos.

Familia I. Pseudomonadaceae	<i>Gluconobacter</i> <i>Pseudomonas</i> <i>Xanthomonas</i> <i>Zoogloea</i>
Familia II. Azotobacteraceae	<i>Azomonas</i> <i>Azotobacter</i> <i>Beijerinckia</i> <i>Derxia</i>
Familia III. Rhizobiaceae	<i>Agrobacterium</i> <i>Rhizobium</i>
Familia IV. Metilomonadaceae	<i>Metilococcus</i> <i>Metilomonas</i>
Familia V. Halobacteriaceae	<i>Halobacterium</i> <i>Halococcus</i>
Otros géneros	<i>Acetobacter</i> <i>Alcaligenes</i> <i>Bordetella</i> <i>Brucella</i> <i>Francisella</i> <i>Thermus</i>

El "Subgrupo 1 de Proteobacterias gramnegativas" también incluye Proteobacterias que se clasifican en este encabezamiento de acuerdo con los criterios utilizados en la clasificación. El encabezamiento también incluye grupos que se clasificaron anteriormente en esta sección, pero ya no están, tales como los géneros *Acidovorax*, *Brevundimonas*, *Burkholderia*, *Hydrogenophaga*, *Oceanimonas*, *Ralstonia* y *Stenotrophomonas*, el género *Sphingomonas* (y el género *Blastomonas*, derivado del mismo), que se creó por los organismos de reagrupamiento que pertenecen a (y anteriormente denominados especies de) el género *Xanthomonas*, el género *Acidomonas*, que fue creado por los organismos de reagrupamiento que pertenecen al género *Acetobacter* como se define en Bergey (1974). Adicionalmente los hospedadores pueden incluir células del género *Pseudomonas*, *Pseudomonas enalia* (ATCC 14393), *Pseudomonas nigrifaciens* (ATCC 19375) y *Pseudomonas putrefaciens* (ATCC 8071), que se han reclasificado, respectivamente, como *Alteromonas haloplanktis*, *Alteromonas nigrifaciens* y *Alteromonas putrefaciens*. Del mismo modo, por ejemplo, *Pseudomonas acidovorans* (ATCC 15668) y *Pseudomonas testosteroni* (ATCC 11996) se han reclasificado desde entonces como *Comamonas acidovorans* y *Comamonas testosteroni*, respectivamente; y *Pseudomonas nigrifaciens* (ATCC 19375) y *Pseudomonas piscicida* (ATCC 15057) se han reclasificado, respectivamente, como *Pseudoalteromonas nigrifaciens* y *Pseudoalteromonas piscicida*. El "Subgrupo 1 de Proteobacterias gramnegativas" también incluye Proteobacterias clasificadas como pertenecientes a cualquiera de las familias: Pseudomonadaceae, Azotobacteraceae (ahora denominada con frecuencia mediante el sinónimo, el "grupo Azotobacter" de Pseudomonadaceae), Rhizobiaceae y Metilomonadaceae (ahora denominada con frecuencia mediante el sinónimo, "Metilococcaceae"). En consecuencia, además de los géneros descritos de otro modo en el presente documento, géneros de Proteobacterias adicionales que pertenecen al "Subgrupo 1 de Proteobacterias gramnegativas" incluyen: 1) bacterias del grupo Azotobacter del género *Azorhizophilus*; 2) bacterias de la familia Pseudomonadaceae de los géneros *Cellvibrio*, *Oligella* y *Teredinibacter*; 3) bacterias de la familia Rhizobiaceae de los géneros *Chelatobacter*, *Ensifer*, *Liberibacter* (también denominado "Candidatus Liberibacter") y *Sinorhizobium*; y

4) bacterias de la familia Methylococcaceae de los géneros *Methylobacter*, *Methylocaldum*, *Methylomicrobium*, *Methylosarcina* y *Methylosphaera*.

En otro ejemplo, la célula hospedadora se selecciona entre el "Subgrupo 2 de Proteobacterias gramnegativas". El "Subgrupo 2 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros (con los número totales de las cepas depositadas enumeradas en el catálogo, disponibles al público, de los mismos, indicados entre paréntesis, todos depositado en ATCC, a menos que se indique lo contrario): *Acidomonas* (2); *Acetobacter* (93); *Gluconobacter* (37); *Brevundimonas* (23); *Beyerinckia* (13); *Derxia* (2); *Brucella* (4); *Agrobacterium* (79); *Chelatobacter* (2); *Ensifer* (3); *Rhizobium* (144); *Sinorhizobium* (24); *Blastomonas* (1); *Sphingomonas* (27); *Alcaligenes* (88); *Bordetella* (43); *Burkholderia* (73); *Ralstonia* (33); *Acidovorax* (20); *Hydrogenophaga* (9); *Zoogloea* (9); *Methylobacter* (2); *Methylocaldum* (1 en NCIMB); *Methylococcus* (2); *Methylomicrobium* (2); *Methylomonas* (9); *Methylosarcina* (1); *Methylosphaera*; *Azomonas* (9); *Azorhizophilus* (5); *Azotobacter* (64); *Cellvibrio* (3); *Oligella* (5); *Pseudomonas* (1139); *Francisella* (4); *Xanthomonas* (229); *Stenotrophomonas* (50); y *Oceanimonas* (4).

Las especies de células hospedadoras de ejemplo del "Subgrupo 2 de Proteobacterias gramnegativas" incluyen, pero no se limitan a las siguientes bacterias (mostrándose los números de ATCC u otro depósito de la cepa o cepas de ejemplo de las mismas entre paréntesis): *Acidomonas methanolica* (ATCC 43581); *Acetobacter acetii* (ATCC 15973); *Gluconobacter oxydans* (ATCC 19357); *Brevundimonas diminuta* (ATCC 11568); *Beijerinckia indica* (ATCC 9039 y ATCC 19361); *Derxia gummosa* (ATCC 15994); *Brucella melitensis* (ATCC 23456), *Brucella abortus* (ATCC 23448); *Agrobacterium tumefaciens* (ATCC 23308), *Agrobacterium radiobacter* (ATCC 19358), *Agrobacterium rhizogenes* (ATCC 11325); *Chelatobacter heintzii* (ATCC 29600); *Ensifer adhaerens* (ATCC 33212); *Rhizobium leguminosarum* (ATCC 10004); *Sinorhizobium fredii* (ATCC 35423); *Blastomonas natatoria* (ATCC 35951); *Sphingomonas paucimobilis* (ATCC 29837); *Alcaligenes faecalis* (ATCC 8750); *Bordetella pertussis* (ATCC 9797); *Burkholderia cepacia* (ATCC 25416); *Ralstonia pickettii* (ATCC 27511); *Acidovorax facilis* (ATCC 11228); *Hydrogenophaga flava* (ATCC 33667); *Zoogloea ramigera* (ATCC 19544); *Methylobacter luteus* (ATCC 49878); *Methylocaldum gracile* (NCIMB 11912); *Methylococcus capsulatus* (ATCC 19069); *Methylomicrobium agile* (ATCC 35068); *Methylomonas methanica* (ATCC 35067); *Methylosarcina fibrata* (ATCC 700909); *Methylosphaera hansonii* (ACAM 549); *Azomonas agilis* (ATCC 7494); *Azorhizophilus paspali* (ATCC 23833); *Azotobacter chroococcum* (ATCC 9043); *Cellvibrio mixtus* (UQM 2601); *Oligella urethralis* (ATCC 17960); *Pseudomonas aeruginosa* (ATCC 10145), *Pseudomonas fluorescens* (ATCC 35858); *Francisella tularensis* (ATCC 6223); *Stenotrophomonas maltophilia* (ATCC 13637); *Xanthomonas campestris* (ATCC 33913); y *Oceanimonas doudoroffii* (ATCC 27123).

En otro ejemplo, la célula hospedadora se selecciona entre el "Subgrupo 3 de Proteobacterias gramnegativas". El "Subgrupo 3 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Brevundimonas*; *Agrobacterium*; *Rhizobium*; *Sinorhizobium*; *Blastomonas*; *Sphingomonas*; *Alcaligenes*; *Burkholderia*; *Ralstonia*; *Acidovorax*; *Hydrogenophaga*; *Methylobacter*; *Methylocaldum*; *Methylococcus*; *Methylomicrobium*; *Methylomonas*; *Methylosarcina*; *Methylosphaera*; *Azomonas*; *Azorhizophilus*; *Azotobacter*; *Cellvibrio*; *Oligella*; *Pseudomonas*; *Teredinibacter*; *Francisella*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

En otra realización, la célula hospedadora se selecciona entre el "Subgrupo 4 de Proteobacterias gramnegativas". El "Subgrupo 4 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Brevundimonas*; *Blastomonas*; *Sphingomonas*; *Burkholderia*; *Ralstonia*; *Acidovorax*; *Hydrogenophaga*; *Methylobacter*; *Methylocaldum*; *Methylococcus*; *Methylomicrobium*; *Methylomonas*; *Methylosarcina*; *Methylosphaera*; *Azomonas*; *Azorhizophilus*; *Azotobacter*; *Cellvibrio*; *Oligella*; *Pseudomonas*; *Teredinibacter*; *Francisella*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

