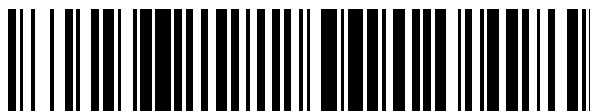


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 931**

51 Int. Cl.:

C12N 15/10 (2006.01)

C07K 16/00 (2006.01)

C40B 50/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.07.2010** **E 16196481 (2)**

97 Fecha y número de publicación de la concesión europea: **12.09.2018** **EP 3156485**

54 Título: **Selección y evolución simultáneas e integradas de rendimiento y expresión de anticuerpos/proteínas en huéspedes de producción**

30 Prioridad:

17.07.2009 US 271168 P

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

26.02.2019

73 Titular/es:

**BIOATLA, LLC (100.0%)
11085 Torreyana Road, Suite 100
San Diego, CA 92121, US**

72 Inventor/es:

SHORT, JAY MILTON

74 Agente/Representante:

CARVAJAL Y URQUIJO, Isabel

ES 2 701 931 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

Selección y evolución simultáneas e integradas de rendimiento y expresión de anticuerpos/proteínas en huéspedes de producción.

Campo de la invención

5 En un aspecto particular, la presente invención es relevante para proteínas y para su optimización mediante evolución de proteínas. Se descubren proteínas terapéuticas, se hacen evolucionar y se fabrican en el mismo huésped usando los mismos sistemas genéticos.

Antecedentes de la invención

10 Se han diseñado, desarrollado e implementado una variedad de sistemas de anticuerpos y otras proteínas que generan moléculas terapéuticas proteicas. Más recientemente, se han desarrollado muchos sistemas de evolución para potenciar la función de las proteínas. Por separado, se han desarrollado sistemas de expresión en mamífero para la producción de alto rendimiento de anticuerpos y otras proteínas para aplicaciones terapéuticas. Hasta la fecha, ningún grupo ha desarrollado un sistema para permitir la generación, evolución de un anticuerpo o proteína, y producción/fabricación de proteínas en un sistema de expresión en mamífero eficaz individual.

15 El documento WO 2002092780 se refiere a métodos para generar conjuntos, o bibliotecas, de ácidos nucleicos que codifican para sitios de unión a antígeno, tales como anticuerpos, dominios de anticuerpo u otros fragmentos de anticuerpo. El método genera estos sitios de unión a antígeno variantes alterando ácidos nucleicos molde usando una técnica seleccionada de mutagénesis por saturación, reensamblaje de ligamiento sintético, o una combinación de los mismos.

20 Rajpal *et al.*, "A general method for greatly improving the affinity of antibodies by using combinatorial libraries", Proc Natl Acad Sci USA., vol. 102, páginas 8466-8471, (2005) dan a conocer un método para mejorar la afinidad de anticuerpos usando la denominada mutagénesis revisada que evalúa y optimiza simultáneamente mutaciones combinatorias de aminoácidos seleccionados en los anticuerpos. El procedimiento alcanza una distribución precisa dentro de uno o más regiones determinantes de la complementariedad del anticuerpo explorando la contribución sinérgica de la química de cadenas laterales de aminoácidos.

25 Iba *et al.*, "Expression vectors for the introduction of highly diverged sequences into the six complementarity-determining regions of an antibody", Gene, vol. 194, páginas 35-46, (1997) dan a conocer tres clases de vectores fagémidos que permiten la introducción simultánea de secuencias con alta divergencia en seis regiones determinantes de la complementariedad de un anticuerpo mediante la reacción en cadena de la polimerasa con cebadores de oligodesoxinucleótidos degenerados. Los vectores fagémidos se pueden usar para expresar o bien la forma de Fv, Fv de cadena sencilla (scFv) o bien Fab de un anticuerpo fusionado con la media molécula de cpIII en la superficie de un fago M13.

30 El documento WO 2005/003345 da a conocer un método de mutagénesis mediante el que se introduce un aminoácido predeterminado en todas y cada una de las posiciones de un conjunto seleccionado de posiciones en una región preseleccionada (o varias regiones diferentes) de un polipéptido para producir una biblioteca de mutantes de polipéptido. Las bibliotecas producidas pueden usarse para estudiar el papel de aminoácidos específicos en una estructura de polipéptido y su función, o para desarrollar polipéptidos nuevos o mejorados tales como anticuerpos, fragmentos de anticuerpo, anticuerpos de cadena sencilla, enzimas y ligandos.

35 El documento WO 2003/023032 da a conocer un procedimiento para la evolución dirigida de alto rendimiento de péptidos y proteínas, particularmente aquellos que actúan en entornos biológicos complejos. Los péptidos y proteínas incluyen, pero no se limitan a, proteínas intracelulares, proteínas de mensajeros/señalización/hormonas y proteínas virales.

40 Se desarrollan muchos anticuerpos usando sistemas de presentación en fago de bacterias en bacterias, mientras que la expresión de anticuerpos de longitud completa se lleva a cabo principalmente en células de mamífero. Esta falta de similitud hace que sea imposible la evolución o selección de clones para expresión. Otras barreras incluyen el requisito tradicional de que se examinen grandes números de variantes usando tecnologías tradicionales y se han optimizado sistemas de expresión en mamífero de alto nivel para la expresión, no la clonación de grandes números de variantes. No se ha diseñado previamente un sistema integrado de selección de proteínas, evolución y expresión en mamífero. El uso de células eucariotas para manejar grandes números, combinado con evolución no estocástica de la proteína en el interior de una célula huésped eucariota optimizada, aumenta la probabilidad de éxito y acelera en gran medida el procedimiento para generar una proteína optimizada que se expresará a niveles suficientemente altos en células eucariotas que se desea fabricar.

Sumario de la invención

55 La presente divulgación proporciona métodos de integración de generación y/o selección de proteínas terapéuticas (incluyendo anticuerpos), evolución y expresión en un huésped eucariota, tal como un huésped de célula de

mamífero o un huésped de célula de levadura, para la fabricación en un sistema individual. Se generan, se optimizan y se fabrican proteínas, que incluyen anticuerpos terapéuticos en el mismo sistema de huésped eucariota. El sistema dado a conocer de Optimización de anticuerpos integrada completa (CIAO!TM) permite la evolución simultánea de la optimización de la expresión y el rendimiento de proteínas.

- 5 En una realización, la divulgación proporciona un método de selección, evolución y expresión de un anticuerpo en un huésped de producción celular eucariota; comprendiendo el método etapas de: hacer evolucionar una proteína molde para producir un conjunto de proteínas mutantes en el huésped de producción celular eucariota; examinar las proteínas mutantes para determinar la al menos una propiedad, característica o actividad predeterminada; seleccionar una proteína mutante superior del conjunto de proteínas mutantes basándose en la al menos una propiedad, característica o actividad predeterminada; y fabricar la proteína mutante superior en el mismo huésped de producción celular eucariota que se usa en la etapa de hacer evolucionar.

15 En un aspecto, el huésped de producción celular eucariota se selecciona de células de fibroblasto de ratón 3T3; células de fibroblasto de hámster sirio BHK21; células epiteliales de perro, MDCK; células epiteliales humanas Hela; células epiteliales de rata canguro PtK1; células plasmáticas de ratón SP2/0; y células plasmáticas de ratón NS0; células de riñón embrionario humano HEK 293; células de riñón de mono COS; células de ovario de hámster chino CHO, CHO-S; células embrionarias de ratón R1; células embrionarias de ratón E14.1; células embrionarias humanas H1; células embrionarias humanas H9; células embrionarias humanas PER C.6; células de levadura *S. cerevisiae*; o células de levadura *Picchia*. En un aspecto particular, el huésped de producción celular eucariota es CHO-S o HEK293. En un aspecto, las etapas de examen utilizan clasificación de células activada por fluorescencia (FACS).

- 20 En un aspecto, la etapa de hacer evolucionar comprende generar n-1 conjuntos independientes de polipéptidos mutantes a partir de la proteína molde, comprendiendo cada conjunto polipéptidos miembros que tienen un número X de residuos de aminoácido diferentes predeterminados en una posición individual predeterminada del polipéptido; en el que cada conjunto de polipéptidos difiere en la posición individual predeterminada; y el número de polipéptidos miembros diferentes generados es equivalente a $[n-1] \times X$. En un aspecto, X representa los 19 residuos de aminoácidos de origen natural no presentes en una posición dada de la proteína molde.

25 En otro aspecto, la etapa de examen comprende someter a ensayo cada polipéptido miembro para determinar al menos una propiedad, característica o actividad predeterminada; identificar cualquier cambio en dicha propiedad, característica o actividad del polipéptido miembro en relación con la proteína molde; crear un mapa funcional en el que el mapa funcional se usa para identificar posiciones y mutaciones en el polipéptido mutante que dan como resultado un mutante superior y/o una mutación silenciosa en comparación con la proteína molde.

30 En otro aspecto, el mapa funcional se usa para identificar uno o más del grupo que consiste en (a) posiciones y mutaciones que no afectan la actividad del polipéptido mutante en comparación con la proteína molde; (b) sitios completamente mutables en comparación con la proteína molde; y (c) posiciones y mutaciones que dan como resultado un mutante superior en comparación con la proteína molde.

- 35 En otro aspecto, la etapa de generación comprende someter un polinucleótido que contiene codones que codifica para dicha proteína molde a amplificación basada en polimerasa usando un oligonucleótido degenerado 64 veces para cada codón que va a mutagenizarse, en el que cada uno de dichos oligonucleótidos degenerados 64 veces se compone de una primera secuencia homóloga y una secuencia de triplete N,N,N degenerada, de modo que se genera un conjunto de polinucleótidos de progenie; y someter dicho conjunto de polinucleótidos de progenie a amplificación clonal de manera que se expresan polipéptidos codificados por los polinucleótidos de progenie.

40 En otro aspecto, la propiedad, característica o actividad predeterminada se selecciona de reducción de la agregación proteína-proteína, potenciación de la estabilidad de proteína, aumento de la solubilidad de anticuerpo, introducción de sitios de glicosilación, introducción de sitios de conjugación, reducción de la inmunogenicidad, potenciación de proteína expresión, aumento de la afinidad antigénica, disminución de la afinidad antigénica, cambio en la afinidad de unión, cambio en la inmunogenicidad y potenciación de la especificidad.

45 En un aspecto, la etapa de selección comprende la creación de un mapa funcional en el que el mapa funcional se usa para identificar posiciones y mutaciones en el polipéptido mutante que da como resultado un mutante superior y/o una mutación silenciosa en comparación con la proteína molde. En otro aspecto, las etapas de selección comprenden la clasificación de células activadas por fluorescencia (FACS).

- 50 En otro aspecto, la etapa de hacer evolucionar comprende una técnica de evolución seleccionada de una de evolución de posición completa (CPE); evolución de inserción de posición completa (CPI); evolución de delección de posición completa (CPD); evolución de posición completa (CPE) seguida por síntesis combinatoria de proteínas (CPS); evolución de delección de posición completa (CPD) seguida por síntesis combinatoria de proteínas (CPS); o evolución de delección de posición completa (CPD) seguida por síntesis combinatoria de proteínas (CPS).

55 En un aspecto, la proteína es un anticuerpo. En otro aspecto, el anticuerpo es un anticuerpo de longitud completa.

En otro aspecto, la al menos una propiedad, característica o actividad predeterminada en la etapa de examen (e) comprende uno o más de (1) examinar para una mutación silenciosa y (2) examinar para una mutación de cambio

de sentido; en comparación con la proteína molde.

En otro aspecto, el anticuerpo es un fragmento de anticuerpo seleccionado de Fc y Fv; región de entramado; y una o más CDR que se modifican en el anticuerpo mutante superior en comparación con el anticuerpo molde que va a hacerse evolucionar.

- 5 En un aspecto adicional, la proteína para la evolución y la fabricación se selecciona de una enzima citosina, receptor, proteína de unión al ADN, agente quelante u hormona.

En otro aspecto, la etapa de hacer evolucionar comprende hacer evolucionar la proteína molde para producir un conjunto de proteínas mutantes en el huésped de producción celular eucariota con presentación en la superficie celular.

- 10 En otra realización, la divulgación proporciona un método de evolución para la expresión y fabricación potenciadas de una proteína humana en un huésped de producción celular eucariota; comprendiendo el método seleccionar una proteína humana molde para evolución; hacer evolucionar la proteína humana molde que comprende la generación de codones mutantes que codifican para la proteína humana molde para producir un conjunto de proteínas humanas mutantes en el huésped de producción; examinar el conjunto de proteínas humanas mutantes para determinar al
15 menos una propiedad, característica o actividad predeterminada y examinar para determinar la expresión potenciada en comparación con la proteína humana molde; seleccionar una proteína humana mutante superior del conjunto de proteínas humanas mutantes basándose en (1) retención u optimización de la al menos una propiedad, característica o actividad predeterminada y (2) expresión potenciada en comparación con la proteína humana molde; y fabricar la proteína humana mutante superior que comprende expresar la proteína humana mutante superior en el
20 mismo huésped de producción que en la etapa de hacer evolucionar.

En un aspecto, los codones mutantes de la proteína humana mutante superior dan como resultado al menos una mutación silenciosa y/o mutación de cambio de sentido. En otro aspecto, los codones mutantes de la proteína humana mutante superior dan como resultado al menos una mutación silenciosa.

- 25 En un aspecto adicional, el anticuerpo humano molde es un fármaco terapéutico de anticuerpo aprobado por un comité ético, y el anticuerpo humano mutante superior es un fármaco biosimilar.

En otro aspecto, la etapa de selección comprende seleccionar una proteína humana mutante superior del conjunto de proteínas humanas mutantes basándose en (1) optimización de la al menos una propiedad, característica o actividad predeterminada y (2) expresión potenciada en comparación con la proteína humana molde.

- 30 En otra realización, la divulgación proporciona un método de identificación y producción de una proteína humana diana, comprendiendo el método generar una biblioteca de proteínas humanas en un huésped de producción celular eucariota con presentación en la superficie celular de proteínas; examinar la biblioteca para determinar al menos una propiedad, característica o actividad predeterminada; identificar una proteína humana diana de la biblioteca basándose en la al menos una propiedad, característica o actividad predeterminada; y expresar la proteína humana diana en el mismo huésped de producción celular eucariota que en la etapa de generación para producir una
35 proteína humana diana. En un aspecto, la proteína humana diana es un anticuerpo. En otro aspecto, el anticuerpo es un anticuerpo de longitud completa.

- En otra realización, la divulgación proporciona un método de evolución de una proteína humana en un huésped de fabricación, comprendiendo el método mutar una proteína humana molde para producir un conjunto de proteínas humanas mutantes en un huésped de fabricación; y examinar el conjunto de proteínas de progenie mutantes para
40 determinar al menos una propiedad, característica o actividad predeterminada. En un aspecto, el método comprende además seleccionar una proteína humana mutante superior del conjunto de proteínas humanas mutantes basándose en la al menos una propiedad, característica o actividad predeterminada. En otro aspecto, el método comprende además fabricar la proteína humana mutante superior que comprende expresar la proteína humana mutante superior en el mismo huésped de producción que en la etapa de mutación. En otro aspecto, la etapa de selección comprende
45 además seleccionar una proteína humana mutante superior del conjunto de anticuerpos humanos mutantes basándose en (1) optimización de la al menos una propiedad, característica o actividad predeterminada en comparación con la proteína humana molde, y (2) expresión modificada en comparación con la proteína humana molde. En un aspecto, la expresión modificada es expresión potenciada.

- 50 En otra realización, la divulgación proporciona un método de evolución y fabricación de una proteína humana en un huésped de fabricación celular; comprendiendo el método mutar una proteína humana molde para producir un conjunto de proteínas humanas mutantes en un huésped de fabricación; examinar el conjunto de proteínas humanas mutantes para determinar al menos una propiedad, característica o actividad predeterminada y examinar para expresión modificada; seleccionar una proteína humana mutante superior del conjunto de proteínas humanas mutantes basándose en (1) optimización de la al menos una propiedad, característica o actividad predeterminada, y
55 (2) expresión modificada en comparación con la proteína humana molde; y fabricar la proteína humana que comprende expresar la proteína humana mutante superior en el mismo huésped de fabricación que en la etapa de mutación.

En una realización adicional, la divulgación proporciona un método de evolución para la expresión y fabricación potenciadas de una proteína humana en un huésped de producción celular eucariota; comprendiendo el método mutar una proteína humana molde que comprende generación de codones mutantes que codifican para la proteína humana molde para producir un conjunto de proteínas humanas mutantes en un huésped de fabricación; examinar el conjunto de proteínas humanas mutantes para determinar al menos una propiedad, característica o actividad predeterminada y examinar para determinar la expresión potenciada en comparación con la proteína humana molde; seleccionar una proteína humana mutante superior del conjunto de proteínas humanas mutantes basándose en (1) retención u optimización de la al menos una propiedad, característica o actividad predeterminada y (2) expresión potenciada en comparación con la proteína humana molde; y fabricar la proteína humana mutante superior en el mismo huésped de fabricación que en la etapa de mutación.

En un aspecto, la etapa de examen comprende examinar el conjunto de anticuerpos humanos mutantes para determinar la al menos una propiedad, característica o actividad predeterminada y examinar para determinar la expresión potenciada simultáneamente.

Breve descripción de los dibujos

La figura 1 proporciona un esquema del método CIAO™ de integración de generación y/o selección de proteínas (por ejemplo, anticuerpos) terapéuticas, evolución y expresión en un huésped eucariota, tal como un huésped de célula de mamífero o un huésped de célula de levadura, para la fabricación en un sistema individual.

La figura 2 ilustra cómo se usa la evolución de posición completa (CPE™) para generar una base de datos específica de moléculas (EvoMap™).

La figura 3 muestra un esquema de la síntesis de posición completa (CPS™) que puede usarse para combinar mutantes superiores de CPE™.

La figura 4 muestra un esquema de la evolución de inserción de posición completa (CPI™).

La figura 5 ilustra una combinación de métodos de evolución: un ácido nucleico prolongado a partir de evolución CPI™ se somete a evolución de posición completa (CPE™) y se usa para generar una base de datos específica de moléculas (EvoMap™).

La figura 6 muestra un esquema de evolución de delección de posición completa (CPD™).

La figura 7 ilustra otra combinación de métodos de evolución: un ácido nucleico acortado a partir de evolución CPD™ se somete a evolución de posición completa (CPE™) y se usa para generar una base de datos específica de moléculas (EvoMap™).

Definición de términos

Para facilitar la comprensión de los ejemplos proporcionados en el presente documento, se describirán determinados métodos y/o términos que aparecen frecuentemente.

El término “agente” se usa en el presente documento para indicar un polipéptido, una mezcla de polipéptidos, un alineamiento de compuestos localizados espacialmente (por ejemplo, un alineamiento de péptido VLSIPS, alineamiento de polinucleótido y/o un alineamiento de molécula pequeña combinatoria), macromolécula biológica, una biblioteca de presentación de péptidos en bacteriófago, un anticuerpo de bacteriófago (por ejemplo, scFv) biblioteca de presentación, una biblioteca de presentación de péptidos polisómicos, o un extracto producido a partir de materiales biológicos tales como bacterias, plantas, hongos o células o tejidos de animales (en particular de mamíferos). Se evalúan los agentes para determinar su posible actividad como antineoplásicos, antiinflamatorios o moduladores de la apoptosis mediante su inclusión en ensayos de examen descritos a continuación en el presente documento. Se evalúan los agentes para determinar su posible actividad como inhibidores específicos de la interacción de proteínas (es decir, un agente que inhibe selectivamente una interacción de unión entre dos polipéptidos predeterminados pero que no interfiere sustancialmente en la viabilidad celular) mediante su inclusión en ensayos de examen descritos a continuación en el presente documento.

El término “aminoácido” tal como se usa en el presente documento se refiere a cualquier compuesto orgánico que contiene un grupo amino (--NH₂) y un grupo carboxilo (--COOH); preferiblemente o bien como grupos libres o bien alternativamente tras condensación como parte de enlaces peptídicos. Los “veinte alfa-aminoácidos que forman polipéptidos codificados de manera natural” se entienden en la técnica y se refieren a: alanina (ala o A), arginina (arg o R), asparagina (asn o N), ácido aspártico (asp o D), cisteína (cys o C), ácido glutámico (glu o E), glutamina (gln o Q), glicina (gly o G), histidina (his o H), isoleucina (ile o I), leucina (leu o L), lisina (lys o K), metionina (met o M), fenilalanina (phe o F), prolina (pro o P), serina (ser o S), treonina (thr o T), triptófano (trp o W), tirosina (tyr o Y), y valina (val o V).

El término “amplificación” significa que aumenta el número de copias de un polinucleótido.

El término “anticuerpo”, tal como se usa en el presente documento, se refiere a moléculas de inmunoglobulina

5 intactas, así como a fragmentos de moléculas de inmunoglobulina, tales como fragmentos Fab, Fab', (Fab')₂, Fv, y SCA, que pueden unirse a un epítipo de un antígeno. Estos fragmentos de anticuerpo, que retienen cierta capacidad para unirse selectivamente a un antígeno (por ejemplo, un antígeno de polipéptido) del anticuerpo del que se derivan, pueden producirse usando métodos bien conocidos en la técnica (véase, por ejemplo, Harlow y Lane, citado anteriormente), y se describen adicionalmente, tal como sigue. Los anticuerpos pueden usarse para aislar cantidades preparativas del antígeno mediante cromatografía de inmunoafinidad. Otros usos diversos de tales anticuerpos son diagnosticar y/o estadificar la enfermedad (por ejemplo, neoplasia) y para aplicación terapéutica para tratar una enfermedad, tal como por ejemplo: neoplasia, enfermedad autoinmunitaria, SIDA, enfermedad cardiovascular, infecciones, y similares. Los anticuerpos quiméricos, similares a humanos, humanizados o totalmente humanos son particularmente útiles para la administración a pacientes humanos.

10 Un fragmento Fab consiste en un fragmento de unión a antígeno monovalente de una molécula de anticuerpo, y puede producirse mediante digestión de una molécula de anticuerpo completo con la enzima papaína, para proporcionar un fragmento que consiste en una cadena ligera intacta y una parte de una cadena pesada.

15 Un fragmento Fab' de una molécula de anticuerpo puede obtenerse tratando una molécula de anticuerpo completo con pepsina, seguido por reducción, para proporcionar una molécula que consiste en una cadena ligera intacta y una parte de una cadena pesada. Se obtienen dos fragmentos Fab' por molécula de anticuerpo tratada de esta manera.

Un fragmento (Fab')₂ de un anticuerpo puede obtenerse tratando una molécula de anticuerpo completo con la enzima pepsina, sin reducción posterior. Un fragmento (Fab')₂ es un dímero de dos fragmentos Fab', mantenidos juntos mediante dos enlaces disulfuro.

20 Un fragmento Fv se define como un fragmento modificado mediante ingeniería genética que contiene la región variable de una cadena ligera y la región variable de una cadena pesada expresadas como dos cadenas.

Un anticuerpo de cadena sencilla ("SCA") es una molécula de cadena sencilla modificada mediante ingeniería genética que contiene la región variable de una cadena ligera y la región variable de una cadena pesada, unidas mediante un ligador polipeptídico flexible adecuado.

25 El término "fármaco biosimilar", también denominado "fármaco biológico de continuación", se refiere a nuevas versiones aprobadas de manera oficial de productos biofarmacéuticos innovadores, tras el vencimiento de una patente o exclusividad.

30 El término "huésped de producción celular", o "huésped de fabricación", se refiere a una línea celular usada para la producción o fabricación de proteínas. Células eucariotas tales como células de mamífero, incluyendo, pero sin limitarse a líneas celulares humanas, de ratón, hámster, rata, mono así como líneas celulares de levaduras, insectos y plantas. En un aspecto, un huésped de producción celular de mamífero se selecciona de un miembro del grupo que consiste en células de fibroblasto de ratón 3T3; células de fibroblasto de hámster sirio BHK21; células epiteliales de perro, MDCK; células epiteliales humanas Hela; células epiteliales de rata canguro PtK1; células plasmáticas de ratón SP2/0; y células plasmáticas de ratón NS0; células de riñón embrionario humano HEK 293; células de riñón de mono COS; células de ovario de hámster chino CHO, CHO-S; células embrionarias de ratón R1; células embrionarias de ratón E14.1; células embrionarias humanas H1; células embrionarias humanas H9; células embrionarias humanas PER C.6. En otro aspecto, el huésped de producción celular es a GS-NS0 o GS-CHOK1 línea celular. En otro aspecto, el huésped de producción celular se selecciona de células de levadura *S. cerevisiae*; y células de levadura *Picchia*.

35 Una molécula que tiene una "propiedad quimérica" es una molécula que es: 1) en parte homóloga y en parte heteróloga a una primera molécula de referencia; mientras que 2) al mismo tiempo es en parte homóloga y en parte heteróloga a una segunda molécula de referencia; sin 3) excluir la posibilidad de ser al mismo tiempo en parte homóloga y en parte heteróloga a todavía una o más moléculas de referencia adicionales. En una realización no limitativa, una molécula quimérica puede prepararse ensamblando una reordenación de secuencias moleculares parciales. En un aspecto no limitativo, puede prepararse una molécula de polinucleótido quimérico sintetizando el polinucleótido quimérico usando una pluralidad de moldes moleculares, de manera que el polinucleótido quimérico resultante tenga las propiedades de una pluralidad de moldes.

40 El término "relacionado" tal como se usa en el presente documento se refiere a una secuencia génica que está relacionada evolutiva y funcionalmente entre especies. Por ejemplo, pero sin limitación, en el genoma humano el gen de CD4 humano es el gen relacionado con el gen de 3d4 de ratón, puesto que las secuencias y estructuras de estos dos genes indican que son altamente homólogos y ambos genes codifican para una proteína que funciona en la señalización de la activación de células T a través de reconocimiento de antígeno restringido a CMH de clase II.

45 El término "escala comercial" significa la producción de una proteína o anticuerpo a una escala apropiada para la reventa.

55 Una "ventana de comparación", tal como se usa en el presente documento, se refiere a un segmento conceptual de al menos 20 posiciones de nucleótido contiguas en el que una secuencia de polinucleótido puede compararse con una secuencia de referencia de al menos 20 nucleótidos contiguos y en el que la parte de la secuencia de

polinucleótido en la ventana de comparación puede comprender adiciones o deleciones (es decir, huecos) del 20 por ciento o menos en comparación con la secuencia de referencia (que no comprende adiciones o deleciones) para la alineación óptima de las dos secuencias. La alineación óptima de secuencias para alinear una ventana de comparación puede llevarse a cabo mediante el algoritmo de homología local de Smith y Waterman (1981) Adv. Appl. Math. 2: 482 mediante el algoritmo de alineación de homología de Needleman y Wuncsch J. Mol. Biol. 48: 443 (1970), mediante el método de búsqueda de similitud de Pearson y Lipman Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444 (1988), mediante implementaciones computerizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el paquete de software Wisconsin Genetics versión 7.0, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o mediante inspección, y se selecciona la mejor alineación (es decir, que da como resultado el mayor porcentaje de homología a lo largo de la ventana de comparación) generada mediante los diversos métodos.

Tal como se usa en el presente documento, el término "región determinante de la complementariedad" y "CDR" se refieren al término reconocido en la técnica tal como se ejemplifica mediante Kabat y Chothia. Las definiciones de CDR también se conocen generalmente como regiones supervariables o bucles hipervariables (Chothia y Leks, 1987; Chothia *et al.*, 1989; Kabat *et al.*, 1987; y Tramontano *et al.*, 1990). Los dominios de región variable comprenden normalmente los aproximadamente 105-115 aminoácidos amino-terminales de una cadena de inmunoglobulina que se produce de manera natural (por ejemplo, aminoácidos 1-110), aunque también son adecuados dominios variables algo más cortos o más largos para formar anticuerpos de cadena sencilla. Las CDR forman parte de inmunoglobulinas que determinan la especificidad de dichas moléculas y entran en contacto con un ligando específico. Las CDR son la parte más variable de la molécula y contribuyen a la diversidad de estas moléculas. Hay tres regiones CDR, CDR1, CDR2 y CDR3, en cada dominio V. CDR-H representa una región CDR de una cadena pesada variable y CDR-L se refiere a una región CDR de una cadena ligera variable. H significa la cadena pesada variable y L significa la cadena ligera variable. Las regiones CDR de una región derivada de Ig pueden determinarse tal como se describe en Kabat (1991). Sequences of Proteins of Immunological Interest, 5ª edic., publicación de NIH n.º 91-3242 Departamento de Salud y Servicios Sociales de EE.UU., Chothia (1987) J. Mol. Biol. 196, 901-917 y Chothia (1989) Nature, 342, 877-883.

El término "completa" se usa en el presente documento para referirse a una técnica de evolución en la que se realizan todos los cambios posibles en cada posición de un polinucleótido molde o polipéptido molde y se somete a prueba el polinucleótido o polipéptido para confirmar que se han realizado los cambios pretendidos.

"Sustituciones de aminoácidos conservativas" se refieren a la capacidad de intercambio de residuos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales de hidroxilo alifáticas es serina y treonina; un grupo de aminoácidos que tienen amida que contienen cadenas laterales es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Grupos de sustituciones de aminoácidos conservativas preferidas son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparagina- glutamina.

El término "corresponde a" se usa en el presente documento para significar que una secuencia de polinucleótido es homóloga (es decir, es idéntica, no relacionada evolutivamente de manera estricta) a la totalidad o una parte de una secuencia de polinucleótido de referencia, o que una secuencia de polipéptido es idéntica a una secuencia de polipéptido de referencia. En contraposición, el término "complementaria a" se usa en el presente documento para significar que la secuencia complementaria es homóloga a la totalidad o una parte de una secuencia de polinucleótido de referencia. Para ilustración, la secuencia de nucleótidos "TATAC" corresponde a una secuencia "TATAC" de referencia y es complementaria a una secuencia de referencia "GTATA".

El término cantidad "efectiva de degradación" se refiere a la cantidad de enzima que se requiere para procesar al menos el 50% del sustrato, en comparación con el sustrato que no entró en contacto con la enzima. Preferiblemente, al menos el 80% del sustrato está degradado.

Tal como se usa en el presente documento, el término "región de entramado de secuencia definida" se refiere a un conjunto de secuencias definidas que se seleccionan con una base no aleatoria, generalmente basándose en datos experimentales o datos estructurales; por ejemplo, una región de entramado de secuencia definida puede comprender un conjunto de secuencias de aminoácidos que se predice que forman una estructura de lámina β o puede comprender un motivo de repetición de héptada de cremallera de leucina, un dominio de dedo de zinc, entre otras variaciones. Un "núcleo de secuencias definidas" es un conjunto de secuencias que engloban un alcance limitado de variabilidad. Mientras que (1) una secuencia decamérica completamente aleatoria de los 20 aminoácidos convencionales puede ser cualquiera de (20)¹⁰ secuencias, y (2) una secuencia decamérica pseudoaleatoria de los 20 aminoácidos convencionales puede ser cualquiera de (20)¹⁰ secuencias pero presentará un sesgo hacia determinados residuos en determinadas posiciones y/o en general, (3) un núcleo de secuencias definidas es un subconjunto de secuencias si se permitiese que cada posición de residuo fuese cualquiera de los 20 aminoácidos convencionales admisibles (y/o amino/iminoácidos no convencionales admisibles). Un núcleo de secuencias definidas comprende generalmente posiciones de residuo variantes e invariantes y/o comprende posiciones de residuo variantes que pueden comprender un residuo seleccionado de un subconjunto definido de residuos de

aminoácido), y similares, o bien de manera segmentada o bien por toda la longitud de la secuencia de miembro de biblioteca seleccionada individual. Los núcleos de secuencias definidas pueden referirse o bien a secuencias de aminoácidos o bien a secuencias de polinucleótido. Como ilustración y sin limitación, las secuencias (NNK)₁₀ y (NNM)₁₀, en las que N representa A, T, G, o C; K representa G o T; y M representa A o C, son núcleos de secuencias definidas.

El término “desinmunización” tal como se usa en el presente documento se refiere a la producción de una variante de la molécula de unión molde, que está modificada en comparación con una molécula de tipo natural original volviendo dicha variante no inmunogénica o menos inmunogénica en seres humanos. Las moléculas desinmunizadas según la invención se refieren a anticuerpos o partes de los mismos (como regiones de entramado y/o CDR) de origen no humano. Ejemplos correspondientes son anticuerpos o fragmentos de los mismos tal como se describen en el documento US 4.361.549. El término “desinmunizado” también se refiere a moléculas, que muestran una propensión reducida a generar epítomos de células T. Según esta invención, el término “propensión reducida a generar epítomos de células T” se refiere a la eliminación de epítomos de células T que conduce a la activación de células T específicas.

Además, propensión reducida a generar epítomos de células T significa sustitución de aminoácidos que contribuyen a la formación de epítomos de células T, es decir sustitución de aminoácidos, que son esenciales para la formación de un epítomo de células T. En otras palabras, propensión reducida a generar epítomos de células T se refiere a inmunogenicidad reducida o capacidad reducida para inducir proliferación de células T independiente de antígeno. Además, propensión reducida a generar epítomos de células T se refiere a desinmunización, que significa pérdida o reducción de posibles epítomos de células T de secuencias de aminoácidos que inducen proliferación de células T independiente de antígeno.

El término “epítomo de células T” tal como se usa en el presente documento se refiere a secuencias peptídicas cortas que pueden liberarse durante la degradación de péptidos, polipéptido o proteínas dentro de células y posteriormente presentarse por moléculas del complejo mayor de histocompatibilidad (CMH) para desencadenar la activación de células T; véase, entre otros, el documento WO 02/066514. Para péptidos presentados por CMH de clase II, tal activación de células T puede inducir entonces una respuesta de anticuerpos mediante la estimulación directa de células B para producir dichos anticuerpos.