En otro ejemplo la célula hospedadora se selecciona entre el "Subgrupo 5 de Proteobacterias gramnegativas". El "Subgrupo 5 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Methylobacter*; *Methylocaldum*; *Methylococcus*; *Methylomicrobium*; *Methylomonas*; *Methylosarcina*; *Methylosphaera*; *Azomonas*; *Azorhizophilus*; *Azotobacter*; *Cellvibrio*; *Oligella*; *Pseudomonas*; *Teredinibacter*; *Francisella*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 6 de Proteobacterias gramnegativas". El "Subgrupo 6 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Brevundimonas*; *Blastomonas*; *Sphingomonas*; *Burkholderia*; *Ralstonia*; *Acidovorax*; *Hydrogenophaga*; *Azomonas*; *Azorhizophilus*; *Azotobacter*; *Cellvibrio*; *Oligella*; *Pseudomonas*; *Teredinibacter*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 7 de Proteobacterias gramnegativas". El "Subgrupo 7 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Azomonas*; *Azorhizophilus*; *Azotobacter*; *Cellvibrio*; *Oligella*; *Pseudomonas*; *Teredinibacter*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 8 de Proteobacterias gramnegativas". El "Subgrupo 8 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Brevundimonas*; *Blastomonas*; *Sphingomonas*; *Burkholderia*; *Ralstonia*; *Acidovorax*; *Hydrogenophaga*; *Pseudomonas*; *Stenotrophomonas*; *Xanthomonas*; y *Oceanimonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 9 de Proteobacterias gramnegativas". El "Subgrupo 9 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Brevundimonas*; *Burkholderia*; *Ralstonia*; *Acidovorax*; *Hydrogenophaga*; *Pseudomonas*; *Stenotrophomonas*; y *Oceanimonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 10 de Proteobacterias gramnegativas". El "Subgrupo 10 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Burkholderia*; *Ralstonia*; *Pseudomonas*; *Stenotrophomonas*; y *Xanthomonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 11 de Proteobacterias gramnegativas". El "Subgrupo 11 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Pseudomonas*; *Stenotrophomonas*; y *Xanthomonas*. La célula hospedadora puede seleccionarse entre el "Subgrupo 12 de Proteobacterias gramnegativas". El "Subgrupo 12 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Burkholderia*; *Ralstonia*; *Pseudomonas*. La célula hospedadora puede seleccionarse entre el "Subgrupo 13 de Proteobacterias gramnegativas". El "Subgrupo 13 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Burkholderia*; *Ralstonia*; *Pseudomonas*; y *Xanthomonas*. La célula hospedadora puede seleccionarse entre el "Subgrupo 14 de Proteobacterias gramnegativas". El "Subgrupo 14 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de los siguientes géneros: *Pseudomonas* y *Xanthomonas*. La célula hospedadora puede seleccionarse entre el "Subgrupo 15 de Proteobacterias gramnegativas". El "Subgrupo 15 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias del género *Pseudomonas*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 16 de Proteobacterias gramnegativas". El "Subgrupo 16 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias de las siguientes especies de *Pseudomonas* (mostrándose el número de ATCC u otros depósitos de la cepa o cepas de ejemplo entre paréntesis): *Pseudomonas abietaniphila* (ATCC 700689); *Pseudomonas aeruginosa* (ATCC 10145); *Pseudomonas alcaligenes* (ATCC 14909); *Pseudomonas anguilliseptica* (ATCC 33660); *Pseudomonas citronellolis* (ATCC 13674); *Pseudomonas flavescens* (ATCC 51555); *Pseudomonas mendocina* (ATCC 25411); *Pseudomonas nitroreducens* (ATCC 33634); *Pseudomonas oleovorans* (ATCC 8062); *Pseudomonas pseudoalcaligenes* (ATCC 17440); *Pseudomonas resinovorans* (ATCC 14235); *Pseudomonas straminea* (ATCC 33636); *Pseudomonas agarici* (ATCC 25941); *Pseudomonas alcaliphila*; *Pseudomonas alginovora*; *Pseudomonas andersonii*; *Pseudomonas asplenii* (ATCC 23835); *Pseudomonas azelaica* (ATCC 27162); *Pseudomonas beyerinckii* (ATCC 19372); *Pseudomonas borealis*; *Pseudomonas boreopolis* (ATCC 33662); *Pseudomonas brassicacearum*; *Pseudomonas butanovora* (ATCC 43655); *Pseudomonas cellulosa* (ATCC 55703); *Pseudomonas aurantiaca* (ATCC 33663); *Pseudomonas chlororaphis* (ATCC 9446, ATCC 13985, ATCC 17418, ATCC 17461); *Pseudomonas fragi* (ATCC 4973); *Pseudomonas lundensis* (ATCC 49968); *Pseudomonas taetrolens* (ATCC 4683); *Pseudomonas cissicola* (ATCC 33616); *Pseudomonas coronafaciens*; *Pseudomonas diterpeniphila*; *Pseudomonas elongata* (ATCC 10144); *Pseudomonas flectens* (ATCC 12775); *Pseudomonas azotoformans*; *Pseudomonas brenneri*; *Pseudomonas cedrella*; *Pseudomonas corrugata* (ATCC 29736); *Pseudomonas extremorientalis*; *Pseudomonas fluorescens* (ATCC 35858); *Pseudomonas gessardii*; *Pseudomonas libanensis*; *Pseudomonas mandelii* (ATCC 700871); *Pseudomonas marginalis* (ATCC 10844); *Pseudomonas migulae*; *Pseudomonas mucidolens* (ATCC 4685); *Pseudomonas orientalis*; *Pseudomonas rhodesiae*; *Pseudomonas synxantha* (ATCC 9890); *Pseudomonas tolaasii* (ATCC 33618); *Pseudomonas veronii* (ATCC 700474); *Pseudomonas frederiksbergensis*; *Pseudomonas geniculata* (ATCC 19374); *Pseudomonas gingeri*; *Pseudomonas graminis*; *Pseudomonas grimontii*; *Pseudomonas halodenitrificans*; *Pseudomonas halophila*; *Pseudomonas hibiscicola* (ATCC 19867); *Pseudomonas huttienensis* (ATCC 14670); *Pseudomonas hydrogenovora*; *Pseudomonas jessenii* (ATCC 700870); *Pseudomonas kilonensis*; *Pseudomonas lanceolata* (ATCC 14669); *Pseudomonas lini*; *Pseudomonas marginata* (ATCC 25417); *Pseudomonas mephitica* (ATCC 33665); *Pseudomonas denitrificans* (ATCC 19244); *Pseudomonas pertucinogena* (ATCC 190); *Pseudomonas pictorum* (ATCC 23328); *Pseudomonas psychrophila*; *Pseudomonas filva* (ATCC 31418); *Pseudomonas monteillii* (ATCC 700476); *Pseudomonas mosselii*; *Pseudomonas oryzihabitans* (ATCC 43272); *Pseudomonas plecoglossicida* (ATCC 700383); *Pseudomonas putida* (ATCC 12633); *Pseudomonas reactans*; *Pseudomonas spinosa* (ATCC 14606); *Pseudomonas balearica*; *Pseudomonas luteola* (ATCC 43273); *Pseudomonas stutzeri* (ATCC 17588); *Pseudomonas amygdali* (ATCC 33614); *Pseudomonas avellanae* (ATCC 700331); *Pseudomonas caricapapayae* (ATCC 33615); *Pseudomonas cichorii* (ATCC 10857); *Pseudomonas ficuserectae* (ATCC 35104); *Pseudomonas fuscovaginae*; *Pseudomonas meliae* (ATCC 33050); *Pseudomonas syringae* (ATCC 19310); *Pseudomonas viridiflava* (ATCC 13223); *Pseudomonas thermocarboxydovorans* (ATCC 35961); *Pseudomonas thermotolerans*; *Pseudomonas thivervalensis*; *Pseudomonas vancouverensis* (ATCC 700688); *Pseudomonas wisconsinensis*; y *Pseudomonas xiamenensis*.

La célula hospedadora puede seleccionarse entre el "Subgrupo 17 de Proteobacterias gramnegativas". El "Subgrupo 17 de Proteobacterias gramnegativas" se define como el grupo de Proteobacterias conocido en la técnica como las

"Pseudomonas fluorescentes", incluyendo las que pertenecen, por ejemplo, a las siguientes especies de *Pseudomonas*: *Pseudomonas azotoformans*; *Pseudomonas brenneri*; *Pseudomonas cedrella*; *Pseudomonas corrugata*; *Pseudomonas extremorientalis*; *Pseudomonas fluorescens*; *Pseudomonas gessardii*; *Pseudomonas libanensis*; *Pseudomonas mandelii*; *Pseudomonas marginalis*; *Pseudomonas migulae*; *Pseudomonas mucidolens*; *Pseudomonas orientalis*; *Pseudomonas rhodesiae*; *Pseudomonas synxantha*; *Pseudomonas tolaasii*; y *Pseudomonas veronii*.

Expresión de polipéptidos recombinantes en una célula hospedadora auxotrófica

En un aspecto de la presente invención, se proporcionan procesos de expresión de polipéptidos recombinantes para su uso en la producción mejorada de proteínas. De acuerdo con un aspecto de la presente invención, se proporciona un método para producir un polipéptido recombinante de interés. El proceso incluye la obtención de una población de células auxotróficas para un primer metabolito y un segundo metabolito en la que el segundo metabolito es un aminoácido natural. Además, el método incluye poner en contacto la población de células con una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica, en el que el marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito y en el que la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito. La población de células se pone en contacto con una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés; y un promotor unido operativamente a la segunda secuencia de ácido nucleico para dirigir la expresión de la segunda secuencia de ácido nucleico. La población de células se somete a un primer medio que carece del primer metabolito en condiciones de manera que se obtengan células transfectadas que tengan la prototrofia restablecida para el primer metabolito. Las células transfectadas se someten a un segundo medio que comprende un aminoácido no natural correlacionado con el segundo metabolito en condiciones de manera que la segunda secuencia de ácido nucleico se exprese para producir el polipéptido recombinante de interés que tiene el aminoácido no natural incorporado en el mismo. En el método, el primer medio y el segundo medio pueden ser iguales o diferentes. En otras palabras, el primer medio puede contener el aminoácido no natural y en tal caso, puede servir como el segundo medio. Como alternativa, el primer medio carece del aminoácido no natural. Una vez que se obtienen las células transfectadas, después estas células se someten a un segundo medio que contiene un aminoácido no natural.

La presente memoria descriptiva describe un método para producir un polipéptido recombinante de interés. El método comprende generalmente: introducir en una célula hospedadora que es auxotrófica para un primer metabolito una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica. El marcador de selección auxotrófica comprende una primera secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito, y la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito a la célula hospedadora auxotrófica. El método incluye adicionalmente introducir en la célula hospedadora: (i) una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés; (ii) una tercera secuencia de ácido nucleico que codifica una ARNt sintetasa ortogonal, (iii) una cuarta secuencia de ácido nucleico que codifica un ARNt ortogonal; y (iv) un promotor unido operativamente a las secuencias de ácido nucleico segunda, tercera y/o cuarta con el fin de dirigir la expresión de las secuencias de ácido nucleico segunda, tercera y/o cuarta en la célula hospedadora auxotrófica. La célula hospedadora auxotrófica se somete a un medio que carece del primer metabolito para seleccionar las células transfectadas. El medio puede comprender o ser suplantado por un segundo medio que comprende un aminoácido no natural en condiciones en las que el ARNt ortogonal expresado que puede interactuar con la ARNt sintetasa ortogonal expresada facilita la incorporación del aminoácido no natural en el polipéptido recombinante de interés durante la traducción.

La presente memoria descriptiva adicionalmente un método para producir un polipéptido recombinante de interés. La realización del método comprende en general a) introducir en una célula hospedadora que es auxotrófica para un primer metabolito necesario para la supervivencia de la célula hospedadora: (i) una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica, en la que el marcador de selección auxotrófica comprende una primera secuencia de nucleico ácido que codifica al menos un polipéptido activo en la biosíntesis del primer metabolito y en la que la expresión del marcador de selección auxotrófica restablece la prototrofia para el primer metabolito a la célula hospedadora auxotrófica; (ii) una segunda construcción de ácido nucleico que comprende una segunda secuencia de ácido nucleico que codifica el polipéptido recombinante de interés; una tercera secuencia de ácido nucleico que codifica una ARNt sintetasa ortogonal, una cuarta secuencia de ácido nucleico que codifica un ARNt ortogonal que puede interactuar con la ARNt sintetasa; y (iv) un promotor unido operativamente a las secuencias de ácido nucleico primera, segunda, tercera y/o cuarta con el fin de dirigir la expresión de la primera secuencia de ácido nucleico, la segunda secuencia de ácido nucleico, la tercera secuencia de ácido nucleico y/o la cuarta secuencia de ácido nucleico en la célula hospedadora auxotrófica. Además, el método comprende b) someter la célula hospedadora auxotrófica a un medio que carece del primer metabolito y/o comprende un aminoácido no natural en condiciones de manera que el segundo ácido nucleico se exprese para producir el polipéptido recombinante que tenga el aminoácido no natural incorporado en el mismo.

Adicionalmente, las secuencias de ácido nucleico primera y segunda están contenidas en la misma construcción.

Preferentemente, el sistema de expresión es capaz de expresar el polipéptido diana a una productividad total de polipéptido de al menos 1 g/l a al menos 80 g/l. En una realización particular, el polipéptido recombinante se expresa a un nivel de al menos 3 g/l, 4 g/l, 5 g/l, 6 g/l, 7 g/l, 8 g/l, 9 g/l, 10 g/l, 12 g/l, 15 g/l, 20 g/l, 25 g/l, 30 g/l, 35 g/l, 40 g/l, 45 g/l, 50 g/l, 60 g/l, 70 g/l o al menos 80 g/l.

5 En una realización, al menos un polipéptido recombinante puede expresarse en una célula que es auxotrófica para un metabolito, en la que la auxotrofia sirve como un marcador de selección para el mantenimiento del vector de expresión de ácido nucleico que codifica el polipéptido de interés y la enzima que permite la prototrofia. Como alternativa, puede expresarse más de un polipéptido recombinante en una célula que es auxotrófica para un
10 metabolito, en la que los ácidos nucleicos que codifican los polipéptidos recombinantes pueden estar contenidos en el mismo vector o, como alternativa, en múltiples vectores.

15 En otra realización más, más de un vector de expresión que codifican diferentes polipéptidos diana pueden mantenerse en una célula hospedadora auxotrófica para al menos un metabolito, en la que un vector de expresión contiene un ácido nucleico que codifica una enzima que permite la prototrofia y un primer polipéptido diana de interés, y un segundo vector de expresión que contiene un ácido nucleico que codifica un marcador de selección no auxotrófico alternativo y un segundo polipéptido de interés.

20 En otra realización, puede expresarse al menos un polipéptido recombinante en una célula que es auxotrófica para más de un metabolito, en la que las múltiples auxotrofias sirven como marcadores de selección para el mantenimiento de vectores de expresión de ácidos nucleicos. Por ejemplo, puede utilizarse un vector de expresión en el que están presentes un primer y un segundo gen marcador de selección que permite la prototrofia. Si ambos genes marcadores están situados en la misma construcción de ADN, la célula hospedadora que contiene la construcción puede mantenerse en una o ambas de las condiciones en las que la supervivencia de la célula
25 hospedadora necesita la presencia del gen o genes marcadores de selección. Cuando solo está presente una de las condiciones de supervivencia dependientes del gen marcador, debe expresarse el gen marcador correspondiente y el otro gen o genes marcadores puede entonces estar activo o inactivo, aunque todos los nutrientes necesarios para los cuales la célula permanece auxotrófica aún serán suministrados por el medio. Esto permite que el mismo gen diana o el mismo conjunto de genes diana unidos covalentemente, que codifican el producto o productos
30 transgénicos deseados y/o la actividad o actividades transgénicas deseadas, se mantengan en la célula hospedadora de forma continua a medida que la célula hospedadora se somete a la transición entre o entre diferentes condiciones. Si cada uno de los dos genes marcadores de selección se encuentra en una construcción de ADN diferente, entonces, con el fin de mantener tanto de las construcciones de ADN en la célula hospedadora, ambas condiciones de supervivencia dependientes del gen marcador están presentes y ambos de los genes
35 marcadores correspondientes deben expresarse. Esto permite que más de un gen diana unido no covalentemente o conjunto de genes objetivo se mantengan separados en la célula hospedadora. La secuencia codificante de cada uno de los genes marcadores de selección elegidos puede estar unido operativamente independientemente a un promotor ya sea constitutivo o regulado.

40 Los ejemplos de genes diana duales de un sistema de genes diana múltiples de este tipo incluyen, pero no se limitan a: (1) sistemas en los que el producto de expresión de uno de los genes diana interactúa con el otro gen diana por sí mismo; (2) sistemas en los que el producto de expresión de uno de los genes diana interactúa con el producto de expresión de otro gen diana, por ejemplo, una proteína y su proteína de unión o los polipéptidos α y β de una proteína α y β ; (3) sistemas en los que los dos productos de expresión de los dos genes interactúan ambos con un
45 tercer componente, por ejemplo, un tercer componente presente en la célula hospedadora; (4) sistemas en los que los dos productos de expresión de los dos genes participan ambos en una vía biocatalítica común; y (5) sistemas en los que los dos productos de expresión de los dos genes funcionan independientemente entre sí, por ejemplo, un sistema de expresión de anticuerpo biclonal.

50 En un ejemplo de un sistema de gen diana dual del tipo enumerado anteriormente (1), un primer gen diana puede codificar una proteína diana deseada, en la que el primer gen diana está bajo el control de un promotor regulado; el segundo gen diana puede entonces codificar una proteína implicada en la regulación del promotor del primer gen diana, por ejemplo, el segundo gen diana puede codificar la proteína activadora o represora del promotor del primer gen diana. En un ejemplo en el que el segundo gen codifica una proteína reguladora del promotor para el primer
55 gen, la secuencia codificante del segundo gen puede estar bajo el control de un promotor constitutivo. En una realización, el segundo gen será parte de una construcción de ADN separada que está mantenida en la célula como una construcción de un número de copias alto con un número de copias de al menos 10, 20, 30, 40, 50 o más de 50 copias que se mantienen en la célula hospedadora.

60 Transformación

La transformación de las células hospedadoras auxotróficas con el vector o vectores puede realizarse usando cualquier metodología de transformación conocida en la técnica y las células hospedadoras bacterianas pueden transformarse como células intactas o como protoplastos (es decir, incluyendo los citoplastos). Las metodologías de
65 transformación de ejemplo incluyen metodologías de poración, por ejemplo, electroporación, fusión de protoplastos, conjugación bacteriana y tratamiento con catión divalente, por ejemplo, tratamiento con cloruro de calcio o

tratamiento con CaCl/Mg^{2+} , u otros métodos bien conocidos en la técnica. Véase, por ejemplo, Morrison, *J. Bact.*, 132:349-351 (1977); Clark-Curtiss y Curtiss, *Methods in Enzymology*, 101:347-362 (Wu et al., Ed, 1983), Sambrook et al., *Molecular Cloning, A Laboratory Manual* (2ª ed. 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y *Current Protocols in Molecular Biology* (Ausubel et al., ed., 1994)).

5

Selección

Preferentemente, las células que no se transforman satisfactoriamente se seleccionan en contra después de la transformación y de forma continua durante la fermentación. El marcador de selección puede ser un marcador de selección auxotrófica y/o un marcador de selección con antibióticos tradicional. En una realización específica, el marcador de selección es un marcador auxotrófico. Cuando la célula es auxotrófica para múltiples compuestos de nutrientes, la célula auxotrófica puede cultivarse en medio complementado con todos aquellos compuestos de nutrientes hasta que se transforme con el vector de restablecimiento de la prototofia. Cuando la célula hospedadora es o se ha hecho defectuosa para múltiples actividades biosintéticas, el sistema o sistemas marcadores restablecedores de la prototofia pueden seleccionarse para restablecer una o más o todas las actividades biosintéticas, compensándose el resto al continuar proporcionando, en el medio, los nutrientes que aún faltan. En los sistemas de marcadores de selección en los que más de una actividad biosintética, y/o más de una prototofia, se restablecen, la pluralidad de genes marcadores de selección pueden expresarse juntos en un vector o pueden coexpresarse por separado en diferentes vectores. Incluso cuando un solo metabolito es la diana del sistema marcador de selección, pueden estar implicadas múltiples actividades biosintéticas en el sistema marcador de selección. Por ejemplo, dos o más genes que codifican las actividades de la misma vía anabólica pueden expresarse juntos en un vector o pueden coexpresarse por separado en diferentes vectores, con el fin de restablecer la prototofia en lo que se refiere a la biosíntesis del compuesto que es el producto de la vía.

Cuando el marcador de selección es un gen de resistencia a antibiótico, el antibiótico asociado puede añadirse al medio para seleccionar contra células no transformadas y revertientes, como son bien conocidas en la técnica.

Fermentación

En la presente memoria, el término "fermentación" incluye ambas realizaciones en las que se emplea la fermentación literal y realizaciones en las que se emplean otros modos de cultivo, no fermentativos. La fermentación puede realizarse a cualquier escala. En una realización, el medio de fermentación puede seleccionarse entre medios ricos, medio mínimo, un medio de sales minerales; puede usarse un medio rico, pero se evita preferentemente. En otra realización se selecciona ya sea un medio mínimo o un medio de sales minerales. En otra realización más, se selecciona un medio mínimo. En otra realización más, se selecciona un medio de sales minerales. Se prefieren en particular los medios de sales minerales.

Antes de la transformación de la célula hospedadora con una construcción de ácido nucleico que codifica una enzima que permite la prototofia, la célula hospedadora puede mantenerse en un medio que comprende un metabolito complementario o análogo del mismo, lo que complementa la auxotofia. Después de la transformación, la célula hospedadora puede cultivarse en un medio que carece del metabolito complementario para el que la célula hospedadora es auxotrófica. De esta manera, las células hospedadoras que no contienen el marcador de selección que permite la prototofia se seleccionan en contra. Los medios de sales minerales consisten en sales minerales y una fuente de carbono tal como, por ejemplo, glucosa, sacarosa o glicerol. Los ejemplos de medios de sales minerales incluyen, por ejemplo, medio M9, medio *Pseudomonas* (ATCC 179), medio Davis y Mingioli (véase, B D Davis y E S Mingioli, en *J. Bact.* 60:17-28 (1950)). Las sales minerales utilizadas para hacer medios de sales minerales incluyen las seleccionadas entre, por ejemplo, fosfatos de potasio, sulfato o cloruro de amonio, sulfato o cloruro de magnesio y minerales traza, tales como cloruro de calcio, borato y sulfatos de hierro, cobre, manganeso y cinc. No se incluye ninguna fuente de nitrógeno orgánico, tal como peptona, triptona, aminoácidos o un extracto de levadura, en un medio de sales minerales. En su lugar, se utiliza una fuente de nitrógeno inorgánico y éste puede seleccionarse de entre, por ejemplo, sales de amonio, amoníaco acuoso y amoníaco gaseoso. Un medio de sales minerales particular, contendrá glucosa como fuente de carbono. En comparación con los medios de sales minerales, los medios mínimos también pueden contener sales minerales y una fuente de carbono, pero pueden complementarse con, por ejemplo, niveles bajos de aminoácidos, vitaminas, peptonas u otros ingredientes, aunque estos se añaden a niveles muy mínimos.

En una realización, los medios pueden prepararse usando los componentes que se enumeran a continuación. Los componentes pueden añadirse en el siguiente orden: primero puede disolverse $(\text{NH}_4)\text{HPO}_4$, KH_2PO_4 y ácido cítrico en aproximadamente 30 litros de agua destilada; después puede añadirse una solución de elementos traza, seguido de la adición de un agente antiespumante, tal como Ucolub N 115. A continuación, después de la esterilización por calor (tal como aproximadamente a 121 grados C), pueden añadirse soluciones estériles de glucosa MgSO_4 y tiamina-HCl. El control del pH a aproximadamente 6,8 puede conseguirse usando amoníaco acuoso. Después puede añadirse agua destilada estéril para ajustar el volumen inicial a 371 menos la solución madre de glicerol (123 ml).

Los productos químicos están disponibles en el mercado de diversos proveedores, tales como Merck. Este medio puede permitir un cultivo de alta densidad celular (HCDC, por sus siglas en inglés) para el crecimiento de especies

de *Pseudomonas* y bacterias relacionadas. El HCDC puede comenzar como un proceso discontinuo que va seguido del cultivo discontinuo alimentado en dos fases. Después de un crecimiento ilimitado en la parte discontinua del proceso, el crecimiento puede controlarse a una tasa de crecimiento específica reducida durante un periodo de 3 tiempos de duplicación en el que la concentración de biomasa puede aumentar varias veces. Se describen detalles adicionales de dichos procedimientos de cultivo por Riesenberg, D.; Schulz, V.; Knorre, W. A.; Pohl, H. D.; Korz, D.; Sanders, E. A.; Ross, A.; Deckwer, W. D. (1991) "*High cell density cultivation of Escherichia coli at controlled specific growth rate*" *J Biotechnol*: 20(1) 17-27.