“Digestión” de ADN se refiere a la escisión catalítica del ADN con una enzima de restricción que actúa sólo en determinadas secuencias en el ADN. Las diversas enzimas de restricción usadas en el presente documento están disponibles comercialmente y sus condiciones de reacción, cofactores y otros requisitos se usaron tal como conocerá el experto habitual en la técnica. Para fines analíticos, normalmente se usa 1 µg de plásmido o fragmento de ADN con 2 unidades de enzima en 20 µl de disolución tampón. Con el fin de aislar fragmentos de ADN para la construcción de plásmidos, normalmente se digieren de 5 a 50 µg de ADN con de 20 a 250 unidades de enzima en un mayor volumen. El fabricante especifica tampones y cantidades de sustrato apropiados para enzimas de restricción particulares. Habitualmente se usan tiempos de incubación de 1 hora a 37°C, pero pueden variar según las instrucciones del proveedor. Después de la digestión, se somete la reacción a electroforesis directamente sobre un gel para aislar el fragmento deseado.

El término “intercambio de ADN” se usa en el presente documento para indicar la recombinación entre secuencias sustancialmente homólogas pero no idénticas, en algunas realizaciones intercambio de ADN puede implicar cruzamiento a través de recombinación no homóloga, tal como a través de sistemas de cer/lox y/o flp/frt y similares. El intercambio de ADN puede ser aleatorio o no aleatorio.

Tal como se usa en esta invención, el término “epítomo” se refiere a un determinante antigénico en un antígeno, tal como un polipéptido de fitasa, al que se une el parátomo de un anticuerpo, tal como un anticuerpo específico de fitasa. Los determinantes antigénicos consisten habitualmente en agrupamientos de superficie químicamente activos de moléculas, tales como aminoácidos o cadenas laterales de azúcar, y pueden tener características estructurales tridimensionales específicas, así como características de cara específicas. Tal como se usa en el presente documento “epítomo” se refiere a aquella parte de un antígeno u otra macromolécula que puede formar una interacción de unión que interacciona con el cuerpo de unión de región variable de un anticuerpo. Normalmente, tal interacción de unión se manifiesta como un contacto intermolecular con uno o más residuos de aminoácido de una CDR.

El término “evolución” se refiere a un cambio en al menos una propiedad, característica o actividad de una proteína o anticuerpo modificado genética o sintéticamente en comparación con una proteína o anticuerpo molde.

Los términos “fragmento”, “derivado” y “análogo” cuando hacen referencia a un polipéptido comprenden un polipéptido que retiene al menos una función o actividad biológica que es al menos esencialmente igual que la del polipéptido de referencia. Además, los términos “fragmento”, “derivado” o “análogo” se ejemplifican mediante una molécula de “pro-forma”, tal como una pro-proteína de baja actividad que puede modificarse mediante escisión para producir una enzima madura con actividad significativamente mayor.

Se proporciona un método en el presente documento para producir a partir de una proteína molde un conjunto de

polipéptidos de progeñie en los que una “gama completa de sustituciones de aminoácidos individuales” está representada en cada posición de aminoácido. Tal como se usa en el presente documento, “gama completa de sustituciones de aminoácidos individuales” hace referencia a los 20 alfa-aminoácidos codificados de manera natural que forman polipéptidos codificados de manera natural, tal como se describe en el presente documento.

- 5 El término “gen” significa el segmento de ADN implicado en la producción de una cadena de polipéptido; incluye regiones que preceden y que siguen a la región codificante (secuencias líder y remolque) así como secuencias intermedias (intrones) entre segmentos codificantes individuales (exones).

10 “Inestabilidad genética”, tal como se usa en el presente documento, se refiere a la tendencia natural de secuencias altamente repetitivas a perderse a través de un procedimiento de acontecimientos reductores que implican generalmente simplificación de secuencia a través de la pérdida de secuencias repetidas. Las deleciones tienden a implicar la pérdida de una copia de una repetición y todo lo que haya entre las repeticiones.

15 El término “heteróloga” significa que una secuencia de ácido nucleico monocatenario no puede hibridar con otra secuencia de ácido nucleico monocatenario o su complemento. Por tanto, zonas de heterología significa que zonas de polinucleótidos o polinucleótidos tienen zonas o regiones dentro de su secuencia que no pueden hibridar con otro ácido nucleico o polinucleótido. Tales regiones o zonas son por ejemplo zonas de mutaciones.

20 El término “homóloga” o “homeóloga” significa que una secuencia de ácido nucleico de ácido nucleico monocatenario puede hibridar con una secuencia de ácido nucleico monocatenario complementaria. El grado de hibridación puede depender de varios factores incluyendo la cantidad de identidad entre las secuencias y las condiciones de hibridación tales como temperatura y concentraciones de sal tal como se comentará más adelante. Preferiblemente, la región de identidad es mayor de 5 pb, más preferiblemente la región de identidad es mayor de 10 pb.

25 El término “humanizado” se usa para describir anticuerpos en los que se combinan regiones determinantes de la complementariedad (CDR) de un animal mamífero, por ejemplo, un ratón, con una región de entramado humana. A menudo, se injertarán polinucleótidos que codifican para las CDR aisladas en polinucleótidos que codifican para una región de entramado de región variable adecuada (y opcionalmente regiones constantes) para formar polinucleótidos que codifican para anticuerpos completos (por ejemplo, humanizados o totalmente humanos), fragmentos de anticuerpo, y similares. En otro aspecto, además de anticuerpos de ratón, pueden humanizarse otras especies, tales como, por ejemplo, otros roedores, camello, conejo, gato, perro, cerdo, caballo, vaca, pez, llama y tiburón. En un aspecto amplio, puede utilizarse cualquier especie que produzca anticuerpos en la producción de anticuerpos humanizados. Adicionalmente, los anticuerpos de la invención pueden ser quiméricos, similares a humanos, humanizados o totalmente humanos, para reducir su posible antigenicidad, sin reducir su afinidad por su diana. Los anticuerpos quiméricos, similares a humanos y humanizados se han descrito generalmente en la técnica. Incorporando la menor cantidad posible de secuencia foránea en el anticuerpo híbrido, se reduce la antigenicidad. La preparación de estos anticuerpos híbridos puede llevarse a cabo mediante métodos bien conocidos en la técnica.

35 En un aspecto alternativo, se adaptan anticuerpos humanos o de ratón a una especie receptora diferente, tal como una especie amenazada, para proporcionar agentes terapéuticos para la especie receptora al tiempo que se protege frente a una respuesta inmunitaria negativa. En este aspecto, se utilizan las regiones de entramado de una especie receptora en combinación con las CDR de anticuerpos de una especie conocida o segunda especie.

40 Una región variable de cadena ligera o pesada de inmunoglobulina consiste en una región “de entramado” interrumpida por tres regiones hipervariables, también denominadas CDR. La extensión de la región de entramado y las CDR se ha definido de manera precisa (véase, “Sequences of Proteins of Immunological Interest”, Kabat *et al.*, 1987). Las secuencias de las regiones de entramado de diferentes cadenas ligeras o pesadas están relativamente conservadas dentro de una especie. Tal como se usa en el presente documento, una “región de entramado humana” es una región de entramado que es sustancialmente idéntica (85 o más, habitualmente 90-95 o más) a la región de entramado de una inmunoglobulina humana que se produce de manera natural. La región de entramado de un anticuerpo, es decir las regiones de entramado combinadas de las cadenas ligeras y pesadas constituyentes, sirve para posicionar y alinear las CDR. Las CDR son principalmente responsables de la unión a un epítipo de un antígeno. Según esta invención, una región de entramado se refiere a una región en el dominio V (dominio VH o VL) de inmunoglobulinas que proporciona un armazón proteico para las regiones determinantes de la complementariedad hipervariables (CDR) que entran en contacto con el antígeno. En cada dominio V, hay cuatro regiones de entramado designadas como FR1, FR2, FR3 y FR4. La región de entramado 1 engloba la región desde el extremo N-terminal del dominio V hasta el comienzo de CDR1, la región de entramado 2 se refiere a la región entre CDR1 y CDR2, la región de entramado 3 engloba la región entre CDR2 y CDR3 y la región de entramado 4 significa la región desde el final de CDR3 hasta el extremo C-terminal del dominio V; véase, entre otros, Janeway, Immunobiology, Garland Publishing, 2001, 5ª ed. Por tanto, las regiones de entramado engloban todas las regiones fuera de las regiones CDR en los dominios VH o VL.

El experto en la técnica se encuentra fácilmente en condiciones de deducir a partir de una secuencia dada las regiones de entramado y las CDR; véanse Kabat (1991) Sequences of Proteins of Immunological Interest, 5ª edic., publicación de NIH n.º 91-3242 Departamento de Salud y Servicios Sociales de EE.UU., Chothia (1987) J. Mol. Biol.

196, 901-917 y Chothia (1989) Nature, 342, 877-883.

Los beneficios de esta invención se extienden a "aplicaciones industriales" (o procesos industriales), término que se usa para incluir aplicaciones en la industria comercial propiamente dicha (o simplemente industria) así como aplicaciones industriales no comerciales (por ejemplo, investigación biomédica en una institución sin ánimo de lucro). Las aplicaciones relevantes incluyen aquellas en las áreas del diagnóstico, la medicina, agricultura, fabricación y enseñanza.

El término "idénticas" o "identidad" significa que dos secuencias de ácido nucleico tienen la misma secuencia o una secuencia complementaria. Por tanto, "zonas de identidad" significa que regiones o zonas de un polinucleótido o el polinucleótido global son idénticas o complementarias a zonas de otro polinucleótido o el polinucleótido.

El término "aislado" significa que el material se retira de su entorno original (por ejemplo, el entorno natural si se produce de manera natural). Por ejemplo, una enzima o un polinucleótido que se produce de manera natural o presente en un animal vivo no está aislado, pero el mismo polinucleótido o enzima, separado de parte o de la totalidad de los materiales coexistente en el sistema natural, está aislado. Tales polinucleótidos podrían formar parte de un vector y/o tales polinucleótidos o enzimas podrían formar parte de una composición, y todavía estar aislados porque tal vector o composición no forma parte de su entorno natural.

Por "ácido nucleico aislado" quiere decirse un ácido nucleico, por ejemplo, una molécula de ADN o ARN, que no es inmediatamente contiguo a las secuencias flanqueantes en 5' y 3' a las que normalmente es inmediatamente contiguo cuando está presente en el genoma que se produce de manera natural del organismo del que se deriva. El término describe por tanto, por ejemplo, un ácido nucleico que se incorpora en un vector, tal como un plásmido o vector viral; un ácido nucleico que se incorpora en el genoma de una célula heteróloga (o el genoma de una célula homóloga, pero en un sitio diferente del que se produce de manera natural); y un ácido nucleico que existe como una molécula independiente, por ejemplo, un fragmento de ADN producido mediante amplificación por PCR o digestión con enzimas de restricción, o una molécula de ARN producida mediante transcripción *in vitro*. El término también describe un ácido nucleico recombinante que forma parte de un gen híbrido que codifica para secuencias de polipéptido adicionales que pueden usarse, por ejemplo, en la producción de una proteína de fusión.

Como se usa en este documento, "ligando" se refiere a una molécula, tal como un péptido aleatorio o secuencia de segmento variable, que es reconocida por un receptor particular. Como reconocerá un experto en la técnica, una molécula (o complejo macromolecular) puede ser tanto un receptor como un ligando. En general, la pareja de unión que tiene un peso molecular menor se denomina ligando y la pareja de unión que tiene un peso molecular mayor se denomina receptor.

"Ligamiento" se refiere al procedimiento de formación de enlaces fosfodiéster entre dos fragmentos de ácido nucleico bicatenarios (Maniatis *et al.*, 1982, pág. 146). A menos que se proporcione de otro modo, puede lograrse el ligamiento usando condiciones y tampones conocidos con 10 unidades de ADN ligasa de T4 ("ligasa") por 0,5 µg de cantidades aproximadamente equimolares de los fragmentos de ADN que van a ligarse.

Tal como se usa en el presente documento, "ligador" o "espaciador" se refiere a una molécula o grupo de moléculas que conecta dos moléculas, tales como una proteína de unión a ADN y un péptido aleatorio, y sirve para colocar las dos moléculas en una configuración preferida, por ejemplo, de modo que el péptido aleatorio puede unirse a un receptor con impedimento estérico mínimo de la proteína de unión a ADN.

El término "presentación en la superficie de una célula de mamífero" se refiere a una técnica mediante la cual una proteína o anticuerpo, o una parte de un anticuerpo, se expresa y se presenta en la superficie de una célula huésped de mamífero con fines de examen; por ejemplo, mediante examen para determinar la unión específica a antígeno mediante una combinación de perlas magnéticas y clasificación de células activada por fluorescencia. En un aspecto, se usan vectores de expresión de mamífero para la expresión simultánea de inmunoglobulinas tanto como una forma secretada como una forma unida a la superficie celular como en DuBridge *et al.*, documento US 2009/0136950. En otro aspecto, se emplean las técnicas de Gao *et al.* para un vector viral que codifica para una biblioteca de anticuerpos o fragmentos de anticuerpo se presentan en las membranas celulares cuando se expresan en una célula como en Gao *et al.*, documento US 2007/0111260. Se conoce la presentación en superficie de IgG completa en células de mamífero. Por ejemplo, Akamatsu *et al.* desarrollaron un vector de presentación en la superficie de una célula de mamífero, adecuado para aislar directamente moléculas de IgG basándose en su afinidad de unión a antígeno y actividad biológica. Usando un vector episómico derivado de virus de Epstein-Barr, se presentaron bibliotecas de anticuerpos como moléculas de IgG completas en la superficie celular y se examinaron para determinar la unión específica a antígeno mediante una combinación de perlas magnéticas y clasificación de células activada por fluorescencia. Se recuperaron los plásmidos que codificaban para anticuerpos con características de unión deseadas a partir de las células clasificadas y se convirtieron en la forma para la producción de IgG soluble. Akamatsu *et al.* J. Immunol. Methods 2007 327(1-2):40-52. Ho *et al.* usaron células de riñón embrionario humano 293T que se usan ampliamente para la expresión transitoria de proteínas para la presentación en la superficie celular de anticuerpos de cadena sencilla Fv para maduración de afinidad. Se enriquecieron las células que expresaban un anticuerpo mutante poco común con mayor afinidad, 240 veces mediante una clasificación celular de un solo paso a partir de un gran exceso de células que expresan anticuerpo WT con una

afinidad ligeramente menor. Además, se obtuvo un mutante altamente enriquecido con aumento de la afinidad de unión por CD22 después de una selección individual de una biblioteca combinatoria aleatorizando un punto caliente de anticuerpo intrínseco. Ho *et al.* Isolation of anti-CD22 Fv with high affinity by Fv display on human cells, Proc Natl Acad Sci U S A 20 de junio de 2006; 103(25): 9637-9642.

- 5 Beerli *et al.* usaron células B específicas para un antígeno de interés que se aislaron directamente de células mononucleares de sangre periférica (PBMC) de donantes humanos. Se generan bibliotecas de Fv de cadena sencilla (scFv) recombinantes, específicos de antígeno a partir de esta combinación de células B y se examinan mediante presentación en la superficie de una célula de mamífero usando un sistema de expresión de virus Sindbis. Este método permite aislar anticuerpos específicos de antígeno mediante una tanda individual de FACS. Las regiones variables (VR) de las cadenas pesadas (HC) y cadenas ligeras (LC) se aislaron de clones positivos y anticuerpos totalmente humanos recombinantes producidos como fragmentos Fab o IgG completas. De esta manera, se aislaron varios anticuerpos de alta afinidad hipermutados que se unen a la partícula pseudoviral (VLP) Q β , un antígeno viral modelo, así como anticuerpos específicos para nicotina. Todos los anticuerpos mostraron altos niveles de expresión en cultivo celular. Los Acm específicos de nicotina humanos se validaron de manera preclínica en un modelo de ratón. Beerli *et al.*, Isolation of human monoclonal antibodies by mammalian cell display, Proc Natl Acad Sci USA. 23 de septiembre de 2008; 105(38): 14336-14341.

También se conoce la presentación en la superficie celular de levaduras, por ejemplo, véanse Kondo y Ueda 2004, Yeast cell-surface display-applications of molecular display, Appl. Microbiol. Biotechnol., 64(1): 28-40, que describe por ejemplo, un sistema de modificación mediante ingeniería de la superficie celular usando la levadura *Saccharomyces cerevisiae*. Varios sistemas de presentación representativos para la expresión en la levadura *S. cerevisiae* se describen en Lee *et al.*, 2003, Microbial cell-surface display, TRENDS in Biotechnol. 21(1): 45-52. También Boder y Wittrup 1997, Yeast surface display for screening combinatorial polypeptide libraries, Nature Biotechnol., 15(6): 553.

El término “fabricación” se refiere a la producción de una proteína en una cantidad suficiente como para permitir al menos pruebas clínicas de fase I de una proteína terapéutica, o una cantidad suficiente para la aprobación reguladora de una proteína de diagnóstico.

El término “mutación de cambio de sentido” se refiere a una mutación puntual en la que se cambia un nucleótido individual, lo que da como resultado un codón que codifica para un aminoácido diferente. Las mutaciones que cambian un aminoácido a un codón de terminación se denominan mutaciones sin sentido.

30 Tal como se usa en el presente documento, una “propiedad molecular que va a hacerse evolucionar” incluye la referencia a moléculas que se componen de una secuencia de polinucleótido, moléculas que se componen de una secuencia de polipéptido, y moléculas que se componen en parte de una secuencia de polinucleótido y en parte de una secuencia de polipéptido. Ejemplos particularmente relevantes, pero no limitativos en modo alguno, de propiedades moleculares que van a hacerse evolucionar incluyen actividades enzimáticas en condiciones especificadas, tales como relacionadas con la temperatura; salinidad; presión; pH; y concentración de glicerol, DMSO, detergente, y/o cualquier otra especie molecular con la que entra en contacto en un entorno de reacción. Ejemplos particularmente relevantes adicionales, pero no limitativos en modo alguno, de propiedades moleculares que van a hacerse evolucionar incluyen estabilidades, por ejemplo, la cantidad de una propiedad molecular residual que está presente después de un tiempo de exposición especificado a un entorno especificado, tal como puede encontrarse durante el almacenamiento.

El término “mutar” se refiere a crear una mutación en una secuencia de ácido nucleico; en el caso en el que la mutación se produce dentro de la región codificante de una proteína, lo que conducirá a un cambio de codón que puede conducir o no a un cambio de aminoácido.

45 El término “mutaciones” significa cambios en la secuencia de una secuencia de ácido nucleico de tipo natural o cambios en la secuencia de un péptido o polipéptidos. Tales mutaciones pueden ser mutaciones puntuales tales como transiciones o transversiones. Las mutaciones pueden ser deleciones, inserciones o duplicaciones.

Tal como se usa en el presente documento, la secuencia de nucleótidos “N,N,G/T” degenerada representa 32 posibles tripletes, donde “N” puede ser A, C, G o T.

50 Tal como se usa en el presente documento, la secuencia de nucleótidos “N,N,N” degenerada representa 64 posibles tripletes, donde “N” puede ser A, C, G o T.

El término “que se produce de manera natural” tal como se usa en el presente documento aplicado al objeto se refiere al hecho de que un objeto puede encontrarse en la naturaleza. Por ejemplo, un polipéptido o secuencia de polinucleótido que está presente en un organismo (incluyendo virus) que puede aislarse de una fuente en la naturaleza y que no se ha modificado de manera intencionada por el hombre en el laboratorio, se produce de manera natural. Generalmente, el término se produce de manera natural se refiere a un objeto presente en un individuo no patológico (no enfermo), tal como sería típico para la especie.

Tal como se usa en el presente documento, una “molécula de ácido nucleico” se compone de al menos una base o

un par de bases, dependiendo de si es monocatenaria o bicatenaria, respectivamente. Además, una molécula de ácido nucleico puede pertenecer exclusivamente o de manera quimérica a cualquier grupo de moléculas que contienen nucleótidos, tal como se ejemplifica mediante, pero sin limitarse a, los siguientes grupos de moléculas de ácido nucleico: ARN, ADN, ácidos nucleicos genómicos, ácidos nucleicos no genómicos, ácidos nucleicos que se producen de manera natural y que no se producen de manera natural, y ácidos nucleicos sintéticos. Esto incluye, a modo de ejemplo no limitativo, ácidos nucleicos asociados con cualquier orgánulo, tal como la mitocondria, ARN ribosómico, y moléculas de ácido nucleico que se componen de manera quimérica de uno o más componentes que no se producen de manera natural junto con componentes que se producen de manera natural.

Adicionalmente, una "molécula de ácido nucleico" puede contener en parte uno o más componentes no basados en nucleótidos tal como se ejemplifica mediante, pero sin limitarse a, aminoácidos y azúcares. Por tanto, a modo de ejemplo, pero no de limitación, una ribozima que se basa en parte en nucleótidos y se basa en parte en proteína se considera una "molécula de ácido nucleico".

Además, a modo de ejemplo, pero no de limitación, una molécula de ácido nucleico que se marca con un resto detectable, tal como un marcador radiactivo o alternativamente un marcador no radiactivo, se considera asimismo una "molécula de ácido nucleico".

Los términos "secuencia de ácido nucleico que codifica para" o una "secuencia codificante de ADN de" o una "secuencia de nucleótidos que codifica para" una proteína particular, así como otros términos sinónimos, se refieren a una secuencia de ADN que se transcribe y se traduce para dar una enzima cuando se coloca bajo el control de secuencias reguladoras apropiadas. Una "secuencia promotora" es una región reguladora de ADN que puede unirse a ARN polimerasa en una célula e iniciar la transcripción de una secuencia codificante posterior (en el sentido de 3'). El promotor forma parte de la secuencia de ADN. Esta región de secuencia tiene un codón de iniciación en su extremo 3'. La secuencia promotora incluye el número mínimo de bases cuando hay los elementos necesarios para iniciar la transcripción a niveles detectables por encima del fondo. Sin embargo, después de unirse la ARN polimerasa a la secuencia e iniciarse la transcripción en el codón de iniciación (extremo 3' con un promotor), avanza la transcripción de manera posterior en el sentido de 3'. Dentro de la secuencia promotora se encontrará un sitio de inicio de la transcripción (definido de manera conveniente mediante mapeo con la nucleasa S1) así como dominios de unión a proteína (secuencias consenso) responsables de la unión de la ARN polimerasa.

Los términos "ácido nucleico que codifica para una proteína" o "ADN que codifica para una proteína" o "polinucleótido que codifica para una proteína" y otros términos sinónimos engloban un polinucleótido que incluye sólo la secuencia codificante para la proteína así como un polinucleótido que incluye secuencia codificante y/o no codificante adicional.

En una realización preferida, una "especie de molécula de ácido nucleico específica" se define mediante su estructura química, tal como se ejemplifica mediante, pero sin limitarse a, su secuencia primaria. En otra realización preferida, una "especie de molécula de ácido nucleico" específica se define mediante una función de la especie de ácido nucleico o por una función de un producto derivado de la especie de ácido nucleico. Por tanto, a modo de ejemplo no limitativo, una "especie de molécula de ácido nucleico específica" pueden definirse mediante una o más actividades o propiedades atribuibles a la misma, incluyendo actividades o propiedades atribuibles a su producto expresado.

La presente definición de "ensamblar una muestra de ácido nucleico de trabajo en una biblioteca de ácidos nucleicos" incluye el procedimiento de incorporar una muestra de ácido nucleico en una colección basada en vector, tal como mediante ligamiento en un vector y transformación de un huésped. A continuación en el presente documento se proporciona una descripción de vectores, huéspedes y otros reactivos relevantes así como ejemplos no limitativos específicos de los mismos. La presente definición de "ensamblar una muestra de ácido nucleico de trabajo en una biblioteca de ácidos nucleicos" también incluye el procedimiento de incorporar una muestra de ácido nucleico en una colección no basada en vector, tal como mediante ligamiento a adaptadores. Preferiblemente, los adaptadores pueden aparearse con cebadores de PCR para facilitar la amplificación por PCR.

Por consiguiente, en una realización no limitativa, una "biblioteca de ácidos nucleicos" se compone de una colección basada en vector de una o más moléculas de ácido nucleico. En otra realización preferida, una "biblioteca de ácidos nucleicos" se compone de una colección no basada en vector de moléculas de ácido nucleico. En aún otra realización preferida, una "biblioteca de ácidos nucleicos" se compone de una colección combinada de moléculas de ácido nucleico que se basa en parte en vector y no se basa en parte en vector. Preferiblemente, la colección de moléculas que comprenden una biblioteca puede someterse a búsqueda y separación según la especie de molécula de ácido nucleico individual.

La presente invención proporciona un "constructo de ácido nucleico" o alternativamente a "constructo de nucleótidos" o alternativamente un "constructo de ADN". El término "constructo" se usa en el presente documento para describir una molécula, tal como un polinucleótido (por ejemplo, un polinucleótido de fitasa) que puede opcionalmente unirse químicamente a uno o más restos moleculares adicionales, tales como un vector, o partes de un vector. En un aspecto específico, pero no limitativo en modo alguno, un constructo de nucleótidos se ejemplifica mediante un constructo de expresión de ADN adecuado para la transformación de una célula huésped.

Un “oligonucleótido” (o de manera sinónima un “oligo”) se refiere o bien a un polidesoxinucleótido monocatenario o bien a dos hebras de polidesoxinucleótido complementarias que pueden sintetizarse químicamente. Tales oligonucleótidos sintéticos pueden tener o no un 5'-fosfato. Los que no lo tienen no se ligarán con otro oligonucleótido sin añadir un fosfato con un ATP en presencia de una cinasa. Un oligonucleótidosintético se ligará con un fragmento que no se ha desfosforilado. Para lograr la amplificación basada en polimerasa (tal como con PCR), se mencionan un “oligonucleótido degenerado 32 veces que se compone de, en serie, al menos una primera secuencia homóloga, una secuencia N,N,G/T degenerada, y una segunda secuencia homóloga”. Tal como se usa en este contexto, “homóloga” hace referencia a homología entre el oligo y el polinucleótido parental que se somete a la amplificación basada en polimerasa.

Tal como se usa en el presente documento, el término “operativamente unido” se refiere a un ligamiento de elementos de polinucleótido en una relación funcional. Un ácido nucleico está “operativamente unido” cuando se coloca en una placid relación funcional con otra secuencia de ácido nucleico. Por ejemplo, un promotor o potenciador está operativamente unido a una secuencia codificante si afecta a la transcripción de la secuencia codificante. Operativamente unido significa que las secuencias de ADN que se unen son normalmente contiguas y, cuando sea necesario unir dos regiones codificantes de proteína, contiguas y en el marco de lectura.

Una secuencia codificante está “operativamente unida a” otra secuencia codificante cuando la ARN polimerasa transcribirá las dos secuencias codificantes en un ARNm individual, que entonces se transcribe para dar un polipéptido individual que tiene aminoácidos derivados de ambas secuencias codificantes. No es necesario que las secuencias codificantes sean contiguas entre sí siempre que las secuencias expresadas se procesen en última instancia para producir la proteína deseada.

Tal como se usa en el presente documento el término “condiciones fisiológicas” se refieren a temperatura, pH, fuerza iónica, viscosidad, y parámetros bioquímicos similares que son compatibles con un organismo viable, y/o que normalmente existen de manera intracelular en una célula de mamífero o célula de levadura cultivada viable. Por ejemplo, las condiciones intracelulares en una célula de levadura hecha crecer en condiciones de cultivo de laboratorio típicas son condiciones fisiológicas. Condiciones de reacción *in vitro* adecuadas para cócteles de transcripción *in vitro* son generalmente condiciones fisiológicas. En general, las condiciones fisiológicas *in vitro* comprenden NaCl o KCl 50-200 mM, pH 6,5-8,5, 20-45°C y catión divalente (por ejemplo, Mg⁺⁺, Ca⁺⁺) 0,001-10 mM; preferiblemente NaCl o KCl 150 mM, pH 7,2-7,6, catión divalente 5 mM, y a menudo incluyen el 0,01-1,0 por ciento de proteína no específica (por ejemplo, BSA). A menudo puede estar presente un detergente no iónico (Tween, NP-40, Triton X-100), habitualmente a del 0,001 al 2%, normalmente al 0,05-0,2% (v/v). Pueden seleccionarse condiciones acuosas particulares por el profesional según métodos convencionales. Para una orientación general, pueden aplicarse las siguientes condiciones acuosas tamponadas: NaCl 10-250 mM, Tris HCl 5-50 mM, pH 5-8, con adición opcional de catión/cationes divalente(s) y/o quelantes metálicos y/o detergentes no iónicos y/o fracciones de membrana y/o agentes antiespumantes y/o centelleadores.

El término “población” tal como se usa en el presente documento significa una colección de componentes tales como polinucleótidos, partes de polinucleótidos o proteínas. Una “población mixta” significa una colección de componentes que pertenecen a la misma familia de ácidos nucleicos o proteínas (es decir, están relacionados) pero que difieren en su secuencia (es decir, no son idénticos) y así en su actividad biológica.

Una molécula que tiene una “pro-forma” se refiere a una molécula que experimenta cualquier combinación de una o más modificaciones químicas covalentes y no covalentes (por ejemplo, glicosilación, escisión proteolítica, dimerización u oligomerización, cambio conformacional inducido por temperatura o pH, asociación con un cofactor, etc.) como vía para lograr una forma molecular más madura que tiene una diferencia de propiedad (por ejemplo un aumento de la actividad) en comparación con la molécula de pro-forma de referencia. Cuando pueden distinguirse dos o más modificaciones químicas (por ejemplo dos escisiones proteolíticas, o una escisión proteolítica y una desglicosilación) como vía para la producción de una molécula madura, la molécula precursora de referencia puede denominarse molécula de “pre-pro-forma”.

Una “propiedad” puede describir cualquier característica, incluyendo una propiedad característica física, química o de actividad de una proteína que va a optimizarse. Por ejemplo, en determinados aspectos, la propiedad, característica o actividad predeterminada que va a optimizarse puede seleccionarse de reducción de la agregación proteína-proteína, potenciación de la estabilidad de proteína, aumento de la solubilidad de proteína, aumento de la estabilidad al pH de proteína, aumento de la estabilidad a la temperatura de proteína, aumento de la estabilidad al disolvente de proteína, aumento de la selectividad, disminución de la selectividad, introducción de sitios de glicosilación, introducción de sitios de conjugación, reducción de la inmunogenicidad, potenciación de la expresión de proteínas, aumento de la afinidad antigénica, disminución de la afinidad antigénica, cambio en la afinidad de unión, cambio en la inmunogenicidad, cambio en la actividad catalítica, optimización del pH o potenciación de la especificidad. Una propiedad “optimizada” se refiere a un cambio deseable en una propiedad particular en una proteína, anticuerpo o célula mutante en comparación con una proteína, anticuerpo, o célula molde, respectivamente.

Tal como se usa en el presente documento, el término “pseudoaleatorio” se refiere a un conjunto de secuencias que tienen variabilidad limitada, de manera que, por ejemplo, el grado de variabilidad de residuos en otra posición,

excepto cualquier posición pseudoaleatoria, se permite cierto grado de variación de residuos, sin embargo circunscrito.

5 “Unidades casi de repetición”, tal como se usa en el presente documento, se refiere a las repeticiones que van a reordenarse y son por definición no idénticas. En efecto, el método se propone no sólo para unidades codificantes prácticamente idénticas producidas mediante mutagénesis de la secuencia de partida idéntica, sino también la reordenación de secuencias similares o relacionadas que pueden divergir significativamente en algunas regiones. No obstante, si las secuencias contienen suficientes homologías como para reordenarse mediante este enfoque, pueden denominarse unidades “casi de repetición”.

10 Tal como se usa en el presente documento “biblioteca de péptidos aleatorios” se refiere a un conjunto de secuencias de polinucleótido que codifican para un conjunto de péptidos aleatorios, y al conjunto de péptidos aleatorios codificados por esas secuencias de polinucleótido, así como a las proteínas de fusión que contienen esos péptidos aleatorios.

15 Tal como se usa en el presente documento, “secuencia peptídica aleatoria” se refiere a una secuencia de aminoácidos que se compone de dos o más monómeros de aminoácidos y se construye mediante un procedimiento estocástico o aleatorio. Un péptido aleatorio puede incluir motivos de almacén o región de entramado, que pueden comprender secuencias invariantes.

20 Tal como se usa en el presente documento, “receptor” se refiere a una molécula que tiene una afinidad por un ligando dado. Los receptores pueden ser moléculas que se producen de manera natural o sintéticos. Los receptores pueden emplearse en un estado inalterado o como agregados con otras especies. Los receptores pueden unirse, de manera covalente o no covalente, a un miembro de unión, o bien directamente o bien a través de una sustancia de unión específica. Los ejemplos de receptores incluyen, pero no se limitan a, anticuerpos, incluyendo anticuerpos monoclonales y antisueros reactivos con determinantes antigénicos específicos (tales como en virus, células, u otros materiales), receptores de la membrana celular, carbohidratos complejos y glicoproteínas, enzimas y receptores hormonales.

25 Proteínas “recombinantes” se refieren a proteínas producidos mediante técnicas de ADN recombinantes, es decir, producidos a partir de células transformadas mediante un constructo de ADN exógeno que codifica para la enzima deseada. Proteínas “sintéticas” son las preparadas mediante síntesis química.

El término “polinucleótidos relacionados” significa que regiones o zonas de los polinucleótidos son idénticas y regiones o zonas de los polinucleótidos son heterólogas.

30 “Reordenación reductora”, tal como se usa en el presente documento, se refiere al aumento de la diversidad molecular que se acumula a través de acontecimientos de delección (y/o inserción) que están mediados por secuencias repetidas.

35 Los siguientes términos se usan para describir las relaciones de secuencia entre dos o más polinucleótidos: “secuencia de referencia”, “ventana de comparación”, “identidad de secuencia”, “porcentaje de identidad de secuencia” e “identidad sustancial”.

40 Una “secuencia de referencia” es una secuencia definida usada como base para una comparación de secuencias; una secuencia de referencia puede ser un subconjunto de una secuencia más grande, por ejemplo, como un segmento de un ADNc de longitud completa o secuencia génica dada en una lista de secuencias, o puede comprender una secuencia génica o de ADNc completa. Generalmente, una secuencia de referencia tiene al menos 20 nucleótidos de longitud, frecuentemente al menos 25 nucleótidos de longitud, y a menudo al menos 50 nucleótidos de longitud. Puesto que dos polinucleótidos pueden comprender cada uno (1) una secuencia (es decir, una parte de la secuencia de polinucleótido completa) que es similar entre los dos polinucleótidos y (2) pueden comprender además una secuencia que es divergente entre los dos polinucleótidos, normalmente se realizan comparaciones de secuencias entre dos (o más) polinucleótidos comparando las secuencias de los dos polinucleótidos a lo largo de una “ventana de comparación” para identificar y comparar regiones locales de similitud de secuencia.

“El índice de repetición (RI)”, tal como se usa en el presente documento, es el número promedio de copias de las unidades casi de repetición contenidas en el vector de clonación.

50 El término “saturación” se refiere a una técnica de evolución en la que se realiza cada cambio posible en cada posición de un polinucleótido molde o polipéptido molde; sin embargo el cambio en cada posición no se confirma mediante pruebas, sino que meramente se asume estadísticamente cuando se estima que se produce la mayoría de los posibles cambios o casi cada cambio posible en cada posición de un molde.

55 El término “identidad de secuencia” significa que dos secuencias de polinucleótido son idénticas (es decir, nucleótido por nucleótido) a lo largo de la ventana de comparación. El término “porcentaje de identidad de secuencia” se calcula comparando dos secuencias alineadas de manera óptima a lo largo de la ventana de comparación, determinando el número de posiciones en las que aparece la base de ácido nucleico idéntica (por ejemplo, A, T, C,

G, U o I) en ambas secuencias para proporcionar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes entre el número total de posiciones en la ventana de comparación (es decir, el tamaño de ventana), y multiplicando el resultado por 100 para proporcionar el porcentaje de identidad de secuencia. Esta "identidad sustancial", tal como se usa en el presente documento, indica una característica de una secuencia de polinucleótido, en la que el polinucleótido comprende una secuencia que tiene una identidad de secuencia de al menos el 80 por ciento, preferiblemente una identidad de al menos el 85 por ciento, a menudo una identidad de secuencia del 90 al 95 por ciento, y de la manera más común una identidad de secuencia de al menos el 99 por ciento en comparación con una secuencia de referencia de una ventana de comparación de al menos 25-50 nucleótidos, en la que el porcentaje de identidad de secuencia se calcula comparando la secuencia de referencia con la secuencia de polinucleótido que puede incluir deleciones o adiciones que suman el 20 por ciento o menos de la secuencia de referencia a lo largo de la ventana de comparación.

El término "mutación silenciosa" se refiere a un cambio de codón que no da como resultado un cambio de aminoácido en un polipéptido expresado y se basan en la redundancia de uso de codones para la inserción de aminoácidos.

Tal como se conoce en la técnica, la "similitud" entre dos enzimas se determina comparando la secuencia de aminoácidos y sus sustitutos de aminoácido conservados de una proteína con la secuencia de una segunda proteína. Puede determinarse la similitud mediante procedimientos que se conocen bien en la técnica, por ejemplo, un programa BLAST (Basic Local Alignment Search Tool (herramienta de búsqueda de alineaciones locales básicas) en el Centro Nacional de Información Biológica).

Tal como se usa en el presente documento, el término "anticuerpo de cadena sencilla" se refiere a un polipéptido que comprende un dominio VH y un dominio VL en ligamiento de polipéptidos, generalmente ligados a través de un péptido espaciador (por ejemplo, [Gly-Gly-Gly-Gly-Ser]_x), y que puede comprender secuencias de aminoácidos adicionales en los extremos amino y/o carboxilo terminales. Por ejemplo, un anticuerpo de cadena sencilla puede comprender un segmento de anclaje para la unión al polinucleótido codificante. Como ejemplo, un scFv es un anticuerpo de cadena sencilla. Los anticuerpos de cadena sencilla son generalmente proteínas que consisten en uno o más segmentos de polipéptido de al menos 10 aminoácidos contiguos codificados sustancialmente por genes de la superfamilia de inmunoglobulinas (por ejemplo, véase Williams y Barclay, 1989, págs. 361-368), codificados lo más frecuentemente por una secuencia génica de cadena pesada o cadena ligera de roedor, primate no humano, de ave de corral, porcina, bovina, ovina, de cabra o humana. Un anticuerpo de cadena sencilla funcional contiene generalmente una parte suficiente de un producto génico de la superfamilia de inmunoglobulinas de modo que se retenga la propiedad de unión a una molécula diana específica, normalmente un receptor o antígeno (epítipo).

Se dice que los miembros de un par de moléculas (por ejemplo, un par anticuerpo-antígeno o un par de ácidos nucleicos) se "unen específicamente" entre sí si se unen entre sí con mayor afinidad que a otras moléculas no específicas. Por ejemplo, un anticuerpo producido contra un antígeno al que se une más eficazmente que a una proteína no específica puede describirse como que se une específicamente al antígeno. (De manera similar, una sonda de ácido nucleico se puede describir como específicamente unida a un objetivo de ácido nucleico si forma un dúplex específico con el objetivo mediante interacciones de emparejamiento de bases (véase arriba)).

"Hibridación específica" se define en el presente documento como la formación de híbridos entre un primer polinucleótido y un segundo polinucleótido (por ejemplo, un polinucleótido que tiene una secuencia distinta pero sustancialmente idéntica al primer polinucleótido), en la que secuencias de polinucleótido sustancialmente no relacionadas no forman híbridos en la mezcla.

El término "polinucleótido específico" significa un polinucleótido que tiene determinados puntos finales y que tiene una determinada secuencia de ácido nucleico. Dos polinucleótidos en los que un polinucleótido tiene la secuencia idéntica como parte del segundo polinucleótido pero diferentes extremos comprenden dos polinucleótidos específicos diferentes.

"Condiciones de hibridación rigurosas" significa que se producirá hibridación sólo si hay una identidad de al menos el 90%, preferiblemente una identidad de al menos el 95% y lo más preferiblemente una identidad de al menos el 97% entre las secuencias. Véase Sambrook *et al.*, 1989.

También se incluyen en la invención polipéptidos que tienen secuencias que son "sustancialmente idénticas" a la secuencia de un polipéptido, tal como una de cualquier SEQ ID NO dada a conocer en el presente documento. Una secuencia de aminoácidos "sustancialmente idéntica" es una secuencia que difiere de una secuencia de referencia sólo en sustituciones de aminoácidos conservativas, por ejemplo, sustituciones de un aminoácido por otro de la misma clase (por ejemplo, sustitución de un aminoácido hidrófobo, tal como isoleucina, valina, leucina o metionina, por otro, o sustitución de un aminoácido polar por otro, tal como sustitución de arginina por lisina, ácido glutámico por ácido aspártico, o glutamina por asparagina).

Adicionalmente una secuencia de aminoácidos "sustancialmente idéntica" es una secuencia que difiere de una secuencia de referencia en una o más sustituciones, deleciones o inserciones no conservativas, particularmente cuando se produce una sustitución de este tipo en un sitio que no es el sitio activo de la molécula, y siempre que el

5 polipéptido conserve esencialmente sus propiedades de comportamiento. Por ejemplo, pueden deleccionarse uno o más aminoácidos de un polipéptido de fitasa, lo que da como resultado la modificación de la estructura del polipéptido, sin alterar significativamente su actividad biológica. Por ejemplo, pueden eliminarse aminoácidos amino o carboxilo terminales que no se requieren para la actividad biológica de fitasa. Tales modificaciones pueden dar como resultado el desarrollo de polipéptidos de fitasa activos más pequeños.

10 La presente invención proporciona una "enzima sustancialmente pura". El término "enzima sustancialmente pura" se usa en el presente documento para describir una molécula, tal como un polipéptido (por ejemplo, un polipéptido de fitasa, o un fragmento del mismo) que está sustancialmente libre de otras proteínas, lípidos, carbohidratos, ácidos nucleicos y otros materiales biológicos con lo cual se asocia naturalmente. Por ejemplo, una molécula sustancialmente pura, tal como un polipéptido, puede ser al menos 60%, en peso seco, la molécula de interés. La pureza de los polipéptidos se puede determinar usando métodos estándar que incluyen, por ejemplo, electroforesis en gel de poliacrilamida (por ejemplo, SDS-PAGE), cromatografía en columna (por ejemplo, cromatografía líquida de alto rendimiento (HPLC)) y análisis de secuencia de aminoácidos amino-terminal.

15 Tal como se usa en el presente documento, "sustancialmente puro" significa que una especie objeto es la especie predominante presente (es decir, en una base molar es más abundante que cualquier otra especie macromolecular individual en la composición), y de manera preferible, fracción sustancialmente purificada es una composición en la que la especie objeto comprende al menos el 50 por ciento (en una base molar) de todas las especies macromoleculares presentes. Generalmente, una composición sustancialmente pura comprenderá más del 80 al 90 por ciento de todas las especies macromoleculares presentes en la composición. Lo más preferiblemente, la especie objeto se purifica hasta homogeneidad esencial (no pueden detectarse especies contaminantes en la composición mediante métodos de detección convencionales) en la que la composición consiste esencialmente en una especie macromolecular individual. Las especies de disolvente, moléculas pequeñas (<500 Dalton) y especies iónicas elementales no se consideran especies macromoleculares.

20 Tal como se usa en el presente documento, "oligopéptido molde" significa una proteína molde para el que se desea una biblioteca secundaria de variantes. Tal como apreciarán los expertos en la técnica, cualquier número de moldes encuentra uso en la presente invención. Se incluyen específicamente dentro de la definición de "proteínas" u "oligopéptidos" fragmentos y dominios de proteínas conocidas, incluyendo dominios funcionales tales como dominios enzimáticos dominios de unión, etc., y fragmentos más pequeños, tales como giros, bucles, etc. Es decir, también pueden usarse partes de una proteína. Además, proteína "como se usa en este documento incluye proteínas, oligopéptidos y péptidos. Además, pueden usarse variantes de proteínas, es decir, estructuras análogas a anticuerpos que no se producen de manera natural.

25 Las proteínas adecuadas incluyen, pero no se limitan a, proteínas industriales y farmacéuticas, que incluyen ligandos, receptores de superficie celular, antígenos, anticuerpos, citoquinas, hormonas, factores de transcripción, módulos de señalización, proteínas citoesqueléticas y enzimas. Las clases adecuadas de enzimas incluyen, pero no se limitan a, hidrolasas tales como proteasas, carbohidrasas, lipasas; isomerasas tales como racemerasas, epimerasas, tautomerases o mutasas; transferasas, quinasas, oxidorreductasas y fosfatasas. Las enzimas adecuadas se enumeran en la base de datos de enzimas Swiss-Prot. Estructuras principales de de proteínas adecuadas incluyen, pero no se limitan a, todas los que se encuentran en la base de datos recopilada y mantenida por la Research Collaboratory for Structural Bioinformatics (RCSB, anteriormente el Laboratorio Nacional de Brookhaven).

30 Tal como se usa en el presente documento, el término "segmento variable" se refiere a una parte de un péptido naciente que comprende una secuencia aleatoria, pseudoaleatoria o de núcleo definido. Un "segmento variable" se refiere a una parte de un péptido naciente que comprende una secuencia aleatoria, pseudoaleatoria o de núcleo definido. Un segmento variable puede comprender posiciones de residuo tanto variantes como invariantes, y el grado de variación de residuos en una posición de residuo variante pueden limitarse: se seleccionan ambas opciones a criterio del profesional. Normalmente, los segmentos variables tienen de 5 a 20 residuos de aminoácido de longitud (por ejemplo, de 8 a 10), aunque los segmentos variables pueden ser más largos y pueden comprender partes de anticuerpo o proteínas receptoras, tales como un fragmento de anticuerpo, una proteína de unión a ácido nucleico, una proteína receptora, y similares.

35 El término "salvaje" o "de tipo natural", significa que el polinucleótido no comprende ninguna mutación. Una proteína "de tipo natural" significa que la proteína será activa a un nivel de actividad que se encuentra en la naturaleza y comprenderá la secuencia de aminoácidos que se encuentra en la naturaleza.

Descripción detallada de la invención

40 La divulgación proporciona un método de integración de generación y/o selección, evolución y expresión de anticuerpos terapéuticos en un huésped mamífero para la fabricación en un sistema individual. En una realización del método CIAO, se generan, se optimizan y se fabrican anticuerpos terapéuticos en el mismo sistema de huésped mamífero. En otra realización, se descubren y se fabrican agentes terapéuticos proteicos en el mismo huésped. Esto optimiza la fabricación desde el mismo comienzo, ahorrando costes (tiempo y recursos).

De manera histórica, el descubrimiento de anticuerpos se ha realizado en huéspedes eucariotas (euc) y procariotas (proc). Normalmente, en bacterias (*E. coli*), se descubren anticuerpos de longitud parcial; por ejemplo, en tecnologías de presentación en fago, se recuperan Fab y a veces se convierten en de longitud completa de manera posterior. Existen varias posibles desventajas en estos enfoques.

5 En un ejemplo, existen algunas pruebas de que regiones de Fc y Fv se comunican para afectar a las propiedades de anticuerpo, tales como unión y expresión. Por tanto, cuando un fragmento de anticuerpo se optimiza para una propiedad tal como expresión, la mejora no siempre se traduce en una expresión mejorada en el anticuerpo ensamblado de longitud completa. Por ejemplo, se creó una biblioteca de Fc como intentos de hallar un Fc "santo grial" que pudiera unirse a cualquier Fv para mejorar la expresión en cualquier huésped.

10 En un aspecto, se realizó mutagénesis de codones en la región constante para la optimización de la expresión en células de mamífero. Específicamente, se crearon 326 mutantes en la región constante y se expresaron en células HEK 293 y CHO-S. Se realizó examen mediante ELISA. Varios Fc cumplían con los criterios de expresión mejorada, e incluso se identificaron determinados Fc optimizados que transferían efectos positivos a través de múltiples líneas celulares; sin embargo, cuando se unió un Fv diferente al Fc, la mejora no se tradujo en la expresión. Esto
15 demuestra que los que Fc y Fv se comunican.

Para evitar resultados inesperados tras la recombinación de fragmentos de anticuerpo, en un aspecto preferido, se usa el método CIAO para descubrir moléculas de anticuerpo de longitud completa. En otro aspecto preferido, el método CIAO utiliza huéspedes euc.

20 En una realización, el sistema eucariota es un sistema de mamífero que se selecciona de uno del grupo que consiste en las líneas celulares CHO, HEK293, IM9, DS-1, THP-1, Hep G2, COS, NIH 3T3, C33a, A549, A375, SK-MEL-28, DU 145, PC-3, HCT 116, Mia PACA-2, ACHN, Jurkat, MM1, Ovcara 3, HT 1080, Panc-1, U266, 769P, BT-474, Caco-2, HCC 1954, MDAMB-468, LnCAP, NRK-49F y SP2/0; y esplenocitos de ratón y PBMC de conejo. En un aspecto, el sistema de mamífero se selecciona de una línea celular CHO o HEK293. En un aspecto específico, el sistema de mamífero es una línea celular CHO-S. En otro aspecto específico, el sistema de mamífero es una línea
25 celular HEK293. En otra realización, el sistema eucariota es un sistema de células de levadura. En un aspecto, el sistema eucariota se selecciona de células de levadura *S. cerevisiae* o células de levadura *Pichia*.

En otra realización, la creación de líneas celulares de mamífero puede realizarse comercialmente por una organización de investigación por contrato o de fabricación personalizada. Por ejemplo, para anticuerpos recombinantes u otras proteínas, Lonza (Lonza Group Ltd, Basilea, Suiza) puede crear vectores para expresar
30 producto usando la tecnología del GS Gene Expression System™ con las líneas celulares huésped o bien CHOK1SV o bien NSO.

En otra realización, puede realizarse evolución en huéspedes proc (tales como *E. coli*) y pueden realizarse exámenes en huéspedes euc (por ejemplo, CHO), siempre que se produzca el examen en el mismo huésped que el huésped de fabricación.

35 Selección de proteína(s) candidato(s) líder

Pueden usarse una variedad de métodos para descubrir, generar y/o seleccionar uno o más candidatos a proteína terapéutica que van a hacerse evolucionar. Las moléculas seleccionadas pueden ser proteínas recombinantes existentes, o pueden proceder de colecciones de proteínas recombinantes, incluyendo enzimas, hormonas y anticuerpos. Las proteínas terapéuticas pueden incluir enzimas y hormonas humanas clonadas. Pueden descubrirse
40 anticuerpos recombinantes usando cualquier número de plataformas de generación y examen disponibles. Los anticuerpos pueden estar en cualquier forma, incluyendo anticuerpos de cadena sencilla, totalmente humanos, Fab, Fv, Fc, bifuncionales, quiméricos, humanizados o totalmente humanos o fragmentos de los mismos. En un aspecto preferido, se usa el método CIAO para descubrir moléculas de anticuerpo de longitud completa. Pueden generarse y examinarse bibliotecas de anticuerpos recombinantes usando sistemas de selección o examen en huéspedes
45 mamíferos optimizados o no optimizados para proporcionar candidatos para la evolución. Varios sistemas de expresión eucariotas están publicados, por ejemplo, sistemas de células de mamífero o levadura.

En el método de la presente invención, tales sistemas de expresión en mamífero, en particular sistemas que usan presentación en la superficie celular de moléculas para examen y selección, se emplean para identificar y seleccionar candidatos para fabricación, o evolución seguida por fabricación. Preferiblemente, tales huéspedes
50 mamíferos son células de fibroblastos (3T3, ratón; BHK21, hámster sirio), células epiteliales (MDCK, perro; Hela, ser humano; PtK1, rata, canguro), células plasmáticas ((SP2/0 y NS0, ratón), células de riñón (293, ser humano; COS, mono), células de ovario (CHO, hámster chino), células embrionarias (R1 y E14,1, ratón; H1 y H9, ser humano; PER C.6, ser humano). Se emplea la tecnología de presentación en la superficie celular para presentar proteínas en la superficie de las células de mamífero para su examen. Se clonan proteínas como fusiones con moléculas de
55 membrana que cuando se expresan presentan las proteínas en la superficie de las células para un examen de alto rendimiento, rápido, por ejemplo.

En determinadas realizaciones, la divulgación proporciona un método de provisión de un anticuerpo optimizado. En una realización, se selecciona un antígeno y se genera una biblioteca de anticuerpos humanos y se expresa en un

sistema de mamífero, por ejemplo, células CHO-S. Se examina la biblioteca para identificar coincidencias de anticuerpo totalmente humano, por ejemplo, mediante clasificación de células activada por fluorescencia (FACS). Las coincidencias de anticuerpo totalmente humano se examinan/caracterizan entonces adicionalmente mediante cualquier ensayo relevante para determinar al menos una propiedad, característica o actividad predeterminada. En un aspecto, se examinan moléculas evolucionadas para determinar múltiples características simultáneamente, por ejemplo, función y expresión mejoradas. El ensayo relevante puede comprender por ejemplo, ELISA, o una tecnología de alineamiento. Se selecciona un anticuerpo molde de las coincidencias de anticuerpo humano. Cualquier método de evolución se realiza con el anticuerpo molde, o un polipéptido fragmento del mismo, para preparar un conjunto de anticuerpos mutantes. En un aspecto, la evolución se realiza mediante un método de modificación mediante ingeniería de proteína completa para producir el conjunto de anticuerpos mutantes. El método de modificación mediante ingeniería de proteína completa puede seleccionarse, por ejemplo, de uno o una combinación de evolución de posición completa (CPE™), síntesis de proteína completa (CPS™), evolución flexible, evolución sinérgica, evolución de inserción de posición completa (CPI™) o evolución de delección de posición completa (CPD™). En otro aspecto, los anticuerpos mutantes se expresan en el mismo sistema de mamífero usado para generar la biblioteca de anticuerpos humanos.

El conjunto de anticuerpos mutantes se caracteriza/examina para determinar al menos una propiedad, característica o actividad predeterminada. En un aspecto, el conjunto de anticuerpos mutantes se examina simultáneamente, por ejemplo, para determinar la función y expresión mejoradas. En un aspecto, se usa una base de datos específica de moléculas en forma de un mapa de posición funcional (un EvoMap™) para optimización adicional mediante una o más técnicas de evolución de proteínas conocidas en la técnica; seguida por identificación y caracterización de mutantes superiores. Se selecciona un anticuerpo optimizado mediante comparación con el anticuerpo molde con respecto a la al menos una propiedad, característica o actividad predeterminada.

En un aspecto, se realiza la expresión a escala de gramos del anticuerpo optimizado seleccionado; seguida por toxicología distinta de GLP. Se realiza la transfección de línea celular estable, el desarrollo del procedimiento y la generación de un banco de células maestro. Se realiza la fabricación con BPF en el mismo sistema de mamífero usado para generar la biblioteca de anticuerpos humanos, por ejemplo, células CHO-S, para dar como resultado Acm recombinantes terapéuticos optimizados.

En otra realización, el método CIAO parte de la selección de un hibridoma molde o anticuerpo recombinante; la evolución del anticuerpo para proporcionar un conjunto de anticuerpos mutantes que se examinan mediante el uso de presentación en la superficie celular de anticuerpos en un sistema de células de mamífero; y se realiza la fabricación en el mismo sistema de células de mamífero usado para el examen.

En otra realización, el hibridoma molde/anticuerpo recombinante seleccionado se humaniza y se examina en el huésped de fabricación, seguido por la producción en el huésped de fabricación, en el que la etapa de optimización (evolución) se omite completamente.

En otros aspectos de la presente invención, se realiza la optimización de la expresión posterior en huéspedes de fabricación haciendo evolucionar la región Fc del anticuerpo, codones silenciosos en el anticuerpo, y/o el vector y/o genes del huésped usados en la expresión de proteínas. En un aspecto, se genera una biblioteca de Fc mediante cualquier técnica evolutiva. En un aspecto específico de optimización de la expresión, se realiza CPE con un dominio Fc de un anticuerpo para crear una biblioteca de mutantes de Fc que pueden usarse para seleccionar una pareja óptima para cualquier Fv. La optimización se diseña para una rápida unión de todas las variantes de CPE de Fc a cada nueva región Fv. Alternativamente, puede usarse un subconjunto de estos Fc para unirse a diferentes Fv. Cada una de estas combinaciones de variante de CPE de Fc/Fv se examina como un anticuerpo de longitud completa expresado en células de mamífero (por ejemplo CHO, medios rentables) para una expresión óptima. Además, puede realizarse CPS para examinar todas las permutaciones teóricas de hasta 12 o más de estas coincidencias de CPE en células de mamífero para determinar la mejora de la expresión. También pueden seleccionarse cambios de codón deseables específicos para identificar clones con aumento de la expresión. Se identifican codones silenciosos y se realiza CPE en estas posiciones. Esta biblioteca de CPE se examina para identificar coincidencias de expresión óptimas. Además, pueden usarse todas las permutaciones teóricas de hasta 12 o más coincidencias de CPE en el procedimiento de CPS para generar una nueva biblioteca que puede examinarse en células de mamífero para determinar una mejora de la expresión. Se usan las coincidencias de mutación silenciosa de CPS superiores para personalizar la proteína para una expresión óptima en una línea celular y medios específicos. Esto proporciona la oportunidad de control de la estructura fina de fármacos biosimilares.

Otras zonas para la potenciación de la expresión incluyen: optimización del vector, incluyendo promotor, sitios de corte y empalme, extremos terminales 5' y 3', secuencias flanqueantes, reducción de la delección y la transposición génica, mejora de actividades génicas de células huésped, optimización de enzimas de glicosilación de huésped, y mutagénesis y selección de células huésped de ancho cromosómico. Se ha demostrado que secuencias de aminoácidos en 5' son importantes para la potenciación de la expresión.

Evolución de candidatos líder

Cualquier método de evolución de proteínas puede emplearse para la evolución simultánea de la optimización del

rendimiento y la expresión de proteínas. La optimización del rendimiento de proteínas puede incluir la mejora de diversas características tales como afinidad, características farmacocinéticas, direccionamiento a tejido, agregación proteína-proteína, dirección de la alta variabilidad de ensayo y modificación de otras características *in vivo*.

5 Los métodos para hacer evolucionar moléculas, incluyendo candidatos seleccionados de la presente invención, incluyen métodos estocásticos y no estocásticos. Los métodos publicados incluyen enfoques de mutagénesis aleatoria y no aleatoria. Cualquiera de estos enfoques puede emplearse para hacer evolucionar propiedades de las proteínas terapéuticas de la presente invención hacia una característica deseada, tal como mejor estabilidad a diferentes entornos de pH o temperatura, o mejor expresión en una célula huésped. Otras propiedades potencialmente deseables, tales como actividad catalítica mejorada, estabilidad de proteína mejorada en diversas condiciones, selectividad y/o solubilidad mejoradas, y resultados de expresión mejorados mediante la mejora de características tales como agregación reducida pueden seleccionarse para experimentos de evolución.

15 Se realiza la evolución directamente en un huésped eucariota, tal como un huésped de célula de mamífero o un huésped de célula de levadura, que se usará para la producción posterior de la proteína terapéutica. Pueden hacerse evolucionar candidatos para una expresión óptima en el mismo huésped usado para examinar y/o hacer evolucionar y para la fabricación. Puede lograrse la optimización de la expresión mediante la optimización de vectores usados (componentes de vector, tales como promotores, sitios de corte y empalme, extremos terminales 5' y 3' y secuencias flanqueantes), modificación génica de células huésped para reducir delecciones y transposiciones génicas, evolución de actividades génicas de células huésped mediante métodos *in vivo* o *in vitro* de hacer evolucionar genes relevantes, optimización de enzimas de glicosilación de huésped mediante la evolución de genes relevantes, y/o mediante estrategias de mutagénesis y selección de células huésped de ancho cromosómico para seleccionar para células con capacidades de expresión potenciadas. Las células huésped se describen adicionalmente en el presente documento.

25 La tecnología de examen y expresión con presentación en la superficie celular (por ejemplo, tal como se definió anteriormente) puede emplearse para examinar bibliotecas de proteínas evolucionadas para determinar candidatos que van a fabricarse.

30 Los métodos actuales de amplio uso para crear proteínas mutantes a partir de una molécula de partida son tecnologías de mutagénesis dirigida por oligonucleótidos, reacciones en cadena de la polimerasa propensas a errores y mutagénesis por casete, en las que la región específica que va a optimizarse se reemplaza por un oligonucleótido mutagenizado de manera sintética. En estos casos, se generan varios sitios mutantes alrededor de determinados sitios en la secuencia original.

35 En mutagénesis dirigida por oligonucleótidos, se reemplaza una secuencia corta por un oligonucleótido mutagenizado de manera sintética. La PCR propensa a errores usa condiciones de polimerización de baja fidelidad para introducir un bajo nivel de mutaciones puntuales de manera aleatoria a lo largo de una secuencia larga. En una mezcla de fragmentos de secuencia desconocida, puede usarse PCR propensa a errores para mutagenizar la mezcla. En mutagénesis por casete, un bloque de secuencias de un molde individual se reemplaza normalmente por una secuencia (parcialmente) aleatorizada.

40 Se han producido genes quiméricos uniendo 2 fragmentos de polinucleótido usando extremos cohesivos compatibles generados por enzima(s) de restricción, en los que cada fragmento se deriva de una molécula progenitora (o parental) independiente. Otro ejemplo es la mutagénesis de una posición de codón individual (es decir para lograr una sustitución, adición o delección de codón) en un polinucleótido parental para generar un polinucleótido de progeie individual que codifica para un polipéptido mutagenizado en el sitio individual.

Además, se han utilizado, sistemas de recombinación específicos del sitio *in vivo* para generar híbridos de genes, así como métodos aleatorios de recombinación *in vivo*, y recombinación entre genes homólogos pero truncados en un plásmido. También se ha notificado mutagénesis mediante extensión solapante y PCR.

45 Se han usado métodos no aleatorios para lograr mayores números de mutaciones puntuales y/o quimerizaciones, por ejemplo se han usado enfoques completos o exhaustivos para generar todas las especies moleculares dentro de una agrupación particular de mutaciones, para atribuir funcionalidad a grupos estructurales específicos en una proteína molde (por ejemplo una posición de aminoácido individual específica o una secuencia que se compone de dos o más posiciones de aminoácidos), y para clasificar y comparar agrupaciones específicas de mutaciones. La patente estadounidense número 7033781 titulada "Whole cell engineering by mutagenizing a substantial portion of a starting genome, combining mutations, and optionally repeating" describe un método de hacer evolucionar un organismo hacia características deseadas. La patente estadounidense número 6764835 titulada "Saturation mutagenesis in directed evolution" y la patente estadounidense número 6562594 titulada "Synthetic linkage reassembly in directed evolution" describe métodos de hacer evolucionar de manera exhaustiva y examinar para determinar características deseadas de moléculas. Cualquiera de tales métodos puede usarse en el método de la presente invención.

55 Existe una diferencia entre los métodos conocidos previamente de "mutagénesis por saturación" y las técnicas de evolución "completa" preferidas en el presente documento. Mutagénesis por saturación se refiere a una técnica de

evolución en la que se realiza cada cambio posible en cada posición de un polinucleótido molde o polipéptido molde; sin embargo el cambio en cada posición no se confirma mediante pruebas, sino que meramente se supone estadísticamente. Evolución completa se refiere a una técnica de evolución en la que se realiza cada cambio posible en cada posición de un polinucleótido molde o polipéptido molde y se somete a prueba el polinucleótido o polipéptido para confirmar que se ha realizado el cambio pretendido.

Los métodos de saturación son métodos inherentemente estadísticos, no completos y tampoco fueron nunca verdaderamente completos a través de todas las etapas (por ejemplo, a través de generación de mutantes, identificación de mutantes, expresión de proteínas mutantes, examen de proteínas mutantes y generación, identificación, expresión y examen de mutantes superiores recombinados). En técnicas completas de evolución, cada molécula se examina y confirma tanto en la primera etapa de mutagénesis, como además en una segunda etapa de recombinación de los mutantes superiores o coincidencias.

A menos que se confirme la mutagénesis por saturación mediante secuenciación o algún otro método, la técnica no puede considerarse que sea completa por varios posibles motivos. Por ejemplo, 1) los sistemas de clonación no son eficaces al 100% debido a errores de clonación o de síntesis, o es difícil clonar las moléculas o 2) algunas proteínas son tóxicas cuando se expresan y por tanto no pueden expresarse eficazmente. Por tanto, es importante confirmar mediante secuenciación, o alguna otra técnica, en cada etapa. Es útil puntuar cada etapa para examinar la expresión, de modo que los clones sin expresión no se designen como "negativos" como en el trabajo anterior, sólo se puntúen como no expresables. Por tanto, se considera que las técnicas completas son un sistema no estocástico más puro que las técnicas de saturación, tal como se confirma mediante la etapa de "confirmación".

20 Evolución de posición completa

Haciendo referencia a la figura 1, usando un péptido lineal como ejemplo sencillo, en una primera etapa, se genera un conjunto de variantes de aminoácido que se producen de manera natural (o un subconjunto de las mismas, o derivados de aminoácido) para cada codón desde la posición 1 hasta la n (correspondiendo n al número de residuos en la cadena de polipéptido) mediante un procedimiento denominado en el presente documento evolución de posición completa (CPE™). Este procedimiento se repite para cada cadena de polipéptido de la molécula diana. Un conjunto mínimo de mutaciones de aminoácido contiene sólo un codón para cada uno de los 19 aminoácidos naturales. Sin embargo, se reconoce que cada sistema de expresión puede experimentar sesgo de codones, en el que combinaciones insuficientes de ARNt pueden conducir a detención de la traducción, terminación prematura de la traducción, desplazamiento de marco de la traducción e incorporación incorrecta de aminoácidos. Por tanto, para la optimización de la expresión, cada conjunto contiene hasta 63 codones diferentes, incluyendo codones de terminación. En la siguiente etapa, se confirman las mutaciones mediante secuenciación de cada nueva molécula. También pueden emplearse otros métodos de confirmación.

Cada conjunto de aminoácidos se examina entonces para determinar al menos uno de:

- Función mejorada
- Mutaciones neutras
- Mutaciones inhibitorias
- Expresión
- Compatibilidad del clon con el sistema huésped.

Preferiblemente, se examinan múltiples características simultáneamente tales como, por ejemplo, función y expresión mejoradas.

Se combinan los datos para cada conjunto para toda(s) la(s) cadena(s) de polipéptido y se genera un mapa funcional detallado (denominado en el presente documento EvoMap™) de la molécula diana. Este mapa contiene información detallada sobre cómo afecta cada mutación al rendimiento/expresión y/o capacidad de clonación de la molécula diana. Permite la identificación de todos los sitios en los que no pueden realizarse cambios sin una pérdida de función de la proteína (por ejemplo, unión antígeno/receptor en caso de anticuerpos). También muestra donde pueden realizarse cambios sin afectar a la función. Identifica además cambios que dan como resultado moléculas que no se expresan en el sistema huésped, y por tanto no evalúa el efecto de la mutación.

En la figura 1, se muestra un esquema de un EvoMap™ hipotético. Cada posición en el molde se identifica como un sitio restringido (no mutable), un sitio totalmente mutable, un sitio parcialmente mutable o un mutante superior para una sustitución de aminoácido específica. Cada sitio parcialmente mutable puede designarse además como adecuado para sustitución, por ejemplo, por un residuo cargado, o una sustitución de residuo apolar, y un clon sin expresión y/o molécula que no puede clonarse en el sistema huésped.

Es posible utilizar el EvoMap™ para reconocer y recombinar sustituciones de aminoácidos individuales beneficiosas, y examinar para optimizar adicionalmente las características deseadas en la molécula diana. Sin embargo, la

evolución de determinadas características puede requerir dos o más mutaciones simultáneas para poderse observar. El EvoMap™ puede aprovecharse para producir eficazmente, y de manera rentable, un conjunto de polipéptidos mutantes multi-sitio de modo no aleatorio. El conjunto de polipéptidos mutantes multi-sitio puede examinarse entonces para determinar mutantes superiores multi-sitio.

5 La CPE permite el mapa de mutaciones de proteína confirmadas *in vivo* completo. La identificación de todo el conjunto de mutantes superiores permite etapa(s) de evolución combinatoria adicional(es). Puede utilizarse CPE para reducir el riesgo de inmunogenicidad de proteínas evolucionadas mediante la selección de mutaciones no en superficie; eliminación de epítomos de células T; y imitación de mutaciones somáticas.