El sistema de expresión de acuerdo con la presente invención puede cultivarse en cualquier formato de fermentación. Por ejemplo, puede emplearse modos de fermentación discontinua, discontinua alimentada, semi-continua y continua en el presente documento.

Los sistemas de expresión de acuerdo con la presente invención son útiles para la expresión del transgén a cualquier escala (es decir, volumen) de fermentación. De este modo, por ejemplo, pueden usarse volúmenes de fermentación a escala de microlitros, escala de centilitros y escala de decilitros; y pueden usarse la escala de 1 litro y volúmenes de fermentación más grandes. En una realización, el volumen de fermentación será igual o superior a 1 litro. En otra realización, el volumen de fermentación será igual o superior a 5 litros, 10 litros, 15 litros, 20 litros, 25 litros, 50 litros, 75 litros, 100 litros, 200 litros, 500 litros, 1.000 litros, 2.000 litros, 5.000 litros, 10.000 litros o 50.000 litros.

En la presente invención, el crecimiento, el cultivo y/o la fermentación de las células hospedadoras transformadas se realiza dentro de un intervalo de temperatura que permite la supervivencia de las células hospedadoras, preferentemente una temperatura dentro del intervalo de aproximadamente 40 °C a aproximadamente 55 °C, inclusive.

En diversas realizaciones, las células hospedadoras auxotróficas se cultivan en altas densidades de células, como se indica en términos de biomasa por volumen, midiéndose la biomasa como el peso de célula seca.

En una realización, la densidad celular será de al menos 20 g/l. En otra realización, la densidad celular será de al menos 25 g/l, 30 g/l, 35 g/l, 40 g/l, 45 g/l, 50 g/l, 60 g/l, 70 g/l, 80 g/l, 90 g/l, 100 g/l, 110 g/l, 120 g/l, 130 g/l, 140 g/l o al menos 150 g/l.

En otras realizaciones, la densidad celular en la inducción será de entre 20 g/l y 150 g/l; 20 g/l y 120 g/l; 20 g/l y 80 g/l; 25 g/l y 80 g/l; 30 g/l y 80 g/l; 35 g/l y 80 g/l; 40 g/l y 80 g/l; 45 g/l y 80 g/l; 50 g/l y 80 g/l; 50 g/l y 75 g/l; 50 g/l y 70 g/l; 40 g/l y 80 g/l.

Expresión mejorada de la proteína recombinante

El método de la invención puede conducir a un aumento de la producción del polipéptido recombinante de interés dentro de la célula hospedadora auxotrófica. El aumento de la producción, como alternativa, puede ser un aumento del nivel de proteína correctamente procesada o de polipéptido por gramo de proteína producida, o por gramo de proteína hospedadora. El aumento de la producción también puede ser un aumento del nivel de proteína recuperable o polipéptido producido por gramo de recombinante o por gramo de proteína de la célula hospedadora. El aumento de la producción también puede ser cualquier combinación de un mayor nivel de proteína total, el aumento del nivel de proteína correctamente procesada o el aumento del nivel de proteína activa o soluble. En esta realización, el término "aumento" se refiere al nivel de la proteína o polipéptido que se produce, adecuadamente procesada, soluble y/o recuperable cuando la proteína o polipéptido de interés se expresa en un sistema de expresión que utiliza marcadores de selección con antibióticos o en un sistema de selección con antibióticos o auxotrófica en ausencia de un análogo del metabolito para el que cual célula es auxotrófica.

Una expresión mejorada de una proteína o polipéptido de interés también puede referirse a un aumento en la solubilidad de la proteína. La proteína o polipéptido de interés puede producirse y recuperarse del citoplasma, el periplasma o el medio extracelular de la célula hospedadora. La proteína o polipéptido puede ser insoluble o soluble.

La proteína o polipéptido puede incluir una o más secuencias de direccionamiento o secuencias para ayudar a la purificación, como se conocen en la técnica.

El término "soluble" como se usa en el presente documento significa que la proteína no se precipitó por centrifugación a entre aproximadamente 5.000 y 20.000 x gravedad cuando se centrifugó durante 10-30 minutos en un tampón en condiciones fisiológicas. Las proteínas solubles no son parte de un cuerpo de inclusión u otra masa precipitada. De forma similar, "insoluble" significa que la proteína o polipéptido puede precipitarse por centrifugación a entre 5.000 y 20.000 x gravedad cuando se centrifuga durante 10-30 minutos en un tampón en condiciones fisiológicas. Las proteínas o polipéptidos insolubles pueden ser parte de un cuerpo de inclusión u otra masa precipitada. La expresión "cuerpos de inclusión" tiene por objeto incluir cualquier cuerpo intracelular contenido dentro de una célula, en el que se ha secuestrado un agregado de proteínas o polipéptidos.

El método de la invención puede producir proteína localizada en el periplasma de la célula hospedadora. En una realización, el método produce proteínas o polipéptidos de interés en la célula procesada correctamente. En otra realización, la expresión del polipéptido recombinante puede producir proteínas activas o polipéptidos de interés en la célula.

5 En una realización, el método produce al menos 0,1 g/l de proteína procesada correctamente. Una proteína correctamente procesada tiene un extremo amino terminal de la proteína nativa. En algunas realizaciones, al menos el 50 % de la proteína o polipéptido de interés comprende un extremo amino terminal nativo. En otra realización, al menos el 60 %, al menos el 70 %, al menos el 80 %, al menos el 90 % o más de la proteína tiene un extremo amino terminal de la proteína nativa. En diversas realizaciones, el método produce de 0,1 a 10 g/l de proteína procesada correctamente en la célula, incluyendo al menos aproximadamente 0,2, aproximadamente 0,3, aproximadamente 0,4, aproximadamente 0,5, aproximadamente 0,6, aproximadamente 0,7, aproximadamente 0,8, aproximadamente 0,9 o al menos aproximadamente 1,0 g/l de proteína procesada correctamente. En otra realización, el total de proteína procesada correctamente o polipéptido de interés producido es de al menos 1,0 g/l, al menos aproximadamente 2 g/l, al menos aproximadamente 3 g/l, aproximadamente 4 g/l, aproximadamente 5 g/l, aproximadamente 6 g/l, aproximadamente 7 g/l, aproximadamente 8 g/l, aproximadamente 10 g/l, aproximadamente 15 g/l, aproximadamente 20 g/l, aproximadamente 25 g/l, aproximadamente 30 g/l, aproximadamente 35 g/l, aproximadamente 40 g/l, aproximadamente 45 g/l, al menos aproximadamente 50 g/l o más. En algunas realizaciones, la cantidad de proteína correctamente procesada producida es de al menos aproximadamente el 5 %, aproximadamente el 10 %, aproximadamente el 15 %, aproximadamente el 20 %, aproximadamente el 25 %, aproximadamente el 30 %, aproximadamente el 40 %, aproximadamente el 50 %, aproximadamente el 60 %, aproximadamente el 70 %, aproximadamente el 80 %, aproximadamente el 90 %, aproximadamente el 95 %, aproximadamente el 96 %, aproximadamente el 97 %, aproximadamente el 98 %, al menos aproximadamente el 99 % o más de proteína recombinante total en una forma correctamente procesada.

25 El método de la invención también puede conducir a un aumento del rendimiento de la proteína o polipéptido de interés. En una realización, el método produce una proteína o polipéptido de interés en al menos aproximadamente el 5 %, al menos aproximadamente el 10 %, aproximadamente el 15 %, aproximadamente el 20 %, aproximadamente el 25 %, aproximadamente el 30 %, aproximadamente el 30 %, aproximadamente el 40 %, aproximadamente el 45 %, aproximadamente el 50 %, aproximadamente el 55 %, aproximadamente el 60 %, aproximadamente el 65 %, aproximadamente el 70 %, aproximadamente el 75 % o más de la proteína celular total (tcp, por sus siglas en inglés). "Porcentaje de proteína celular total" es la cantidad de proteína o polipéptido en la célula hospedadora como un porcentaje de la proteína celular total. La determinación del porcentaje de proteína celular total es bien conocida en la técnica.

35 En una realización particular, la célula hospedadora puede tener un nivel de expresión de polipéptido recombinante, polipéptido, proteína o fragmento de los mismos de al menos el 1 % de tcp y una densidad celular de al menos 40 g/l, cuando se cultiva (es decir, dentro de un intervalo de temperatura de aproximadamente 4 °C a aproximadamente 55 °C, incluyendo aproximadamente 10 °C, aproximadamente 15 °C, aproximadamente 20 °C, aproximadamente 25 °C, aproximadamente 30 °C, aproximadamente 35 °C, aproximadamente 40 °C, aproximadamente 45 °C y aproximadamente 50 °C) en un medio de sales minerales. En una realización particularmente preferida, el sistema de expresión tendrá un nivel de expresión de la proteína o polipéptido de al menos el 5 % tcp y una densidad celular de al menos 40 g/l, cuando se cultiva (es decir, dentro de un intervalo de temperatura de aproximadamente 4 °C a aproximadamente 55 °C, inclusive) en un medio de sales minerales en una escala de fermentación de al menos aproximadamente 10 litros.

50 En algunas realizaciones, la proteína también puede producirse en forma activa. El término "activo" significa la presencia de actividad biológica, en la que la actividad biológica es comparable o se corresponde sustancialmente con la actividad biológica de una proteína o polipéptido nativo correspondiente. En el contexto de proteínas esto normalmente significa que un polinucleótido o polipéptido comprende una función biológica o efecto que tiene al menos aproximadamente el 20 %, aproximadamente el 50 %, preferentemente al menos aproximadamente el 60-80 % y más preferentemente al menos aproximadamente el 90-95 % de la actividad en comparación con la proteína o polipéptido nativo correspondiente usando parámetros convencionales. La determinación de la actividad de la proteína o polipéptido puede realizarse utilizando ensayos biológicos comparativos dirigidos, convencionales correspondientes, para proteínas o polipéptidos particulares. Una indicación de que una proteína o polipéptido de interés mantiene la actividad biológica es que el polipéptido sea inmunológicamente reactivo de forma cruzada con el polipéptido nativo.

60 La invención también puede mejorar la recuperación de proteína o polipéptido activo de interés. Las proteínas activas pueden tener una actividad específica de al menos aproximadamente el 20 %, al menos aproximadamente el 30 %, al menos aproximadamente el 40 %, aproximadamente el 50 %, aproximadamente el 60 %, al menos aproximadamente el 70 %, aproximadamente el 80 %, aproximadamente el 90 % o al menos aproximadamente el 95 % de la de la proteína o polipéptido nativo del que deriva la secuencia. Adicionalmente, la especificidad de sustrato (kcat/Km) es opcionalmente sustancialmente similar a la de la proteína o polipéptido nativo. Normalmente, kcat/Km será de al menos aproximadamente el 30 %, aproximadamente el 40 %, aproximadamente el 50 %, aproximadamente el 60 %, aproximadamente el 70 %, aproximadamente el 80 %, al menos aproximadamente el

90 %, al menos aproximadamente el 95 % o más. Los métodos de ensayo y cuantificación de las mediciones de la actividad de la proteína y el polipéptido y de la especificidad de sustrato (kcat/Km), son bien conocidos por los expertos en la materia.