10 En un aspecto, puede usarse CPE para generar una biblioteca de hasta 5, 10 ó 15 aminoácidos, o hasta la totalidad de los 19 aminoácidos. Se realizan cambios en cada posición en la proteína y se examina para determinar una característica deseable, tal como la afinidad de unión o expresión, y se crea el Evomap™. Pueden usarse tandas posteriores de mutación y examen para generar los datos para los 19 aminoácidos. A partir del mapa, se identifican sitios totalmente mutables. Estos sitios son útiles para identificar posiciones que pueden modificarse para crear una nueva colección de moléculas que pueden producirse y someterse a prueba para determinar características. Por ejemplo, puede emplearse informática para identificar haplotipos de HLA en la secuencia, y pueden realizarse cambios deseados para evitar estos haplotipos realizando cambios dirigidos específicos en sitios "neutros" ("totalmente mutables") identificados a partir del mapa, en los que la característica neutra no se verá afectada. Esto podría reducir potencialmente el riesgo de inmunogenicidad (podrían seleccionarse mutaciones no en superficie, eliminar epítomos de células T, imitar mutaciones hipersomáticas). Además, el mapa puede mostrar sitios para modificaciones específicas de sitio (glicosilación y conjugación química) para mejorar diversas características. Además, la optimización de mutaciones silenciosas puede mejorar la expresión de proteínas en una variedad de huéspedes.

Evolución sinérgica

25 En una realización de la presente invención, se genera un EvoMap™ y se utiliza para evolución sinérgica, tal como se muestra en la figura 2. En la evolución sinérgica, puede combinarse la mutación simultánea en 2-20 sitios seleccionados para producir un efecto combinatorio. Se usa el EvoMap™ del polipéptido molde para seleccionar mutaciones puntuales de aminoácido específicas individuales para el ensamblaje en mutaciones de polipéptido multi-sitio.

30 En la evolución sinérgica, se seleccionan mutaciones puntuales de aminoácido no desactivantes de dentro de sitios parcialmente mutables que están cerca de sitios no mutables en el EvoMap™. En un aspecto, las mutaciones puntuales no desactivantes seleccionadas son adyacentes a sitios no mutables. En la evolución sinérgica, se realiza mutación simultánea de aminoácidos en de dos a 20 de los sitios seleccionados para efectos combinatorios. En un aspecto, se usa recombinación de dos a 20 mutaciones seleccionadas para producir una biblioteca de variantes de codón que codifica para una población de polipéptidos mutantes multi-sitio. Después de la clonación y expresión, entonces se examinan los polipéptidos mutantes multi-sitio producidos para determinar al menos una propiedad, característica o actividad predeterminada en comparación con el polipéptido molde. De esta manera, pueden identificarse polipéptidos mutantes superiores multi-sitio. En un aspecto, se producen polipéptidos mutantes multi-sitio mediante síntesis combinatoria de proteínas. Una ventaja de la evolución sinérgica es que no requiere una estructura cristalina de rayos X de la proteína para dirigir la evolución del polipéptido molde. Esta técnica es útil particularmente para proteínas con alta variación de ensayo y otros efectos multi-sitio.

40 Según la presente invención, las aplicaciones de la evolución sinérgica incluyen, pero no se limitan a evolución de cambios mecanísticos moleculares complejos, evolución de proteínas con alta variación de ensayo, evolución de la especificidad de proteína, mejora de la expresión en diversos huéspedes de expresión, mejora de la actividad catalítica de la proteína, estabilidad y optimización del pH. La evolución sinérgica es aplicable a todos los tipos de proteínas terapéuticas, incluyendo, pero no limitado a hormonas, enzimas, citoquinas y anticuerpos.

45 En un aspecto de la presente invención, puede usarse evolución sinérgica para optimizar uno o más aspectos de un polipéptido que forma parte de una molécula de proteína. La molécula de proteína puede ensamblarse ligando uno o más ácidos nucleicos mutantes que codifican para polipéptidos con cero, uno o más ácidos nucleicos que codifican para polipéptidos de región de entramado para crear una proteína variante mediante técnicas de clonación, traducción y expresión conocidas en la técnica. En un aspecto, un polipéptido de región de entramado se deriva de una molécula de proteína de tipo natural. En este aspecto, puede usarse la evolución sinérgica junto con técnicas de humanización de anticuerpos. Por ejemplo, puede seleccionarse un anticuerpo monoclonal de ratón para evolución y humanización. Se clonan y se secuencian las regiones CDR del anticuerpo y pueden sintetizarse regiones CDR individuales (CDR1, CDR2, CDR3) y ligarse a otros nucleótidos que codifican para polipéptidos de región de entramado de anticuerpo humano, seguido por la producción de una biblioteca de IgG variante humana. Se examina entonces la biblioteca de IgG variante humana para determinar al menos una propiedad, por ejemplo dos o más propiedades incluyendo función y expresión mejoradas, en comparación con el Acn de ratón. En otro aspecto, un polipéptido de región de entramado es un polipéptido de armazón artificial. En la sección de ejemplos, se presentan técnicas específicas de preparación de fragmentos de ADN bc, ligamiento y ensamblaje de ácidos nucleicos, clonación, transfección, expresión, síntesis en fase sólida de bibliotecas, síntesis en fase de disolución de

bibliotecas, evolución de posición completa, síntesis combinatoria de proteínas, cuantificación de la expresión mediante cuantificación por ELISA y ensayo de β -galactosidasa y ELISA funcional.

5 En otra realización de la invención, puede usarse la evolución sinérgica para potenciar la afinidad de unión de un anticuerpo. En esta realización, puede realizarse la optimización de la región variable del anticuerpo. Por ejemplo, para la producción de mutantes de anticuerpo, se realiza CPE para regiones variables de cadena ligera y cadena pesada de un anticuerpo seleccionado y se genera un EvoMap™. Se seleccionan mutantes para reensamblaje; por ejemplo, se seleccionan variantes de la cadena ligera y se seleccionan variantes de la cadena pesada para el ensamblaje. Se seleccionan mutaciones puntuales de aminoácido no desactivantes de dentro de sitios parcialmente mutables que están cerca de sitios no mutables. Puede usarse la tecnología de reensamblaje que utiliza CPS para
10 crear una biblioteca de cadenas pesadas. Pueden combinarse las variantes de cadena ligera con las variantes de cadena pesada, clonarse, expresarse y se examinan las variantes como IgG completas de sobrenadantes de líneas celulares eucariotas. Se evalúa la afinidad de unión para determinadas variantes mediante, por ejemplo, el uso de ensayos de instrumentación ELISA, BIAcore y/o Sapidyne, u otras técnicas conocidas por un experto en la técnica.

Evolución flexible

15 En otra realización, puede usarse CPE/EvoMap para identificar y aprovechar sitios totalmente mutables. En un aspecto, el aprovechamiento de múltiples sitios totalmente mutables se denomina evolución flexible y se usa para realizar cambios dirigidos tales como introducción de sitios para glicosilación (por ejemplo codones para aminoácidos para glicosilación unida a N u O; Asn dentro de la secuencia consenso Asn-Aa-Ser-Thr o Ser/Thr) y conjugación química. También puede usarse evolución flexible en el diseño de sitios de escisión con proteasa,
20 introducción de etiquetas para purificación y/o detección, marcaje específico del sitio, y similares. Además, puede utilizarse la optimización de codones de mutaciones silenciosas para la mejora de la expresión de proteínas. En esta realización, denominada evolución flexible, tras la expresión de proteínas, las bibliotecas de polipéptidos mutantes producidas se vuelven a examinar para determinar al menos una propiedad, característica o actividad predeterminada en comparación con el polipéptido molde. En un aspecto, la propiedad predeterminada incluye
25 reducción de la agregación proteína-proteína, potenciación de la estabilidad de proteína, o aumento de la solubilidad de proteína. En un aspecto, las bibliotecas de polipéptidos mutantes se examinan para determinar dos o más propiedades simultáneamente. En otro aspecto, puede usarse cualquier sistema de expresión eucariota que glicosila para la introducción de sitios de glicosilación, tal como, por ejemplo, líneas celulares de mamíferos, plantas, levaduras e insectos.

30 En la evolución flexible, la evaluación de la bioinformática y estructuras cristalinas de rayos X de proteína de proteínas relacionadas, o la proteína o polipéptido molde, es útil para la optimización de moldes. En un aspecto, sitios seleccionados no son residuos en contacto. En otro aspecto, la selección de mutaciones de proteína no en superficie permite un riesgo reducido de inmunogenicidad.

35 Las aplicaciones de la evolución flexible incluyen, pero no se limitan a, reducción de la agregación proteína-proteína, mejora de la solubilidad de proteínas, optimización de la farmacocinética a través de bibliotecas de glicosilación, optimización de la estructura secundaria y terciaria de proteínas y desinmunización de sitios antigénicos directamente a través de o bien conjuntos de mutación o bien indirectamente a través de enmascaramiento de glicosilación.

40 En un aspecto de la evolución flexible, se utiliza un EvoMap™ para identificar sitios totalmente mutables, se realiza generación de CPS con inserción de residuos de glicosilación en sitios totalmente mutables (o mutaciones silenciosas para efectos de traducción), y se realiza el examen de la biblioteca glicosilada combinatoria mediante análisis analítico (por ejemplo análisis mediante espectrometría de masas, dispersión dinámica de la luz), reducción de la inmunogenicidad (mediante bioinformática o ensayo) y/o análisis farmacocinético (por ejemplo en ratones Foxn1nu).

45 En un aspecto, puede usarse evolución flexible para desinmunización para eliminar la inmunogenicidad mientras que se mantiene la función. Puede realizarse desinmunización mediante evolución flexible enmascarando la inmunogenicidad con glicosilación, identificación de sustituciones de aminoácidos de espectros de mutaciones hipersomáticas humanas que pueden eliminar la inmunogenicidad mientras que se mantiene la función, reducción de la dosis para eludir una posible inmunogenicidad, y minimización de cambios de residuo de aminoácido no en
50 superficie. Además, pueden usarse bases de datos de inmunogenicidad y algoritmos para identificar y reemplazar posibles epítopos de unión a CMH. En un aspecto, se acopla la predicción de modificación *in silico* con datos de CPE/CPS para generar variantes.

Pueden medirse la propensión reducida a generar epítopos de células T y/o la desinmunización mediante técnicas conocidas en la técnica. Preferiblemente, puede someterse a prueba *in vitro* la desinmunización de proteínas
55 mediante un ensayo de proliferación de células T. En este ensayo, se examinan PBMC de donantes que representan > 80% de alelos de HLA-DR en el mundo, para determinar la proliferación en respuesta a polipéptidos o bien de tipo natural o bien desinmunizados. De manera ideal, sólo se detecta proliferación celular tras la carga de las células presentadoras de antígeno con péptidos de tipo natural. Los ensayos adicionales para determinar la desinmunización incluyen ensayos de reestimulación *in vitro* de PBMC humanos (por ejemplo ELISA para interferón-

gamma (TH1) o IL4 (TH2)). Alternativamente, puede someterse a prueba la desinmunización mediante la expresión de tetrámeros de HLA-DR que representan todos los haplotipos. Para someter a prueba si se presentan péptidos desinmunizados en haplotipos de HLA-DR, puede medirse la unión de, por ejemplo, péptidos marcados con fluorescencia en PBMC. La medición de ratones transgénicos para HLA de clase I y clase II para determinar las respuestas al antígeno diana (por ejemplo interferón-gamma o IL4). Alternativamente, el examen de biblioteca de epítomos con células T educadas (CMHI noamérico; CMHII icosamérico) de ensayos de PBMC y/o ratón transgénico. Además, puede probarse la desinmunización determinando si se han generado anticuerpos contra las moléculas desinmunizadas después de la administración en pacientes.

En otra realización, las técnicas flexibles de evolución de la presente invención pueden utilizarse para la optimización de la expresión. En un aspecto, la presente invención da a conocer la utilización de métodos de modificación mediante ingeniería de proteínas para desarrollar variantes de Fc con codones optimizados con mutación silenciosa con expresión mejorada en células eucariotas. Una mutación silenciosa es una en la que una variación de la secuencia de ADN no da como resultado un cambio en la secuencia de aminoácidos del anticuerpo. En un aspecto, se realiza mutagénesis de codones en la región constante para la optimización de la expresión en células eucariotas. Una variante de Fc con codones optimizados con propiedades de expresión mejorada al tiempo que conserva la capacidad para mediar en funciones efectoras mejora la producción de anticuerpos terapéuticos. En este aspecto, por ejemplo, puede hacerse evolucionar una región constante de una molécula de anticuerpo para examinar en diferentes huéspedes de expresión, por ejemplo, examen de la expresión de líneas celulares de mamífero utilizando CHO, HEK293 y COS-7. En la figura 3, se muestra un ejemplo de optimización de la expresión mediante mutagénesis de codones en la región constante para la expresión en células de mamífero y se describe en el ejemplo 19. Los niveles de expresión mostrados son cada uno un promedio de 4 puntos de datos, y se confirmaron a lo largo de múltiples experimentos. Se demostró la capacidad de múltiples líneas celulares para el primer mutante sometido a prueba en sistemas de expresión de línea celular HEK293 y CHO.

Además, puede usarse el EvoMap™ para generar modelos moleculares tridimensionales computacionales del oligopéptido, o regiones específicas del mismo, para explorar los mecanismos estructurales implicados en, por ejemplo, estabilidad y especificidad anticuerpo-epítipo. Se muestra un EvoMap™ tridimensional hipotético en la figura 13.

También puede combinarse la información en EvoMap con información estructural (si está disponible) para seleccionar por ejemplo, sólo residuos de superficie para mutaciones para aumentar la solubilidad/disminuir la agregación.

Evolución de inserción de posición completa

En una realización, la divulgación proporciona métodos de identificación de y mapeo de polipéptidos mutantes formados a partir de, o basados en, un polipéptido molde. Haciendo referencia a la figura 8, usando un péptido lineal como ejemplo sencillo, en una primera etapa, se genera un conjunto de variantes de aminoácido que se producen de manera natural (o un subconjunto de las mismas, o derivados de aminoácido) para cada codón desde la posición 2 hasta la n (correspondiendo n al número de residuos en la cadena de polipéptido) mediante un procedimiento denominado en el presente documento as evolución de inserción de posición completa (CPI™).

En la CPI™, se inserta un aminoácido después de cada aminoácido en la totalidad de un polipéptido molde uno cada vez para generar un conjunto de polipéptidos alargados. Puede usarse CPI para insertar 1, 2, 3, 4, o hasta 5 nuevos sitios cada vez. Cada uno de los 20 aminoácidos se añade en cada nueva posición, uno cada vez, creando un conjunto de 20 moléculas diferentes en cada nueva posición añadida en el molde. En este caso, se salta la posición 1, que es metionina e invariante. Este procedimiento se repite para cada cadena de polipéptido de la molécula diana. Un conjunto mínimo de mutaciones de aminoácido contiene sólo un codón para cada uno de los 20 aminoácidos naturales.

La presente invención se refiere a métodos de identificación de y mapeo de polipéptidos mutantes formados a partir de, o basados en, un polipéptido molde. Normalmente, el polipéptido comprenderá n residuos de aminoácido, en el que el método comprende (a) generar $n+[20 \times (n-1)]$ polipéptidos independientes, en los que cada polipéptido difiere del polipéptido molde en que se ha insertado después de cada posición en el molde cada uno de los 20 aminoácidos uno cada vez (tal como se ilustra en la figura 1); someter a ensayo cada polipéptido para determinar al menos una propiedad, característica o actividad predeterminada; y (b) para cada miembro identificar cualquier cambio en dicha propiedad, característica o actividad en relación con el polipéptido molde.

En una realización, se seleccionan una o más regiones para mutagénesis para añadir una posición cada vez tal como se describió anteriormente. En tal caso, n representa un subconjunto o región del polipéptido molde. Por ejemplo, cuando el polipéptido es un anticuerpo, todo el anticuerpo o una o más regiones determinantes de la complementariedad (CDR) del anticuerpo se someten a mutagénesis para añadir una posición cada vez en el polipéptido molde después de cada posición.

La invención incluye por tanto métodos de mapeo de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde que tiene al menos una, y preferiblemente seis, regiones determinantes de la complementariedad

(CDR), comprendiendo las CDR en conjunto n residuos de aminoácido, comprendiendo el método (a) generar $n+[20 \times (n-1)]$ anticuerpos independientes, en el que cada anticuerpo difiere del anticuerpo molde en que se ha insertado una posición predeterminada individual, una cada vez, después de cada posición en la secuencia molde; (b) someter a ensayo cada conjunto para determinar al menos una propiedad, característica o actividad predeterminada; y (c) para cada miembro identificar cualquier cambio en una propiedad, característica o actividad en relación con el polipéptido molde. Para anticuerpos, la propiedad, característica o propiedad predeterminada pueden ser la afinidad de unión y/o la inmunogenicidad, por ejemplo.

Además, se proporcionan métodos de producción de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde que tiene al menos una región determinante de la complementariedad (CDR), comprendiendo la CDR n residuos de aminoácido, comprendiendo el método: (a) generar $n+[20 \times (n-1)]$ anticuerpos independientes, en los que cada anticuerpo difiere del anticuerpo molde en que tiene un aminoácido extra añadido en una posición individual predeterminada de la CDR. En otra realización, el anticuerpo comprende seis CDR, y en conjunto las CDR comprenden n residuos de aminoácido.

En otra realización, los nuevos polipéptidos alargados descritos anteriormente se mutan adicionalmente y se mapean después del examen para identificar un cambio en una propiedad, característica o actividad en relación con el polipéptido acortado. Normalmente, el polipéptido alargado comprenderá n residuos de aminoácido, en el que el método comprende (a) generar n ($n-1$ en el caso en el que el residuo inicial sea metionina) conjuntos independientes de polipéptidos, comprendiendo cada conjunto polipéptidos miembro que tienen un número X de residuos de aminoácido diferentes predeterminados en una posición individual predeterminada del polipéptido; en el que cada conjunto de polipéptidos difiere en la posición individual predeterminada; someter a ensayo cada conjunto para determinar al menos una propiedad, característica o actividad predeterminada; (b) para cada miembro identificar cualquier cambio en dicha propiedad, característica o actividad en relación con el polipéptido molde; y opcionalmente (c) crear un mapa funcional que refleje tales cambios. Preferiblemente, el número de polipéptidos miembros diferentes generados es equivalente a $n \times X$ (o $[n-1] \times X$, según sea el caso).

Como alternativa, el método comprende generar una población individual que comprende los conjuntos de polipéptidos mutados de los polipéptidos alargados. En esta realización, se examina toda la nueva población, se identifican los miembros individuales y se genera el mapa funcional.

Normalmente, cuando se usa cada aminoácido que se produce de manera natural, X será 19 (representando los 20 residuos de aminoácido que se producen de manera natural y excluyendo el residuo particular presente en una posición dada del polipéptido molde). Sin embargo, puede usarse en la totalidad cualquier subconjunto de aminoácidos, y cada conjunto de polipéptidos pueden sustituirse por la totalidad o un subconjunto del X total usado para toda la población.

Sin embargo, se reconoce que cada sistema de expresión puede experimentar sesgo de codones, en el que combinaciones insuficientes de ARNt pueden conducir a detención de la traducción, terminación prematura de la traducción, desplazamiento de marco de la traducción e incorporación incorrecta de aminoácidos. Por tanto, para la optimización de la expresión, cada conjunto contiene hasta 63 codones diferentes.

Cada conjunto de aminoácidos se examina entonces para determinar al menos una, y preferiblemente dos o más, características deseables, tales como función mejorada; mutaciones neutras, mutaciones inhibitoras y expresión.

En un aspecto, los polipéptidos alargados pueden mapearse para identificar un cambio en una propiedad, característica o actividad que da como resultado los polipéptidos acortados en relación el "tipo natural". Se combinan los datos para cada conjunto para todo el polipéptido, o "molécula diana". Pueden usarse entonces las coincidencias del examen de los polipéptidos alargados (moléculas diana) para la mutagénesis completa adicional de cadena(s) y el examen tal como se describe en el presente documento. Los datos de mutagénesis proporcionan que se genere un mapa funcional detallado (denominado en el presente documento EvoMap™) de la molécula diana. Este mapa contiene información detallada sobre cómo afecta cada mutación al rendimiento/expresión de la molécula diana. Permite la identificación de todos los sitios en los que no pueden realizarse cambios sin una pérdida de la función de la proteína (o unión a antígeno/receptor en el caso de anticuerpos). También muestra donde pueden realizarse cambios sin afectar a la función.

En otro aspecto, puede usarse CPE para generar una biblioteca de 5, 10, hasta 15, o hasta los 19 aminoácidos en cada posición de interés.

Evolución de delección de posición completa

La evolución de delección de posición completa (CPD™) se refiere a métodos de identificación de y mapeo de polipéptidos mutantes formados a partir de, o basados en, un polipéptido molde. La evolución de CPD deleciona cada aminoácido a través de la proteína una posición cada vez. Normalmente, el polipéptido comprenderá n residuos de aminoácido, en el que el método comprende (a) generar $n-1$ ($n-2$ en el caso en el que el residuo inicial es metionina) polipéptidos independientes, en el que cada polipéptido difiere del polipéptido molde en que carece de una posición predeterminada individual; someter a ensayo cada polipéptido para determinar al menos una propiedad, característica o actividad predeterminada; y (b) para cada miembro identificar cualquier cambio en dicha

propiedad, característica o actividad en relación con el polipéptido molde.

En una realización de evolución de CPD, se seleccionan una o más regiones para mutagénesis para eliminar una posición cada vez. En tal caso, n representa un subconjunto o región del polipéptido molde. Por ejemplo, cuando el polipéptido es un anticuerpo, todo el anticuerpo o una o más regiones determinantes de la complementariedad (CDR) del anticuerpo se someten a mutagénesis para eliminar una posición cada vez en el polipéptido molde.

En una realización, CPD incluye por tanto métodos de mapeo de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde que tiene al menos una, y preferiblemente seis, regiones determinantes de la complementariedad (CDR), comprendiendo las CDR en conjunto n residuos de aminoácido, comprendiendo el método (a) generar $(n-1)$ anticuerpos independientes, en el que cada anticuerpo difiere del anticuerpo molde en que carece de una posición predeterminada individual; (b) someter a ensayo cada conjunto para determinar al menos una propiedad, característica o actividad predeterminada; y (c) para cada miembro identificar cualquier cambio en una propiedad, característica o actividad en relación con el polipéptido molde. Para anticuerpos, la propiedad, característica o propiedad predeterminada puede ser la afinidad de unión y/o la inmunogenicidad, por ejemplo.

Un aspecto de la evolución de CPD incluye métodos de producción de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde que tiene al menos una región determinante de la complementariedad (CDR), comprendiendo la CDR n residuos de aminoácido, comprendiendo el método: (a) generar $n-1$ anticuerpos independientes, en el que cada anticuerpo difiere del anticuerpo molde en que carece de una posición predeterminada individual de la CDR. En otra realización, el anticuerpo comprende seis CDR, y en conjunto las CDR comprenden n residuos de aminoácido.

En otra realización de la evolución de CPD, los nuevos polipéptidos acortados descritos anteriormente se mutan y se mapean adicionalmente tras el examen para identificar un cambio en una propiedad, característica o actividad en relación con el polipéptido acortado. Normalmente, el polipéptido acortado comprenderá n residuos de aminoácido, en el que el método comprende (a) generar n ($n-1$ en el caso en el que el residuo inicial es metionina) conjuntos independientes de polipéptidos, comprendiendo cada conjunto polipéptidos miembros que tienen un número X de residuos de aminoácido diferentes predeterminados en una posición individual predeterminada del polipéptido; en el que cada conjunto de polipéptidos difiere en la posición individual predeterminada; someter a ensayo cada conjunto para determinar al menos una propiedad, característica o actividad predeterminada; (b) para cada miembro identificar cualquier cambio en dicha propiedad, característica o actividad en relación con el polipéptido molde; y (c) crear un mapa funcional que refleje tales cambios. Preferiblemente, el número de polipéptidos miembros diferentes generados es equivalente a $n \times X$ (o $[n-1] \times X$, tal como puede ser el caso).

Como alternativa, el método de CPD comprende generar una población individual que comprende los conjuntos de polipéptidos mutados a partir de los polipéptidos acortados. En esta realización, se examina toda la nueva población, se identifican los miembros individuales y se genera el mapa funcional. Normalmente, cuando se usa cada aminoácido que se produce de manera natural, X será 19 (representando los 20 residuos de aminoácido que se producen de manera natural y excluyendo el residuo particular presente en una posición dada del polipéptido molde). Sin embargo, puede usarse en la totalidad cualquier subconjunto de aminoácidos, y cada conjunto de polipéptidos puede sustituirse por la totalidad o un subconjunto del X total usado para toda la población.

Puede usarse cualquier medio de mutación o sintético para generar el conjunto de mutantes en la evolución de CPD. En una realización, la generación de polipéptidos comprende (i) someter un polinucleótido que contiene codones que codifica para el polipéptido molde a amplificación basada en polimerasa usando un oligonucleótido degenerado 64 veces para cada codón que va a mutagenizarse, en el que cada uno de los oligonucleótidos degenerados 64 veces se compone de una primera secuencia homóloga y una secuencia de triplete N,N,N degenerada, de modo que se genera un conjunto de polinucleótidos de progenie; y (ii) someter el conjunto de polinucleótidos de progenie a amplificación clonal de manera que se expresan polipéptidos codificados por los polinucleótidos de progenie.

En una realización de evolución de CPD, todo el polipéptido acortado se somete a mutagénesis por saturación. En otra realización, se seleccionan una o más regiones para mutagénesis por saturación. En tal caso, n representa un subconjunto o región del polipéptido molde. Por ejemplo, cuando el polipéptido es un anticuerpo, todo el anticuerpo o una o más regiones determinantes de la complementariedad (CDR) del anticuerpo se someten a mutagénesis por saturación.

La divulgación de la evolución de CPD incluye por tanto métodos de mapeo de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde acortado que tiene al menos una, y preferiblemente seis, regiones determinantes de la complementariedad (CDR), comprendiendo las CDR en conjunto n residuos de aminoácido, comprendiendo el método (a) generar n conjuntos independientes de anticuerpos, comprendiendo cada conjunto anticuerpos miembros que tienen un número X de residuos de aminoácido diferentes predeterminados en una posición individual predeterminada de la CDR; en el que cada conjunto de anticuerpos difiere en la posición individual predeterminada; y el número de anticuerpos miembros diferentes generados es equivalente a $n \times X$; (b) someter a ensayo cada conjunto para determinar al menos una propiedad, característica o actividad predeterminada; (c) para cada miembro identificar cualquier cambio en una propiedad, característica o actividad en relación con el polipéptido molde; y (d) crear un mapa de posición estructural de tales cambios. Para anticuerpos, la propiedad,

característica o propiedad predeterminada puede ser la afinidad de unión y/o la inmunogenicidad. Tal como se expuso anteriormente, como alternativa puede generarse una población individual que comprende todos los conjuntos de anticuerpos mutados.

5 Además, se proporcionan métodos de producción de un conjunto de anticuerpos mutantes formados a partir de un anticuerpo molde acortado que tiene al menos una región determinante de la complementariedad (CDR), comprendiendo la CDR n residuos de aminoácido, comprendiendo el método: (a) generar n conjuntos independientes de anticuerpos, comprendiendo cada conjunto anticuerpos miembros que tienen un número X de residuos de aminoácido diferentes predeterminados en una posición individual predeterminada de la CDR; en el que cada conjunto de anticuerpos difiere en la posición individual predeterminada; y el número de anticuerpos miembros diferentes generados es equivalente a $n \times X$. En otra realización, el anticuerpo comprende seis CDR, y en conjunto las CDR comprenden n residuos de aminoácido.

15 El método de evolución de CPD™ incluye un mapa de posición funcional (EvoMap™) producido mediante los métodos descritos en el presente documento. En una realización adicional, determinados residuos particularmente sensibles al cambio pueden indicarse así en el EvoMap™. Puede implementarse una optimización adicional realizando cambios de mutación por adición en posiciones fuera de estas posiciones sensibles. También es posible utilizar el EvoMap™ para reconocer y recombinar sustituciones de aminoácidos individuales beneficiosas, y examinar para optimizar adicionalmente las características deseadas en la molécula diana, en un procedimiento denominado síntesis combinatoria de proteínas (CPS™).

Síntesis combinatoria de proteínas

20 La síntesis combinatoria de proteínas (CPS™) implica combinar coincidencias individuales de CPE, CPI, CPD, o cualquier otra técnica de evolución para combinar dos o más mutaciones. Se usa CPS para sintetizar proteínas con mutaciones combinadas que luego se examinan para determinar características de genes y proteínas optimizadas. En la figura 3 se muestra un esquema de CPS™. En un aspecto, se combinan en la CPS dos o más mutaciones puntuales que dan como resultado mutantes superiores o mutaciones neutras.

25 En una realización, se combina CPE con CPS para crear mutantes, que se examinan para detectar la propiedad deseada. En un aspecto, pueden ahorrarse tiempo y recursos en el procedimiento de CPE cambiando 2 aa o 3 aa o 4 aa cada vez frente a uno cada vez; de modo que si el número de aa en la proteína es N, el número total generado y examinado para 2 aa cada vez sería $(20^2) \times 1/2N$; 3 cada vez sería $(20^3) \times 1/3N$, etc. Por ejemplo, en un aspecto específico, (en el ejemplo de 2 aa): se combina el 1º aa en la 1ª posición de aa con los 20 en la 2ª posición de aa y los demás aa permanecen igual, luego se combina el 2º aa en la 1ª posición de aa con los 20 en la 2ª posición de aa y los demás aa permanecen iguales. Se examina toda la población para detectar mutantes superiores y luego se realiza mutación en el segundo conjunto de los siguientes dos aa más adelante. En un aspecto similar, esto puede realizarse para 3 aa cada vez o 4 aa cada vez. En otro aspecto, se sigue opcionalmente el procedimiento de CPE con CPS de mutantes superiores (incluyendo cualquier subconjunto de los mismos).

35 En un aspecto, pueden incorporarse aminoácidos no naturales en el procedimiento (de modo que los otros 19 aminoácidos, o un subconjunto de los mismos, más aminoácidos no naturales) usando tecnologías novedosas tales como el codón cuadruplete descrito en los artículos adjuntos y relacionados. Neumann *et al.* Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome Nature 464, 441-444 (14 de febrero de 2010). En este aspecto se realiza CPE/CPS para la incorporación de aminoácidos no naturales.

40 En un aspecto adicional, toda la biblioteca de CPE se crea de manera sintética (sintetizando todas las moléculas en máquinas disponibles comercialmente). En el caso que de la máquina de síntesis no pueda crear hebras suficientemente grandes, se sintetizan fragmentos y luego se ligan para generar moléculas de longitud completa. Esta biblioteca se examina y se sigue con CPS para combinar mutaciones deseadas. Este es un procedimiento en dos etapas en el que la CPE va seguida por CPS, no una etapa de sólo CPE.

45 En otro aspecto, se genera una biblioteca de CPE y se examina, seguido luego por combinar mediante CPS mutantes superiores tal como sigue: si hay 10 mutantes superiores, se somete a prueba una molécula individual con los 10 cambios, luego se someten a prueba todas las versiones de 9 mutaciones, luego 8, 7, 6, 5 etc. hasta que uno de los grupos no encuentra una molécula mejorada con respecto a cualquiera en el grupo anterior. Una vez que se identifica una molécula mejorada, puede terminarse el procedimiento.

50 En un aspecto adicional, se realiza CPE para identificar mutantes superiores y mutaciones neutras para determinar la afinidad y expresión, luego se realiza CPS con combinaciones de mutantes superiores y mutaciones neutras, y vuelve a examinarse la biblioteca para detectar mejoradas adicionales en características tales como función, afinidad y/o expresión.

En un aspecto adicional, se realiza CPE en codones del Fc u otro dominio para detectar cambios de glicosilación.

55 En otro aspecto, puede realizarse CPE o CPE combinada con CPS de microARN o intrones.

En un aspecto adicional, se realiza CPE o CPE combinada con CPS de CDR de anticuerpos de roedor, luego se

examinan para detectar mutantes superiores, seguido por humanización.

En un aspecto, se realiza CPE o CPE combinada con CPS para producir nucleótidos intermedios alternativos que conducen a la mutación deseada en la reacción final, por ejemplo, una citosina metilada que se convierte en un uracilo.

- 5 En un aspecto, se utiliza CPE o CPE combinada con CPS más informática para convertir CDR de ratón en CDR humanas y viceversa.

En un aspecto, se utiliza CPE o CPE combinada con CPS con 2 y 3 mutaciones separadas por todo el anticuerpo.

En otro aspecto, se usa CPE o CPE combinada con CPS en un vector de cadena doble para evaluación de examen para detectar un aumento de sensibilidad.

- 10 En un aspecto adicional, se utiliza CPE o CPE combinada con CPS para seleccionar cambios alostéricos.

Puede usarse cualquiera de varias técnicas de examen para evaluar mutantes de CPE/CPS. En un aspecto pueden secretarse mutantes de CPE o CPE combinada con CPS y presentarse en huéspedes eucariotas. Alternativamente, pueden producirse mutantes de CPE o CPE combinada con CPS en *E. coli* y examinarse en huéspedes eucariotas.

- 15 En otro aspecto, se realiza CPE partiendo de 15 aa o 10 aa y seguida por CPS; luego seguida por el resto de los 19 aa restantes. En otro aspecto, se utiliza CPE o CPE combinada con CPS para hacer evolucionar proteínas específicamente con cambios de aminoácidos que no están en la superficie. En un aspecto, puede utilizarse CPE para mapeo epitópico multidimensional. En otro aspecto, puede realizarse examen mediante CPE o CPE combinada con CPS de manera transitoria en células eucariotas. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, luego se realiza la secuenciación y el alineamiento de todos los clones, por ejemplo, en un formato basado en chip o basado en pocillo para la expresión y el examen. En otro aspecto, se utiliza CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 20 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 25 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 30 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 35 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 40 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 45 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 50 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- 55 En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS para hacer evolucionar la coordinación de iones de metal seleccionando concentraciones de iones variables. En un aspecto adicional, se realiza CPE o CPE combinada con CPS, y se expresan y examinan proteínas en condiciones libres de células y en organismos vivos no humanos. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS de células madre para determinar efectos variables sobre la diferenciación y la expresión de proteínas y ARN y ARNm. En un aspecto adicional, se realiza el examen múltiplex mediante CPE o CPE combinada con CPS para detectar múltiples características de proteínas como expresión y unión. En otro aspecto, se realiza CPE o CPE combinada con CPS en moléculas molde que participan en el transporte cerebral y el cruce de membranas; y se examinan mutantes para detectar características mejoradas. En un aspecto, se examinan mutantes de CPE o CPE combinada con CPS para detectar características higroscópicas de proteínas. En otro aspecto, se someten a ensayo mutantes de CPE o CPE combinada con CPS para seleccionar proteínas dinámicas. En un aspecto, se realiza el examen mediante CPE o CPE combinada con CPS fuera de la condición objetivo para identificar mutantes dentro de la condición objetivo y viceversa.