5 Aislamiento y purificación

Las proteínas recombinantes producidas de acuerdo con la presente invención pueden aislarse y purificarse a una pureza sustancial mediante técnicas convencionales bien conocidas en la técnica, incluyendo, pero no limitadas a, precipitación con sulfato de amonio o etanol, extracción con ácido, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de níquel, cromatografía de hidroxilapatita, cromatografía de fase inversa, cromatografía de lectina, electroforesis preparativa, solubilización con detergente, precipitación selectiva con sustancias tales como cromatografía en columna, métodos de inmunopurificación y otros. Por ejemplo, las proteínas que tienen propiedades de adhesión molecular establecidas pueden fusionarse reversiblemente con un ligando. Con el ligando apropiado, la proteína puede adsorberse selectivamente en una columna de purificación y, después, liberarse de la columna de una forma relativamente pura. La proteína fusionada se retira después por la actividad enzimática. Además, la proteína puede purificarse usando columnas de inmutafinidad o columnas Ni-NTA. Se describen adicionalmente técnicas generales en, por ejemplo R. Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag: N.Y. (1982); Deutscher, *Guide to Protein Purification*, Academic Press (1990); Patente de los EE.UU. N.º 4.511.503; S. Roe, *Protein Purification Techniques: A Practical Approach (Practical Approach Series)*, Oxford Press (2001); D. Bollag, et al., *Protein Methods*, Wiley-Lisa, Inc. (1996); A K Patra et al., *Protein Expr Purif*, 18(2): p/ 182-92 (2000); y R. Mukhija, et al., *Gene* 165(2): p. 303-6 (1995). Véase también, por ejemplo, Ausubel, et al. (1987 y suplementos periódicos); Deutscher (1990) "*Guide to Protein Purification*", *Methods in Enzymology* vol. 182 y otros volúmenes de esta serie; Coligan, et al. (1996 y suplementos periódicos) *Current Protocols in Protein Science* Wiley/Greene, NY; y bibliografía del fabricante acerca del uso de productos de purificación de proteínas, por ejemplo, Pharmacia, Piscataway, N.J. o Bio-Rad, Richmond, Calif. La combinación con técnicas recombinantes permite la fusión a segmentos apropiados, por ejemplo, a una secuencia Flag o un equivalente que puede fusionarse a través de una secuencia extraíble por proteasa. Véase también, por ejemplo, Hochuli (1989) *Chemische Industrie* 12: 69-70; Hochuli (1990) "*Purification of Recombinant Proteins with Metal Chelate Absorbent*" in Setlow (ed.) *Genetic Engineering, Principle and Methods* 12:87-98, Plenum Press, NY; y Crowe, et al. (1992) *QIAexpress: The High Level Expression & Protein Purification System* QIAGEN, Inc., Chatsworth, Calif.

La detección de la proteína expresada se consigue por métodos conocidos en la técnica e incluye, por ejemplo, radioinmunoensayos, técnicas de transferencia Western o inmunoprecipitación.

La enzima producida y expresada de forma recombinante puede recuperarse y purificarse a partir de los cultivos de células recombinantes mediante numerosos métodos, por ejemplo, puede emplearse cromatografía líquida de alto rendimiento (HPLC) para las etapas de purificación finales, según sea necesario.

Ciertas proteínas expresadas en la presente invención pueden formar agregados insolubles ("cuerpos de inclusión"). Varios protocolos son adecuados para la purificación de proteínas a partir de cuerpos de inclusión. Por ejemplo, la purificación de cuerpos de inclusión normalmente implica la extracción, separación y/o purificación de los cuerpos de inclusión mediante la ruptura de las células hospedadoras, por ejemplo, por incubación en un tampón de TRIS 50 mM/HCl pH 7,5, NaCl 50 mM, MgCl₂ 5 mM, DTT 1 mM, ATP 0,1 mM y PMSF 1 mM. La suspensión celular normalmente se lisa con 2-3 pases a través de una prensa francesa. La suspensión de células también puede homogeneizarse usando un Polytron (Brinknan Instruments) o sometiéndola a ultrasonidos en hielo. Los métodos alternativos de lisado de bacterias son evidentes para los expertos en la materia (véase, por ejemplo, Sambrook et al., citado anteriormente; Ausubel et al., citado anteriormente).

Si es necesario, los cuerpos de inclusión pueden solubilizarse y la suspensión de células lisadas normalmente puede centrifugarse para retirar la materia insoluble no deseada. Las proteínas que formaban los cuerpos de inclusión pueden renaturalizarse por dilución o diálisis con un tampón compatible. Los disolventes adecuados incluyen, pero no se limitan a urea (de aproximadamente 4 M a aproximadamente 8 M), formamida (al menos aproximadamente al 80 %, en base volumen/volumen) y clorhidrato de guanidina (de aproximadamente 4 M a aproximadamente 8 M). Aunque el clorhidrato de guanidina y agentes similares son desnaturizantes, esta desnaturización no es irreversible y la renaturalización puede producirse después de la retirada (por diálisis, por ejemplo) o dilución del desnaturizante, permitiendo la re-formación de la proteína inmunológicamente y/o biológicamente activa. Otros tampones adecuados son conocidos por los expertos en la materia.

Como alternativa, es posible purificar las proteínas o péptidos recombinantes del periplasma del hospedador.

Después de la lisis de la célula hospedadora, cuando la proteína recombinante se exporta al periplasma de la célula hospedadora, la fracción periplásmica de las bacterias puede aislarse por choque osmótico en frío además de otros métodos conocidos por los expertos en la materia. Para aislar proteínas recombinantes del periplasma, por ejemplo, las células bacterianas pueden centrifugarse para formar un sedimento. El sedimento puede resuspenderse en un tampón que contenga sacarosa al 20 %. Para lisar las células, las bacterias pueden centrifugarse y el sedimento

puede resuspenderse en MgSO₄ 5°mM enfriado con hielo y mantenerse en un baño de hielo durante aproximadamente 10 minutos. La suspensión celular puede centrifugarse y el sobrenadante se decanta y se guarda.

5 Las proteínas recombinantes presentes en el sobrenadante pueden separarse de las proteínas del hospedador mediante técnicas de separación convencionales bien conocidas por los expertos en la materia.

10 Un fraccionamiento salino inicial puede separar muchas de las proteínas no deseadas de la célula hospedadora (o proteínas derivadas del medio de cultivo celular) de la proteína recombinante de interés. Un ejemplo de este tipo puede ser el sulfato de amonio. El sulfato de amonio precipita proteínas mediante la reducción eficaz de la cantidad de agua en la mezcla de proteínas. Después, las proteínas precipitan basándose en su solubilidad. Cuanto más hidrófoba sea una proteína, más probable es que precipite a concentraciones de sulfato de amonio inferiores. Un protocolo frecuente incluye añadir sulfato de amonio saturado a una solución de proteínas de modo que la concentración de sulfato de amonio resultante sea de entre el 20-30 %. Esta concentración precipitará la más hidrófoba de las proteínas. Después, se desecha el precipitado (a menos que la proteína de interés sea hidrófoba) y se añade sulfato de amonio al sobrenadante hasta una concentración conocida para precipitar la proteína de interés.

15 Después, el precipitado se solubiliza en tampón y el exceso de sal se retira si es necesario, ya sea a través de diálisis o diafiltración. Otros métodos que se basan en la solubilidad de las proteínas, tales como la precipitación con etanol frío, son bien conocidos por los expertos en la materia y pueden usarse para fraccionar mezclas complejas de proteínas.

20 El peso molecular de una proteína recombinante puede usarse para aislarla de proteínas de mayor y menor tamaño usando ultrafiltración a través de membranas de diferente tamaño de poro (por ejemplo, membranas Amicon o Millipore). Como primera etapa, la mezcla de proteínas puede ultrafiltrarse a través de una membrana con un tamaño de poro que tiene un peso molecular de corte menor que el peso molecular de la proteína de interés. El retenido de la ultrafiltración después puede ultrafiltrarse contra una membrana con un peso molecular de corte mayor que el peso molecular de la proteína de interés. La proteína recombinante pasará a través de la membrana hacia el filtrado. Después, el filtrado puede someterse a cromatografía como se describe a continuación.

25 30 Las proteínas recombinantes también pueden separarse de otras proteínas basándose en su tamaño, carga superficial neta, hidrofobia y afinidad por ligandos. Además, pueden conjugarse anticuerpos generados contra las proteínas con matrices de columna y las proteínas pueden inmunopurificarse. Todos estos métodos son bien conocidos en la técnica. Será evidente para un experto que las técnicas cromatográficas pueden realizarse a cualquier escala y usando equipos de muchos fabricantes diferentes (por ejemplo, Pharmacia Biotech).

35 Renaturalización y replegamiento

40 En algunas realizaciones de la presente invención, más del 50 % del polipéptido transgénico, el polipéptido, la proteína o fragmento de la misma expresados producidos puede producirse de una forma renaturalizable en una célula hospedadora. En otra realización aproximadamente el 60 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 % de la proteína expresada se obtiene en o puede renaturalizarse en la forma activa.

45 La proteína insoluble puede renaturalizarse o replegarse para generar la conformación estructural de la proteína secundaria y terciaria. Pueden usarse etapas de replegamiento de proteínas, según sea necesario, para completar la configuración del producto recombinante. El replegamiento y la renaturalización pueden conseguirse usando un agente conocido en la técnica por promover la disociación/asociación de proteínas. Por ejemplo, la proteína puede incubarse con ditiotreitol seguido de incubación con sal disódica de glutatión oxidado seguido de incubación con un tampón que contiene un agente de replegamiento tal como la urea.

50 55 La proteína o polipéptido de interés también puede renaturalizarse, por ejemplo, dializándola frente a solución salina tamponada con fosfato (PBS) o acetato de Na 50°mM, tampón pH 6 más NaCl 200°mM. Como alternativa, la proteína puede replegarse mientras está inmovilizada en una columna, tal como la columna Ni NTA utilizando un gradiente lineal de urea 6 M-1 M en NaCl 500°mM, glicerol al 20 %, Tris/HCl pH 7,4 20 mM, que contiene inhibidores de proteasa. La renaturalización puede realizarse durante un periodo de 1,5 horas o más. Después de la renaturalización las proteínas pueden eluirse mediante la adición de imidazol 250°mM. El imidazol puede retirarse mediante una etapa final de diálisis frente a PBS o tampón pH 6 de acetato de sodio 50°mM más NaCl 200°mM. La proteína purificada puede almacenarse a 4 °C o congelarse a -80 °C.

60 Otros métodos incluyen, por ejemplo, aquellos que pueden describirse en M H Lee et al., *Protein Expr. Purif.*, 25(1): p. 166-73 (2002), W. K. Cho et al., *J. Biotechnology*, 77(2-3): p. 169-78 (2000), Ausubel, et al. (1987 y suplementos periódicos), Deutscher (1990) "*Guide to Protein Purification*", *Methods in Enzymology* vol. 182, y otros volúmenes de esta serie, Coligan, et al. (1996 y suplementos periódicos) *Current Protocols in Protein Science* Wiley/Greene, NY, S. Roe, *Protein Purification Techniques: A Practical Approach (Practical Approach Series)*, Oxford Press (2001); D. Bollag, et al., *Protein Methods*, Wiley-Lisa, Inc. (1996).