- El término "molde" puede referirse a un polipéptido base o a un polinucleótido que codifica para tal polipéptido. Tal como apreciaría un experto en la técnica, puede usarse cualquier molde en los métodos y las composiciones de la presente invención. Pueden usarse moldes que pueden mutarse y de ese modo hacerse evolucionar para guiar la síntesis de otro polipéptido o biblioteca de polipéptidos tal como se describe en la presente invención. Tal como se describe en más detalle en el presente documento, el molde que puede hacerse evolucionar codifica para la síntesis de un polipéptido y puede usarse posteriormente para descodificar la historia de síntesis del polipéptido, para amplificar indirectamente el polipéptido, y/o para hacer evolucionar (es decir, diversificar, seleccionar y amplificar) el polipéptido. El molde que puede hacerse evolucionar es, en determinadas realizaciones, un ácido nucleico. En determinada realización de la presente invención, el molde se basa en un ácido nucleico. En otras realizaciones, el molde es un polipéptido.

- Los moldes de ácido nucleico usados en la presente invención están compuestos por ADN, ARN, un híbrido de ADN y ARN, o un derivado de ADN y ARN, y pueden ser mono o bicatenarios. La secuencia del molde se usa para codificar para la síntesis de un polipéptido, preferiblemente un compuesto que no es, o no se parece a, un ácido nucleico o análogo de ácido nucleico (por ejemplo, un polímero no natural o una molécula pequeña). En el caso de determinados polímeros no naturales, el molde de ácido nucleico se usa para alinear las unidades de monómero en la secuencia en la que aparecen en el polímero y ponerlos en proximidad estrecha a unidades de monómero adyacentes a lo largo del molde de modo que reaccionarán y se unirán mediante un enlace covalente. En otras determinadas realizaciones, el molde puede utilizarse para generar polímeros no naturales mediante amplificación por PCR de una biblioteca de molde de ADN que consiste en una región aleatoria de nucleótidos.

- Se apreciará que el molde puede variar enormemente en el número de bases. Por ejemplo, en determinadas realizaciones, el molde puede ser de 10 a 10.000 bases de longitud, preferiblemente entre 10 y 1.000 bases de

longitud. La longitud del molde dependerá por supuesto de la longitud de los codones, la complejidad de la biblioteca, la longitud del polímero no natural que va a sintetizarse, la complejidad de la molécula pequeña que va a sintetizarse, el uso de secuencias espaciadoras, etc. La secuencia de ácido nucleico puede prepararse usando cualquier método conocido en la técnica para preparar secuencias de ácido nucleico. Estos métodos incluyen métodos tanto *in vivo* como *in vitro* incluyendo PCR, preparación de plásmidos, digestión con endonucleasas, síntesis en fase sólida, transcripción *in vitro*, separación de hebras, etc. En determinadas realizaciones, el molde de ácido nucleico se sintetiza usando un sintetizador de ADN automatizado.

Tal como se comentó anteriormente, en determinadas realizaciones de la invención, el método se usa para sintetizar polipéptidos que no son, o no se parecen a, ácidos nucleicos o análogos de ácido nucleico. Por tanto, en determinadas realizaciones de la presente invención, el molde de ácido nucleico comprende secuencias de bases que codifican para la síntesis de un polímero no natural o molécula pequeña. El mensaje codificado en el molde de ácido nucleico comienza preferiblemente con un codón específico que lleva a su lugar un sitio químicamente reactivo a partir del cual puede tener lugar la polimerización, o en el caso de sintetizar una molécula pequeña el codón de "iniciación" puede codificar para un anti-codón asociado con un armazón de molécula pequeña o un primer reactante. El codón de "iniciación" de la presente invención es análogo al codón de "iniciación", ATG, que codifica para el aminoácido metionina.

En aún otras realizaciones de la invención, el propio molde de ácido nucleico puede modificarse para que incluya un sitio de iniciación para la síntesis de polímeros (por ejemplo, un nucleófilo) o un armazón de molécula pequeña. En determinadas realizaciones, el molde de ácido nucleico incluye un bucle en horquilla en uno de sus extremos que termina en un grupo reactivo usado para iniciar la polimerización de las unidades de monómero. Por ejemplo, un molde de ADN puede comprender un bucle en horquilla que termina en un grupo amino en 5', que puede estar protegido o no. A partir del grupo amino puede comenzar la polimerización del polímero no natural. El grupo amino reactivo también puede usarse para unir un armazón de molécula pequeña sobre el molde de ácido nucleico con el fin de sintetizar una biblioteca de moléculas pequeñas.

Para terminar la síntesis del polímero no natural, debe incluirse un codón de "terminación" en el molde de ácido nucleico preferiblemente en el extremo de la secuencia codificante. El codón de "terminación" de la presente invención es análogo a los codones de "terminación" (es decir, TAA, TAG, TGA) que se encuentran en transcritos de ARNm. Estos codones conducen a la terminación de la síntesis de proteínas. En determinadas realizaciones, se elige un codón de "terminación" que es compatible con el código genético artificial usado para codificar para el polímero no natural. Por ejemplo, el codón de "terminación" no debe entrar en conflicto con ningún otro codón usado para codificar para la síntesis, y debe ser del mismo formato general que los otros codones usados en el molde. El codón de "terminación" puede codificar para una unidad de monómero que termina la polimerización al no proporcionar un grupo reactivo para la unión adicional. Por ejemplo, una unidad de monómero de terminación puede contener un grupo reactivo bloqueado tal como una acetamida en vez de una amina primaria. En aún otras realizaciones, la unidad de monómero de terminación comprende un extremo terminal biotinilado que proporciona un modo conveniente de terminación de la etapa de polimerización y purificación del polímero resultante.

En una realización, se usan productos de ADN mutagenizado directamente como molde para la síntesis *in vitro* de las correspondientes proteínas mutantes. Debido a la alta eficacia con la que pueden generarse sustituciones de los 19 aminoácidos en un residuo individual, es posible realizar mutagénesis por saturación en numerosos residuos de interés, o bien independientemente o bien en combinación con otras mutaciones dentro de la proteína. Tal como se usa en el presente documento, mutagénesis "por saturación completa" se define como sustituir un aminoácido dado dentro de una proteína por los otros 19 aminoácidos que se producen de manera natural. Por ejemplo, la mutagénesis por saturación de sitio génico, que explora de manera sistemática mínimamente todas las posibles sustituciones de aminoácidos individuales a lo largo de una secuencia de proteína, se da a conocer en Kretz *et al.*, *Methods in Enzymology*, 2004, 388:3-11; Short, patente estadounidense n.º 6.171.820; y Short, patente estadounidense n.º 6.562.594.

En un aspecto, esta invención prevé el uso de cebadores de codón (que contienen una secuencia N,N,G/T degenerada) para introducir mutaciones puntuales en un polinucleótido, de modo que se genera un conjunto de polipéptidos de progeie en los que una gama completa de sustituciones de aminoácidos individuales está representada en cada posición de aminoácido (véase la patente estadounidense n.º 6.171.820; véase también la patente estadounidense n.º 5.677.149). Los oligos usados se componen de manera contigua de una primera secuencia homóloga, una secuencia N,N,G/T degenerada, y preferiblemente pero no necesariamente una segunda secuencia homóloga. Los productos de traducción de la progeie posterior a partir del uso de tales oligos incluyen todos los posibles cambios de aminoácido en cada sitio de aminoácido a lo largo del polipéptido, debido a que la degeneración de la secuencia N,N,G/T incluye codones para los 20 aminoácidos.

El uso de codones es uno de los factores importantes en la expresión génica eucariota. Las frecuencias con las que se usan diferentes codones varían significativamente entre diferentes huéspedes, y entre proteínas expresados a altos o bajos niveles dentro del mismo organismo. El motivo más probable de esta variación es que codones preferidos se correlacionan con la abundancia de ARNt relacionados disponibles dentro de la célula. Es posible que el uso de codones y las concentraciones de aceptores de ARNt hayan evolucionado conjuntamente, y que la presión de selección para esta coevolución sea más pronunciada para genes altamente expresados que para genes

expresados a bajos niveles.

En un aspecto, se usa un oligo degenerado de este tipo (que se compone de un casete de N,N,G/T degenerado) para someter cada codón original en un molde de polinucleótido parental a una gama completa de sustituciones de codones. En otro aspecto, se usan al menos dos casetes de N,N,G/T degenerados, o bien en el mismo oligo o bien no, para someter al menos dos codones originales en un molde de polinucleótido parental a una gama completa de sustituciones de codones. Por tanto, más de una secuencia N,N,G/T puede estar contenida en un oligo para introducir mutaciones de aminoácidos en más de un sitio. Esta pluralidad de secuencias N,N,G/T puede ser directamente contigua, o estar separada por una o más secuencia(s) de nucleótidos adicional(es). En otro aspecto, pueden usarse oligos utilizables para introducir adiciones y deleciones o bien solos o bien en combinación con los codones que contiene una secuencia N,N,G/T, para introducir cualquier combinación o permutación de adiciones, deleciones y/o sustituciones de aminoácidos.

En otro aspecto, la presente invención proporciona el uso de casetes degenerados que tienen menos degeneración que la secuencia N,N,G/T. Por ejemplo, puede ser deseable en algunos casos usar (por ejemplo, en un oligo) una secuencia de triplete degenerada que se compone de sólo un N, en el que dicho N puede estar en la primera, segunda o tercera posición del triplete. Cualquier otra base incluyendo cualquier combinación y permutación de las mismas puede usarse en las dos posiciones restantes del triplete. Alternativamente, puede ser deseable en algunos casos usar (por ejemplo, en un oligo) una secuencia de triplete N,N,N degenerada.

Sin embargo, se aprecia que el uso de un triplete N,N,G/T degenerado tal como se da a conocer en el presente documento es ventajoso por varios motivos. En un aspecto, esta invención proporciona un medio para generar de manera sistemática y bastante fácil la sustitución de la gama completa de posibles aminoácidos (para un total de 20 aminoácidos) en todas y cada una de las posiciones de aminoácido en un polipéptido. Por tanto, para un polipéptido de 100 aminoácidos, la presente invención proporciona un modo para generar de manera sistemática y bastante fácil 2000 especies distintas (es decir, 20 posibles aminoácidos por posición X 100 posiciones de aminoácido). Se aprecia que se proporciona, a través del uso de un oligo que contiene un triplete N,N,G/T degenerado, 32 secuencias individuales que codifican para 20 posibles aminoácidos. Por tanto, en un recipiente de reacción en el que se somete una secuencia de polinucleótido parental a mutagénesis por saturación usando un oligo de este tipo, se generan 32 polinucleótidos de progenie distintos que codifican para 20 polipéptidos distintos. En cambio, el uso de un oligo no degenerado en mutagénesis dirigida al sitio conduce a sólo un producto de polipéptido de progenie por recipiente de reacción.

Por tanto, en una realización preferida, cada recipiente de reacción de mutagénesis por saturación contiene polinucleótidos que codifican para al menos 20 moléculas de polipéptidos de progenie de manera que los 20 aminoácidos están representados en una posición de aminoácido específica correspondiente a la posición de codón mutagenizada en el polinucleótido parental. Los polipéptidos de progenie degenerados 32 veces generados a partir de cada recipiente de reacción de mutagénesis por saturación puede someterse a amplificación clonal (por ejemplo, clonarse en un huésped de *E. coli* adecuado usando un vector de expresión) y someterse a examen de expresión. Cuando se identifica mediante examen un polipéptido de progenie individual que presenta un cambio en una propiedad (en comparación con el polipéptido molde), puede secuenciarse para identificar la sustitución de aminoácido responsable de tal cambio contenido en el mismo.

El polipéptido molde puede ser cualquier proteína, sin embargo, se prefieren las proteínas que tienen un ensayo conveniente para la actividad tal como la actividad catalítica o la unión al ligando. Como se usa en este documento, un ligando es cualquier molécula que se une específicamente a una más grande, tal como la unión de una molécula pequeña a una proteína. Los ejemplos representativos de las interacciones objetivo incluyen catálisis, interacciones enzima-sustrato, interacciones proteína-ácido nucleico, interacciones receptor-ligando, interacciones proteína-metal e interacciones anticuerpo-antígeno. Las proteínas diana representativas incluyen enzimas, anticuerpos, citoquinas, receptores, proteínas de unión al ADN, agentes quelantes y hormonas.

Puede usarse cualquier método mutagénico recombinante o de síntesis química para generar la población de polipéptidos mutantes. La práctica de la presente invención puede emplear, a menos que se indique otra cosa, técnicas convencionales de biología celular, cultivo celular, biología molecular, biología transgénica, microbiología, ADN recombinante e inmunología, que están dentro del conocimiento de la técnica. Tales técnicas se explican con más detalle en la bibliografía. Véase, por ejemplo, *Molecular Cloning A Laboratory Manual*, 2ª ed., ed. por Sambrook, Fritsch y Maniatis (Cold Spring Harbor Laboratory Press: 1989); *ADN Cloning*, volúmenes I y II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis *et al.*, patente estadounidense n.º 4.683.195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription and Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); el tratado, *Methods in Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller y M. P. Cabs eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, volúmenes 154 y 155 (Wu *et al.* eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer y Walquer, eds., Academic Press, Londres, 1987); *Handbook Of Experimental Immunology*, volúmenes I-IV (D. M. Weir y C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

En una realización, el polipéptido molde es un anticuerpo. El anticuerpo se somete a los métodos descritos en el presente documento para, por ejemplo, mapear y entender qué posiciones dentro de la CDR efectúan la afinidad de unión. Las técnicas para preparar y usar diversos constructos basados en anticuerpos y fragmentos de los mismos se conocen bien en la técnica. Un aspecto importante de la presente invención es la identificación de residuos que desempeñan, o es probable que desempeñen, un papel en la interacción de interés (por ejemplo, interacción antígeno-anticuerpo, quelación de metales, unión a receptor, unión a sustrato, etc.). Puede usarse cualquier anticuerpo o fragmento de anticuerpo según la presente invención.

En una realización, puede usarse cualquiera de las plataformas de evolución CPE, CPI, CPD y CPS para generar anticuerpos agonistas, es decir anticuerpos activantes. Estas tecnologías de evolución permiten la generación de anticuerpos agonistas más allá de la activación del tipo de reticulación de proteínas más sencilla y en particular permiten la activación de receptores tales como GPL-1 o 2 que se activan tradicionalmente por péptidos.

En un aspecto, se seleccionan anticuerpos mediante FACS o microscopía o equivalente para detectar anticuerpos débilmente activantes usando células con señales fluorescentes que fluorescen cuando se activa el receptor de superficie celular. Posteriormente, se usan las herramientas de evolución para potenciar esta activación. Entonces se utiliza la tecnología de CPS para combinar mutantes superiores.

En otro aspecto, se selecciona un anticuerpo que se une al sitio de activación del receptor tal como se determina mediante mapeo de epítomos. Se usan técnicas de CPE, CPI y/o CPD para seleccionar mutantes que provocan la estimulación del receptor tal como se determina mediante una lectura intracelular tal como fluorescencia en respuesta a la liberación de iones de calcio u otros ensayos que se conocen bien en la técnica. Entonces se utiliza la tecnología de CPS para combinar mutantes superiores.

En un aspecto particular, algunas de las ventajas clave de CPI con inserciones de aminoácidos individuales, dobles o triples son que estos aminoácidos insertados pueden extenderse en la cavidad de unión del receptor para activar el receptor. En otro aspecto particular, la CPD puede remodelar y/o reposicionar aminoácidos que interactúan con el receptor para mejorar o efectuar la activación y finalmente la CPE puede realizar cambios relativamente más pequeños para efectuar la activación del receptor.

La especificidad de un anticuerpo está determinada por las regiones determinantes de la complementariedad (CDR) dentro de las regiones variables de cadena ligera (VL) y regiones variables de cadena pesada (VH). El fragmento Fab de un anticuerpo, que tiene aproximadamente un tercio del tamaño de un anticuerpo completo contiene las regiones variables de cadena pesada y ligera, la región constante de cadena ligera completa y una parte de la región constante de cadena pesada. Las moléculas Fab son estables y se asocian bien debido a la contribución de las secuencias de región constante. Sin embargo, el rendimiento de Fab funcional expresado en sistemas bacterianos es menor que el del fragmento Fv más pequeño que contiene sólo las regiones variables de las cadenas pesada y ligera. El fragmento Fv es la parte más pequeña de un anticuerpo que todavía retiene un sitio de unión a antígeno funcional. El fragmento Fv tiene las mismas propiedades de unión que el Fab, sin embargo, sin la estabilidad conferida por las regiones constantes, las dos cadenas del Fv pueden disociarse de manera relativamente fácil en condiciones diluidas.

En un aspecto, pueden fusionarse regiones VH y VL por medio de un ligador polipeptídico (Huston *et al.*, 1991) para estabilizar el sitio de unión a antígeno. Este fragmento Fv de polipéptido individual se conoce como anticuerpo de cadena sencilla (scFv). Las VH y VL pueden disponerse con cualquier dominio en primer lugar. El ligador une el extremo carboxilo-terminal de la primera cadena con el extremo amino-terminal de la segunda cadena.

Un experto en la técnica reconocerá que los fragmentos Fv de cadena sencilla, Fv de cadena pesada o ligera o Fab pueden usarse con este sistema. Una cadena pesada o ligera puede mutagenizarse seguido por la adición de la cadena complementaria a la disolución. Entonces se permite que las dos cadenas se combinen y formen un fragmento de anticuerpo funcional. La adición de secuencias de cadena pesada o ligera no específicas aleatorias permite la producción de un sistema combinatorio para generar una biblioteca de diversos miembros.

Generalmente, se genera un polinucleótido de expresión de cadena sencilla. Este polinucleótido de expresión contiene: (1) un casete de anticuerpo de cadena sencilla que consiste en un dominio V_H, un péptido espaciador y un dominio V_L operativamente unidos para codificar para un anticuerpo de cadena sencilla, (2) un promotor adecuado para la transcripción *in vitro* (por ejemplo, promotor de T7, promotor SP6, y similares) operativamente unido para garantizar la transcripción *in vitro* del casete de anticuerpo de cadena sencilla que forma un ARNm que codifica para un anticuerpo de cadena sencilla, y (3) una secuencia de terminación de la transcripción adecuada para funcionar en una reacción de transcripción *in vitro*. Opcionalmente, el polinucleótido de expresión puede comprender también un origen de replicación y/o un marcador seleccionable. Un ejemplo de un polinucleótido de expresión adecuado es pLM166.

Las secuencias de V_H y V_L pueden obtenerse convenientemente de una biblioteca de secuencias de V_H y V_L producidas mediante amplificación por PCR usando cebadores específicos de la familia del gen de V o cebadores específicos del gen de V (Nicholls *et al.*, J. Immunol. Met., 1993, 165: 81; documento WO93/12227) o se diseñan según métodos convencionales conocidos en la técnica basándose en la información de secuencia disponible.

- Normalmente, se aíslan secuencias de V_H y V_L de ratón o ser humano. Entonces se ligan las secuencias de V_H y V_L, habitualmente con una secuencia espaciadora intermedia (por ejemplo, que codifica para un espaciador peptídico flexible en marco), formando un casete que codifica para un anticuerpo de cadena sencilla. Normalmente, se usa una biblioteca que comprende una pluralidad de secuencias de V_H y V_L (algunas veces también con una pluralidad de especies de péptido espaciador representadas), en el que la biblioteca se construye con una o más de las secuencias de V_H y V_L mutadas para aumentar la diversidad de secuencia particularmente en residuos de CDR, algunas veces en residuos de región de entramado. Las secuencias de región V pueden clonarse convencionalmente como ADNc o productos de amplificación por PCR para células que expresan inmunoglobulinas. Por ejemplo, pueden usarse células de hibridoma humano, o linfoma, u otra línea celular que sintetiza inmunoglobulina o bien secretada o bien de superficie celular para el aislamiento de ARN de poliA+. El ARN se usa entonces para la síntesis de ADNc cebado con oligo dT usando la enzima transcriptasa inversa (para métodos generales véanse Goodspeed *et al.*, Gene 1989, 76: 1; Dunn *et al.*, J. Biol. Chem., 1989, 264: 13057). Una vez aislado el producto de PCR o ADNc de región V, se clona en un vector para formar un casete de anticuerpo de cadena sencilla.
- Para lograr la construcción de anticuerpos y fragmentos de anticuerpo, se aíslan los genes codificantes y se identifican. Los genes pueden modificarse para permitir la clonación en un vector de expresión o una transcripción/traducción *in vitro*. Aunque pueden usarse métodos tales como estudiar con sonda el ADN para V_H y V_L a partir de ADNc de hibridoma (Maniatis *et al.*, 1982) o construir un gen sintético para V_H y V_L (Barbas *et al.*, 1992), un modo conveniente es usar métodos dirigidos por molde para amplificar las secuencias de anticuerpo. Puede amplificarse una población diversa de genes de anticuerpo a partir de una muestra de molde diseñando cebadores para las secuencias conservadas en los extremos 3' y 5' de la región variable conocida como región de entramado o para las regiones constantes del anticuerpo (Iverson *et al.*, 1989). Dentro de los cebadores, pueden colocarse sitios de restricción para facilitar la clonación en un vector de expresión. Dirigiendo los cebadores a estas regiones conservadas, se mantiene la diversidad de la población de anticuerpos para permitir la construcción de bibliotecas diversas. La especie y clase específicas de anticuerpo pueden definirse mediante la selección de las secuencias de cebador tal como se ilustra por el gran número de secuencias para todos los tipos de anticuerpos facilitados en Kabat *et al.*, 1987.
- Puede usarse ARN mensajero aislado del bazo o la sangre periférica de un animal como molde para la amplificación de una biblioteca de anticuerpos. En determinadas circunstancias, cuando es deseable presentar una población homogénea de fragmentos de anticuerpo sobre la superficie celular, puede aislarse ARNm de una población de anticuerpos monoclonales. Puede prepararse ARN mensajero a partir de cualquier fuente mediante métodos convencionales y usarse directamente o para la preparación de un molde de ADNc. La generación de ARNm para fines de clonación de anticuerpos se logra fácilmente siguiendo los procedimientos bien conocidos para la preparación y caracterización de anticuerpos (véase, por ejemplo, Antibodies: A Laboratory Manual, 1988).
- La generación de anticuerpos monoclonales (Acm) sigue generalmente los mismos procedimientos que aquellos para preparar anticuerpos policlonales. En resumen, se prepara un anticuerpo policlonal inmunizando un animal con una composición inmunogénica acorde y recogiendo antisueros de ese animal inmunizado. Puede usarse una amplia gama de especies animales para la producción de antisueros. Normalmente el animal usado para la producción de antisueros es un conejo, un ratón, una rata, un hámster, una cobaya o una cabra. Debido al volumen de sangre relativamente grande de los conejos, se prefieren habitualmente conejos para la producción de anticuerpos policlonales.
- Las composiciones inmunogénicas a menudo varían en la inmunogenicidad. A menudo es necesario por tanto reforzar el sistema inmunitario del huésped, tal como puede lograrse acoplado un inmunógeno peptídico o polipeptídico a un portador. Portadores a modo de ejemplo y preferidos son hemocianina de lapa californiana (KLH) y albúmina sérica bovina (BSA). Otras albúminas tales como ovoalbúmina, albúmina sérica de rata o albúmina sérica de conejo también pueden usarse como portadores. Se conocen bien medios reconocidos para conjugar un polipéptido a una proteína transportadora e incluyen glutaraldehído, éster de m-maleimidobenzoi-N-hidroxisuccinimida, carbodiimidas y bencidina bis-diazotada.
- La inmunogenicidad de una composición de inmunógeno particular puede potenciarse mediante el uso de estimuladores no específicos de la respuesta inmunitaria, conocidos como adyuvantes. Los adyuvantes a modo de ejemplo y preferidos incluyen adyuvante completo de Freund (un estimulador no específico de la respuesta inmunitaria que contiene *Mycobacterium tuberculosis* inactivada), adyuvante incompleto de Freund y adyuvante de hidróxido de aluminio.
- La cantidad de composición de inmunógeno usada en la producción de anticuerpos policlonales varía con la naturaleza del inmunógeno así como el animal usado para la inmunización. Puede usarse una variedad de vías para administrar el inmunógeno (subcutánea, intramuscular, intradérmica, intravenosa e intraperitoneal). La producción de anticuerpos policlonales puede monitorizarse extrayendo muestras de sangre del animal inmunizado en diversos puntos de tiempo tras la inmunización. También puede administrarse una segunda inyección, de refuerzo. El procedimiento de refuerzo y titulación se repite hasta que se logra un título adecuado. Cuando se obtiene un nivel deseado de inmunogenicidad, puede extraerse sangre del animal inmunizado y aislarse el suero, almacenarse y recogerse el bazo para el aislamiento de ARNm a partir de la respuesta policlonal o puede usarse el animal para

generar Acm para el aislamiento de ARNm a partir de una población homogénea de anticuerpos.

Pueden prepararse Acm fácilmente a través del uso de técnicas bien conocidas, tales como las ejemplificadas en la patente estadounidense n.º 4.196.265. Normalmente, esta técnica implica inmunizar un animal adecuado con una composición de inmunógeno seleccionado, por ejemplo un hapteno de molécula pequeña conjugado con un portador, una proteína, un polipéptido o péptido purificado o parcialmente purificado. La composición de inmunización se administra de una manera eficaz para estimular las células productoras de anticuerpos. Roedores tales como ratones y ratas son animales usados frecuentemente; sin embargo, también es posible el uso de células de conejo, oveja, rana. El uso de ratas puede proporcionar determinadas ventajas (Goding, págs. 60-61, 1986), pero se prefieren ratones, particularmente el ratón BALB/c ya que este es el más usado de manera rutinaria y generalmente proporciona un mayor porcentaje de fusiones estables.

Tras la inmunización, se seleccionan células somáticas con el potencial de producir anticuerpos, específicamente linfocitos B (células B), para su uso en el protocolo de generación de Acm. Estas células pueden obtenerse a partir de bazo, amígdalas o ganglios linfáticos que se han sometido a biopsia, o a partir de muestras de sangre. Son preferibles células de bazo y células sanguíneas, las primeras debido a que son una fuente rica de células productoras de anticuerpos que están en la fase de plasmablasto en división, y las últimas porque puede accederse fácilmente a la sangre. A menudo, se habrá inmunizado un panel de animales y se extraerá el bazo del animal con el título de anticuerpos más alto y se obtendrán los linfocitos del bazo homogeneizando el bazo con una jeringa. Normalmente, un bazo de un ratón inmunizado contiene de aproximadamente 5×10^7 a 2×10^8 linfocitos.

Los linfocitos B productores de anticuerpos del animal inmunizado se fusionan entonces con células de una célula de mieloma inmortal, generalmente una de la misma especie que el animal que se inmunizó. Líneas celulares de mieloma adecuadas para su uso en procedimientos de fusión de producción de hibridomas preferiblemente no son productoras de anticuerpos, tienen alta eficacia de fusión y deficiencias enzimáticas que las hacen incapaces de crecer en determinados medios selectivos que soportan el crecimiento sólo de las células fusionadas deseadas (hibridomas).

Puede usarse una cualquiera de varias células de mieloma, tal como conocen los expertos en la técnica (Goding, págs. 65-66, 1986; Campbell, 1984). Por ejemplo, cuando el animal inmunizado es un ratón, puede usarse P3-X63/Ag8, X63-Ag8.653, NS1/1.Ag 4 1, Sp210-Ag14, FO, NSO/U, MPC-11, MPC11-X45-GTG 1.7 y S194/5XX0 Bul; para ratas, puede usarse R210.RCY3, Y3-Ag 1.2.3, IR983F y 4B210; y U-266, GM1500-GRG2, LICR-LON-HMy2 y UC729-6 son todas útiles en relación con fusiones de células humanas.

Una célula de mieloma murina preferida es la línea celular de mieloma NS-1 (también denominada P3-NS-1-Ag4-1), que está fácilmente disponible del NIGMS Human Genetic Mutant Cell Repository solicitando el número de depósito de línea celular GM3573. Otra línea celular de mieloma de ratón que puede usarse es la línea celular no productora SP2/0 de mieloma murino de ratón resistente a 8-azaguanina.

Los métodos para generar híbridos de células de mieloma y células de ganglios linfáticos o bazo productoras de anticuerpos comprenden habitualmente mezclar células somáticas con células de mieloma en una proporción 2:1, aunque la proporción puede variar entre 20:1 y 1:1, respectivamente, en presencia de un agente o agentes (químicos o eléctricos) que promueven la fusión de membranas celulares. Se han descrito métodos de fusión usando virus Sendai por Kohler & Milstein (1975; 1976), y los que usan polietilenglicol (PEG), tal como PEG al 37% (v/v), por Gefter *et al.*, 1977). El uso de métodos de fusión inducida eléctricamente también es apropiado (Goding págs. 71-74, 1986).

Los procedimientos de fusión producen habitualmente híbridos viables a bajas frecuencias, de 1×10^{-6} a 1×10^{-8} . Sin embargo, esto no supone un problema, ya que los híbridos viables, fusionados se diferencian de las células parentales, no fusionadas (particularmente las células de mieloma no fusionadas que normalmente seguirían dividiéndose indefinidamente) mediante cultivo en un medio selectivo. El medio selectivo es generalmente uno que contiene un agente que bloquea la síntesis *de novo* de nucleótidos en los medios de cultivo tisular. Agentes a modo de ejemplo y preferidos son aminopterina, metotrexato y azaserina. La aminopterina y el metotrexato bloquean la síntesis *de novo* tanto de purinas como de pirimidinas, mientras que la azaserina bloquea sólo la síntesis de purinas. Cuando se usa aminopterina o metotrexato, el medio se complementa con hipoxantina y timidina como fuente de nucleótidos (medio HAT). Cuando se usa azaserina, el medio se complementa con hipoxantina.

El medio de selección preferido es HAT. Sólo células que pueden llevar a cabo rutas de salvamento de nucleótidos pueden sobrevivir en medio HAT. Las células de mieloma son defectuosas en enzimas clave de la ruta de salvamento, por ejemplo, hipoxantina fosforribosil transferasa (HPRT), y no pueden sobrevivir. Las células B pueden llevar a cabo esta ruta, pero tienen una vida limitada en cultivo y generalmente mueren en el plazo de aproximadamente dos semanas. Por tanto, las únicas células que pueden sobrevivir en los medios selectivos son los híbridos formados a partir de células B y de mieloma.

Este cultivo proporciona una población de hibridomas a partir de los cuales se seleccionan hibridomas específicos. Normalmente, la selección de hibridomas se realiza cultivando las células mediante dilución de clones individuales en placas de microtitulación, seguido por someter a prueba los sobrenadantes clonales individuales (después de

aproximadamente de dos a tres semanas) para detectar la reactividad deseada. Los ensayos simples y rápidos incluyen radioinmunoensayos, inmunoensayos enzimáticos, ensayos de citotoxicidad, ensayos en placa, ensayos de unión inmunológica puntual, y similares.

5 Los hibridomas seleccionados se diluyen en serie y se clonan para dar líneas celulares productoras de anticuerpos individuales a partir de las cuales pueden propagarse entonces clones indefinidamente para proporcionar Acm. Las líneas celulares pueden aprovecharse para la producción de Acm en dos modos básicos. Puede inyectarse una muestra del hibridoma (a menudo en la cavidad peritoneal) en un animal histocompatible del tipo que se usó para proporcionar las células de mieloma y somáticas para la fusión original. El animal inyectado desarrolla tumores que secretan el anticuerpo monoclonal específico producido por el híbrido celular fusionado. Los líquidos corporales del animal, tales como suero o líquido ascítico, pueden extraerse entonces para proporcionar Acm a alta concentración. 10 Las líneas celulares individuales también podrían cultivarse *in vitro*, en donde los Acm se secretan de manera natural al medio de cultivo a partir del cual pueden obtenerse fácilmente en altas concentraciones. Los Acm producidos por cualquier medio pueden purificarse adicionalmente, si se desea, usando filtración, centrifugación y diversos métodos cromatográficos tales como HPLC o cromatografía de afinidad.

15 Tras el aislamiento y la caracterización del anticuerpo monoclonal deseado, el ARNm puede aislarse usando técnicas bien conocidas en la técnica y usarse como molde para la amplificación de la secuencia diana.

Están disponibles varios procedimientos dependientes de molde para amplificar las secuencias diana antes y después de la mutagénesis. Uno de los métodos de amplificación mejor conocidos es la reacción en cadena de la polimerasa (denominada PCR) que se describe en detalle en las patentes estadounidense n.ºs 4.683.195, 4.683.202 20 y 4.800.159, y en Innis *et al.*, 1990. En resumen, en la PCR, se preparan dos secuencias de cebador que son complementarias a regiones en hebras complementarias opuestas de la secuencia diana. Se añade un exceso de desoxinucleósidos trifosfato a una mezcla de reacción junto con una ADN polimerasa, por ejemplo, Taq polimerasa. Si la secuencia diana está presente en una muestra, los cebadores se unirán a la diana y la polimerasa provocará que los cebadores se extiendan a lo largo de la secuencia diana añadiendo nucleótidos. Elevando y disminuyendo la temperatura de la mezcla de reacción, los cebadores extendidos se disociarán de la diana para formar productos de reacción, los cebadores en exceso se unirán a la diana y a los productos de reacción y se repite el procedimiento. Preferiblemente puede realizarse un procedimiento de amplificación por PCR con transcriptasa inversa con el fin de cuantificar la cantidad de diana amplificada. Se conocen bien en la técnica metodologías de reacción en cadena de la polimerasa. Usando técnicas de amplificación enzimáticas tales como PCR, pueden diseñarse elementos de control deseados en el cebador y, por tanto, se incorporarán en el producto de ADN. 30

Otro método para la amplificación es la reacción en cadena de la ligasa ("LCR"), dada a conocer en el documento EPA n.º 320 308. En la LCR, se preparan dos pares de sondas complementarias, y en presencia de la secuencia diana, cada par se unirá a hebras complementarias opuestas de la diana de manera que quedan contiguas. En presencia de una ligasa, los dos pares de sondas se unirán para formar una unidad individual. Mediante ciclos de temperatura, como en la PCR, las unidades ligadas unidas se disocian de la diana y entonces sirven como "secuencias diana" para el ligamiento de pares de sondas en exceso. La patente estadounidense n.º 4.883.750 describe un método similar a LCR para unir pares de sondas a una secuencia diana. 35

También puede usarse Qbeta replicasa, descrita en la solicitud PCT n.º PCT/US87/00880, como método de amplificación. En este método, se añade una secuencia de replicación de ARN que tiene una región complementaria a la de una diana a una muestra en presencia de una ARN polimerasa. La polimerasa copiará la secuencia de replicación que entonces puede detectarse. 40

Un método de amplificación isotérmica, en el que se usan endonucleasas de restricción y ligasas para lograr la amplificación de moléculas diana que contienen 5'-[alfa-tio]-trifosfatos de nucleótido en una hebra de un sitio de restricción también puede ser útil en la amplificación de ácidos nucleicos (Walquer *et al.*, 1992).