65

Polipéptidos recombinantes

La presente invención proporciona la producción de proteínas mejorada en sistemas de expresión de proteínas. Los ejemplos de polipéptidos recombinantes que pueden usarse en la presente invención incluyen polipéptidos derivados de organismos procariotas y eucariotas. Dichos organismos incluyen organismos del dominio Arquea, Bacteria, Eucaria, incluyendo organismos del Reino Protista, Hongos, Plantas y Animales.

Los péptidos recombinantes que se han de expresar de acuerdo con la presente invención pueden expresarse a partir de polinucleótidos en los que la secuencia codificante del polipéptido diana está unida operativamente a elementos reguladores de la transcripción y la traducción para formar un gen funcional a partir del cual la célula hospedadora puede expresar la proteína o péptido. La secuencia codificante puede ser una secuencia codificante nativa para el polipéptido diana, si está disponible, pero más preferentemente será una secuencia codificante que haya sido seleccionada, mejorada u optimizada para su uso en la célula hospedadora de expresión seleccionada. El gen o genes que se produzcan como resultado se habrán construido dentro o se insertarán en uno o más vectores, que después se transformarán en la célula hospedadora de expresión. Un ácido nucleico o un polinucleótido que se dice que se proporciona en una "forma expresable" significa un ácido nucleico o un polinucleótido que contiene al menos un gen que puede ser expresado por la célula hospedadora de expresión bacteriana seleccionada.

Existe, ampliamente disponible para el público, información de secuencias exhaustiva necesaria para la genética molecular y las técnicas de ingeniería genética. El acceso a secuencias completas de nucleótidos de mamífero, así como seres humanos, genes, secuencias de ADNc, secuencias de aminoácidos y de genomas, puede obtenerse de GenBank en la dirección URL www.ncbi.nlm.nih.gov/Entrez. También puede obtenerse información adicional de GeneCards, una enciclopedia electrónica que integra información sobre genes y sus productos y aplicaciones biomédicas del Instituto Weizmann de Ciencia, Genoma y Bioinformática (<http://lbiinformatics.weizmann.ac.il/cards/>), también puede obtenerse información de secuencias de nucleótidos de la Base de Datos de Secuencias de Nucleótidos EMBL (www.ebi.ac.uk/embl/) o del Banco de datos de ADN de Japón (DDBJ, www.ddbj.nig.ac.jp/); los sitios adicionales para obtener información sobre secuencias de aminoácidos incluyen el sitio web de recursos de información de proteínas de Georgetown (www.nbrf.georgetown.edu/pir/) y Swiss-Prot (au.expasy.org/sprot/sprot-top.html).

El método de la presente invención es útil para producir altos niveles de proteína o polipéptido correctamente procesados de interés en un sistema de expresión celular. La proteína o polipéptido de interés (también denominados en el presente documento como "proteína diana" o "polipéptido diana") puede ser de cualquier especie y de cualquier tamaño. Sin embargo, en ciertas realizaciones, la proteína o polipéptido de interés es una proteína o polipéptido terapéuticamente útil. En algunas realizaciones, la proteína puede ser una proteína de mamífero, por ejemplo una proteína humana y puede ser, por ejemplo, un factor de crecimiento, una citocina, una quimiocina o una proteína sanguínea. La proteína o polipéptido de interés pueden procesarse de una manera similar a la proteína o polipéptido nativos. En ciertas realizaciones, la proteína o polipéptido no incluyen una señal de secreción en la secuencia codificante. En ciertas realizaciones, la proteína o polipéptido de interés tienen menos de 100 kD, menos de 50 kD o menos de 30 kD de tamaño. En ciertas realizaciones, la proteína o polipéptido de interés es un polipéptido de al menos aproximadamente 5, 10, 15, 20, 30, 40, 50 o 100 aminoácidos.

Existe, ampliamente disponible para el público, información de secuencias exhaustiva necesaria para la genética molecular y las técnicas de ingeniería genética. El acceso a secuencias completas de nucleótidos de mamífero, así como seres humanos, genes, secuencias de ADNc, secuencias de aminoácidos y de genomas, puede obtenerse de GenBank en la dirección URL www.ncbi.nlm.nih.gov/Entrez. También puede obtenerse información adicional de GeneCards, una enciclopedia electrónica que integra información sobre genes y sus productos y aplicaciones biomédicas del Instituto Weizmann de Ciencia, Genoma y Bioinformática (<http://lbiinformatics.weizmann.ac.il/cards/>), también puede obtenerse información de secuencias de nucleótidos de la Base de Datos de Secuencias de Nucleótidos EMBL (www.ebi.ac.uk/embl/) o del Banco de datos de ADN de Japón (DDBJ, www.ddbj.nig.ac.jp/); los sitios adicionales para obtener información sobre secuencias de aminoácidos incluyen el sitio web de recursos de información de proteínas de Georgetown (www.nbrf.georgetown.edu/pir/) y Swiss-Prot (au.expasy.org/sprot/sprot-top.html).

Los ejemplos de proteínas que pueden expresarse en la presente invención incluyen moléculas tales como, por ejemplo, renina, una hormona de crecimiento, incluyendo la hormona de crecimiento humana; la hormona de crecimiento bovina; factor de liberación de hormona de crecimiento; hormona paratiroidea; hormona estimulante de la tiroides; lipoproteínas; α -1-antitripsina; cadena A de insulina; cadena B de insulina; proinsulina; trombopoyetina; hormona folículo estimulante; calcitonina; hormona luteinizante; glucagón; factores de coagulación tales como el factor VIIIc, el factor IX, el factor tisular y el factor de von Willebrand; factores anticoagulantes tales como la Proteína C; factor de natriurético atrial; tensioactivo pulmonar; un activador del plasminógeno, tal como urocinasa o activador del plasminógeno de tipo urinario o tisular humano (t-PA); bombesina; trombina; factor de crecimiento hematopoyético; factor de necrosis tumoral alfa y beta; encefalinasa; una albúmina sérica tal como la albúmina sérica humano; sustancia inhibidora de Müller; cadena A de relaxina; cadena B de relaxina; prorelaxina; polipéptido asociado a gonadotropina de ratón; una proteína microbiana, tal como beta-lactamasa; Dnasa; inhibina; activina; factor de crecimiento endotelial vascular (VEGF); receptores para hormonas o factores de crecimiento; integrina;

proteína A o D; factores reumatoides; un factor neurotrófico tal como el factor neurotrófico derivado del cerebro (BDNF), neurotrofina-3, -4, -5 o -6 (NT-3, NT-4, NT-5 o NT-6) o un factor de crecimiento de los nervios tal como NGFD- \square ; cardiotrofinas (factor de hipertrofia cardiaca) tales como cardiotrofina-1 (CT-1); factor de crecimiento derivado de plaquetas (PDGF); factor de crecimiento de fibroblastos tal como aFGF y bFGF; factor de crecimiento epidérmico (EGF); factor de crecimiento transformante (TGF) tal como TGF-alfa y TGF- \square , incluyendo TGF- \square 1, TGF- \square 2, TGF- \square 3, TGF- \square 4, o TGF- \square 5; factor de crecimiento similar a la insulina I y II (IGF-I e IGF-II); des(1-3)-IGF-I (IGF-I de cerebro), proteínas de unión al factor de crecimiento similar a la insulina; Proteínas CD tales como CD-3, CD-4, CD-8 y CD-19; eritropoyetina; factores osteoinductivos; inmunotoxinas; una proteína morfogénica ósea (BMP); un interferón tal como interferón-alfa, beta y gamma; factores estimulantes de colonias (CSF), por ejemplo, M-CSF, GM-CSF y G-CSF; interleucinas (IL), por ejemplo, IL-1 a IL-10; anticuerpo anti-HER-2; superóxido dismutasa; Receptores de linfocitos T; proteínas de membrana de superficie; factor acelerador del desgaste; antígeno viral tal como, por ejemplo, una porción de la cubierta del SIDA; proteínas de transporte; receptores de migración dirigida; adhesinas; proteínas reguladoras; anticuerpos; y fragmentos de cualquiera de los polipéptidos enumerados anteriormente.

En ciertas realizaciones, la proteína o polipéptido pueden seleccionarse entre IL-1, IL-1a, IL-1b, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-12elasti, IL-13, IL-15, IL-16, IL-18, IL-18BPa, IL-23, IL-24, VIP, eritropoyetina, GM-CSF, G-CSF, M-CSF, factor de crecimiento derivado de plaquetas (PDGF), MSF, ligando FLT-3, EGF, factor de crecimiento de fibroblastos (FGF, por ejemplo, \square -FGF (FGF-1), \square -FGF (FGF-2), FGF-3, FGF-4, FGF-5, FGF-6 o FGF-7), factores de crecimiento similares a la insulina (por ejemplo, IGF-1, IGF-2); factores de necrosis tumoral (por ejemplo, TNF, linfotoxina), factores de crecimiento de los nervios (por ejemplo, NGF), factor de crecimiento endotelial vascular (VEGF); interferones (por ejemplo, IFN- \square , IFN- \square , IFN- \square); factor inhibidor de la leucemia (LIF); factor neurotrófico ciliar (CNTF); oncostatina M; factor de células madre (SCF); factores de crecimiento transformante (por ejemplo, TGF- \square , TGF- \square 1, TGF- \square 2, TGF- \square 3); superfamilia TNF (por ejemplo, LIGHT/TNFSF14, STALL-1/TNFSF13B (Bly5, BAFF, THANK), TNFalfa/TNFSF2 y TWEAK/TNFSF12); o quimiocinas (BCA-1/BLC-1, BRAK/Kec, CXCL16, CXCR3, ENA-78/LIX, Eotaxina-1, Eotaxina-2/MPIF-2, Éxodo-2/SLC, fractalcina/Neurotactina, GROalpha/MGSA, HCC-1, I-TAC, Linfotactina/ATAC/SCM, MCP-1/MCAF, MCP-3, MCP-4, MDC/STCP-1/ABCD-1, MIP-1.cuadratura., MIP-1.cuadratura., MIP-2.cuadratura./GRO.cuadratura., MIP-3.cuadratura./Exodus/LARC, MIP-3/Ex-3/ELC, MIP-4/PARC/DC-CK1, PF-4, RANTES, SDF1, TARC o TECK).

En una realización de la presente invención, la proteína de interés puede ser una proteína o polipéptido de múltiples subunidades. Las proteínas de múltiples subunidades que pueden expresarse incluyen las proteínas homoméricas y las heteroméricas. Las proteínas de múltiples subunidades pueden incluir dos o más subunidades, que pueden ser iguales o diferentes. Por ejemplo, la proteína puede ser una proteína homomérica que comprende 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 o más subunidades. La proteína también puede ser una proteína heteromérica que incluye 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 o más subunidades. Las proteínas de múltiples subunidades de ejemplo incluyen: receptores incluyendo los receptores de canal iónico; proteínas de la matriz extracelular incluyendo condroitina; colágeno; inmunomoduladores incluidas las proteínas del CMH, anticuerpos de cadena completa y fragmentos de anticuerpos; enzimas que incluyen ARN polimerasas y ADN polimerasas; y proteínas de membrana.