45 La amplificación por desplazamiento de hebra (SDA) es otro método de llevar a cabo la amplificación isotérmica de ácidos nucleicos que implica múltiples tandas de desplazamiento de hebra y síntesis, es decir, traducción de mellas. Un método similar, denominado reacción en cadena de reparación (RCR) implica aparear varias sondas a lo largo de toda una región seleccionada como diana para la amplificación, seguido por una reacción de reparación en la que sólo dos de las cuatro bases están presentes. Las otras dos bases pueden añadirse como derivados biotinilados para lograr una fácil detección. Se usa un enfoque similar en SDA. También pueden detectarse secuencias específicas de diana usando una reacción de sonda cíclica (CPR). En CPR, una sonda que tiene secuencias en 3' y 5' de ADN no específico y secuencia media de ARN específico se hibrida con ADN que está presente en una muestra. Tras la hibridación, se trata la reacción con ARNasaH, y se identifican los productos de la sonda como productos distintivos que se liberan tras digestión. El molde original se aparea con otra sonda de ciclado y se repite la reacción. 55

Se describen otros métodos de amplificación en la solicitud de GB n.º 2 202 328, y en la solicitud PCT n.º PCT/US89/01025, pueden usarse según la presente invención. En la primera solicitud, se usan cebadores "modificados" en una síntesis dependiente de molde y enzima, similar a PCR. Los cebadores pueden modificarse mediante marcaje con un resto de captura (por ejemplo, biotina) y/o un resto detector (por ejemplo, enzima). En la

última solicitud, se añade un exceso de sondas marcadas a una muestra. En presencia de la secuencia diana, la sonda se une y se escinde catalíticamente. Tras la escisión, se libera la secuencia diana intacta a la que va a unirse la sonda en exceso. La escisión de la sonda marcada señala la presencia de la secuencia diana.

5 Otros procedimientos de amplificación de ácidos nucleicos incluyen sistemas de amplificación basados en transcripción (TAS), incluyendo amplificación basada en la secuencia de ácido nucleico (NASBA) y 3SR (Kwoh *et al.*, 1989). En NASBA, los ácidos nucleicos pueden prepararse para la amplificación mediante extracción con fenol/cloroformo convencional, desnaturalización térmica de una muestra clínica, tratamiento con tampón de lisis y columnas de minicentrifugación para el aislamiento de ADN y ARN o extracción de ARN con cloruro de guanidinio. Estas técnicas de amplificación implican aparear un cebador que tiene secuencias específicas de diana. Tras la polimerización, se digieren los híbridos de ADN/ARN con RNasa H mientras que se desnaturalizan térmicamente de nuevo moléculas de ADN bicatenarias. En cualquier caso, el ADN monocatenario se hace completamente bicatenario mediante la adición de un segundo cebador específico de diana, seguida por polimerización. Las moléculas de ADN bicatenario se transcriben entonces de manera múltiple mediante una polimerasa tal como T7 o SP6. En una reacción cíclica isotérmica, los ARN se transcriben de manera inversa para dar ADN bicatenario, y se transcriben de nuevo con una polimerasa tal como T7 o SP6. Los productos resultantes, ya estén truncados o completos, indican secuencias específicas de diana.

20 Davey *et al.*, documento EPA n.º 329 822 dan a conocer un procedimiento de amplificación de ácido nucleico que implica sintetizar de manera cíclica ARN monocatenario ("ARNmc"), ADNmc, y ADN bicatenario (ADNbc), que pueden usarse según la presente invención. El ARNmc es un primer molde para un primer oligonucleótido cebador, que se alarga mediante la transcriptasa inversa (ADN polimerasa dependiente de ARN). El ARN se elimina entonces del dúplex de ADN:ARN resultante mediante la acción de ribonucleasa H (RNasa H, una RNasa específica de ARN en dúplex con o bien ADN o bien ARN). El ADNmc resultante es un segundo molde para un segundo cebador, que también incluye las secuencias de un promotor de ARN polimerasa (ejemplificada por ARN polimerasa de T7) en 5' con respecto a su homología al molde. Este cebador se extiende entonces mediante ADN polimerasa (ejemplificada por el fragmento "Klenow" grande de ADN polimerasa I de *E. coli*), dando como resultado una molécula de ADN bicatenario ("ADNbc"), que tiene una secuencia idéntica a la del ARN original entre los cebadores y que tiene adicionalmente, en un extremo, una secuencia de promotor. Esta secuencia de promotor puede usarse por la ARN polimerasa apropiada para realizar muchas copias de ARN del ADN. Estas copias pueden reintroducirse entonces en el ciclo conduciendo a una amplificación muy rápida. Con la elección apropiada de enzimas, esta amplificación puede hacerse de manera isotérmica sin adición de enzimas en cada ciclo. Debido a la naturaleza cíclica de este procedimiento, puede elegirse que la secuencia de partida esté en forma de o bien ADN o bien ARN.

35 Miller *et al.*, solicitud PCT WO 89/06700 dan a conocer un esquema de amplificación de secuencias de ácido nucleico basado en la hibridación de una secuencia de promotor/cebador a un ADN monocatenario diana ("ADNmc") seguido por transcripción de muchas copias de ARN de la secuencia. Este esquema no es cíclico, es decir, no se producen nuevos moldes a partir de los transcritos de ARN resultantes. Otros métodos de amplificación incluyen "race" y "PCR unilateral" (Frohman, 1990; O'Hara *et al.*, 1989).

También pueden usarse métodos basados en el ligamiento de dos (o más) oligonucleótidos en presencia de ácido nucleico que tiene la secuencia del "di-oligonucleótido" resultante, amplificando de ese modo el dioligonucleótido, en la etapa de amplificación (Wu *et al.*, 1989).

40 Los productos de amplificación pueden analizarse mediante electroforesis en gel de agarosa, agarosa-acrilamida o poli(acrilamida) usando métodos convencionales (véase, por ejemplo, Maniatis *et al.*, 1982). Por ejemplo, puede usarse un gel de agarosa al 1% teñido con bromuro de etidio y visualizarse bajo luz UV. Alternativamente, los productos de amplificación pueden marcarse de manera integral con nucleótidos radiomarcados o marcados fluorométricamente. Entonces pueden exponerse geles a película de rayos x o visualizarse bajo los espectros de estimulación apropiados, respectivamente.

Los procedimientos mutagénicos de la presente invención pueden comprender cualquier enfoque mutagénico que pueda adaptarse a un sitio particular en un gen, es decir, mutagénesis específica de sitio o dirigida al sitio. Debido a que la presente invención se basa en mutagénesis completa, la presente invención contempla como realizaciones preferidas los procedimientos mutagénicos que son rápidos, eficientes y económicos.

50 En una realización, el procedimiento mutagénico utiliza técnicas de síntesis química. Al hacer esto, es posible colocar exactamente la sustitución en una o más ubicaciones particulares dentro del gen, y también definir específicamente la naturaleza de las alteraciones. Se conocen bien métodos de síntesis química para ADN dentro de la técnica. Se prefieren en este sentido técnicas de fase sólida.

55 Una ventaja del método de fase sólida de síntesis génica es la oportunidad de mutagénesis usando técnicas de síntesis combinatoria. Las técnicas de síntesis combinatoria se definen como las técnicas que producen grandes colecciones o bibliotecas de compuestos simultáneamente, uniendo secuencialmente diferentes unidades estructurales. Pueden construirse bibliotecas usando compuestos libres en disolución, pero preferiblemente el compuesto se une a un soporte sólido tal como una perla, una partícula sólida o incluso se presenta sobre la superficie de un microorganismo.

Existen varios métodos para la síntesis combinatoria (Holmes *et al.*, 1995; Burbaum *et al.*, 1995; Martin *et al.*, 1995; Freier *et al.*, 1995; Pei *et al.*, 1991; Bruce *et al.*, 1995; Ohlmeyer *et al.*, 1993), incluyendo síntesis dividida o síntesis paralela. La síntesis dividida puede usarse para producir pequeñas cantidades de un número relativamente grande de compuestos, mientras que la síntesis paralela producirá mayores cantidades de un número relativamente pequeño de compuestos. En términos generales, usando la síntesis dividida, se sintetizan compuestos sobre la superficie de una micropartícula. En cada etapa, las partículas se reparten en varios grupos para la adición del siguiente componente. Los diferentes grupos se recombinan entonces y se reparten para formar nuevos grupos. El procedimiento se repite hasta que se completa el compuesto. Cada partícula contiene varias copias del mismo compuesto permitiendo una fácil separación y purificación. La síntesis dividida sólo puede realizarse usando un soporte sólido.

Una técnica alternativa conocida como síntesis paralela puede realizarse o bien en base sólida o bien en disolución. Usando la síntesis paralela, se sintetizan diferentes compuestos en receptáculos independientes, usando a menudo automatización. La síntesis paralela puede realizarse en una placa de microtitulación en donde pueden añadirse diferentes reactivos a cada pocillo de una manera predefinida para producir una biblioteca combinatoria. La síntesis paralela es el enfoque preferido para su uso con técnicas enzimáticas. Se entenderá bien que existen muchas modificaciones de esta técnica y pueden adaptarse para su uso con la presente invención. Usando métodos combinatorios, puede sintetizarse un gran número de moldes de genes mutantes.

También pueden generarse genes mutantes mediante métodos semisintéticos conocidos en la técnica (Barbas *et al.*, 1992). Usando las regiones conservadas de un fragmento de anticuerpo como entramado, pueden insertarse regiones variables en combinaciones aleatorias una o más cada vez para alterar la especificidad del fragmento de anticuerpo y generar sitios de unión novedosos, especialmente en la generación de anticuerpos frente a antígenos no propicios para la inmunización tales como compuestos tóxicos o lábiles. Del mismo modo, puede variarse una secuencia de anticuerpo conocida introduciendo mutaciones aleatoriamente. Esto puede lograrse mediante métodos bien conocidos en la técnica tal como el uso de PCR sensible a errores.

Usando los cebadores oligonucleotídicos apropiados, se usa PCR para la síntesis rápida del molde de ADN que contiene una o más mutaciones en el gen de la proteína de unión. La mutagénesis específica de sitio es una técnica útil en la preparación de péptidos individuales, o proteínas o péptidos equivalentes funcionales biológicamente, a través de mutagénesis específica del ADN subyacente. La técnica proporciona además una fácil capacidad para preparar y someter a prueba variantes de secuencia, incorporando una o más de las consideraciones anteriores, introduciendo uno o más cambios de secuencia de nucleótidos en el ADN. La mutagénesis específica de sitio permite la producción de mutantes a través del uso de secuencias de oligonucleótidos específicas que codifican para la secuencia de ADN de la mutación deseada, así como un número suficiente de nucleótidos adyacentes, para proporcionar una secuencia de cebador de tamaño y complejidad de secuencia suficientes para formar un dúplex estable en ambos lados de la unión de delección que está atravesándose. Normalmente, se prefiere un cebador de 17 a 25 nucleótidos de longitud, con de 5 a 10 residuos en ambos lados de la unión de la secuencia que está alterándose.

En general, la mutagénesis dirigida al sitio se realiza obteniendo en primer lugar un vector monocatenario, o fusionando dos hebras de un vector bicatenario que incluye dentro de su secuencia una secuencia de ADN que codifica para la proteína deseada. Se prepara de manera sintética un cebador de oligonucleótido que lleva la secuencia mutada deseada. Este cebador se aparea entonces con la preparación de ADN monocatenario, teniendo en cuenta el grado de apareamiento erróneo cuando se seleccionan condiciones de hibridación, y se somete a enzimas de polimerización de ADN tales como fragmento Klenow de polimerasa I de *E. coli*, con el fin de completar la síntesis de la hebra que lleva la mutación. Por tanto, se forma un heterodúplex en el que una hebra codifica para la secuencia no mutada original y la segunda hebra lleva la mutación deseada. Este vector de heterodúplex se usa entonces para transformar células apropiadas, tales como células de *E. coli*, y se seleccionan clones que incluyen vectores recombinantes que llevan la disposición de secuencia mutada.

La preparación de variantes de secuencia del gen seleccionado usando mutagénesis dirigida al sitio se proporciona como medio de producción de especies potencialmente útiles y no pretende ser limitativa, ya que hay otros modos en los que pueden obtenerse variantes de secuencia. Por ejemplo, pueden tratarse vectores recombinantes que codifican para el gen deseado con agentes mutagénicos, tales como hidroxilamina, para obtener variantes de secuencia.

En determinadas aplicaciones, sustitución de aminoácidos mediante mutagénesis dirigida al sitio, se aprecia que se requieren condiciones de menor rigurosidad. En estas condiciones, puede producirse hibridación aun cuando las secuencias de sonda y hebra diana no son perfectamente complementarias, sino que tienen apareamientos erróneos en una o más posiciones. Las condiciones pueden hacerse menos rigurosas aumentando la concentración de sal y disminuyendo la temperatura. Por ejemplo, podría proporcionarse una condición de rigurosidad media mediante NaCl de 0,1 a 0,25 M a temperaturas de 37°C a 55°C, mientras que podría proporcionarse una condición de baja rigurosidad mediante sal de 0,15 M a 0,9 M, a temperaturas que oscilan entre 20°C y 55°C. Por tanto, las condiciones de hibridación pueden manipularse fácilmente, y por tanto serán generalmente un método de elección dependiendo de los resultados deseados.

En otras realizaciones, puede lograrse la hibridación en condiciones de, por ejemplo, Tris-HCl 50 mM (pH 8,3), KCl 75 mM, MgCl₂ 3 mM, ditioneitol 10 mM, a temperaturas de entre aproximadamente 20°C y 37°C. Otras condiciones de hibridación utilizadas podrían incluir Tris-HCl aproximadamente 10 mM (pH 8,3), KCl 50 mM, MgCl₂ 1,5 μM, a temperaturas que oscilan entre aproximadamente 40°C y 72°C. También pueden usarse formamida y SDS para alterar las condiciones de hibridación.

En una realización particular, puede emplearse PCR de solapamiento. En resumen, se usa un plásmido como molde para la primera tanda de PCR. Los productos de PCR de la primera tanda se purifican y se usan, junto con cebadores externos, en la reacción de PCR de extensión de solapamiento. Los productos finales contienen la sustitución dirigida al sitio de un aminoácido dado por todos los demás residuos de aminoácido posibles.

El molde de ADN mutagenizada para el polipéptido de interés se puede clonar en un plásmido para la transcripción/traducción in vitro o en la realización preferida, los elementos de control apropiados se incluyen dentro del producto de la PCR para la transcripción/traducción directa in vitro. La transcripción/traducción in vitro de genes utiliza extractos libres de células para proporcionar las enzimas, los ribosomas y los factores proteicos necesarios. La síntesis de proteínas está dirigida por ARNm sintetizado a partir de los moldes de ADN deseadas. El molde de ADN debe contener los elementos de control apropiados para el sistema utilizado, incluido un sitio de unión a ribosoma y una secuencia promotora. Un experto en la técnica reconocerá claramente los elementos requeridos apropiados para cada sistema.

Los polipéptidos mutantes generados mediante la presente invención pueden caracterizarse usando una variedad de técnicas. En general, pueden analizarse productos proteicos para determinar el peso molecular aparente correcto usando SDS-PAGE. Esta proporciona una indicación inicial de que el polipéptido, de hecho, se sintetizó. Cuando se compara con la molécula natural, también indica si está teniendo lugar un plegamiento o procesamiento normal con el mutante. En este sentido, puede comprobarse que es útil marcar el polipéptido. Alternativamente, el polipéptido puede identificarse mediante tinción del gel.

Más allá de la mera síntesis, las proteínas pueden caracterizarse según diversas propiedades y una gama extensa de funciones. Las propiedades incluyen punto isoeléctrico, estabilidad térmica, velocidad de sedimentación y plegamiento. Una manera de examinar el plegamiento es la capacidad para reconocerse por una pareja de unión relacionada. El principal ejemplo de esta función es la interacción anticuerpo-antígeno. Está disponible una amplia variedad de diferentes formatos de inmunoensayo para este fin y se conocen bien en la técnica. Principalmente, pueden determinarse cambios en o bien la afinidad o bien la especificidad cuando se pone en contacto la proteína con un ligando específico o paneles de ligandos relacionados.

Los inmunoensayos pueden dividirse generalmente en dos tipos: ensayos heterogéneos que requieren múltiples etapas de separación, y ensayos homogéneos que se realizan directamente. Los inmunoensayos heterogéneos en general implican un ligando o un anticuerpo inmovilizado sobre una matriz sólida. Se pone en contacto una muestra que contiene un ligando con el anticuerpo inmovilizado y se determina la cantidad de complejo formado sobre el soporte de matriz a partir de un marcador unido directa o indirectamente al complejo inmovilizado. Tal como se usa en el contexto de la presente invención, ligando se define como una especie que interacciona con una molécula no idéntica formando un complejo estable, fuertemente unido. Para fines prácticos, la afinidad de unión es habitualmente mayor de 10⁶ M⁻¹ y está preferiblemente en el intervalo de 10⁹ - 10¹⁵ M⁻¹. El ligando puede ser cualquiera de varios tipos de moléculas orgánicas, incluyendo hidratos de carbono alicíclicos, compuestos aromáticos polinucleares, compuestos halogenados, bencenoides, hidratos de carbono polinucleares, compuestos heterocíclicos de nitrógeno, compuestos heterocíclicos de azufre, compuestos heterocíclicos de oxígeno, e hidratos de carbono de alcano, alqueno, alquino, etc. Son de particular interés moléculas biológicas, incluyendo aminoácidos, péptidos, proteínas, líquidos, sacáridos, ácidos nucleicos y combinaciones de los mismos. Por supuesto se entenderá que estos son a modo de ejemplo sólo y que pueden aplicarse métodos de inmunoensayo contemplados para detectar una gama extraordinariamente amplia de compuestos, siempre que pueda obtenerse un anticuerpo que se une al ligando de interés.

Pueden realizarse inmunoensayos heterogéneos como ensayos de tipo sándwich en los que una molécula de interés se hace reaccionar con un anticuerpo inmovilizado que se une específicamente a esa molécula con una afinidad alta. En una segunda etapa, un conjugado formado a partir del mismo o diferente anticuerpo frente al antígeno y una molécula marcadora se hace reaccionar con el complejo antígeno-anticuerpo sobre la matriz de inmovilización. Tras la eliminación del conjugado de marcador libre en exceso, se mide el conjugado de marcador unido, que es proporcional a la cantidad de ligando en la muestra.

La detección de la formación de inmunocomplejos se conoce bien en la técnica y puede lograrse a través de la aplicación de numerosos enfoques. Estos enfoques se basan normalmente en la detección de una marca o un marcador, tal como cualquiera de las etiquetas o marcadores radioactivos, fluorescentes, quimioluminiscentes, electroquimioluminiscentes, biológicos o enzimáticos conocidos en la técnica. Las patentes estadounidenses referentes al uso de tales marcadores incluyen las patentes estadounidense n.ºs 3.817.837; 3.850.752; 3.939.350; 3.996.345; 4.277.437; 4.275.149 y 4.366.241. Por supuesto, pueden encontrarse ventajas adicionales a través del uso de un ligando de unión secundario tal como un segundo anticuerpo o una disposición de unión de ligandos de biotina/avidina, tal como se conoce en la técnica.

Los métodos preferidos para la detección incluyen radioinmunoensayo (RIA) o ensayo de inmunoabsorción ligado a enzimas (ELISA), siendo ELISA el más preferido debido a un aumento general de la sensibilidad. Los ELISA se usan de manera extensa en aplicaciones de biotecnología, particularmente como inmunoensayos para una amplia gama de sustancias antigénicas. La sensibilidad del ELISA se basa en la amplificación enzimática de la señal.

5 Otras proteínas preferidas contempladas para uso de acuerdo con la presente invención son aquellas que tienen un ensayo conveniente para la actividad. Los ejemplos representativos de las interacciones objetivo incluyen catálisis, interacciones enzima-sustrato, interacciones proteína-ácido nucleico, interacciones receptor-ligando e interacciones proteína-metal. En estos ensayos, las proteínas mutantes pueden compararse con la proteína de tipo natural por los cambios en la capacidad para realizar cualquiera de las funciones anteriores.

10 Tal como se usa en el presente documento, el término "poner en contacto" se define como llevar los componentes de la reacción a proximidad estrecha suficiente entre sí para permitir que se produzca la interacción deseada. La puesta en contacto puede lograrse mezclando los componentes en disolución, por ejemplo, o mediante interacción heterogénea tal como mediante contacto de flujo a través de una columna o matriz de inmovilización que se une a uno de los componentes.

15 Para proteínas mutantes que tienen una actividad catalítica, la reacción apropiada puede monitorizarse para detectar un cambio en la tasa catalítica o una alteración en la especificidad.

Los anticuerpos producidos y aislados mediante el método de la invención se seleccionan para unirse a una diana predeterminada. Normalmente, la diana predeterminada se seleccionará en vista de su aplicabilidad como diana de diagnóstico y/o terapéutica. La diana predeterminada puede ser un epítipo conocido o desconocido. Los anticuerpos se unen generalmente a un antígeno predeterminado (por ejemplo, el inmunógeno) con una afinidad de al menos $1 \times 10^7 \text{ M}^{-1}$, preferiblemente con una afinidad de al menos $5 \times 10^7 \text{ M}^{-1}$ más preferiblemente con una afinidad de al menos $1 \times 10^8 \text{ M}^{-1}$ a $1 \times 10^9 \text{ M}^{-1}$ o más, algunas veces hasta $1 \times 10^{10} \text{ M}^{-1}$ o más. Frecuentemente, el antígeno predeterminado es una proteína humana, tal como por ejemplo un antígeno de superficie celular humano (por ejemplo, CD4, CD8, receptor de IL-2, receptor de EGF, receptor de PDGF), otra macromolécula biológica humana (por ejemplo, trombosmodulina, proteína C, antígeno de hidrato de carbono, antígeno de Lewis sializado, L-selectina), o macromolécula asociada a enfermedad no humana (por ejemplo, LPS bacteriano, proteína de la cápside de virión o glicoproteína de la envuelta) y similares.

En otro ejemplo, se han publicado varios informes de la utilidad de diagnóstico y terapéutica de scFv (Gruber *et al.*, 1994 op. cit.; Lilley *et al.*, 1994 op. cit.; Huston *et al.*, Int. Rev. Immunol 1993, 10:a 195, Sandhu JS, Crit. Rev. Biotechnol., 1992,12: 437).

Pueden diseñarse por ingeniería genética anticuerpos de alta afinidad de la especificidad deseada y expresarse en una variedad de sistemas. Por ejemplo, se han producido scFv en plantas (Firek *et al.* (1993) Plant Mol. Biol. 23: 861). Además, los anticuerpos de cadena sencilla pueden usarse como base para construir anticuerpos completos o diversos fragmentos de los mismos (Kettleborough *et al.*, Euro J. Immunol., 1994, 24: 952). La secuencia que codifica para la región variable puede aislarse (por ejemplo, mediante amplificación por PCR o subclonación) y someterse a corte y empalme para dar una secuencia que codifica para una región constante humana deseada para codificar para un anticuerpo de secuencia humana más adecuado para usos terapéuticos en seres humanos en donde se minimiza preferiblemente la inmunogenicidad. El/los polinucleótido(s) que tiene(n) la(s) secuencia(s) codificante(s) completamente humana(s) pueden expresarse en una célula huésped (por ejemplo, a partir de un vector de expresión en una célula eucariota) y purificarse para dar una formulación farmacéutica.

Los constructos de expresión de ADN incluirán normalmente una secuencia de ADN de control de la expresión operativamente unida a las secuencias codificantes, incluyendo regiones promotoras heterólogas asociadas de manera natural. Preferiblemente, las secuencias de control de la expresión serán sistemas promotores eucariotas en vectores que pueden transformar o transfectar células huésped eucariotas. Una vez que el vector se ha incorporado en el huésped apropiado, el huésped se mantiene en condiciones adecuadas para una expresión de alto nivel de las secuencias de nucleótidos, y la recolección y purificación de los anticuerpos "diseñados por ingeniería genética" mutados.

Tal como se estableció anteriormente, las secuencias de ADN se expresarán en huéspedes después de que las secuencias se hayan unido operativamente a una secuencia de control de la expresión (es decir, situada para garantizar la transcripción y traducción del gen estructural). Estos vectores de expresión normalmente pueden replicarse en los organismos huésped o bien como episomas o bien como una parte integral del ADN cromosómico del huésped. Comúnmente, los vectores de expresión contendrán marcadores de selección, por ejemplo, tetraciclina o neomicina, para permitir la detección de las células transformadas con las secuencias de ADN deseadas (véase, por ejemplo, la patente estadounidense n.º 4.704.362).

55 Además de microorganismos eucariotas tales como levaduras, también pueden usarse cultivos celulares de tejidos de mamífero para producir los polipéptidos de la presente invención (véase, Winnacker, "From Genes to Clones", VCH Publishers, N.Y., N.Y. (1987)). Se prefieren células eucariotas, porque se han desarrollado en la técnica varias líneas de células huésped adecuadas que pueden secretar inmunoglobulinas intactas, e incluyen las líneas celulares

CHO, diversas líneas celulares COS, células HeLa, líneas celulares de mieloma, células B o hibridomas. Los vectores de expresión para estas células pueden incluir secuencias de control de la expresión, tales como un origen de replicación, un promotor, un potenciador (Queen *et al.*, Immunol. Rev. 1986, 89: 49), y sitios de procesamiento de información necesarios, tales como sitios de unión al ribosoma, sitios de corte y empalme de ARN, sitios de poliadenilación y secuencias de terminador de la transcripción. Secuencias de control de la expresión preferidas son promotores derivados de genes de inmunoglobulina, citomegalovirus, SV40, adenovirus, virus del papiloma bovino, y similares.

Puede aumentarse la transcripción de ADN eucariota insertando una secuencia de potenciador en el vector. Los potenciadores son secuencias de actuación en cis de entre 10 y 30 pb que aumentan la transcripción por un promotor. Los potenciadores pueden aumentar eficazmente la transcripción cuando están o bien en 5' o bien en 3' con respecto a la unidad de transcripción. También son eficaces si se ubican dentro de un intrón o dentro de la propia secuencia codificante. Normalmente, se usan potenciadores virales, incluyendo potenciadores de SV40, potenciadores de citomegalovirus, potenciadores de poliovirus y potenciadores de adenovirus. También se usan comúnmente secuencias de potenciador de sistemas de mamífero, tales como el potenciador de cadena pesada de inmunoglobulina de ratón.

Los sistemas de vector de expresión de mamífero también incluirán normalmente un gen marcador seleccionable. Los ejemplos de marcadores adecuados incluyen el gen de dihidrofolato reductasa (DHFR), el gen de timidina cinasa (TK) o genes procariotas que confieren resistencia a fármacos. Los primeros dos genes marcadores prefieren el uso de líneas celulares mutantes que carecen de la capacidad para crecer sin la adición de timidina al medio de crecimiento. Entonces pueden identificarse células transformadas por su capacidad para crecer en medios no complementados. Los ejemplos de genes de resistencia a fármacos procariotas útiles como marcadores incluyen genes que confieren resistencia a G418, ácido micofenólico e higromicina.

Los vectores que contienen los segmentos de ADN de interés pueden transferirse a la célula huésped mediante métodos bien conocidos, dependiendo del tipo de célula huésped. Por ejemplo, puede usarse tratamiento con fosfato de calcio, lipofección o electroporación para otras células huésped. Otros métodos usados para transformar células de mamífero incluyen el uso de polibreno, fusión de protoplastos, liposomas, electroporación y microinyección (véase, en general, Sambrook *et al.*, citado anteriormente).

Una vez expresados, los anticuerpos, cadenas de inmunoglobulina mutadas individuales, fragmentos de anticuerpo mutados, y otros polipéptidos de inmunoglobulina de la invención pueden purificarse según procedimientos convencionales de la técnica, incluyendo precipitación con sulfato de amonio, cromatografía en columna de fraccionamiento, electroforesis en gel y similares (véase, en general, Scopes, R., Protein Purification, Springer-Verlag, N.Y. (1982)). Una vez purificados, parcialmente o hasta la homogeneidad si se desea, los polipéptidos pueden usarse entonces terapéuticamente o en el desarrollo y la realización de procedimientos de ensayo, tinciones inmunofluorescentes, y similares (véase, en general, Immunological Methods, volúmenes I y II, Eds. Lefkovits and Pernis, Academic Press, N.Y. N.Y. (1979 y 1981)).

Los oligopéptidos de la presente invención pueden usarse para diagnóstico y terapia. A modo de ilustración y no de limitación, pueden usarse anticuerpos para tratar cáncer, enfermedades autoinmunitarias o infecciones virales. Para el tratamiento de cáncer, los anticuerpos se unirán normalmente a un antígeno expresado preferentemente en células cancerosas, tales como erbB-2, CEA, CD33, y muchos otros antígenos bien conocidos por los expertos en la técnica. Para el tratamiento de enfermedad autoinmunitaria, los anticuerpos se unirán normalmente a un antígeno expresado en células T, tal como CD4, el receptor de IL-2, los diversos receptores de antígenos de células T y muchos otros antígenos bien conocidos por los expertos en la técnica (por ejemplo, véase Fundamental Immunology, 2ª ed., W. E. Paul, ed., Raven Press: Nueva York, N.Y.). Para el tratamiento de infecciones virales, los anticuerpos se unirán normalmente a un antígeno expresado sobre células infectadas por un virus particular tal como las diversas glicoproteínas (por ejemplo, gB, gD, gE) de virus del herpes simple y citomegalovirus, y muchos otros antígenos bien conocidos por los expertos en la técnica (por ejemplo, véase Virology, 2ª ed., B. N. Campos *et al.*, eds., (1990), Raven Press: Nueva York, N.Y.).

Las composiciones farmacéuticas que comprenden anticuerpos de la presente invención son útiles para administración parenteral, es decir, por vía subcutánea, por vía intramuscular o por vía intravenosa. Las composiciones para administración parenteral comprenderán comúnmente una disolución del anticuerpo o un cóctel del mismo disuelto en un portador aceptable, preferiblemente un portador acuoso. Pueden usarse una variedad de portadores acuosos, por ejemplo, agua, agua tamponada, solución salina al 0,4%, glicina al 0,3% y similares. Estas disoluciones son estériles y generalmente están libres de materia particulada. Estas composiciones pueden esterilizarse mediante técnicas de esterilización convencionales, bien conocidas. Las composiciones pueden contener sustancias auxiliares farmacéuticamente aceptables según se requiere para aproximarse a las condiciones fisiológicas tales como agentes de ajuste y tamponamiento del pH, agentes de ajuste de la toxicidad y similares, por ejemplo acetato de sodio, cloruro de sodio, cloruro de potasio, cloruro de calcio, lactato de sodio, etc. La concentración de los anticuerpos mutantes en estas formulaciones puede variar ampliamente, es decir, desde menos del 0,01%, habitualmente al menos el 0,1% hasta tanto como el 5% en peso y se seleccionará principalmente basándose en los volúmenes, viscosidades de fluidos, etc., según el modo de administración particular seleccionado.

Por tanto, podría constituirse una composición farmacéutica típica para inyección intramuscular para que contuviese 1 ml de agua tamponada estéril, y 1 mg de anticuerpo mutante. Una composición típica para infusión intravenosa puede constituirse para que contenga 250 ml de solución de Ringer estéril, y 10 mg de anticuerpo mutante. Los expertos en la técnica conocerán o les resultarán evidentes métodos reales para preparar composiciones que pueden administrarse por vía parenteral y se describen en más detalle en, por ejemplo, Remington's Pharmaceutical Science, 20ª Ed., Mack Publishing Company, Easton, Pa. (2000).

Fabricación de candidatos y/o candidatos evolucionados

En una realización, el sistema eucariota es un sistema de mamífero seleccionado de uno del grupo que consiste en líneas celulares CHO, HEK293, IM9, DS-1, THP-1, Hep G2, COS, NIH 3T3, C33a, A549, A375, SK-MEL-28, DU 145, PC-3, HCT 116, Mia PACA-2, ACHN, Jurkat, MM1, OvcAR 3, HT 1080, Panc-1, U266, 769P, BT-474, Caco-2, HCC 1954, MDAMB-468, LnCAP, NRK-49F y SP2/0; y esplenocitos de ratón y PBMC de conejo.

En una realización, pueden usarse una amplia variedad de células huésped de mamífero en la fabricación del candidato, incluyendo células de fibroblastos (3T3, ratón; BHK21, hámster sirio), células epiteliales (MDCK, perro; Hela, ser humano; PtK1, rata canguro), células plasmáticas ((SP2/0 y NS0, ratón), células de riñón (293, ser humano; COS, mono), células de ovario (CHO, hámster chino), células embrionarias (R1 y E14,1, ratón; H1 y H9, ser humano; PER C.6, ser humano).