En otra realización, la proteína de interés puede ser una proteína sanguínea. Las proteínas sanguíneas expresadas en esta realización incluyen, pero no se limitan a proteínas transportadoras, tales como la albúmina, incluyendo albúmina humana y bovina, transferrina, medias moléculas de transferrina recombinante, haptoglobina, fibrinógeno y otros factores de la coagulación, componentes del complemento, inmunoglobulinas, inhibidores enzimáticos, precursores de sustancias tales como la angiotensina y la bradicinina, insulina, endotelina y globulina, incluyendo alfa, beta y gamma-globulina, y otros tipos de proteínas, polipéptidos y fragmentos de los mismos que se encuentran principalmente en la sangre de mamíferos. Se han notificado las secuencias de aminoácidos para numerosas proteínas sanguíneas (véase, S. S. Baldwin (1993) *Comp. Biochem Physiol.* 106b:203-218), incluyendo la secuencia de aminoácidos para la albúmina sérica humana (Lawn, L. M., et al. (1981) *Nucleic Acids Research*, 9:6103-6114) y la transferrina sérica humana (Yang, F. et al. (1984) 81:2752-2756).

En otra realización, la proteína de interés puede ser una enzima o cofactor recombinante. Las enzimas y cofactores expresados en la presente realización incluyen, pero no se limitan a aldolasas, amina oxidasas, aminoácido oxidasas, aspartasas, enzimas dependientes de B12, carboxipeptidasas, carboxiesterasas, carboxilasas, quimi tripsina, enzimas que requieren CoA, cianhidrina sintetasas, cistationina sintetasas, descarboxilasas, deshidrogenasas, alcohol deshidrogenasas, deshidratasas, diaforasas, dioxigenasas, enoato-reductasas, epóxido hidrasas, fumarasas, galactosa oxidasa, glucosa isomerasas, glucosa oxidasas, glicosiltransferasas, metiltransferasas, nitrilo hidrasas, nucleósido fosforilasas, oxidorreductasas, oxinitilasas, peptidasas, glicosiltransferasas, peroxidasas, enzimas fusionados con un polipéptido terapéuticamente activo, activador del plasminógeno tisular; urocinasa, reptilasa, estreptocinasa; catalasa, superóxido dismutasa; ADNasa, aminoácido hidrolasas (por ejemplo, asparraginasas, amidohidrolasas); carboxipeptidasas; proteasas, tripsina, pepsina, quimi tripsina, papaina, bromelina, colagenasa; neuraminidasa; lactasa, maltasa, sacarasa y arabinofuranosidasas.

En otra realización, la proteína de interés puede ser un anticuerpo monocatenario, de fragmento Fab y/o de cadena completa o fragmentos o porciones de los mismos. Un anticuerpo monocatenario puede incluir las regiones de unión a antígeno de los anticuerpos en una sola cadena polipeptídica plegada de forma estable. Los fragmentos Fab

pueden ser un trozo de un anticuerpo particular. El fragmento Fab puede contener el sitio de unión de antígeno. El fragmento Fab puede contener 2 cadenas: un fragmento de cadena ligera y uno de cadena pesada. Estos fragmentos pueden unirse a través de un enlazador o un enlace disulfuro.

5 En ciertas realizaciones, la proteína de interés es, o es sustancialmente homóloga a, una proteína nativa, tal como una proteína nativa de mamífero o humana. En estas realizaciones, la proteína no se encuentra en una forma concatémica, pero está unida solo a una señal de secreción y opcionalmente una secuencia marcadora para la purificación y/o el reconocimiento.

10 Aminoacil-ARNt sintetasas ortogonales (O-RS, por sus siglas en inglés)

15 Como se ha descrito anteriormente, la presente memoria descriptiva describe construcciones de ácido nucleico que se han modificado por ingeniería genética para expresar ARNt ortogonales y ARNt sintetasas ortogonales que permiten la introducción de un aminoácido no natural en un polipéptido recombinante de interés. Con el fin de incorporar específicamente un aminoácido no natural en una proteína o polipéptido de interés, en una célula, la especificidad de sustrato de la sintetasa se altera de manera que solo el aminoácido no natural deseado, pero ninguno de los 20 aminoácidos frecuentes se carguen al ARNt. Si la sintetasa ortogonal es promiscua, dará como resultado proteínas mutantes con una mezcla de aminoácidos naturales y no naturales en la posición diana.

20 Se describe en el presente documento una célula que incluye una aminoacil-ARNt sintetasa ortogonal (O-RS). La O-RS preferentemente aminoacila un ARNt ortogonal (O-ARNt) con un aminoácido no natural en la célula. En ciertos ejemplos, la O-RS utiliza más de un aminoácido no natural, por ejemplo, dos o más, tres o más, etc. Por tanto, una O-RS puede tener la capacidad de aminoacilar preferentemente un O-ARNt con aminoácidos no naturales diferentes. Esto permite un nivel adicional de control mediante la selección de qué aminoácido no natural o
25 combinación de aminoácidos no naturales se ponen con la célula y/o mediante la selección de las diferentes cantidades de aminoácidos no naturales que se ponen con la célula para su incorporación.

30 Una O-RS tiene opcionalmente una o más propiedades enzimáticas mejoradas o potenciadas para el aminoácido no natural en comparación con un aminoácido natural. Estas propiedades incluyen, por ejemplo, Km mayor, Km menor, kcat mayor, kcat menor, kcat/km menor, kcat/km mayor, etc., para el aminoácido no natural, en comparación con un aminoácido de origen natural, por ejemplo, uno de los 20 aminoácidos frecuentes conocidos.

Opcionalmente, la O-RS puede proporcionarse a la célula mediante un polipéptido que incluye una O-RS y/o
35 mediante un polinucleótido que codifica una O-RS o una porción de la misma.

Véase las Publicaciones de Patente de los EE.UU. N.º 20100093082 y 20080118464 para información adicional acerca de tRNA ("ortogonales") y sintetasas modificados.

40 En un ejemplo, una célula comprende una aminoacil-tRNA sintetasa ortogonal (O-RS), un ARNt ortogonal (O-ARNt), un aminoácido no natural y un ácido nucleico que comprende un polinucleótido que codifica un polipéptido de interés, polinucleótido que comprende un codón selector que es reconocido por el O-ARNt. Los ejemplos de un codón selector considerado pueden incluir, pero no se limitan a, un codón de parada o tal vez un codón de de cuatro bases que se puede asociar al O-ARNt. El polipéptido recombinante puede codificar uno o más de los codones selectores. La O-RS aminoacila preferentemente el ARNt ortogonal (O-ARNt) con el aminoácido no natural en la
45 célula y la célula produce el polipéptido de interés en ausencia del aminoácido no natural con un rendimiento que es, por ejemplo, inferior al 30 %, inferior al 20 %, inferior al 15 %, inferior al 10 %, inferior al 5 %, inferior al 2,5 %, etc., del rendimiento del polipéptido en presencia del aminoácido no natural.

ARNt ortogonales

50 En la memoria descriptiva se describen células que incluyen un ARNt ortogonal (O-ARNt). El ARNt ortogonal media la incorporación de un aminoácido no natural en una proteína que está codificada por un polinucleótido que comprende un codón selector que es reconocido por el O-ARNt, in vivo. Un O-ARNt como se describe media la incorporación de un aminoácido no natural en una proteína, por ejemplo, al menos un 40 %, al menos un 45 %, al
55 menos un 50 %, al menos un 60 %, al menos un 75 %, al menos un 80 % o incluso un 90 % o más tan eficientemente como el ARNt que comprende o se procesa en una célula.

Pares de ARNt ortogonal y aminoacil-tRNA sintetasa ortogonal

60 Un par ortogonal se compone de un O-ARNt, por ejemplo, un ARNt supresor, un ARNt de desplazamiento de marco de lectura, o similar, y una O-RS. El O-ARNt no es acilado por sintetasas endógenas y es capaz de mediar la incorporación de un aminoácido no natural en una proteína que está codificada por un polinucleótido que comprende un codón selector que es reconocido por el O-ARNt in vivo. La O-RS reconoce el O-ARNt y preferentemente aminoacila el O-ARNt con un aminoácido no natural en una célula. Se describen métodos para la producción de
65 pares ortogonales junto con pares ortogonales producidos mediante dichos métodos y composiciones de pares

ortogonales para su uso en células. El desarrollo de múltiples pares de ARNt/sintetasa ortogonales puede permitir la incorporación simultánea de múltiples aminoácidos no naturales usando diferentes codones en una célula.

Un par O-ARNt/O-RS ortogonal en una célula puede producirse mediante la importación de un par, por ejemplo, un par supresor no codificante, de un organismo diferente con aminoacilación ineficiente cruzada entre especies. El O-ARNt y la O-RS se expresan eficientemente y se procesan en la célula y el O-ARNt se exporta eficientemente desde el núcleo al citoplasma. Por ejemplo, uno de dichos pares es el par tirosil-ARNt sintetasa/tRNA.sub.CUA de *E. coli* (véase, por ejemplo, H. M. Goodman, et al., (1968), *Nature* 217:1019-1024; y D. G. Barker, et al., (1982), *FEBS Letters*. 150:419-23). La tirosil-ARNt sintetasa de *E. coli* aminoacila eficientemente su afín tRNA.CUA de *E. coli* cuando ambos se expresan en el citoplasma de *S. cerevisiae*, pero no aminoacila el ARNt de *S. cerevisiae*. Véase, por ejemplo, H. Edwards y P. Schimmel, (1990), *Molecular & Cellular Biology* 10:1633-1641; y H. Edwards, et al. (1991), *PNAS United States of America* 88:1153-6. Además, el tirosil tRNA.sub.CUA de *E. coli* es un mal sustrato para las aminoacil-tRNA sintetetas de *S. cerevisiae* (véase, por ejemplo, V. Trezeguet, et al., (1991), *Molecular & Cellular Biology* 11: 2744-51), pero funciona eficientemente en la traducción de proteínas en *S. cerevisiae*. Véase, por ejemplo, H. Edwards y P. Schimmel, (1990) *Molecular & Cellular Biology* 10:1633-1641; H. Edwards, et al., (1991), *PNAS United States of America*. 88:1153-6; y, V. Trezeguet, et al., (1991), *Molecular & Cellular Biology* 11:2744-51. Además, TyrRS de *E. coli* no tiene un mecanismo de edición para corregir un aminoácido no natural unido al tRNA.

El O-ARNt y la O-RS pueden ser de origen natural o pueden derivar mediante la mutación de un ARNt y/o RS de origen natural, que genera bibliotecas de ARNt y/o bibliotecas de RS, a partir de una diversidad de organismos. Véase la sección titulada "Fuentes y hospedadores" en el presente documento. En diversos ejemplos, el O-ARNt y la O-RS derivan de al menos un organismo. En otro ejemplo, el O-ARNt deriva de un tRNA de origen natural o de origen natural mutado de un primer organismo y la O-RS deriva de una RS de origen natural o de origen natural mutado de un segundo organismo. En un ejemplo, los organismos no vertebrados primero y segundo son los mismos. Como alternativa, los organismos no vertebrados primero y segundo pueden ser diferentes.

Véase también, la Solicitud de Patente Internacional WO 2002/086075, titulada "*Methods and compositions for the production of orthogonal tRNA-aminoacyl-tRNA synthetase pairs*".

Los siguientes ejemplos se ofrecen a modo de ilustración y no a modo de limitación.