En determinados aspectos, los anticuerpos recombinantes se producen en líneas celulares CHO y NS0 y SP2/0. En un aspecto específico, el sistema de mamíferos es una línea celular CHO-S. Los sistemas de vector de expresión más frecuentemente usados son sistemas de expresión de glutamina sintetasa y otros basados en genes de dihidrofolato reductasa.

En otra realización, el sistema eucariota es un sistema de células de levadura. En un aspecto, el sistema de células de levadura se selecciona de células de *S. cerevisiae* o *Picchia*.

En el método de la presente invención, los huéspedes usados para examinar moléculas evolucionadas son iguales que los huéspedes usados para la fabricación de coincidencias. En otro aspecto de la presente invención, el sistema genético usado para el descubrimiento y la evolución de proteínas es exactamente igual que el sistema genético usado para fabricar la proteína para aplicaciones comerciales.

Fármacos biosimilares

Los fármacos biosimilares son productos terapéuticos basados en proteína que tienen una secuencia de aminoácidos idéntica (es decir composición química) que un fármaco ético aprobado que ya no está protegido por una patente. En un aspecto, el método de CIAO es particularmente relevante para fármacos biosimilares. Aunque es esencial producir la proteína terapéutica en una formulación y composición equivalentes, para que sea competitivo en el mercado, el fármaco biosimilar debe fabricarse tan rápida y económicamente como sea posible. Los medios de cultivo celular y el desarrollo de procedimientos son algunas de las partes más costosas y que requieren más tiempo de la preparación y producción de un fármaco biosimilar.

El cambio de los codones de mutación silenciosa dentro de una proteína terapéutica cambia el codón usado para la traducción de la proteína pero conserva la secuencia de aminoácidos dentro de la proteína. Estos cambios de codones en una variedad de posiciones dentro de una molécula, particularmente en el extremo amino-terminal pueden tener un impacto significativo sobre la expresión y en algunos casos incluso la glicosilación. En un aspecto, el método de CIAO se usa para seleccionar y hacer evolucionar los codones de mutación silenciosa en una proteína dentro de una célula huésped similar a la usada en última instancia para la fabricación. Por tanto, puede reducirse el tiempo de procesamiento debido a los mayores rendimientos de la proteína y al hecho de que la proteína se seleccionó para expresarse en la línea celular huésped, de modo que la mayoría de los problemas de fabricación tradicionales se han descartado de la molécula. Además, seleccionando las moléculas en medios de cultivo baratos, libres de suero, pueden seleccionarse moléculas con codones que permiten una fabricación y purificación baratas.

Sin elaboración adicional, se cree que un experto en la técnica puede utilizar, usando la descripción anterior, la presente invención en su extensión más completa.

Ejemplos

Ejemplo 1. Generación y examen de una biblioteca de anticuerpos

Este ejemplo describe el método de generar y examinar una biblioteca de anticuerpos humanos con presentación en la superficie de células de mamíferos para aislar anticuerpos humanos con actividad de unión a un antígeno diana usando la combinación de clasificación citométrica de flujo y ELISA.

Examen de biblioteca mediante análisis citométrico de flujo

1. Generar bibliotecas de anticuerpos humanos integrados de manera estable en células de mamífero tal como se describe en los apéndices 1.2, 1.3 y 1.4 a continuación.

ES 2 701 931 T3

2. Expandir clones de biblioteca de anticuerpos totalmente humanos estables antes del análisis citométrico de flujo.
3. En el día del análisis citométrico de flujo, lavar 1×10^7 células con 1 x PBS.
4. Desprender las células con medio de desprendimiento celular Detachin y recoger las células en 1 x PBS.
5. Centrifugar las células a 3000 rpm durante 5 minutos. Retirar el sobrenadante.
- 5 6. Resuspender el sedimento celular en 1 ml de 1 x PBS frío y centrifugar a 3000 rpm durante 5 minutos.
7. Retirar el sobrenadante y resuspender el sedimento celular en 500 μ l de 2 μ g/ml de proteína 001 humana purificada en 1 x PBS frío.
8. Incubar en hielo durante 1 hora con mezclado ocasional a mano.
9. Centrifugar las células a 3000 rpm durante 5 minutos. Retirar el sobrenadante.
- 10 10. Resuspender el sedimento celular en 1 ml de 1 x PBS frío y centrifugar a 3000 rpm durante 5 minutos.
11. Repetir las etapas 7 y 8.
12. Retirar el sobrenadante y resuspender el sedimento celular en 500 μ l de 1 ng/ml de anticuerpo policlonal de conejo anti-001 humana en 1 x PBS frío con suero de cabra al 10%.
13. Incubar en hielo durante 30 minutos con mezclado ocasional a mano.
- 15 14. Centrifugar las células a 3000 rpm durante 5 minutos. Retirar el sobrenadante.
15. Resuspender el sedimento celular en 1 ml de 1 x PBS frío y centrifugar a 3000 rpm durante 5 minutos.
16. Repetir las etapas 7 y 8.
17. Retirar el sobrenadante y resuspender el sedimento celular en 500 μ l de anticuerpo de cabra anti-conejo conjugado con FITC y anticuerpo de cabra anti-Fc humano conjugado con ficoeritrina en 1 x PBS frío con suero de cabra al 10%.
- 20 18. Incubar en hielo durante 30 minutos con mezclado ocasional a mano.
19. Centrifugar las células a 3000 rpm durante 5 minutos. Retirar el sobrenadante.
20. Resuspender el sedimento celular en 1 ml de 1 x PBS frío y centrifugar a 3000 rpm durante 5 minutos.
21. Repetir las etapas 7 y 8.
- 25 22. Retirar el sobrenadante y resuspender el sedimento celular en 1 ml de 1 x PBS frío con suero de cabra al 2%.
23. Continuar con el análisis citométrico de flujo usando MoFlo de Dako.
24. Preparar una ventana de clasificación para incluir el 0,1% superior de las células totales en lo que se refiere a la razón de fluorescencia PE/FITC. Recoger las células que se encuentran dentro de la ventana de clasificación en placas de 96 pocillos con 100 μ l de medios de crecimiento.
- 30 Recuperación de secuencias de región variable de cadena ligera y cadena pesada
1. Expandir los clones a partir de placas de 96 pocillos a placas de 6 pocillos. Cuando las células alcanzan el 80% de confluencia en las placas de 6 pocillos, continuar con el aislamiento de ADN genómico usando el kit DNeasy Tissue de Qiagen.
- 35 2. Retirar por aspiración los medios de las células. Añadir 500 μ l de 1 x PBS a cada 6 pocillos. Raspar las células de la placa con puntas de pipeta estériles. Transferir las células raspadas en PBS a un tubo de microcentrífuga estéril.
3. Centrifugar las células durante 5 minutos a 3000 rpm.
4. Retirar el sobrenadante y resuspender el sedimento celular en 200 μ l de 1 x PBS.
5. Añadir 20 μ l de proteinasa K y 200 μ l de tampón AL a la muestra, mezclar concienzudamente mediante agitación con vórtex e incubar a 56°C durante 10 minutos.
- 40 6. Añadir 200 μ l de etanol a la muestra y mezclar concienzudamente mediante agitación con vórtex.
7. Pipetear la mezcla de la etapa 6 al interior de una columna de centrifugación. Centrifugar a 8000 rpm durante un

ES 2 701 931 T3

minuto. Desechar el flujo pasante.

8. Añadir 500 µl de tampón AW1 y centrifugar durante un minuto a 8000 rpm. Desechar el flujo pasante.

9. Añadir 500 µl de tampón AW2 y centrifugar durante 2 minutos a 14.000 rpm. Desechar el flujo pasante. Centrifugar de nuevo durante un minuto a 14.000 rpm. Asegurarse de que la membrana está completamente seca.

5 10. Colocar la columna de centrifugación en un tubo de microcentrífuga estéril y pipetear 200 µl de tampón AE directamente sobre la membrana.

11. Incubar a temperatura ambiente durante un minuto y centrifugar durante un minuto a 8000 rpm para eluir el ADN genómico.

12. Someter a QC el ADN genómico preparando las siguientes reacciones en tubos de microcentrífuga de 1,5 ml:

ADNg	5 µl
10x tampón de carga de muestra	5 µl
Volumen total	10 µl

10 Cargar en un gel de agarosa al 0,8% - TAE bromuro de etidio 0,5 µg/ml. Usar marcador de tamaño molecular de ADN de 1 kb como patrón. Correr el gel a 100 V durante 20-30 minutos en 1X tampón TAE.

13. Preparar las siguientes reacciones de PCR en tubos de PCR estériles:

ADNg	1 µl
2x mezcla maestra HotStar Taq	12,5 µl
Cebador directo de dominio variable*	0,5 µl
Cebador inverso de dominio variable*	0,5 µl
H2O	10,5 µl
Volumen total	25 µl

*véase el apéndice 1.2

14. Colocar los tubos de PCR en el ciclador térmico y comenzar el programa de ciclación. Etapas de activación inicial: 15 minutos, 95°C, ciclación de 3 etapas

15 Desnaturalización: 40 segundos, 94°C

Apareamiento: 40 segundos, 55°C

Extensión: 2 minutos, 72°C

Número de ciclos: 30

Etapas de extensión final: 10 minutos, 72°C

20 15. Someter a QC las reacciones de PCR preparando las siguientes reacciones en tubos de microcentrífuga de 1,5 ml:

Reacción de PCR	5 µl
10x tampón de carga de muestra	5 µl
Volumen total	10 µl

Cargar en un gel de agarosa al 1% - TAE bromuro de etidio 0,5 µg/ml. Usar marcador de tamaño molecular de ADN de 1 kb como patrón. Correr el gel a 100 V durante 20-30 minutos en 1X tampón TAE.

16. Preparar las siguientes reacciones de clonación en tubos de microcentrífuga de 1,5 ml usando el kit TOPO 2.1

ES 2 701 931 T3

de Invitrogen:

Reacción de PCR	4 μ l
Solución salina	1 μ l
Vector TOPO	1 μ l
Volumen total	6 μ l

17. Mezclar las reacciones suavemente e incubar durante 5 minutos a temperatura ambiente.

18. Añadir 2 μ l de la reacción de clonación TOPO de la etapa 17 al interior de un vial de *E. coli* químicamente competente One Shot y mezclar suavemente.

5 19. Incubar en hielo durante 30 minutos.

20. Someter las células a choque térmico durante 30 segundos a 42°C.

21. Transferir los tubos a hielo e incubar durante 2 minutos.

22. Añadir 250 μ l de medio S.O.C. a temperatura ambiente.

23. Agitar los tubos horizontalmente a 37°C durante una hora a 200 rpm.

10 24. Extender 10 μ l de la transformación en una placa de LB-carbenicilina calentada nuevamente.

25. Incubar la placa durante la noche a 37°C.

26. Recoger 6 clones de cada transformación para la secuenciación.

27. Analizar las secuencias de región variable de cadena pesada y cadena ligera. Continuar a la segunda ronda de examen usando el método de ELISA.

15 Digerir vector y clones de anticuerpos humanos con enzimas de restricción

Las reacciones dependerán de la(s) enzima(s) de restricción elegida(s) y según las instrucciones del fabricante; se facilitan ejemplos en el presente documento: Preparar las siguientes reacciones de digestión en un tubo de microcentrífuga en hielo:

ADN de vector (2 μ g)	x μ l
10x tampón Rest Enz x	10 μ l
Agua libre de nucleasa	c.s. hasta 97 μ l
Rest Enz 1 (10 U/ μ l)	3 μ l
Rest Enz 2 (10 U/ μ l)	3 μ l
Volumen de reacción total	100 μ l
Clones de anticuerpos humanos (5 μ g)	x μ l
10X tampón Rest Enz x	10 μ l
Agua libre de nucleasa	c.s. hasta 97 μ l
Rest Enz 1 (10 U/ μ l)	3 μ l
Rest Enz 2 (10 U/ μ l)	3 μ l
Volumen de reacción total	100 μ l

ES 2 701 931 T3

1. Mezclar suavemente y centrifugar brevemente (5 s) en microcentrífuga.
 2. Incubar la reacción a 37°C durante la noche
- Digerir mediante CIP vector y purificar con el kit de purificación de PCR QIAquick
3. Añadir 2 µl de Apex fosfatasa al tubo de microcentrífuga que contiene la reacción de digestión del vector.
 - 5 4. Incubar a 37°C durante 10 minutos.
 5. Calentar a 70°C durante 5 minutos para inactivar la Apex fosfatasa.
 6. Añadir 500 µl de tampón PBI a la microcentrífuga.
 7. Mezclar mediante vórtex y centrifugar rápidamente.
 8. Cargar 750 µl cada vez en una columna.
 - 10 9. Centrifugar a 12.000 x g durante 1 minuto y decantar el líquido del tubo de recogida.
 10. Repetir hasta que se haya procesado toda la muestra.
 11. Lavar con 750 µl de tampón PE (con etanol añadido).
 12. Centrifugar a 12.000 x g durante 1 minuto y decantar el líquido del tubo de recogida.
 13. Colocar la columna de nuevo sobre el tubo de recogida y centrifugar de nuevo.
 - 15 14. Poner la columna sobre nuevos tubos de microcentrífuga y eluir con 50 µl de tampón EB.
- Purificar en gel clones de anticuerpos humanos digeridos con enzimas de restricción
1. Preparar las siguientes reacciones en un tubo de microcentrífuga de 1,5 ml:

Clones de anticuerpos totalmente humanos digeridos con Rest Enz	100 µl
10x tampón de carga de muestra	3 µl
<hr/>	
Volumen total	103 µl
2. Cargar en un gel de agarosa al 1%-TAE bromuro de etidio 0,5 µg/ml. Usar marcador de tamaño molecular de ADN de 1 kb como patrón. Correr el gel a 100 V durante 20-30 minutos en 1X tampón TAE.
 - 20 3. Recortar las bandas correspondientes a las regiones variables de cadena pesada (HC) y cadena ligera (LC) y purificarlas utilizando el kit de extracción de gel QIAquick.
 4. Añadir 3 volúmenes de tampón QG a 1 volumen de gel.
 5. Incubar a 50°C durante 10 minutos hasta que el trozo de gel se haya disuelto completamente. Añadir 1 volumen de gel de isopropanol a la muestra y mezclar.
 - 25 6. Colocar una columna de centrifugación QIAquick en un tubo de recogida de 2 ml facilitado.
 7. Aplicar la muestra a la columna QIAquick y centrifugar durante 1 minuto.
 8. Desechar el flujo pasante y colocar de nuevo la columna QIAquick en el mismo tubo de recogida.
 9. Añadir 0,75 ml de tampón PE a la columna QIAquick y centrifugar durante 1 minuto.
 - 30 10. Desechar el flujo pasante y centrifugar la columna QIAquick durante 1 minuto adicional a 17.900 x g (13.000 rpm).
 11. Colocar la columna QIAquick en un tubo de microcentrífuga de 1,5 ml limpio.
 12. Añadir 52 µl de tampón EB al centro de la membrana QIAquick y centrifugar la columna durante 1 minuto. Dejar reposar la columna durante 1 minuto y entonces centrifugar durante 1 minuto.

ES 2 701 931 T3

Ligar el dominio variable de HC y LC humano en ADN de vector digerido

Preparar la siguiente reacción de ligamiento en un tubo de microcentrifuga en hielo:

ADN de vector digerido (100 ng)	x μ l
Dominio variable de HC y LC humano	y μ l
5X tampón ligasa de T4	4 μ l
Agua libre de nucleasa	c.s. hasta 19 μ l
Ligasa de T4 (2.000 U/ μ l)	1 μ l
<hr/>	
Volumen de reacción total	20 μ l

1. Mezclar suavemente y centrifugar brevemente (5 s) en microcentrifuga.
 2. Incubar a temperatura ambiente durante 2 horas a 16°C durante la noche.
 - 5 3. Transformar cada una de las mezclas de la reacción de ligamiento en células *E. coli* supercompetentes.
 4. Enfriar previamente tubos de fondo redondo de polipropileno BD Falcon de 14 ml en hielo. Preparar medio SOC hasta 42°C.
 5. Descongelar las células supercompetentes en hielo. Cuando estén descongeladas, mezclar suavemente y llevar alícuotas de 100 μ l de células al interior de cada uno de los tubos enfriados previamente.
 - 10 6. Añadir 1,7 μ l de β -mercaptoetanol a cada alícuota de células. Incubar las células en hielo durante 10 minutos, agitando suavemente cada 2 minutos.
 7. Añadir 2 μ l de la mezcla de reacción de ligamiento a una alícuota de células. Golpear los tubos suavemente.
 8. Incubar los tubos en hielo durante 30 minutos.
 9. Calentar por impulsos los tubos en un baño de agua a 42°C durante 45 segundos.
 - 15 10. Incubar los tubos en hielo durante 2 minutos.
 11. Añadir 900 μ l de medio SOC precalentado e incubar los tubos a 37°C durante 1 hora con agitación a 225-250 rpm.
 12. Sembrar en placa 20 μ l y 200 μ l de la mezcla de transformación en placas de agar LB que contienen carbenicilina.
 - 20 13. Incubar las placas a 37°C durante la noche.
 14. Contar colonias en las placas y recoger 6 colonias para secuenciación y examen de PCR.
 15. Elegir un clon con la secuencia correcta, preparar ADN de plásmido y continuar con la transfección en células 293F.
- Transfección de células 293F
- 25 1. Una semana antes de la transfección, transferir células 293F a cultivo en monocapa en medio Eagle modificado por Dulbecco (D-MEM) complementado con suero.
 2. Un día antes de la transfección, sembrar en placa 0,1 x 10⁵ células en 100 μ l de D-MEM complementado con suero por muestra de transfección en formatos de 96 pocillos.
 3. Para cada muestra de transfección, preparar complejos de ADN-Lipofectamine.
 - 30 4. Diluir 0,2 μ g de ADN en 50 μ l de medio de suero reducido Opti-MEM. Mezclar suavemente.

ES 2 701 931 T3

5. Diluir 0,125 µl de Lipofectamine en 50 µl de medio de suero reducido Opti-MEM. Mezclar suavemente e incubar durante 5 min a temperatura ambiente.
6. Combinar el ADN diluido con la Lipofectamine diluida. Mezclar suavemente e incubar durante 20 min a temperatura ambiente.
- 5 7. Añadir los 100 µl de complejos de ADN-Lipofectamine a cada pocillo que contiene células y medio. Mezclar suavemente sacudiendo la placa hacia delante y hacia atrás.
8. Incubar las células a 37°C en una incubadora con el 5% de CO₂.
9. Añadir 100 µl de D-MEM complementado con suero a cada pocillo tras 6 horas. Incubar las células a 37°C en una incubadora con el 5% de CO₂ durante la noche.
- 10 10. Retirar por aspiración el medio en cada pocillo. Lavar cada pocillo con 100 µl de 293 SFM II con L-glutamina 4 mM. Añadir 100 µl de 293 SFM II con L-glutamina 4 mM a cada pocillo.
11. Recoger el sobrenadante para ELISA a las 96 horas tras la transfección.

Apéndice 1.1: Fórmulas de tampón

- 1 x PBS con suero de cabra al 2%
- 15
 - 2 ml de suero de cabra
 - 98 ml de 1 x PBS

50X tampón TAE

 - 242 g de base Tris
 - 57,1 ml de ácido acético glacial - 20
 - 37,2 g de Na₂EDTA-2H₂O
 - Añadir H₂O destilada hasta el volumen final de 1 litro

1X tampón TAE

 - 20 ml de 50X tampón TAE
 - 800 ml de H₂O destilada - 25
 - Gel de agarosa al 0,8% con bromuro de etidio
 - 0,8 g de agarosa LE
 - 100 ml de 1X tampón TAE
 - Fundir la agarosa en un horno microondas y agitar para garantizar un mezclado uniforme
 - Enfriar la agarosa hasta 55°C
 - 30
 - Añadir 2,5 µl de bromuro de etidio 20 mg/ml a la agarosa
 - Verter sobre una plataforma de gel

Gel de agarosa al 1% con bromuro de etidio

 - 1 g de agarosa LE
 - 100 ml de 1X tampón TAE - 35
 - Fundir la agarosa en un horno microondas y agitar para garantizar un mezclado uniforme
 - Enfriar la agarosa hasta 55°C
 - Añadir 2,5 µl de bromuro de etidio 20 mg/ml a la agarosa
 - Verter sobre una plataforma de gel

LB

- 10 g de NaCl
- 10 g de triptona
- 5 g de extracto de levadura

- 5
- Añadir H₂O destilada hasta un volumen final de 1 litro
 - Ajustar el pH a 7,0 con NaOH 5 N
 - Someter a autoclave

Agar LB-carbenicilina

- 10
- 10 g de NaCl
 - 10 g de triptona
 - 5 g de extracto de levadura
 - 20 g de agar
 - Añadir H₂O destilada hasta un volumen final de 1 litro
 - Ajustar el pH a 7,0 con NaOH 5 N
- 15
- Someter a autoclave
 - Enfriar hasta 55°C
 - Añadir 10 ml de carbenicilina 10 mg/ml esterilizada por filtración
 - Verter en placas Petri (placas de 25 ml/100 mm)

Medio SOC

- 20
- 0,5 g de NaCl
 - 20 g de triptona
 - 0,5 g de extracto de levadura
 - 2 ml de glucosa al 20% esterilizada por filtración
 - Añadir H₂O destilada hasta un volumen final de 1 litro
- 25
- Someter a autoclave
 - Añadir 10 ml de MgCl₂ 1 M esterilizado por filtración y 10 ml de MgSO₄ 1 M esterilizado por filtración antes de su uso

Disolución de lavado

- Tween-20 al 0,05% en PBS

30 Disolución de bloqueo

- Leche descremada Carnation al 2% en PBS

Suero bovino fetal inactivado por calor

- 500 ml de suero bovino fetal inactivado por calor en la botella original del vendedor
- Calentar durante 30 minutos a 56°C con mezclado cada 5 minutos

- 35
- Preparar alícuotas de 50 ml y almacenar a -20°C

Medio Eagle modificado por Dulbecco complementado con suero

- 500 ml de medio Eagle modificado por Dulbecco

- 50 ml de suero bovino fetal inactivado por calor
 - 5 ml de aminoácidos no esenciales-MEM 10 mM
- 293 SFM II con L-glutamina 4 mM

- 500 ml de SFM II

- 5 • 10 ml de L-glutamina 200 mM

Disolución de DEAE-dextrano

- 1 g de DEAE-dextrano (dietilaminoetil-dextrano)
- Disolver en 100 ml de agua destilada
- Esterilizar con filtro

- 10 Apéndice 1.2: Construcción de biblioteca de anticuerpos totalmente humanos

Pueden obtenerse todas las regiones V de cadena ligera kappa (V_k) y cadena pesada (V_H) de línea germinal humana funcionales de la base de datos Vbase (<http://vbase.mrc-cpe.cam.ac.uk/>) y alinearse. Entonces pueden analizarse las alineaciones con respecto a la diversidad, especialmente en las tres regiones de entramado. Entonces puede seleccionarse el número deseado de genes de V_H y V_k de las agrupaciones de secuencias resultantes para la construcción de la biblioteca.

- 15

Clonación de la región V

Se amplifican genes de la región V de cadena pesada y ligera (incluyendo las regiones de entramado 1, 2 y 3 y CDR1, CDR2 y CDR3) a partir de ADN genómico humano en dos fragmentos usando cebadores específicos de genes. Entonces se combinan genes parciales de la región V mediante PCR de solapamiento. Se añade un ligador a las regiones V de LC de longitud completa mediante PCR anidada antes de la clonación en un vector de expresión de mamífero. Se confirma la secuencia de los dominios variables LC clonados (produciendo clones V de LC). Se clonan mediante TOPO los dominios variables de la cadena pesada y se confirma la secuencia (clones V de HC mediante TOPO). Se amplifican las regiones V de HC de secuencia confirmada a partir de los plásmidos correspondientes, se añade un ligador al extremo 3' y se clonan los productos de PCR resultantes en los clones de dominio variable V de LC para formar combinaciones V_k/V_H y en un vector de mamífero.

- 20

- 25

Expresión de IgG de longitud completa

La expresión de cadena pesada de IgG1 y cadena ligera capa de longitud completa en el vector de mamífero deseado puede dirigirse a partir de un único promotor, por ejemplo, un promotor de CMV. Cada cadena va precedida por una señal de secreción que dirige la cadena de polipéptido naciente al retículo endoplasmático (ER). Una señal de anclaje puede fusionarse al extremo C-terminal de la cadena pesada. Esta señal se escinde y se sustituye por un anclaje que une la IgG de longitud completa al exterior de la membrana celular tras la secreción.

- 30

Apéndice 1.3: Generación de biblioteca de anticuerpos totalmente humanos estables en células de mamífero mediante transfección

- 35

1. Una semana antes de la transfección, transferir células CHO-S a cultivo en monocapa en medio Eagle modificado por Dulbecco (D-MEM) complementado con suero.

2. Un día antes de la transfección, sembrar en placa 6×10^6 células en 15 ml de D-MEM complementado con suero por muestra de transfección en una placa de cultivo tisular de 10 cm. Preparar diez placas de 10 cm.

3. Para cada placa de 10 cm, preparar complejos de ADN-Lipofectamine siguiendo las etapas 4-7.

- 40

4. Diluir 25 μ g de ADN de plásmido de biblioteca de anticuerpos totalmente humanos Maxi-prep en 1,5 ml de medio de suero reducido Opti-MEM. Mezclar suavemente.

5. Diluir 60 μ l de Lipofectamine en 1,5 ml de medio de suero reducido Opti-MEM. Mezclar suavemente e incubar durante 5 min a temperatura ambiente.

6. Combinar el ADN diluido con la Lipofectamine diluida. Mezclar suavemente e incubar durante 20 min a temperatura ambiente.

- 45

7. Añadir los 3 ml de complejos de ADN-Lipofectamine a cada placa que contiene células y medio. Mezclar suavemente sacudiendo la placa hacia delante y hacia atrás.

8. Incubar las células a 37°C en una incubadora con el 5% de CO₂ durante la noche.

ES 2 701 931 T3

9. Cambiar el medio de cada placa por 15 ml de D-MEM complementado con suero. Incubar las células a 37°C en una incubadora con el 5% de CO₂ durante 48 horas.
10. Desprender las células con medio de desprendimiento celular Detachin y resuspender las células en D-MEM complementado con suero.
- 5 11. Sembrar en placa 0,4 x 10⁶ células en 10 ml de D-MEM complementado con suero con G418 800 µg/ml en una placa de cultivo tisular de 10 cm. Transferir todas las células transfectadas dando como resultado 150 placas de 10 cm.
12. Sembrar en placa 0,4 x 10⁶ células CHO-S no transfectadas en 10 ml de D-MEM complementado con suero con G418 800 µg/ml en una placa de cultivo tisular de 10 cm.
- 10 13. Alimentar las células con D-MEM complementado con suero con G418 800 µg/ml cada 4 días.
14. Tras 14 días, inspeccionar las placas con células CHO-S no transfectadas. No debe haber células vivas en las placas.
15. Desprender las células transfectadas restantes con medio de desprendimiento celular Detachin y congelar las células en medios de congelación a 1x10⁷ células/ml.
- 15 Apéndice 1.4: Generación de biblioteca de anticuerpos totalmente humanos estables en células de mamífero mediante infección retroviral
 1. Un día antes de la transfección, sembrar en placa 6 x 10⁶ células EcoPack-2 293 en 15 ml de D-MEM complementado con suero por muestra de transfección en una placa de cultivo tisular de 10 cm. Preparar diez placas de 10 cm.
 - 20 2. Prepare el medio que contiene MBS. Esto se realiza inmediatamente antes de la transfección. Para cada placa de cultivo tisular de 10 cm, deben prepararse 12 ml de medio que contiene MBS.
 3. Añadir 12 ml de medio que contiene MBS a cada placa de 10 cm y devolver las placas a la incubadora a 37°C. Esto debe realizarse 20-30 minutos antes de la adición de la suspensión de ADN.
 - 25 4. Resuspender 10 µg de sedimento de ADN de plásmido de biblioteca de anticuerpos totalmente humanos Maxi-prep en 450 µl de H₂O estéril y transferir el ADN a tubos de fondo redondo de poliestireno BD Falcon de 5 ml separados.
 5. Añadir al ADN 50 µl de disolución I y 500 µl de disolución II del kit de transfección en mamíferos Transfection MBS de Stratagene.
 6. Resuspender suavemente cualquier precipitado en la suspensión de ADN pipeteando la suspensión hacia arriba y hacia abajo con una pipeta fijada a 500 µl.
 - 30 7. Incubar la suspensión de ADN a temperatura ambiente durante 10 minutos.
 8. Retirar las placas de 10 cm que van a transfectarse de la incubadora y añadir la suspensión de ADN sobre las placas en una forma gota a gota, agitando suavemente para impedir que las células se suban por la placa y para distribuir la suspensión de ADN uniformemente.
 - 35 9. Devolver las placas de cultivo tisular a la incubadora a 37°C.
 10. Tras incubar durante 3 horas, retirar el medio de las placas y sustituirlo por 4 ml de D-MEM complementado con suero complementado con cloroquina 25 µM. Devolver las placas a la incubadora a 37°C.
 11. Tras incubar durante 6-7 horas adicionales, retirar el medio de crecimiento que contiene cloroquina 25 µM y sustituirlo por 4 ml de D-MEM complementado con suero sin cloroquina.
 - 40 12. Incubar la placa en la incubadora a 37°C durante la noche.
 13. Retirar el medio de las placas y sustituirlo por 3,0 ml de D-MEM complementado con suero reciente. Devolver las placas a la incubadora a 37°C.
 14. Retirar las células de empaquetamiento productoras de virus de la incubadora.
 - 45 15. Recoger el sobrenadante que contiene virus de la primera placa y filtrarlo a través de un filtro de 0,45 µm al interior de un tubo cónico estéril de 50 ml.
 16. Llevar alícuotas del sobrenadante viral al interior de criovirales de 1,5 ml y someter a congelamiento rápido en un

ES 2 701 931 T3

baño de hielo seco/etanol. Almacenar el sobrenadante viral a -80°C.

17. Sembrar en placa 0,5 x 10⁶ células NIH-3T3 en 10 ml de D-MEM complementado con suero en una placa de cultivo tisular de 10 cm. Sembrar en placa 102 placas de cultivo tisular de 10 cm.

18. Someter a descongelación rápida el sobrenadante mediante agitación rápida en un baño de H₂O a 37°C.

5 19. Diluir el virus en DMEM complementado con suero de ternero con disolución de DEAE-dextrano con el título de 0,3 x 10⁵ partículas virales/ml. Preparar 3 ml de virus diluidos por placa de 100 mm que va a infectarse (se infectará el 20% de las células). Preparar "cóctel simulado" de medio de crecimiento más DEAE-dextrano que va a usarse como control negativo.

10 20. Retirar por aspiración el medio de las células NIH-3T3. Para cada placa, extender 3,0 ml virus diluidos uniformemente sobre las células. Devolver las placas a la incubadora a 37°C durante 3 horas.

21. Tras la incubación de 3 horas, añadir 7,0 ml adicionales de D-MEM complementado con suero de ternero a cada placa y devolver las placas a la incubadora a 37°C. Incubar la placa durante 48 horas.

22. Sustituir el medio por 10 ml de D-MEM complementado con suero de ternero con G418 800 µg/ml en cada placa de cultivo tisular de 10 cm.

15 23. Alimentar las células con D-MEM complementado con suero de ternero con G418 800 µg/ml cada 4 días.

24. Tras 14 días, inspeccionar las placas con células NIH-3T3 infectadas de manera simulada. No debe haber células vivas en la placa.

25. Desprender las células transfectadas restantes con medio de desprendimiento celular Detachin y congelar las células en medio de congelamiento at 1 x 10⁷ células/ml.

20 **Ejemplo 2. Reacciones para la evolución de posición completa (CPE)**

Reacción de mutagénesis

Se diseña un par de cebadores (mezcla de cebadores 1 y mezcla de cebadores 2) para cada codón que va a insertarse. El diseño dependerá de la secuencia génica, y pueden usarse bases de datos de análisis de secuencia tales como Sequencher (Gene Codes Corporation) o VectorNTI® (Life Technologies) para diseñar los cebadores. Para CPE, se diseña un par de cebadores para cada codón que va a mutarse. Un codón diana degenerado (NNK o NNN) está en el centro, flanqueado por 20 bases en cada lado (longitud total del cebador: 43 bases, clones 9f para secuenciación para identificar mutantes únicos). El ADN de molde es ADN de vector con gen(es) diana.

25

Preparar las siguientes reacciones en placas de PCR de pared delgada de 96 pocillos o tubos de PCR de pared delgada de 0,2 ml en hielo:

Mezcla de cebadores 1 (2,5 µM)	5 µl
Mezcla de cebadores 2 (2,5 µM)	5 µl
10X tampón de ADN polimerasa Pfu turbo	2,5 µl
Molde de ADN (5, 10 25 ng)	x µl
dNTP	2 µl
Agua libre de nucleasa	c.s. hasta 24,5 µl
ADN polimerasa <i>Pfu</i> turbo	0,5 µl
Volumen de reacción total	25 µl

30 1. Preparar una reacción de control negativo por una placa de 96 pocillos (sustituir cebadores por tampón TE).

2. Mezclar suavemente y centrifugar brevemente (5 s) en centrífuga de sobremesa.

3. Ciclar las reacciones usando los parámetros de ciclación explicados resumidamente a continuación:

ES 2 701 931 T3

Segmento	Ciclos	Temperatura	Tiempo
1	1	95°C	30 segundos
2	18	95°C	30 segundos
		55°C	1 minuto
		68°C	16 minutos

Análisis de control de calidad

1. Para someter a QC las reacciones de amplificación, preparar las siguientes reacciones en placas de PCR de pared delgada de 96 pocillos o tubos de PCR de pared delgada de 0,2 ml:

Reacción de mutagénesis	5 µl
Agua	4 µl
Tampón de carga de muestra	1 µl
Volumen	10 µl

- 5 2. Cargar 10 µl en un gel de agarosa al 1%-TAE con bromuro de etidio 0,5 µg/ml. Usar un marcador de tamaño molecular de ADN de 1 kb como patrón. Correr el gel a 100 V durante 20-30 minutos en 1X tampón TAE.