EJEMPLOS EXPERIMENTALES

Las cepas de *P. fluorescens* auxotróficas para metionina DC454, DC485, DC552, DC556 y DC568 (delección de metC) que llevan la construcción citoplasmática-IFN-beta pIFN-001 se analizaron como se indica a continuación.

Brevemente, se inocularon colonias individuales en medio M9 complementado con glucosa al 1 % y elementos traza además de metionina (250 ug/ml) y se cultivaron durante la noche a 30 °C con agitación. Los cultivos de siembra se usaron para inocular 200 ml de Medio Dow HTP (un medio de sales mínimo definido con glicerol como fuente de carbono), además de metionina (250 ug/ml) hasta una DO600 de ~0,1. Los matraces se incubaron durante la noche a 30 °C con agitación. Tras una fase inicial de crecimiento de 24 horas, las células se recogieron por centrifugación y se resuspendieron en el mismo medio HTP sin metionina y se incubaron a 30 °C durante 30 minutos. Las células se recogieron por centrifugación y se resuspendieron en 200 ml de medio HTP ya sea con metionina (250 ug/ml) o azidohomoalanina (AHA) (1 mg/ml) como aditivo y se indujo la expresión a través del promotor Ptac con la adición de isopropil- β -D-1-tiogalactopiranosido (IPTG) 0,3 mM. Cuando fue apropiado, la expresión del modulador de plegado se indujo con manitol al 1 %. Los cultivos se muestrearon en el momento de la inducción (I0) y a las 6 (I6), 12 (I12) y 24 horas (I24) después de la inducción. La densidad celular se midió mediante densidad óptica a 600 nm (DO600). Después de la centrifugación las células se resuspendieron en PBS y la densidad celular se ajustó a DO600 = 20. Se transfirieron alícuotas de 200 μ l a tubos recién preparados y se congelaron a -80 °C para su posterior procesamiento.

Se prepararon fracciones solubles e insolubles por ultrasonidos seguidos de centrifugación. Se descongelaron muestras de caldo de cultivo diluido (400 μ l) y se sometieron a ultrasonidos con un Cell Lysis Automated Sonication System (CLASS, Scinomix) con un cuerno de punta de la sonda 24. Los lisados se centrifugaron a 14.000 rpm durante 20 minutos (4 °C) y los sobrenadantes se recogieron (fracción soluble). Los sedimentos (fracción insoluble) se congelaron y más tarde se descongelaron para su procesamiento. Se retiró el sobrenadante residual del sedimento mediante la recentrifugación a 20.800 xg durante 20 minutos a 4 °C. Los sedimentos se resuspendieron en tampón fosfato salino (PBS), pH 7,4, mediante ultrasonidos.

La expresión de IFN-beta en las cepas auxotróficas para metionina se evaluó en la escala de 1L (200 ml de volumen de trabajo). Se cultivaron matraces de agitación por triplicado para cada una de las cinco cepas construidas (véase la Tabla 3) y se indujeron como se ha descrito en Materiales y Métodos. Dos de los matraces de agitación se usaron para el análisis de la incorporación de AHA, mientras que el matraz restante se usó como control con metionina como aditivo.

Las cepas de expresión de IFN-beta crecieron como se esperaba, alcanzando densidades celulares de aproximadamente 10 unidades de DO600 después de un periodo inicial de crecimiento de 24 horas y de 15-30 unidades después de un periodo de inducción de 24 horas. En general, el crecimiento de las cepas después de la adición de AHA y metionina fue similar.

5 Las muestras tomadas a I0, I6, I12 y I24 se normalizaron a una densidad celular de 20 unidades de DO600 y se fraccionaron. Se analizó la proteína de las fracciones tanto solubles como insolubles resultantes mediante SDS-PAGE y transferencias Western (ejemplo que se muestra en la Figura 2). En todas las cepas, se expresó IFN-beta en presencia de metionina o AHA después de la inducción y, de forma coherente con las observaciones realizadas en el estudio inicial, el IFN-beta expresado se acumuló principalmente en la fracción celular insoluble. Los perfiles y niveles de expresión de IFN-beta fueron similares entre las cepas analizadas. Dentro de cada cepa, los matraces duplicados con adición de AHA proporcionaron resultados muy similares. Y en el caso de WT, FMO-16 y FMO-18, parecía que se había expresado más proteína IFN-beta en las muestras que contenían AHA que en la muestra de control que contenía metionina. Basándose en los resultados de SDS-PAGE, se estimó que el rendimiento de la proteína diana expresada estaba en el intervalo de 100-300 mg/l de cultivo.

Tabla 3. Números de los matraces de agitación de las cepas para el análisis de incorporación de AHA

Número del matraz	Cepa hospedadora	Nombre del hospedador	Aditivo de aminoácido
1	DC454	WT	Met
2			AHA
3			AHA
4	DC485	PD-17	Met
5			AHA
6			AHA
7	DC552	FMO-16	Met
8			AHA
9			AHA
10	DC556	FMO-12	Met
11			AHA
12			AHA
13	DC568	FMO-18	Met
14			AHA
15			AHA

REIVINDICACIONES

1. Un método para producir un polipéptido recombinante de interés que comprende:

- 5 a) obtener una población de células auxotróficas para un primer metabolito y un segundo metabolito, en el que dicho segundo metabolito es un aminoácido natural;
- b) introducir en células de dicha población una primera construcción de ácido nucleico que comprende un marcador de selección auxotrófica, en el que dicho marcador de selección auxotrófica comprende una primera
10 secuencia de ácido nucleico que codifica al menos un polipéptido activo en la biosíntesis de dicho primer metabolito, y en el que la expresión de dicho marcador de selección auxotrófica restablece la prototrofia para el primer metabolito;
- c) introducir en células de dicha población: (i) una segunda secuencia de ácido nucleico que codifica dicho polipéptido recombinante de interés; y (ii) un promotor unido operativamente a dicha segunda secuencia de ácido nucleico para dirigir la expresión de la segunda secuencia de ácido nucleico;
- 15 d) someter dicha población de células a un primer medio que carece de dicho primer metabolito en condiciones de manera que se obtengan células transformadas que tengan la prototrofia restablecida para el primer metabolito; y
- e) someter dichas células transformadas a un segundo medio que comprende un aminoácido no natural correlacionado con dicho segundo metabolito en condiciones de manera que dicha segunda secuencia de ácido
20 nucleico se exprese para producir dicho polipéptido recombinante de interés que tenga dicho aminoácido no natural incorporado en el mismo; en el que dicho primer medio y dicho segundo medio son iguales o diferentes; en el que dicha segunda secuencia de ácido nucleico se proporciona en dicha primera construcción de ácido nucleico y en el que dicha célula hospedadora se selecciona entre el Subgrupo 18 de Proteobacterias gramnegativas.
- 25 2. El método de la reivindicación 1, en el que el primer medio y el segundo medio son diferentes.
3. El método de la reivindicación 1, en el que dicho primer medio y dicho segundo medio son iguales.
- 30 4. El método de la reivindicación 1, en el que dicho primer metabolito es un aminoácido o en el que dicho primer metabolito es un nucleósido, nucleótido o base nitrogenada o en el que dicho primer metabolito es una fuente de carbono.
5. El método de la reivindicación 1, en el que dicha primera secuencia de ácido nucleico codifica una enzima biosintética de tipo pirimidina o una enzima biosintética de tipo purina, o una enzima de aminoácidos biosintéticos o una enzima de utilización.
- 35 6. El método de la reivindicación 1, en el que dicha población de células comprende células hospedadoras bacterianas.
- 40 7. El método de la reivindicación 6, en el que dichas células hospedadoras bacterianas comprenden *Pseudomonas fluorescens* o *E. coli*.
8. El método de la reivindicación 1, en el que dicha población de células carece de una construcción de expresión que comprende un marcador de selección con antibióticos.
- 45 9. El método de la reivindicación 1, en el que dicho primer medio carece de un antibiótico.
- 50 10. El método de la reivindicación 1, en el que dicho segundo metabolito comprende metionina.

Crecimiento de auxótrofos para metionina en Met/AHA

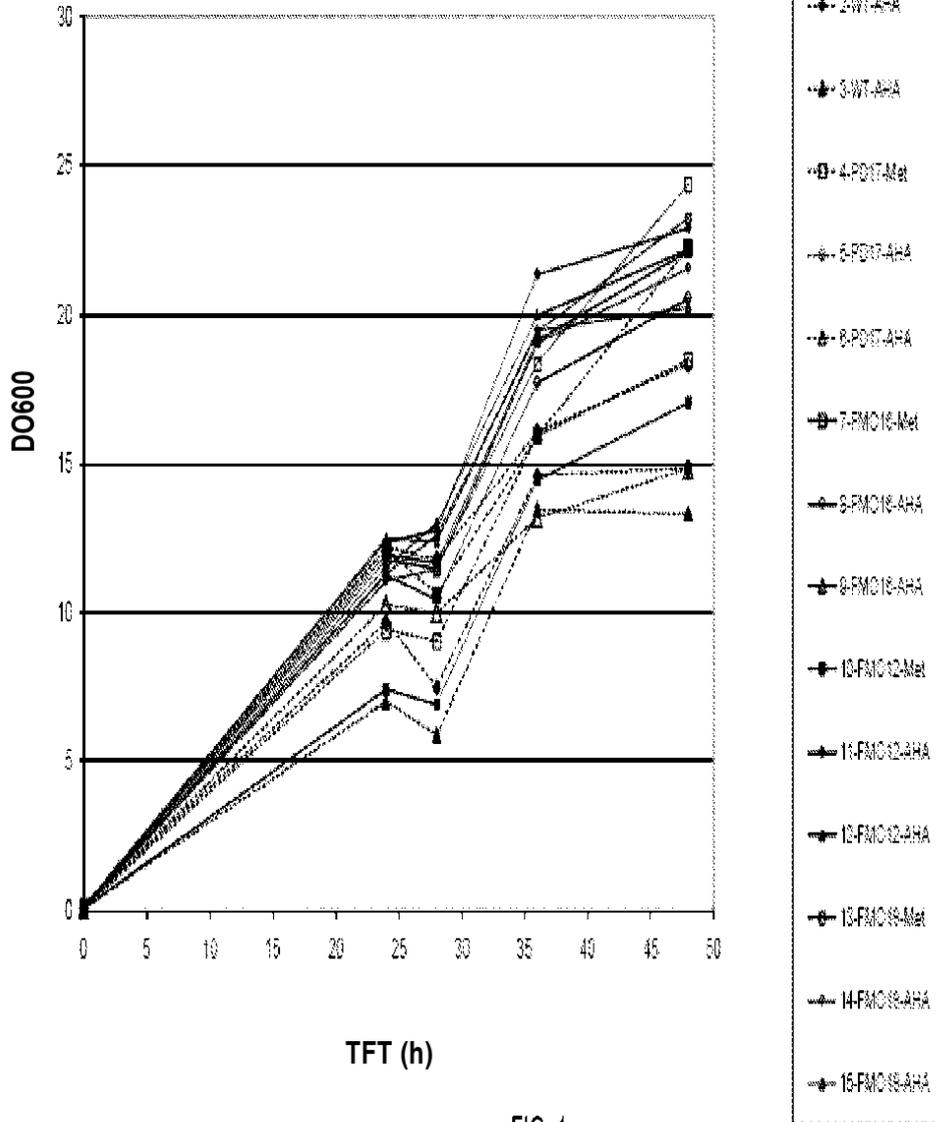


FIG. 1