Digerir las reacciones de mutagénesis con enzimas de restricción apropiadas para clonar en ADN de vector - Ejemplo para enzima de restricción DpnI

1. Añadir 0,5 µl de la enzima de restricción DpnI (10 U/µl) directamente a cada reacción.
- 10 2. Mezclar suavemente y centrifugar brevemente (5 s) en una centrifuga de sobremesa.
3. Incubar a 37°C en máquinas de PCR durante 2 horas.
4. Transformar 6 mezclas de reacción de cada una de las placas de 96 pocillos en células supercompetentes XLI Blue. Almacenar el resto de las reacciones a -20°C.
5. Enfriar previamente tubos de PCR de 0,2 ml en hielo. Calentar medio SOC hasta 42°C.
- 15 6. Descongelar las células supercompetentes XLI Blue en hielo. Cuando estén descongeladas, mezclar suavemente y llevar alícuotas de 50 µl de células al interior de cada uno de los tubos enfriados previamente.
7. Añadir 0,8 µl de beta-mercaptoetanol a cada alícuota de células. Incubar las células en hielo durante 10 minutos, agitando suavemente cada 2 minutos.
8. Añadir 2 µl de la mezcla de reacción a una alícuota de células. Golpear los tubos suavemente.
- 20 9. Incubar los tubos en bloques fríos durante 30 minutos.
10. Calentar por impulsos los tubos en un baño de agua a 42°C durante 45 segundos.
11. Incubar los tubos en hielo durante 2 minutos.
12. Añadir 100 µl de medio SOC precalentado e incubar los tubos a 37°C durante 1 hora con agitación a 225-250 rpm.
- 25 13. Sembrar en placa la mezcla de transformación completa en placas de agar LB que contienen carbenicilina.
14. Incubar las placas a 37°C durante la noche.
15. Contar colonias en las placas y recoger 12 colonias de cada reacción de transformación para minipreparación y secuenciación.

Transformación a gran escala

ES 2 701 931 T3

1. Descongelar las células supercompetentes XLI Blue en hielo. Descongelar 20 tubos de células competentes para 96 reacciones. Cuando estén descongeladas, añadir 4 µl de β-mercaptoetanol a cada tubo de 250 µl de células competentes. Incubar las células en hielo durante 10 minutos, agitando suavemente cada 2 minutos.
2. Enfriar previamente tubos de PCR de 0,2 ml en hielo. Calentar medio SOC hasta 42°C.
- 5 3. Llevar alícuotas de 50 µl de células al interior de cada uno de los tubos enfriados previamente.
4. Añadir 2 µl de la mezcla de reacción a una alícuota de células. Golpear los tubos suavemente.
5. Incubar los tubos en bloques fríos durante 30 minutos.
6. Calentar por impulsos los tubos en un baño de agua a 42°C durante 45 segundos.
7. Incubar los tubos en hielo durante 2 minutos.
- 10 8. Añadir 100 µl de medio SOC precalentado e incubar los tubos a 37°C durante 1 hora con agitación a 225-250 rpm.
9. Sembrar en placa la mezcla de transformación completa en placas de agar LB que contienen carbenicilina.
10. Incubar las placas a 37°C durante la noche.
11. Hacer crecer las células en bloques de 96 pocillos para minipreparación.
12. Preparar minipreparación de ADN usando el kit QIAVac 96 siguiendo el protocolo de fabricación.

15 **Ejemplo 3. Examinar para determinar la mejora en la afinidad de los anticuerpos**

Transfección

Una semana antes de la transfección, transferir células 293F a cultivo en monocapa en medio Eagle modificado por Dulbecco (D-MEM) complementado con suero. Un día antes de la transfección, sembrar en placa 0,2 x 10⁵ y 0,4 x 10⁵ células en 100 µl de D-MEM complementado con suero por muestra de transfección en formatos de 96 pocillos.

- 20 1. Para cada muestra de transfección, preparar complejos de ADN-Lipofectamine.
2. Diluir 0,2 µg de ADN en 50 µl medio de suero reducido Opti-MEM. Mezclar suavemente.
3. Diluir 0,125 µl de Lipofectamine en 50 µl de medio de suero reducido Opti-MEM. Mezclar suavemente e incubar durante 5 min a temperatura ambiente.
- 25 4. Combinar el ADN diluido con la Lipofectamine diluida. Mezclar suavemente e incubar durante 20 min a temperatura ambiente.
5. Añadir los 100 µl de complejos de ADN-Lipofectamine a cada pocillo que contiene células y medio. Mezclar suavemente sacudiendo la placa hacia delante y hacia atrás.
6. Incubar las células a 37°C en una incubadora con el 5% de CO₂.
7. Añadir 100 µl de D-MEM complementado con suero a cada pocillo tras 6 horas. Incubar las células a 37°C en una incubadora con el 5% de CO₂ durante la noche.
- 30 8. Retirar por aspiración el medio en cada pocillo. Lavar cada pocillo con 100 µl de 293 SFM II con L-glutamina 4 mM. Añadir 100 µl de 293 SFM II con L-glutamina 4 mM a cada pocillo.
9. Recoger el sobrenadante para ELISA a las 96 horas tras la transfección.

ELISA funcional

- 35 1. Recubrir placas de 96 pocillos Nunc-Immuno Maxisorp con 100 µl de antígeno 2 µg/ml en disolución de recubrimiento.
2. Cubrir las placas con elementos de sellado e incubar durante la noche a 4°C.
3. Decantar las placas y extraer con golpes suaves el líquido residual.
4. Añadir 200 µl de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.
- 40 5. Decantar las placas y extraer con golpes suaves el líquido residual.
6. Añadir 200 µl de disolución de bloqueo. Agitar a 200 rpm durante 1 hora a temperatura ambiente.

ES 2 701 931 T3

7. Decantar las placas y extraer con golpes suaves el líquido residual.
8. Añadir duplicados de 100 µl/pocillo de anticuerpo control (2 µg/ml) en disolución de bloqueo a las placas.
9. Añadir duplicados de 100 µl de sobrenadante procedente de la transfección (SOP 5A) a las placas.
10. Agitar a 200 rpm durante una hora a temperatura ambiente.
- 5 11. Decantar las placas y extraer con golpes suaves el líquido residual.
12. Añadir 200 µml de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.
13. Repetir las etapas 11-12 3 veces.
14. Añadir 100 µl de dilución 1:5000 de anticuerpo de cabra anti-ser humano purificado mediante afinidad conjugado con HRP en disolución de bloqueo a cada pocillo.
- 10 15. Agitar a 200 rpm durante una hora a temperatura ambiente.
16. Decantar las placas y extraer con golpes suaves el líquido residual.
17. Añadir 200 µl de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.
18. Repetir las etapas 17-18 3 veces.
- 15 19. Añadir 100 µl de sustrato TMB de Sigma a cada pocillo. Incubar a temperatura ambiente y comprobar cada 2-5 minutos.
20. Añadir 100 µl de HCl 1 H para detener la reacción.
21. Leer a 450 nm.

ELISA de cuantificación

- 20 1. Recubrir placas de 96 pocillos Nunc-Immuno Maxisorp con 100 µl de anticuerpo de cabra anti-IgG humana específico de Fc purificado mediante afinidad 10 µg/ml en disolución de recubrimiento.
2. Cubrir las placas con elementos de sellado e incubar durante la noche a 4°C.
3. Decantar las placas y extraer con golpes suaves el líquido residual.
4. Añadir 200 µl de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.
5. Decantar las placas y extraer con golpes suaves el líquido residual.
- 25 6. Añadir 200 µl de disolución de bloqueo. Agitar a 200 rpm durante 1 hora a temperatura ambiente.
7. Decantar las placas y extraer con golpes suaves el líquido residual.
8. Añadir duplicados de 100 µl/pocillo de concentración normalizada de IgG sérica humana purificada en disolución de bloqueo a las placas.
9. Añadir duplicados de 100 µl de sobrenadante procedente de la transfección (SOP 5A) a las placas.
- 30 10. Agitar a 200 rpm durante una hora a temperatura ambiente.
11. Decantar las placas y extraer con golpes suaves el líquido residual.
12. Añadir 200 µl de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.
13. Repetir las etapas 11-12 3 veces.
- 35 14. Añadir 100 µl de dilución 1:5000 de anticuerpo de cabra anti-ser humano purificado por afinidad conjugado con HRP en disolución de bloqueo a cada pocillo.
15. Agitar a 200 rpm durante una hora a temperatura ambiente.
16. Decantar las placas y extraer con golpes suaves el líquido residual.
17. Añadir 200 µl de disolución de lavado. Agitar a 200 rpm durante 5 min a temperatura ambiente.

18. Repetir las etapas 17-18 3 veces.

19. Añadir 100 µl de sustrato TMB de Sigma a cada pocillo. Incubar a temperatura ambiente y comprobar cada 2-5 minutos.

20. Añadir 100 µl de HCl 1 N para detener la reacción.

5 21. Leer a 450 nm.

Ejemplo 4. Generación y examen de una biblioteca de variantes de codones de Fc para expresión óptima de anticuerpos

10 El presente ejemplo proporciona métodos para generar una biblioteca de variantes de codones de Fc y métodos de examen para obtener variantes de Fc optimizadas para expresión mejorada en células huésped de producción en comparación con la forma parental del polipéptido Fc.

A. Diseño y construcción de una biblioteca de variantes de codones de Fc

15 Para cada codón en la zona diana (en este caso la parte Fc de la molécula de IgG1 humana) se diseña un par de cebadores degenerados (directo e inverso) que incluye el codón diana y 20 bases en cada lado. La 3ª posición del codón diana (posición de titubeo) contiene bases mixtas (tabla 3) que permiten la generación de todas las mutaciones silenciosas en la posición diana usando el mismo codón (ejemplo A). Se diseña un segundo conjunto de cebadores degenerados para la misma posición de codón si el aminoácido correspondiente puede codificarse por otro codón (ejemplo B). Se mezclan los cebadores degenerados directo e inverso correspondientes 1:1, se aparean con el molde y se extienden para obtener los productos de longitud completa mediante desplazamiento de hebra usando una ADN polimerasa termoestable. El molde se digiere con DpnI y se transforman los transforman productos de extensión de longitud completa en *E. coli*. Se secuencian hasta 12 colonias por reacción de mutagénesis. Se disponen los mutantes confirmados en cuanto a secuencia en placas de 96 pocillos y se guardan en glicerol. Las reservas en glicerol se usan para realizar minipreparaciones de ADN de plásmido para su transfección en células de mamífero y su examen.

Tabla 3: Códigos para bases degeneradas en oligonucleótidos sintéticos

Símbolo	Base mixta
R	A,G
Y	C,T
M	A,C
K	G,T
S	C,G
W	A,T
H	A,C,T
B	C,G,T
V	A,C,G
D	A,G,T
N	A,C,G,T

25 **Ejemplo A:** codón diana = CCC (prolina)

→cebador directo: CCD, cebador inverso: HGG

Ejemplo B: codón diana = TCG (serina)

→ cebador directo 1: TCH, cebador inverso 1: DGA

→ cebador directo 2: AGY, cebador inverso 2: RCT

30 No se muestran las 20 bases que flanquean el codón diana. Longitud de cebador total: 43 bases.

B. Expresión y examen basado en ELISA de la biblioteca de variantes de codones de Fc

5 Se transfectaron clones de la biblioteca de variantes de codones de Fc en una línea celular de mamífero. Se produjeron IgG de longitud completa y se secretaron al medio. Se examinaron los sobrenadantes de las variantes de codones de Fc expresadas para determinar el nivel de expresión de IgG superior al del clon parental usando ensayo de ELISA. Se normalizaron los datos de ELISA con ensayo de beta-galactosidasa que mide la eficacia de transfección. Las mejores coincidencias identificadas en el examen primario volvieron a transfectarse y a examinarse tres veces para confirmar el nivel de expresión aumentado.

REIVINDICACIONES

1. Un método de evolución de una proteína en un huésped de producción celular eucariota; comprendiendo el método:
- 5 a. hacer evolucionar una proteína molde para producir un conjunto de proteínas mutantes en el huésped de producción celular eucariota;
- b. examinar el conjunto de proteínas mutantes para al menos una propiedad, característica o actividad predeterminada;
- c. seleccionar una proteína mutante superior del conjunto de proteínas mutantes basándose en la al menos una propiedad, característica o actividad predeterminada; y
- 10 d. fabricar la proteína mutante superior que comprende expresar la proteína mutante superior en el mismo huésped de producción celular eucariota como el usado en la etapa a de hacer evolucionar
2. El método de la reivindicación 1, en el que la etapa de selección (c) comprende:
- seleccionar la proteína mutante superior del conjunto de proteínas mutantes basado en
- 15 i. optimización de la al menos una propiedad, característica o actividad predeterminada seleccionada de la reducción de la agregación proteína-proteína, mejora de la estabilidad de la proteína, solubilidad de la proteína mejorada, introducción de sitios de glicosilación, introducción de sitios de conjugación, mejora de la expresión de la proteína, afinidad por el antígeno mejorada, un cambio en la afinidad de unión, un cambio en la inmunogenicidad, un aumento de la selectividad y un aumento de la especificidad, en comparación con la proteína molde; y
- ii. expresión modificada en comparación con la proteína molde.
- 20 3. El método de la reivindicación 1, en el que la etapa de selección (c) comprende:
- seleccionar la proteína mutante superior del conjunto de proteínas mutantes basado en
- i. optimización de la al menos una propiedad, característica o actividad predeterminada seleccionada de una estabilidad mejorada del pH de la proteína, una estabilidad mejorada de la temperatura de la proteína, una estabilidad mejorada del disolvente de la proteína, una selectividad reducida, una afinidad reducida por el antígeno, un cambio en la actividad catalítica, optimización del pH, y una reducción de la inmunogenicidad; y
- 25 ii. expresión modificada cuando se compara con la proteína molde.
4. El método de una cualquiera de las reivindicaciones 2-3, en el que la expresión modificada es una expresión mejorada en el huésped de producción de células eucariotas.
5. El método de la reivindicación 1, en el que la etapa a. de hacer evolucionar comprende una de evolución de posición completa; evolución de inserción de posición completa; evolución de delección de posición completa; evolución de posición completa seguida por síntesis combinatoria de proteínas; evolución de delección de posición completa seguida por síntesis combinatoria de proteínas; o evolución de delección de posición completa seguida por síntesis combinatoria de proteínas.
- 30 6. El método de la reivindicación 1, en el que la proteína molde se selecciona de una enzima, una citoquina, un receptor, una proteína de unión al ADN, un agente quelante, un anticuerpo y una hormona.
7. El método de la reivindicación 1, en el que la proteína molde es una enzima, una citoquina o un receptor.
8. El método de la reivindicación 1, en el que la etapa de hacer evolucionar a. comprende hacer evolucionar la proteína molde para producir un conjunto de proteínas mutantes en el huésped de producción de células eucariotas con visualización de la superficie celular.
- 40 9. El método de la reivindicación 1, en el que la proteína es un fármaco terapéutico y la proteína mutante superior es un fármaco biosimilar.
10. El método de la reivindicación 1, en el que uno o más codones de mutación silenciosos dentro de la proteína se seleccionan y evolucionan.
11. El método de la reivindicación 1, que comprende además las etapas de:
- 45 generar una biblioteca de proteínas en un huésped de producción de células eucariotas;
- examinar la biblioteca para al menos una propiedad, característica o actividad predeterminada; y

seleccionar la proteína molde de la biblioteca basándose en al menos una propiedad, característica o actividad predeterminada.

12. El método de la reivindicación 10, en el que la etapa de generación comprende generar uno de:

5 n-1 conjuntos independientes de proteínas mutadas mediante la mutación de cada residuo de 2 a n de la proteína parental de uno en uno, teniendo cada proteína mutada n residuos de aminoácidos, comprendiendo cada conjunto de proteínas mutadas proteínas mutadas que tienen un número X de diferentes residuos de aminoácidos predeterminados en una única posición predeterminada de 2 a n de la proteína mutada; en el que cada conjunto de proteínas difiere de la proteína parental solo en la posición predeterminada única; y el número de diferentes proteínas mutadas generadas es equivalente a $(n-1) \times X$,

10 ii. n-1 o n-2 en el caso de que el residuo inicial sea metionina, separar proteínas mutadas eliminando cada residuo de 2 a n en la proteína parental de una en una, en donde cada proteína difiere de la proteína parental en que un aminoácido se elimina en una posición predeterminada única, y

15 iii. $20 \times (n-1)$ proteínas mutadas separadas insertando cada uno de los 20 aminoácidos naturales en cada posición entre dos residuos de aminoácidos adyacentes en la proteína parental, uno a la vez, en donde cada proteína mutada difiere de la proteína parental en que se ha insertado después de una posición específica en el parental uno de los 20 aminoácidos naturales.

13. El método de la reivindicación 11, que comprende además las etapas de

ensayar cada proteína mutada para al menos una propiedad, característica o actividad predeterminada;

20 para cada proteína mutada que identifica un cambio en dicha propiedad, característica o actividad en relación con una misma propiedad, característica o actividad de la proteína parental;

crear un mapa funcional que correlaciona cada mutación en cada una de dichas posiciones de la proteína parental con el cambio identificado en dicha propiedad, característica o actividad de cada proteína mutada en relación con la misma propiedad, característica o actividad de la proteína parental; y

utilizar dicho mapa funcional para dicha etapa de hacer evolucionar.

25 14. El método de la reivindicación 1, en el que el huésped de producción de células eucariotas se selecciona de células de fibroblastos de ratón 3T3; Células de fibroblastos de hámster sirio BHK21; MDCK, células epiteliales de perro; células epiteliales humanas Hela; células epiteliales de rata canguro PtK1; células plasmáticas de ratón SP2/0; y células plasmáticas de ratón NS0; células renales embrionarias humanas HEK 293; células de riñón de mono COS; CHO; células de ovario de hámster chino CHO-S; células embrionarias de ratón R1; células embrionarias de ratón E14.1; células embrionarias humanas H1; células embrionarias humanas H9; PER C.6, células embrionarias humanas; células de levadura de *S. cerevisiae*; y células de levadura *picchia*.

30

15. El método de la reivindicación 1, en el que el huésped de producción de células eucariotas se selecciona de células de riñón embrionarias humanas HEK 293 y células de ovario de hámster chino CHO-S.

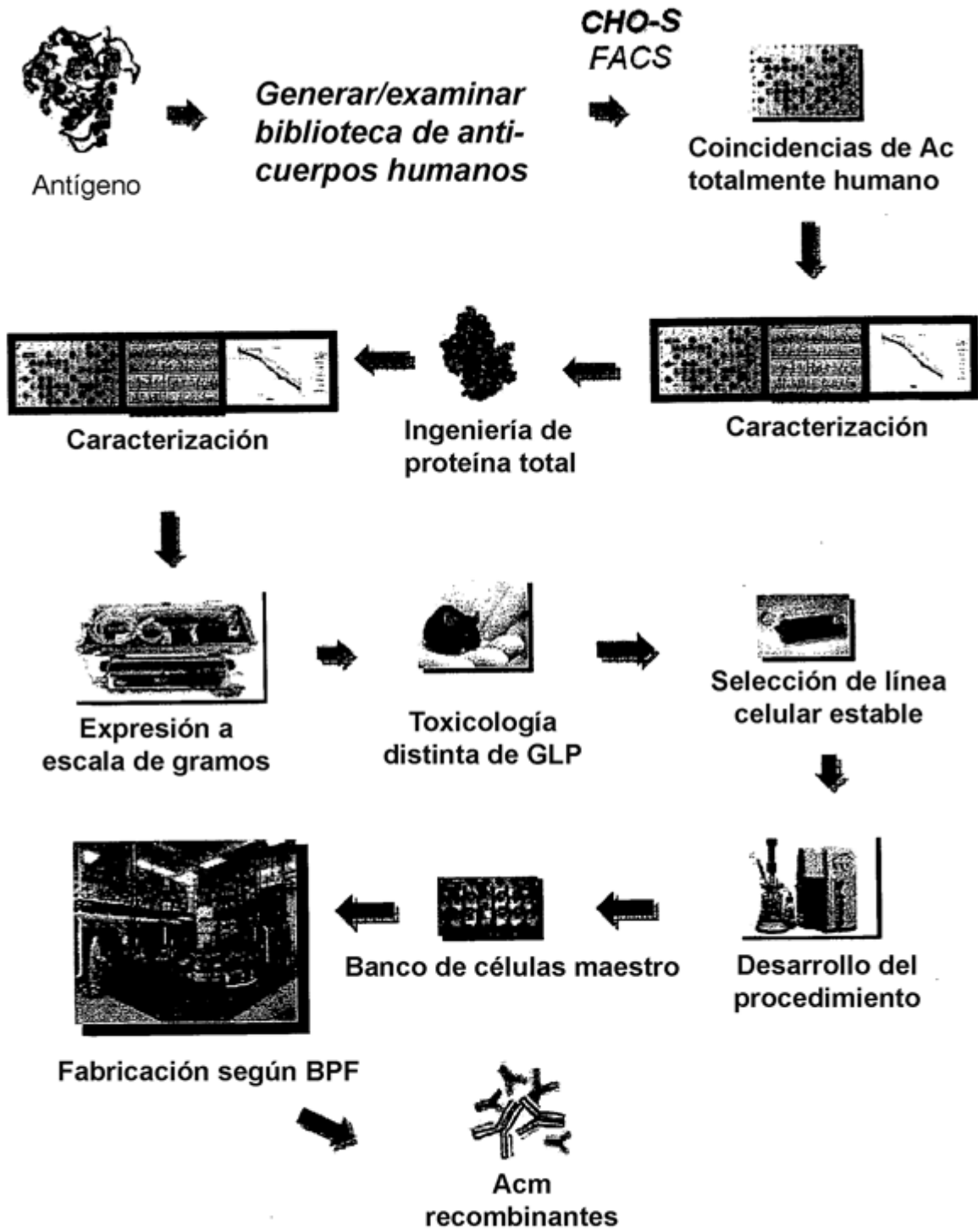


FIGURA 1

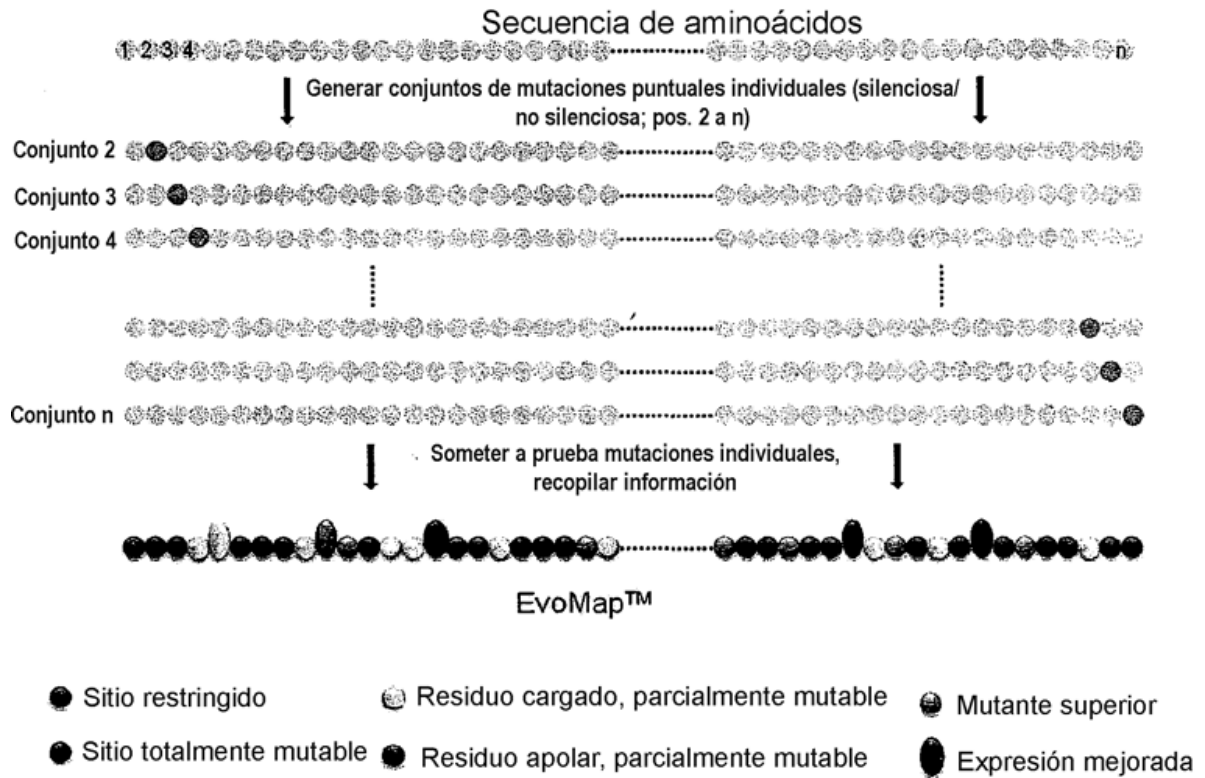


FIGURA 2

Síntesis combinatoria de proteínas **CPSTM**

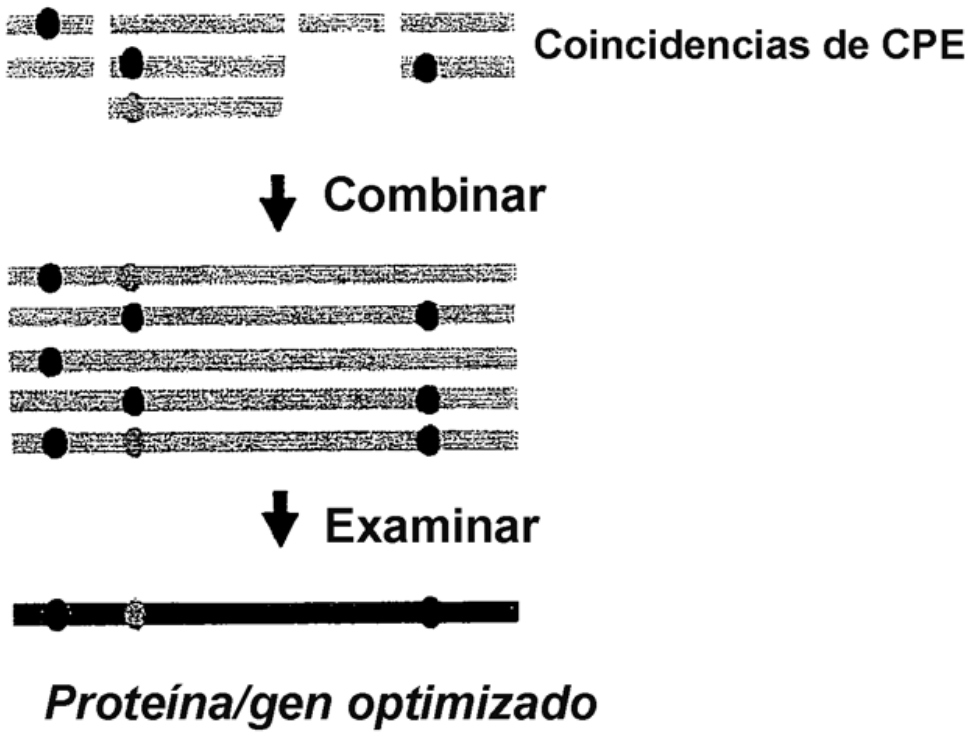


FIGURA 3

CPI™- Evolución de inserción de posición completa
Inserta hasta 20 aminoácidos individuales en cada posición
de aminoácido individual en la totalidad de la proteína

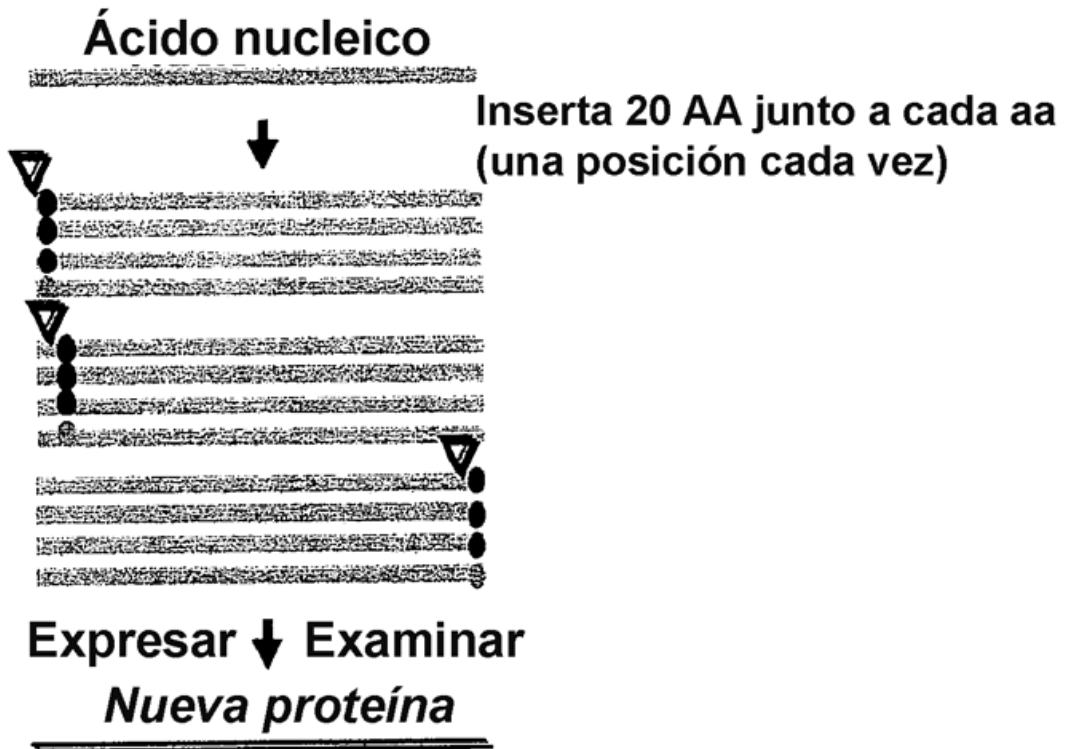


FIGURA 4

Evolución de posición completa (CPE™)

Ácido nucleico alargado de CPI

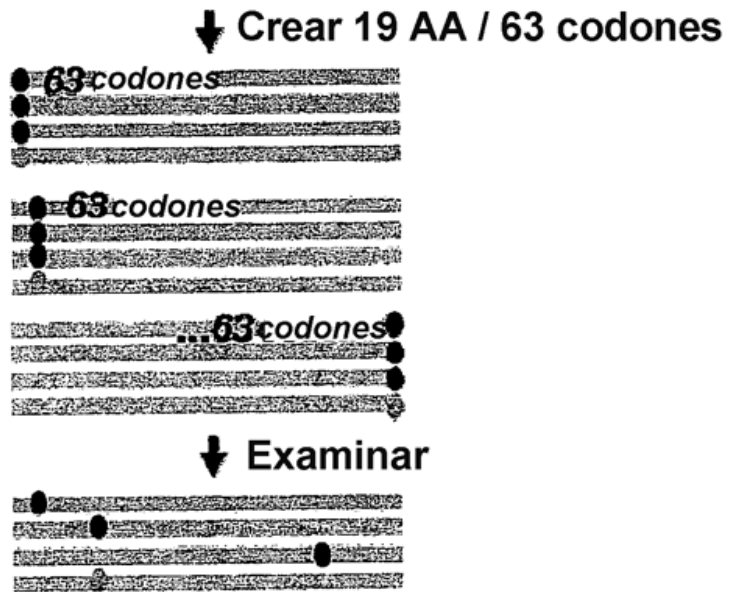


FIGURA 5

CPD™- Evolución de deleción de posición completa

Deleciona cada aminoácido a través de una proteína una posición cada vez

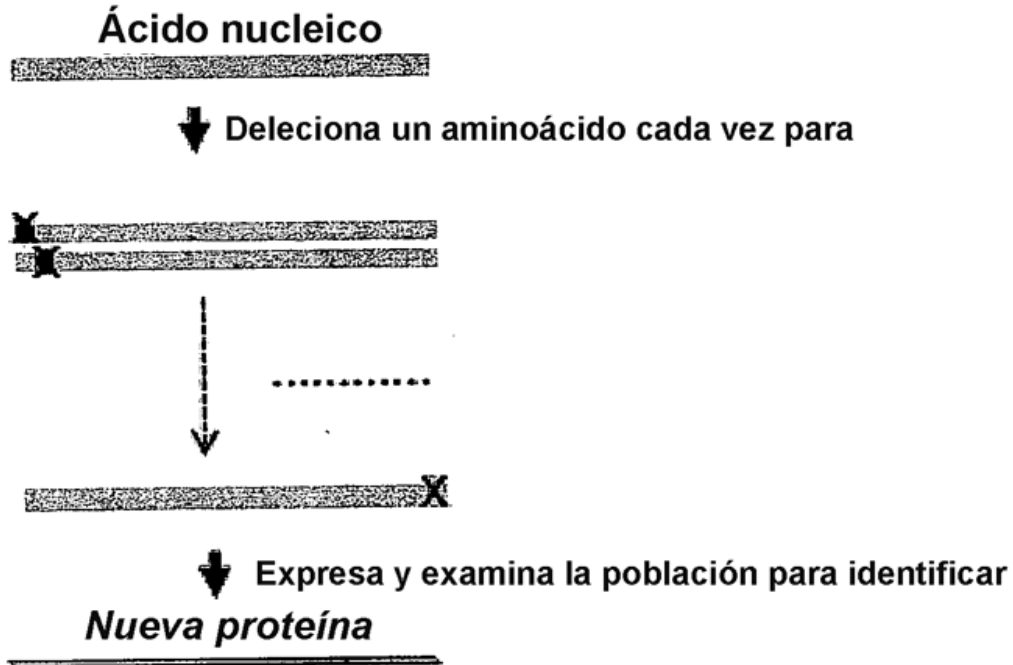


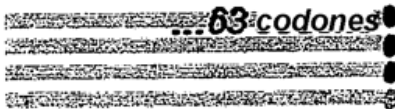
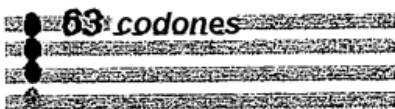
FIGURA 6

Evolución de posición completa (CPE™)

Ácido nucleico
acortado de CPD



↓ Crear 19 AA / 63 codones



↓ Examinar



Proteína/gen optimizado

FIGURA 7